



UNIVERSIDADE CATÓLICA PORTUGUESA

Finance and Artificial Intelligence

The Effect of AI on Ethics in the Financial Sector

Final Work in Academic Context presented to
Universidade Católica Portuguesa
In order to obtain the master's degree in Finance

by

Marta Kirsch Moreira Ferreira

Under the guidance of
(PhD) Luís Borges Gouveia
(PhD) Paulo Alves

Católica Porto Business School, Universidade Católica Portuguesa
April 2024

Acknowledgements

Gostaria de expressar o meu profundo agradecimento ao Professor Paulo Alves por orientar a minha dissertação num tema novo e complexo, e aos restantes professores que contribuíram para a minha formação académica.

Um agradecimento especial ao Professor Luís Gouveia por ter aceitado, sem hesitar, o desafio de me apoiar na orientação deste tema.

Obrigada, também, ao meu grupo de amigos, de longa data, em especial à Beatriz Cruz, Francisca, Joana Lopes, Joana Sá, Joana Soares, Nuno, Paula, Tiago, pela constante presença e apoio. Ao grupo das Tainadas, um obrigada por tornarem este processo não só de aprendizagem, mas também de bons convívios e grandes amizades. Agradeço ao Católica Porto Investment Club por todas as oportunidades e amigos que fiz no clube, em especial, à Inês, Gonçalo, Henrique e Mafalda pelo apoio contínuo.

O meu profundo agradecimento à minha família, que sem ela nada disto seria possível, em particular, ao meu pai, que me motivou ao longo destes meses e insistiu para escrever sobre o tema em questão, à minha mãe por me ter acompanhado e apoiado, à minha irmã por ter estado sempre presente em todos os momentos, bem como à minha tia e avó Rosalina, que me ajudaram sempre no que foi preciso.

Por fim, e não menos importante, um obrigada à minha avó Marianne, por me ter proporcionado ao longo destes anos diversas oportunidades.

Agradeço ainda a todos os outros familiares e amigos que contribuíram de alguma forma para este processo.

Resumo

A presente dissertação tem como objetivo investigar o impacto da implementação da Inteligência Artificial (IA) nas Finanças, com destaque para áreas que são suscetíveis de serem profundamente afetadas por esta tecnologia. Através do estudo das implicações da IA, é apresentada uma explicação sobre as mudanças e os avanços que já se encontram em curso na implementação destas ferramentas no setor financeiro.

Após a realização da revisão da literatura, tornou-se perceptível a presença de um gap na literatura. Com base da plataforma *Scopus* foi realizada uma pesquisa sobre a interdisciplinaridade entre as finanças, IA e a ética verificando-se a quase inexistência de trabalhos para análise. Para dar resposta a este *gap* foram investigadas quatro áreas relevantes nas finanças: *trading*, gestão de portfólios, detecção e prevenção de fraudes e gestão de risco, principalmente, na gestão de risco de crédito e as ocorrências e questões associadas com o uso e exploração da IA.

Seguidamente, procedeu-se à análise de desafios que a IA veio impor à indústria financeira tais como os custos de implementação dos sistemas e desafios de liquidez. Foram ainda considerados os desafios que realçam a multidisciplinariedade entre as finanças, a ética e a IA. Estes desafios levantam questões éticas, a título de exemplo, a potencial perda de postos de trabalho, a dificuldade de transparência e facilidade de explicação dos modelos de IA, implementados nas organizações, bem como a segurança da informação e privacidade dos dados utilizados pelos modelos de IA. A discussão realizada focou com maior detalhe a gestão de risco de crédito.

Os resultados revelam que os benefícios da IA são inúmeros, mas também acrescenta muitos desafios às instituições financeiras. Para que se garanta que a implementação destas ferramentas seja feita com sucesso, tanto no curto, como

no médio prazo, é crucial encontrar um equilíbrio entre a tecnologia, os reguladores e os financeiros. É essencial a aplicação de uma abordagem multidisciplinar que englobe as diversas áreas envolvidas no desenvolvimento, implementação, utilização e supervisão destes sistemas. Somente, desta forma, será possível manter a confiança, valor que está na base das finanças e dos mercados financeiros. Em especial, a oferta de ferramentas de IA associadas com às Finanças introduzem um potencial de desintermediação e de alteração das relações entre o cliente, as instituições financeiras e mesmo o papel dos especialistas, por via do uso e exploração da IA – o que acarreta riscos adicionais para o setor.

Palavras-chave: Inteligência Artificial, Finanças, Princípios éticos, riscos e desafios, estratégias para mitigar riscos, risco de crédito.

Abstract

This dissertation aims to investigate the impact of implementing Artificial Intelligence (AI) in Finance, with emphasis on areas that are likely to be profoundly affected by this technology. Through the study of the implications of AI, an explanation is presented about the changes and advances that are already underway in the implementation of these tools in the financial sector.

After carrying out the literature review, the presence of a gap in the literature became noticeable. Based on the *Scopus* platform, research was carried out on the interdisciplinarity between finance, AI and ethics, verifying that there was almost no work for analysis. To respond to this gap, four relevant areas in finance were investigated: trading; portfolio management; fraud detection and prevention; and risk management; mainly in credit risk management and the occurrences and issues associated with the use and exploitation of AI.

Next, we analysed the challenges that AI imposed on the financial industry, such as the costs of implementing systems and liquidity challenges. Such issues highlight the interdisciplinarity between finance, ethics and AI. These challenges raise ethical questions, for example, the potential loss of jobs, the difficulty of transparency and ease of explanation of AI models implemented in organizations, as well as information security and privacy of data used by AI models. The discussion held focused in greater detail on credit risk management.

