POLITECNICO DI MILANO

Management Engineering Department

ARTIFICIAL INTELLIGENCE IN THE SPORT INDUSTRY

Supervisor: Prof. Emanuele Lettieri

Co-Advisor: Dott. Claudio Rorato

Ing. Federico Iannella

Master Degree Thesis of:

Jacopo Mosele 875678

Academic year 2017/2018

SUMMARY

IMMAGES SUMMARY

ABSTRACT-EN

This work aims to 1. Mapping the current state-of-art of artificial intelligence-based systems in the sport industry and 2. Investigate how artificial intelligence can become a competitive advantage in the sports organizations' strategy, identifying the main benefits that such solutions can generate. The work comprises an introduction of literature review, that allowed to detect the main AI classes and their evolution. After that, the current state of sport business has been presented, highlighting dynamics, macro-trend and future perspectives. The central body represents the mapping phase of the AI-based solutions currently available on the market, dedicated to sport. Through the analysis of the common features, it was possible to identify the main application areas (optimization of sport's performances, support to managerial processes, and, tool for the value creation in the relationship with fans). It has been also presented the working methodologies, artificial intelligence technologies leveraged, and business' processes involved. From this analysis it appears that the main benefits of these solutions are related to their employment as tool, in supporting decision processes, currently characterized by uncertainty and complexity. It follows that artificial intelligence systems are already nowadays source of value, both in term of effectiveness and reliability in the physical performance management, injuries management, in the tactical analysis of sport team dynamics, in the economic evaluation of players and in development of club-fans relationship. To conclude, some guidelines have been provided that can be used by a generical sport organisation in the evaluation of the artificial intelligence-based systems' introduction.

ABSTRACT-ITA

Questo lavoro di tesi si prefigge gli obiettivi di: 1. mappare l’attuale stato dell’arte dei sistemi di intelligenza artificiale nel mondo dello sport e 2. Comprendere come l’intelligenza artificiale possa diventare fonte di vantaggi competitivi nella strategia di un club sportivo, identificando i benefici principali che tali soluzioni possono generare. Il lavoro si compone di una fase introduttiva di analisi bibliografica che ha permesso l’identificazione delle principali classi di intelligenza artificiale e della loro evoluzione storica. Successivamente è stato introdotto lo stato attuale dell’industria sportiva, presentandone le principali dinamiche, macro-trend e prospettive future. Il corpo centrale è rappresentato dalla mappatura delle soluzioni attualmente disponibili di intelligenza artificiale nel contesto sportivo. Attraverso l’analisi degli aspetti comuni alle diverse realtà sul mercato, sono state individuate le principali area di applicazione (ottimizzazione delle performance sportive, supporto ai processi manageriali e strumento di creazione di valore nella relazione con il tifoso), i metodi di lavoro, le classi di intelligenza artificiale alla base, e il loro impatto sulle performance dei processi aziendali interessati. Da questa analisi si ricava che i benefici di tali soluzioni sono massimi quando essa viene impiegata come strumento di supporto a processi decisionali attualmente caratterizzati da incertezza e complessità. Si evince quindi, che strumenti di intelligenza artificiale rappresentano già oggi fonte di valore, sia in termini di efficacia che di affidabilità nella gestione delle performance atletiche, nella gestione degli infortuni, nell’analisi tattica di sport di squadra, nella valutazione del valore economico degli atleti, nella stima del valore delle sponsorizzazioni e nella valorizzazione del rapporto club-tifoso. Per concludere, sono state individuate delle linee guida utilizzabili da un generico club sportivo per la gestione di un’eventuale introduzione di simili sistemi nei suoi processi.

EXECUTIVE SUMMARY

The following document will be divided into three parts, that represent the as many logical blocks that compose the work.

In Part I, Artificial Intelligence and Sport Industry will be presented, sequentially and as standing-alone subjects.

Part II is the result of the mapping phase of all the cases in which AI has been employed regarding sport business, along with the common features between the different cases, the working methodologies and the value's propositions.

In Part III, three case studies about Italian organizations that are developing AI-based solutions in the sport industry will be presented, along with the conclusions about the opportunities and the main benefits that AI technologies offer, at this moment, for sport clubs.

Part I has been written using a methodologic approach, since the arguments are presented mainly employing academic models.

Part I is aimed to create a notional background useful to walk through the followings parts, that are more practical, in the sense of based on the analysis of real-world cases.

Regarding Artificial Intelligence, the handling has been initially dedicated to the development of a logical framework that permits to answer to the following question "*What an artificial intelligence system is? And how can I recognize it?".*

To answer this question, the definitions' matrix proposed by J. Stuart Russel and Peter Norvig in "*Artificial Intelligence: A Modern Approach*" in 1995 has been presented. It distinguishes within the Artificial Intelligence world, four main development flows, that differ in function to the different dimension of analysis: one is Human Inspired- Not Human Inspired, and the other is, Able to think – Able to Act. In this way, the categories arise as: 1. Acting Humanly (Human Inspired, Able to Act), 2. Thinking Humanly (Human Inspired, Able to think) 3. Thinking Rationally (Not Human Inspired, Able to Think) 4. Acting Rationally (Not Human Inspired, Able to Act).

The Acting Rationally group is the more interesting for the purposes of this thesis, mainly for the potential to create value within real-world organization, because, these systems develop a form of intelligence, that authors call “Rationality”, able to perform tasks that humans cannot.
This part has permitted to understand over which artificial intelligence systems focus the work. The following logical step is answering the question *“What does it means Acting Rationally for a machine?”* or, in other words, which are the traits that mark the intelligence for an artificial system.

The starting point for answering this second question has been the definition proposed by Nils J. Nilson that defines artificial intelligence *“as the quality that enables an entity to function appropriately and with foresight in its environment”.*
To match the *“function appropriately and with foresight in its environment”* condition, the system is expected to *“Perceive”* the environment around, to *“comprehend”* it and to *“act”* consequently[1]. The ability to perform these three logical steps should be embodied in any form of artificial intelligence. Once qualitatively defined what Artificial Intelligence is, it is possible to describe which is its role in the real-world organizations. It has been identified that artificial intelligence systems generate the maximum value, when they work for supporting inefficient decision-making processes. In fact, artificial intelligence systems permit to overcome the limits of human bounded rationality that are caused by: 1. Restricted computational power, 2. Limited time available and the 3. Presence of cognitive biases in our mental processes. For these reasons, human operators systematically take sub-optimal decisions. Artificial Intelligence address this specific issue.
Indeed, it is possible to notice that artificial intelligence has been mainly applied to: 1. Identify and quantify undiscovered causal links between variables. 2. Forecast future outcomes. Two activity that are difficult to be performed by the human operators.

After that, the main artificial intelligence methodologies, intended as algorithms’ classes, have been presented, along with their historical evolution.

[1] *Perception – Reasoning – Acting scheme proposed in the Accenture report “Why is Artificial Intelligence is the future of growth?”*

Currently, the most promising applications of AI are linked to Machine Learning's techniques. Machine Learning refers to solutions that can learn from experience expressed in form of data, and, perform tasks for which they are not expressively programmed. Following it, the three main types of Machine Learning has been presented: 1. Supervised-Learning, meaning that the system is trained with human-made instructions about how to perform the task. The system is required to apply these rules in new cases 2. Unsupervised Learning, where the system isn't supported by any human interventions but, instead, it is expected to develop autonomously rules and procedures to perform effectively the task and 3. Reinforcement Learning, an algorithms classes able to change autonomously its working methodology as a result of an optimization process of its performances.

After it, the Artificial Neural Network (ANN) paradigm has been introduced. It is a logical working methodology, that enables to maximize the learning abilities of the AI systems thanks to the fact that it achieves the solutions decomposing the initial issue into simpler causal links. Indeed, it expressed the problem as: Input- (Many Variables)- Output.

At this point, the deep learning evolution has been introduced. Deep Learning refers to the employment of very deep neural networks. Deep means very large and hyperconnected. This characteristic permits to Deep Learning algorithms to achieve a high level of abstraction in facing the problems, and, so, develop, the ability to solve more complex issues.

The last section is a comparison between the human performances and the performances of the best algorithms, regarding three different tasks with, so many, different outcomes. In one task humans are still better than artificial intelligence, in one currently humans and AI are at the same level and in one artificial intelligence outperform humans' abilities.

The linking point between artificial intelligence and sport is represented by an historical digression on two well-known episodes in the artificial intelligence history: the chess match between Deep Blue of IBM and the world champion Garry Kasparov, and, the match between AlphaGo by Google and Lee Sedol in the game of Go. They are both artificial intelligence applications superficially build up to compete with the best human players in two strategic and complex games. They both resulted as winner. This digression helps to understand better the evolutions of artificial intelligence systems in the last decades, and, it uses best-in-class examples to present some of the concepts introduced before.

Part 1B is an introduction about modern sport industry. To perform it, the conceptual model proposed by the Osservatorio Innovazione Digitale nell'Industria dello Sport of Politecnico di Milano, has been employed. This model has been developed to map the areas where digital technologies can impact the sport business. In this work, it has been used to introduce the current Sport Industry. The model individualizes four main macro-areas of activity: 1. Fan experience 2. Athletic Performance 3. Management of sport event 4. Organization management. Each of them is composed into sub-sections that have been deeply explained.

The takeaway of this part is that modern sport environment is getting highly complex. Sport organizations must deal with a fan base worldwide spread, always more demanding in term of engaging initiatives. The industry is also facing a great level of competition from a sport performance point of view, that forces clubs to find always new competitive advantages respect to other teams.

The activities of sport clubs have been broadened, since now they must generate on-line contents, manage communication platforms and big physical facilities like the arenas. In this context, the digital technologies play a central role in supporting the clubs in this dynamic environment.

Part II is the result of a research's activity whose aim is mapping the applications' areas of AI in the sport business.

From the analysis of the cases in which AI has been applied in this industry, it has been possible to individualize four main applications areas. The first one is about the Activity and Performance management of players, which is further divided into: 1. Physical Performance optimization, that are systems, that basing on biomedical and biomechanical data analysis, aim to improve the training effectiveness, perform injury management and optimize the overall condition of the players 2. Development of virtual training environment, that are artificial intelligence systems, along with other technologies like VR and AR, that are dedicated to the creation of virtual training sessions 3. Team performance analysis intended as development of tools to assess collective tactical actions.

The second area is Management and Organization. It can be referred to the 1. Team management, that are artificial intelligence solutions devoted to scouting and players' investment evaluation and optimization. 2. Artificial Intelligence employed as support of business strategies like sponsorships' fair value computation and dynamic ticketing process. 3. Artificial intelligence employed in the fan engagement and media creation initiatives. For example, it has been used in media content creation,

for delivery more engaging contents, and in arenas' management, for making the experience more valuable for fans.
The last category is called external stakeholders, and it comprises 1. Artificial intelligence applications for betting purposes, and 2. Solutions dedicated to support the referees' tasks.

Part III comprises: 1. The case studies. They are dedicated to two Italian already developed start-ups: Math&Sport and Wallabies, and, one to a start-up that has been just founded, Playerank. Math&Sport has developed a software for the football tactical analysis, Wallabies created a system dedicated to scouting and for the players evaluation and Playerank, has the objective to create an AI-based system to correctly evaluate players performances, in relation to their contribution to the results.
2. The conclusions, where it will be presented why artificial intelligence systems represent an opportunity for sport organizations, which benefits are generated along with a framework that can support a club in planning the preliminary moves towards the introduction of such systems in its operations.

PART IA - ARTIFICIAL INTELLIGENCE

INTRODUCTION

1A.1 DEFINITIONS OF ARTIFICIAL INTELLIGENCE: THE RUSSEL AND NORVIG 'S MATRIX

The term Artificial Intelligence was first coined in 1956 by john McCarthy[2], during a conference organized at the Dartmouth College by some of the maximum experts of this rising research field. Although it is more than 60 years that AI exist as a standing-alone subject, a "*precise and universally accepted definition of AI does not yet exist"[3].*

In the years, many scholars proposed various definitions, highlighting each time different nuances of meaning and inspiration's sources to delineate AI.

In this work, the topic will be presented following the logic matrix, proposed by J. Stuart Russel and Peter Norvig in "*Artificial Intelligence: A Modern Approach*" in 1995, showed in the figure 1.

---

[2] *Machine Who Thinks: A Personal Inquiry into the History and Prospects of Artificial Intelligence Pamela McCorduck*

[3] *Artificial Intelligence and Life in 2030, One-hundred-year study on Artificial Intelligence, report of the 2015 study panel*

Thinking Humanly	Thinking Rationally
"The effort to make computers think… machines with minds, in the full and literal sense" Haugeland, 1985) "[the automation of] of activities that we associate with human thinking, activities such as decision-making, problem solving, learning…" (Bellman, 1987)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998) "AI, broadly defined, is concerned with intelligent behaviours in artefacts." (Nilsson, 1998)

Figure 1: The matrix proposed by Russel and Norvig in their" Artificial Intelligence: A Modern Approach" in 1995 to categorized different types of Artificial Intelligence systems. For each category identified, some definition of different authors have been presented to clarify the meaning.

The model is a bit dated but, nevertheless, it shines for clarity. According to the authors, it is possible to distinguish, within the general topic of AI, four main flows, in function of two dimensions of analysis:

- Vertically speaking, on the top of the figure 1, there are definitions concerned with thoughts' processes and reasoning, whereas the ones on the bottom, address the behavior of the system. So, the idea is distinguishing between systems that can act from systems that can just process information.

- Horizontally speaking, on the left there are systems that are human-inspired, and, thus, highly faithful to human performance. On the right, the systems, instead, follow an ideal standard of intelligence. This ability is called by the authors "*Rationality".* They also added that a system is rational if it does the "*right thing*", given what it knows, or in other terms, if it will take the best decision in the specific environment which it is placed in, that must be uncertain, by definition. The meaning of "*Take the best decision*" will be explained later in a more detailed way. For now, it is possible to link the idea of "*best decision*" with two aspects.
    1. Acting to meet a certain objective, so in other words, the action must be the result of an optimization.
    2. Perceiving the environment in which the system is involved.

Now, the four categories will be explained more in details.

1A.1.1 ACTING HUMANLY

This first category looks at AI as a branch of computer science, dealing with the imitation of human behaviors. The goal is creating systems that can replicate humans' abilities and, potentially, replace a human operator at that specific task.
Alan Turing (1950), proposed the so-called "Turing Test", to prove the operational intelligence of a machine. He affirmed that a machine, to be considered intelligent, have to pass it. In developing this definition, he also provided a perfect example of what means "Acting Humanly".

Turing suggested that a computer would pass the "Turing Test", only if a human interrogator, after posing some written questions to the computer itself, cannot tell whether the written responses come from a person or from a computer[4]. Rephrasing it in modern terms, if a computer could pass for a human in, for example, an on-line chat, it should be considered as intelligent, because it is able to act as a human[5].

[4] *Artificial Intelligence: A modern approach, Stuart J. Russel, Peter Norvig*
[5] *The concept can be generalized for every task, not only answering questions' tasks*

1A.1.2 SEARLE ANTITHESIS: THE CHINESE ROOM ARGUMENT

The results and the conclusions achieved by the "Turing Test" have been partially disproved by Searle (1980), when he proposed the "Chinese Room Argument".

The Searle argument arrived 30 years later respect to the "Turing Test", so in comparing them, it is necessary to consider that in this time frame, the frontiers of AI have moved far ahead, and with them, also our ideas of AI have changed.

Researches call this phenomenon the "Odd Paradox" and it simply "*states that when AI brings a new technology into the common fold, people become accustomed to this innovation, and it stops immediately to be considered as AI, and newer technology emerges*[6].

Searle argued that passing the "Turing Test" is not a sufficient condition to consider a machine as intelligent, because, instead, what really matters is how the machine passes the Test. The machine can be considered intelligent only if it has a sort of "awareness" of what is going on around it.

To support this thesis, he created this example, demonstrating that a person can pass to know Chinese even if he does know anything of Chinese. He asked to imagine a native English speaker who knows no Chinese, in a room, together with a book of instructions for manipulating the symbols, and a person outside the room, who sends in the room questions, written in Chinese.

The book is the equivalent of a computer program, made of instructions that enables the person in the room to correctly answer the questions, having no idea of what he is writing.

The book is potentially made of thousands of lines like this one: "if you receive a message like this one" 你会吗？", so write this one as answer "我会说中文"". The program could easily enable the person in the room to pass the Turing Test of Chinese with the person outside, even if he does not understand Chinese.

The critics that Searle moved was that Artificial Intelligence, at least, is not only a matter of what the system is able to perform, but what really counts, is how the system perform the tasks. Generalizing, if

---

[6] *Artificial Intelligence and Life in 2030, One-hundred-year study on Artificial Intelligence, report of the 2015 study panel*

a machine can perform a very complex task, but just following strictly an immense database of rules and instructions human-written, thanks to his large computational power, can we consider it still AI? The answer of Searle is probably no.

1A.1.3 THINKING HUMANLY: THE COGNITIVE MODELLING APPROACH

This AI stream deals mainly with systems aiming to imitate the cognitive abilities of humans. The object, here, is not making the systems able to solve problems, but, on contrary, the aim is creating systems able to see at the problems, following exactly the path in which human subjects would have faced the same problems. In other words, to copy how humans think.

Their main application is building up precise and testable theories about how the human mind works and predicting it. This stream is, naturally, mainly dedicated to psychological and neuro-scientific researches.

1A.1.4 THINKING RATIONAL

This category comprises systems that are logically programmed. The idea behind is that all the problems faced by the system should be first translated and expressed in logical terms and, then, solved, using logical inferences. Unfortunately, they have not much application's room in the real world, since inference is just one of several possible mechanisms for achieving rationality. For this reason, they are mainly used for academic purposes to create models about how humans should "think".

1A.1.5 ACTING RATIONALLY

In this conceptual category lay the most interesting and valuable solutions for real-world applications. Here, all the efforts are *"…devoted to making machines intelligent, and intelligence [is intended as] that quality that enables an entity to function appropriately and with foresight in its environment"*[7].

---

[7] *Nils J. Nilsson, The Quest for Artificial Intelligence: A History of Ideas and Achievements (Cambridge, UK: Cambridge University Press, 2010*

Systems that aim to act rationally differ completely from all the others operational model proposed until now.

First, here there are no direct references to human intelligence as benchmark. Machines can be inspired by how people acts or thinks, but it is not a prerequisite. The computer systems are supposed to create their own "intelligence". "*This separation of AI from considering human Intelligence as benchmark and source of inspiration, has permitted to the field to grow up*"[8].

As highlighted by the authors in their book, this change in the way we think about AI was revolutionary. To explain the implications of this change, they made this example about the development of airplanes. They pointed out that all the attempts for "Artificial Flight" succeeded only when the *"Wright brothers stopped imitating birds, and, started using wind tunnels and learning about aerodynamics. Aeronautical engineering texts do not define the goal of their field as making machines that fly so exactly like pigeons, that they can fool even other pigeons"*[9].

Computer agents are, now, expected to do more than just imitating how humans act or think and eventually, perform it better. Indeed, they are required to operate autonomously, perceiving the environment around them, adapting to change, creating and pursuing goals.

For the developing of the thesis, it could be useful to focus on what *"functions appropriately and with foresight in its environment"* means, to define better, later on, the activities that AI systems can be required to do.

A general scheme to define how a system should face the surrounding uncertain environment is the "Sense-Comprehend-Act"[10] framework, showed in figure 2. It defines the logical steps that bring to the understanding of a real issue.

It has been used as conceptual benchmark to define how AI systems should behave appropriately and which activities are required to do it.

[8] *Artificial Intelligence: A modern approach, Stuart J. Russel, Peter Norvig*

[9] *Artificial Intelligence: A modern approach, Stuart J. Russel, Peter Norvig*

[10] *Accenture report "Why is Artificial Intelligence is the future of growth?"*

In the part on right, the figure shows the name of some AI solutions that can meet and perform that particular task. They will be deepened in the section of the thesis specifically dedicated to map the current systems of AI.

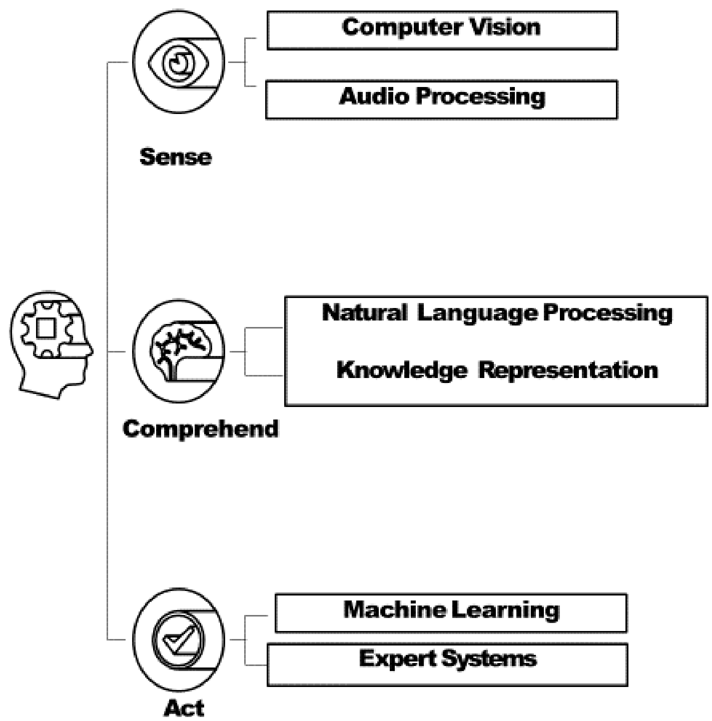


Figure 2: Perception – Reasoning – Acting scheme proposed in the Accenture report "Why is Artificial Intelligence is the future of growth?"

This scheme highlights that to act rationally are required the abilities to "*perceive and respond to sensory inputs, synthesize and summarize information, reason and achieve goals*"[11].

All these activities imply the idea of planning, solving problems, think abstractly and learn from experience.

[11] *Nils J. Nilsson, The Quest for Artificial Intelligence: A History of Ideas and Achievements (Cambridge, UK: Cambridge University Press, 2010*

The second aspect over which it is necessary to focus to define Rationality is related to the relationship between the system and the environment. To consider a system as a rational agent, the external environment must be dynamic and prone to change. The system should be able to perceive the environment and own the ability to change, if the surrounding environment does the same. In other words, *"Intelligence is [also] the ability to adapt to change"*[12].

The aim of this part was giving a qualitative definition and explanation of how it is possible to conceptually distinguish AI systems and which traits must be considered when an AI solution is evaluated.

Once identified them, the following logical step is defining which could be the general objectives and the strategic goals of AI systems in a given organization. Leaving the field of abstract definitions about the general scope and objectives, to move to more concrete applications.

1A.2 THE VALUE OF AI SYSTEMS: ENHANCE THE AGENTS' DECISION ABILITY

"We're at an inflection point where artificial intelligence can help people make faster, better and cheaper decisions."-The Human Factor, working with machines to make big decisions – PwC

It has been highlighted that the most interesting AI solutions are the ones that are able to "Act Rationally".

This paragraph will explain why a system that "*functions appropriately and with foresight in its environment"* can create the maximum value for an organization. For example, systems that "Act Humanly" can represent a source of efficiency and a costs' reduction for a company. They can perform the same task of humans, without getting tired or stopping. On the other hand, systems that can "Act Rationally" must have a different and more ambitious goal.

To define the role of these system, it is possible to take inspiration looking at the general economic theory and its evolutions in past decades. Initially, the neo-classical theory assumed that all the agents,

---

[12] *Stephen Hawking, https://www.brainyquote.com/quotes/stephen_hawking_378304*

when confronted with various alternatives, always choose the one that maximizes their individual utility[13].This theory lies on three fundamental assumptions: 1. Agents are rational, 2. Individual choices are consistent with expected utility, 3. People correctly update their opinions and beliefs based upon additional information received[14].

These hypotheses have been partially disproved by modern theories that, instead, consider these assumptions not true in the real world, where the rationality of agents' is, basically, bounded due to the limited time and computational power available and, also, because humans often face systematic cognitive biases, due to emotions and feelings.

Between these two definition lays the higher object for AI systems in human organizations. AI should mainly focus on extend human-agents rationality, to permits decision-makers to take better informed decisions.

Dedicating some effort to develop a clear definitions' framework of what is AI, is a valuable activity for this work. Only after having defined what is AI and what should be its purposes, it is possible to start to look at its applications in the real world.

1A.3 WEAK AND STRONG ARTIFICIAL INTELLIGENCE

Some academics have introduced a further classification to distinguish conceptually AI systems between Weak AI and Strong AI. Even if this new partition has not been specifically created as continuation of the Russel and Norvig's Matrix, it is possible to find several linking points between the two models.

Weak AI refers to systems that were previously approximately classified under the category of "Acting Rationally". The word "Weak" means that the system has just narrow applications. The large majority

---

13 *From Rational Choice to Behavioural Economics, Klaus Mathis and Ariel David Steffen.*

14 *Behavioral Economics vs. Traditional Economics: what is the difference? Dr. Vicki Bogan*

of AI systems now on the market or, anyway, used operationally by companies can make decisions and solve problems, but only in a very limited area of activities.

On contrary, Strong AI refers to future states of AI that will be theoretically able to face broad and different tasks, using the same physical and software infrastructure. Systems with a great versatility and resilient to face any type of problem. These two states must be considered as the development path that AI should follow in the next years: "*The evolution of weak systems will be the creation of strong systems or artificial general intelligence, which aims to perform, replicate and improve any intellectual task that human can do. This technology has yet to arrive and now has just a long-term business potential*"[15].

1A.4 THE EVOLUTION OF AI SYSTEMS

The discipline of Artificial Intelligence is broad, and it has seen several evolutions in his history. Nowadays the techniques that are most used in real-world applications are linked to the concept of Machine Learning and with his most innovative development, the Deep Learning. They are the approaches that show the most remarkable capacities to generate value.

This figure 3 shows how AI evolved in the last decades. With breath is intended the numbers of solutions existing. 'Artificial Intelligence' refers to all the solutions and categories expressed in the Russel Matrix at the beginning. 'Machine Learning' and, consequently 'Deep Learning', are sub-categories of 'Artificial Intelligence' one. They are referring to "Acting Rationally" approaches, since Machine Learning intends systems that can "learn" from experience, carrying out a task even if not specifically programmed to do it, just learning how a human operator does it. Deep Learning is another sub-category of Machine Learning. The difference between machine Learning and Deep Learning is that Deep Learning refers to systems that use a learning paradigm based on long and complex networks. From this, the name "Deep".

[15] *The Rise of Artificial Intelligence: Future Outlook and emerging risks, Allianz Global Corporate & Specialty*

The concepts of Neural Networks and Deep Learning will be explained later in a more detailed way later.

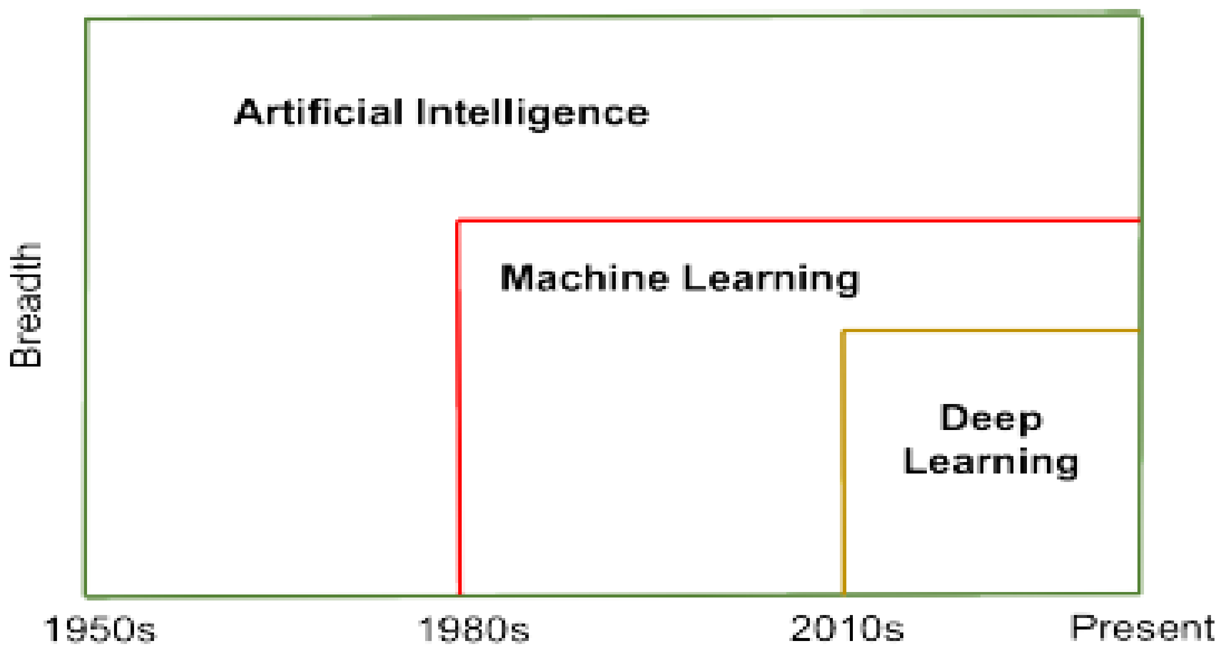


Figure 3: The chronological evolution of Artificial Intelligence considering Machine Learning and Deep Learning systems. Source: Deep Learning and the Artificial Intelligence Revolution, MongoDB White Paper, August 2017

1A.5 MACHINE LEARNING

Machine learning, as research's field, has been created to overcome the limits that the traditional computer science met facing complex problems.

"[At the beginning] advances in information technology have focused on codifying existing knowledge and procedures and embedding them in machines"[16]. *"Traditional approaches to programming rely on hardcoded rules, which set out how to solve a problem, step-by-step"*[17].

[16] *The business of Artificial Intelligence, what it can-and cannot-do for your organization, Harvard Business Review, Erik Brynjolsson and Andrew Mc Afee*

[17] *Deng, L.; Yu, D. (2014). "Deep Learning: Methods and Applications" (PDF). Foundations and Trends*

"*This approach had a fundamental weakness: much of the knowledge we have [as human beings] is tacit, meaning that we can't fully explain it verbally, because it is based on our personal experience, insights or intuitions*"[18] This phenomenon is called Polanyi's Paradox, and, it substantially says, that we know more than we can say or explain[19].

Machine Learning, conceptually, represents the attempt to overcome those limits, enabling computers to learn from examples and experience, expressed in form of data, through the creation of structured feedbacks to solve problems[20].

The working methodology of Machine Learning is completely different respect to traditional approaches. Machine Learning systems are just set a task or, let's say, an objective to reach. Then, they are given a large amount of data to use as examples of how this task can be performed, or, from which to detect patterns, and so "ideas", useful to reach the object itself. The system, in this way, learns the best way to achieve the desired output autonomously.

The increasing availability of data, the increasing computer processing power and the definition of better performing algorithms have permitted to Machine Learning to grow in recent years. In fact, some systems can outperform humans in several tasks. All of us now interact with systems based on Machine Learning every day, for example in image recognition systems, such as those used on social media; voice recognition systems, used by virtual personal assistants; and recommender systems, such as those used by online retailers. These are all examples of well-functioning systems based on Machine Learning technologies.

Coming back for a while at the Polanyi's Paradox, it is known that the need of a Learning approach arises from two major issues: the possible complexity in explaining a given procedure to solve a

in Signal Processing

[18] *The business of Artificial Intelligence, what it can-and cannot-do for your organization, Harvard Business Review, Erik Brynjolsson and Andrew Mc Afee*

[19] *https://medium.com/@jrodthoughts/can-computers-have-common-sense-polanyis-paradox-judgment-and-artificial-intelligence-4119c4199a1d*

[20] *Machine Learning, the power and promise of computers that learn by examples, The Royal Society*

problem, and, second, the need for adaptively, since the structure of the problem may change. For these two main reasons, it is impossible write down a program to deal specific problems.
Regarding the complexity in describing methods to reach a solution, the issue is articulated in two main aspects.

1. There are activities that are very difficult to be formalized and expressed through a rule-based model. "*This inability to articulate our knowledge meant we couldn't automate these tasks*"[21] Examples of such tasks are driving, speech recognition and image understanding. For them the unique approach is an experience-based learning. To carry out these tasks is basically required the capacity of abstraction of rules and methods. An example is getting the difference between the two images in figure 4 and understanding that only the left one can represent a face. It is impossible to formulate a rule that can be applied "mechanically" it to identify a face just defining the position of all the elements within it. It is something that requires the understanding of the abstract concept of what a face is.

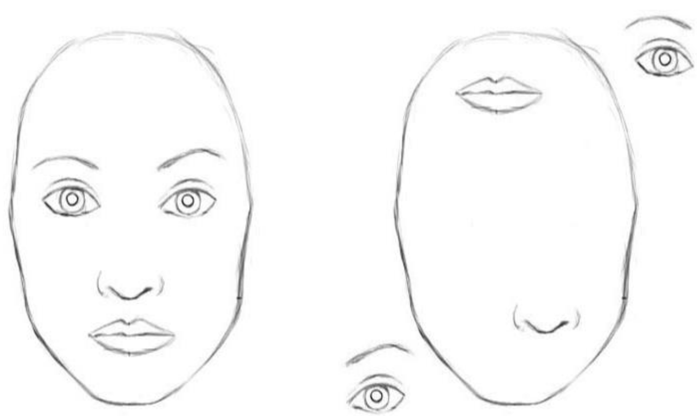


Figure 4: this figure represents two imagines. One is a human face the other image has the same elements of the first one but placed in a different and obviously wrong order. It has been used to show graphically the complexity to create a step-by-step rule to identify a face. Source: Understanding Hinton's Capsule Networks. Part I: Intuition. Max Pechyonkin, Medium

[21] *The business of Artificial Intelligence, what it can-and cannot-do for your organization, Harvard Business Review, Erik Brynjolsson and Andrew Mc Afee*

2. There are also tasks that goes beyond human capabilities. For example, analysing and detecting very large dataset of data and figuring out immediately insights and possible patterns, is something that is too complicated to us to do, and so, even almost impossible to be explained through a rule. In this field, machines are required to work completely autonomously.

Concerning adaptation, one limiting feature of programmed tools is their rigidity. Once the program has been written down and installed, it stays unchanged. However, it may happen that the tasks that the system is required to perform can change over time. Machine learning tools are, by nature, adaptive to changes in the environment they interact with, because, as said before, they have not to follow a code. Examples of successful applications of machine learning to such problems include programs that, for example, decode handwritten text, where they can adapt to variations between the handwriting of different users.

In this first part, Machine Learning has been introduced as general discipline pointing out for which reasons it has been developed and which are its main characteristics.
In the next paragraph, the different learning paradigms will be explained. Learning paradigms must be intended as the processes that make the machine learn.

1A.5.1 TYPES OF LERNING PARADIGMS

There are three main branches of machine learning, in function of the nature and of the feedback's generations' process.

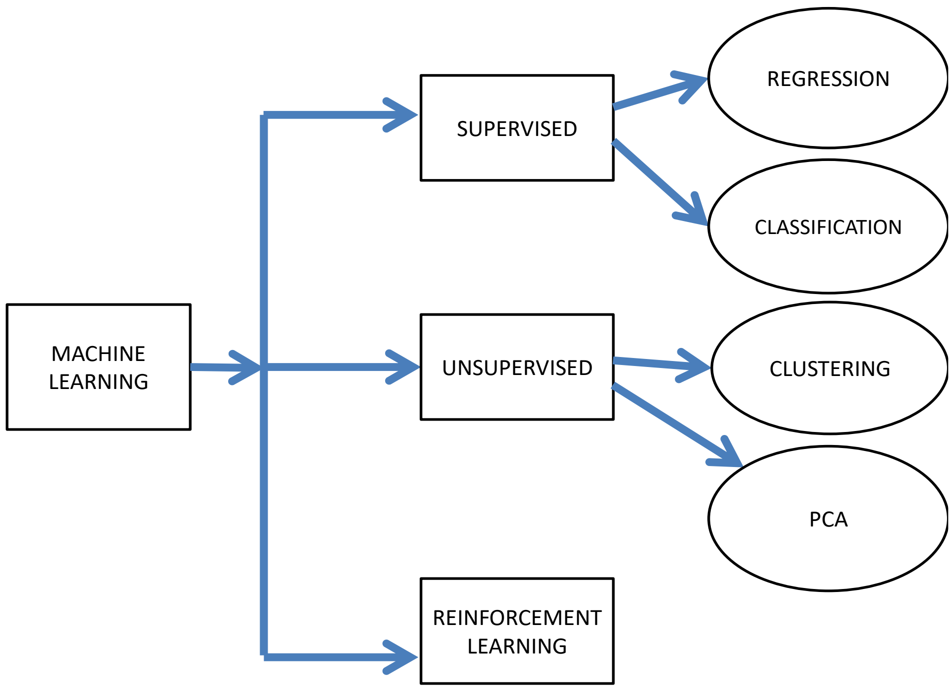


Figure 5: it is a scheme of the Machine Learning main branches, in function of the nature of feedbacks over which the systems are based. It shows also some algorithms examples belonging to each brach.

Supervised

A supervised learning approach is a methodology where the machine is trained with data that has been already labelled. The labels categorize each data pointing it, into one or more groups. An example, just to make it clearer can be categorize all the image between 'photo that contains an apple' from 'photo that contains an orange'. This activity of labelling input data is performed by human operators.

Then, the system looks at how this data are categorized, understand the differences and uses this knowledge to predict the category of new data, attaching to it a new label. So, in other words, the system *"on the basis of the training set, figures out a rule for labelling a new data"*[22].

[22] *Machine Learning, the power and promise of computers that learn by examples, The Royal Society*

Examples of supervised learning systems are the spam email filtering. The spam detecting usually "*do not require specifying rules to work. Instead, looking at the email categorized as spam by the user, it learns the rules that can classify effectively a new incoming email between spam and not spam"*[23].

To make an example, a type of rule that a spam detecting system can probably figure out from the labelled data (all the email categorized as spam by a given user) is that *"All the email that contains the word cheap and that have no title are spam"*[24]. In this way, algorithms 'learn' to predict outputs based on previous examples of relationships between input data and outputs.

Two classical example of Supervised Learning algorithms are:

- Regression: The system creates a curve that fit the distribution of the data, and that can be used to predict the value of one variable, given the value of another one.
- Classification: It is the example of the spam detection. Here, a database of labelled data is used to train the algorithm. This phase is properly called Training Period. The algorithm extracts rules that explain the classification and, then, it applies them to label a new object.

In figure 6, some real-world applications of supervised learning are presented.

INPUT	**OUTPUT**	**APPLICATION**
Voice Recognition	Transcript	Speech Recognition
Historical market data	Future market data	Trading bots
Photograph	Caption	Image tagging
Drug chemical properties	Treatment efficiency	Pharma R&D
Store transaction details	Is the transaction fraudulent?	Fraud detection
Recipe ingredients	Customers reviews	Food recommendation
Purchase histories	Future purchase behaviour	Customer retention
Car locations and speed	Traffic flow	Traffic lights
Faces	Names	Face recognition

[23] *Machine Learning, the power and promise of computers that learn by examples, The Royal Society*
[24] *Machine Learning, the power and promise of computers that learn by examples, The Royal Society*

Figure 6: the table represents some of real-world application of supervised learning. Source: The business of Artificial Intelligence, what it can-and cannot-do for your organization, Harvard Business Review, Erik Brynjolsson and Andrew Mc Afee

Unsupervised Learning

In unsupervised learning instead, there is no distinction between training data and test. Also, the desired output is unknown. It is called unsupervised because there are no already existing rules that the system can follow to carry out the task. There are no labelled data that can direct the work, and, for this reason, they are called unlabelled data. The systems are required to process the input data, with the goal of coming up with parameters that can categorize that set of data. The classical example of unsupervised learning algorithm is Clustering. The idea is that the system is given a big and raw set of data and it is required to recognize groups within it, basing on similarities. In this case there is no human interaction in the entire process. The algorithm carries out the task completely alone.

Another technique that belongs to Unsupervised Learning approach, is the Principal Component Analysis (PCA). It is a tool used to reduce the complexity of wide datasets. It is used to find the minimum number of uncorrelated variables that can explain the data.

Reinforcement Learning

In this case, the learning methodology is completely different respect the first two. In Reinforcement Learning there are no labeled or unlabeled data. The system, instead, can be considered as a sort of decision-making agent. Reinforcement learning is an approach in which the machine is not told how to carry out a task, but, instead, it knows how to evaluate the gains and/or losses associated with its actions and it is free to discover, autonomously, how to maximize the correspondent reward. In this case, the focus is not on the working methodology but on the reward/punishment framework, that can make the system autonomously judge how good is an output. The system will act basing on its ability to evaluate the different options available.

The reinforcement learning is a paradigm that can effectively overcome the complexity issues of the Polanyi's Paradox. In fact, since our knowledge in performing some tasks is not expressible through a well-defined set of rules, it is impossible to code it. So, the solution has been thought the system the objectives of the tasks and let it free to autonomously choose the option that can maximize the reward. An additional aspect is that "*usually the actions, may not only affect the immediate reward, but also the next situation and, through that, all the subsequent situations*"[25]. For this reason, the reward structure is not limited to the short-term, but it also considers the long-term future consequences of every single action of the system.

In searching for the option that maximize the reward, the reinforcement learning system face a trade-off between two concepts that from this point ahead will be present in all the AI's applications. The concepts are 1. <u>Exploration</u> and 2. <u>Exploitation</u>. To get rewards an agent will tend to prefer actions that it has already tried in the past in similar situation, and found to be effective in producing rewards (exploitation), but, to improve the performances, it has to try actions that it has not selected before (exploration). The agent has to exploit what it has already experienced but, it also has to explore new options to make better actions in the future.

To conclude, it is possible to highlight that Reinforcement Learning has represented a considerable improvement for machine learning, because it introduced the idea of experience-driven decision-making processes.

[25] *Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. Vol. 1. No. 1. Cambridge: MIT press, 1998.*

1A.6 NEURAL NETWORK: A NEW PARADIGM FOR MACHINE LEARNING

The Artificial Neural Networks (ANNs) is a new paradigm for implement machine learning solutions. it refers to the conceptual structure of the algorithm, that results to be better performing in case of complex systems with numerous variables that have reciprocal effect on each other.

Artificial Neural Networks find inspiration in the brain's structure[26], that consist of many basic computing devices (neurons) that are connected to each other in a complex communication network, through which the brain is able to carry out highly complex computations.

ANN can be described, as showed in figure 7, as a graph whose nodes correspond to neurons, and, edges correspond to links between them. Each single neuron is modelled as a non-linear function which transform a set of input variables in an output variable[27]. Each edge in the graph links the output of some neurons to the input of another one. The input of a neuron might be obtained in a very simple way by taking a weighted sum of the outputs of all the neurons to its incoming edges[28], or, it can be the outcome of a more complex mathematical operation.

ANN create this alternative computational paradigm, in which the solution is basically achieved, decomposing the initial problem into its simpler causal links. This is, conceptually, what differentiate ANN from all the others Machine Learning approaches.

---

[26] *A logical calculus of the ideas immanent to nervous activity ", Warren McCulloch & Walter Pitts, 1943*
[27] *A logical calculus of the ideas immanent to nervous activity ", Warren McCulloch & Walter Pitts, 1943*
[28] *Understanding Machine Learning, c 2014 by Shai Shalev-Shwartz and Shai Ben-David*

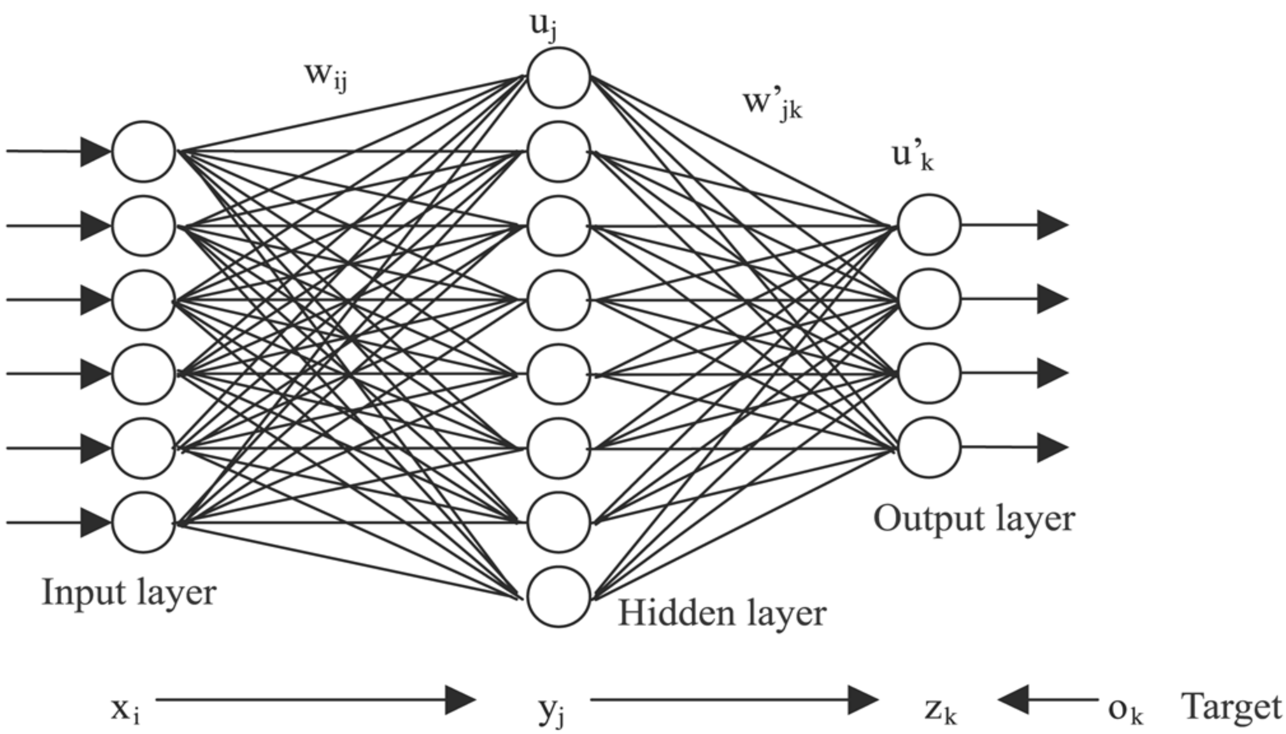


Figure 7: The figure represents the logical functioning of Neural networks algorithms. It shows three different layers, the links between the neurons of the Network and how the back-propagation algoritm works to train the Network.