The results reveal that the benefits of AI are numerous, but it also adds many challenges to financial institutions. To ensure that the implementation of these tools is successful, both in the short and medium term, it is crucial to find a balance between technology, regulators and finance. It is essential to apply a multidisciplinary approach that encompasses the different areas involved in the development, implementation, use and supervision of these systems. Only in this way will it be possible to maintain trust, a value that is the basis of finance and

financial markets. In particular, the offer of AI tools associated with Finance introduces a potential for disintermediation and change in relationships between the customer, financial institutions and even the role of specialists, through the use and exploitation of AI – which entails additional risks for the sector.

Keywords: Artificial Intelligence, AI models, Finance, Ethical standards, Practices, Risks and Challenges, and Mitigation Strategies, Credit Risk

Table of Contents

Acknowledgements	iii
Resumo	v
Abstract	viii
List of Figures	xiv
List of Tables	xvi
List of abbreviations	xvii
Chapter 1: Introduction	1
1.1. Motivation	2
1.2. Work Relevance	3
1.3. Research Question	3
Chapter 2: Literature Review	4
2.1. Proposed areas for the multidisciplinary study.....	4
2.1.1 On the definition of Finance.....	4
2.1.2 On the definition of AI.....	5
2.1.3 On the definition of ethics	7
2.2. Contextualization of AI and Ethics in Finance.....	9
Chapter 3: Methodology	12
3.1 Collection of Data	12
3.2 Research Design and Goals	13
Chapter 4: Applications of AI	15
4.1 The AI potential	16
4.2 Considered areas within Finance	18
4.2.1 Fraud prevention and detection.....	18
4.2.2 Risk management	19
4.3 Further discussion about AI and Risk Management.....	22
Chapter 5: Challenges and Ethical Concerns	24
5.1 Challenges.....	24

5.2. Ethical concerns	24
5.2.1 Ethics and AI	24
5.2.2 Job losses due to the adoption of AI	26
5.2.3 Transparency and explainability	29
5.2.4 Human control and alignment	32
Chapter 6: Conclusions.....	34
6.1. Discussion of results.....	34
6.2. Contributions	37
6.3. Future Work	37
References	39
Appendix.....	46
Appendix A. Trading	46
Appendix B. Portfolio management	50
Appendix C. Challenges	51
Appendix D. Security and safety issues	54
Appendix E. Selected papers	59
Appendix F: A selected list of Financial related AI Tools	64

List of Figures

Figure 1 – AI subsets.-----	7
Figure 2 – Examples of how AI is being used in financial market activities.	17
Figure 3 – CEOs anticipate reducing headcount by 5% or more in 2024 due to Gen AI -----	26
Figure 4 - Trading and AI Historical Evolution -----	46
Figure 5 - Taxonomy of responsible AI -----	56
Figure 6 – How a blockchain works. -----	57

List of Tables

Table 1 – Details of the research conducted on the Scopus database.....	13
---	----

List of abbreviations

- AI – Artificial Intelligence
- DeepLIFT – Deep Learning Important FeaTures
- DL – Deep Learning
- DNN – Deep neural networks
- DRL – Deep reinforcement learning
- GDPR – General Data Protection Regulation
- Gen AI – Generative Artificial Intelligence
- IMF – International Monetary Fund
- LIME – Local Interpretable Model-Agnostic Explanations
- TF – Trend Following
- WEF – World Economic Forum
- XAI – Explainable Artificial Intelligence

Chapter 1: Introduction

In the current technological age, the relationship between artificial intelligence (AI) and finance has proven itself to be extremely promising, opening the way for a new era of opportunities and challenges. Similar to the beginnings of computers, mobile phones, and televisions, owning one was seen as a luxury, identically, developing an AI system or purchasing one, it is also viewed as a luxury, although, nowadays, the use of this technology is becoming increasingly common. In the current environment, a wide range of these technologies are easily accessible to the public with little effort, as some are free, and others have small usage costs. A great illustration of this is ChatGPT (<https://openai.com/chatgpt>).

In reaction to the new technology used by other industries, including gaming, transportation, communications and healthcare, most of them will have to change and reinvent itself, turned out to be a historically shift. This adjustment can call to reconsider ethical principles or regulatory frameworks to better suit the markets and the consumer needs.

Not immune to this trend, the financial sector has shown a rise in innovation and positioned itself in line with new developments in the market. The technology is expected to yield significant benefits, but caution must be paid to the risks, like data security and privacy.

Humankind has a lot of power, and it is up to everyone to choose the best way to use it and how far they may carry it. An analogy could be made with the development of the atomic bomb, in 1945. After its creation, it became necessary to create specific regulation controlling its use, that which involved a careful consideration of the ethical implications within society (dual use). However, we find ourselves in the beginning phase of developing this discussion and fundamental regulations. First moves towards regulating AI use, in European

space, have already been taken by the European Union. On December 8, 2023, the EU reached the first agreement on AI regulation as part of its digital strategy.

My goal through this master thesis is to investigate the use of AI in finance. AI was the focus of much news at the time this topic was chosen. Still, I found a gap in the literature, as there was a limited exploration of ethical aspects and regulation of AI within the field of finance. This seemed like an interesting opportunity to think, critically, about these problems and go further into the challenges this technology presents, especially in the ethical sphere, in particular, considering to widen perspectives, mainly by evaluating the opportunities and risks of applying AI to finance.