This framework is revolutionary because it expresses no more the problem with an input-output approach, but instead, it uses a model that is input- (numerous variables) - output. Introducing these steps between inputs and output, it is possible for the system to discover and manage complex problems, with multiple interdependences.

It is possible to assume that the network is organized in layers. The calculation happens in a layer by layer manner. The system calculates the outputs of the neurons at layer t + 1 and then it passes to layer t and so on. Except for the first and the last one, the Layers are often called hidden layers. The "depth" of the network refers to the number of layers in the network itself.

As said before, a neuron takes in input the output of all the neurons of its incoming edges. To each linking edge is associated a weight, Wk (that can be positive, negative or zero), that, all together, represent the parameters of the model. They basically explained how relevant a particular input variable respect to the output is. These weights are learned and adjusted during the training phase, to get the required output. They are not defined by human operator a priori.

An example of how the weights can be set is The Back-propagation Algorithm (Rumelhart and McClelland, 1986). The idea is that, while the artificial neurons send their signals “forward”, the errors are propagated backwards. The back-propagation algorithm uses supervised learning, which means that the algorithm is provided with examples of the outputs the network is expected to compute, and then the error (difference between actual and expected results) is calculated. The idea of the back-propagation algorithm is minimizing this error find the most suitable weights’ values. As the error flow moves back, it changes the weights between the neurons. This operation is called training the network.

1A.7 THE EVOLUTION OF THE TRAINING: EVOLUTIONARY ALGORITHMS

It is possible to consider the Back Propagation as the traditional way to train ANNs. Nowadays there are also more advanced solutions to train networks and to make them adaptive to possible changes in inputs’ nature. An improvement is represented by Evolutionary Algorithms. They basically mimic the process of learning how to adapt to the environment of living being: evolving from one generation to another. The development of these typologies of algorithms have fundamentally introduced the “survival of the fittest” principle, in the designing and optimizing processes of algorithms. They have three main characteristics:

- Population-based: they maintain a group of solutions, called populations, to face at the same time the problem in parallel way. So more than just one type of algorithm that work on the same issue.
- Fitness-oriented: every solution in a population is called individual, every individual is evaluated and only the bests, the fittest ones, will be held for the second generation.
- Variation-driven: individual will undergo a few variations in the way they operate, to find if better solutions exist[29].

[29] *Introduction to Evolutionary Algorithms, Xinjie Yu, Mitsuo Gen, 2010*

The system is now intended as a group of algorithms that works in parallel. Each of them tries to optimize its own performance in solving a problem, then, they are supposed to change partially their structure to search for better way to face the problem, with the goal to make emerge, at the end, the fittest solution with the environment. This model assures also a good adaptation to mutations in the environment itself.

1A.8 DEEP LEARNING: THE TOP PERFORMING NEURAL NETWORKS

Deep Learning is that specific field of Machine Learning that, as his name can suggest, works with very deep networks, so with systems composed by many layers. This conceptual framework permits Deep Learning algorithms to leverage on representation-based learning methods.

This means that each layer transforms the representation, the information, received in input into a higher-level representation in output, slightly more abstract. "*An image, for example, comes in the form of an array of pixel values. In the first layer of representation [the system identifies] the presence or the absence of edges at a particular location of the image. The second layer detects motifs by spotting arrangements of the edges, regardless of the small variations in the edge positions. The third layer may assemble motifs into larger combinations that corresponds to parts of familiar objects, and the subsequent layer would detect objects at combination of these parts*"[30].

[30] *http://www.cedar.buffalo.edu/~srihari/CSE574/Chap5/Chap5.8-DeepLearning.pdf*

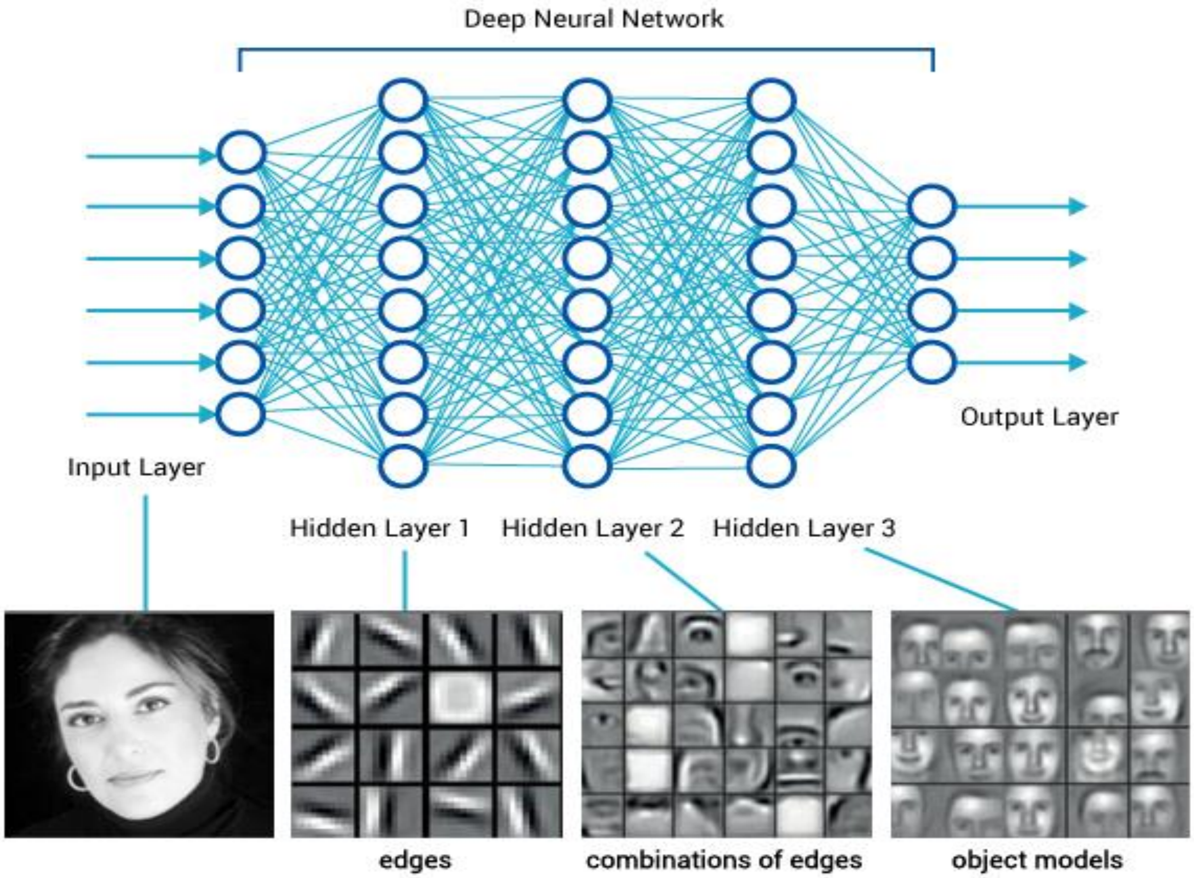


Figure 8: The figure represents how Deep Learning algorithms works in the field of Image Recognition. And how it can get to a solution leveraging on the recognition of macro-features present in the image itself.

So, it has been said that Deep Learning aims to express complex representations in term of other simpler ones, enabling to disentangle the so called “Factor of Variation”. The variation factor refers to the fact that in real world there are a lot of elements that affect the final output’s representation, for example, speaking about image identification, the fact that the shape of a face depends on the viewing angle or on the lights.

The systems must be able to manage them and to not get fooled. To deal with these complex tasks it is not possible to use a microscopic approach, using a network’s that analyse every single pixel in input as an independent unit, searching for pattern that can bring back to an already-known subject, because, for example, two images can be totally different in terms of colours, lights and intensity but they can show the same subject. To cope this factor of variation, machines need to think in terms of high-level features recognition and building up the output, leveraging on these abstract notions. This is exactly how Convolution Neural Networks (CNNs), a class of Deep Learning algorithms, work. They are widely used in image recognition field and, instead of analysing singularly each unit, they divide the whole pixel array in input in sub-regions, in areas of space, that will be inspected to find certain features. The final output will be the result of the comprehensive analysis of all these features and how they are placed respect to each other. CNNs, detecting macro-features within the image, develop a representation that makes object information, increasingly explicit along the process[31]. A graphical representation of how they work is expressed in figure 9.

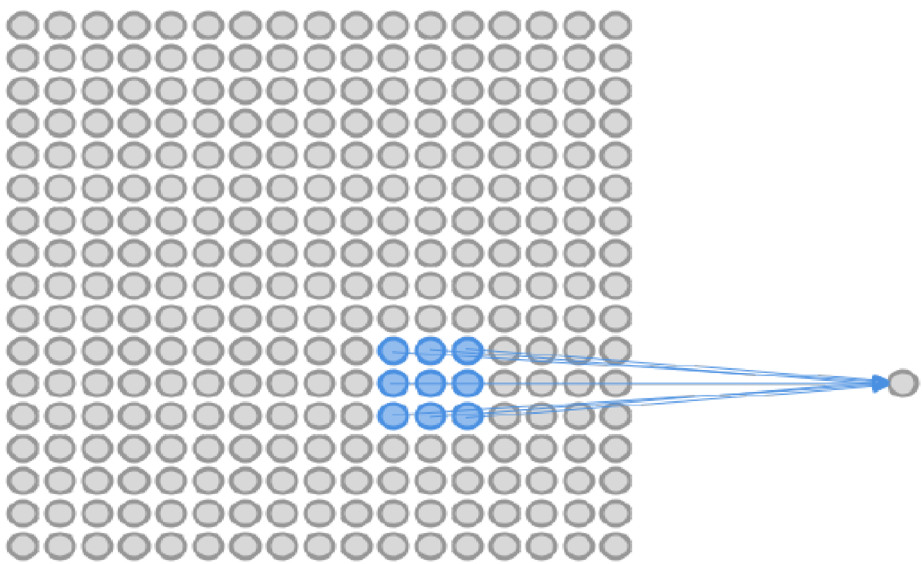


Figure 9: Convolution Neural Network working methodology. It is based on localized connection of an area of input to a unique neuron in output. Source: Deep Learning Quantification, Master Graduation Thesis by Andrea Azzini, Giovanni Battista Conserva, Politecnico di Milano, Academic Year 2015-2016

A different example of Deep Learning algorithm are the pooling layers. They are used to speed the computation, reducing the size of analysis. The idea behind Pooling is leaving out of the analysis some information after that features are detected. Once recognized a feature, the system saves just an approximation of its characteristics, leaving out from the network the detailed information.

An example of how Pooling algorithms works is showed in the image below. It represents the functioning process of the Max-Pooling algorithm. To reduce the dimension of the analysis, it divided

[31] *Texture synthesis and the controlled generation of natural stimuli using convolutional neural networks. Gatys, L. A., Ecker, A. S. & Bethge, M.*

all the input in sub-regions and will pass to the next level a smaller representation, made off the highest number in that specific sub-region.

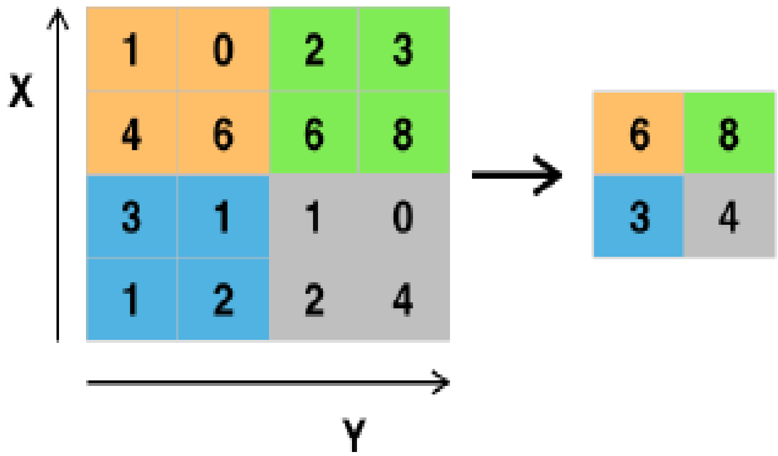


Figure 10: the function of a max-pooling layers is showed. The idea is that for every sub-area just the maximum number will be kept and passed the next layer. Source: Deep Learning Quantification, Master Graduation Thesis by Andrea Azzini, Giovanni Battista Conserva, Politecnico di Milano, Academic Year 2015-2016

Thanks to the methodologies explained before and also to more complex and innovative algorithm classes, Deep Learning application deal with problems leveraging on a hierarchical system of representation, where each level (that are the layers on the network), transform the concept of the underlying level in more abstract one. This permits to disentangle the Variation factors that are present in the real-world and focusing on the essence, leaving out all the “Noise”.

Now, to give a real-world example of the extraordinary performances of current Deep Learning algorithms and, to show at which level of understanding and abstraction they can yearn for the resolution of complex problems, the work developed in the paper *“A neural algorithm of artistic style*” by Leon A. Gatys, Alexander S. Ecker and Matthias Bethge will be briefly presented. The objective of this Deep Learning based system is basically understanding the artistic style of a given painting and reproducing it. It is able to manage the difference between the content of image, from the artistic style used to representing it. An idea of how it works is showed in the figure 11. Where the system has created new works, that combine the style of several well-known paintings, with the content of an arbitrarily chosen photograph. This system is able to manage the abstract concepts of the style of a painters and transfer it in another context, creating a new one. The figure A is a common photo of small village. All the other images represent the output of the system. The smaller figures at the lower left side is the painting from which the system has extracted the style to develop the image itself. This is just an example of the results, in art field, that a representation-based learning approach can reach.

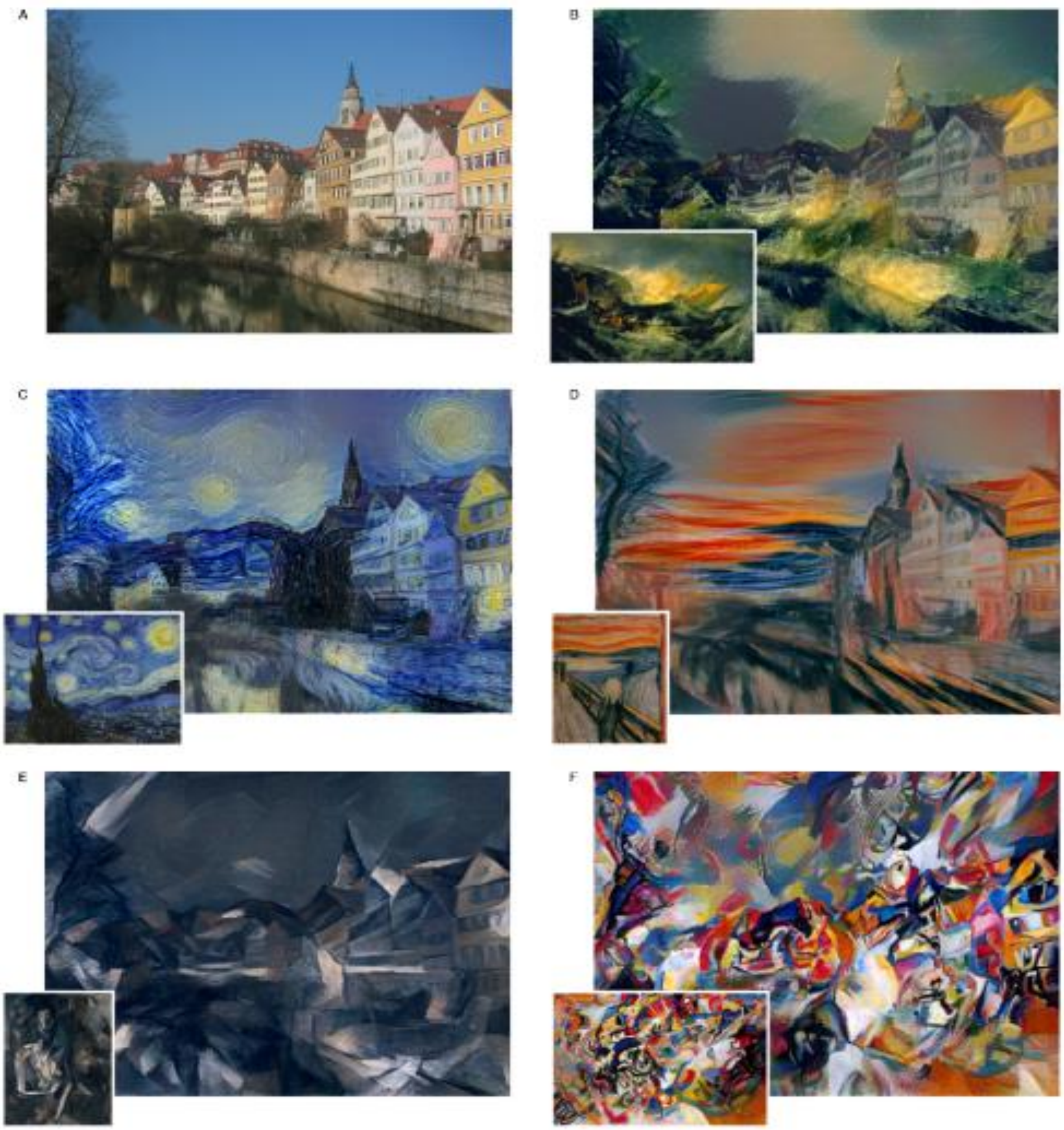


Figure 11: the image shows the functioning of the Deep Learning system presented in "A neural algorithm of artistic style" by Leon A. Gatys, Alexander S. Ecker and Matthias Bethge. The image combines the content of a photograph with the style of several well-known artworks. The big images are created by matching the content of the photograph and the style representation of the artwork. The original photograph depicting the Neckarfront in T¨ubingen, Germany, is shown in A (Photo: Andreas

Praefcke). The painting that provided the style for the respective generated image is shown in the bottom left corner of each panel. B The Shipwreck of the Minotaur by J.M.W. Turner, 1805. C The Starry Night by Vincent van Gogh, 1889. D Der Schrei by Edvard Munch, 1893. E Femme nue assise by Pablo Picasso, 1910. F Composition VII by Wassily Kandinsky, 1913.Source: A Neural Algorithm of Artistic Style, Leon A. Gatys, Alexander S. Ecker, Matthias Bethge

1A.9 RISKS AND LIMITATIONS OF ARTIFICIAL INTELLIGENCE SYSTEMS

Even if machine learning systems enabled many advances, they are subject to some limitations, that is useful to highlight.

- Supervised systems rely on the accessibility of significant amounts of labelled training data, whose creation can be resource-intensive and time-consuming.
- It is difficult to develop systems with contextual understanding of a problem, or, in other words, with "common sense". Machine Learning algorithms make assumptions about the 'best' function that fits the data. Thus, it is possible to have the situation there are more than just one function that can fit properly our set of data, and some of them, can be spurious and have no logical sense to exist. A logical evaluation usually is missing in these machines. This can bring to two main problems: Over-fitting and Under-fitting. Over-fitting means that the model uses complex hypothesis and focuses on irrelevant factors in the training set, limiting the ability of generalizing when faced with new data. Under-fitting means that the model only considers simple hypothesis and therefore excludes the real relationships and is not able to generate the real value from data[32].
- Problems with "Interpretability", meaning that, sometimes, humans have difficulties figuring out how the systems reached their decisions. In particular, machines lack the ability to "explain" rationally the reason of their actions. So, we are facing a sort of inverse Polanyi's Paradox, in which the machine "*knows more than they can explain*"[33]. This problem is amplified when the machine is

[32] *Machines that learn in the world, Machine learning capabilities, limitations and implications, Nesta, July 2015*
[33] *The business of Artificial Intelligence, what it can-and cannot-do for your organization, Harvard*

working with hidden biases like over-fitting, under-fitting or spurious or wrong relationship. "*For instance, if a system learns which job applicants to accept for an interview by using data made by humans in the past, it may inadvertently learn to perpetuate their racial, gender, ethnic or other biases*"[34]. This problem is mainly related with Deep Learning applications that usually works as a black box, with several level of abstractions. In these cases, it is almost impossible to make clear the reasons for which the systems give a certain output.

1A.10 AI PERFORMANCE COMPARED TO HUMANS

Since now it has been presented the potential of artificial intelligence. Some artificial intelligence systems can outperform humans in many tasks, but in others, humans' performances are still higher. This paragraph wants to roughly compare the performances of AI and humans. In this field, a general rule has emerged about the improvements of AI systems. At the beginning usually, AI starts well below the human level, but, once AI finds the proper logical structure to represent the problem, it can reach our performances, and then, easily overpass them[35]. The difference is that computer systems are equipped with self-learning methods able to reduce the errors in a systematic way (like the minimization of a loss function). We, usually, lack the possibility of continuous improvements. This means that once reached a certain threshold, our performance cannot increase more due to natural limits. In fact, in the next graphs Humans' performances are presented a straight line.
In the following part, three arbitrary chosen case will be presented. One about a task where AI outperformed humans, one where it just reached humans' level, and one where mankind is still better.

The figure 12 illustrates the historical evolution of image recognition performances. Here computer systems have already overpassed the human benchmark. The data are taken basing on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), a competition that evaluates algorithms for object detection and image classification at large scale.

Business Review, Erik Brynjolsson and Andrew Mc Afee

[34] *The business of Artificial Intelligence, what it can-and cannot-do for your organization, Harvard Business Review, Erik Brynjolsson and Andrew Mc Afee*

[35] *https://aiindex.org/*

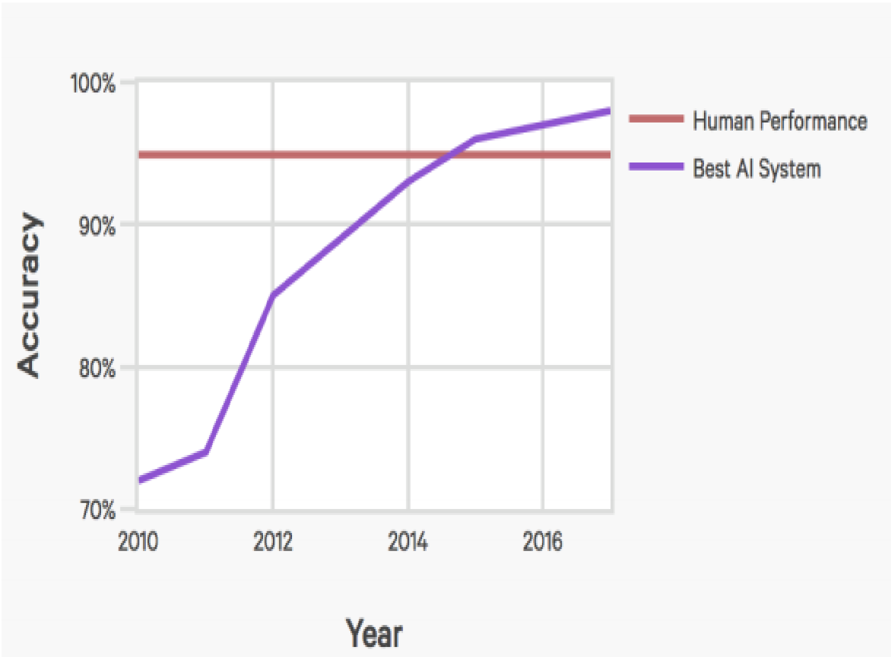


Figure 12: the image represents human performance in image recognition task compered with the best AI system. It also show the historic trend of evolution. Source: Olga Russakovsky, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution)* ***ImageNet Large Scale Visual Recognition Challenge****. IJCV, 2015*

Instead in the field of Speech Recognition, the performance of AI systems just reached human level. Speech recognition is measured as recognizing speech from call audio, humans and algorithms are asked to listen phone calls and transcribe conversation into text[36]. (Figure 13)

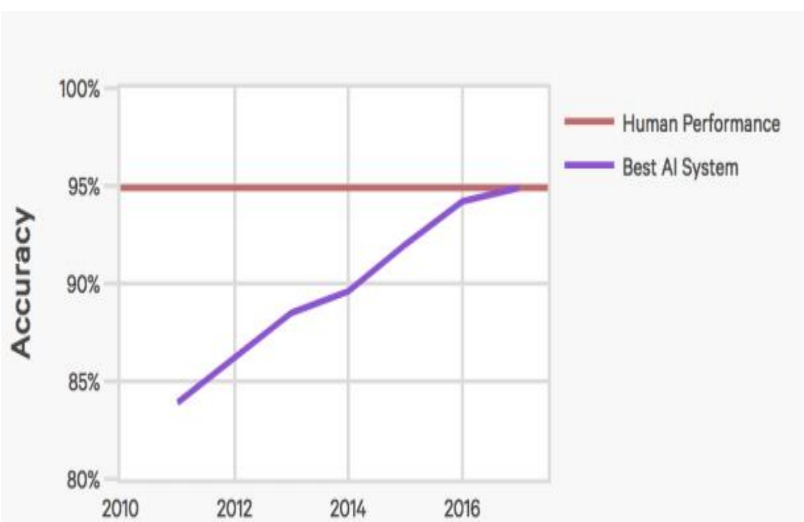


Figure 13: the image represents the human performance in Speech recognition compared to the best AI system. Source: Lingustic Data Consortioum. 2000 HUB5 Englidh Evaluation Speech. Philadelphia: Linguistic Data Consortium. 2002

[36] *https://catalog.ldc.upenn.edu/LDC2002S09*

The last case that is presented shows a situation in which humans still outperformers AI systems. It is about question answering. The test used the Stanford Question Answering Dataset (SQuAD) to make humans and algorithms competing. It is a reading comprehension test, consisting of questions posed about articles. Given a question about the content of an article, the task is about identifing the answer within the text.[37] (Figure 14)

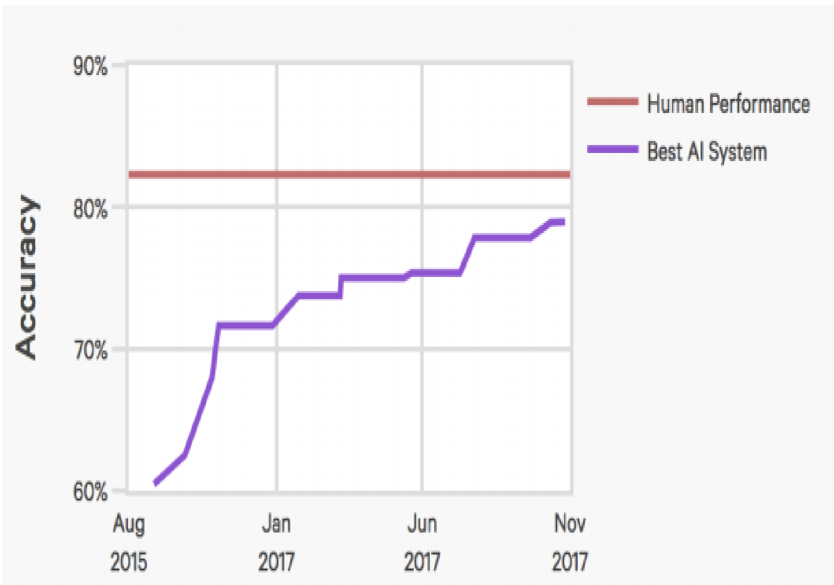


Figure 14: question answering performance between humans and best AI systems. Source: stanford-qa.com

1A.11 FIRST PART'S END

Now the first part about the AI's introduction as general topic can be considered finished. In this brief argumentation the objectives and the most important features of artificial intelligence solutions have been proposed, along with the historical developments of AI. Also, the most promising solutions have been presented with the results they can achieve.

The digression of the following paragraph must be considered as the conceptual bridge between the theoretic dissertation about the AI and the sport environment. It will be used to enter the sport topic through the description of two of the best examples of AI. The first is about Deep Blue, a chess player developed by IBM, and the second, is AlphaGo, a system able to play the game of Go, developed by

37 https://rajpurkar.github.io/SQuAD-explorer/

Google DeepMind. Through these two real-world examples, some of the solutions introduced before will be presented operationally. They will be also useful to contextualize, in practical terms, their use.

1A.12 TWO HISTORIC EPISODES: DEEPBLUE VS KASPAROV AND ALPHAGO VS LEE SEDOL

Deep Blue was a chess-playing computer developed by IBM. In 1997, it defeated on a six-game match the world champion Garry Kasparov, becoming the first computer to defeat a reigning world champion in a match, under standard chess tournament rules[38].
AlphaGo, instead, is a software that plays the game of Go. It was developed by Google DeepMind. In 2015 it become the first computer to beat a human professional Go player and in 2016 it beat one of the best human players in the world.

Between these two systems there are interesting common points, but also profound differences, in how they reached these remarkable results.

Deep Blue was a specialized, purpose-built computer, capable of examining 200 million moves per second. This means that in the three minute, granted by regulation, available for examining a position it could weight 36 billion different moves[39]. Each move was evaluated using a reward function that tried also to consider the future consequences of that specific move on the board, some moves ahead.
The machine, then, toted up the expected value of each one and chose the highest, making the move. In his working methodology it is possible to recognize a Reinforcement Learning system. In this case the approach is called properly “Brute Force Approach” just because it deals with much more complexity, but conceptually, the framework is the same. For Deep Blue the effectiveness of his next moves depended mainly of the reward function (instructed by engineers) about how to evaluate correctly the expected gain associated with a specific move.

---

[38] *https://en.wikipedia.org/wiki/Deep_Blue_(chess_computer)*
[39] *Forbes, The brute Force of IBM Deep Blue And Google DeepMind, Gil Press, Feb 7, 2018*

Deep Blue won against Kasparov because it was able to evaluate properly the large majorities of its moves. Deep Blue was said to be able to forecast every possible board configuration of the next 6 or 7 moves and to compute the associated cumulative utility of that specific path. "*Up to a certain horizon, it was omniscient"*[40]*.*

Sometimes, Deep Blue succeeded in doing moves that were considered unusual and strange, that later resulted in unexpected gains. These moves were the output of two possible events: 1. The computer figured out something that was difficult to see for human thanks to his computation power. 2. A bug or in any case a failure of the system. In fact, if unable to select a move, Deep Blue, was programmed by default to pick a play completely at random[41].

As highlight by Murray Campbell, one of the chief engineers of the Deep Blue project,[42] sometimes it was hard even for the engineers' team to figure out if the moves were the results of the first or the second option. A real-world example of what has been presented before in 10A.6 as Inverse Polanyi's Paradox.

AlphaGo, instead, was the first program to defeat a Go world champion. The game of Go *"has a whole other level of complexity [respect to chess] …the number of possible positions on a Go board are [endlessly higher] …the upshot is that, unlike in chess, players, whether human or machine, can't [try to forecast] the outcome of potential move*"[43].

This made evident that a "Brutal Approach", developed to analyse all the possible outcomes was not possible in this environment. For this reason, the researchers needed a way to make the software analysing only the moves that could be more prominent and successful and do not dedicate computational power to analyse moves that are obviously meaningless.

The object is to mimic in the computer the intuitive understanding that a professional player has, looking at the board and focusing just on a specific set of moves and, select between them, the best

---

[40] WeeklyStandard, Be Afraid, Charles Krauthammer, 1997

[41] The Signal and the Noise: Why Most Predictions Fail – but Some Don't, Nate SIlver

[42] The Signal and the Noise: Why Most Predictions Fail – but Some Don't, Nate SIlver

[43] What the AI behind alphago can teach us about being human, Edward c. monaghan, wired

without considering all the others. AlphaGo was able to do exactly this thing. It relied on two different processes: a tree search and a mechanism to guide the search, to reduce the effective depth and breadth of the search tree[44]. To understand which moves could be the best one, the machines faced the trade-off between exploration and exploitation, explained in 10A.5, when reinforcement learning has been introduced.

Regarding exploitation, it used experience to identify the set of the best possible moves in that specific situation. The machine was able to do it because it had analysed and memorized millions of games of professional players. Machine was basically able to "remember" or, in other words, compute statistically which move an average professional player would have done in the same situation.

This learning approach is called "Imitation Learning" and it makes machines automatically learn from good policy from observed behaviour.

This lead to limit the breath of the analysis, ignoring last majority of the possible theoretical moves available. To limit the depth of the analysis instead the machine estimates, basing on experience, the probability of victory associated to that specific branch of the tree, basing on experience. The idea is that the machine does not go to analyse all the possible future developments that a move can generate but instead, basing on past cases, it computed the expected outcome related to that choice.

Instead, regarding the exploration: the machine was able to evaluate the possible gain associated with moves that, statistically, were not often used, but that could lead to higher reward. The effectiveness of the software lies in the capacity to manage and make "profitable" this trade-off.

So, to recap, first, the AlphaGo selects the set of moves that are more likely to bring to end victory basing on experience and then, comparing them with some outliers, it selects the best one.

Thanks to this mechanism "*AlphaGo evaluated thousands of time fewer positions than Deep Blue did in its match against Kasparov, compensating by selecting those positions more intelligently… and evaluating them more precisely…Furthermore, while Deep Blue relied on a handcrafted evaluation function, AlphaGo is trained directly from gameplay, through learning methods*"[45].

---

[44] *Mastering the game of Go with Deep neural networks and tree search, Nature, David Silvers, Aja Huang and others*

[45] *Mastering the game of Go with Deep neural networks and tree search, Nature, David Silvers, Aja Huang and others*

This digression gives also the opportunity to exploit additional insights from the first part. Now, it is possible to derive a high-level example of the debate between “Turing Test” and the “Chinese Room Argument”. Both of the machine would have passed the “Turing Test”, but, following what Searle said, AlphaGo showed a higher “level of intelligence”.

Second, these machines provide two examples to understand and consolidate what “Act Rationally” means, to enter, later on, the second part of the thesis.

Just to conclude, it is necessary to highlight that Deep Blue came twenty years early respect to AlphaGo, so no direct comparison is allowed. As said before, when the “Odd Paradox” has been introduced in 1A.1, when the AI frontiers goes ahead, also our idea of AI evolves with them.

AlphaGo can be considered as an evolution respect to Deep Blue, that learnt how to manage in a more efficient way its computation power leveraging on the experience value, showing a sort of awareness of what it is going on in the game.

1.B SPORT BUSINESS INTRODUCTION

1B.1 FOUR MACRO-AREAS BASED MODEL

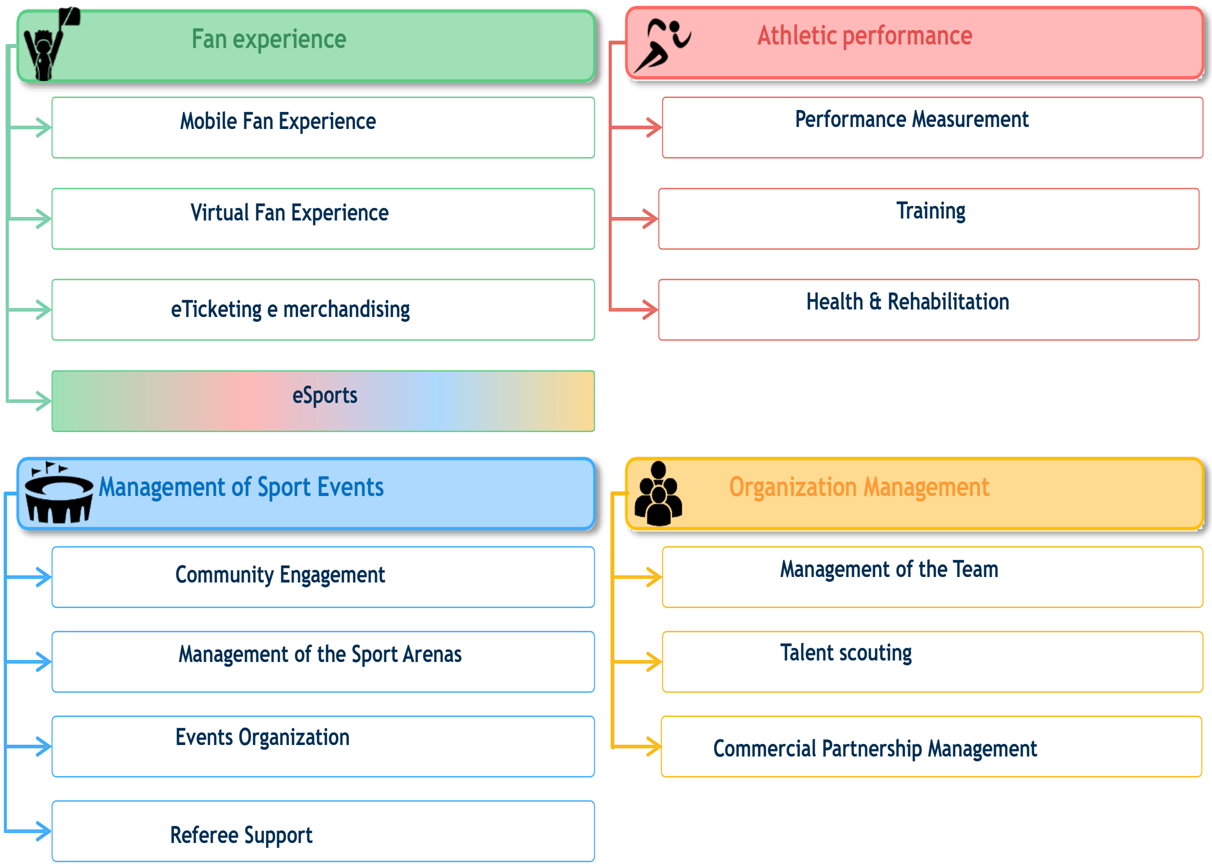


Figure 15: this scheme has been developed and created by the Osservatorio Innovazione Digitale nell'Industria dello Sport of Politecnico di Milano to map the application fields of technological innovation in sport business.

This section of the thesis will be used to introduce the current world of sport business from the perspective of a generic team. The model employed to perform this task has been developed by the Osservatorio Innovazione Digitale nell'Industria dello Sport of Politecnico di Milano (figure 15). The model has been used to map the areas where digital technologies can be introduced in sport. In this work, it will be used to try to draw a clear big picture of the main trends that are characterising sport business.

The model divides the processes that can be affected by digital innovation in 4 main areas:

1. <u>Fan Experience</u>
2. <u>Athletic Performance</u>
3. <u>Management of Sport Events</u>
4. <u>Organization Management</u>

The peculiarity of sport business is that all these 4 areas are completely different form one to another. They have different drivers, objectives and characteristics. The difficulties are related with managing all of them coherently and in an efficient and effective way.

Just to contextualize the sport industry with some data and macro-trends, it is possible to assert that it is a fast-growing industry. It has reached in 2017 a global value of US $ 700 bn, equal to 1% of the global GDP[46]. Sports tourism is the fastest growing sector in global travel industry according to the World Travel Organisation, as, for example, football brings on average 800,000 visitors to Britain every year[47], that represents a big economic opportunity. Speaking about Professional teams' sport, clubs are extending their market presence and size, engaging with fans world-wide and entering new business' areas that are making them at the same time entertainment, media, real-estate and even tech companies. Thus, now there are increasingly more stakeholders in the sport business than ever. Fans, sponsors, investors, local communities, media broadcaster, commercial partners expect to be involved in club management processes. From this, the necessity to look at sport with business methodologies aimed to maximize the value creation of every single decision.

[46] *Sport Tech Innovation in the start-up nation, Deliotte Report, March 2017*
[47] *The future of Sport fan, Performance Communication and Canvas8*

1B.1.1 FAN EXPERIENCE

Fan engagement is the central element in modern sport industry. Fans are the final clients of the products and services offered by the teams and, at the same time, they are the target of media companies and sponsors' initiatives. A higher fan base permits a higher sponsors' brand exposure that can be monetized from sport organization. So, fans are the drivers of all the revenues in sport environment.

The relationship between sports' clubs and fans has changed impressively in the last years. We are witnessing to a situation in which clubs have their more numerous fan-base geographically placed far away from the city of the team itself. A stunning example is the situation of Manchester united, that have more than 650 million fans worldwide and just less than 1% come from Manchester.[48]
The fan-base of sports' club in not just geographically changed but it also embraces new fans' segments, reaching categories that in the past has been marginalized, like women in football who represents the most-growing fans category.[49]
As in every other human context, diversity and differences are, at the same time, great opportunity for growth but they also represent challenges. Since every different group of fans has unique feature and, thus, it must be engaged in a unique way, the sport' organization should be agile in adapting and shaping their offers for each of them.
Establishing well-functioning relationships with all these different fans' segments and in general with the club's stakeholder is not complex just from a managerial point of view, but, it can also create real tensions and frictions. Real-world examples of this is what happened in 2014, when Bank of Abu Dhabi and Real Madrid made a deal that impose the Club to remove the Christian cross from their badge. This decision created dissatisfaction in a part of fans, or the protests that raised in 2011 when Barcelona decided to add for the first time in history a t-shirt sponsor. Also in this case, part of the fan base protested the decision seen as a misalignment with the tradition. Creating dissatisfaction in the fan base is absolutely not in the clubs' interests. Even if we will not enter the cases, we can just takeaway that broadening the fan base, create managerial but also relationship-based complexity, and clubs must

---

[48] *The future of Sport fan, Performance Communication and Canvas8*
[49] *The future of Sport fan, Performance Communication and Canvas8*

succeed in avoiding any type of conflicts between two stakeholders' category because it will damage the club's interests.

Respect to digital innovation, the fan experience can be divided into two main flows: 1 At home fan experience and 2. Live. In this section, it will be presented home experience, leaving the argumentation of in-stadium operations to the following one. Home experience have two declinations, one physical and one temporal. The word home experience is used to address the fact that, a fan can be geographically far away from where the match is happening, but he still wants to be engaged actively. This is the geographical connotation. The temporal one refers to the fact that fan must be engaged 24/7 even if no official events are happening.

Regarding home experience, the digital channel is, for its very nature, the most appropriate one to carry on the relationship. Club need to create personalized marketing strategies, differentiating a "commodity experience", like watching a match at home on TV, in an inclusive experience, that could permit to every fan cluster to live it as they most like, leveraging on different additional services. This phenomenon of differentiation of the experiences will be also present in the Arenas Management section, where we will see that the trend is creating several different sectors within stadium, each of them tailored-made for different categories of fans with the aim of enhancing the experience's value. Sport home experience is mainly related to the delivery of media content to fans. Now there are different media channels that clubs use to engage fans:

- On-line streaming
- Mobile app
- 3D & VR
- Interactive Platform
- TV media[50]

[50] *Sports: the most disrupted of all industries? Pwc's Sport Survey 2017, Pwc Report*

To contextualize, Media rights and related sponsorship account for more than 70% of sport revenues. "*Globally, 127,000 hours of sports programs are available on TV (up 160% since 2005) and more than 30 billion hours are spent viewing sports annually. In 2015, 93 of the 100 most viewed live programs were sports events*"[51].

In the next part, all the media channels and digital touch-points between clubs and fans are clustered in two main categories, Mobile and Virtual, in relation to the service offered. Mobile refers mainly to exchanging data, information and media content in the traditional digital way, through an app for example, while Virtual refers to the creation of new experience thanks to technologies like VR and AR.

Mobile Fan Experience

It is now estimated that smart-phones' ownership in the United States has reached 72% of the population, and in Europe it is estimated to be around 68%[52]. So, every single person that potentially can be considered a sport fan has a mobile device with him/her.

The objective is the make the fan use their devices to stay in touch h24 with their teams. Some successful examples of initiatives will be presented to clarify what fans appreciate. The Leitmotiv is turn fans' behavior from passive to active. In our mind, being a sport fan is something mainly passive. The unique exception is when fans in stadium cheering have the possibility to affect players performance, but looking at sport at home, the behavior is completely passive. With mobile app, clubs can change this situation.

For example, the NFL has introduced the possibility for fans to vote about several aspects of the games. They now collect 35 million fan votes every day about team and players' performance, and, fans have even the opportunity to pose press conference questions[53].

Probably the best example of making fans active is what has been proposed by Formula E' FanBoost. This application permits fan to literally boost three drivers voting through a mobile device. "*The three*

---

[51] *Sport Tech Innovation in the start-up nation, Deloitte Report, March 2017*
[52] *Sport Tech Innovation in the start-up nation, Deloitte Report, March 2017*
[53] *The future of Sport fan, Performance Communication and Canvas8*

drivers with the most votes are awarded a significant burst of power, which they can deploy during the second half of the race."[54] This represents a completely change in the concept of fans-athletes relationship, because, in this way, fans are becoming an integral part of the game. As highlighted by Luca Colajanni, the media delegate for Formula E "The *idea is controversial and maybe difficult to accept, but we need to learn from the success of the reality talent shows and make fans the protagonist"[55].*

This is another common point in current sport environment: team, athletes and organization are all looking at other industries and environment to take inspiration for innovations.

Mobiles application permits also to extend the life of the live-sport events, increasing the sponsors' brand exposure and the number of interactions with fans. In professional sports there is now an immense availability of data about every single aspects of performances. The idea is to use these data, to establish a connection with fans that starts before the game and finishes later respect to the end of the match. So, fan should be provided with information that can be meaningful to make them understand which type of performance will happen, and, after the match, to analyze like as real coaches, how the performance has gone.

Sports like rugby and cycling are already experimenting microchips in athlete gear to monitor movement, heart rate and fatigue. They're sharing this data to make the fan experience more entertaining[56].

As highlighted before the success of every activity depend on the ability to make the fan active. So not just providing data but use them in a meaningful way.

Examples in which sharing common players' stats with fans create an active relationship, are the app GAME golf[57], a system that allows professionals golf players to track their swing and share the stats

---

[54] *https://fanboost.fiaformulae.com/*

[55] *The future of Sport fan, Performance Communication and Canvas8*

[56] *The future of Sport fan, Performance Communication and Canvas8*

[57] *https://www.gamegolf.com/home/en-us?v=095483d*

with fans, and the social network Strava[58], that permits to amateur cyclers to compete with professional, retracing their routes and seeing how their times changed.

It is necessary to highlight that the tech giants are looking with interest to sport rights and broadcasting. Since their business model is extrapolate and deliver value through data, it is likely to expect that if they enter this market, the idea of providing always more data to fans and create new digital relationships between them and professionals' team will increase. In this sense, Facebook has been already used for live-stream sport events. Twitter won a $ 10 million deal to live-stream NFL's Thursday game and Amazon is planning to bid for Premier League Streaming Rights[59].
Also, several sport organizations are launching their own digital mobile channel to stream and create stronger relationships with fans. Examples are WWE Networks, F1, MLB and NBA.

Virtual Fan Experience

Virtual Reality (VR) is a "*computer-generated scenario that simulate experience through sense and perception*"[60] while Augmented Reality (AG) "*is the integration of digital information with user's environment in real time"*[61]. These two technologies have the potential to revolutionize the sport fan experience. From Goldman Sachs' researches, it is expected that the VR+AR market will worth $80 billion by 2025[62].

It is necessary to highlight that these two technologies are still in their initial phase of development, so it is necessary to distinguish between short-term and long-term potential.
What technology is currently able to deliver to a mass market are media contents, that can be highly personalized by the users, in terms of what fans are watching and which information they want to see.

---

[58] *https://www.strava.com/*
[59] *https://www.bloomberg.com/news/articles/2018-01-05/amazon-is-said-to-plan-bid-for-premier-league*
[60] *https://en.wikipedia.org/wiki/Virtual_reality*
[61] *https://en.wikipedia.org/wiki/Augmented_reality*
[62] *Global Media Sector Trends 2018, DLA PIPER INTERNATIONAL LLP*

The idea is making the fans completely autonomous and free to choose the best way to live the sport event from home, changing the perspective as they want.
An example of how clubs are planning to employ VR/AR in the short-term is, for example, what Aston Villa, an English football club, is considering to do. This club has a wide fan base in China and to engage it more actively, the club started to think about selling virtual seasons stadium tickets[63]. In this way, China-based fans can buy virtual tickets and look at the matches in an immersive environment like if they would have been really in the stadium.
What is stunning are the long-term application of these technologies. Thanks to advancements in technologies for capturing live actions like cameras, wearable and sensors and, at the same time, the developments in viewing technologies, media broadcaster will be able to deliver completely new type of contents, in the future.
The idea is that every single fan, at home or even in stadium, can have the visual perspective of players on the field. This result can be achieved thanks to advanced cameras systems, but, to make a tangible example the outcome could be looking at a match on TV as if the camera would be set on the front of a players. This will never happen for safety reasons but a very similar results can be obtained thanks to multiple cameras and 3D-triangulation operations.
The one just explained is one of the possible future developments of this technologies in the field of media broadcaster. Also in this case, the idea is creating more immersive experiences.

Similar applications are even more interesting for sport like cycling or motor racing, where one of the big limitation is that fans can see just a small part of the total performance. Introducing media system that enable fans to have the same view of the riders will create opportunities for the development of more interesting contents.

E-ticketing and Merchandising

Ticketing and merchandising represent two relevant sources of revenue for sport organizations. About ticketing we are witnessing to a trend that is showing the raising of sports' ticket price in every

[63] *The future of Sport fan, Performance Communication and Canvas8*

competition and in every country. In USA, the price of the cheapest ticket for live events have raised at twice the rate of cost of living since 2011[64]. On one side the fans are willing to pay always more to join live events for the high-level experience they can live, but on the other side, the fan base has broadened while the size of facilities has remained almost the same. These are the two main reasons behind this phenomenon.

In general terms a raising in the prices is profitable for the sellers, and it can generate positive return in the short-term for clubs. Different situation in the long-term, where constant high prices can make a substantial part of the fan base loosing interests in the clubs itself. In particular, the less-wealthy part of fans can feel left out, and they will abandon the club. This will jeopardize the future club growth, that mainly depends on the overall size of the fan base. Furthermore, it is necessary also to consider the social aspects of sport, and the fact that, it is fair to give everybody the possibility to access sports' events.

The solution can be achieved introducing dynamic price systems, like the ones used by airlines companies to sell their tickets. Dynamic price has the capacity to better fill the capacity respect to a fixed price system. (despite what has been said before, some sport clubs have problems to fill their facilities for some official matches).

Dynamic pricing permits to optimize revenues flows, because it can ideally charge to every single clients the maximum price he is willing to pay. On other hand, it also creates opportunities (like in flights-booking process) to buy cheap tickets in advance or concurrently with particular situations. For example, it is possible to image that the tickets of a specific sport arena are dynamically sold, and they will cost less if the weather conditions are expected to be bad or if the same day of the match, another big event (like a famous singer's concert) is happening in the same city. This is just an example, but in general terms, the algorithms that define the tickets' prices can take into consideration dozens of different variables.

There are also other approaches to ticketing aimed to maximize the return for the clubs. An example is the so-called "banking system membership", that has just been introduced by New Jersey Devils, an US

---

[64] *The future of Sport fan, Performance Communication and Canvas8*

ice-hockey team. It basically permits people to make an advance deposit into a sort of pre-paid club' s account to buy tickets and products for the entire season at discounted prices[65].

The idea is that customers will have the opportunity to pay less tickets and other products for the entire season, loading money on their club pre-paid card at the beginning of the season. This represents also an opportunity for teams, that can collect money months ahead respect to when they will deliver the service or sell the products.

Since always more clubs are dealing with financial markets to raise capital, the possibility to collect cash inflows earlier respect to costs can positively impact their financial situation with the consequent benefits in term of markets' reputation[66].

A different phenomenon that now is affecting the operations of club is the re-selling, where always more people are looking at events' ticket secondary market with speculation purposes, exploiting the big difference between demand and offer. Clubs need to monitor, control and avoid illegal activities even introducing authorized systems where people can exchange their tickets in case of giving up.

E-sports

E-sports now represent a relevant business for many sport clubs and organizations in the world. The evidence of the growing interest is that in the Pyeong-Chang winter Olympics game of 2018 have been organized parallel events of e-sport during the racing period[67].

As it is possible to notice, in the figure 15, the e-sports box is colored contemporarily with the tones of all the four macro-areas. This is the graphical representation of the fact that it is a complex sector for classical sport' organizations since it involves the same benefits and criticalities of all the macro-areas together. In fact, it is possible to highlight that, for example for a football club, create a e-sport team is like to create a basketball team, meaning a completely different business area and it requires the creation of a new business unit in the organization. E-sports' athletes need to be trained as

---

[65] *Deloitte's sport industry starting lineup, trends expected to disrupt and dominate 2018, Deliotte Report*

[66] *One of the companies that, before anyone else, leveraged on this practice to grow was Starbucks, that in 2017 has reached more than $ 1.3 bn of deposits for pre-paid products.*

[67] *https://www.ilfattoquotidiano.it/2018/02/09/esports-alle-olimpiadi-a-pyeongchang-si-sono-mossi-i-primi-passi-con-gli-intel-extreme-masters-di-starcraft-ii/4147819/*

professionals, with dedicated training programs and ad-hoc facilities. To be highly competitive it is required a team of technicians and experts to support their activities. Since the matches are completely digital it is also required the use of software to analyze the performances and to improve the skills. The technological infrastructure to develop a team is relevant.

E-sports require ad-hoc marketing strategies to engage fans that, at least in some cases, do not support the “main team”. In other words, it is possible that the fans of the e-sports’ team do not care about the football’ team of the same club. So, differentiation in engaging strategies is necessary.

E-sports’ activities are currently used to drive re-branding strategies of sport organizations. The idea is that the introduction of these teams will permit sports’ organizations to engage with new supporters and enlarge the fan base.

E-sports team requires also in some case new facilities, like new arenas where to place the events. These buildings must be totally different respect to the already-owned facilities of the traditional sports’ organizations. This increase the complexity, also in economic term, of the e-sport team’s introduction.

In other cases, clubs decided to manage the ‘old’ and the ‘new’ team trying to exploit possible synergies. So, the idea is to match traditional sport’ events and e-sports’ ones in the same arena and simultaneously. Also in this case, the arena must be equipped with the necessary digital infrastructures to allows fans to assist properly to the e-sport events and, thus, relevant investments are required.

Despite of the complexity, e-sports now represent a big opportunity for traditional sports’ organizations. In fact, many European soccer leagues are rumored to be launching e-sport leagues soon. Some of top European clubs have already invested in teams like Schalke 04, PSG, Valencia and West-Ham and, in Italy, clubs like Roma and Cagliari have organized e-sports teams. In USA the level is even more advanced since a relevant part of professional clubs have created digital sport teams.

The global e-sport audience reached around 350 million in 2017 with a turnaround linked just to professional e-sport world of almost 1 bn €[68].

[68] *http://www.ilsole24ore.com/art/tecnologie/2018-02-22/in-crescita-ma-ancora-immaturo-nuovi-numeri-mercato-esport-093400.shtml?uuid=AEYZ5Q4D*

The general interest is also demonstrated by the number of acquisitions that are occurring in this sector. Notable example is the acquisition of Twith, a streaming platform specifically dedicated to videogames, by Amazon for $ 970 million[69]. Facebook and the Blizzard Entertainment have reached in 2016 an agreement[70] that enables Facebook to become the platform for sharing the live streaming of the e-sport events organized by Blizzards.

1B.1.2 ATHLETIC PERFROMANCE

Performance Measurement

A different area where digital technologies can impact the sports clubs' activities, is the one of performance management and measurement. The following image represent the share of clubs in the top 4 major professional leagues in USA that use advanced analytics for the performance management of their teams. The figure is dated on 2016 and probably the percentage now is significantly higher.

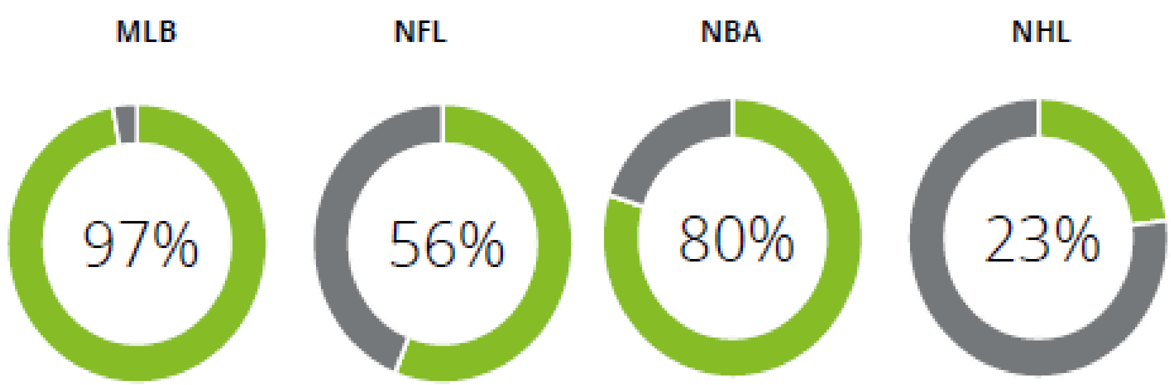


Figure 16: percentage of teams employing analytics and professional consultants in USA sport environment. Source: SportTech Innovation in the start-up Nation, Deloitte, March 2017

69 http://www.businessinsider.com/amazon-buys-twitch-2014-8?IR=T

70 https://blizzard.gamespress.com/BLIZZARD-ENTERTAINMENT-AND-FACEBOOK-TEAM-UP

Potentially sport performances can generate an immense amount of data. *“An average Major League Baseball’s game can produce 7 terabytes of data. For comparison, the Hubble Space Telescope generates 10 terabytes of new data in an entire year”*[71]. Digital technologies can play a central role in gathering and analysing this information.

The value of data is related to the fact that sport’ team can leverage on them to better understand and assess the performances of their athletes in an objective and evidence-based way. For taking more informed decision is necessary to have historical dataset of past performances and so every type of performance tracking is necessary.

This concept will be in-depth analyzed in the Part II of the work, since it is one of the main applications of AI in sport.

Training

The central activity of every sport organization is the training. Now professional sport’ organization have advanced facilities and the most innovative technologies for train at best their athletes.

The underpinned concept of any improvements in the sport training is the awareness that every athlete is different in the way in which his/her body will face and react to the training session. For this reason, the current trend aims to create ad hoc training plans for every single athlete, even if they are part of the same team. Digital technologies can support team managers to properly assess the conditions of every single players and create for them a training plan that could fit perfectly. In this case, innovation mainly deals with wearables and tracking technologies.

Generally speaking, the sport environment has become incredibly competitive. In the last Olympic games, the difference between gold and silver medal in some cases has been as little as three thousandth of a second.

In this context, every solution that can improve, even slightly, the performance of athletes can make the difference between losing and winning. This is even more relevant considering that to certain sport

[71] The secret of winning with data science: use the right playbook, Kinduct, https://www.kinduct.com/wp-content/uploads/2017/02/Kinduct-White-Paper_Data-Science-Mar1.pdf

victories are associated turnover of billions. So, every activity that can impact the final performance is extremely valuable for clubs.

Health & Rehabilitation

Athletes are now monitored 24/7, to ensure that every aspect of their life can positively impact their ability to perform at best. Athletes are subject to biological test to control their health. They wear sensors and other smart accessories that can track every single movement and the evolution of specific values. All these data are used by the medical equip to build up a holistic analysis of the current state of the athlete, with two main objectives. The first to optimizing the performances and the second is preventing any injury. Digital technologies as before are a valuable support in the tracking and in the analysis of these values.

1B.1.3 MANAGEMENT OF SPORT EVENTS

Community Engagement

One of the most important classes of stakeholder for a professional sport clubs is the community where the club is placed. The concept of community and fan base can be similar. In the past the two words were synonymous, but now, since fans are worldwide spread, the two concepts have assumed different nuances. The concept of community refers to the fans that live close to the club's facilities, but also, all the person that worldwide feel "belonging" to that club. These individuals cannot be considered simple fans, and consequently, also the relationship between them and clubs must be different.
To explain better the dynamics of the community engagement some notable cases will be presented. An example, in which digital technologies have been employed for community engagement purposes is the social platform MAKEACHAMP[72]. It is a crowdfunding platform specifically dedicated to sport,

[72] *https://makeachamp.com/it/*

where people can philanthropically support athletes and sport-related project. It raised more than $ 2 million since it was founded[73].

There are also cases in which fans decided autonomously to finance some clubs' activities, like in Norway the team Kristiansund bought a player thanks to the crowdfunding of their fans[74].

It is also notable to consider the fact that, at least in Europe, some sport clubs are directly owned by their fans through a share-based system. In this case, the shares are not traded on a stock exchange, but they are purchased with a subscription model. This additional level of engagement fan-clubs generates values for the club both from a social perspective, since it becomes a central entity in the life of fans, and also from economic perspective, because it can eventually be sustained by the fans. At the same time, it introduces additional level of managerial complexity which can generate conflicts.

A completely different phenomena is that professional sport clubs worldwide are starting to engage their fans far from the "headquarter" not just in a digital way but also with a physical presence. There are clubs and organization that are opening in other geographical markets, physical venues where fans can buy branded products, consume coffee bar products and recreate a stadium-like atmosphere during the official matches, recreating in every city a sort of community. Real-world examples are the NBA café in Barcellona[75] and the future opening of café in the Asian market managed by joint-venture between Juventus and Segafredo[76].

Management of the Sport Arenas

it is always more evident that the sport facilities represent a clear competitive advantage and a critical success factor for every sport organization. The revenues generated directly and indirectly by the facilities-related activities have been increasing their specific weights on balance sheets of clubs. There

[73] *https://makeachamp.com/it/*
[74] *https://www.dreamteamfc.com/c/archives/news-gossip/133225/top-flight-norwegian-side-kristiansund-bk-use-crowdfunding-sign-new-defender/*
[75] *http://barcelona.nbacafe.com/*
[76] *http://www.calcioefinanza.it/2018/04/12/juventus-caffetterie-segafredo-asia/*

are numerous cases in which the clubs have improved their brand and their financial position thanks to a properly managed facility.

For these reasons is now crucial to manage effectively them to maximize the potential return. Stadiums must become a technological platform[77], that can support the creation and the delivery of value to fans though new services. Clubs must change the way in which they have traditionally managed the arenas, moving toward a more entertainment-inspired offer. This situation is affecting also the way teams are building up their venues. The Palau Blaugrana, the new arena of Barcelona for indoor sports, has been "*expressly designed to create an intense spectator experience and to intimidate the visiting team, amplifying the crowd noise*"[78].

Venues in the next years are expected to offer truly unique experiences, also considering the previous arguments about the increasing tickets' price. We will probably assist to a situation in which stadiums will be filled up with lights, massive screens and large projections to enhance the atmosphere's engagement level.

In parallel, we are witnessing to a differentiation of the arena experience, with the creation of different sectors (or versions of the same products, that is assisting to the sport match). The versioning of the tickets, from high-level to basic one, permits to differentiate the price, enabling each client to choose to most appropriate offer for him. In this way clubs can extrapolate the maximum value from ticketing. Digital technologies can play a central role in enhancing experience value and deliver always more services and information to fans.

Events' Organization

Event organization refers to all the activities and operations that are not directly comprises in the management of the arenas experience, so not related with the core event.

Since in every stadium there is a relevant flow of people, security is a central issue. Clubs must mandatorily ensure that everything will happen in the maximum level of security. For this purpose, new

[77] *The stadium as a platfrom, A new model for integrating venue technology into sport business, Deloitte Report*

[78] *http://www.hok.com/about/news/2016/03/16/hok-reveals-design-details-for-fc-barcelonas-new-palau-blaugrana-arena/*

digital technologies like face recognition, digital fingerprint-based access and crowd analytics can enhance the control level within venues.
The management of the flows like traffic, parking and venues' access can be optimized to make the experience even more valuable. The idea is the deliver information to tickets' owner about which root to travel to reach the arena, where search for parking and at which access get in the venue with the objective to optimize all these streams through a central aggregated management.

Referee' Support

In this case, technology will be employed to support the referee in taking more objective decisions. Fan, athletes and clubs want more transparency and objectivity in the referee decisions. in some sports, the judgment system is still based on a jury's grade and this make opaque some decisions.
To solve this problem a lot of world sport organization have introduced technologies to help and support the referees. Examples are the Hawk-eye in tennis, the VAR and goal-line technology in football, and the cameras-based systems in basketball and cycling. There is also technology in fencing, where sensors indicate the precise timing and location of hits.

1B.1.4 ORGANIZATION MANAGEMENT

Management of the Team

The management of the team, intended as the management of the group of the athletes that form the team, is an activity that has been widely impacted by digital technologies. As said before, teams now collect large amounts of data about the health and the performances of athletes. This is a central activity for the team management. Indeed, teams started to adopt real Information Systems, dedicated to the management of the athletes. The number of athletes associated to a single club, considering first teams and young ones, is relevant and so, the necessity for a centralized management of the data arises.

Talent Scouting

Traditionally talent scouting was an activity performed by the groups of scouts that attended thousands of matches and gathered information about relevant events that happened. All these stats were later aggregated and used to study the next opponent or to evaluate a possible future player to sign. In the recent years, all these activities turned digitally. Now there are professional scout platforms for every major sport in the world, that make available for their clients, wide databases with thousands of stats for each player and video of past matches. Clubs can access these platforms, via cloud, and extract specific information for preparing the next matches or for recruitment purposes.

This is an area in which AI can play a relevant role. We will speak about AI-based solutions in details in Part II.

Commercial Partnership Management

The global sports sponsorship market had an estimated value of almost $ 70 billion in 2017[79], with an annual growth rate of almost 15%[80]. One of the main drivers of this trend is the Digital channel growth that enables brand to reach a wide audience. Modern analytics tools give companies the possibility to create effectively personal communication campaigns through sport events, ad-hoc marketing strategies and a better exploitation of the digital touch points.

Sponsor brands also need measurable return on sport sponsorships' investments to justify their spending. Digital tool enables this kind of valuations, measuring the reactions of fans to the brand exposure.

Currently the sport sponsorships are moving away from traditional TV-based format to focus the efforts on social media platform and on-line channel more in general.

According to PwC's global entertainment and media outlook 2017–2021, global internet advertising surpassed global TV advertising in 2015, with mobile online advertising set to count for more than 70% of the total online advertising in late 2018.

---

[79] Sports: the most disrupted of all industries? Pwc's Sport Survey 2017, Pwc Report

[80] *https://www.livemint.com/Sports/EoSxiHNmogPEnStRxP95NP/Growth-in-sports-sponsorships-slowed-to-141-in-2017-Repor.html*

PART II - ARTIFICIAL INTELLIGENCE IN SPORT

After having introduced the current AI environment and sport industry, Part II is the convergence of the two sectors in which Part I was divided. Here the mapping analysis of all the artificial intelligence-based solution in the sport environment will be presented.

In figure 17, the graphical results of these analysis have been presented. It is possible to distinguish different applications' areas and the sub-categories for each of them.

MACRO-AREAS OF APPLICATIONS	CATEGORIES IN WHICH THE MACRO-AREA IS DIVIDED	
ACTIVITY & PERFORMANCE		• **ATHLETIC PERFORMANCES OPTIMIZATION**: processing biomedical and biomechanical data to improve the training effectiveness, injuries management and pattern recognition. Support of wearables devices. • **TRAINING ENVIRONMENT:** AI together with other technologies like VR to create virtual and 3D training sessions. • **TEAM PERFORMANCE ANALYSIS:** match analysis and definition of new metrics.
MANAGEMENT & ORGANIZATION	**TEAM-RELATED ACTIVITIES**	• **SCOUTING & PLAYERS' INVESTMENT EVALUATION AND OPTIMIZATION**
	SUPPORT TO COMMERCIAL ACTIVITIES	• **SPONSORSHIP VALUE ESTIMATION:** to track social value of signage and monitor real-time fans reactions to brands. • **NON-TRADITIONAL FAN MONETIZATION:** smart-ticketing, segmentation.
FAN & MEDIA		• **MEDIA AND BROADCASTING:** automatic media creation and distribution, home-fan can actively "direct the show". Exploit the concept of data-driven storytelling. • **SMART ARENA AND STADIUM AS A PLATFORM:** stadium as a technological and commercial platform. AI for fan engagement in relation with "secondary screen experience", for the management of the event and of the facility.
EXTERNAL STAKEHOLDER		• **BETS AS ASSET CLASS:** improve the forecast ability of future matches for betting purposes. • **REFEREE SUPPORT:** leverage on AI to make referees' judgments more objective.

Figure 17: this table represent the AI applications' field in Sport.

2.1 ACTIVITY AND PERFORMANCE

This category comprises all the solutions whose applications' field are the training session, or more in general, all the activities that occur in preparation of an official performance. It has been modelized that the sports performance depends on these four factors' categories:

- PHYSICAL. It involves both the **physical performance optimization** and the **injuries management** related activities.
- TECHNICAL.
- TACTICAL.
- COGNITIVE.[81]

There are AI solutions that cover each of these 4 main areas, that affect the final performance of the athlete. While regarding the first three areas there are some already developed solutions, the cognitive analysis represents the frontier of the current sports' knowledge. So, in this case, only a general description of the topic will be provided.

2.1.1 PHYSICAL PERFORMANCE OPTIMIZATION

Objectives of the Systems and Current Limits

The goal of these systems, is making the athletes able to arrive at the official performance at the peak of their athletic conditions, structuring the appropriate training programme, over the length of the entire competition. To do this, it is necessary coupling progressive overload period, with recovery ones, in relation with official performances, that are planned to be delivered at fixed dates like Olympic games, play-off and finals.

The objective is making the athletes in condition to perform at best, and then, to recover the at highest level possible for the following performance. The basic assumption is recognizing that the training

[81] *http://www.paulbull.co.uk/index.php/my-blog/28-coaching/34-ttpp-technical-tactical-physiological-psychological.html*

capacity and adaptation are highly individualized, also because these attitudes depend on many factors like health, injuries, nutrition, sleep, training experience, psychological condition and stress that are highly specific for each person. All these factors can explain and participate in the definition of the performance.

In this field, currently all the knowledge available comes from the Medicine studies. Even if, they are valid, they lack the precision to explain the single case because, as said before, they body's reaction is highly individualized[82]. For example, from medical knowledge, it is known that stress affect negatively the physical performance. But as commonly known, everybody reacts to stress in a different way and what AI is asked to do is quantifying how much, and, in which terms, stress affects only a single athlete and do it for everyone.
So, if it would have been possible to elaborate a clear matrix of analysis, to define the causal links between all these variables and the single athletes' condition, we could have in a second moment, the opportunity to leverage on it to improve the athlete's health, and to avoid underperformance in the future, thanks to the predictive power of such instrument.
This is exactly what the AI solutions are trying to do, concerning the physical aspects of sport management.

AI Systems' Features

Usually these systems are based on biological data tracking and analysis, that are monitored by wearables. These instruments are small electronic devices, that can be worn or attached to various body's parts and they comprise sensors that have the computational capabilities to collect data on various performance like cardiovascular or hydration values. The fact that the input of the AI systems, that is, by definition, an algorithm or a software, is measured by another machine, introduces an additional level where errors and variations can lay. In fact, the perfect functioning of these AI solutions depends on the reliability of these detections tools' output.

[82] *Paolo Cintia, during a case study*

The classical working methodology of AI-based products is the continuous h24 monitoring of the strain and the overall health's conditions of the athletes, to give personalized training recommendations and designing real-time changes in training sessions' plan, to have a better adherence and resilience to the athletes' physical condition. They are also expected to assess all the external factors already mentioned like nutrition, sleep, stress etc. evaluating for each of them, their impact on the physical performance.

The values that can be monitored are numerous and every AI solution's producer, usually track a specific set of data which, at least in the opinion of the developer company, shows a correlation with specific performance drivers.

For example, the company WHOOP created a product that monitors h 24 cardiovascular values and heart rate. From this data, it extrapolates the body's physiological workload and gives insights about the needs of the body, pointing out if it is in condition to have specific training session or, for example, if it needs to rest. Additionally, it creates tailored-made recommendations for the diet and the consumptions of specific food&beverage products. It contemplates also functions for optimizing the sleep cycle within a day.

The approach of WHOOP, that it is also used by other companies like Fatigue Science[83], Catapults[84] and Orreco[85] is a supervised learning methodology. The producers know that a correlation exists between a specific pattern in a physiological activity (that could the hearth rate variability) and the athletic performance (that could be the ability of the body to reap the benefits from training at a given point in time).

This is proved scientifically by medicals equips or scientific research, and it is not a task that is carried out by the algorithm. The system is, instead, trained to recognize a similar pattern in a future case. In other case, more rarely, it is possible to see also unsupervised learning-based analysis. Where, a complete new set of data are given in input to the system (pre and during performance) and the machine is expected to find correlations within them. In this case, the system can create more valuable

[83] *https://www.fatiguescience.com/*
[84] *https://www.catapultsports.com/*
[85] *https://orreco.com/*

insights, because they are supposed to be 'new' but, of course, arises the need to validate, later on, these results in a scientific way.

AI Systems Working Methodology

Now the working methodologies of some of the functionalities offered by the artificial intelligence-powered platform of WHOOP will be presented more in detail, to make a real-world example of which are the activities that artificial intelligence can perform in this field. The concepts and the principles explained can be generalize in the entire athletic field, to other solutions.

Figure 18 has been taken from the scientific papers over which the WHOOP technologies are based. It shows the correlation between the Hearth Rate Recovery (HRR), that is defined as the percentage decrease in the Hearth rate caused by resting, 30 seconds immediately following a near-maximal exercise (on y-axis) and the recovery effectiveness (on x-axis). This is something related with body's biological aspects. It is also known by medical theory that the HRR is affected by a variety of factors, like accumulation of fatigue, workout intensity, sleep cycle and diet[86]. The system is basically required to predict the recovery effectiveness knowing the HHR. This means give real-time insights to coaches about how the training is affecting the athletes.

[86] Case study // the science and application of heart rate recovery, Christopher Allen, Emily Breslow, department of physiology and analytics Whoop, inc.

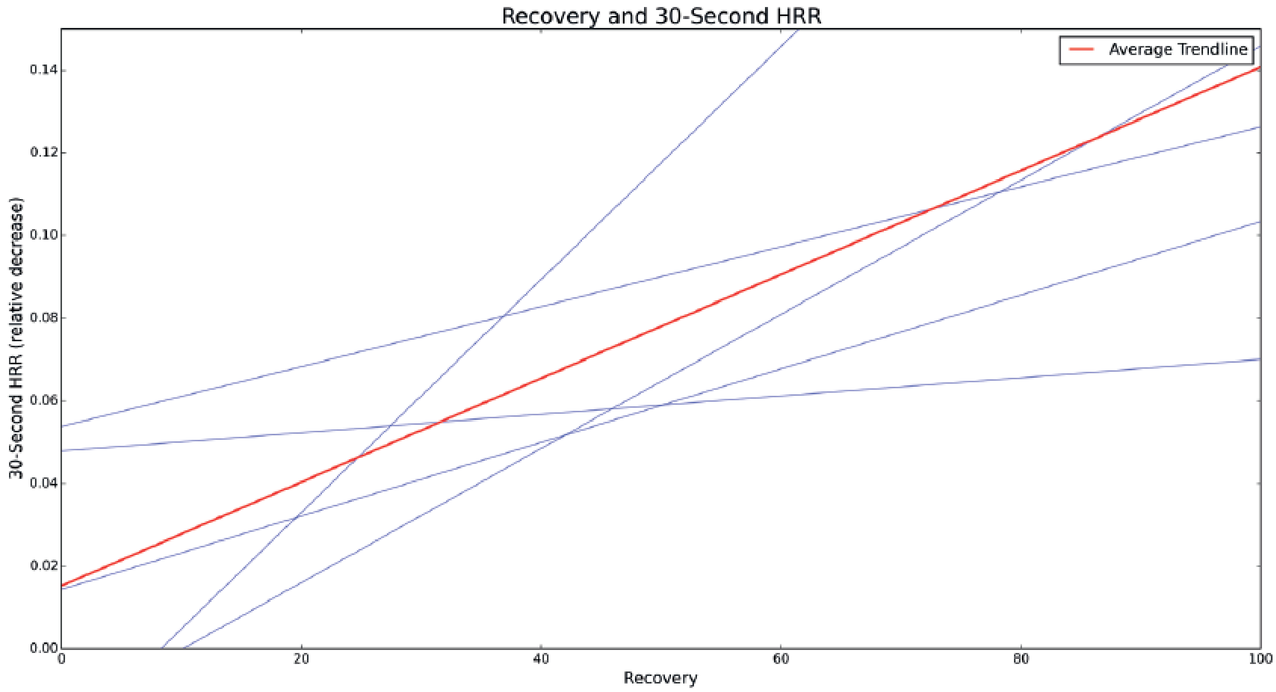
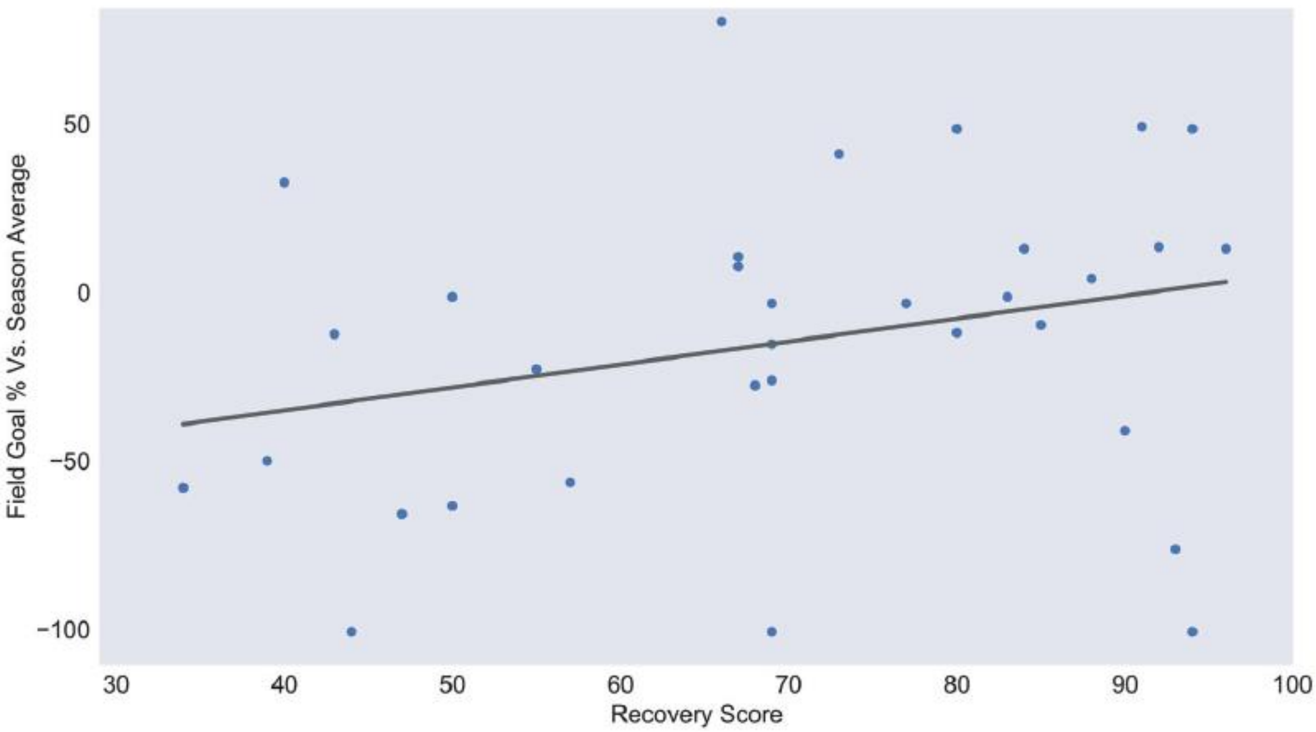


Figure 18: "The relationship between Recovery and 30-second HRR after running workouts between 6.01.2016 and 12.01.2016 for five WHOOP users. A linear regression of the day's Recovery on HRR was fit for each user and plotted in blue. The red line represents the population average." Source: Case Study // The science and application of heart rate recovery, CHRISTOPHER ALLEN, EMILY BRESLOW DEPARTMENT OF PHYSIOLOGY AND ANALYTICS WHOOP, INC.

Figure 19 in another example of the working methodology of WHOOP systems. The data are taken from a U.S.A Division I college basketball team throughout 24 games of their 2015-2016 season.

The blue dots represent the performance of an athlete during one of the 24 games analysed. The y-axis shows the percent difference between the shooting accuracy for that game respect to the athlete's average performance for the season. The x-axis shows the Recovery Score, an index developed by the company WHOOP, that merges different metrics as resting hearth rate, hearth rate variability and sleep management. It is supposed to be predictive of an athlete's readiness. So, assuming that this statement is true (the purpose of this example is not validating theories about the human body function, but just show the qualitative functioning of the algorithms), the system has to track the Recovery Score for every player for some days before a match and give insights to the coaches about which are the players that in that specific moment are more prepared to play, and, statistically, can have an higher accuracy in shooting.



87

Figure 19: it shows the relationship between Recovery Score and points prediction. "The blue dots represent the performance of an athlete during one of 24 game analysed in this report. The y-axis shows the percent difference between the field goal shooting accuracy for that game, and the athlete's average performance for the season. The x-axis shows the WHOOP Recovery Score obtained the night before the game. The solid line is a least squares linear fit through the points." Source: Case study // whoop recovery score as a predictor of basketball performance in NCAA, division I collegiate ahtletes, Janathan Lansey, Gary Power, Emily Breslow, department of physiology and analytics whoop, inc.

It is also possible to notice that a consistent share of the market is focusing on the sleeping management to optimize the performance. The sleep has a significant impact on the athletes' psychological and

[87] *Case study // whoop recovery score as a predictor of basketball performance in ncaa, division i collegiate athletes Jonathan Lansey, Gary Power, Emily Breslow department of physiology and analytics Whoop, inc.*

physical readiness to perform. Sleep optimization' systems are already common in USA professional team sport, where the clubs play matches in different time zones, and thus, the effects of sleep are more significant.

Even if a lot of organizations have already adopted such systems, it is possible to notice some criticalities in term of how to monitor effectively the sleep cycle. A lot of devices lack of the adequate accuracy and it is not yet scientifically defined what is the best way to monitor sleep: position, breath, cardiac beat all are presented as meaningful proxy of the sleep utility but, in general, the solutions lack of a medical and scientific validation of the thesis. In this case, the problem that could arise is the over-fitting. That means that the Machine Learning system would focus on irrelevant and spurious relationships in evaluating the impact of sleeping on the athletes' performance.
In fact, for example, David Dario Cioni, technical director of Team sky[88], confirmed that, even if his organization recognizes the value of sleep monitoring, they still have not found a solution that guarantee an adequate reliability.

It is worth highlighting that AI-based solutions in the athletic performance optimization, are not a prerogative of only big clubs. A lot of companies are proposing both the elite and the basic version of their products, and other have a price that cannot be consider prohibitive also for small organizations[89].

Even if every company proposed products with distinctive features and strategic target, it is possible to follow a general model to define the role of these systems in the training environment[90]. The model consists of the following steps:

- Identify long-term strategic objectives. This planning activity has an extended time horizon, usually the entire competition. Usually the objective is bringing the athletes at the end of the competition at the condition's highest peak, without jeopardize previous results with

[88] *Team Sky is one of the most important cycling team in the world*
[89] *https://www.kinduct.com/*
[90] *Computational intelligence in sports: Challenges and opportunities within a new research domain, IztokFisterJr., KarinLjubiˇc, Ponnuthurai Nagaratnam Suganthan 2015*

underperformances. it depends on several factors like age, sex, nutrition, maturity and training predisposition.

- Identify shot-term objectives: they usually are the adaptation of the long-term strategy over a restricted time horizon. They depend on the specific athletic conditions in that timeframe.
- Measuring the training intensity: the measuring is typically performed with wearables.
- Data analysis: it comprehends data processing and visualization that are later used to re-adapt and re-plan the long-term and short-term objective.[91]

Another characteristic is that all these systems are trying to use a holistic approach for performance optimization, following the idea that the athletic performance is influenced by many factors and, some of them, are unrelated to the training. Nowadays there are the technologies available to measure all these factors in a continuous and reliable way. For example, the wearables can be worn by athletes without any problems thanks to their small size. Until some time ago, it was not possible to measure real-performances due to this big limitation. All the measurement had to be done in a lab, with great problem in generalize the results. Other technologies like GPS and Bluetooth had also increased the effectiveness of the measurements. There are also devices that can monitor blood pressure, hydration and other biological aspects without bother the athletes during the performance. This permits to measure the effects of diet, sleep and the of the every-day-life activities on the performance. Combining all this information is possible to draw a holistic assessment of the player.

2.1.2 INJURIES MANAGEMENT

Why Injuries Management is a Source of Value

Injuries are one of the biggest determinants of performance in sports. Injuries have severely limited many athletes in achieving their best in almost every sport, and, they can represent a relevant discriminating factor between competitive success or failure. In modern sport environment, injuries are

---

[91] *Computational intelligence in sports: Challenges and opportunities within a new research domain, IztokFisterJr., KarinLjubiˇc, Ponnuthurai Nagaratnam Suganthan 2015*

also a source of economic losses. Injuries can cost teams millions in salary and lost revenues, but they also destroy the players residual value (usually serious injuries reduce the market price for players). Considering, for example, the football market, where the valuations of players are getting relevant, the residual value can become an important source of profits for team (through capital gain after the sale).

This situation makes injuries' prevention a central activity in the management of a team and the introduction of matrixes to assess injuries risk can become a driver for the sport success.

The aim of artificial intelligence-based systems for injuries management solutions is not addressing predicting injuries. They depend by too many factors that, in most of the cases, are unforecastable (consider for example rude tackles). Instead, these systems' aims are computing and monitoring the risk of injuries, answering to questions like "*Knowing an individual's current physiological condition and with a knowledge of their expected workload, what is the likelihood that the player will be injured?*"[92]. The idea is analysing in conjunction data about training regimes and game frequency to determine the likelihood of future injuries. Clubs can also go further, collecting data about each player' training history over his career with the club, to create a database to discover more meaningful and long-term correlations[93].

How AI systems Can Mitigate These Issues

Injury management is one of the fields where AI can be employed with higher success thanks to the enormous set of data available and the nonlinear relations existing between variables. Basing on historic time series Machine Learning algorithms can figure out insights and correlation not discovered yet. In other words, AI systems can monitor all the data related to players injuries and try to find correlations between the occurrences of injuries and the specific patterns in the workloads of the players.

---

92 An enhanced metric of injury risk utilizing Artificial Intelligence, Calham Dower, Abdul Rafehi, Jason Weber, Razali Mohamad Alerte Digital Sport

93 Scout 7, changing the game, Intel 2017

For example, Alerte Digital Sport has developed a solution for injuries management. It proposes an AI software that takes data from a multitude sources e.g. GPS, RPE (Rate of Perceived Exertion), well-being, sleep and integrates all of them with current and future workloads, to develop unique models for each player." *With the purpose to provide users with a risk assessment per player per day*"[94].

Another company that is creating solutions to monitor injuries is ThermoHuman, a Spain based company, that is using infrared thermography as tool to assess the sport performances (if the athletes condition is in line with the training objective) with specific application in injuries treatment and prevention. They developed an algorithm based on deep learning technologies, that goes through the thermic images of the players to evaluate the risk of injuries and the effectiveness of trainings and rehabilitations[95].

Decreasing the percentage of injuries will not only impact the situation of sport clubs but it will significantly affect the athletes' health avoiding possible grave consequences.
For example, in "*Deep learning approach to predict mild traumatic brain injury in contact sports*", the authors[96] tried to create a deep learning system to predict TBI injuries (Traumatic Brain Injuries), basing on neuroimages. Even if they recognized to not have found an "optimal" solution, they discovered that their deep learning model got the highest results respect to all the previous employed system to recognize in advance this specific type of injuries.

The topic of injuries management will be presented more in detail in the case study about Luca Pappalardo and Paolo Cintia (Section 3.2.1). They are both researcher at the CNR of Pisa and they developed a system able to predict the injuries' occurrences.

---

94 https://www.alerteds.com/

95 http://www.thermohuman.com/

96 Yunliang Cai, Wei Zhao, Zhigang Li, Songbai Ji

2.1.3 TECHNICAL PERFORMANCE ANALYSIS

In this category lay solutions are specifically dedicated to improving the technical gesture of the athletes. The technologies are based on biomechanical analyses, to give trainers kinematic and dynamic data to help optimizing movements or for adopting new better-performing techniques. In this section, will be considered also all the solutions that, leveraging on VR and AR, creates virtual environment for training.

Concerning the first case, the solutions are usually made off high-performing cameras and wearable that monitors the kinematics and the anatomy during specific athletic gestures, with the purpose to enhance performance and avoiding injuries.
Basing on the classification matrix proposed by Russel and Norvig in 1A.1, it is possible to cluster them in the “Acting Humanly” category. The artificial intelligence systems carry out a task that also a human can do, but, obviously thanks to technology enhancement, they can do more precisely. They can monitor the mechanic stress that the movements can create, the velocity, the load and create a digital 3D model of the players’ movement to improve the analysis.