1.1. Motivation

In recent years, the convergence of AI with the financial industry has created new opportunities for efficiency, decision-making and cost reduction. The motivation for writing this master's thesis comes from this recent rise of AI, in different industries, with exponential advancements in the last year. According to Schwab (2016), this is characterised as the fourth revolution, where the pace of innovation in AI is exponential and not linear anymore. Advances in this field have been noticeable and with a lot of potential. However, it becomes critical to scrutinize the ethical considerations that these steps are completed in a secure and transparent way, to maintain the credibility and trust, which is of critical importance for organizations, employees, and investors in their investment decisions.

This concerns serves as the foundations for this work that explores the relationship of AI and Finance, with an emphasis on ethical implications.

1.2. Work Relevance

This work is especially significant because it sheds lights on a quite underexplored intersection field as of current research time.

As financial institutions gradually are trying to adopt AI algorithms for areas ranging from risk management to portfolio optimization, the need to address the ethical consequences of these technologies grows. This thesis provides useful insights into emerging challenges like algorithm transparency, bias, regulation development and opportunities, such as cost efficiency, faster analysis of a large quantity of data and, with that, finally help industry players in developing informing strategies and responsible technological integrations.

1.3. Research Question

How does the implementation of AI models in risk management affect the ethical standards and practices of financial institutions, and what are the possible strategies to mitigate the associated risks and challenges?

This research question is focused on a specific aspect: credit risk management. It specifies the main variables of interest (AI models, ethical standards and practices, risks and challenges, and mitigation strategies). It is also relevant to current state of the field and the society, as AI models, that take advantage of Generative AI, are increasingly used in the financial sector and raise ethical and practical concerns. It is also researchable, as it can be answered by collecting and analysing primary and/or secondary data from various sources, such as literature, interviews, surveys, case studies, among other possible sources for gathering data.

Chapter 2: Literature Review

2.1. Proposed areas for the multidisciplinary study

In today's fast-paced research scene, the convergence of finance, AI, and ethics has emerged as critical area. This promotes a deep investigation of the complex relationship among these three areas.

2.1.1 On the definition of Finance

To obtain a better understanding of the remaining master thesis, it is necessary to contextualise the concept of Finance.

Finance is a field that covers the management, generation and the study of monetary resources and investments. It is based on the time value of money, establishing links with factors such as interest rates and future cash flows associated with projects. Usually, finance is divided into three distinct categories: public, corporate and personal finance.

According to the CFI¹, Corporate Finance, can be defined as a domain within Finance that focus on the development and changes connected to an organization's capital structure. Fundamentally, it concerns with varied approaches of creating and increasing value for a company, also based on issues of financing and cost of capital, with the main aim being the maximization of the company's value. Within the domain of corporate finance, three key activities can be identified: investment decisions, associated with future cash flows and return on invested capital, as well as, capital budgeting, capital financing, concerning with obtaining the necessary cash for investments, and dividends and return of capital, associated with shareholder remuneration decisions.

¹ CFI – [What is Corporate Finance? \(corporatefinanceinstitute.com\)](http://corporatefinanceinstitute.com)

Lastly, in the realm of personal finance, faced with unlimited necessities and scarce resources, it is imperative to ensure that each individual's needs are met optimally, a comprehensive planning and effective management are required. This includes aspects like pensions, savings, loans, among others.

2.1.2 On the definition of AI

There are several definitions of AI, (OECD, 2021) defines it as being machine-based systems. These systems have different degrees of independence, which use substantial amounts of data and sources, in other words, big data. They predict outcomes, offer recommendations, and make decisions for certain goals, defined by humans (OECD, 2021).

Russel and Norving (2009) list eight distinct definitions, which are categorized into four areas: thinking rationally, acting rationally, thinking humanely, and acting humanely.

The first step is to understand human thought process, which (Russel and Norving, 2009) refer to: as "Thinking Humanly". This process can be conducted by human brain observation, the author defines it as "brain imaging", or as we reflect on our own thoughts, this approach is known as "introspection". The last method involves conducting psychological experiments which involves the analysis of an individual performing an action. Following this, it is necessary to accurately simulate the human mind in a computer programme. Thus, going beyond simple behavioural imitation, the "cognitive modelling approach", enables a more sophisticated comprehension of human thought and reasoning.

The Turing Test is then introduced in "Acting Humanly" which is a try to mimic human behaviour and intelligence. Alan Turing developed this test in 1950, and this appears to remain a current test, when we are talking about AI. Also, the definition of intelligence needs to be clarified because it is highly ambiguous, in this perspective and is where the Turing test can be used, as a tool

for defining intelligence. When using this test, it is considered successful when the answers to the questions are so similar, making it difficult to determine if they were written by a human or a computer. After discussing human cognitive processes and mimicking human behaviour, the idea of AI is further separated into two categories: "Thinking rationally" and "The laws of thought approach". With this method, an attempt is made to mimic a rational thought process by using Aristotle's concept of logic (Russel and Norvig, 2009).

At last, "Acting rationally: the rational agent approach" is predicted on the idea of a rational agent, that acts in an uncertain situation, with the intention of reaching the best result. This method places a strong emphasis on creating intelligent systems that act to maximise the end goal (Russel and Norvig, 2009).

Numerous definitions of AI exist, yet it is equally crucial to establish its goals. According to Choithani et al. (2022), the goal of AI is to simulate machines that mimic human beings and show that they have enough ability to function and reason on their own, without the assistance of humans. The author highlights a very important point, the application of AI is making humans smarter.

Machine Learning (ML) models are also crucial to note in the financial industry, the algorithms in these models could be trained to act as a human and think as one (Choithani et al., 2022). Algorithms help develop client-centric strategies (Garg et al., 2023) and in the identification of strange transactions (Aleksandrova et al., 2023).

Finally, Deep Learning (DL) is a subset of ML (Ahmed et al., 2022) and can be thought as a symbolic representation of how human brain works (Choithani et al., 2022).

Generative Artificial Intelligence (Gen AI) is a subset of AI Deep Learning and these models are gaining prominence in finance, giving to finance processes efficiency, real-time market analysis, simplify processes like data entry and accuracy (Rane, 2023; Lee, 2020)

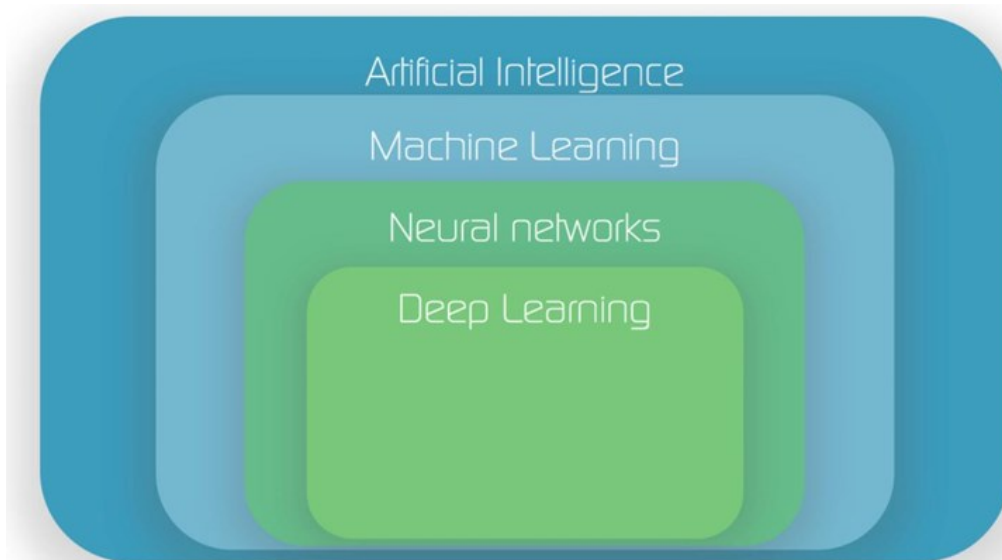


Figure 1 – AI subsets.
Source: OECD (2021)

2.1.3 On the definition of ethics

Ethics according to (Fisher, 2004) could be defined as *“an individual’s personal beliefs regarding what is right and wrong or good or bad”*. This paper also mentions a distinction between the ethics, in the context of organizations and in individuals. Regarding organizations, ethics is a concept that doesn’t exist since organizations themselves don’t think or have beliefs. However, those that drive these companies have beliefs that come from diverse backgrounds, which leads us to the concept of ethics, but in the perspective of individuals. Individuals are the ones who introduce the concept of ethics into companies in a wide range of sectors, being one of them the financial sector. The ethics, thus, has to do with the conduct of people within the organisation. In this way, organizations have responsibilities in terms of ethics like social responsibilities, as they must assess the impact of their actions are having on society.

Melé et al. (2017) defined ethics, as going beyond rules, requires a mutual and interdependent system of values, virtues and principles. The growing importance of ethics in financial aspects is also a key aspect, which becomes even

more pertinent, when even the International Monetary Fund (IMF) is concerned about these issues. Problems related to ethical issues have become more prominent since the financial crises, IMF stated.

The concept of ethics linked with the financial markets isn't one that creates a general agreement about its application. Some experts mention that what investors care about is profits and the maximization of them, and they do not care about the ethics behind these investments but if that's the case, we will have agents being characterised by looking only at themselves and the competition will be more aggressive (Dobson, 1997.). Melé et al. (2017) defend that according to the economic model, the firm's theory focus on the goal of maximizing value. This decision rule is promoted because, under certain assumptions, it leads to socially efficient decisions, but the rationalize behind the maximization of profits as a first goal will lead us to not have the best decisions in terms of ethics behaviour. However, on the other hand, there are other researchers that believe in the connection between the ethics and the financial sector and in the decisions that are made by the investors. These ones argue that there is a whole reality behind the single goal of maximising profits in which concepts such as moral values or the principles by which societies are governed must be considered because this are the core of a good financial practice and ethics. Therefore, the concept of ethics will come into play when analysing the different options on the market (San-Jose & Retolaza, 2018). Despite differing viewpoints, the author highlights a common concept: financial sector relies on trust. Economic crises underscore how a lack of confidence can have a negative impact for both the sector and consumer perceptions. The concept of trust is only applied if the ethic's concept is applied (Melé et al., 2017).

2.2. Contextualization of AI and Ethics in Finance

Given the magnitude and ongoing changes in the financial industry, it is imperative to make choices quickly and efficient to create value. AI appears to be an effective tool for meeting these goals. Although ethical concerns are not recent in the finance sector and have been used to do many regulations in the sector, however this is a field that are not much studied (Boatright, 2018).