There are also companies that are creating totally-virtual training environment, accessible through virtual 3D viewer. An example is STRIVR[97]. The company records the matches in 360° and make them available to players in a second moment. Athletes can re-live episodes in first-hand through VR viewer, having the opportunity to be back on the field in with the same perspective of a real-match.
These solutions require big investments in terms of cameras and hardware. Even if these solutions are still unfavourable from an economic point of view, it is possible, in any case, recognize the value that such solutions can create for the players’ performance enhancement. These technologies will be part also of the fan engagement section, since they can deliver completely new and high-value media contents. It is notable to highlight that the producers of these systems are not specifically dedicated to sport. Usually, they have born in completely different industries and, then, eventually adapt their

[97] *https://www.strivr.com/*

products to sport. An example is Eon Reality[98]. It is a company that was born to create virtual working environment for Oil&Gas applications. The idea was training the worker before making them work in dangerous situations that are common in this industry. Just recently it started to develop systems with sport applications, more precisely, for baseball[99].

2.1.4 TACTICAL PERFROMANCE ANALYSIIS

Tactics can be defined as "*an action or strategy carefully planned to achieve a specific end*"[100]. In almost every sport it has been achieved a high complexity in the tactical development. For simplicity and following the academic's interest, this section will deal mainly with two sports: football and basketball.

In these sports, tactics is mainly related to how "*teams manage space, time and individual actions to win a game*"[101] *." Space refers to which pitch's area a team want to occupy, time, in contrast, describes variables like frequency, events duration (like ball possession) or how quick the actions are being initiated"*[102].

In these two sports, exist subgroups of players that, in certain situation, are expected to play in a highly coordinated way. So, tactics goes also to assess and plan the synchronization of their moves. To conclude, "*individual actions specify technical gestures and personal decisions made by single player, in relation with the state of the game*"[103].

As said before, tactics in professional sports is already highly studied and exploited, but until now, most of tactical evaluation are still undertaken in a subjective and qualitative way by the coaches. So, AI can innovate this field basically in two ways:

---

[98] *https://www.eonreality.com/*

[99] *ttps://www.eonreality.com/press-releases/eon-sports-vr-provides-state-art-baseball-training-system-yokohama-dena-baystars/*

[100] *https://en.oxforddictionaries.com/definition/tactic*

[101] *Garganta J (2009) Trends of tactical performance analysis in team sports: bridging the gap between research, training and competition. And, Fradua L, Zubillaga A, Caro O, Ivan Fernandez-Garcia A, Ruiz-Ruiz C, Tenga A (2013) Designing small-sided games for training tactical aspects in soccer: extrapolating pitch sizes from full-size professional matches. J Sports Sci 31(6):573–581*

[102] As note 85

[103] *Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science, Robert Rein* and Daniel Memmert, 2016*

1. Bringing in this field, a more quantitative and evidence-based approach to assess and evaluate the benefits of tactical decisions and the consequences of teams' and players' behaviours.
2. Creating systems with predictive ability to forecast how opponent will play, and, consequently, create valuable competitive advantages.

Entering in a more detailly way the field, it is possible to distinguish in sport tactics development process, two main steps, in relation with the different timing:

- First, there is an a-priori decision that is a "plan of actions", made before the match, with respect to how the team wants to play.
- The second aspect is related to how players behave singularly and collectively during a match, and, how their choices or performances, affect the overall performance of the team, considering also the presence and the interactions with the opponents. This refers to the real-time interpretation made of single players of the a-priori strategy, in relation with the specific game situation[104].

Currently, professional football and basketball's teams, as highlighted in 1B.1.2, have availability to an immense amount of data about players, and, they have already started analysing them to improve their tactical performance.
In fact, a professional profile that just some years ago did not even exist at all, and now, it is present in all the professional sport organizations is the one of the match analysts. A match analyst has assigned the task to manage this information and provide insights and statistics for the coaches, giving them a higher understanding of the games.

The analyses usually focus on three main objectives. 1. The first is to understand how opponents' teams play and which tactical behaviour they adopt more frequently. The forecast of the behaviours of the future opponents permits to create tactics that can be as much as possible effective in facing that team. If you know how your competitor will play, you have a great advantage.

---

[104] *Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science, Robert Rein* and Daniel Memmert, 2016*

2. The second area is dedicated to the post-match analysis of the own team. This activity is very similar to a sort of Variance Analysis of the budgeting process in business companies.
Teams, before the games defined a strategy about how the matches will be played, which behaviours will have to be followed to create certain situations that, in theory, can lead to a victory (for examples the use of schemes to sign goals). After the match, an analysis must be performed to understand why a possible variance respect to what was planned, happened.
3. The third area of analysis is the evaluation of the single players' performances, through a large set of indicators.

The last innovations in the tactical analysis of the game, is the use of technologies like advanced cameras and GPS-based systems, to create spatiotemporal database of the matches. These instruments can track all the spatial locations of players' movements, along with the time in which they happened. The purpose is to represent players during a match as point-like objects that moves on a two-dimension plan.
This is a significant innovation in the sport analytics, because it enables us to track dynamically and constantly how players behave on the game field.
With these modern technologies it is possible to broaden the analysis and start to evaluate players game in a more holistic way, tracking collective dynamics that, until now, have been impossible to be measured, and, thus, expressing the evolution of matches in rigorous terms.

Current Sport Analytics State-of-Art and Possible Future Development vs The Creation of New Tactical Tools

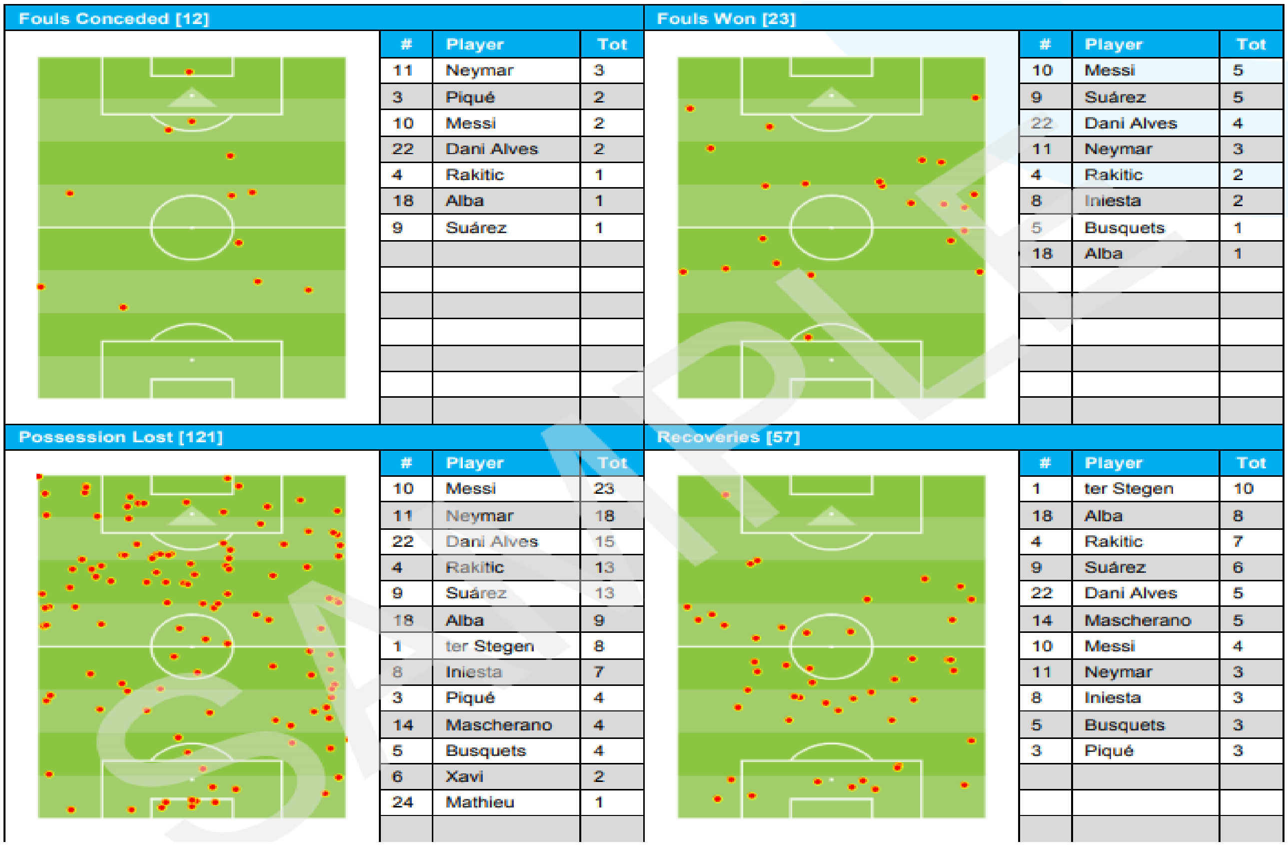
The following logical approach has been directly inspired my Mr. Crivaro[105], in one of meeting in which he explained the limitations of the current tactical analytics. Here, it has been adapted to the thesis purposes.

[105] *Founder and CEO of Math&Sport, a start-up to which one of the case studies has been dedicated*

To show how AI's applications, basing on players' movements tracking can create value for the tactical management of teams, a digression about the sport analytics state-of-art is useful to highlight the conceptual difference between how currently, data, generate value for sport teams, and how they can be used in the future.

The figure 20 shows an example of a report created by Optapro[106] for the analysis of the match of Barcelona against Juventus of 06 June 2015. Optapro is a private company that provides football teams and media broadcaster match analysis services. It is one of the biggest players on the market.
Usually these analyses are well-structured reports of more than 20 pages. They can be used for the analysis of the single match or, if based on aggregated data, to analyse the tactical behaviour of a team along the entire season.
Speaking about football, these analyses cover numerous aspects, like number of passes, kicks, ball-recoveries, average distance between players, distribution of the play, average position and physical performance analysis. Figure 20 showed below is just a small extract from such report.

[106] *https://www.optasportspro.com/it/*



Fouls Conceded [12]

#	Player	Tot
11	Neymar	3
3	Piqué	2
10	Messi	2
22	Dani Alves	2
4	Rakitic	1
18	Alba	1
9	Suárez	1

Fouls Won [23]

#	Player	Tot
10	Messi	5
9	Suárez	5
22	Dani Alves	4
11	Neymar	3
4	Rakitic	2
8	Iniesta	2
5	Busquets	1
18	Alba	1

Possession Lost [121]

#	Player	Tot
10	Messi	23
11	Neymar	18
22	Dani Alves	15
4	Rakitic	13
9	Suárez	13
18	Alba	9
1	ter Stegen	8
8	Iniesta	7
3	Piqué	4
14	Mascherano	4
5	Busquets	4
6	Xavi	2
24	Mathieu	1

Recoveries [57]

#	Player	Tot
1	ter Stegen	10
18	Alba	8
4	Rakitic	7
9	Suárez	6
22	Dani Alves	5
14	Mascherano	5
10	Messi	4
11	Neymar	3
8	Iniesta	3
5	Busquets	3
3	Piqué	3

Figure 20: Sample of one Optapro match analysis's report. The match is Barcelona vs. Juventus of 06 June 2015

The main limitation with current approaches, is that usually these data are purely descriptive, meaning that are based on 'simple' statistics and counts.

For example, speaking about ball possession lost, it is possible to see that in image 20, in the lower left section, there are all the locations of possession losses of Barcelona in that specific game. In the report it is also possible find the number of wrong passes for every player.

But what really is missing in these analyses is a valuation about how these events happened and why. These analyses don't give any indications about the reasons behind these events. Currently the process of understanding the data, figuring out a reason and preparing a tactical approach basing on these evaluations is made off coaches' staff in subjective way.

So, as suggested by Mr. Crivaro, such data are not valuable for the preparation of a strategy to face a hypothetical next match against Barcelona. There is no predictive value in these data and the is no causal explanation and, thus, they cannot really support decision's processes of coaches. These data have just a descriptive power of past events.

Instead, the analysis would become valuable if they would provide, the reasons for which Barcelona lost ball possessions, in which circumstances these events happened and how our own team can behave to make it happen again in the next match.

Only if the system can provide these additional information, it will become useful for the preparation of tactics' aspect of matches, because it would basically enhance the decision ability of coaches. AI systems in the sport tactical analysis should be employed to transform purely descriptive statistics into predictive and meaningful tools to support the decisions of coaches. This exactly the main objective of AI systems, highlighted in the first pages of the work.
Following the conceptual path described above, in the last years several models that apply AI's tools have been introduced to, on one hand, enhance the understanding of the tactical phenomena and, on the other hand, to develop a forecast ability in this field.

These tools are individual or collective performance evaluation indexes, that go to assess episodes, movements and decisions of players, that have been ignored by the classical match analysis frameworks, basically, for the impossibility to find a way to measure that tactical aspects.

In fact, to make an example, it is easy to measure the results of an athletic performances, like for a speed run, it is sufficient to measure the time needed to run through a certain distance. Different thing is measuring if the five players of a basketball team have defended in the optimal way against the opponents in a specific moment of a match, regardless how the action ended, in other words, without considering if the opponent team has scored. Until now, it has been possible to do it only subjectively, looking at the game, highlighting where, according to us, an error may have happened.

In the next pages, some tools of AI-based application for tactical analysis will be presented. All these concepts are not 'new'. Professional coaches already knew these notions and they already exploit them to try to win the matches. But until now it has been possible to judge their effectiveness only in qualitative and subjective way. The tools introduce an instrument to measure them quantitatively and objectively.

The followings part about the tactical analysis will be organized in two sections. In the first one will be presented some models divided respect to the sport target, football and basketball. They are mainly performance evaluation tools. In the second part the "Ghosting" method will be presented. It is an AI-based technique that can be applied in both sports.

Football

In football most of indicators that are currently used to look at the matches and the players' performances are, firstly, individual and, secondly, only specifically dedicated to with-ball dynamics. We completely lack valuable indicators that show how a player behaves without the ball, and how it faces off-ball dynamics. The figure 21 is taken from WyScout[107], one of the most-advanced analytics' platform for football in the world. The player whom these statistics are referring to is Francesc Fabregas of Chelsea F.C.

Match	Chelsea - Huddersfield Town 1:1
Competition	England. Premier League
Goals	0
Assists	0
Shots / on target	2
	0
Passes / accurate	150
	131
Long passes / accurate	21
	13
Crosses / accurate	6
	3
Dribbles / successful	4
	2
Aerial duels / won	1
	1
Interceptions	3
Losses / own half	7

[107] *https://wyscout.com/*

Figure 21: The table above is taken from Wyscout and it refers to the match of Francesc Fabregas of Chelsea F.C. against Huddersfield Town of 09 May 2018

These data should express how Fabregas played that match. As it is possible to see, all of them are based on with-ball situations. There are no indicators on off-ball ones. The unique measures commonly used that are not specifically dedicated to with-ball dynamics are the physical performance indicators, like distance run, that in most cases lack of adequacy and meaningfulness.
This situation, in football, is almost paradoxically, considering that, statistically, a player has the ball for 3 minutes over 90 of a match[108]. So, doing this, even a professional scouting platform is implicitly excluding everything that happen in the 96% of the time, basing all its matrixes on what happen in the in the remaining 4%.
A similar approach has relevant limits, furthermore, if it is used on a scouting platform, where managers are supposed to gather information for taking decisions about new players' investments that can worth millions of Euros.
In fact, WyScout to solve this situation, is proving videos of all the past matches of players. Doing this, managers can look at this video and evaluate personally the information that are missing.

A said before, the impossibility to create matrixes that can cover these topics does not mean that they are underestimated: everybody is aware of the importance of off-ball situations.
Indeed, looking at the football's evolutions in the last two decades, it is possible to highlight that elite teams have adopted playing style in which off-ball situation are always more important. What is lacking is just a way to express this complexity methodologically. Machine Learning and Artificial Neural networks can deal with such complexity and, so, they can be useful for this purpose.

In the paper "*Wide Open Spaces: a statistical technique for measuring space creation in professional soccer*", Javier Fernandez and Luke Bornn, have introduced a system to evaluate quantitative and systematically how well players behave, when they do not have the ball. The intuition is measuring the

---

[108] *Johann cruyff, https://www.pastemagazine.com/articles/2015/02/25-johan-cruyff-quotes.html*

effectiveness of off-ball movements, quantifying the value of space on the pitch that a player, or aggregating eventually a team, have occupied after a specific movement.

To do it the authors, first, defined the concept of “Player Influence Area”. It is basically the space over which a player has a significant probability of controlling the ball. it is influenced by many factors like the location and the speed of the ball and of the mutual interaction with opponents. All these factors draw an area around the players over which it is possible to assume that they have the control.
The model is able to aggregate all the influence areas of each player of both teams and provide a graphical representation. An example is the figure 22. It shows the Pitch control surface for team in red. Arrows show players velocities, and contour lines allow to visualize the surface of control. Numbers in white indicate the pitch control value at their drawing location. Axis dimensions are in meters.

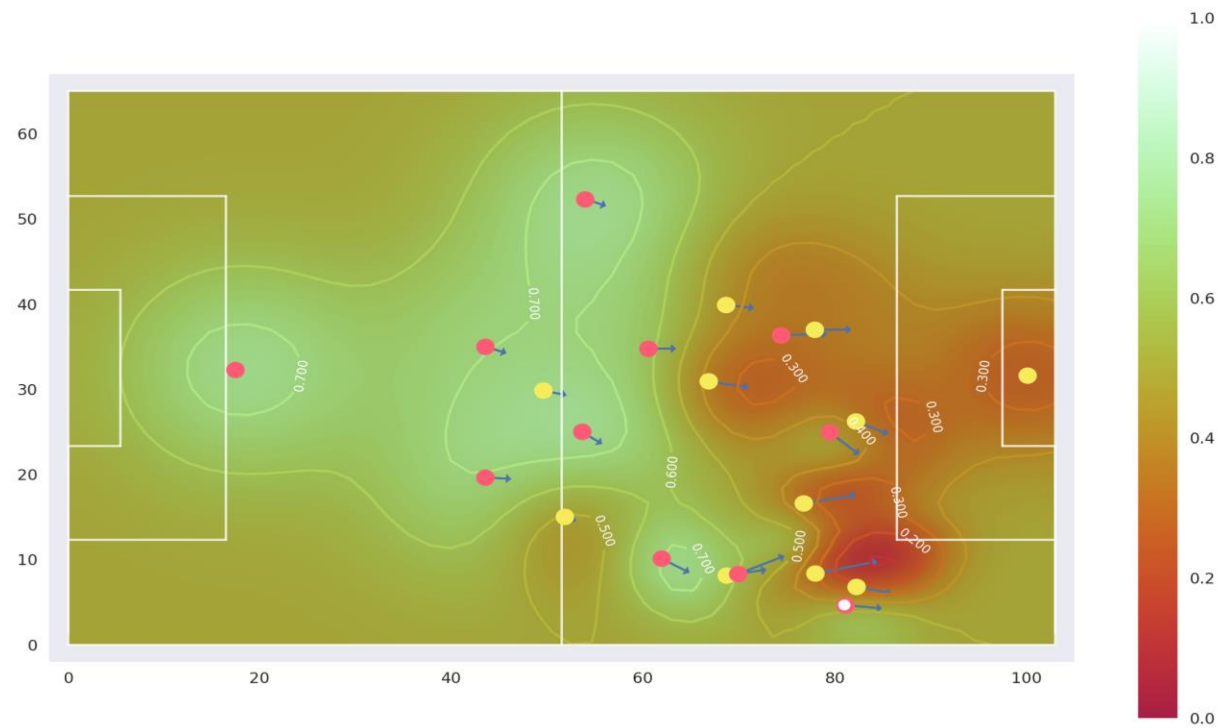


Figure 22: “Pitch control surface indicating the degree of control for team in red. Arrows show players velocities, and contour lines allow to visualize the surface geometry. Numbers in white indicate the pitch control value at their drawing location. Axis dimensions are in meters” Source: Wide Open Spaces: a statistical technique for measuring space creation in professional soccer”, Javier Fernandez and Luke Bornn

After having defined what is the portion of space controlled, the second step, is evaluating the value associated to that area. "*The value or the importance to control an area of the pitch in football changes dynamically, depending on multiple positional factors, such as the position of the ball and the players*"[109]. To face this complexity, the authors employed a Neural Network, that, analysing the spatiotemporal data of matches, learned how to compute the value of space autonomously. The starting hypothesis was that, considering a sufficiently high number of situations, the defending teams in matches distributes themselves through-out the field in a manner which covers the highest spatial value. Studying the behaviour of thousands of teams, the systems learned and figured out an evaluation tool for quantify space's value.

The last step is aggregating the amount of controlling space with the value of the space itself, and, the system is able to compute the punctual evaluation of the space occupied by a team.

The authors proposed also to compute the value of the space that a player is able to generate for a teammate during an open play. They made this distinction, between "*space occupation, that is the space occupied for oneself, respect to space generation, that refers to opening-up space for teammates by attracting opponents out of position"*[110].

In this last case, the concept is more complex because there is the presence of more players and moves. "*A space gain happens when a player makes a move, dragging an opponent away from his previous position, that was close to one or more other teammates. Doing so, a player, without even show interest in the ball can create space for a teammate useful to sign a goal*"[111]. The system is able to associate a utility to these dynamics. In fact, it is able to autonomously recognize when these situations happen and apply the process explained before.

[109] *Wide Open Spaces: a statistical technique for measuring space creation in professional soccer", Javier Fernandez and Luke Bornn*

[110] *Wide Open Spaces: a statistical technique for measuring space creation in professional soccer", Javier Fernandez and Luke Bornn*

[111] *Wide Open Spaces: a statistical technique for measuring space creation in professional soccer", Javier Fernandez and Luke Bornn*

This paper created a tool to evaluate something that could be "highlighted" by a match analyst only looking at the match. The issue is that the dynamics of space creation are so frequent in a game that for an analyst is impossible to get them all. The implementation of a similar system in a real football team would permit to track the consequences every single action.

Figure 23 below shows a new tool proposed by the authors. It is a matrix that represent the total times in which space was generated by generators (y-axis) to receivers (x-axis). Employing this tool for an entire match, it can quantitatively show how much value a single player has create during a match for the rest of the team.

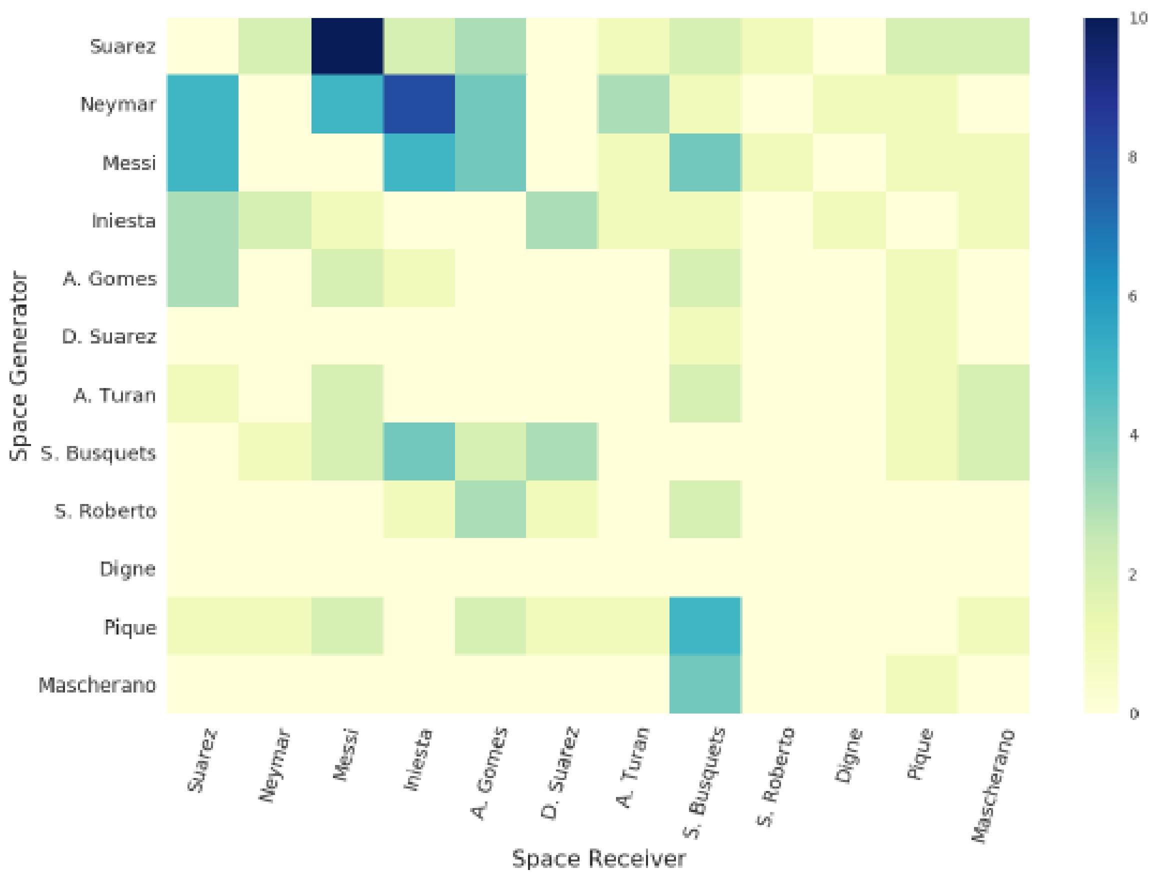


Figure 23: A heatmap showing the total times space was generated by generators (y-axis) for receivers(x-axis) Source: "Wide Open Spaces: a statistical technique for measuring space creation in professional soccer", Javier Fernandez and Luke Bornn

Another model that goes to identify and to quantitatively measure the value of collective action in football is the one presented in *"Quantifying the Value of Transitions in Soccer via Spatiotemporal Trajectories Clustering"* by Jennifer Hobbs, Paul Power, Long Sha, Hector Ruiz and Patrick Lucey. The authors created a tool to identify autonomously and to measure the effectiveness of counter-attacks. Counter-attacks in football are specific transitions, characterized by the fact that they imply the transition from a defensive to an offensive situation in a "*fast, aggressive and direct way*"[112].

The input of the system are the spatiotemporal data of the matches. First, the system recognizes when a counter-attack happened. This is done employing a machine learning tool that is based on both supervised and unsupervised learning approach. The idea is that the system gets in input the trajectories of the players of an entire match, from them it is able to cluster the actions that refers to only counterattacks, because it has received in input human-made labelled examples of counterattacks. The second step is defining the effectiveness of counterattacks. The authors decided to measure it introducing the concept of "Defensive disorder". The starting point is that, a counterattack is effective because it takes by surprise the opponent team that has no time to organize a defence. So, they associated the effectiveness of a counterattack with the deviation of the opponent respect to the formation they would have used to face that situation. In this way, the model can associate to every counterattack a value about its effectiveness and, consequently, the possibility to assess them properly.

Basket

Basketball has seen the rise of analytics before respect to football. The reasons lay in two aspects. The first is related to the nature of the sport itself. In basketball there are less players over a field that is extremely smaller respect to football one. In an average basketball match there are 60-70 points per team respect to 1-2 goals per football match. There is more room for statistics. Also, the presence of technical errors is very different between the two sports. In basketball you play with hands so there are less probabilities to fail as much as in football, where players run and play with feet.

---

[112] *"Quantifying the Value of Transitions in Soccer via Spatiotemporal Trajectories Clustering" by Jennifer Hobbs, Paul Power, Long Sha, Hector Ruiz and Patrick Lucey.*

The second is related to the fact that basketball is most played in USA where the cultural environment and the size of the industry, have helped the introduction of new analysis tools early respect to Europe. In the football section, the focus of the new matrixes proposed, was analyzing the off-ball situations. In basketball, instead, the academic interest is aimed to create tools to evaluate the decision ability of players and if that decisions are optimal for the team.

In "*predicting points and valuing decisions in real time with NBA optical tracking system*", Dan Cervone, Alexander D'Amour, Luke Bonn and Kirk Goldsberry created a model specifically aim to quantify the value of each decision made during a basketball game. With decision they intend every possible play, so pass, dribble or shot. The idea is to evaluate if the player in that circumstance, so given the position of all the other players on the pitch, have taken an optimal decision. To do it they proposed the creation of an indicator called "*Expected Possession Value*" (EPS) for evaluate the offensive actions of teams. It basically assigns a value to every tactical option available to a player at each moment of the possession.

A graphical representation of what EPS means is showed in the figure 24. The players are represented as point-like objects on the pitch. The player number 2 has the possession of the ball. To each option available, is associated a probability to perform that choice correctly, and the expected value associated to it. The probabilities of the events and the estimation of the value is learned employing AI systems on very large amounts of data of past games. With this method all the possible solutions available are showed and it is possible to compare them with respect to their value creation.

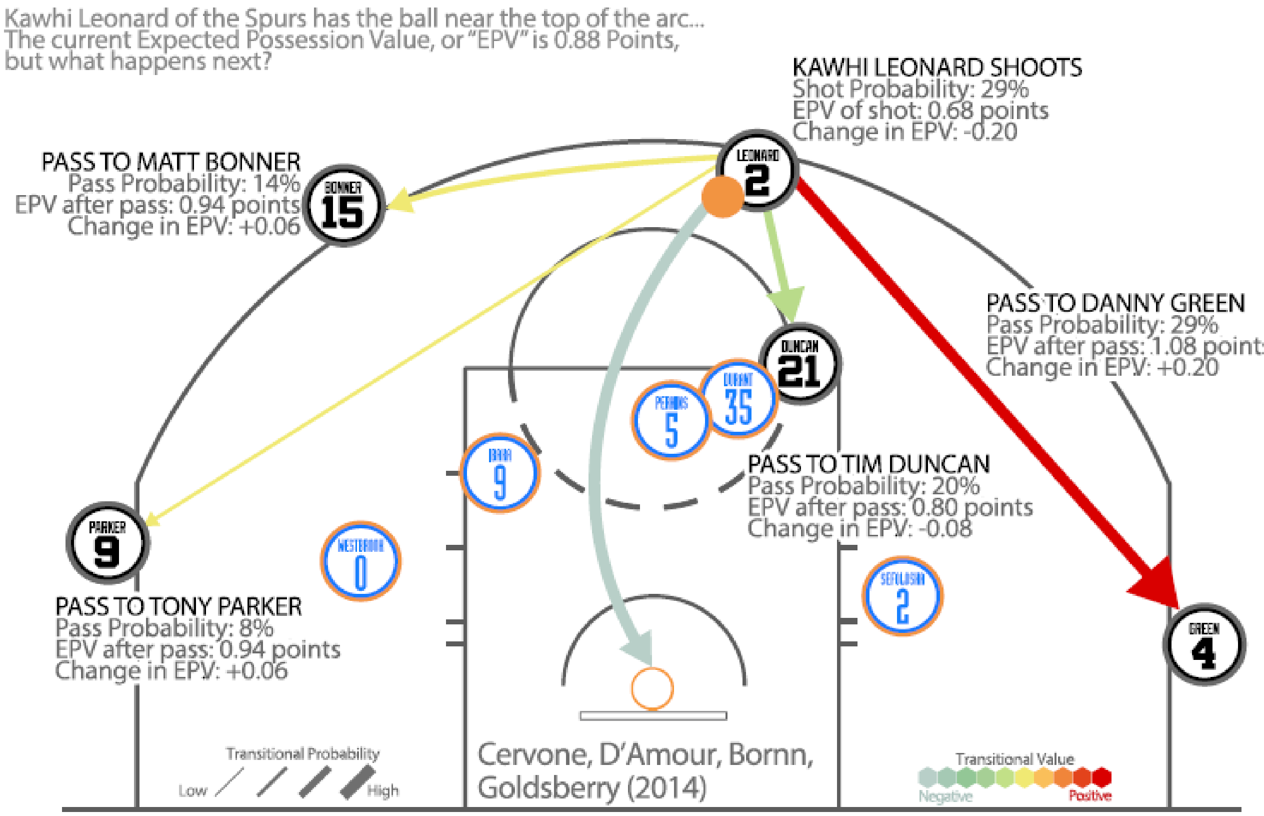


Figure 24: Diagram of EPV as a weighted average of the values of the ballcarrier's (Leonard's) decisions and the probability of making each decision. Source: In "predicting points and valuing decisions in real time with NBA optical tracking system", Dan Cervone, Alexander D'Amour, Luke Bonn and Kirk Goldsberry

A further characteristic is that this model is unaffected by the downstream events and outcomes. The idea is that if a player decided to exploit a shoot opportunity at which is associated a value lower respect to pass option, it will be evaluated as bad, regardless how the situation ends, so regardless, if did shoot right.

In fact, as highlighted by Matt Goldman and Justin Rao[113], the shoot decision is a problem that involves weighing the continuation value of the possession, so, the value associated to a possible teammate shooting, with the expected points of the immediate shooting.

[113] *"Allocative and dynamic efficiency in NBA decision making" Matt Goldman and Justin Rao*

A further application of this work is when EPV is aggregated and calculated continuously. Doing this, it is possible to represent the value evolution during a possession, showing an instantaneous snapshot of the utility of the action. The image 25 represents the evolution of an NBA real possession. In the upper part there are the four events and decisions that happened in that action. In the graph below instead is plotted the EPV of the action. So, it is possible to notice how it varies and understand what make it increase or decrease, regardless how the action ended.

If the EPS is computed for all the matches in a season, it is possible to understand which behaviours and moves are associated to the highest value of EPS and exploit them systematically.

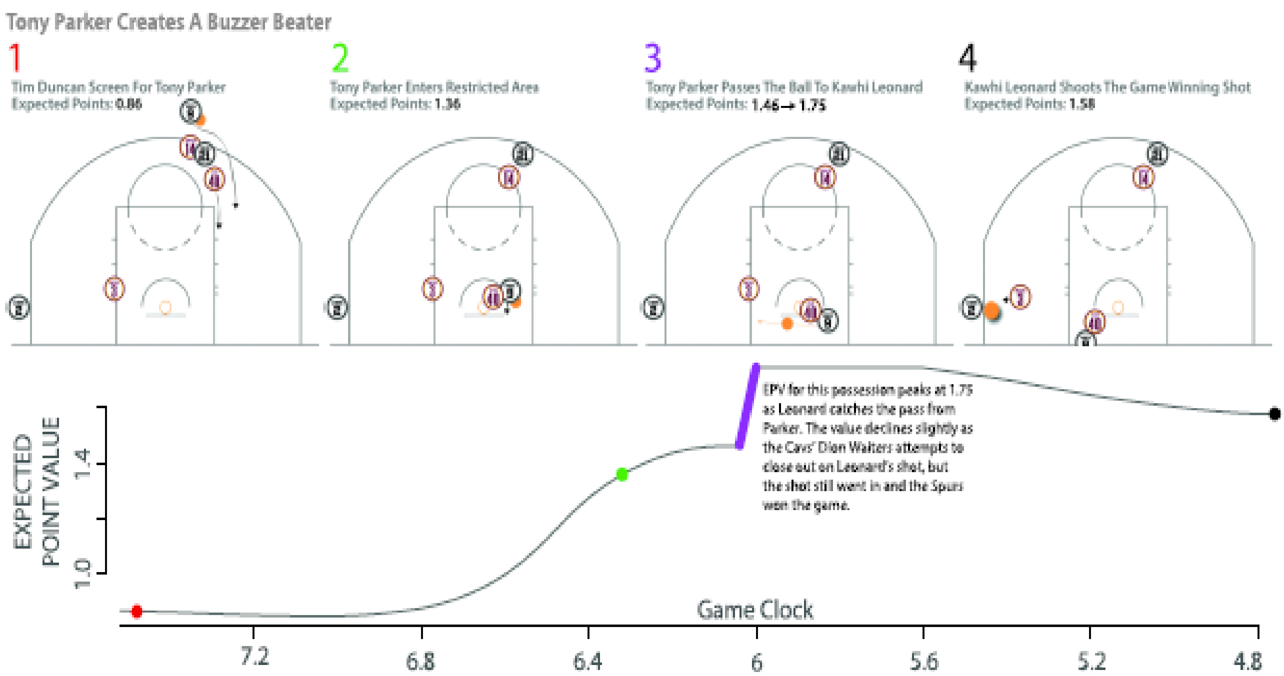


Figure 25: EPV throughout a possession, with annotations of major events. Source: In “predicting points and valuing decisions in real time with NBA optical tracking system”, Dan Cervone, Alexander D’Amour, Luke Bonn and Kirk Goldsberry

Data-Driven Ghosting

The idea of ghosting has been introduced in 2013 in the NBA team of Toronto Raptors[114], where the coaches' staff created a model to show the differences between what players did in a given situation, and what they should have done according to the coaches' ideas. The system was a sort of dynamic white board that had the ability to show how the coach would have expected to manage entire possession or a defensive scheme.

Starting from this idea, some researchers have started to develop new tools, mainly based on Machine Learning, to compare the current performances of players, with a sort of benchmark, an expected one. The main difference is that in the case of Toronto Raptor explained above, the coaches created the ideal model, the benchmark. In other words, they defined how each player would have been expected to play and insert this information in the computer system that just show the difference between the ideal case and the real one. This situation can be prone to misunderstandings and the biases of coaches. The intuition of these researchers was creating the benchmark, the expected behaviour, in an objective and potentially error-free way, learning how an "average" hypothetical player would have done in the same situation. This concept is a sort of application of the principles of the "Wisdom of the Crowd" that affirms that large groups of people are collectively smarter decision-makers respect to any individual when an aggregation happens. So, if in a specific game situation, the 90% of players did a certain play, we are implicitly assuming that that play could be the right one.

Even if data-driven ghosting was born in basketball, also some application in football have been developed using the same concepts. An example is the one below.

In *'Data-Driven Ghosting Using Deep Imitation Learning'*, by Hoang M. Lee, Peter Carr, Yisong Yue and Patrick Lucey, has been presented an application of ghosting systems dedicated to football. The system has been created to build highly detailed models of players and team defensive behaviour.

Conceptually the defence in football is really complicated to be modelled because it is a collective and synchronised action of many players and, further, because it is difficult to properly evaluate the effectiveness of single defensive player's move: even if a team is not able to take the possession it can

---

[114] *http://grantland.com/features/the-toronto-raptors-sportvu-cameras-nba-analytical-revolution/*

"*force the opponent players in a poor location of the field, where they cannot anymore be dangerous"*[115].

The system presented in this paper is based on Deep Learning algorithms. In particular, it uses an Imitation approach. The concept of Imitation Learning has been already presented in the paragraph 1A.12, where the functions of AlphaGo has been presented.

The first objective is make the computer system able to manoeuvre a football team autonomously on a 2-D field while it is defending against a real opponent. To teach the system how to do, the authors used an Imitation Learning approach. In fact, "explaining" to the system how to do it, writing down the rules of football defensive art would have been a task too much difficult for the complexity of the matter and for the completeness of the information. They faced the limit of the "Polanyi's Paradox": explain how to behave on a football field is something that we are not able to express verbally. Even the human young players, when they approach football, they learn how to play imitating adults or peers.

The same conceptual approach has been used for this model. The model was trained giving in input information about the defensive spatiotemporal position of all the players in 100 football matches. Given these information, the system understood how to play football. So, if the system is capable of "playing" football, it can also in real-time guess the future evolution of the game, basing on the recent history of all players moves. The authors introduced also a "Training system", meaning a technique that can make the system improve from its own prediction mistakes, minimizing the opponents' probability to sign a goal.

With this approach the authors created a system "*that learned to maintain solid defensive formation and structure… in a manner that shows spatial and formational awareness"*[116].

After having created a system able to learn how to "play" football the applications are numerous, and they differ mainly in function of the data used to initiate it.

---

[115] *'Data-Driven Ghosting Using Deep Imitation Learning', by Hoang M. Lee, Peter Carr, Yisong Yue and Patrick Lucey*

[116] *Data-Driven Ghosting Using Deep Imitation Learning', by Hoang M. Lee, Peter Carr, Yisong Yue and Patrick Lucey*

If the objective is evaluating in absolute term how a team defend, it is possible to initiate the system with data from many teams and many matches and so create a benchmark that can represent, at best, the average team, as explained at the beginning of this paper's explanation.
But it is also possible giving in input to the system the data from a unique team. Doing it, the aim is basically build up a system able to forecast and reproduce, how that specific team will probably face a certain situation, and, this could represent a big competitive advantage. It potentially means knowing in advance how the opponent will probably behave.

Coming back to Basketball applications, in "*Bhostgusters: Realtime Interactive Play Sketching with Synthesized NBA Defenses*", Thomas Seidl, Aditya Cherukumudi, Andrew Hartnett, Peter Carr and Patrick Lucey introduce a system, taking inspiration from the work explained before, solving, at the same time, some of its criticalities. They created a system based on data-ghosting for in-game decision. In other words, a system where the coaches though a digital interface can draw down their schemes and plays and instantly see how the opponents are likely to respond.
Basket is a sport that is strongly based on pre-determined plays. This system permits coaches to prove quantitatively their intuition about possible plays to engage in a match, testing them against a system that "learned" how the other team play.
In this case, the system is specifically created to reproduce and forecast the behaviour of a single team. They improved the model because they introduced relevant factors in the analysis like score situation and fatigue.

The last example of ghosting applications is about how to evaluate players' decisions in NBA. It is presented in "*The advantage of doubling: a deep reinforcement learning approach to studying the double team in the NBA*" by Jiaxuan Wang, Ian Fox, Jonathan Skaza, Nich Linck, Satinder Singh and Jenna Wiens. The paper proposed a deep learning model expected to find the optimal strategy to doubling in basketball. Doubling is a defensive strategy that brings two players to guard a single offensive player. The decision of doubling is basically about solving a trade-off between doubling, and increasing the possibility to steal the possession, and the risk of giving up easy shots to the offensive players that remained free. As usual, the systems get in input the spatiotemporal data of the numerous NBA in which a doubling strategy has been applied by the defence. On these data, the reinforcement learning framework started to quantify the relationship between the court's configurations and the decision to double or not and whom to leave open, and the outcome, in terms of point per possession. From this, the system was learned to minimize the points conceded to the offensive team. Doing this, the system can develop new strategies for doubling and understand which one is the optimal in any specific situation.

The input data can be enriched with several new information like the player heights, weights, shooting abilities, shot clock and game clock.

The system reached significant results. For example, in case in which the ball-handler is far from the basket, it noticed that the common rule to leave free the offensive player much far from the ball-handler is not so effective, while it is better to leave alone the player closer to the basket and use two players to block potential passes to the open man. It also learned that statistically is better to not double a star player, instead it is more effective double normal players. The explanation proposed by the authors is that a top-player is able to face a doubling strategy, and, in any case, he has good probabilities to emerge as victorious, signing points, while an average player can face more difficulties to overcome a doubling strategy and, thus, the defence can be more effective.

All these models are examples of academic works in tactical analysis based on artificial intelligence systems. They have been presented to show the main features and objectives of AI-based system in this field. There are also already existing companies that are offering products, specifically dedicated to support the tactical development of the team.

Even if some of them can operate in a slightly different way, the general approach is the same. Starting from the spatiotemporal data of the players, the system developed useful insights for the coaches about the 1. the decision-making capacity of the players, 2. The quantitative evaluation of singular and collective moves and 3. The forecast of the opponents' behaviour. Examples are Metrica Sport[117] for football, Icerberg analytics[118] for hockey, Second-Spectrum[119] for basket and Math&Sport[120] for volleyball and football.

---

[117] *http://metrica-sports.com/*

[118] *https://www.iceberg.hockey/*

[119] *https://www.secondspectrum.com/*

[120] *http://www.mathandsport.com/*

2.1.5 COGNITIVE TRAINING

The athlete final performance, as said highlighted in section 2.1, is affected by the physical condition, by the tactical preparation and the technical abilities, but also by the psychological readiness and preparation of players. Cognitive abilities depend on different aspects that can be internal or external. Internal mainly refers to talent of the players. External ones are, instead, related to the environment situation like pressure, stress, importance of the match, relevance of the opponents etc.

This is the next level of performance assessment. Once it is possible to create matrixes that can show in an omni-comprehensive way the abilities of players, the second step is evaluate why they changed in specific matches and in relation to which factors. This is what, for example, Math&Sport and Wallabies are involving in their products. They will be presented more in detail in the case studies.

2.2 MANAGEMENT AND ORGANIZATION

2.2.1 SUPPORT TO TEAM-RELATED ACTIVITIES

Scouting and Player Investment Evaluation and Optimization

in this category lay all the solutions that are dedicated to the corporate management of the athletes. They mainly deal with scouting activities and players-related investment optimization.

Sport players' market is a market where the level of imperfections is high. This makes it inefficient and so the price of players, usually does not reflect their intrinsic value. Some of the reasons are listed below:

- The asymmetric information between the players' seller and the buyer clubs is relevant, in terms of knowledge of physical conditions and technical/tactical behaviours of the player. This means that the buyer knows much less respect to the seller about the expected performances of the players. Imaging a limit situation, in which a player suffers physical problems, and, the seller will not reveal his condition to the buyer, before the closing of the deal, to not lose potential extra earnings.

- The low liquidity of the market, intended as the fact that the transactions happen una-tantum, and, through a private negotiation, between the two counterparties. So, it does not exist a sort of public benchmark to evaluate the fair value of a player. In other words, does not exist something like a stock exchange for players, where a public price, that reflects all the information available, is uploaded in real-time by the actions of all the agents.
- The evaluation of players depends on many variables. Some of them are easy-to-evaluate like age, role, personal history, injuries etc. other are more difficult to be analysed before the beginning of the season. For example, there are several cases in which a player, that was used to play very good in a team, decided to change the club, and, it happened that he was no more able to express what was used to do before the trade.
  This can be due to many reasons, like psychological aspects, different environment and lifestyle. But also, because the performance of players is highly related to the other players' playing style. So, it could happen that in the playing environment of the buyer team, the new player cannot express his maximum value.
- The fact that even, if there are millions of football players, there are few cases in which two players are 'Perfect Substitutes'. This means that the needs of a specific team, can be satisfied just with some few players. For example, if a team needs a midfielder with the ability to kick the free kicks, with the constraint to be, for example, Italian, for limits imposed by the regulation, the available choice will be no more than few players. Considering that some of them, for different reasons, will not accept the trade, it is likely that there will be just one or two players, that meet all the constraints. In this case, the seller is in a monopolistic position and can fix the price with no upper limits.

All these issues create big misalignments between the fair price of players and what club actually pay for them. The difference is even larger if the wages of players, that are getting always richer, are considered.

The artificial intelligence can be successfully employed to solve these issues, giving precise insights to clubs about the fair value of players.
In fact, artificial intelligence can bring into the players evaluation process 3 main benefits:

1. Objectivity in the evaluation of performances. As it has been highlighted, in the section 2.1, artificial intelligence systems have been employed to develop tools that can evaluate in objectives terms what until now was possible to do only in a heuristic way. If these matrixes will be applied on third-party players, buyers' team could get clear information about the right price of players that are under consideration to be acquired. Ideally, it is also possible to run simulations about the future level of integration of team and the expected performances of that player in the new environment. The goal is avoiding any error in the forecast of the players' potential impact.
2. Starting from these considerations and, monitoring all the transaction in the markets, artificial intelligence systems could suggest and compute an ideal fair price, for the purchase and for the annual remuneration of players[121].
3. Enable also the inverse process, meaning that the team could no more searching for players, but instead insert in a query-platform the features that it is searching for, like age, role and other characteristics. The system will give as output the name of the possible candidates. This last application is more related with scouting.

This is exactly what Wallabies, a Milan-based start-up is trying to develop. They are working on a system able to identify for any players and all his comparable in term of technical performances. The idea is to use the market values of the comparable to give a price to the initial players. The activities and operations of Wallabies will be deepened later, since one of the case studies has been dedicated to them.

The first case in sport's history in which a team exploited the inefficiencies of the players' markets, to outperform competitors was in 2002 when Billy Baene[122], the general manager of Oakland Athletics, an US baseball team, employed in a systematic way quantitative analysis, to find and sign the players that would have been more valuable for his team. He was able to achieve successfully results, respect to the

[121] https://www.wallabies.it/

[122] *MoneyBall, Micheal Lewis, 2003*

other teams in the leagues in term of capital efficiency. He obtained more points at the end of the season, with an initial budget that was much smaller respect to others.

In modern day, the multinational information company Sap, has an information system, specifically dedicated to football clubs, in which there is a functionality dedicated to scouting supported by machine learning[123].

Scisport[124] is a Netherlands-based company, that offers an online platform for scouting purposes. It is equipped with machine learning algorithms, able to:

1. search for players with the inverse process explained in point 3 above.
2. They offer a functionality to discover undervalued players worldwide
3. A new evaluation tool based on the influence of players on their team results.[125]

2.2.2 SUPPORT TO COMMERCIAL ACTIVITIES

Sponsorship Value Estimation

This category has as target all the processes that aim to monetize directly the activity related to the sport team. AI-based solutions have been mainly employed for sponsorships evaluation and their value optimization.

The traditional sponsorship model is getting outdated for both sport teams and commercial partners. The pillar over which the traditional sponsorship contracts are build up is: 'the more the team win, the more the commercial partner will pay for the sponsorship'. The assumption is that the exposure of the sponsors' brand depends directly on the success of the team, and so the value of the contract. This was true in an off-line world, where sport media contents were transmitted just on TV. The televisions were used to broadcast the more successful teams and, so, the previous assumption was right.

[123] *https://www.sap.com/products/sports-one.html*
[124] *http://www.scisports.com/products*
[125] *http://www.scisports.com/products*

But now, sport contents are mainly distributed and consumed throughout digital channels. Internet allows to watch sports' event whenever fans want.
The value of a sport sponsorships, now, do not depend anymore only on the direct exposition on the TV and, so, the statement that links sport success to brand exposure cannot be anymore the basic assumption over which develop the partnerships. This creates the necessity to change the paradigm behind sport sponsorship value estimation.

A practical example is that the highlights of significant moment of the sport matches are shared and seen through social medias. So, a team can lose a game, but, if one of its players will perform a high-quality technical gesture, the video will be probably shared world-wide for weeks, and, the exposure of the sponsor over the t-shirt of that player will be significant higher respect to an "anonymous" win. New sponsorships paradigm must consider the real and the digital exposure of a brand during an event.

Another criticality with the traditional way in which sponsorships are evaluated is that the commercial partners, the ones that finance the sponsorship contracts, are required to justify their decision and bring evidence that a certain sponsorship contract will have a tangible return [126]. So, both the counterparties involved in sport sponsorships, clubs on one side, and commercial partners on the other, have the necessity and the incentive to introduce more detailed and objective matrixes in the evaluation process, that can better fit the new environment.

In practical terms, social media-based data creates the opportunity to have real-time insights about the reaction of the public to the specific brand. So, now it is possible to evaluate the answer of people to brands exposure.
Examples of companies that are operating in this sector is GumGum[127]. A computer vision and AI-based analytics company that created a system to monitor over all the channels the overall exposure of a brand monitoring TV broadcasts, streaming video and social conversation[128]. In fact, all the media content in which appear the brand of a commercial partner must be considered in the evaluation of the

[126] *Commercial trends in sport 2017, Nielsen*
[127] *https://gumgum.com/*
[128] *From the official website*

marketing efforts, with the aim to compute, in a holistic way, the evaluation of the total brand exposure.

The figure 26 represents the new conceptual approach to sponsorship evaluation. If in the past the value of the sponsorship was directly related to the victories of the sport club, the approached proposed by the newcomers is different.

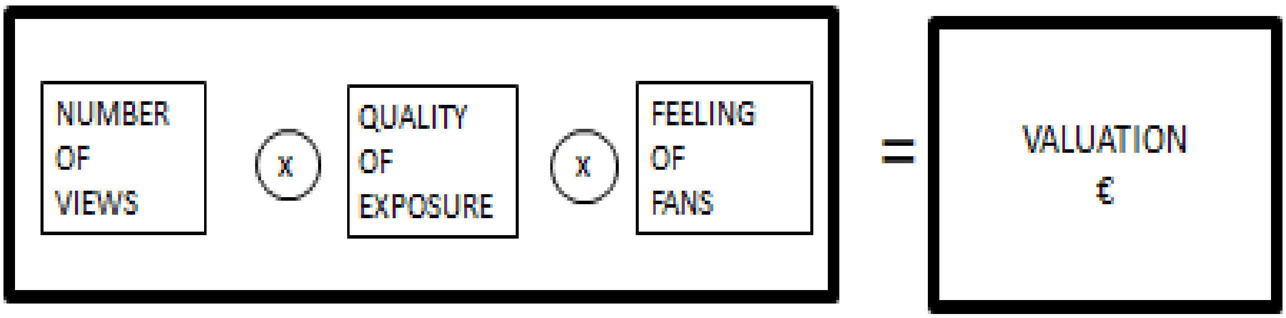


Figure 26: the formula represents the conceptual approach used for sponsorship evaluation.

Here the sport results even do not enter in the computation. The idea is computing the number of time that the brand has been seen by the public. Potentially, it is possible to consider every type of media content. Real-time TV' shots, post-match highlights, social media videos and every other type of means through which a company's brand can be view from the public. AI can play a fundamental role in estimating the real number of views and attention generated by media contents, in which a specific brand is present. The idea is to sift all the channels, TV, social media, online platforms and counts the viewers.

The second element to take into considerations is the quality of the brand exposure. Measuring the clarity of the exposure, so how clear people can see the brand image, for example evaluating the size of the logo compared to the rest of the frame. If it is in front and centre, the visibility changes respect if it is partially hidden in the background. For example, GumGum consider the percentage of logo that can be seen, and, if other brands are in the same frame as a meaningful proxy of the exposure quality[129]. Considering all these aspects, the AI system should measure the level of visibility.

The last step is measuring in monetary term the benefits generated for a company through his brand exposure. It can be done evaluating the feeling of fans regarding that brand. The idea is to leverage on AI for performing in-depth sentiment analysis and about fans feelings and response to brand exposure. If a brand has received consistent appreciation after the exposure in an official match, is something that must be taken into account for the final evaluation. On this topic, AI solutions have already demonstrated their value. Companies like Stride[130] and Aylien[131] have created solutions (the application regard other industries, but the working methodologies is the same) able of carrying out successfully sentiment analysis analysing written posts.

This is the qualitative explanation of how sponsorships' value can be better estimated using AI-based analysis. The idea is transforming real-time the brand time exposure and the people reaction into their equivalent economic values for companies.

Regarding the sponsorship evaluation, Other companies like BlocksixAnalytics[132] and Hookit[133] are proposing similar AI-powered services. On the field of fans' reaction analysis, probably, the best player on the market is IBM with its Watson applications' system. It can analyze millions of unstructured data taken from any on-line and off-line channel, and, develop meaningful insights. For example, IBM has run a partnership with Wimbledon to revolutionize the way in which the tennis fans are engaged[134]. IBM employed its systems to analyse all the tweets and other social media posts related to Wimbledon,

---

[129] *The social side of sport sponsorships, GumGum Sport*
[130] *https://stride.ai/*
[131] *https://aylien.com/*
[132] *https://www.blocksixanalytics.com/*
[133] *http://www.hookit.com/*
[134] *https://www.ibm.com/blogs/watson/2017/07/ibm-watsons-ai-is-powering-wimbledon-highlights-analytics-and-a-fan-experiences/*

providing information to the organizers about how to maximize the online traffic and the transaction volume. Similar partnerships are carried out also with US Tennis Open[135] and main golf events[136].

Non-Traditional Fan Monetization

In this section some AI applications will be presented that are not yet widely used in the sport industry. They are inspired by other businesses, but, with some adaptations, could be employed successful employed also in sport one.

The first solution is related to dynamic pricing. Dynamic pricing is a technique widely used, between others, in the airlines and in the hotels.
Companies can change prices based on real-time algorithms-based evaluations, that take into account several factors, like competitor prices, weather conditions, market trends and potentially all the variables that directly or indirectly affect the demand curves (the supply usually is fixed).
It is necessary to highlight that the sport ticketing can be considered less complex respect to the airline ones, for example. Airlines tickets' demand is more elastic respect to the sport tickets' one. If a person is planning a vacation and the plane tickets are too much expensive, he can also consider other destinations. If a person is a fan of a specific club he will not go to see a different club's match just because tickets are cheaper.
Nevertheless, also in the sport business we are witnessing to an increase complexity related to ticketing processes. For some events there is a demand that is much higher respect to supply, and, for others, the situation is the exact opposite. In the first cases, there are problems linked, for example, to re-selling activities that can damage the club brand, and in the second, there are empty seats at the events, that represent a loss.
Introducing dynamic ticketing in sport industry could on one side, maximize the economic return of organizations, charging, potentially, the highest prices people are willing to pay, and on the other side, decreasing prices when the demand is not high, incentivize fans to go to stadiums.

[135] *https://www.ibm.com/sports/usopen*
[136] *https://www.ibm.com/sports/masters*

AI can be employed to compute the dynamic optimum price thanks to the algorithms capacity to analysed big amount of data and to be reactive to various factors.
An example of companies that are developing smart ticketing products for sport industry is Qcue[137] and Digonex[138].

Another AI-powered solution common in different industries but still not introduced in sport are the Chatbots. Chatbots are computer programs able to conversation with people. They can be introduced in the official app of sport organizations, to support all the processes, from ticketing, to online purchases and the arena experience. A further idea is linking the club' app with external stakeholders and use the official chatbot to give information about traffic to people that are going to the stadium or inform about offers in the arena's sales points and about other initiatives, eventually, not directly linked to sport activities.
Chatbots can be also used to solve on-line problems and to managing the messages' applications answering to commonly questions. Chatbot for their nature of dealing with potentially new problems must be powered by AI technologies.

AI can be also used as feedback instrument. Systems can track huge volumes of on-line activities on the social networks and extrapolate feelings and opinions. These insights can be used to evaluate the effectiveness of all the sport clubs' initiatives, measuring the response of fans. An example is what is doing Symanto, a German start-up[139], that was born to develop AI-systems capable to decipher the humans' feelings and emotions through the analysis of written text, like messages. Now they are collaborating with some German sport organizations, like Schalke 04 and Bayern Munchen to "*gather feedbacks from fans and understand their need*"[140] analysing on-line fans' conversations on the most common social netoworks.

[137] *http://www.qcue.com/*
[138] *https://www.digonex.com/*
[139] *http://fanalytics.symanto.net/*
[140] *http://fanalytics.symanto.net/*

2.3 FAN & MEDIA

2.3.1 MEDIA AND BROADCASTING AND DIGITAL ENGAGMENT

The modern trend in sport business is considering it always more as entertainment, introducing in sport broadcasting and sport-related media creation some of the best practices of other similar sectors, like cinema, music events management and commercial expositions.

AI and other technologies like AR, VR, IoT can be used for both supporting the creation of traditional media content, like video, but also to deliver new kind of media contents, with the aim to deliver higher value to customers. AI, in particular, can enhance the value of the media content, making them more meaningful and engaging for the fans. This scenario creates the opportunity for sport teams to increase the interactions with final clients.

Automatic Media Content Creation

An example is WSC SPORT[141]. It has developed a Learning Machine system to analyse live sport events and create automatically, and in real-time, short videos. The input of the system is the live TV coverage of a sport event. The system, basing on several factors like cheering of fans, commentator analysis and players stats, is able to distinguish and isolate only the most valuable moments; The ones that will be probably most appreciated by fans, and create sharable videos completely autonomously. Similar solutions are also provided by Wildmoka[142] and Reely[143]. These products create advantages in term of efficiency, because they perform the task of media creation completely autonomously without human intervention, with a consequent cost reduction effect. They also perform effectively in term of “time to market” because they can do it in real-time, making the content available on social networks immediately after when it happened in the reality. In fact, some of the companies presented above offer also, as additional service, the management of the club’s social platforms.

[141] *https://wsc-sports.com/*
[142] *https://wildmoka.com/*
[143] *http://www.reely.ai/*

Even if there is no public information about the price of these service, it is possible to notice that these companies collaborate manly with the top sport' clubs in the world, with big sport events' organizer, or with multinational media providers. They are solution that are targeted for the top-level actors in the market and are not economically feasible for smaller organizations.

Media Technologies for Amateurs

On the opposite side, there is a segment of the market that is specifically dedicated to small clubs even belonging to non-professional levels. An example is the solution proposed by Veo[144], a Copenhagen-based start-up, that have developed an AI-powered video camera, that automatically follows the actions on sport pitches. The system is based on a machine learning algorithm able to follow the ball on the field and, thus, record the most interesting actions. It is created to be affordable. The idea is to offer media coverage also to events, that have no the appeal to attract big crowds. Solutions like this one are targeting the non-professional levels and they can have a big impact in the way in which amateur sport or young leagues are managed. A similar solution could enable also in these categories the digital revolution. They can impact also the sport that are commonly considered as minors, giving them media coverage at a very low costs. Obviously, these video recording must be shared on social networks like Facebook to be easily accessible to everyone.

Media Contents Based on Athletes' Data

AI permits also to deliver completely new media contents. All the data, analysis tools and indicators presented in the section dedicated to the athletes' performance (2.1), can also be also delivered to the fans for enabling a higher understanding of the game. Potentially, all the data that could be useful for coaches, have value also for commentators and fans. For this reason, Sportlogiq[145] is a sort of hybrid solution in the sport environment. It is a company that offers both AI-powered solutions to support the

---

[144] *https://www.veo.co/*
[145] *http://sportlogiq.com/en/*

team managers about the definition of the tactics and, at the same time, it aspires be a partner of the media broadcasters' companies, to developed more engaging and meaningful media.
The idea is exploiting scope economies and the convergences present in these two areas. The underlying data are the same. On one side they can be exploited by teams, and, on the other side, by broadcaster for fan engagement purposes. In fact, the company collects athletes' performance data and their historical evolution. Basing on these data, it employs AI-based systems to get insights that are meaningful to sport organizations. But, also fans could be interested to have a higher understanding of the athletes' condition and knowing possible future states of the match. So, these analyses can be used to enhance the value of the media delivered. Exactly for these reasons, the company decided to target both.

Active Media Contents

The data collected during the matches can also use to propose active media content to fans. This idea idea has been introduced in the paper "*Bhostgusters: real-time interactive play sketching with synthesized NBA defences*". It has been introduced in section 2.1.4, speaking about tactical applications of AI. The work proposed a Machine Learning system able to understand how a specific team plays, and, use this predictive capacity to simulate how that specific team will probably react to certain schemes or plays. Starting from this idea, the authors imagined a hypothetical media broadcaster providing to the fans a tool like the one just presented. The system proposed can guess with a high probability how a team would have behaved in facing a certain situation in basketball. If media broadcasters make available this system, the fans would have the opportunity to analyse "actively" real-game situations of the matches and explored the "what if" scenarios. It is possible imaging that a certain basketball player decided to shoot the ball and he failed. The fans could re-live that moment through a digital interface like the one proposed in figure 24 and fake to pass the ball, taking a different decision respect to the one took by the player in reality, and see how the situation would have probably evolved. Similar technologies can be used also to create and develop videogames that can be always more loyal to reality.

Tracab[146] is another big player in this market. It relies on an advanced camera-based tracking system to deliver high value media content. The camera system can track every move of the players on the field and then the computation system elaborates this information to create advance media content, like, for example, heatmaps of players presence, or, the trajectories followed during certain actions, in a completely automatic way.

VR-Based Media

Together with the other technologies already mentioned, AI can also provide innovative media content. Example is Headcase[147]. It is a VR company that creates 360° and 3D media. The company is specifically dedicated to the creation of story-telling short films in Virtual Reality. another example is Taqtile[148], a company specialized in augmented-reality based products, that started to create holographic maps available through specific viewer of golf field. They collaborated with Microsoft for PGA golf tour to create holographic maps of the golf courses as additional service for high level customers.

A solution that has also great value in the fan engagement is STRIVR[149]. It has been already mentioned in the technical analysis section (2.1.3). The company provides a technology that allows athletes to re-live practices and games through VR viewer, giving them the opportunity to be back on the field in first-hand and have the same visual field they had in the real time. The same system can be tailored for customers to live the sport event completely immersed in it. It is possible to image to watch a match from the perspective, or, with "the eye" of a player and not with the traditional TV shootings.
For example, Intel[150] is developing a specific set of camera-based technologies dedicated to the digitalization of sports, with the objective to improve the fan experience. The company is developing media content based on VR, where fans can immerse themselves in the game, visualize on the screen statistics and data that they want, change the point of view, and, recreate at home the view field of players.

---

146 https://chyronhego.com/products/sports-tracking/tracab-optical-tracking/
147 http://www.headcasevr.com/
148 https://taqtile.com/
149 https://www.strivr.com/
150 https://www.intel.com/content/www/us/en/sports/sports-overview.html

The process of digitalization of sport does not pass only through new ways to reproduce images and videos but also a through the delivery of new engaging information to the public. So, it is probably to assist to the introduction of always new sensors or detectors that can measures a wide set of performances and transmit them immediately to the fans in real-time.
The objective must be making every fan free to choose the most appropriate way to live the sport event. Nowadays, with global markets also in the sport business, not all the fans are the same, or want to get engaged in the same way. Teams need to segment the fans base and create ad hoc digital strategy for each category.

2.3.2 SMART-STADIUM AND STADIUM-AS-A-PLATFORM PARADIGM

In the arena management, AI has the potentialities for offering solutions that can successfully address the needs of fans. As it has been also seen in the media section, AI enables the delivery of much more value to clients, thanks to innovative services, personalization and targeted activities. On the other hand, it can also facilitate the management of the crowd flows, inside and outside the arenas, optimize the management of the traffic and the implementation of securities systems not invasive for the fans.

In this market, the players are mainly the big companies that developed technological systems to manage complex facilities like IBM[151], Intel[152] and Parasonic[153]. They usually are required to develop the information system of the stadium and, in some cases, they have also introduced owned or third-party AI-based solutions. It is possible to highlight that majorities of these solutions are polyvalent, meaning that can be applied also in other contexts without problems, like commercial buildings and hospitals.

[151] *https://www.ibm.com/industries/government/infrastructure-citizen-services/buildings-facilities-management*
[152] *https://www.intel.com/content/www/us/en/sports/sports-overview.html*
[153] *https://na.panasonic.com/us/industries/sports-entertainment-media*

For example, the access to the venues can be speed up by AI-based security systems that enable clients to get in, using, for example, digital fingerprint recognition, or even face recognition. Solutions like these ones would avoid all the securities and tickets checks.
There are solutions[154] that can support the smart parking management, for example. The idea is to direct fans arriving to arenas to free parking spaces, without losing time to search for it. Other similar applications can split in a homogenous way the cars on streets, to avoid traffic jam. Before the game, clubs can also create promotions and events to make fans come early to the stadium enhancing revenues for the sales of products and reducing traffic congestion. AI can be employed to find the best way to meet these purposes, analyzing fans behaviors and preferences.
AI can support the proper targeting of the commercial promotions and deliver insights and suggestions to fans, within the arena over which shops to visit.

During the event, the new paradigm is related to the "Second screen experience". So as said during the section dedicated to media (2.3.1), stadium apps should also become the vehicle to deliver other media content like replays and video. In this case, AI should play the role to permits to fans of enabling the personalization of the contents.

The direct digital interaction with fans permits also to carry out local sentiment and emotional analysis on fans feelings. A company specialized in local social media analysis is Ampsy[155]. It was born in the music and entertainment industry and it just entered the sport business. "*it provides two things for brands: high-level media insights as well as sentiment analysis around an experience"*[156]. The sentiment analysis is done through a partnership with IBM Watson technologies. The analysis does not stop at the end of the game, but ii goes on until weeks after the game, to get a comprehensive insight of the event impact on fans[157].

---

[154] *http://vimoc.com/product-2/*
[155] *http://ampsy.com/*
[156] *https://www.sporttechie.com/ampsy-helps-sports-teams-gain-insights-into-audiences-using-social geofencing/*
[157] *https://www.sporttechie.com/ampsy-helps-sports-teams-gain-insights-into-audiences-using-social-geofencing/*

NAME OF THE START-UP	WHAT DOES THE START-UP?	AI TECHNOLOGY	BENEFITS FOR CLUBS
VIRTEX APP	Solution that enable multiplayers AR competitions, set in the arenas, where fans can challenge each other.	AR / VR	+ FAN ENGAGEMENT + SECOND SCREEN PARADIGM
IDEMIA-INDENTGO	security-related products, as Digital fingerprinting and face recognition	Face Recognition	+FLOWS OPTIMIMAZTION
SNAPTIVITY	Robotic cameras, installed in stadium that are triggered by sensory networks, which capture only candid fan reactions at stadium. Each fan at a stadium received his/hers best in-game photos.	Face Recognition	+ SECOND SCREEN PARADIGM + FAN ENGAGEMENT + ADDITIONAL PAY SERVICES
BRIZICAM	fan controlled photo cameras, installed inside stadium to get picture from different point of view.	Face Recognition	+ SECOND SCREEN PARADIGM + FAN ENGAGEMENT + ADDITIONAL PAY SERVICES
SATISFILABS	Chatbots for sport events	Chatbots	+ FAN ENGAGEMENT + SECOND SCREEN PARADIGM + DELIVERY OF AD-HOC MARKETING STRATEGIES

Figure 27: Some AI-based solutions for arena management

This concept is linked to the one expressed in figure 26, where the new matrix to quantify the economic impact of sponsorship has been introduced. The same concept can also be applied for in stadium sponsorships. The idea is to employ the same formula expressed in figure 26, but for in-stadium brand exposure.

In the next table some AI-based companies, that have developed solutions for arena management, has been presented, along with their technologies.

As highlighted also in the section 1B.1, where the sport arena management has been introduced, the experience has become the fundamental part of every engagement activities with fans or clients. People are willing to pay for meaningful experience. Sport clubs need to find new ways to make the spectators' experience unforgettable.

Basing on these considerations, it is likely that in the next years it is possible to witness to the introduction of big entertainment systems in the sport arenas. Audio systems, huge screens, lighting systems and even special effects systems will be probably used in sports' arenas to generate strong emotions in spectators. The Artificial Intelligence will be probably employed as support for management of these systems. The inspiration source should be what is now offered in the thematic parks to the public[158].

2.4 EXTERNAL STAKEHOLDERS

2.4.1 SPORT BETS AS AN ASSET CLASS

The increasing amount of data available on any aspect of the matches, and the numerous platforms through which it is possible to bet about sport events' outcomes, have induced some companies to look at sports' bet as a purely financial investment. This is a sector in which AI and Machine Learning applications have created the premises for a fast development.

As it has been seen in the performance analysis part (2.1), it is possible to employ models to forecast how a certain sport match or competition will end. The sport betting system, now, offers a myriad of other events over which it is possible to bet, not only the results. Thus, the possibilities of betting are numerous.

For these reasons, some companies have developed systems, that leveraging on AI, try to forecast the outcomes of sport events and earn the betting premium over it.

Priomha Capital is one of the biggest edge funds in the world specifically dedicated to the sport betting as asset class. Brendan Poots, Priohma's founders, pointed out that the main opportunity related with sport betting is that "*they are a completely uncorrelated asset class. No political, economic and financial events affect the results. They are also supposed to be recession-proved*"[159] .

Their long-term vision is using sporting betting to diversify the financial risk in portfolio management thanks to the reasons explained before.

[158] https://www.disneyresearch.com/labs/

[159] Economist, hedge funds try to promote sports betting a san asset class

Another company that is operating in this market is the London-based start-up Strategem[160]. It has a wider product range. It is a sport betting trading-and-investment platform. Clients can receive insights and recommendation based on Machine Learning algorithms over the future outcomes of sport events and investing money in real betting. It has also a section dedicated to asset management specifically dedicated to institutional investors.

Another actor in this market is Unanomous AI[161]. It is a company that basically offer forecasting services, but its approach is quite unique because it leverages on swarm AI algorithms. Swarm Intelligence refers to the fact that it is possible to exploit the collective behaviour of people to predict some event.
The starting points of this method is the collective wisdom created by online human groups. In "*Amplifying prediction accuracy using swarm AI*", a white paper sponsored by the company, is showed that individuals who averaged 55% accuracy when working alone were able to amplify their accuracy to 72% by forming swarms, leveraging on company's owned algorithms.
So, in this case, Artificial Intelligence systems are not used to forecast an event's outcome but to extrapolate useful information by common people interactions.

Regarding the thematic of collective intelligence, there are academic works, in which have been proposed systems that leveraging crowds' behaviour, are able to develop effective predictive systems. In "*beating the bookmakers: leveraging statistics and twitter micro posts for predicting soccer results*", Freeric Gdin, Jasper Zuallaert, Baptist Vandersmissen, Wesley De Neve and Rik Van de Walle developed a method to aggregate and extract the information contained in over 50 million of twitter's micro post to predict the outcomes of soccer games, realizing theoretically a profit of 30% in betting. As in the example before, the algorithms are employed to get insights from the crowd, not to operationally forecast a sport outcome.

[160] *http://www.stratagem.co/*
[161] *https://unanimous.ai/*

There are also cases in which researchers have tried to develop artificial intelligence systems specifically dedicated to the forecast of the results itself. Naturally the outcomes of sports events are complex to be forecasted due to the numerous variables that can affect the result. So, for these applications the unique approach possible is based on neural network and, in particular, deep learning systems, whose multi-layer structure permits to build up causal links between basic information and the outcome[162].

The firsts attempt to use neural network for sport forecasting are dated at the end of the '90, when Purucker (1996) conducted one of the initial studies on predicting results in the National Football League (NFL) using artificial neural network models[163].

Tax and Joustra (2015) tried to create a prediction system on Dutch Eredivise's outcomes basing on thirteen seasons past match data, with the specific aim to create profitable betting strategies. Davoodi and Khanteymoori (2010) created a model to predict results in horse racing.

In "*A machine learning approach to predicting winning patterns in track Olympic cycling omnium*", the authors[164] proposed a model whose aim was to define which were the characteristic that most were required to win a medal for this Olympic discipline. There are also cases related to swimming (Edelmann-Nussser 2012) and golf (Wiseman 2016).

To the best of my knowledge, the unique paper written whose objective was comparing the performances of these algorithms against best human' experts is *"Human decision making and artificial intelligence: a comparison in the domain of sports prediction"* where Arnu Pretorious and Douglas A. Parry compared a particular class of artificial intelligence algorithms (Random Forest Classification) with the performances of human experts. It was written in 2015 and it demonstrated that specifically for the forecast of the rugby world cup of 2015, there were no evidences to suggests that the human prediction capabilities outperformed the machine learning approach.

So, if the traditional improvement curve of artificial intelligence systems presented in 1A.11 is assumed as valid, it is possible to approximately affirm that in 2015 AI-based systems reached the human

[162] *Artificial Intelligence in Sports Prediction, Alan McCabe, Jarrod Trevathan*

[163] *A Machine Learning Approach to Predicting Winning Patterns in Track Cycling Omnium, Bahadorreza Ofoghi, John Zeleznikow, Clare MacMahon, and Dan Dwyer*

[164] *Bahadorreza Ofoghi, John Zelezznikow, Clare MacMahon and Davon Dwyer*

performance level, and so, probably, the 2018 systems are generally better than human experts to forecast sport's outcomes.

2.4.2 REFEREE DECISION SUPPORT

An application field where AI can have relevant and positive consequences is the related to Referee support. It has been already said that the referee tasks in sport is controversial. They cannot reach a 100% accuracy in decisions' and their activities can be highly impacted by subjectivity. In some sports, like artistic gymnastics or figure skating, the results of competition are the outcomes of juries votes and the process can be quite opaque.

The AI has the big advantage to make objective what since now has been possible to be judged only on subjective way. The technologies have the potential to learn the rules to evaluate certain movements and associate them a score, in an objective and scientific way. AI have the potential to perfectly oversee the games and take always the more appropriate decision regarding what is happening.

Currently the role of referee is too much complex to think about replacing it with AI systems. The audience and probably even the technologies are not ready to run this change. But in any case, the situation is evolving. The International Gymnastic Federation (FIG) is planning to introduce an AI system to assist jury with scoring at Tokyo 2020 Olympic Game. The development is made in partnership with the IT company Fujitsu. Even if the system is still in a phase of research and development, it is said to work with a 3D sensory camera system, able to capture all the athlete's performance[165].

The systems aim to speed up the judgment process, make it more objective, decrease the biases and increase the coherence between different votes.

The system will probably have some difficulties to be effectively used, because gymnasts, trying to impress the jury, introduce always innovative moves that has never been seen by the computer, so theoretically it should be able to evaluate completely new actions maintaining the same judgement parameter.

---

[165] https://www.theguardian.com/sport/blog/2017/nov/04/ai-judges-gymnastics-olympics

2.5 PART II CONCLUSION

From the Part I it emerged how AI's value is related mainly to his ability to support human decision-processes and overcome the limits of bounded rationality.
Part II represents the mapping phase of all the AI-based solutions currently available in the sport business and how these systems have been employed to improve processes within sport organizations in all the areas found.

To address the purpose of the thesis and validate the knowledge acquired in this first two parts, is it necessary to deepen the knowledge of some real-world cases, to better delineate the working methodologies and the value proposition. For this reason, the next part will be dedicated to some case studies.

The first one will be dedicated to Luca Pappalardo and Paolo Cintia. They are both researchers and they have developed some academic works regarding AI in sport business. They are planning to launch a start-up, Playerank, to exploit AI system for the creation of meaningful football players' matrixes. This solution can be employed for Scouting & Player Investment Evaluation and Optimization or in the Tactical Analysis.
The second case is dedicated to the Milan-based start-up Wallabies. It is an AI-system specifically dedicated to Scouting & Player Investment Evaluation and Optimization, since it aims to introduce the Pricing Market Methodology in the football players' transfer market.
The last one is dedicated to Math&Sport start-up, also based in Milan. It developed an AI-powered system to enhance and bring to a completely new level the tactical analysis in the football field.

PART III

3.1 RESEARCH METHOD

3.1.1 OBJECTIVES OF THE RESEARCH

The objective of this work is analyzing the current state-of-art of artificial intelligence applications in the sport business, investigating on what is their potential value for sport organizations.

More in details, the present works aims to achieve the followings goals:

1. Mapping the current artificial intelligence solutions on the market.
2. Provide an explanation about their working methodology, their value proposition and their business area.
3. Embody the perspective of a generic sport club and define if these solutions are valuable for the business purposes of the organization itself. In other words, answering to the following question: *"If I am a generic football club, playing in the Italian Serie A, Can Artificial Intelligence be helpful for me? and if yes, How?".* This is final objective of the work.

   The fact that in the question above it is mentioned "a football club playing the Serie A", does not imply that the work has been designed for the football field. Everything that has been presented until now, and that has been written in the conclusions, can be applied in every sport. it has been defined an objective like this one, since a relevant share of the applications is dedicated to football, and since, all the case studies deal with football. But this does affects the applicability of the conclusions in other sports.

3.1.2 METHODOLOGY

To achieve the goals explained before, the following logical framework has been adopted:

1. Model what is an artificial intelligence solution, in term of:
   - Existence conditions. What are the minimum conditions that an artificial intelligence solution need to meet to be properly consider as artificial intelligence.

- Scope. What is the high-level scope of an artificial intelligence solution in a generic organization.
- Operational characteristics. What are the operational processes that distinguish an artificial intelligence solution.

2. Present the main classes of artificial intelligence algorithms, in term of historical evolution and application fields.
3. Present parallelly the current sport environment, highlighting the complexity, the issues and the evolutions that sport industry is facing.
4. Mapping the artificial intelligence applications in sport industry, focusing on the business area, processes, procedures and activities that these solutions are targeting.
5. Develop some case studies to present more in detail some applications.
6. Answering to the questions presented into the 3.1.1 section.

3.1.3 DATA RESEARCH

PART I- Literary Review of AI and Sport Industry

For Part 1A, the one about Artificial Intelligence as generic topic, the approach used was the one of a classical Literature Review. The peculiarity was that the argument is unusual and, furthermore, the research on the topic of artificial intelligence, is often developed within companies or private research centers. So, most of the information are not public available. The second big issue is that artificial intelligence is a complex topic. It mainly deals with advanced computer science, mathematical engineering and statistics. Subjects that need a specific background be properly handled. At the same time, the thesis had a scope that was related to a qualitative analysis of the market sector. So, for this entire part, it was necessary to manage the trade-off between, developing the thesis aligned with its purposes, and, deal with artificial intelligence in a rigorous way. Usually this trade-off has been faced employing real-world examples of artificial intelligence applications, that have been used to present, from time to time, the characteristics, the features and the working methodologies of the algorithms presented.

Important literature sources that have been used in this part are:

- Google Scholar
- Sciencedirect
- ResearchGate

At the same type also several books, newspaper articles and video conferences have been exploited as sources of information.

For Part 1B, the sources were mainly related to the findings of the Osservatorio Innvazione Digitale nell'Industria dello Sport of Politecnico di Milano and commercial reports on the topic of modern sport industry, founded online.

PART II- The Mapping Phase of AI Solutions in Sport Industry

For Part II, where the artificial intelligence solutions have been mapped and presented, the methodology used was the creation of a big Excel database as starting point. In this file, all the companies that were offering artificial intelligence solutions, were listed.
In the final version of the database, around 100 cases in which AI has been employed in the sport business have been listed. Since the topic deals with advanced technological innovation, large majority of the organizations were start-ups.
The collection of start-ups' data relied mainly on online databases such as Crunchbase and CB Insights. Also, websites dedicated to innovation have been considered. Examples are Sport Techie, Medium and general sports newspapers.
Not all the start-ups seen have been inserted in the database. Even if constraints have not been strictly defined, usually the attention was mainly focused only on start-ups that 1. Have a size that can show reliability 2. Have a clear business model. These limits were imposed to select only the most prominent solutions. The other limits were related to the funding date, since only start-ups with at least one year of life have been generally considered.
For each start-up, have been reported the name, a brief description of the activity, the nationality, the business model description, an indication about costs, the technology over which it is based and the

sport target and information about the growth.

It happened that also big companies offered artificial intelligence-based services. In these cases, it was more difficult to disentangle data for the single service.

On the 100 cases founded, around the 80% of them are startups that are launching or plan to launch AI solution dedicated to sport industry. The exact number depends also on what is intended with startups. Some of them are traditional startups, like, for example, the ones to which the case studies have been dedicated. There are a small number of founders, they are raising money following the traditional founding-raising path and they are autonomous entities. There are also cases in which the organization that is developing the AI-based solution is a spin-off of a bigger company or it is a startup with already several years of activity, that can be considered a common business.

Providing some additional information, the 100 organizations that have employed AI in sport business come from 11 different countries (USA, Canada, Australia, Denmark, UK, Netherlands, Italy, Israel, France, Ireland and Spain). If we cluster them on macro-geographic areas, we obtain that the large majorities of cases are in North-America with 70 cases if we consider also the multinational like Microsoft, Intel and IBM that are operating in this sector. The EMEA region counts for around 25 cases. The remaining cases come from Oceania and South-America.

The last majorities of the case have been developed in the last 3 years.

Regarding the reliability of the startups, the main issue was related to the fact that additional information on the organization were usually not public or not free. So, time by time, it has been used as a proxy for the reliability of the startup the number of employees, the reliability of commercial partners, the press sources that dedicated space to this startup, and eventually, the size of the investments when the information were available.

Instead, in the best of my knowledge, there is no cases in which a sport organization have developed an artificial intelligence solution in-house.

The objective of this database was to identify patterns, relations and common aspects between all the analyzed solutions. Basing on these common characteristics, a sort of clustering process has been performed.
Indeed, all the solutions have been grouped in new categories that could represent the entire market. The underpinned assumption was to use the start-ups contained in the excel database as a meaningful sample of the entire market.
As usual when happen an aggregation, it is possible to lose the highly-detailed granularity to assess every single case, but, on the other side, it is possible to create categorizations, that can be effective in drawing a snapshot of the current state-of-art of the artificial intelligence solutions in sport industry.

The structure of figure 17, that it the starting point of the entire Part II, has been developed basing on the results obtained through the clustering process. The categories in which figure 17 is divided are the categories that better divide all the solutions found in the excel database.

In the following writing phase, only some cases, between all the cases present in the excel file, have been used as example for the entire category.
At the same time, the argumentation has been enriched with all the academic papers available which were presenting artificial intelligence systems belonging to one of the categories already mentioned. Unfortunately, it was not possible find academic works for each category. These results were anyway predictable since some category, like fan engagement or media creation do not belong to any faculty of the traditional academical system and, thus, it lacks a structured research system.

The category, for which was possible to find the highest number of academic paper was the tactical analysis, a subcategory of Activity and Performance. In this case, the argumentation has been entirely developed as a literature analysis of the numerous papers found. This was done for two reasons:

1. It seemed more appropriate present the topic in this way, showing the academic interest in this field.
2. The solutions that belong to these category, defined themselves generally as "*AI-powered sport*

analytics company"[166] or "*deep learning algorithms able to come up with powerful details and explanations that actually mean something to coaches"*[167]. The risk of presenting them without a consistent academic argumentation as backbone, was to not be able to fully express the potentialities of such solutions.

A further evidence that the work based on the analysis of the excel database was correct is that all the papers found fit perfectly in just one category. This could be seen as a qualitative validation of the categorization process.

3.1.4 CASE STUDIES METHODOLOGY

To deal with the case studies, the Methodology proposed by professor Michela Arnaboldi (Management Engineering Department, 2016) has been followed as guideline. The case studies have been organized as classical interviews with important personalities within the selected organizations. The questions were based on the work done until that moment. The idea was using the case studies to have access to information that were not available on the classical channels, like on-line and academics works.
What was impossible to discover was how the companies, that are offering artificial intelligence solutions, work in a more operational way. For these reasons, the case studies were necessary to complete the work. At the same time, it was useful for the purposes of the thesis to speak directly with people directly involved into the business, to compare their future perspective and expectations, with the ones that were emerging from the thesis itself. The opinions of the interviewed were used as a sort of validation of the resilience and of the correctness of this work.

The practical approach to case studies has been inspired by the work done until this moment. The objective was using some introductive questions to run through again all the logical steps developed until now. They are presented as follows:

[166] *http://sportlogiq.com/en/*
[167] *https://www.iceberg.hockey/technology*

1. Since it has been recognized that the main value generation is associated with "*Acting Rational"* systems, at the beginning, it was important to understand if the solutions about which the case study has been dedicated respects the "*Acting Rational"* definition proposed, to have a double-check of the applicability of the framework. It was done verifying that: 1. the solution would respect the Perception-Reasoning-Acting scheme (1A.5), 2. the degree of dynamicity of the environment in which it is supposed to work and 3. its ability to adapt to it.
2. Since the objective of artificial intelligence solutions is support human decision-makers, the second point was evaluating the ability of the solution to help decision-makers to take more informed decision. This activity has been usually done comparing how the task is performed now in sport organizations, meaning without the support of the artificial intelligence solution, and in case in which the human operator could benefit the insights generated by the artificial intelligence system.
3. Defining to which algorithms class the solution belongs. For example, if the solution is based on supervised learning or unsupervised, or, if it is possible to consider the solution deep learning or not. The idea behind was classifying the solution in function of artificial intelligence technology.
4. At this point, the solution has been addressed to one of the categories of figure 17.
5. The last logical step was evaluating how much the solution of the case study differs respect to the solutions found until now and define the reasons why.

The questions have changed from one case to another, also to respect the willingness of the interviewed about dealing with some topics. All the organizations are start-ups that have constraints about the confidentiality of some information. So, in some cases was necessary to not deal with specific topics for request of the interviewer.

Once transcribed entirely the text, the analysis process started. The objective was identified theories that could be in common in all the case studies to use them as starting point for the writing of the conclusions.

Even if the questions were prepared before the interview, in some cases, thanks to the answers, new and unplanned questions raised. This brought the case studies in new topics areas each time.

With a personal choice, it has been decided to dedicate the case studies to Italian entities. The idea was picking-up one case study for each of the macro-areas of figure 17. So, starting from the excel data base, all the Italian companies have been selected and contacted.

3.2 CASE STUDIES

In this section will be presented the three case studies employed to deepen the operation knowledge of artificial intelligence systems. The case studies are will be presented with the name of the interviewed and the name of the company. They are:

1. Luca Pappalardo, Paolo Cintia and the Playerank case.
2. Luigi Libroia, Pietro Tartella and the Wallabies case.
3. Ottavio Crivaro and the Math&Sport case.

3.2.1 LUCA PAPPALARDO, PAOLO CINTIA AND THE PLAYERANK CASE

NAME OF THE START-UP	PLAYERANK
FOUNDATION DATE	2018
FOUNDATION PLACE	Pisa, Italy
VALUE PROPOSITION	Football player ranking system
BUSINESS MODEL	On-line Platform
“ACTING RATIONAL APPLICATION”	Yes
FOLLOW THE “PERCEPTION-REASONING-ACTING” SCHEME	Yes
ENHANCE DECISION-MAKING PROCESS EFFECTIVENESS	Yes
MACRO-AREA OF APPLICATION (FIGURE 17)	• Management & Organization • Activity & Performance
CATEGORIES	• Team-based activity as scouting & players’ investment evaluation and optimization • Tactical performance analysis
INTERVIEWED	• Luca Pappalardo, Co-Founder, Data Scientist at CNR Pisa • Paolo Cintia, Co-Founder, Data Scientista at CNR Pisa

Introduction

"Artificial Intelligence should explain us what we are not able to comprehend" Paolo Cintia

Luca Pappalardo is a researcher in Data Science at the ISTI Institute of CNR in Pisa. He is also a member of the research European infrastructure SoBigData.
Paolo Cintia is a researcher in Data Science at CNR of Pisa. He belongs to the research Lab of KDD.

They have developed some academic works in the field of AI in sport business, and, they just founded a start-up, Playerank s.r.l, whose aim is leveraging AI to objectively evaluate players. It is still in a very early stage of development since it has been officially founded in June 2018.

Their initial research field was the study of human mobility. After a while, they found that between mobility studies and sport there are some common points, mainly because in both the topics, it is necessary to analyse and model how the agents move. In fact, their first step in the sport environment has been in cycling. Regarding cycling, they tried to re-adapt the knowledge developed earlier, to create predictive systems of the athletes' performance. Basing on metabolisms and energy's consumption data, they tried to develop a system to estimate the possible future states of the race. To make an example, the system should give insights about the ability of the cycler to end the race with the same current rhythm, or, if he, physiologically, will be obliged to decrease the pace. Obviously, this information can be fundamental to plan in real-time future strategies in cycling.

Academic Works

Since they are both researchers, the starting point of this case study will be the analysis of their academic papers.