According to Fong et al. (2021) the introduction of AI has led to value creation inside the industry and the banking sector expects to gain an additional trillion dollars from the application of AI.

As a result, several beneficial aspects of these programmes are emphasised, including the ability to analyse trends and patterns among an extensive amount of information. This data comes from a variety of sources, and it is characterized by its complexity. Real-time evaluation of complex data also drives the industry and its profits. Therefore, decisions made in real-time are more cost-efficient, quicker, and more accurate. Thus, applying AI becomes a way to gain competitive advantage over competitors in the market.

These efficient gains and cost reductions also make it possible to lower investment risks because of the improved accuracy in predicting prices, trends, and portfolio construction (Choithani et al., 2022). On the other hand, it also helps the industry to overcome the world-imposed challenges, these problems involve the need for dealing with financial fraud, addressing cybersecurity to support in global financial market fluctuations, and fulfilling the industry's rising need for sustainability and social responsibility (Al-Sartawi et al., 2022).

When addressing this convergence, it is crucial to highlight some key methods that enable all the advantages mentioned before, such as Neural Networks, that are applied for credit risk estimation, while Support Vector Machines (SVM) are employed for the analysis of consumer transactions (Königstorfer & Thalmann, 2020). Further, robo-advisors contribute to portfolio

management (Nguyen et al., 2023) by either reducing or moderating cognitive biases in decision-making (Hasan et al., 2023) or just use these agents to begin the application of AI, performing tasks that involves costumer interaction.

Regardless of what is stated in theory, it is necessary to understand whether it is feasible in practice and there are several examples demonstrating successful AI applications, in well-known international corporations, such as JP Morgan, Citigroup, and Goldman Sachs. These businesses already use AI to detect fraud, by looking for patterns in emails, for example. Based on a study published by Kamalnath et al. (2023) it has been applied in operations that were highly reliant on human capital. In these banks, the manual method resulted in higher expenses and longer task completion times.

Despite that, this convergence also involves some conflicts for companies, between the models and methods used by AI and, for instance, the strict regulations, which the transparency of information is a major requirement. This transparency implies that AI models, which by nature are considered back-boxes models, can't be used. Given the latest advancements in technology, explainable artificial intelligence (XAI) models have developed in response to the lack of transparency in AI. XAI approaches are strategies used to convert back-box models into easily interpretable ones, so that people can understand the processes involved and how the outcomes were derived. This is an important method especially in fields like risk management (Weber et al., 2023). Furthermore, the financial system has continuously depended on a core idea across time: trust. Applying AI might potentially undercut this idea, but since XAI depends on the explainability of judgements, it guarantees that applying AI is possible and ethical concerns may be solved. XAI methodologies including logistic/linear and visual explanation, could be used to get around these problems. Thus, the financial sector is once again positioned as one that values

transparency. Transparency is the basis for the fair, to see if markets are making unfair transactions, like insider trading.

Chapter 3: Methodology

This chapter describes the methodological approach used in the research for the writing of these master's thesis, providing a detailed overview of the steps taken for writing the report and the sources of information used to make the research. During data collection for writing the report, a gap in the literature was identified that work aims to contribute to its resolution. A qualitative research approach was taken to establish connections between AI, Finance, and Ethics, with focus on risk management and credit risk management.

3.1 Collection of Data

The collection of data was conducted to ensure that the information gathered was relevant to the topic under study.

Firstly, emerged the idea of connecting two fields: Finance and AI, which led to define a broad topic. However, after completing the literature review, it was clear that a third area may be added to the two previously considered: Ethics.

The next step was to make the revision of the literature and it was conducted in the *Scopus* platform (<https://www.scopus.com/>), it was chosen because of the diversity articles and because it is a reliable source of information. *Scopus* was used to select the main information for writing this report.

Given the relationship between AI and many other areas, it makes sense to filter results into two specific areas: "*Economics, Econometrics and Finance*" and "*Business, Management and Accounting*". After applying these filters, the research was conducted using expressions such as "*«Artificial Intelligence» AND «Finance»*" or "*«Artificial Intelligence» AND «Finance» AND «Ethics»*". The results can be analysed in Table 1.

Key words	Total number of articles found	Selected articles	Date
“Artificial” AND “Intelligence” AND “Finance”	634	16	14/10/2023
“Artificial intelligence” AND Finance	623	6	16/10/2023
“Artificial Intelligence” AND Finance AND Ethics	6	0	18/10/2023

Table 1 – Details of the research conducted on the Scopus database

Table 1 shows that this master’s literature review found on 22 certified papers. Nevertheless, as it can be seen there were no articles related to the three previously mentioned areas (within the *scopus* research base), which allowed to define the research question and identify the research gap.

Further to this data set, other sources of information were explored to make the revision of the literature on the identification of different types of AI models used to assess credit risk, including websites and organizational reports from renowned organisations such as OECD, PwC, Federal Reserve, KPMG, McKinsey, Fobes and IBM. From these sources, information was taken from articles, news, market studies, surveys and available statistics.

3.2 Research Design and Goals

In the beginning, as mentioned in section 3.1 a review of the literature was conducted to find the research question and the best articles to support the writing. The next step was to figure out which AI model would be most appropriate for risk management and, in particular, for its application to credit risk management. The methods chosen to address these issues were those related to Explainable Artificial Intelligence (XAI) models. These AI models would allow for a solution of ethical issues and the challenges presented by this technology.

Following that, research was carried out into ethical behaviours and effects, looking particularly at when the AI models are employed, how they affect

explainability, transparency, and the confidence, in other words, how the presence of AI models affects ethics.