AI FOR REPLICATING HUMAN RANKING

In "*Human Perception of Performance*" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi have developed a Machine Learning system able to rank football players performances, assigning them a grade between 0 and 10.

The peculiarity of this work was not developing a system that could rank players' performance in the best, objective and unbiased way, considering all the variables available. Instead "*the system was created to imitate and replicate the human mechanism of performances evaluation*"[168]. The human concerning with this work are the Italian football journalists that after the matches, judge and ranks all the players' performances.

"*The journalist is asked to summarize 90 minutes of match, with thousands of events per players, both with and without ball in a unique final grade. The mechanism used by journalists to evaluate players is basically unknown since it is highly subjective, personalized and informal*"[169].

"Formally the authors assume the existence of a function F representing the rating process. They use a machine learning system to infer a function f, which approximates the unknown function F, representing the human process"[170].

So, the idea behind was developing an algorithm "*that evaluates players in the same way in which a human would have done*"[171]. Thanks to this definition, it is possible to place this system into the *"Thinking Humanly"* category, presented at the beginning of Part IA. Just as reminder, Thinking Humanly systems aim to imitate the cognitive abilities of humans, replicating them artificially, with the scope of understanding how we think.

The objective of this work was basically "*demonstrating how biased is the human judgment and [how it is] based only on noticeable events"[172].*

This is done to understand the "*criteria behind human performance judgement"[173]* and quantify the level of objectivity, and, the presence of bias in our judgement process, with the final purpose to teach

[168] *"Human Perception of Performance" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017*

[169] *"Human Perception of Performance" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017*

[170] *"Human Perception of Performance" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017*

[171] *"Human Perception of Performance" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017*

[172] *"Human Perception of Performance" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017*

[173] *"Human Perception of Performance" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017*

us how to improve it.

The system works on three logical steps: observe, predict, explain. "*Observe means collect detailed data that allow to measure objectively all the different aspects of the performance that the human judge is expected to evaluate. Predicting means that leveraging these observations, the system can construct a predictive model using machine learning techniques. These models should perform the tasks that humans do: assign to any performance an evaluation. The scope is accurately predicting what would be the human judge evaluation for a given performance, providing an artificial intelligence proxy of human expertise. The last step is explaining what the machine learning system obtained discovering and understanding the features that most influence the evaluation outcome and highlighting the rules adopted by the model to score a performance, thus reproducing the logic of the human evaluator".*

The observation phase was performed collecting the ratings, assigned to every player of the Italian Serie A, by the three most important Italian sports newspaper (La Gazzetta dello Sport, Il Corriere dello Sport e Tuttosport). In parallel the authors also collected a great number of information about a single players performance describing any quantifiable aspects of the match.

They use these two types of information to train a machine learning system, that could learn the relationship between technical performance and final rating, trying to approximate the human evaluation process. They found out that the technical features, like quality of the passes, numbers of goals etc.., alone cannot fully explained the human evaluation process. In other words, they discovered that "*the sport journalists do not base their evaluations on the objective evaluations of the players' performance"*[174]*.*

In fact, just after the *"inclusion of contextual information, not directly related to the performance of the athlete, the artificial intelligence system produced better results in term of alignments with the humans' grade"*[175]*.*

Contextual information is information like nationality, age, the club of the player, the expected game

[174] *Human Perception of Performance*" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017

[175] *Human Perception of Performance*" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017

output as estimated by bookmakers, the actual game outcome and whether a game is played home or away. So, in other word something that should not be included into the performance ranking process. Only adding these additional information, the machine learning system was able to replicate accurately the journalists' ratings system.

The system was not able to forecast extreme grades like 10 or 3, that happened only along with rare events." *For example, a 10, has been assigned when Gonzalo Higuain become the best seasonal scorer ever in the Italian Serie A. This is a something absolutely unrelated to player's performance of that specific match and, thus, it was impossible to be forecasted by the machine learning system*"[176].

They also found that sport journalisms' ranks are based on different features in relation with the role of the players. So, the same feature has different importance for different roles.

"For example, while goalkeepers and forward are evaluated for technical features, midfielder and defenders are evaluated just by collective features like the team goals difference"[177].

This is probably due to the fact that goalkeepers and forwards are directly associated with events, such as goals, which attract the attention of the observers, like journalists or fans. In contrast, midfielders and defenders are involved in normal events and so, attracting less attention, are evaluated in a more passive way.

After having developed a system that is able to replicate accurately how Italian sport journalists rank players performances, the last step, is explaining the human behaviour. Analysing how the machine learning system works, they discovered that human evaluation follows a "*simplistic cognitive process, based on a noticeably heuristic: judges select a limited number of features which attract their attention and then, they rate a performance based on the presence of noticeable values, features values far from the norm that can be easily brought to mind"*[178].

A graphical representation of how it works is presented in the figure 28. It shows the rating process for goalkeepers and forwards. The features highlighted are the ones that are most considered in the rating process. For example, for forwards three features matter the most: the number of goals scored

---

[176] *Human Perception of Performance*" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017

[177] *Human Perception of Performance*" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017

[178] *Human Perception of Performance*" Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti ed Albert-Laszalo Barabasi, 2017

(performance-based indicator), the game goal difference (contextual) and the expected outcome as estimated by bookmakers before the match (contextual).

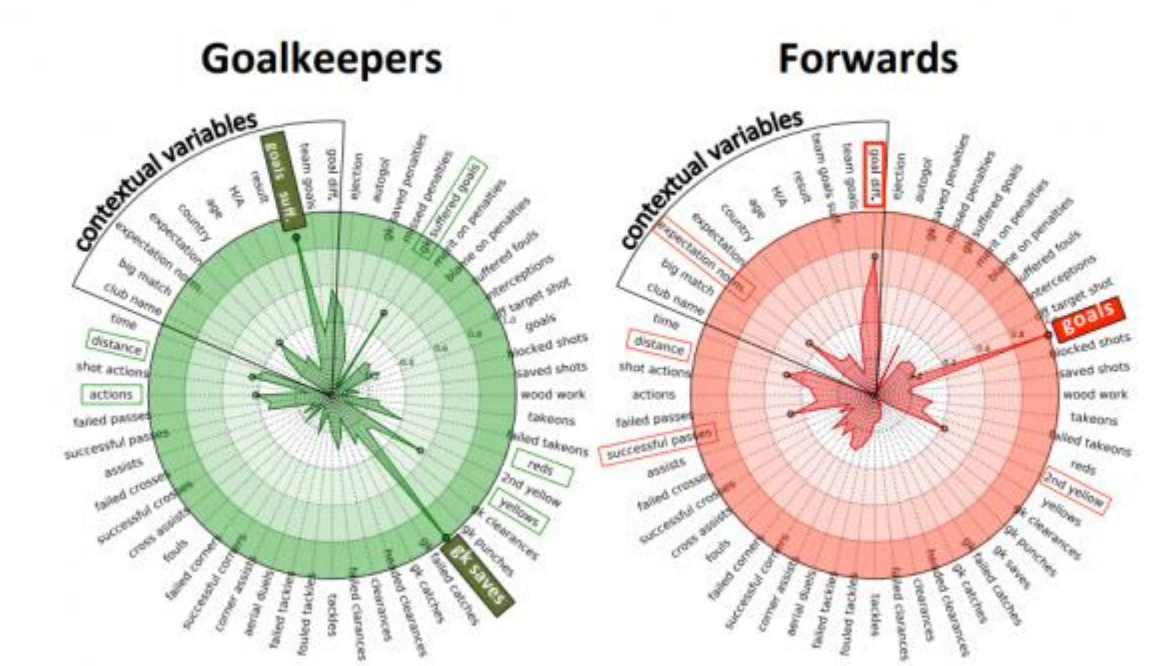


Figure 28: the radar chart represents the importance of every feature, normalized in a range of [0, 1], to human rating process for Goalkeepers and Froward. Features with an importance higher than 0.4 are highlighted and contextual features are grouped in the left upper corner of the two-radar chart. The plot indicates that, for example, for forward three features matter the most: the number for goals scored (performance), the game goal difference (contextual) and the expected outcome as estimated by bookmakers before the match(contextual). Source: Human Perception of Performance, Luca Pappalardo, Paolo Cintia, Dino Pedreschi, Fosca Giannotti, Albert-László Barabási

INJURIES MANAGEMENT

The second academic work deals with artificial intelligence in the field of injury management. As seen saw in the section 2.1.2, *"Injuries have a great impact on professional soccer, due to their large influence on team performance"* [179]. Furthermore, the impacts of injuries are not limited to the on-field

[179] *Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017*

performances, but they also affect the financial achievements of the club. "*the costs associated with the complex process of recovery and rehabilitation for the player is often considerable, both in terms of medical care and missed earnings deriving from the popularity of the player himself*"[180].

"Recent research demonstrates that injuries in Spain cause an average of 16% of season absence by players, corresponding to a total cost estimation of 188 million euros just in one season"[181].

For all these reasons, any tool that could predicts the injuries' occurrences, would be a valuable instrument in the hands of clubs, that could basically prevent injuries, operating corrective actions in case of high likelihood that a player could get injured.

As highlighted in the paper, the relationship between training workload and injury risk has been investigated by some studies. These studies demonstrated that players have an increasing risk of injuries when their workloads increase above certain thresholds[182]. Other researchers discovered that the highest probability to get injured is associated to weeks in which the workload is higher respect to the last month average, defining a clear causal link between fatigue and injuries[183].

In other academic works has been demonstrated that, if the workload is higher respect to the athletes' ability to recover, it is likely to enter in the so-called "overtraining syndrome" that is associated to high probabilities of injuries[184].

These are just few examples of the wider research field about sport injuries. All these studies demonstrated that there are causal links that explain why and when injuries happened. Their main limitation is that they have not predictive power. They explained why injuries occurred in the past, but,

---

[180] Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017

[181] Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017

[182] Gabbett TJ, Jenkins DG. Relationship between training load and injury in professional rugby league players. Journal of Science and Medicine in Sport. 2011;14: 204209.

[183] Hulin BT, Gabbett TJ, Blanch P, Chapman P, Bailey D, Orchard JV. Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. Br J Sports Med. 2014;48:708-712.

[184] Foster C. Monitoring training in athletes with reference to overtraining syndrome. Med Sci Sports Exerc. 1998;30:11641168.

they are not able to tell a coach if a player is likely to suffer an injury in the next training session, given his recent history.

This is exactly the purpose of the work presented in this paper. This work is one of the first ever attempt to create systems with a predictive ability for injuries in soccer. As it has been also highlighted in different parts of this thesis, such systems could not exist in the past for the limited availability of data about the players' performances. Just in recent years, thanks to the data revolution enabled by wearables and Internet of Things have been possible to start to think about developing predictive system based on historical data.

The work is about non-contact injuries, that are defined as *"any tissue damage sustained by a player that causes the absence in the next sport activities for at least the day after the day of the onset"*[185].

The system is based on automatic data collection through GPS technologies about the training workload of players. From these data, the authors extract a set of features used in sports science to describe kinematic, metabolic and mechanical aspects of training workload. The authors collected data from an Italian professional football club for an entire season.

In other words, they transformed GPS-based information in data about the three aspects quoted before.

The second step was integrating into the system the data about all the injuries that happened along the season and let the machine learning system find meaningful relationships between the training workload and the injuries' occurrences.

"The construction of such injury prediction models poses many challenges, on one hand, an injury prediction model must be highly accurate. Predictions which rarely forecast injuries are useless for coaches, as well as predictors which frequently produce false alarms… on the other hand, a 'black box' approach is less desirable for practical use since it is fundamental to understand the complex relationships between the performance of players and their injury risk through simple, interpretable and

[185] *Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017*

easy-to-use tools. An interpretable model reveals the influence of variables to injuries and the reason behind them, allowing to react in time by modifying properly training plans"[186]*.*

To resume the system to be valuable must be:

1. Highly precise to predict injuries
2. Easy to understand for a coach. Linking this part with what has been said in section 1A.6, an inverse Polanyi's Paradox is not desirable since the system is not just required to predict the injuries' risk, but also to make the reasons clear for coaches that can re-adapt the training plans to the new findings.

The system produced interesting results. Basing on just few variables that analysed the workload volatility and intensity, the system was able to detect 76% of the injuries, with a 94% of precision along an entire season[187]. (it has been tested in a professional Italian football team)

The false positive rate is very small, meaning that it is rare that the system predict that an injury will happen when it will not. This is important for professional soccer clubs *"because the scarcity of players can negatively affect the performance of a team, both in single match and during the season"*[188].

False negative rate is also moderate low, meaning that *"situations where a player that will get injured is classified as out of risk are infrequent"*[189]*.*

The authors investigate also the timing necessary to implement the solution in a real-world case. They assumed that, at the beginning of the season a club starts to record training workload and, at the same time, employ their model. At the beginning the machine learning system need a training phase to get the parameters of the variables involved. The authors observed that after 16 weeks of data collection the performances of the system stabilize within an acceptable range of effectiveness, reaching a precision of 0.8.

---

[186] *Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017*

[187] *Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017*

[188] *Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017*

[189] *Effective injury prediction in professional soccer with GPS data and machine learning, Alessio Rossi, Luca Pappalardo, Paolo Cintia, F. Marcello Iaia, Javier Fernandez, Daniel Medina, 2017*

PLAYERANK S.R.L

The idea to found Playerank comes from the work on players' evaluation. Since it has been demonstrated that humans lack the ability to judge players in an objective and omni-comprehensive way the players' performances, Luca Pappalardo and Paolo Cintia saw the commercial opportunity to develop a machine learning-based system, that can support human operator in judging football payers correctly and without biases.

The correct and fair evaluation of players can benefit two main players in a football team:

1. Coaches
2. Managers that deal with football players trading.

The system is based on the spatiotemporal data of the matches as highlighted in figure 29 and 30. The system receives in input data and statistics about every event that happen during a football match, for each player. The Working methodology is based on two different steps, that are showed in figures.

The first one is analysing a big sample of football matches and figuring out which are the behaviours of the players that most affect positively the capacity of sign a goal for a team. Once defined in an unbiased way the cause-effect correlation between players behaviours and results, the following logical step is assessing the single players' stats for each of the events previous mentioned during a certain time frame that could be a game or an entire season, and figure out, what has been their contribution to win the games.

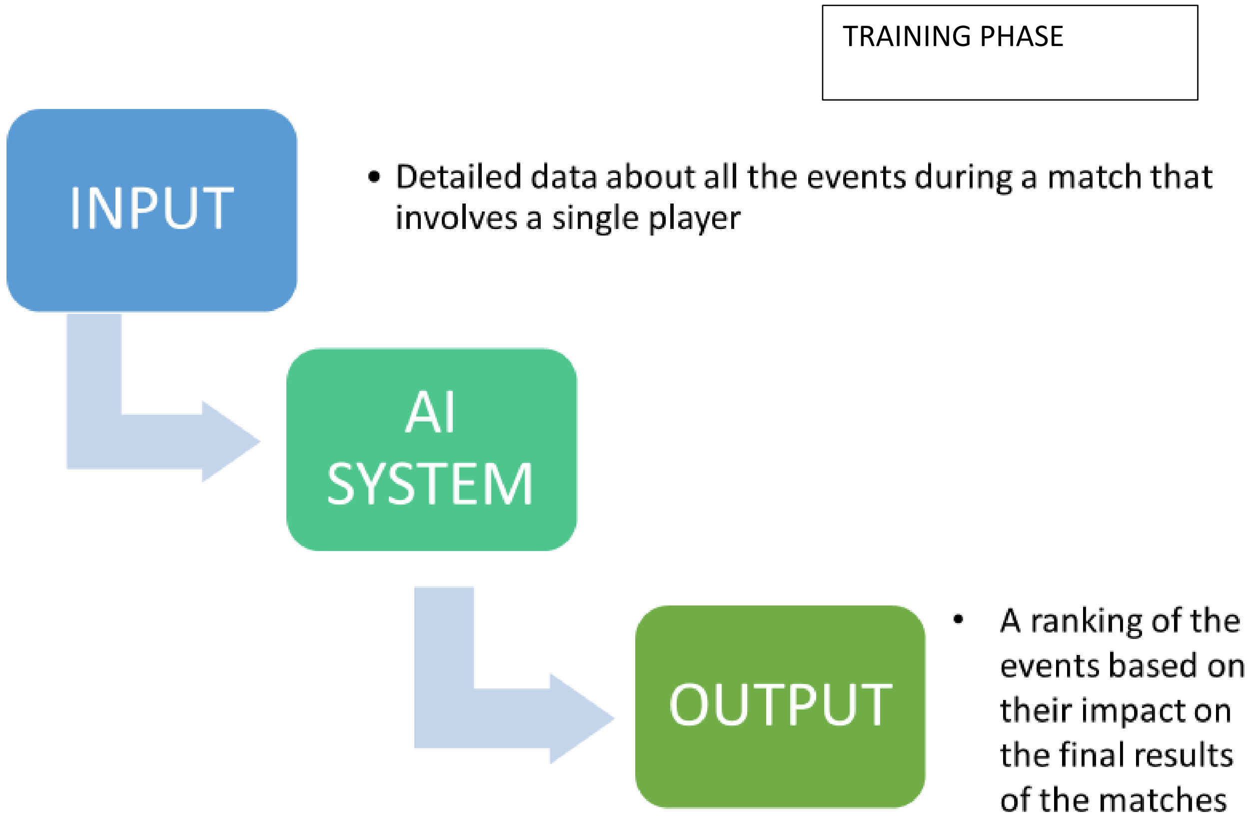


Figure 29: the training phase of Playerank

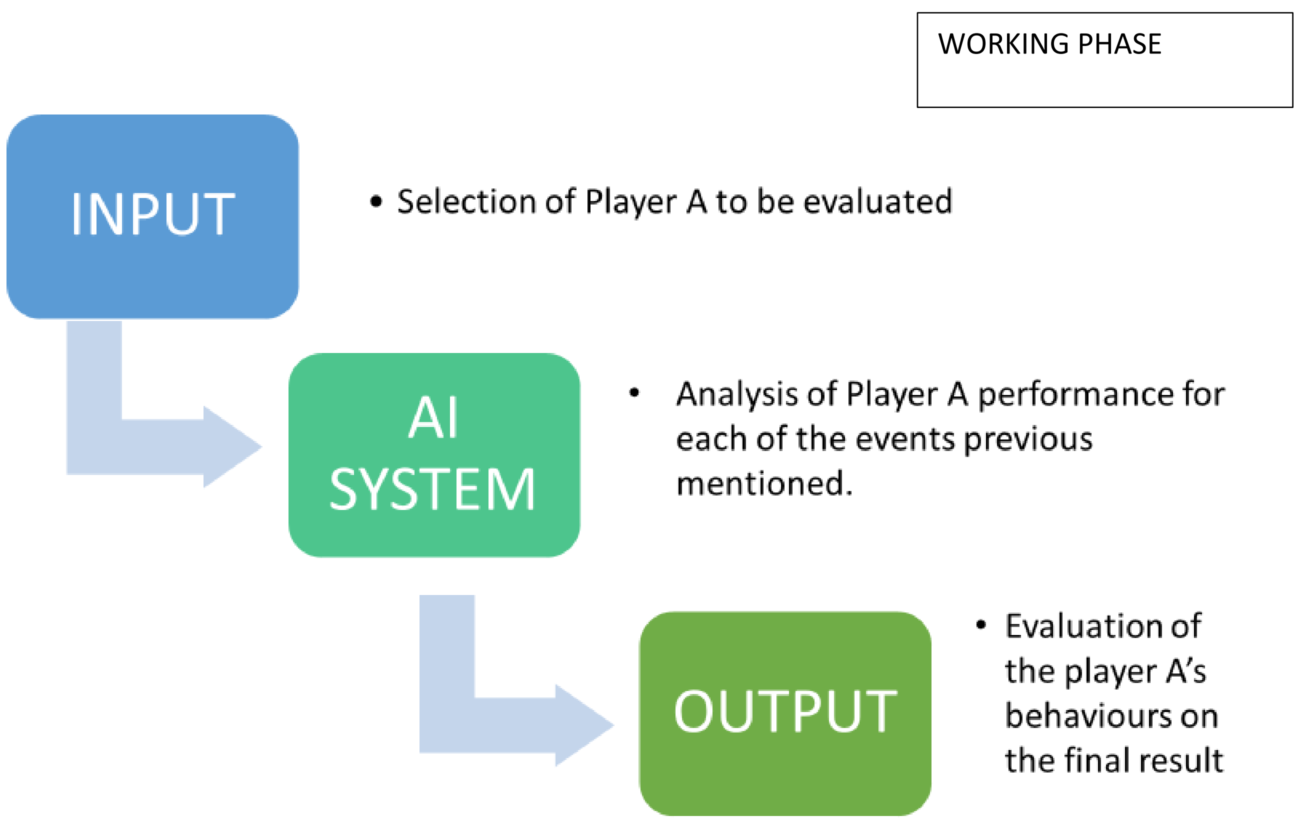


Figure 30: The working methodology of Playerank.

The main goal of Playerank is creating a system that can evaluate a player, 1. regardless how the match ends (it has been demonstrated in the paper "*Human Performance Perception*" that result is one the most important driver used by human judges to evaluate players performances), and 2. regardless to the quality of the other players in the same or in the opponent team. (the idea is that if a player plays a high-quality team his performances will be overestimated probably).
The features of this process must be: 1. **Objectivity** 2. **Data-Driven** 3. **Not related with downstream events**, in other words, unaffected by contextual information and 4. **Goal oriented**, meaning that the evaluation must be linked to the ability of the player to contribute to the winning possibilities of the matches.
To meet these objectives the judgement process must be completely human-free, because, as demonstrated in "Human Perception of Performance" it is highly biased and, thus, not reliable.
To do it, the founders have created a machine learning system that is able to analyse a big database of football matches data and, define, autonomously how each event (like ability to cross, ability to shoot, off-ball moves successful etc) impact the results. The system, in the mind of the founders, is basically able to rank the different characteristic of each players in relation to their impact to the team performance.

Performing effectively this task, permit further developments like:

1. Tactical Analysis. In fact, if the system is able to judge objectively each player in function of the effectiveness to affect positively the team performances, it is possible to aggregate for an entire team and, understand how the "Dangerousness of the Game" change during a match. This can make the team understand which the common features are (like collective moves, personal technical gestures etc.) that are associated to events that correspond to a high probability to sign a goal.

   The same analysis can be performed in a passive way, looking at the events that are linked to the highest dangerousness of the opponents, to understand and improve the weaknesses of the team.

   If the analysis is performed on opponents, it can permit to teams to prepare better the future

matches.
2. It can be used for scouting purposes and to better evaluate football players related investments. The idea is to create a sort of research engine for players in the long-term.

A characteristic of playerank is the fact that the variables taken into consideration for evaluate players are dynamic and not equal for everyone.
Basing on the average position of the player on the field, the system is able to understand the abilities of the players that most impact on the final performance of the team for players that play in the field zone.
This permits also to study the versatility of players. In fact, the basic idea is that if the system knows what the most valuable characteristics are to play in a specific role, it is possible to evaluate if a player that plays in a different role could have the ability to also in another one.

Even if the project is still in a very early phase, the founders think that the best strategy is to remain totally independent and offer support to clubs.

3.2.2 LUIGI LIBROIA, PIETRO TARTELLA AND THE WALLABIES CASE

NAME OF THE START-UP	WALLABIES
FOUNDATION DATE	2016
FOUNDATION PLACE	Milano, Italy
VALUE PROPOSITION	Football player ranking system with Market Methodology
“ACTING RATIONAL APPLICATION”	Yes
FOLLOW THE “PERCEPTION-REASONING-ACTING” SCHEME	Yes
ENHANCE DECISION-MAKING PROCESS EFFECTIVENESS	Yes
BUSINESS MODEL	On-line Platform
MACRO-AREA OF APPLICATION (FIGURE 17)	• Management & Organization
CATEGORIES	• Team-based activity as scouting & players’ investment evaluation and optimization

INTERVIEWED	• Luigi Libroia, Co-Founder and Business consultant • Pietro Tartella, Data scientist at Wallabies

"Artificial Intelligence can overcome humans' failures, because it permits us to reach a limitless learning ability and to exploit an unrestricted data-storage capacity" Luigi Libroia, Co-Founder of Wallabies

Luigi Libroia is one of the co-founders of the start-up Wallabies. He is a business consultant and accountant.
Pietro Tartella is a data science researcher, that joined the company some months ago in supporting the development of the algorithms.
They are the ones that have been interviewed during the case study.

Introduction

Wallabies is a Milan-based start-up, founded in 2016. It has been created to solve the problems related to football players' evaluation. Mr. Libroia is a business accountant, and, he was asked to evaluate the market value of players, to elaborate the balance sheets of football clubs. The players represent football clubs' biggest asset category, and so, it was important for him, to evaluate them properly. In facing this issue, he developed the idea to apply the Market Methodology in evaluating football players' process. The market methodology defines the price of an asset, looking at which price similar assets have been exchanged on the market. This methodology is widely used in other industries, for example, to fix the price of consumer products, or in the financial world, to assign a first price to the shares of a company, just listed on the stock exchange.
Wallabies was the first ever organization that tried to apply this concept in football. As it has been presented in section 2.2.1, the football players market is highly inefficient and, thus, it might happen that the market prices are not aligned with the fair value of the players.

The company has been created to answer at a similar question: *"Given this specific player, what is his fair value on the market?".*

To answer this question more effectively respect to how it is currently done, the process should have some characteristics:

1. it should be **Objective**. Meaning that it should be totally based on evidences, leaving out from the computation, every type of guess about future performances or personal judgements 2. **Data-driven**. To meet the objective at point 1, the unique approach possible is be based on data. These data should be related just to the performance of players. The objective is leaving out all the “Noise”. In other words, focus the attention just on the specific players and not considering all the contextual information that could affect the evaluation. 3. **Unbiased**, meaning that the solution should overcome the limits of humans bounded rationality and deliver a solution that could be the closest possible to the optimum one. 4. **Speed**. The system should give results in short time.

Working Methodology

The starting point is developing a Database that can contain complete information. The company have access to data of an external provider that are highly detailed in reporting every possible event that happens during a football match, with probably the highest level of granularity reachable.

The company has access to an immense amount of data, considering that now, it can use the information related to all the matches of 11 leagues (9 from Europe and 2 from south America) for several seasons.

From this data, 580 variables have been extracted as showed in figure 31. These variables can describe in a holistic and omni-comprehensive way the performance of the single player during that match. These variables are, just to make some examples, the number of goals, the number of passes, the number of assists, the number of lost balls and so on.

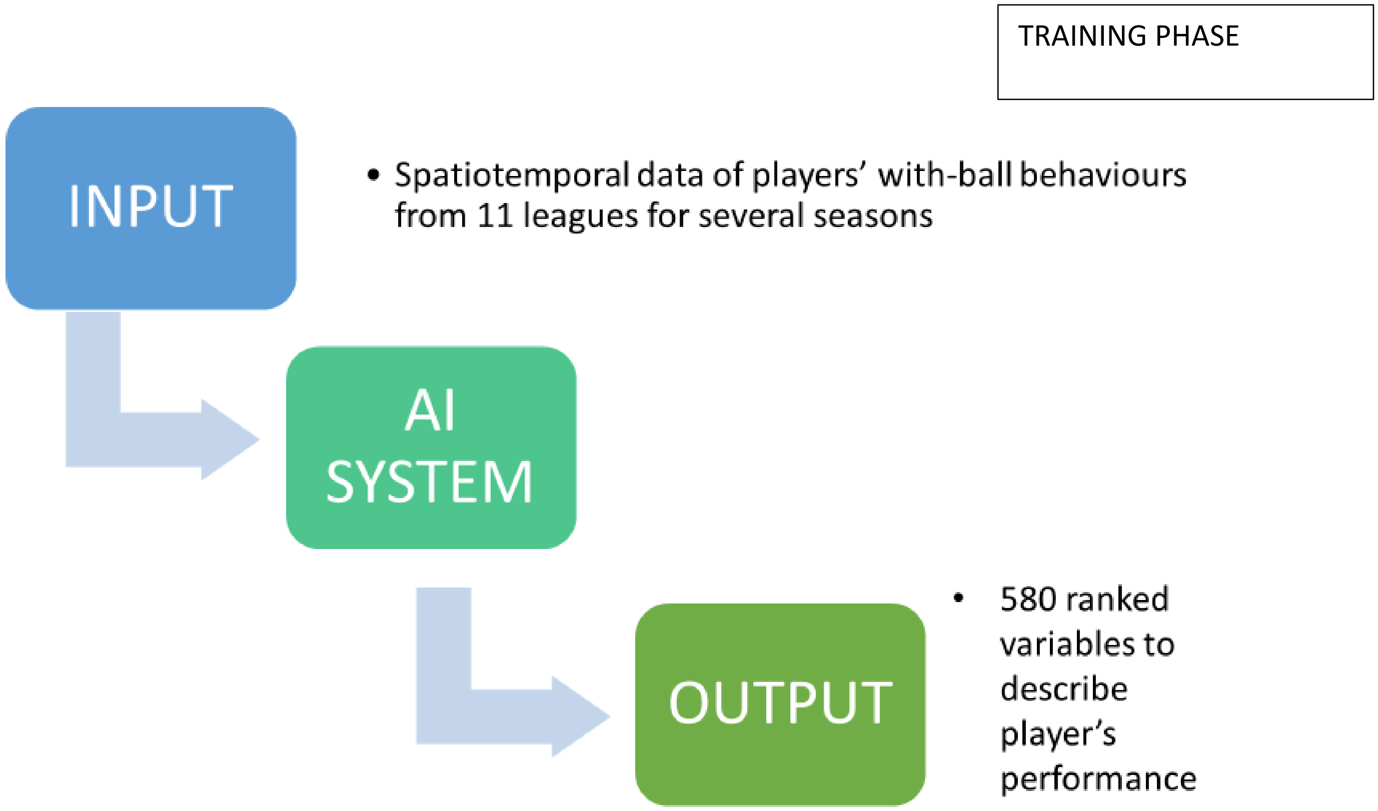


Figure 31: the training phase of Wallabies

To answer to the question expressed before using a market methodology, it is necessary to find comparable, or in other words, players that are really like the one that is necessary to evaluate. To do it two logical steps are necessary:

1. Evaluate in absolute term how an athlete plays
2. Find the most similar player and eventually look his market evaluation

The 580 variables are selected to perform these two tasks. Obviously, all these variables cannot concur in the final evaluation with the same magnitude. Some variables are more important than other in defining the performance, and the consequent value of a player.

To do it in mathematical term it is necessary to assign to every variable a weight. The problem is that these weights cannot be assigned by human experts otherwise the system would have embodied their biases. To be totally objective in the evolution, no human interact is allowed. With this logical step, Wallabies make the final result objective and un-biased.

The system should be able to derive autonomously these weights, basing on the analysis of the immense database mentioned before. Doing this, the company employed a data-driven approach. In fact, the weights are derived from the analysis of all the events matches and their impacts on the final results of the players. There no assumption about it. Looking at thousands of matches, the system recognized autonomously that one feature is more important than another in defining the performance of the player.

The last step in defining the market value, is finding two or more players that can be as much as similar and evaluating the market transactions that happened involving them, as showed in figure 32 and 33. Figure 33 represents the case in which the system has found 4 comparable players for player A around the globe. To each of them is associated a Comparability Rate and each of them has been traded at a given value. Toting up all this data the system can figure out accurately the value of the Player A.

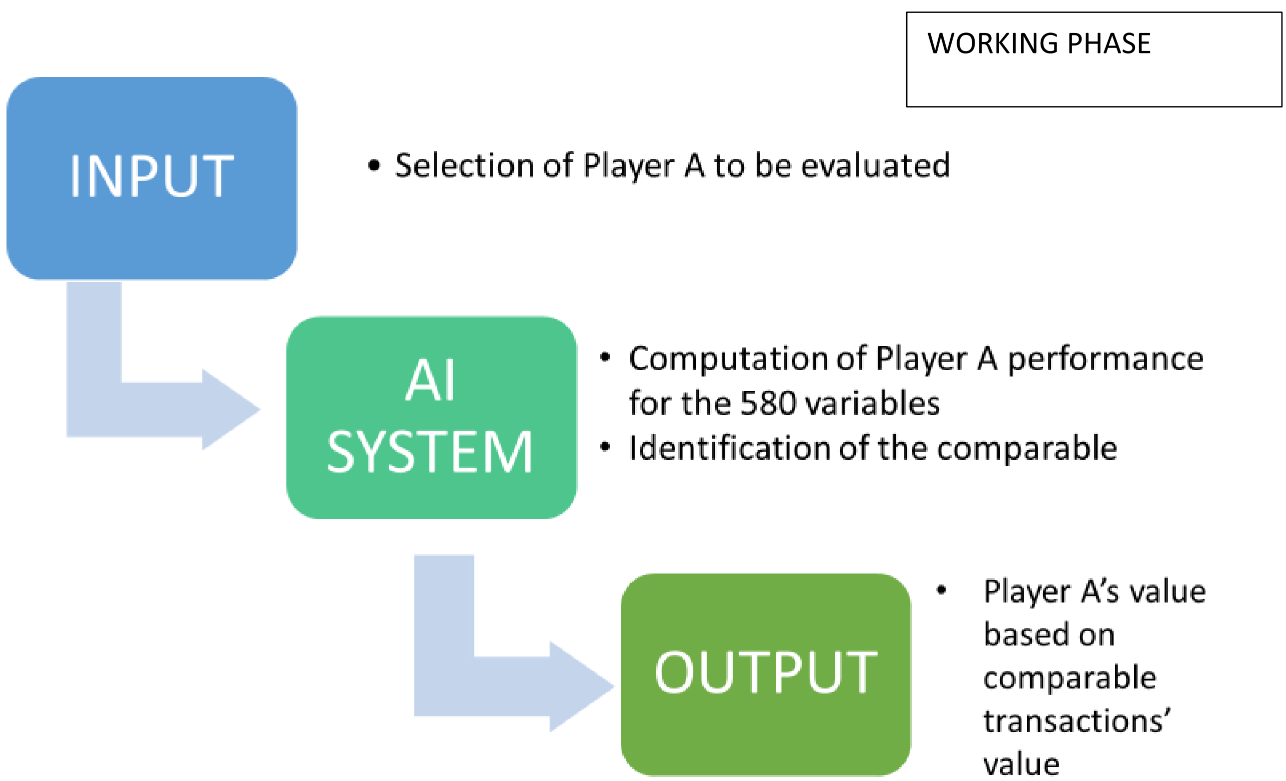


Figure 32: the working phase of wallabies- Source Wallabies Company

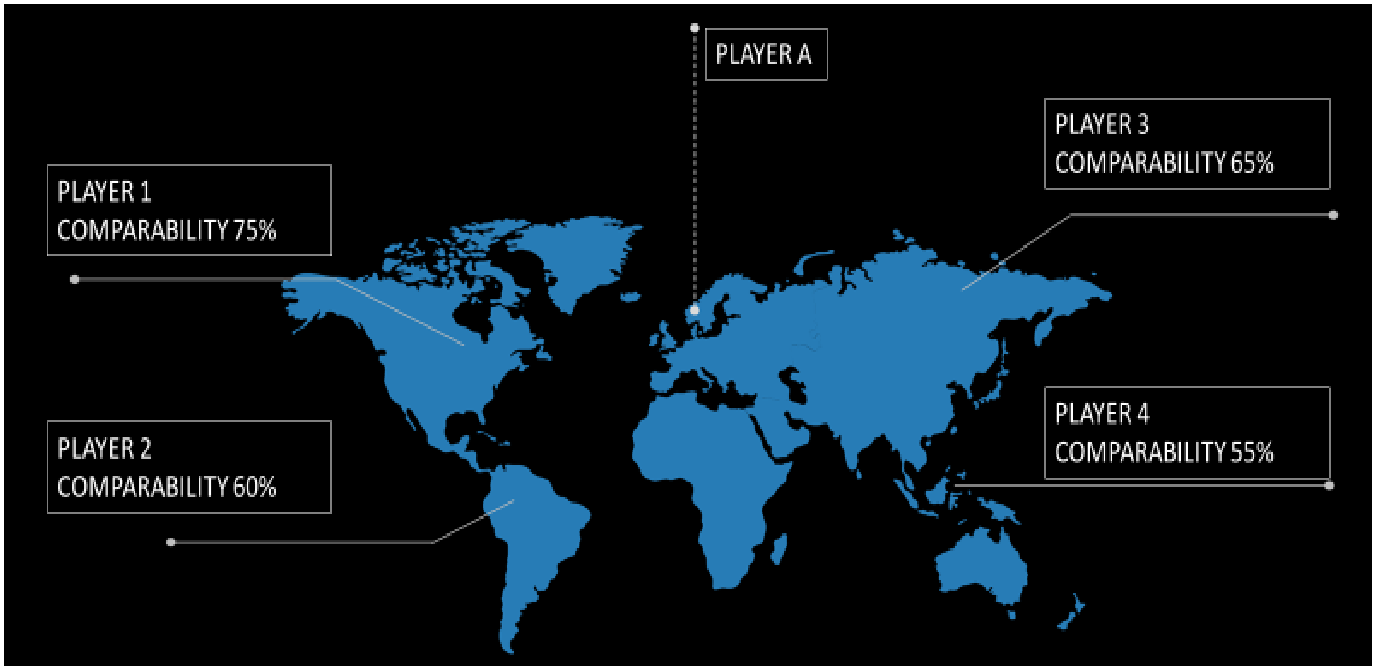


Figure 33: Wallabies working Phase- Source Wallabies

A similar system, if well-functioning opens infinite opportunities in the world of football. Some of them are listed:

1. It enables to compare players considering all the relevant aspects, leveraging on a system that has the 4 already mentioned features. This has great potential in the scouting activities, when clubs have to compare and decide which players to sign. As highlighted by Wallabies' staff, human scouts can effectively compare just two/ three players per time. These tasks require a lot of time, since dozens of matches for each player must be seen to take informed decision and, so, a great mnemonical capacity is required. For this reason, it could easily happen that they will ignore some factors. The Wallabies's system could avoid these problems, comparing effectively 80 000 players.
2. It can compare two players in different time frame. The idea is that wallabies have the records of the last years of matches so potentially it is possible to compare the player performance in the past with players performances now.
3. If the records of the performance are available, it is possible to make a step further in analysing why they changed. This is an application that is dedicated to training and improvements of

players. The idea is search for the reasons for which the performance, or even the single aspect of the overall performances, changed and understand the reasons behind. The potentiality is that it makes possible to find out that a player tends to perform better along with specific external characteristics.

4. Knowing all the nuances in which the players' performance is declined, it is possible to make further data-driven analysis. So, it is possible to group or cluster some of these features, that can explain a generic macro-ability of a players. For example, the system can recognize that the ability of a left back defender to be effective in attack, is related to his performances in a specific cluster of the 580 initial features. So, this enable also punctual evaluations of specific behaviours of players.
5. Probably the long-term objective is creating e a platform with simulation abilities. Teams can use it to try to forecast the outcomes of specific choices about players, and their impact on the team performance.

A beta version of the Wallabies algorithms has been finished in February 2017. it has already showed a deep understanding of the game. This ability will be presented through a brief anecdote. During the training phase, the system was working in defining the weights of all the 580 features. In other words, it was analysing thousands of matches to quantify how much each of the features was relevant in defining the performance of a player. The particularity was that it assigned a null weight to the "assist" category. This means that assist as statistics, according to the system, is useless in evaluating the performance of a player. In other words, there is no sense to evaluate a player basing on the number of assist. This outcome seemed odd, even the Wallabies' team. Then they figured out that it was not the results of a bug, but instead the reason was that an assist is "a *play in which a players touches the ball and this leads to another player on the same team scoring a goal or a point*"[190].

So, basically the probability that an assist occurred, is conditioned by the second player's ability to sign the goal. So, theoretically, it is not a measurement of the performance of the first players, but rather a measurement of the scoring ability of the second one.

Everybody still considers the number of assist a meaningful statistic to evaluate a player without

[190] *https://dictionary.cambridge.org/dictionary/english/assist*

thinking about it. The Wallabies' system was able to understand it autonomously and "teach" it to the developers. This is just an episode, but it is meaningful of the potential of AI and of how it can be exploited to help human operators to understand better what is happing around and also to improve their abilities, having a system that can systematically show when and how they did wrong.

3.2.3 OTTAVIO CRIVARO AND THE MATH&SPORT CASE

NAME OF THE START-UP	MATH&SPORT
APPLICATION NAME	Smart Data Platform
APPLICATION LAUNCH	2018
FOUNDATION PLACE	Milano, Italy
VALUE PROPOSITION	Football tactical optimization system
BUSINESS MODEL	On-line Platform
"ACTING RATIONAL APPLICATION"	Yes
FOLLOW THE "PERCEPTION-REASONING-ACTING" SCHEME	Yes
ENHANCE DECISION-MAKING PROCESS EFFECTIVENESS	Yes
MACRO-AREA OF APPLICATION (FIGURE 17)	• Activity & Performance
CATEGORIES	• Tactical performance analysis
INTERVIEWED	• Ottavio Crivaro, Founder and CEO of Math&Sport

"*Artificial Intelligence is not required to discover new things, but to make us clear and evident what we cannot notice*". *Ottavio Crivaro is the CEO and founder of Moxoff and Math&Sport.*

Math&Sport is a Milan-based start-up that is developing AI-based systems, for tactical analysis of different sports. This case will deal with their new product, SDP (Smart Data Platform), an application dedicated to football.

Before describing the features of the product, it is useful to present their conceptual model about artificial intelligence. It will be presented in relation to football, but it can be easily generalized to every other field or industry.

Considering a generic natural event, that in this case is a football match, the understanding of that event can be reached only passing through two logical steps:

1. **Description of the event**. The first step is describing the event in a formal language, and thus, employing mainly mathematics and statistic. They are the unique method in which it is possible to quantify the variations during the evolution of an event.
   So, dealing with football, the idea behind is describing a football match, creating statistics and indicators as the number of passes, the number of goals, the number of tackles etc. that can represent whatever happen during the 90 minutes. The objective is creating a record of all the events that occur in that match.
2. **Interpretation of the data**. Once having described the match using statistics, the second step is searching for the causal links that make the events to happen, with that specific occurrences. So, making an example about football, the first step, the descriptive one, will give that a certain player has the 75% of success in the passes in the last match, the objective of this second step is making emerge the reasons behind the numbers, and, explain why they happened. The idea is to apply the Root-Cause analysis in football. It is a backword analysis, that starting from the outcome, goes back to find the causes that defined that outcome.

This approach to the analysis of a problem is the unique available when it is necessary to deal with complex problems, in which there are numerous variables.

SMART DATA PLATFORM

As introduced before, SDP is the AI-powered platform of Math&Sport dedicated to football. It can perform autonomously both the logical steps presented before.

Figure 33 and 34 represents graphically the functions of the system.

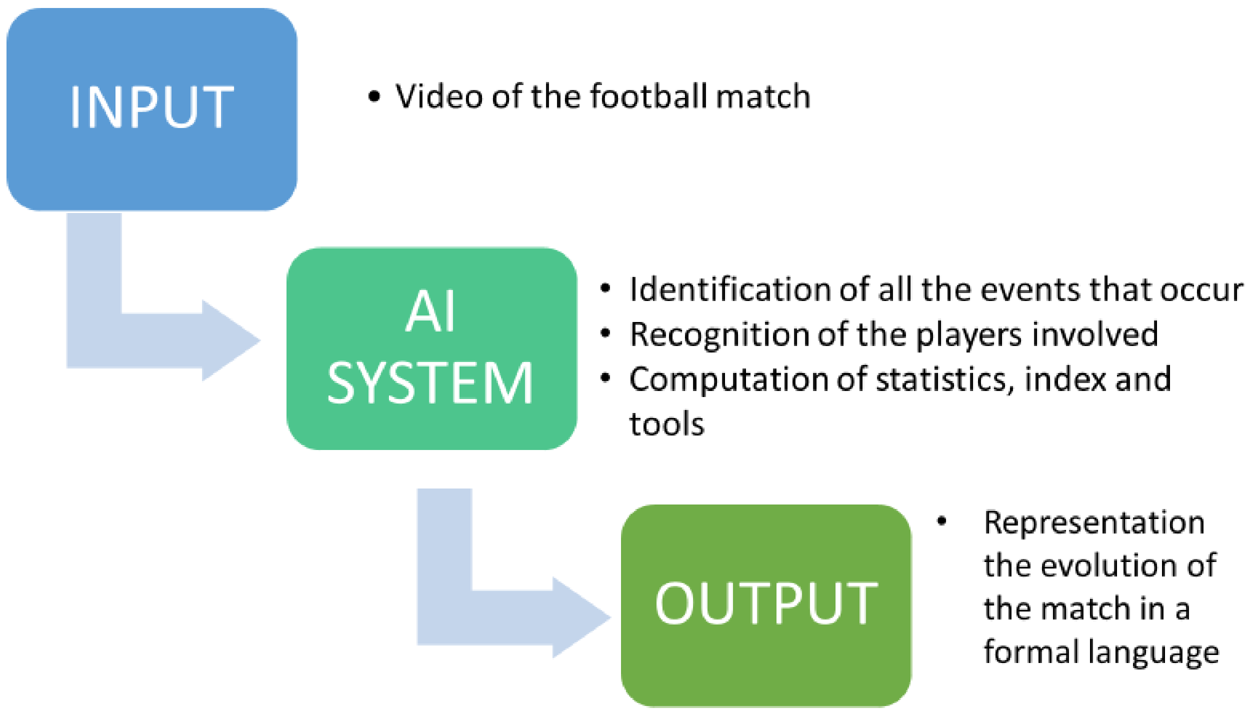


Figure 34: Phase 1: Description of the event. Source: Math&Sport documents

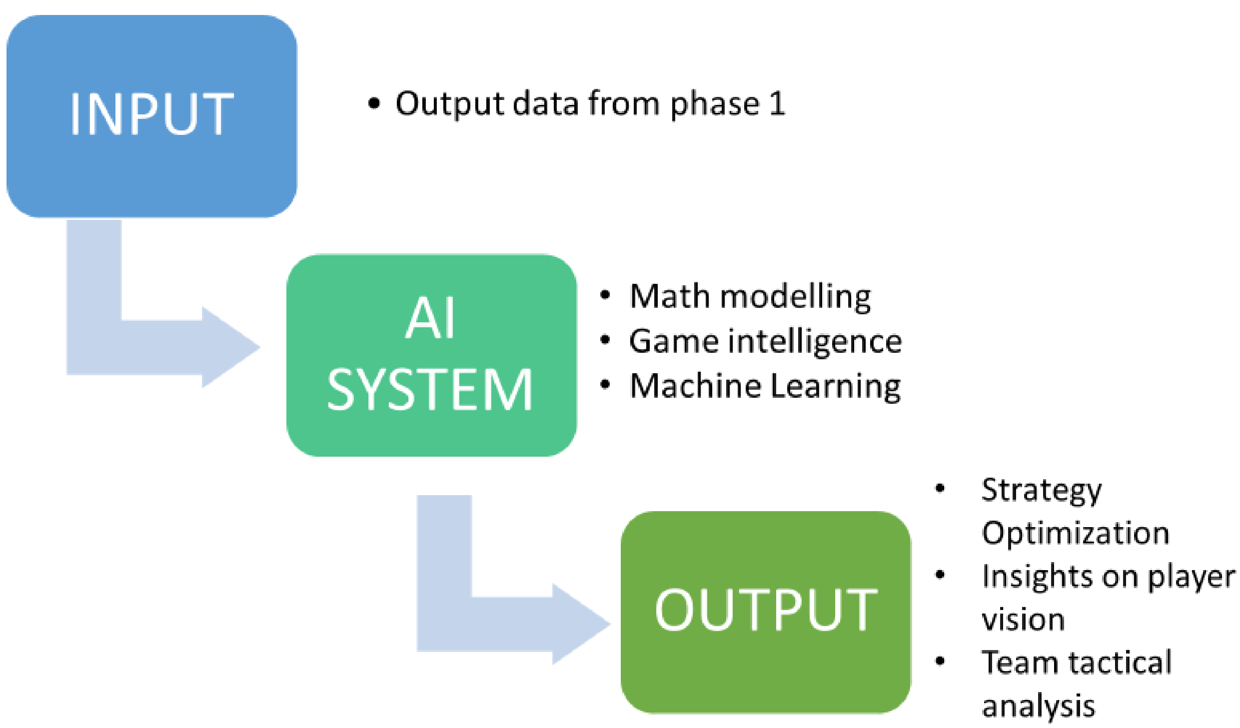


Figure 35: Phase 2: Interpretation of the data. Source: Math&Sport private documents

For these reasons, it represents a great source of value for teams, that can perform the two tasks immediately.

Regarding the activity of describing the matches employing statistics, it is a task that currently is already performed in all professional football clubs. They employ human operators, that looking at the video of the matches, fill in a big table all the events' occurrences.
Math&Sport employed artificial intelligence systems to automatize this activity. SDP leverages on a system that is based both on the video and the players' spatiotemporal data analysis. Form these sources, it can automatically identify, classify and count all the events that happen in a football match. The events, that are the game situation, that the system can recognize are thousands. They can be personal technical gestures, (tackles, passes, shoots…) but also more complex and collective dynamics like of-sides, high-pressing, low-pressing, forward actions, backward actions…).
This can make evident the completeness of analysis that the system can achieve.
The system looks at the video of the matches and identify in every single moment what is happing, giving it a 'name'. the name expresses a concept that is comprehensible for humans, it is not a sort of abstraction.
To perform this task, it was necessary to translate in mathematical terms all the possible situation that can happen in a football match.
This process has been done with the support of Adriano Bacconi. He is a football coach and he helped Math&Sport, in defining all the possible features and events can be detected in a football match, trying to code this knowledge.
This have permitted to define a mathematical model, that the system can employ to recognize and classify all the football events.

From this definition, it is possible to include such system in the "Acting Humanly" category presented in 1A.1. Just as reminder, acting humanly systems are systems that are able to perform a task that also humans can potentially do. But the system can perform them systematically faster and with less errors respect to a human operator.

The system is now capable to perform autonomously the point 1 of the process presented before.

To perform the task required for the second step, it is necessary to make the system able to search for the causes that generated that events. In other words, the system is asked to interpret the results, expressed in terms of statistics.

Math&Sport decided to perform this task, computing an indicator that could explain what is happing in an objective and omni-comprehensive way on the football field. Basing on the variations of the value of this indicator, it is possible to define the causes for which all the features' value changed.

The indicator that has been chosen is the "**Game Effectiveness**". It is a weighted average of several factors, that measure as many aspects of the game, like the distance from the opponent goal, the goal visibility, the opponents' pressing near the ball, the number of pass' options available, the velocity of the player and of the ball, and, several other factors. It potentially considers all the variables can affect the performances of a player or of a team and, thus, consequently affect the statistics of the point 1. Since the objective of football is winning, all the factors that are considered in the "Game effectiveness" formula, are highly related to the capacities of a team to win the match.

The magnitude in which these factors affect the punctual value of the "Game Effectiveness" is a secret of the company and it is the most important "component" od the entire product. Only if it is capable to describe successfully and rigorously the effectiveness of a football action, the system can really create valuable insights.

Basing on the fluctuations of the "Game Effectiveness" value, the system explains what it is happening on the football field. Now an example will be provided to clarify the working system. With a similar application is possible for a team A to analyse all the past matches of team B, that is Team A's next opponents. Doing it, it is possible to discover that all the actions in which team B's reached the highest levels of "Game Effectiveness" are associated with a specific game situation, in which are involved two team B's players that exchange their positions quickly.

This make possible to Team A coaches to prepare adequately a scheme that can neutralize this tactic.

Obviously, an indicator like "Game Effectiveness" can enhance the level of individual performance analysis. It is possible associate to each player's behaviour the correspondent level of "Game Effectiveness". This will immediately highlight how much the player is capable to carry out, for example, offensive tasks or defensive ones.

The insights that this system is used to provide are not concepts that are completely unknown to human football experts. However, as seen in section 2.1.4. they systems is able to recognize them in a systematic way and make clear to the coaches' staff what they may have not noticed looking at the match. In fact, a coach must take into consideration too many variables during a football match, and so, it will surely happen that some of them will be undervalued. Artificial intelligence should be employed to make it not happen.

In the next figures, a tangible example of how the Game Effectiveness index works. All these data have been directly provided by Math&Sport itself.

Figure 36:

$$GAME\ EFFECTIVENESS = x * D + y * V + w * P + z * O$$

- D = Distance from the goal
- V= Visibility of the goal
- P= Pressing around the ball
- O= Passing options available
- x, y, w, z = the weights of the formula. As explained before they are a secret of the company

this graph represents the evolution of the Game Effectiveness's values for just a timeframe of a match, between minute 1 and 22.
From an analysis like this one, it is possible to highlight all the cases in which the value has been higher than, for example, 75%.

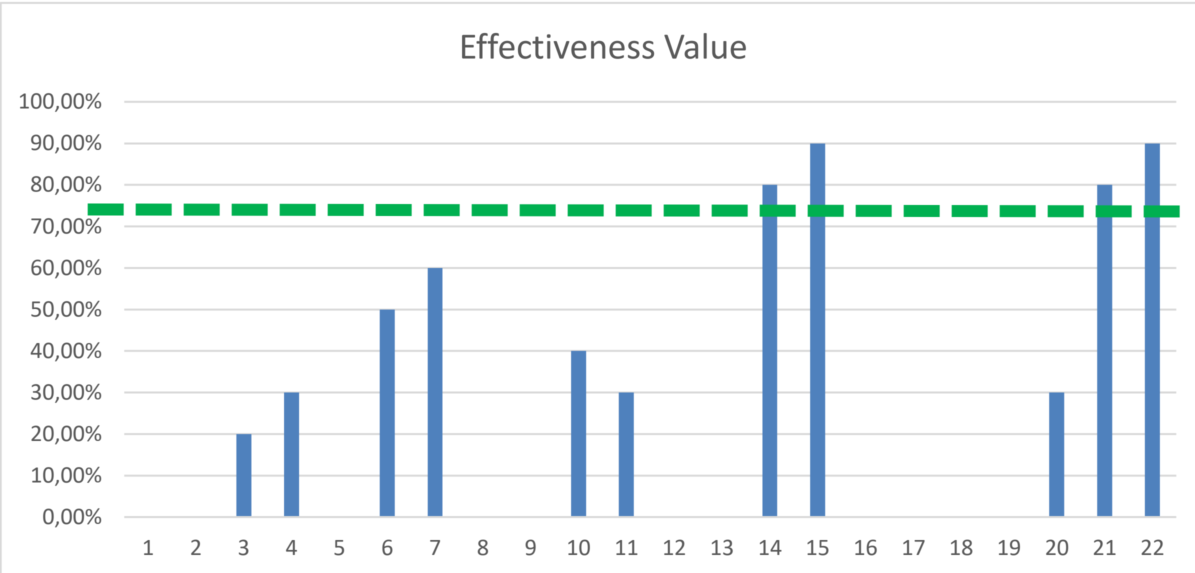


Figure 36: Game Effectiveness graph. All this information has been provided confidentially by Math&Sport.

In all these cases, the system is able to detect some recurrent patterns like, that are for example:

1. Involved players 19,16,18
2. One player checking back towards the ball and the other in checking up the field
3. All actions starting from recoveries in the middle of the Pitch

In this way it is possible to develop a deeper knowledge of the game and have a methodological approach in the analysis of the football games, knowing exactly why and how a team has been effective in developing his play.

The last step of analysis that SDP can offer is a cognitive analysis of the player. It tries to combine psychological aspects with technical ones. The idea is analysing how the players performances change in relation to contextual aspects. So, trying to understand how external variables like, for example, the pressure, the stress or the opponent football style can change the behaviours of players.
A practical example is measuring the number of times in which a player make himself available for a pass in different matches and look at how this number have changed. This could show that in certain

situation the player can be less prone to take initiative, and, this is a meaningful proxy of his state of mind.

Future Development

Math&Sport has already developed a functioning product, and, the aim of the company is offering it from the next season to professional Italian football clubs.

The first commercial version will be a simplified version, as highlighted by the company itself. It has been decided to make it easier to facilitate the introduction in teams' operations. Even in the one that are not used to deal with informatic systems. This will permit to better understand the product, and, get used to its functionalities in a gradual way.

The application will be on-line available. So, it does not require any physical infrastructure. An additional way to make it easier.

To further simplify the introduction, the first version will not include functionalities related to the step 1 of the process. This means that Math&Sport will not sell the application to describe the matches autonomously. In fact, clubs have the possibility to upload their own data on the SDP platform and making analysis basing on them.

The product enables a high level of customization. In fact, it is possible to adapt it to the coaches' football philosophies and ideas. The personalization will happen changing the weights on the "Game Effectiveness" formula. The idea is that a specific coach can preferer some aspects of the game, rather than others. He personally evaluates that specific variables are more important than other in his idea of "Game Effectiveness". It is possible to translate these preferences varying the weights' values in an arbitrary way. So, this system does not aim to be a substitute of a coach. On contrary it can be a valuable tool in amplifying his talent, and, maximize his ability to understand the game.

3.3 CASE STUDIES' CONCLUSIVE ANALYSIS

The objectives of the cases studies were basically two: 1. Deepen the knowledge from an operation point of view and with a first-hand experience of how the organization that are proposing AI-solution in the sport business works and which are their future perspectives, 2. Show the state-of-art of the

Italian ecosystem of artificial intelligence application in the sport business.

The choice of the companies to which dedicate the cases direct by the fact that they are the unique case of start-ups that can be considered active in Italy in this sector.

The first consideration is that, coherently with the 1A.2 chapter, all the solutions aim to deliver value to sport organization trying to limit the negative effect of the bounded rationality of the decision-makers in the respective application fields. (Playerank and Wallabies in the player evaluation process and Math&Sport in the tactical analysis).

The founders have understood the structural limits of the current way to manage some processes within sport clubs. For this reason, they saw the market opportunity to create a start-up to solve these issues.
All of them agree that AI algorithms are the unique techniques able to perform the tasks required to offer a solution to the problems that they were trying to solve. In particular, Luigi Libroia, in his work as business consultant, have noticed that there are no tools to evaluate objectively a football player value. This had created him several problems in his work, since he was asked to evaluate commercially some football clubs. For this reason, he created Wallabies.
Similarly, Luca Pappalardo and Paolo Cintia have demonstrated in their academic works how the journalists' judgement is biased in evaluating players' performances. They discovered that professional sport journalists adopt a very simple process to evaluate players. They consider just few of the total features of a players and judge them basing on noticeable events. So, they have identified space for create Playerank, bringing in the players' evaluation more objectivity and data-driven tools. Ottavio Crivaro and the team of Math&Sport have noticed how much the football tactical decisions are based on personal evaluation of the coaches, without using a cause-effect reasoning approach. So, they developed the platform SDP, to address this issue.

It is useful to notice that no one of these solutions has already a commercial version at June 2018. All the products are still in a R&D phase. This is not a peculiarity of the Italian situation. Very few systems that leverage on AI are already operationally used in the sport business. it is still a sector that, generally, it is still in a research or testing phase. For this reason, and since the algorithms are covered by secrecy to avoid spillovers, an operational analysis of the working methodology has not been possible to be performed.

A common point in all the three cases is the start-ups consider their products as a valuable support for the current club management process and not as a system to automatize activities performed by humans. So, it is important to highlight that their purpose is not substitute human operators in football clubs, but instead, providing them the necessary tools to work with a better-performing methodology and make emerge their talents.

In doing this, they are implicitly introducing in the sport business the idea of value maximization of every decision. The football industry, and in general the sport one, has been traditionally always managed without applying the basic principles of business management like planning the future activities, the research of source of effectiveness and efficiency in the operations and the possibility to measure the performances and the events that occur within the organization.
Artificial intelligence can introduce also in this industry the tools necessary to manage it with the proper business principles because it permits to measure and calculate what until now has been impossible to do.
A further peculiarity is that all the solutions presented need to work a high-detailed database to train the system. The initial database has been bought directly from third-party data provider, as in the case of Wallabies and Playerank, or, it can be automatically generated as in the case of Math&Sport.
In any case, the success of the solution depends massively on the quality of the input data. An interesting point is that all the three systems presented in the so-many cases, work with similar input data.

- Playerank: “detailed data about all the events that occur during a match involving a single player”
- Wallabies: “spatiotemporal data of players’ with ball behaviors”
- SDP: “representation of the evolution of the matches in a formal language”

What is astonishing to highlight is that, even starting with data that are similar, they developed completely different systems, both in term of objectives and in term of analysis performed. In fact, from these data, the systems, perform different activities: Playerank aims to find a framework that can judge a football player just for his contribution to the final result of his team, Wallabies instead want to create a system able to give clear indications to managers about the price at the which buy/sell football players. Math&Sport instead aim to develop a platform that can be used by coaches to exploit at the best the talents of their players and understand the tactical evolution of the football matches.

It is also important to highlight that this typology of data is already collected by all professional football clubs in Italy. So, it is possible to assert that, at least from a technological point of view, the football industry is ready to introduce AI systems.

A further common point in all the cases is the high personalization level of the products. All the companies will create systems that can be personalized by the user, following their preferences and personal settings. This will force to think that AI introduction will not be as a sort of “commodity” product, the same in all the clubs. Instead, every club will probably use the same platform but with the opportunity to personalize it at best, and so it will become a completive advantage being able to use it at the best of its potentiality. Only the clubs able to extract more value from data and to exploit at best the functionalities of AI solution will succeed.

For this reason, all the interviewed see the necessity to develop a close relationship with their future clients and to make them really capable to exploit all the functionalities of their products. Only doing it they will have commercial success.

Passing to the future strategy of the start-ups, all of them want to remain completely autonomous in their working methodology, providing their services to all the clubs interested, and not limiting their partnerships with a small number of clubs. This can force to assume that in the sport business, in the future, there will be the development of AI applications that will support clubs’ activities and there will be several companies operating in this sector in competition between them.

3.3.1 ASSESSMENT OF THE MAPPING PHASE MODEL AFTER THE ANALYSIS OF THE CASE STUDIES

It could be useful at this point, verify how the model derived from the mapping phase of part II, is eventually able to classify also the cases just explained in the case studies. This step cab be seen as a validation of the model itself. In fact, if the model can be employed to classify successfully new cases, it means the starting assumptions are true and the model is valid.

It is possible to assert all the three case studies can be placed within the areas resulted from the mapping phase without relevant issues. So, in general, the model can be considered totally valid.
Nevertheless, some criticalities have been raised and it is useful to highlight them. The first one is related to the fact that all the case studies deal with start-ups that are in the early-stages of life and, thus, their business model is still partially unclear and prone to change. This was the main reason for which it was impossible to place Playerank in just one category for example.
The second peculiarity is that all the three cases studies deal only with football. In this case, it emerged that the AI systems that are employed in the tactical analysis are similar to the ones employed for scouting purposes, at least from an input-data point of view. This point can make to hypothesize that in the future a convergence between tactical analysis and the players' investment optimization areas will happen, and only a combined category of AI systems, will exist.

3.3 CONCLUSIONS

The main objective of the thesis is the evaluation of the current state-of-art of artificial intelligence applications in the sport industry. The work aims to be a research that can be potentially employed by a generic sport organization to plan, operationally, the preliminary steps of its future strategy, regarding this field. To achieve this goal, it is necessary to pass through two logical steps:

1. Map where artificial intelligence has been employed in sport industry, defining:
    - 1.1 How artificial intelligence solutions work from a generic point of view and what are their high-level objectives.
    - 1.2 Which are the AI applications in the sport industry and how they work from an operational point of view.
    - 1.3 Which business areas are impacted by AI solutions and how sport clubs can leverage on these systems.

2. Knowing, on one side, how artificial intelligence works and what it can achieve, and on the other side, how currently the clubs are managed, the second step is evaluating:
    - 2.1 When Artificial intelligence represents a relevant opportunity for sport organizations.
    - 2.2 Considering all the AI categories (figure 17) and their applications areas in the sport industry, which are the main benefits for a sport club.

To answer the entire first question, it is necessary to recap the work done until now in the thesis. This activity will be done in the 3.3.1 part.

The 2.1 question will be addressed in the 3.3.2 part, in which a generic framework to assess the possible opportunities linked to the introduction of artificial intelligence solutions will be provided.

The 2.2 question instead will be answered in the 3.3.3 part.

Figure 36 represents graphically the logical flow followed during the entire thesis, and in which terms every part of the thesis has concurred in answering to the research questions.

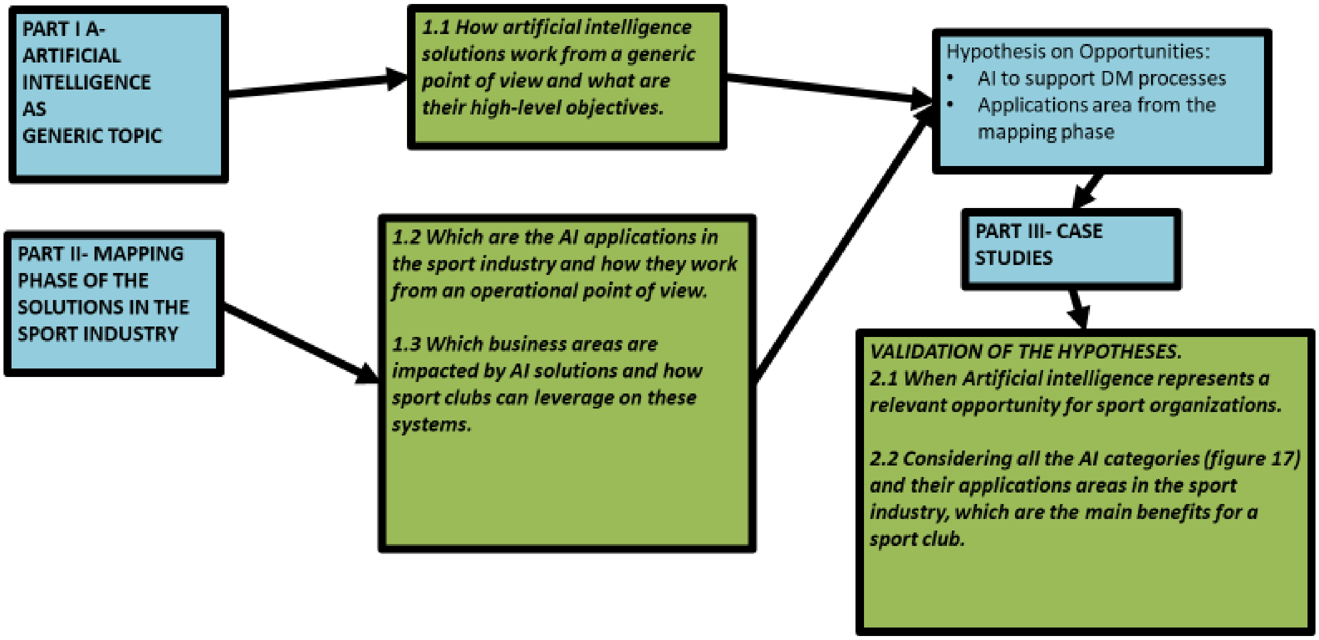


Figure 37: Logical Flow followed in developing the thesis (DM = Decision-Making). In green the research's questions

3.3.1 RECAP OF THE WORK

To describe how artificial intelligence's solutions work from a generic point of view, (first part of 1.1 question) it is necessary to recap what it has been said about the artificial intelligence world and its historical evolution. The first part of the thesis has been totally dedicated to the definition of a clear framework, that could help defining what is Artificial Intelligence. This could be considered as a trivial work, but it is not. The world of artificial intelligence is complex and highly dynamic, and, so, it is hard to be formalized.

To deal with this complexity, the first step has been the use of the definition model presented in the book "*Artificial Intelligence: A Modern Approach*", that showed how it is possible to divide the entire world of artificial intelligence employing a matrix, with two dimensions of analysis. One is about differentiating between systems that can act, or in other words, that can take a decision, from systems that can just process information. The other dimension of analysis is about the very nature of the artificial intelligence, since, on one side, there are systems that are human-inspired and that are build up to imitate how humans act or think. On the other side, there are systems that are required to perform tasks that humans can't. To do it they cannot be human-inspired and so, AI systems are expected to create their 'own' intelligence to face the problem. The authors called this ability "Rationality".

After that, it has been highlighted that the most interesting category is the "**Acting Rationally**" one. It comprises systems able to act, and that are targeting activities, that humans cannot perform. Following this, it has been defined the logical scheme that these systems need to follow in carrying out their tasks. It is the **Perception-Reasoning-Acting scheme**. So, they are required to perceive the environment, to reason about it or analyse it, and, to consequently act.

Passing to describe the operation side of the artificial intelligence world, some issues arise. In fact, it has been said that the value of artificial intelligence systems lay in their ability to perform tasks that human can't. But, the problem is now related to *"How can we teach a system to perform something that we can't?"*. This has been the limit of artificial intelligence in his early stages because it is impossible to codify all the existing knowledge and, it has no sense to try to codify a problem that is not clear. Machine learning was born to overcome these limits. It represented a monumental event in the history of artificial intelligence, because, with machine learning, the systems can learn autonomously, from the data, how to carry out specific tasks. This changed the issues related to the question presented before, since it is no more necessary to teach systems how to act, because now, it is possible to make them able to learn autonomously how to solve the problems. Thanks to machine learning, all the limits of computer science have been continuously and repeatedly moved ahead and, now, machines can learn to solve highly complicated problems.
A further improvement has been presented with Reinforcement Learning, with which artificial systems started to judge their final performances, with the aim to optimize their working methodology, in relation to a specific objective.

The last part has showed how the algorithms' modern logical structures, like Artificial Neural Networks and Deep Learning, can maximize their learning abilities.

Throughout this first part, it has been presented how artificial intelligence systems can really create value for human operators. The value of artificial intelligence system is related to their ability to support human decision-making processes. Human operators suffer of bounded rationality, since their rationality, intended as the discretion to deal with problems following a cause-effect analysis, is highly limited for different reasons:

1. **Complexity**. Several decisions are taken in an environment that is too complex to be understood. We lack an adequate computational power and the necessary mnemonical capacities to deal with this complexity.
2. **Time constraint**. Even if we can understand the environment, we have not time to evaluate all the options.
3. **Cognitive biases**. That are *"systematic deviation from our rationality*"[191]. Our feelings, our emotions and the structure of our mental processes do not permit us to take rational and objective decisions.

Considering all these factors, the optimal decision is a utopia. In fact, business practitioners learn that *"Decision-makers do not aim at optimum, but they settle for satisfactory decision"*[192]. Artificial intelligence, instead, can help decision-makers to aim to take better decisions.

[191] *https://en.wikipedia.org/wiki/Cognitive_bias*

[192] *La gestione dell'impresa: organizzazione, processi decisionali, marketing, acquisti e supply chain, Gianluca Spina, Raffaella Cagliano, Stefano Ronchi, Matteo Kalchschmidt, Federico Caniato, Francesca Bodini e Davide Luzzini, 2006*

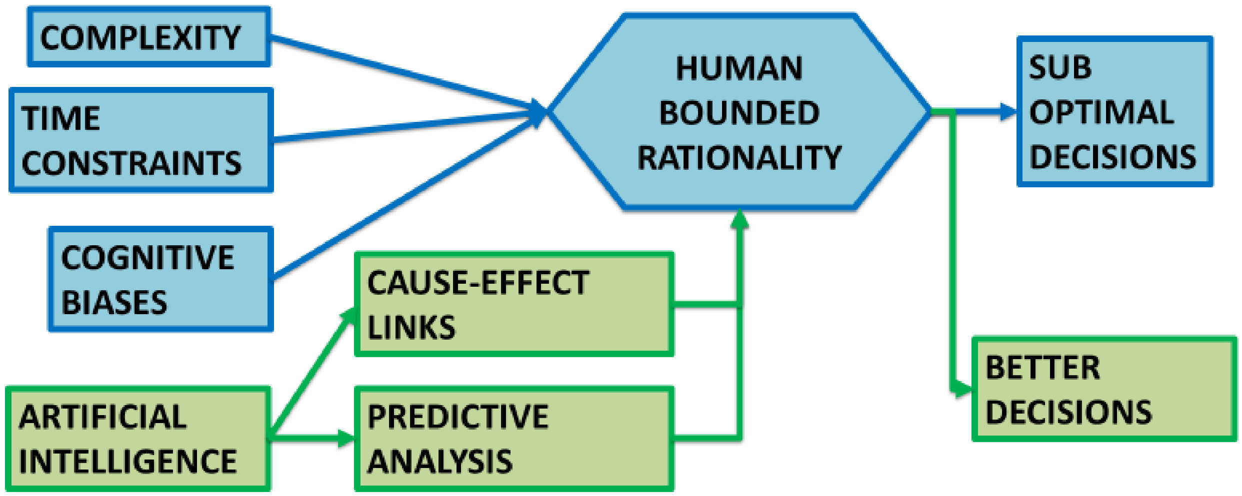


Figure 38: The Value of AI in human organizations

The following part has been employed to present the current world of sport industry, with the related challenges, problems and future perspective. To do it, the "4 Macro-areas" model, created by the Osservatorio Innovazione Digitale nell'Industria dello Sport of Politecnico di Milano, has been employed. It showed how sport industry has become a highly competitive environment, where clubs need innovation in all their operations to create durable competitive advantages.

1.1 How artificial intelligence solutions work from a generic point of view…

- They receive in input big amount of data and they can extrapolate from them relationships, patterns and cause-effect links with a high predictive power. These tasks are performed:
  - Autonomously.
  - With the support of self-learning systems aimed to the continuous improvement.
  - With a remarkable ability to adapt to input nature changes.

…and what are their high-level objectives.

- The high-level objectives of AI systems are supporting the decision-making processes within organizations.

Part II is the result of a research aimed to discover all the cases in which artificial intelligence has been already employed in the sport industry. This section has been used to answer to the 1.2 question, since it shows all the applications and the respective business areas that are impacted in the sport industry, along with their operational tasks. At the same time, also the working methodologies of these solutions and the common features within certain areas have been deducted and presented.

1.2 Which are the AI applications in the sport industry and how they work from an operational point of view.

To answer this question all the AI applications in the sport business has been clustered respect to the nature of the input data.

Figure 35 shows all the applications of AI that leverage on personal data on the health, on the athletic condition and on the training workload adaptation of the single athlete. The output activities are intended as the categories of the solutions. They are employed for: **Athletic Performance Optimization**, **Injuries Management, Technical Gesture Optimization**, **Virtual Training Environment Creation** and for **Scouting & Players Investments Optimization**.

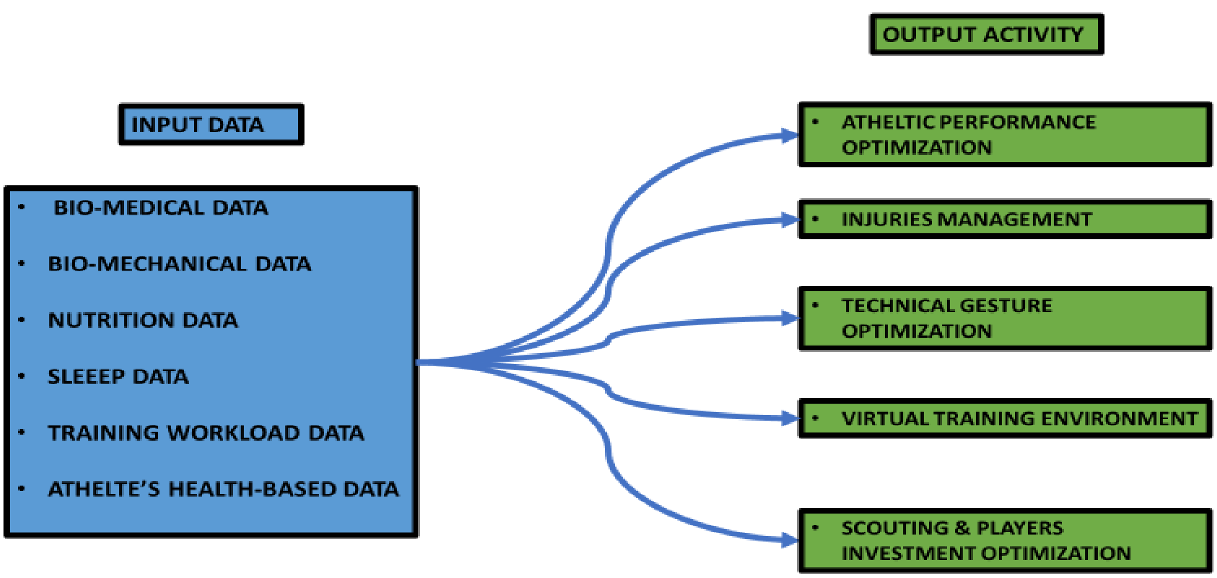


Figure 39: Applications that work with data based on the heath, on the athletic condition and on the training workloads.

Figure 36 shows all the applications that works with data taken from official matches. They are mainly used to figure out in quantitative and objective tools aim to rank and judge the performances of players and to provide meaningful explanations about the reasons for which certain events occurred during a game. This info can be provided to coaches, but also to fans, to increase the quality of the media contents. The applications are: **Tactical Strategy Optimization**, **Scouting & Players Investments Optimization** and **Fan Engagement**.

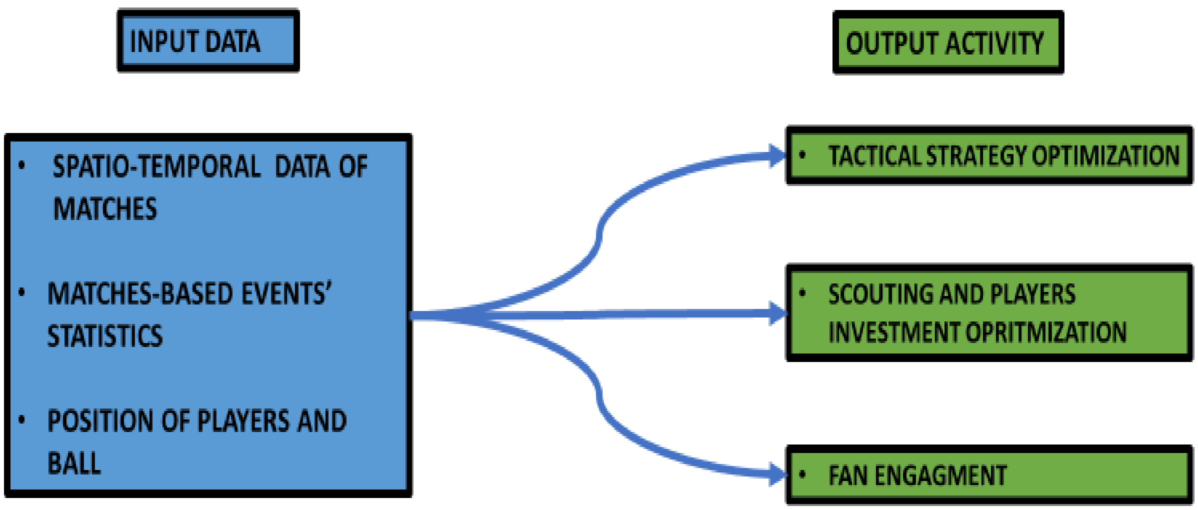


Figure 40: Applications that works with individual and collective data taken from official performances

Figure 37 shows all the AI solutions that can be implemented within a sport organization and are based on data coming from external sources like fans, online activities, facility-based data. They are employed to create and improve new revenues flows for the club itself, in particular for: **Sponsorship Value Optimization**, **Fan Engagement**, **Dynamic Ticketing** and **Smart Arenas Strategy Implementation**.

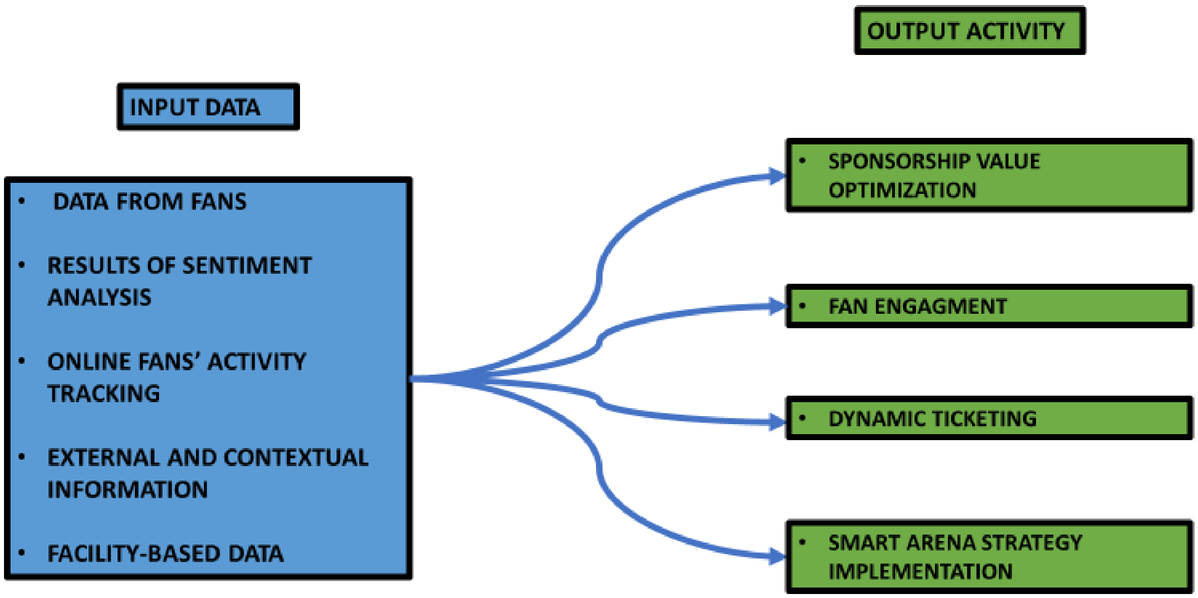


Figure 41: Applications that work with data that come from fans or from sources that are external respect to the sport organization itself.

To answer to the 1.3 question, it is necessary to show the working methodology of these solutions and how a club can leverage on it to improve the effectiveness of its operations along with the business processes impacted.

1.3 Which business areas are impacted by AI solutions…

The business areas that are impacted by the AI solutions are:

- **ACTIVTY & PERFROMANCE OF THE PLAYERS' TEAM, Physical Condition Optimization, Injuries Management, Technical Performance Analysis and Tactical Strategy Optimization.**
- **THE CORPORATE MANAGEMENT & ORGANIZATION OF THE CLUB.** In this case, it is intended as a support for the **Players-Related Investments' Decisions** and as tools for **Enable New Revenues' Streams.**
- **FAN & MEDIA RELATED ACTIVITIES:** AI employed as support for the **Media Broadcasting**

Activities and in the **Arena Management.**

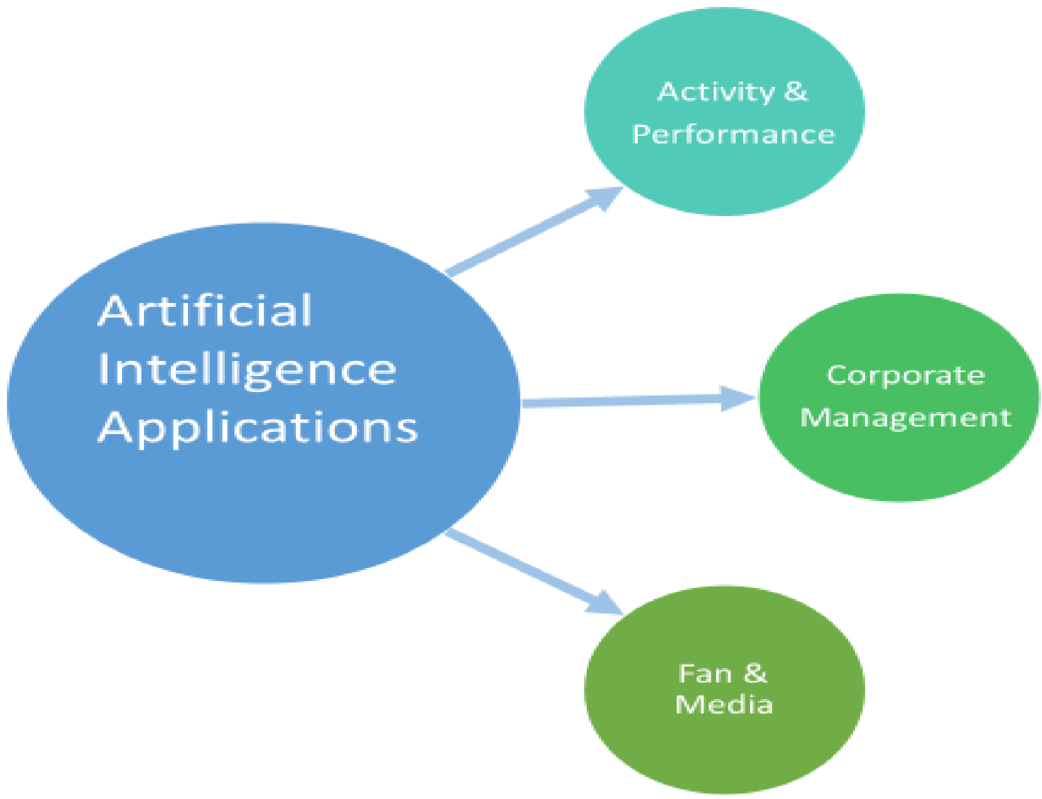


Figure 42: Business Areas impacted by Artificial Intelligence

…and how sport clubs can leverage on these systems.

1. **ACTIVITY & PERFORMANCE**: Solutions dedicated to improving athletes' performances. The solutions can be clustered in relation to their functions in:
    - **Physical Performance Optimization**: Leverage on wearables and the h24 continuous monitoring of the athletes' biological processes, to optimize the athletic condition. The solutions get in input data about the workload, the recovery capacity, the sleeping effectiveness, the blood-related values and other indicators that can be potentially meaningful representation of the athletes' health to forecast players physical performance and to optimize them. The optimization can be done through:

1. The creation of personalized training programs.
2. The possibility of exploit targeted interventions to improve the physical abilities of the athletes, leveraging on generic factors like the one explained before (improving in the diet, sleeping, blood-values, recovery ability…)

- **Injury Management**: Analyse the past training workload and eventually other factors, to forecast the injury risk for athletes. The system takes in input all the past training cycle of the athlete and the occurrences of injuries and find undiscovered cause-effect relation between the intensity of the training and the likelihood of get injured. Potentially, the input data can be enriched with other factors to create more holistic analyses. These data can be collected through GPS systems, since it has been demonstrated in the case study of Luca Pappalardo and Paolo Cintia that the information about speed and acceleration are sufficient to have valuable outputs in this field. The output is an indication about the probability of that a specific player will suffer an injury, and so, a proxy of his need to recover and not play/train.
- **Technical Performance Analysis**: Artificial intelligence system, with the support of other technologies like advanced cameras systems and 3D-based technologies, can be employed to better assess the technical gesture of athletes and the creation of virtual training environment. The systems receive in input videos of past matches/training session from different angulation and it recreate an environment in 3D in which the athlete can re-live in first-hand past plays and, potentially, correct errors that would be impossible to notice in other ways. In this case the application requires a great investment in physical supports like cameras and hardware.
- **Tactical Strategy Optimization**: Artificial intelligence employed to create players' evaluation tools always more omni-comprehensive and meaningful indicators about the decisions and behaviours of players and teams. The input data are the spatiotemporal data of the players and the ball, and the punctual description of all the events that occur during a game. The system applies on this data math

modelling algorithms, unsupervised learning analysis to understand the effectiveness and the consequences of all the events on the results. These concepts have been highlighted, with different nuances, in all the three case studies, where AI-based systems have been employed to evaluate players in an un-biased and objective way considering only their impact of the winning ability of the team. This analysis will be performed to provide insights and optimization to coaches in the preparation of the future matches.

2. **MANAGEMENT & ORGANIZATION**: Solutions dedicated to support the corporate activities of sport organization.
    - **Team-related activity**: Artificial intelligence employed for **scouting** purposes, and for the computation of the fair economic evaluation of **players-related investments**, based only on athletes' performances. The input are data taken from the official matches, and the output shows in an un-biased and objective way the performances of the players. Basing on this information, it is possible to evaluate in a methodological way the value of players in relation to their possible contribution and compare the value at with similar players have been settled.
    - **Commercial-related activity**: Artificial intelligence used to:
        - <u>Support the Evaluation of Sponsorship's Contracts</u>. The systems take in input the real brand exposure of the sponsors and the analysis fan reaction on social media channels to create an estimation of the real value of that sponsorship, that can be computed dynamically. This can permit to sport clubs to sell sponsorships contracts at a fair price and working to increase their value.
        - <u>Support non-Traditional Fan Monetization Activities</u>. The input of the system is the analysis of the fans' activities, feelings and opinions to developed services that can match their requirements at best and, thus, increase the revenues' opportunities for clubs. Examples are like smart-ticketing and the analysis of fan base for targeted marketing strategies.
3. **FAN & MEDIA**: we have two macro-areas:
    - **Media & Broadcasters**: Artificial intelligence employed for the automatic

creation of media content, and for the delivery of more meaningful and engaging media. The idea is providing to the audience of sport events personalized media contents and a large set of information that can make the experience more inclusive and valuable for the clients. The data are the ones related to the team and players performances.

- **Smart Stadium:** Artificial intelligence can be employed to optimize several activities in and around the stadium (traffic, access, security etc.) and improve the fan-club relationship thanks to the delivery of personalized insights.

There are also cases in which AI has been employed by external stakeholders of sport organizations.

4. **EXTERNAL STAKEHOLDERS:**
   - **Outcome Forecast:** Artificial intelligence is employed by external stakeholder for betting purposes. In particular, some hedge-funds and start-ups started to exploit the predicative powers of AI solutions to create investments strategies in the sport betting field.
   - **Referee Support**: Artificial intelligence can be also used as tool for increase objectivity in the judgments of referees.

This recap of all the work done until now, have permitted to completely solve the first bulleted point presented at the beginning of chapter 3.3. The following part will be employed to answer the remaining research questions.

3.3.2 A QUALITATIVE FRAMEWORK FOR AI SOLUTION INTRODUCTION

Now, it is evident how artificial intelligence systems works, which sport business' areas are impacted, and, in which terms, AI applications can be exploited by sport organizations.
The following point is answering to the question 2.1.