The goal of this research is to find a relationship between the ethical behaviours in financial institutions and the importance of trust as a fundamental value shared by both industries. It also emphasises how imperative it is to maintain this value as AI is incorporated into the financial industry.

Furthermore, aims to assess the effectiveness of the previous chosen models in adhering to AI regulations. Financial organisations' ethical principles and norms were examined and contrasted using a qualitative comparing study. The research design chosen employs a qualitative methodology to comprehensively investigate the ethical implications of AI models inclusion in financial sector, especially in credit risk management. A qualitative method is based on natural language, in contrast to quantitative that is based on numbers and statistical models (Gerring, 2017). In addition, a qualitative method according to Busetto et al. (2020) is a method focused on comprehending the nature of phenomena, through contextual perspectives, and different manifestations in which they arise. Considering this, the proposed research question does not make sense to follow a quantitative approach, given that there are a very few credible quantitative data available.

Finally, although applying AI in risk management provides significant opportunities, it also requires a careful approach to mitigate risks and ensure that benefits exceed challenges, which are also other research goal of this master thesis. The search of a balance between innovation and responsibility is crucial for the success of this integration.

Chapter 4: Applications of AI

This chapter is focused in highlighting the function of AI in the financial industry. AI has revolutionized how financial organizations function, make decisions, and drive innovation. As Charles Dickens stated, *this is the best of times and the worst of times* (Dickens, 1859).

AI can be applied in a variety of sectors within the financial industry, including the buy-side of the market through asset allocation or stock selection, enabling the optimization of operational workflows, finding complicated correlations among vast amounts of data, and also helping in risk management (OECD, 2021). Applications supporting greater financial inclusion, decision-making support, real-time data analysis, high precision, and quick analysis of large volumes, are just a few examples, of how it can help customise services provided and adapt to the specific needs of each client. By guaranteeing security, these tools help professionals and investors to make more precise and informed decision.

As we have seen thus far, AI is here to stay, redefining paradigms and going beyond what's conventional – especially in an era where digital significance is noteworthy. The purpose of this chapter is to clarify two important financial industry activities, fraud prevention and detection and a more deep analysis of risk management. Trading techniques and portfolio management are also mentioned in Appendix A. Trading and Appendix B. Portfolio management

4.1 The AI potential

(Georgieva, 2024) stated that the potential of AI is to accelerate growth globally, while raise incomes all over the world.

AI has a tremendous potential, but this potential might be positive, but also negative. The IMF developed an index, named AI Preparedness Index, to assess if the country is ready to start implementing AI, evaluating parameters such as digital infrastructure, innovation, economic integration, regulations and ethics, and the conclusions are that wealthier economies tend to be better in these components than the poorer countries. Countries that performed better in this index are Singapore, United States and Denmark. Also, according to an article published by Rao et al. (2017), the world GDP has the potential to grow by up to 14% higher, by 2030, because of the effect of AI. Moreover, this study mentions that the application of AI in different sectors will be a game changer contributing up to \$15,7 trillion to the global economy.

Starting with of the main potential of AI, the increasing productivity, not only in the financial sector, but also in industries such as transportation or healthcare. Based on this study, areas such as personalised financial services, fraud detection, anti-money laundering and process automation are three areas of high advantages in implementing AI. Putting this potential in a line of time, in a medium-term perspective, it is expected that algorithms will understand what the preferences of the clients are and improve costumer recommendations based on their continuous answers and questions. In a long-term perspective, the potential of AI can go further, for example, predictions in credit risk assessments, based on a continuous evaluation of data, structured and unstructured, of client's information.

This promotes productivity in business, freeing up analyst's time for creativity. However, it is important to understand that these models are literal and designed

for particular phases of decision-making. Customers benefit with solutions more personalised in a less space of time, with less costs.

As Figure 2 illustrates, there is a wide range of applications for AI, from front-office to back-office work. This insight suggests that the impact on all professions will be different. Later, a balance will be made, to determine if this intensity tends considering ethical issues.

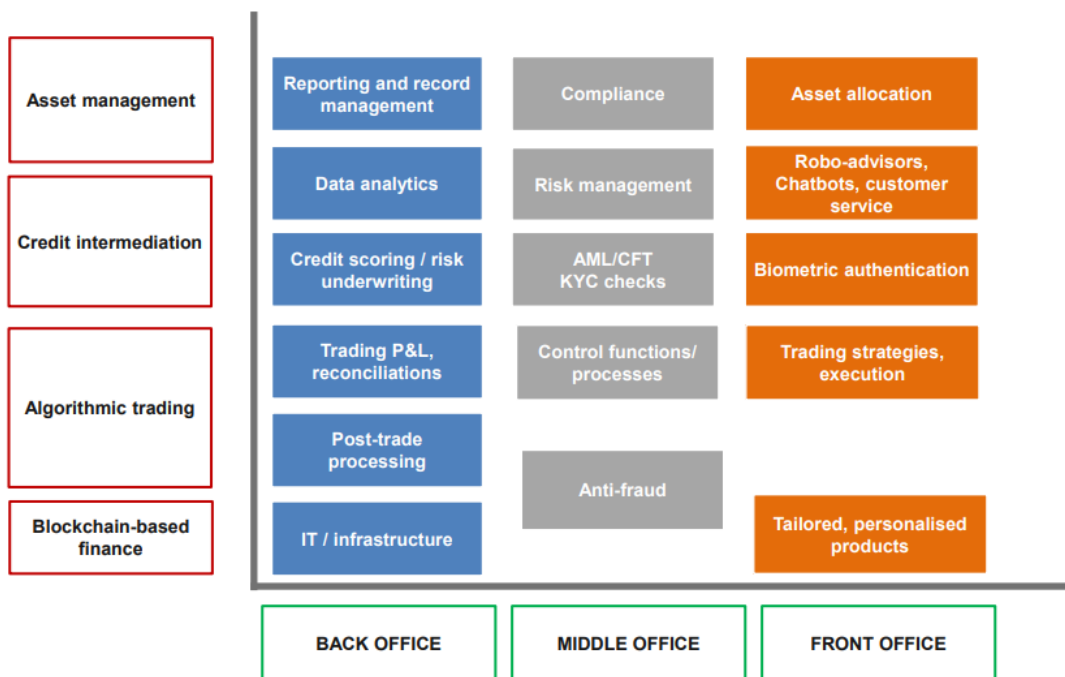


Figure 2 – Examples of how AI is being used in financial market activities.
Source: OCDE (2021)

Moreover, another powerful tool of AI is their DRL this could be a solution for the typical biases, cognitive and emotional, of human behaviour and potentialize the final outcomes.

To conclude, this chapter will analyse the potential of AI in two areas: fraud prevention and detection and risk management, another time, with more emphasis in risk management.

4.2 Considered areas within Finance

4.2.1 Fraud prevention and detection

Over time, extensive research has been conducted on fraud detection in the financial sector, and AI has emerged as a valuable tool in this field (Choithani et al., 2022). Detecting fraud means finding unauthorised transaction on accounts (Weber et al., 2023). With the increasing amount of data in circulation, the organisations in charge of these assessments are finding it increasingly difficult to properly analyse all the information accessible, in a real-time. In this context, AI has played a key role, frequently interacting with other systems, such as XAI to explain what are behind these models. XAI models enable greater transparency in AI-driven judgements, mitigating concerns such as discrimination.

The models act as analysts and alerts when there are strange transactions occurring, that could be fraudulent, like the use of client's information for activities that they are not allowed to. When talking about fraud prevention and detection, in finance the number of false positives is large, and need to be minimised, and this is where AI comes in, with their high potential of decreasing false alerts (OECD, 2021). An example of a fraudulent conduct is when there is an attack to our credit cards. If banks cancel the card, most of the times, data is still available and information these days is worth a lot of money. Behaviours such as this one occurs because of the lack of cyber-security. In these contexts, AI have also a strong role, giving alerts to the financial institutions, so that they know what is happening in a real-time and act in the moment of the threat. One method of alert is giving a score for the level of fraud detected, so that institutions can act according to the urgency of the problem (Lee, 2020). However, human supervisory of models are essential to understand if in the context that the machine is thinking, there are a fraudulent action or is it is actually crime

(Aleksandrova et al., 2023). Another time, transparency of the models is a must to understand if the systems is giving a good line of reasoning or not.

In other scenarios, AI systems are also worthy for following complex schemes of money laundering, through the analysis of structured and unstructured data (textual data, emails, ...). Methods such as Natural Language Processing (NLP), to search for non-linear relationships have been applied. In addition, analysis financial statements through methods of AI could also prevent fraudulent behaviours (Königstorfer & Thalmann, 2020).

Finally, the use of AI models in fraud prevention and detection have boosted the resilience of the financial system against more regular threats.

4.2.2 Risk management

Risk management, according to the CFA², involves recognising, measuring and create strategies to try to mitigate risks while ensuring that the risks taken fit with the desired goals and this process, generally, are based in models, which are attempts at representing reality.

In the field of financial risk management, particularly credit risk management and default prediction that entails forecasting the likelihood of borrowers failing to repay loans (Weber et al., 2023), the accuracy of the assessment is crucial in various circumstances, whether in bank loans or investments. The lower the perceived risk, the better for both financial institutions and investors. Moreover, the report (Königstorfer & Thalmann, 2020) states that AI systems are less prone to biases such as those connected to race and gender when compared to traditional methodologies such as logistic regression, underlining the benefits of more effective credit risk management. However, the conclusions of many studies are that is beneficial for the sector to combine AI methods, like neural

² CFA – [Measuring and Managing Market Risk \(cfainstitute.org\)](https://www.cfainstitute.org)

networks, with traditional systems for evaluating the credit risk (Goodell et al., 2021).

When the organizations like banks or regulators apply AI to risk management, there are some advantages for institutions such as its cost reduction (Svoboda, 2023) or has emerged as a tool that allows risks to be identified and goes further analysing these risks from the perspective of how severe they are for bank's balance sheets, for example and provides us with probabilities of these same events happening. This predictive analytical capability not only predicts adverse scenarios, but also prepares more efficient mitigation strategies, avoiding significant financial losses. Furthermore, maintaining a company's reputation, an intangible yet invaluable asset, is strongly protected (Božić, 2023).

In a globalised world, information reaches the market through a diverse set of channels, from videos and images to detailed texts, AI emerges as a tool for understanding and analysing enormous amounts of data with unprecedented efficiency. This unique ability to scan and interpret structured or unstructured information allows for a deeper understanding of financial markets, revealing hidden risks and complicated patterns that human eye, limited by time and processing capacity, may easily miss (Božić, 2023).

Examining a concrete examples of risk management in different types of risks that highlights the benefits of human and AI collaboration, credit risk management stands out as one of the risks that is important to consider.