2.1 When Artificial intelligence represents a relevant opportunity for sport organizations.

For simplicity, it can be rephrased as follow:

(premise: I am a sport club and I have understood that artificial intelligence can represent an opportunity for me), "*how can I evaluate the benefits of the introduction of artificial intelligence system in my operations?*". This is the question presented at point 2.1.

To answer to this question, it is necessary to provide a qualitative framework that can be employed by a sport club, to evaluate the potential benefits of the introduction of artificial intelligence system in its operations.

The idea behind is that a lot of AI-based solutions exist regarding sport business. They address different processes and activities. Nevertheless, currently, most of clubs still do not employ any of these systems though. So, it could be useful to provide some guidelines to sport clubs in the selection of the area in which firstly introduce AI system.

This framework is thought to support a club in the identification of the processes in which artificial intelligence solutions can create the biggest benefits for them. This is the first step to introduce artificial intelligence solutions. Once identified that a specific activity can be highly improved thanks to the support of artificial intelligence, the club should compare the different solutions on the market, and, choose the one that better fit their requirements.

The objective of this framework is just providing guidelines in the preliminary evaluation of the artificial intelligence potential.

Artificial Intelligence has been defined as a tool to support decision processes. This concept has been also profoundly highlighted in the case studies. All the systems presented are solutions that try to support the decision processes within sport organizations, limiting the negative effect of the human bounded rationality.

Once it has been analysed the value proposition of the companies presented in the case studies, and it has been highlighted how they aim to deliver value to sport club, it is possible to apply an inverse

process and reach the reasons for which clubs would mostly need these solutions.
The idea is starting from:

- The high-level objectives of the AI start-ups analysed in the case studies.
- The needs that the companies have identify as currently unsatisfied and, thus, consequently, they have decided to develop a product to solve these issues.

From these analyses it is possible to show that when in sport clubs rise the same issues that the AI companies aim to solve, the situation will perfectly match, and the benefits will be as much as possible.

So, it is possible to assert that the more the current decision process is sub-optimal, the bigger are the opportunities to introduce artificial intelligence systems. So, clubs should evaluate their current decision processes, and, select the more critical ones. These will become the first candidates to be improved by artificial intelligence's introduction. To evaluate the level of criticalities related to a process, it is possible to look at some features that characterize the process itself.

The more the process is **complex**, in term of the number of variables that must be considered to take the final decision, the more it is likely that the final decision can be wrong and, so, the higher the potential for artificial intelligence's introduction. An example is the evaluation of the physical condition of the athletes that are impacted by many factors, as it has been highlighted in section 2.1. The higher the number of variables to be considered, the lower is the probability to have omni-comprehensive evaluations without the support of computer systems, and so, the potential benefits associated to AI are relevant.

Another feature to be considered is the **level of subjectivity** in the final decision. if it is the result of subjective evaluations, instead of formalized models, it is more likely that it is biased. For example, the process of players' value estimation is, at least in football, highly based on personal evaluations of the coaches and managers, and this can result in overpaying some players. In this case, artificial intelligence could bring objectivity in the evaluation process, and so, it represents a relevant improvement in the corporate management of the team.

The last aspect that must be necessarily considered in the evaluation of the benefits of the introduction if artificial intelligence solutions, is the **relevance of the process itself**, both in term of reversibility and in term of operational and economic impact. The lower is the reversibility, the higher the importance to take correct decision. The higher is the impact of the processes on the sport and economic team's performance, the higher is the value of managing those processes properly. An example could be the injuries management process. Supposing that the best player in a team is in overtraining situation, so if he will train again, he will have high risk of get injuries. If the team management decide to train him and he gets injured, the decision is not reversable, since he needs to stop and take care of the injure, and the decision has also a great impact on the performance of team since he is one of best players. The impact is financially relevant too, since he will still earn a rich wage without playing. So, it follows, that it is an area in which artificial intelligence can generate immediate benefits for a team.

The figure 38 is a graphical represenation of the framework just proposed.

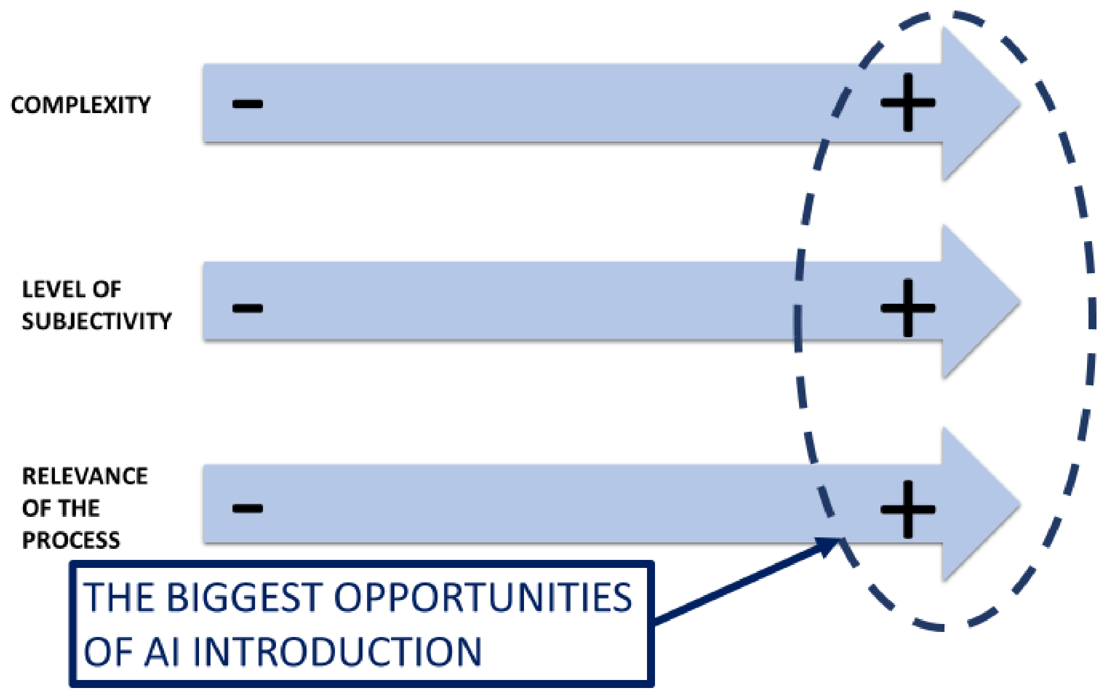


Figure 43: scheme of the qualitative framework

2.1 When Artificial intelligence represents a relevant opportunity for sport organizations.

The introduction of Artificial Intelligence systems in the sport organization can represent a relevant opportunity when the process that the AI system aim to support is currently managed in a sub-optimal way. To define when it is managed in a sub-optimal way it is possible considering the level of complexity, the level of subjectivity and the relevance of the process itself.

3.3.3 SO WHAT? THE MAIN ARTIFICIAL INTELLIGENCE'S BENEFITS

It has been just explained that the opportunities of the AI introduction are related to the fact that these systems will support decision processes within sport organizations, that are currently managed in a sub-optimal way.

This will create great benefits in the sport clubs because this will enable the introduction of business management principles also in the sport industry, a sector generally managed without taking them into consideration.

As highlighted in three case studies, the value of the solutions proposed is mainly related to the fact that they can support the management of sport clubs, to take decisions with the aim of value maximization. The common point in all the case studies is the capacity of the solutions to make the agents taking more informed decisions.

As explained in the section 3.3 (the case studies conclusive analysis), the football industry, and the sport one in general, has been traditionally always managed without applying the basic principles of business management, like future planning, the research of efficiency and effectiveness sources in all the processes and the possibility to measure the performances and the events that occur within the organization.

Most of the decisions that are taken in football clubs are not the result of processes in which the trade-offs between costs and benefits, threats and opportunities are weighted and assessed.

This has done because the decisions in the football industry are impacted by external variables that, at least, traditionally, has been considered as unpredictable or anyway affected by a high level of uncertainty. For example, it can happen that a player will be not in the condition to play at best of his abilities and this is difficult to be forecasted employing the traditional tools , or, that for some reasons, a lot of players are affected by injuries and this affect negatively the team's performance, or, that a

player that has been bought through a relevant investment will not meet the expectations.
The risks and the uncertainty linked to all these external factors made impossible to manage a football club with the traditional business principles.

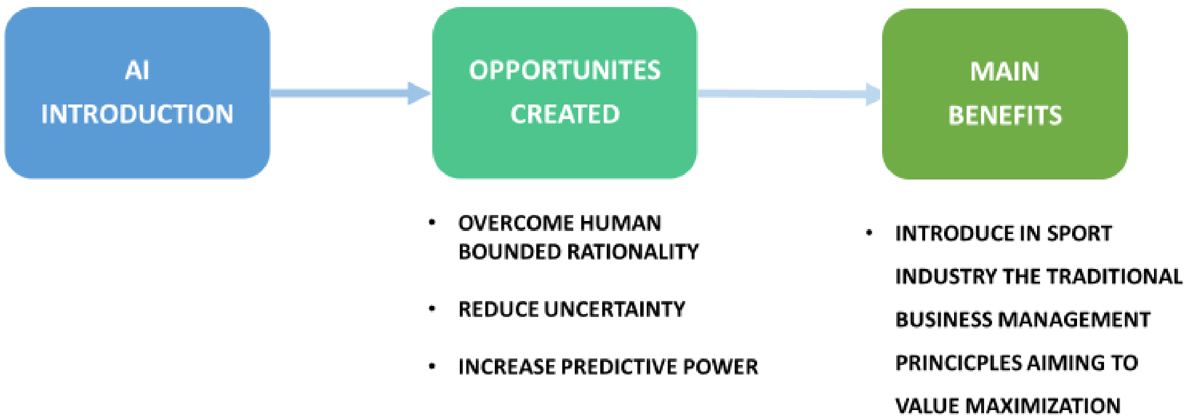


Figure 44: the value related to AI systems introduction.

Through the literature analysis, the mapping phase and the development of the case studies, it emerged that AI systems can provide the necessary tools, to manage the uncertainty that derives from all these external factors. It is possible since AI can figure out cause-effect relationships, patterns and rules in the data received in input, providing to the sport clubs' management the necessary tools to interpret the unknown. From this, it derives that if the sport clubs' top management can reduce the uncertainty and unpredictability that characterized the industry, they can start to manage the organizations with a more traditional business approach, and take decisions aiming to maximize always the value that are the results of the analysis of all the related costs and the potential benefits.

This is the central point that emerged from the combined analysis of the literature, of the mapping phase and of the case studies.
Considering the previous work of the thesis, and the guidelines just presented, it is possible to go through all the areas in which artificial intelligence has been employed regarding sport industry (Figure 17) and, define where lay the main benefits of AI solutions and, thus, answer to the last research question.

2.2 Considering all the AI categories (figure 17) and their applications areas in the sport industry, which are the main benefits for a sport club.

Athletic Performance Optimization: In this category the benefits could be numerous. As it has been seen, the athletic condition of players is impacted by several variables (diet, sleep cycle, biological process…) and each of them impacts non-linearly and, in different magnitude athletes, since body's needs and reactions, are highly individualized. So, the introduction of artificial intelligence systems, jointly with wearables, can support the monitoring of all the relevant factors that can affect the physical condition of players. The system can analyse all the data-records to find the causal links between these variables and the final performance of the single athlete. This can enable the team management to leverage on them in a systematic and highly personalized way, to improve the health and the ability to perform of the team and reduce at the same time the occurrences of unpredictable under-performance. The level of subjectivity of the current process is also necessarily high in case in which teams do not use an individualized approach in the management of physical performances. To conclude, also the impact of an improvement of the processes thanks to AI-based systems, can be highly positive in modern sport environment, where the competition is getting always more relevant and to win it is necessary to have athletes at best of their abilities.

Injuries Management: Injury management, intended as the estimation of the probability to suffer an injury for an athlete, basing on past and future workload analysis, is a field in which the potential of AI-powered system can be notable. The variables that should be consider for drawing a complete picture, are numerous and, furthermore, they impact every single athlete in an individualized magnitude. The impact of an effective injury management is relevant for teams, both in term of sport performances and financial position because it enables to have a reliable predictive power to forecast future states. Artificial intelligence's ability to deal with complexity and variability can create the opportunity to develop powerful tools, that can automatically process all this information and provide a highly calibrated results for the single athletes, that can help the staff to better manage athletes' work cycles.

Tactical Performance Management: Tactical analysis is a task that is highly complex. Coaches are asked to evaluate the mutual interactions of 22 players (speaking about football) and exploit them to have

competitive advantages for winning matches. The judgment can be highly biased, for this complexity and for the duration of the matches. (potentially coaches are asked to monitor all the tactical aspects for 90 minutes). Artificial Intelligence in this field can be successfully employed to support the tactical preparation, making clear and quantifiable the evolution of matches. It can also permit to coaches to not neglect some details, and, to introduce in this field a casual approach. AI-powered systems can also provide an omni-comprehensive evaluation of players' and teams' performances, considering more aspect respect traditional matrixes.
As highlighted also in the Math&Sport case, the introduction of AI system can be a valuable tool in the hand of coaches to optimize the tactical strategy. The AI systems can be highly personalized, following the preferences and the personal evaluations of each coach. This will create the premise to make the coach even more relevant than today, since it can exploit at best its talent.

Scouting and player investment optimization: In this field artificial intelligence systems can be employed to support the scouting processes, since it can manage it in a more methodological and in an evidence-based way respect to the current methods, that are still based on personal evaluations of scouts and managers. Furthermore, in the process of players' value estimation, the potential of AI is relevant, since it can introduce objectivity in players' values definition, linking them just to the effective performances. This could represent a big advantage for teams, since signing new players, is always more expensive and, so, errors can have a profound impact on teams. As showed by the Wallabies case, this can represent a big opportunity for club. The occurrences of bad transfers will be lower, and the management will have all the information it needs to evaluate properly the players and their future impact in the new team.

Sponsorship value estimation: Artificial intelligence can be used in evaluation the media sponsorship contract, in a dynamic and quantitative way. As highlighted in 2.2.2, there are numerous variables that affect the fair economic value of sponsorships, like quality of the exposure, reaction of fans and other contextual aspects. Artificial intelligence can be successfully employed to make these evaluations more precise, enabling sport clubs to sell them at a fair price.

Non-traditional fan monetization: The applications in this field are mainly related to ticketing process and fans' behaviours analysis for marketing purposes. In the first case, artificial intelligence systems have been employed to identify patterns in the tickets' purchase process, that can be exploited for maximizing the economic return for club, through real-time changes in tickets' prices. In this case, the utility of artificial intelligence is related to situation in which there is a high variability of volumes without a clear explanation.

Regarding fans' behaviour analysis, through the mobile devices and on-line activities, artificial intelligence is potentially able to detect regularities, patterns and trend that can be used by the club to offer personalized contents, marketing promotions and initiatives to every single fans. The importance of delivering individualized engaging strategies is always more relevant, since the fan bases of the biggest sport organizations are worldwide spread and made of different clusters. To maximize the effectiveness the retention and the fans' fidelization it is necessary to perform ad hoc activities.

Smart-stadium: Artificial intelligence systems have been employed to improve the clubs' facilities management. Artificial intelligence can provide useful insights for fans to make experience more valuable. There are cases in which artificial intelligence has been used in the crowd management within arenas, in the flows management for accessing the facility, and, in the security system. The common feature of all the artificial intelligence systems related to arenas is that a centralized management can make everything more efficient. So, ideally, the traffic flows can be optimized, giving indication to fans about which root is better to take, the queues can be speed up thanks to an efficient utilization of the accesses, and, the security checks can become less invasive and faster.

In this specific area, majority of artificial intelligence solutions existing in the world have been developed by IT companies specialized in the digital management of big facilities, also because even if the facilities are fundamental assets for sport organizations, their operational management is not part of the core business. So, it is unlikely to see artificial intelligence solution in this field, developed directly by sport organizations.

Fan engagement: Artificial intelligence systems can be employed as a tool for supporting the choice of the best engagement strategy to adopt with each fan. it can also be employed to develop new media contents to fans. It is used to deliver new data and information about the teams to fans, before, during and after the match. The idea is to make the fan always more active, enabling them to have a first-hand relationship with the club and creating immersive experiences, thanks also to other technologies like VR and AR.

To conclude, it is necessary to dedicate attention to the eventual introduction phase of artificial intelligence technologies in a real sport organization. As highlighted by all the case studies, the introduction of a technology like artificial intelligence, can generate conflicts and tensions within an organization. In fact, it has the potential to modify profoundly how tasks have been performed until that moment.

So, to create value in long-term, the introduction must be gradual and step-by-step. Human operators have the necessity to get used to the new working methodology that artificial intelligence system requires, in an accommodating way.

To do it, all the interviewed, agrees that the central point is to make human trust the artificial intelligence system. To do it, it is necessary that the working methodology of the artificial intelligence system must be comprehensible. The idea is that, the human operator should be able to retrace the steps that brought the machine to take a certain decision, and, recognize that they are valuable and correct. So, the system should not work as a black-box. Instead, it should have a high level of interpretability, meaning that humans should have no difficulties to figure out how the machine reached a solution (this was one of the issues highlighted also on 1A.9).

This is necessary, because human operators should be able to improve their performances thanks to AI introduction. This means that AI, in providing the explanation of its results, can highlight to human operator where are his possible faults, and, make him consider them the following time.

It is important that within the sport organization AI is perceived as a collective improvement tool, that can amplify the talent and not substitute humans.

A successful introduction of AI-based system can permit to the sport organization to get immediately the benefits presented before, and, consolidate the competitive advantages that artificial intelligence can generate.

3.3 FURTHER DEVELOPMENTS

In writing this thesis, some interesting points for further deployments have been found. The first topic is related to the strategy that teams can adopt in the future to develop and improve artificial intelligence systems.
In particular, it is related to the possibility to adopt strategies of Co-opetition between different teams. Indeed, the effectiveness of artificial intelligence systems partially depends on the training set. So, the best way to increase the training set is use the same system in more than one team. An example could be a machine learning application, that is adopted and improved, at the same time, by more than one clubs. in fact, it is in the interests of the clubs to collaborate to improve its performance.

A second topic is that currently F.I.F.A, the global regulatory body of football, is starting to develop regulatory frameworks that are dedicated to the management of the players' data. As highlighted by Paolo Cintia in the case study, the FIFA is thinking to introduce rules that can regulate the sharing of players-based data between clubs. One of the ideas is to impose clubs to share all the data related to a player when he is sold to another club, in order to make all the historic record available also for the buyer. This will create the regulatory base, that will basically force sport teams to compete in the data analysis, because if team will have access to always bigger databases and they are obliged to share them. For this reason, the value will lay in how the data will be analysed.

As highlighted by Mr. Crivaro, a third aspect is that probably now, at least, football is a breakthrough point. FIFA have allowed for the first time in history the use of real-time data by coaches in the FIFA World Cup 2018's matches. So, the institutions have defined clearly that in the future in football will be possible the real-time data analysis. This will definitely open the market, moving it out from the current a grey area of deregulation. Math&Sport is already evaluation this potential market since it is developing application able to real-time analysis. This will be possible thanks the new 5G network that is expanding in Italy and that will provide the adequate technological infrastructure to these new applications to work properly.
Mr. Crivaro highlighted that the algorithms and the new internet speed standard commonly available, will also provide rooms for augmented reality application in the fan engagement field, mainly related

with the “What if?” concept explained in section 2.3.1.

3.4 LIMITS OF THE RESEARCH

The main limit of the research is related to the impossibility to analyze more real-world cases that could have generate a higher knowledge of the topic in terms of benefits and issued related to the introduction of AI-based systems in leading sport organization around the world.
In general, the lack of detailed information has limited all the phases of the work development. In fact, dealing mainly with start-ups, there are no information released to the public from these entities and, to protect themselves for opportunistic behavior of competitors, they tend to not share outside any relevant information.

The second limit is that, to evaluate in a comprehensive way the impact of artificial intelligence solutions, it is necessary to know exactly how sport organization currently are managed. Even if it is possible to find qualitative information about the current state-of-art sport management, to create a comprehensive work, it is required to compare deeply the status-quo with what could be with AI.

The last limitation is related to the fact that there are some activities in the operations of sport organizations that are not related to an engineering background. For example, to deal properly with the fields of health management, physical condition evaluations and injuries related considerations, it should be suitable to have specific competencies in that fields.

BIBLIOGRAPHY

- 2018 media and entertainment industry outlook, reaching new heights through personalization and mobility, Deloitte report
- a deep learning approach to predict mild traumatic brain injury in contact sports, Yunliang Cai, Wei Zhao, Zhigang Li, Songbai Ji, 2017
- A Machine Learning Approach to Predicting Winning Patterns in Track Cycling Omnium, Bahadorreza Ofoghi, John Zeleznikow, Clare MacMahon, Dan Dwyer,
- A Machine Learning Framework for Sport Result Prediction, Rory Bunker, Fadi Thabtah, 2017
- A Neural Algorithm of Artistic Style, Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, 2015
- A New Application of Linear Modeling in the Prediction of College Football Bowl Outcomes and the Development of Team Ratings, Brady T. West, Madhur Lamsal, 2008
- A Study on Overfitting in Deep Reinforcement Learning, Chiyuan Zhang, Oriol Vinyals, RemiMunos, Samy Bengio, 2018
- Acceleration in the NBA: Towards an Algorithmic Taxonomy of Basketball Plays, Philip Maymin, 2013
- Allocative and Dynamic Efficiency in NBA Decision Making, Matt Goldman, Justin M. Rao, 2011
- Amplifying Prediction Accuracy using Swarm A.I., Louis Rosenberg, Niccolo Pescetelli, 2017
- An Economic Evaluation of the Moneyball Hypothesis, Jahn K. Hakes and Raymond D. Sauer, 2006
- An enhanced metric of injury risk utilizing Artificial Intelligence, Calham Dower, Abdul Rafehi, Jason Weber, Razali Mohamad, Alerte Digital Sport, 2018
- An introduction to Artificial Neural Network, Kuldeep Shiruru, 2016
- An introduction to Neural Networks and Deep Learning Talk given at the Department of Mathematics of the University of Bologna, Andrea Asperti , February 20, 2018
- An Introduction to Particle Swarm Optimization, Matthew Settles, 2005
- Analyzing In-Game Movements of Soccer Players at Scale, László Gyarmati, Mohamed Hefeeda, 2016

- Analyzing In-Game Movements of Soccer Players at Scale, László Gyarmati, Mohamed Hefeeda, 2016

- Application of Neural and Regression Models in Sports Results Prediction, Adam Maszczyka, Artur Gołaśa, Przemysław Pietraszewskia, Robert Rocznioka, Adam Zająca, Arkadiusz Stanulaa, 2013
- Application of Neural Networks to Analyze Performance in Sports, Perl Jürgen, Baca Arnold

- Artificial Intelligence A Modern Approach, third edition, Stuart Russell, Peter Norvig, 1995
- Artificial Intelligence and Economic Theories, Tshilidzi Marwala, Even Hurwitz, March 2017
- artificial intelligence and life into 2030, one hundred year study on artificial intelligence | report of the 2015 study panel | september 2016
- Artificial Intelligence for executives, Integrating AI into your analytical strategy, Sas
- Artificial Intelligence in Sports Prediction, Alan McCabe, Jarrod Trevathan, 2014
- Artificial Intelligence in the 21st Century, XIangijie Kong, Ivan Lee, Feng Xia, 2018
- Artificial Intelligence Index 2017 Annual report, Yoav Shoham (chair), Raymond Perrault, Erik Brynjolfsson, Jack Clark, 2017
- Artificial Intelligence vs. Human Intelligence Man vs. Machine, Discussion paper on Artificial Intelligence, Human Intelligence and the rapid transformation of digitalization and datafication that our current society is going through, Lucas Kromann Bøgehøj Nielsen, December 2016
- Artificial Intelligence, Automation, and the Economy National Science and Technology Council PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE, 2016
- Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care?, Stefano A. Bini, 2018
- Artificial Intelligence, the next digital frontier ? McKinsey global Insitute, Discussion paper, June 2017
- Artificial Neural Networks for Beginners, Carlos Gershenson
- Artificial Swarm Intelligence, a Human-in-the-Loop Approach to A.I., Louis Rosenberg, 2016
- At the gate and beyond, outlook for the sports market in north America through 2019, Pwc
- Beating the Bookmakers: Leveraging Statistics and Twitter Microposts for Predicting Soccer Results, Fréderic Godin, Jasper Zuallaert, Baptist Vandersmissen,
- Behavioural economics: development, condition and perspectives, Dimitar Kanev, Venelin Terziev, 2017 Belief Networks, By Honglak Lee, Roger Grosse, Rajesh Ranganath, and Andrew Y. Ng, 2014
- Beyond Expected Goals, William Spearman, 2018
- Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science, Robert Rein and Daniel Memmert, 2016
- Bounded Rationality, Coralio Ballester, Penelope Hernandez, 2012
- Building high-level features using large scale unsupervised learning, Quoc V. Le, Rajat Monga, Matthieu Devin, Greg Corrado, Kai Chen, Marc'Aurelio Ranzato, Je_ Dean, Andrew Y. Ng, 2011
- Chapter 3, From Rational Choice to Behavioural Economics, Theoretical Foundations, Empirical Findings and Legal Implications, laus Mathis and Ariel David Steffen,

- Commercial Trends in sport 2017, Nielsen Sport, 2017
- Complex Adaptive System, Serena Chan, 2001
- Complex Adaptive systems, The health foundation, 2010
- Complex Systems in Financial Economics, Applications to Interbank and Stock Markets, Daan Laurens in 't Veld, 2014
- Computational intelligence in sports: Challenges and opportunities within a new research domain, Iztok Fister Jr, Karin Ljubiˇc, Ponnuthurai Nagaratnam Suganthan, Matjaž Percd, Iztok Fister, 2015
- Connected Sports Fans 2016, Avaya, 2016
- Consumer experience quality: A review and extension of the sport management literature, Masayuki Yoshida, 2017
- Deep Blue, Christoph Schulze,
- Deep Learning and the Artificial Intelligence Revolution, MongoDB White Paper, 2017
- Deep Learning for Causal Inference, Vikas Ramachandra, 2018
- Deep Learning of Player Trajectory Representations for Team Activity Analysis, Nazanin Mehrasa*, Yatao Zhong*, Frederick Tung, Luke Bornn, Greg Mori, 2017
- Deep Learning Quantification, Master thesis by Andrea Azzini, Giovanni Battista Conserva, ay 2015-2016
- Deep Learning, Yann LeCun, Yoshua Bengio & Geoffrey Hinton, 2015
- Deloitte's sports industry starting line-up, Trends expected to disrout and dominate 2018, Deloitte
- Disruption ahead, Deloitte's point of view on IBM Watson, Deloitte, March 2015
- Empathic media: the rise of emotion AI, Andrew McStay, 2017
- European Sportech Report, leAD sports tech X, 2018
- FanRevolution, Global fans in the information age. Who are we? What motivates us? How do we create value?, Repucom
- Feature Selection in Supervised Learning, Jie Shi, Robin M. Gubela, 2018
- Future of Sport Report, nbn Australia's broadband network
- Game on! Winning at the business of sports, IBM, 2015
- Global media sector trend 2018, DLA piper
- Goal!! Event detection in sports video, Grigorios Tsagkatakis, Mustafa Jaber, Panagiotis Tsakalides, 2017
- Hierarchical Imitation and Reinforcement Learning, Hoang M. Le, Nan Jiang, Alekh Agarwal, Miroslav Dud´ık, Yisong Yue, Hal Daum´e III, 2018
- How deep learning works — The geometry of deep learning, Xiao Dong, Jiasong Wu, Ling Zhou, 2017

- How push notifications impact sports&recreation app retention rates, Urban airship
- Human Decision Making and Artificial Intelligence - Comparison in the Domain of Sports Prediction, Arnu Pretorius, Douglas A. Parry, 2016
  in U.S. sports fans' tweets, Yang Yu, Xiao Wang, 2015
- Introduction to Deep Neural Network, Liwei Ren, Trend Micro, 2016
- Machine Learned Learning Machines, Leigh Sheneman, Arend Hintze, 2017
- Machine learning approach to model sport training, Edward Mezyk, Olgierd Unold, 2011
- Machine Learning: An Applied Econometric Approach, Sendhil Mullainathan and Jann Spiess, 2017
- Machine learning: things are getting intense, Deloitte
- Machine Who Think, Pamela McCorduck, 2004
- Machines that learn in the wild, Machine learning capabilities, limitations and implications, Nesta 2015
- Mastering the game of Go with deep neural networks and tree search, David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, Demis Hassabis, 2016
- Match analysis in football: a systematic review, hugo sarmento, rui marcelino, m. teresa anguera, jorge campaniço, nuno matos, and josé carlos leitão, 2014
- Move Prediction in Go - Modelling Feature Interactions Using Latent Factors, Martin Wistuba and Lars Schmidt-Thieme
- Naturalness of Artificial Intelligence, Jan Romportl, 2014
- Notes from the AI frontier: Insights from hundreds of use cases, Mckinsey Global Institution, Discussion Paper April 2018
- Optimal Forecast Combination Based on Neural Networks for TimeSeries Forecasting Lin Wang, Zhigang Wang, Hui Qu, Shan Liu, 2017
- POINTWISE: Predicting points and valuing decisions in real time with NBA optical tracking data, Dan Cervone, Alexander D'Amour, Luke Bornn†, and Kirk Goldsberry, 2014
- Predicting The Dutch Football Competition Using Public Data: A Machine Learning Approach, Niek Tax and Yme Joustra, 2015
- Probabilistic Supervised Learning, Frithjof Gressmann, Franz J. Király, Bilal Mateen, and Harald Oberhauser, 2018

- Quantifying the Value of Transitions in Soccer via Spatiotemporal Trajectory Clustering, Jennifer Hobbs, Paul Power, Long Sha, Hector Ruiz, Patrick Lucey, 2018
- Redefining home field advantage, A game plan for enhancing engagement, Deloitte
- Reinforcement Learning: An Introduction, Richard S. Sutton and Andrew G. Barto,2014
- SAP for sport sports & entertainment, SAP
- Scout7 changing the game, Intel, 2016
- Smart shopping, a playbook by the store WPP in partnership with IBM, how artificial intelligence is transforming the retail conversation, David Roth, Karen Lomas,
- Smart stadiums take the lead in profitability, fan experience, and security, Intel, 2016
- Sport Insights Central Report, IBM
- Sport tech market trend overview, Sport Innovation Lab, 2018
- Sports tech innovation in the start-up nation, Deloitte
- Sports: the most disrupted of all industries?, Pwc’s sport survey 2017, Pwc
- Swarm Intelligence in humans: a perspective of emergent evolution, Yong Tao, 2017
- Technology, Media and Telecommunications Predictions, Deloitte, 2017
- The 5 essential components of a Data Strategy, Sas
- The Advantage of Doubling: A Deep Reinforcement Learning Approach to Studying the Double Team in the NBA, Jiaxuan Wang, Ian Fox, Jonathan Skaza, Nick Linck, Satinder Singh, Jenna Wiens, 2017
- The business of sport, playing to win as the game unfurls, KPMG, 2016
- The Chinese Room argument, first published Fri Mar 19, 2004 substantive revision Wed Apr 9, 2014
- The concept of Intelligence in AI, Ipke Wachsmuth,
- The European football club, a data driven analysis focused on football and social media, iquii sport, 2018
- The future of Sport fan, Performance Communication and Canvas8,
- The human factor: working with machines to make big decisions, pwc report
- The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation, Future of -Humanity Institute, University of Oxford, Centre for the Study of Existential Risk, University of Cambridge, Center for a New American Security, Electronic Frontier Foundation, OpenAI, February 2018
- The New Wave of Artifcial Intelligence, whitepaper, Every
- The rise of artificial intelligence: future outlook and emerging risks, Allianz Corporate&Specialty,
- The science of sleep and workplace fatigue, the risks and the costs of human fatique, and how one technological solution is helping predict and prevent it, Fatigue Science
- The social side of sports sponsorship, GumGum report

- The stadium as a platform, A new model for integrating venue technology into sport business, Deloitte
- The total fan engagement playbook, Parasonic
- Thinking machines, Mattia Grava, 2017
- Understanding Machine Learning: From Theory to Algorithms, Shai Shalev-Shwartz and Shai Ben-David, 2014
- Unified measurement, defining a new sponsorship currency, Nielsen Sport,
- Unsupervised Learning of Hierarchical Representations with Convolutional Deep
- Using Computer Vision and Machine Learning to Automatically Classify NFL Game Film and Develop a Player Tracking System, Omar Ajmeri, Omar Ajmeri, 2018
- Using Digital Signals To Measure Audience Brand Engagement At Major Sports Events: The 2015 MLB Season, Peter Ibarra, Dstillery, Peter E. Lenz, Dstillery, 2017
- Views into the Chinese room: new essays on Searle and Artificial Intelligence, John Mark Bishop, 2002
- Visual soccer match analysis using spatiotemporal positions of players, Vinicius Machado, Roger Leite, Felipe Moura, Sergio Cunha, Filip Sadlo, João L.D. Comba, 2017
- Wage against the machine: a generalized deep-learning market test of dataset value, Philip Z. MAymin,
- Weak and Strong Computational Creativity, Mohammad Majid al-Rifaie and Mark Bishop
- Why artificial intelligence is the future of growth?, accenture report,
- Why is artificial intelligence important? Accenture Report
- Wide Open Spaces: A statistical technique for measuring space creation in professional soccer, Javier Fernandez, Luke Bornn, 2018
- World Cup 2014 in the Twitter World: A big data analysis of sentiments
- World Sport Review, Nielsen Sport, 2016

SITOGRAPHY

- http://ampsy.com/
- http://behindthepicks.com/
- http://deloitte.wsj.com/cmo/2018/02/01/considering-cognitive-tech-look-beyond-digital/?lipi=urn%3Ali%3Apage%3Ad_flagship3_profile_view_base_recent_activity_details_all%3B5ujgKEeOQRasHRSoOoaazw%3D%3D
- http://epsi.eu/

- http://geosurfaces.com/
- http://iceberg.hockey/
- http://insytesports.com/
- http://journals.sagepub.com/doi/pdf/10.1177/1354856515587163
- http://matchanalysis.com/?lipi=urn%3Ali%3Apage%3Ad_flagship3_company%3B%2FxHixpZWRCeujRPcMOv6mA%3D%3D
- http://motusglobal.com/
- http://motusglobal.com/motusbaseball.html
- http://mrblairhughes.com/how-sports-teams-are-using-virtual-reality-to-enhance-the-fan-experience/
- http://optitrack.com/motion-capture-movement-sciences/
- http://searchcio.techtarget.com/definition/AI
- http://sportlogiq.com/en/
- http://sportsscientists.com/
- http://whoop.com/
- http://www.biometricupdate.com/201710/idemia-to-equip-levis-stadium-with-biometrics-based-technology-deploy-trusted-fan-program
- http://www.digitalreasoning.com/
- http://www.financialexpress.com/industry/technology/how-artificial-intelligence-can-predict-biological-age-based-on-smartphone-wearables-data/1116344/
- http://www.getdashtag.com/#dash
- http://www.headcasevr.com/?lipi=urn%3Ali%3Apage%3Ad_flagship3_pulse_read%3BfFjFV9KhRym6rKphhendzQ%3D%3D
- http://www.livelikevr.com/
- http://www.livelikevr.com/
- http://www.moxoff.com/portfolio/la-matematica-al-servizio-della-pallavolo-2
- http://www.optasportspro.com/products/scouting-reports-and-data-feeds/
- http://www.p3.md/
- http://www.reely.ai/
- http://www.reengineeringlabs.com/soccer-tactical-analysis-tech/
- http://www.solos-wearables.com/
- http://www.thermohuman.com/
- http://www.thevisionary.com/computer-vision/moneyball-3-0-how-visual-data-is-supercharging-sabermetrics-in-sports/?utm_campaign=The%20Visionary&utm_source=hs_email&utm_medium=email&utm_content=57358154&_hsenc=p2ANqtz-8FIvpKaGnByJG7-Mr3J07dzo7myAJtWS-UaxqxnppyCvE-32jQw_Mc_YDZ4i7cZeGVXJ9csFAkCgVDOUDdjehvUbL-7w&_hsmi=57398876
- http://www.tis.tv/video/how-to-use-babolat-play/
- http://www.tis.tv/video/mounir-zok-sport-technologist/

- http://www.vaynersports.com/
- http://www.virtexapps.com/virtex-arena/
- https://about.grabyo.com/
- https://aspirevc.com/
- https://austinstartups.com/sxsw-startups-dashtag-tracks-soccer-f7779d5042b5
- https://aws.amazon.com/it/machine-learning/what-is-ai/
- https://biztechmagazine.com/article/2016/07/stadiums-need-physical-and-digital-security-keep-players-and-fans-safe
- https://blog.starters.co/how-the-next-generation-of-sports-will-be-driven-by-machine-learning-53a0aa3dd812
- https://businessservices24.com/240769/sports-analytics-market-is-expected-to-grow-37-86-cagr-till-2025/
- https://channels.theinnovationenterprise.com/articles/analytics-and-crowd-control
- https://chyronhego.com/products/sports-tracking/tracab-optical-tracking/
- https://deadspin.com/just-what-the-hell-is-real-plus-minus-espns-new-nba-s-1560361469
- https://deepdreamgenerator.com/
- https://deeplearning4j.org/neuralnet-overview
- https://deepmind.com/research/alphago/
- https://dzone.com/articles/sports-industry-shine-through-ai-chatbots
- https://edjsports.com/
- https://gumgum.com/sports
- https://gumgum.com/sports
- https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence
- https://iq.intel.com/
- https://iq.intel.com/widgets-to-wickets-cricket-tech-immerses-fans-in-the-game/
- https://medium.com/@ben.bland/connecting-emotions-to-extreme-sports-next-gen-media-a5553a86da79
- https://medium.com/@imomisport.ai/imomi-sensor-artificial-intelligence-sport-225a40e622b2
- https://medium.com/@nigewillson/what-is-deep-learning-92956a5273aa
- https://medium.com/sportstechx
- https://newsroom.intel.com/press-kits/digitization-of-sports/
- https://nextgenstats.nfl.com/charts/list/all
- https://openai.com/research/
- https://orreco.com/
- https://piq.com/
- https://plato.stanford.edu/entries/chinese-room/
- https://pubs.acs.org/doi/abs/10.1021/acsnano.7b02182?src=recsys&journalCode=ancac3
- https://sensum.co/experiential-marketing

- https://sensum.co/releases/sync
- https://sensum.co/releases/the-hero-feeling
- https://spectrum.ieee.org/biomedical/devices/a-temporary-tattoo-that-senses-through-your-skin
- https://sportradar.com/
- https://sportthinking.it/
- https://strivr.com/360-demo
- https://strivr.com/sports/
- https://swisscognitive.ch/
- https://taqtile.com/
- https://thelocker.whoop.com/2018/03/21/for-pro-tennis-player-nicole-gibbs-balance-is-key-to-recovery/?lipi=urn%3Ali%3Apage%3Ad_flagship3_company%3BKGOFTYP2SxST9SWCGaT1GQ%3D%3D
- https://towardsdatascience.com/building-a-deep-neural-network-to-play-fifa-18-dce54d45e675
- https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464
- https://www.asensei.com/
- https://www.asmag.com/project/Vertical_Solution_Event_Venue/
- https://www.blocksixanalytics.com/
- https://www.catapultsports.com/blog/la-liga-basketball-connect-fans
- https://www.cbinsights.com/research-ai-100
- https://www.eff.org/ai/metrics
- https://www.eonreality.com/portfolio-items/immersive-3d-training-environment/
- https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#58c5ca14f5d8
- https://www.forbes.com/sites/gilpress/2018/02/07/the-brute-force-of-deep-blue-and-deep-learning/#13310b7849e3
- https://www.forbes.com/sites/gilpress/2018/02/07/the-brute-force-of-deep-blue-and-deep-learning/#7e4baac949e3
- https://www.forbes.com/sites/gilpress/2018/02/07/the-brute-force-of-deep-blue-and-deep-learning/#14ed5d0a49e3
- https://www.forbes.com/sites/kevinmurnane/2016/04/01/thirteen-companies-that-use-deep-learning-to-produce-actionable-results/#a55381233b84
- https://www.hykso.com/
- https://www.ibm.com/blogs/internet-of-things/watson-iot-sports/
- https://www.ibm.com/blogs/think/2016/02/ibms-watson-helping-sports-teams-make-talent-decisions/

- https://www.ibm.com/cognitive/de-de/outthink/stories/cognitive-sports/
- https://www.ibm.com/watson/stories/basketball-with-watson.html
- https://www.mc10inc.com/
- https://www.nature.com/articles/nature24270.epdf?author_access_token=VJXbVjaSHxFoctQQ4p2k4tRgN0jAjWel9jnR3ZoTv0PVW4gB86EEpGqTRDtpIz-2rmo8-KG06gqVobU5NSCFeHILHcVFUeMsbvwS-lxjqQGg98faovwjxeTUgZAUMnRQ
- https://www.nature.com/search?q=artificial+intelligence
- https://www.nbn23.com/en/
- https://www.quora.com/How-well-do-the-modern-deep-learning-algorithms-perform-compared-to-a-brute-force-algorithm-for-eg-Deep-Blue-supercomputer-in-chess
- https://www.reddit.com/r/MachineLearning/comments/4a7lc4/alphago_vs_deep_blue/
- https://www.satis.fi/
- https://www.si.com/mlb/2016/08/26/statcast-era-data-technology-statistics
- https://www.si.com/tech-media/2016/07/28/nasl-san-francisco-deltas-soccer-team-technology-virtual-reality
- https://www.sjearthquakes.com/post/2017/07/21/news-earthquakes-partner-second-spectrum-inc-artificial-intelligence-and-machine
- https://www.sportsilab.com/
- https://www.sporttechie.com/ampsy-helps-sports-teams-gain-insights-into-audiences-using-social-geofencing/
- https://www.sporttechie.com/artificial-intelligence-sports/
- https://www.sporttechie.com/grabyo-opta-sports-create-automated-highlights/
- https://www.sporttechie.com/microsofts-head-of-sports-discusses-influence-of-predictive-analytics-for-top-football-league-la-liga/
- https://www.sporttechie.com/super-bowl-lii-technology-patriots-eagles/?utm_campaign=The%20Visionary&utm_source=hs_email&utm_medium=email&utm_content=60415474&_hsenc=p2ANqtz-_0vARlDeRup4d-5s1uC12CNfC6-J2nRQrGddYM5svaClRjKxV9H4rSAyzitlDb7Y4Vf2wgOBvLlHqUEN2PWgN4lFqyrA&_hsmi=60415474
- https://www.theguardian.com/football/blog/2016/mar/29/san-francisco-deltas-nasl-soccer-team-virtual-reality
- https://www.upgraded-inc.com/
- https://www.vicon.com/motion-capture/biomechanics-and-sport
- https://www.wareable.com/sport/wearable-tech-at-rio-olympics-2016-2097
- https://www.wareable.com/sport/wearable-tech-injuries-us-pro-sports
- https://www.washingtonpost.com/sports/olympics/virtual-reality-helps-us-olympic-ski-team-prepare-for-pyeongchangs-great-unknown/2018/01/23/45085e20-ff73-11e7-bb03-722769454f82_story.html?utm_term=.3933da4e2003
- https://www.wearit.net/#home

- https://www.weeklystandard.com/be-afraid/article/9802
- https://www.wired.com/2012/09/deep-blue-computer-bug/
- https://www.wired.com/2012/09/usa-chess-texas-tech-computer/
- https://www.wired.com/2016/05/google-alpha-go-ai/
- https://www.wired.com/story/ai-beat-humans-at-reading-maybe-not/
- https://www.youtube.com/watch?v=fZrZWeXA9KU
- https://www-01.ibm.com/software/ebusiness/jstart/portfolio/usacycling.html

https://www2.deloitte.com/us/en/pages/consumer-business/topics/sports.html

ACKNOWLEDGMENTS

I would like to thank Professor Emanuele Lettieri, Claudio Rorato, Federico Iannella and the entire Osservatorio, for having given me the opportunity to work for my final thesis on this topic. It has been really interesting discover the most innovative solutions in sport and study their impact on the overall industry.

I would like also to thank Mr. Crivaro, Mr. Libroia, Mr. Pappalardo and Mr. Cintia for their availability demonstrated in the developing the case studies.

As these are the last words I am writing as student, I would like to use them to thank my family, my mum Elisabetta, my father Bruno and my sister Arianna, for having supported and helped me in these years even at cost of sacrifices. I hope that this work could worth as reward for their hard commitments.

A special though is dedicated to my grandparents, Nino and Chicca, for have always advised me wisely when I was in need, and for having always demonstrated their support. Their teachings, recommendations and advices are some of my most precious memories.

To conclude, the last heartfelt thanks is dedicated to my uncle Fabio, my aunt Graziella and my cousin Daniele, for have strongly believed in me since the beginning and for have constantly pushed me, providing all the support, help and the encouragements I needed.

I would like also to thank all my friends with whom I shared unforgettable experiences and great time in these years…Thank you all.

Milano, July 2018