Credit risk, according to the Federal Reserve³, refers to the possibility that an individual will not fulfil their financial obligations. Traditionally, credit risk has been assessed using logistic regressions, probit, and similar methods, although these techniques have provided a foundation for informed decision-making in the financial sector, and they have limitations as compared to AI models, which are considered more complete than the traditional ones. AI has

³ Federal Reserve – [The Fed – Supervisory Policy and Guidance Topics – Credit Risk Management \(federalreserve.gov\)](https://www.federalreserve.gov/monetarypolicy/supervisory-policy-and-guidance-topics-credit-risk-management)

the potential of improving the traditional models, providing a semantic understanding of unstructured data, representing a significant improvement in analysis and interpretation skills. This comes to help with the analysis of complex data and relationships that may not be obvious with traditional models, efficiency, and allows for a depth analysis, allowing the financial institutions to make a more informed decision. ML methods, like decision trees and SVM, are great methods to implemented in this field to help with the decision-making process, however there is a line that distinguished ML and AI, and according to (Aziz & Dowling, 2019) this has to do with automated decision-making process in lending. Moreover, with AI we can monitor in real-time the risks that the financial statements face, allowing for adjusts in risk management strategies.

As an example of the use of AI in this area of finance, the company UpStart⁴, uses AI models to the lending process. This company uses AI to try to prevent fraud and to make a more accurate, efficient, and inclusive lending.

The combination of human skills in interpreting subjective situations and AI's analytical ability in assessing objective factors, results in a more comprehensive and precise assessment of credit risk and in the estimation of the probability of default of the borrower, allowing the institutions to reduce their exposure to the credit risk. Concerning this topic, AI, also, helps in the context of credit scoring. Financial institutions, through institutions such as World Bank, which develop methods to evaluate the credits presented in each portfolio to be able to assess the risks to which they are exposed to and mitigate them.

Market risk is another one that needs to be considered. This risk is, according to CFA institute, related to markets that are volatile, with frequent fluctuations in stock prices, interest rates, exchange rates and commodity prices, leading to the emergence of this risk and the companies' exposure to the risk. Methods such as neural networks and deep learning algorithms show a significant ability to

⁴ Upstart – [Upstart Powered Loans: Personal, Car Refinance & Consolidation](#)

capture complicated market patterns in data, but to go a step further into AI it is necessary to implement the reinforcement learning (Aziz & Dowling, 2019). This is an effective approach for continuously assessing financial markets, especially in dynamic and stressful environments. Advanced AI algorithms can process large amounts of data in real-time, this enables a fast-decision-making process for potential risks caused by these fluctuations, the ability analyse and recognise patterns on information. The reinforcement learning (Aziz & Dowling, 2019) appears has a huge advantage because it provides the algorithms to make better decisions and prediction in the future. Similarly, to credit risk management, the introduction of unstructured data, such as texts, social media, provides to the market important warnings as soon as the information becomes available, otherwise it would be nearly impossible for humans to keep up with information, process it, and take action to mitigate the risks. The automations of this tasks show a strong support for humankind.

4.3 Further discussion about AI and Risk Management

AI has become a worth tool for the microprudential supervisors and not as suitable for macroprudential. The use of AI in macroprudential context faces significant challenges, firstly due to the uncertainty of future systemic risks and when it comes to systemic risk, the past is not always an indicator of future risks.

AI tools are beneficial for microprudential operations, focusing on individual financial institutions. Among many applications of AI, the ability to detect maturity adjustments in portfolios, evaluate the efficiency of liquidity guarantees, and identify similar trading patters amongst models stands out. Also, the distinction between exogenous and endogenous risks are important to address, being the first ones quantifiable, but the second ones are not easy to measure, when using AI algorithms, it's crucial to treat risks together, rather than independently, otherwise we could have an increment of all the risks and not a

decrease. Analysing and defining strategies, should be treated as a single optimization problem (Danielsson et al., 2017), this leads to an error that should be avoided – put the risks as an exogenous factor.

Although the exponential potential of AI in these models, considerations about the risks should not be ignore, particularly in the context of ethics. In the further chapters, these risks will be evaluated.

Chapter 5: Challenges and Ethical Concerns

5.1 Challenges

As previously mentioned, there are several challenges, such as implementation costs and liquidity challenges discussed in Appendix C. Challenges This chapter will only mention ethical concerns.

5.2. Ethical concerns

In recent years, with the fast technological growth, ethical issues arise as a big debate. Ethics play a key role in shaping the society and have significant impacts on the development of technology. This debate goes beyond individual industries, affecting our interconnected world. As technologies such as AI, big data, become incorporated into everyday life, worries about data privacy, model bias, and responsible use of this achievements have come to the forefront. Finding a balance between using the benefits of technology advancement and guaranteeing ethical governance is a world challenge.

With this, ethical questions arise and according to Cortella, M. S. (2020), there are three important questions in ethics for the implementation of new systems such as AI: *Do I want to? Can I do it? And Should I do it?* If all the answers were yes, then they should move forward with implementation.

5.2.1 Ethics and AI

Nowadays, there is a raise in enthusiasm when the subject of AI comes up, but on the date of selecting this topic, there was no paper published on the *Scopus* platform that combined AI, finance, and ethics.

The financial sector is characterized by the high level of trust that people place in it, as individuals commit their savings to banks and invest their money, all because of their effort and work to improve their quality of life, now and in the future. Thus, when we employ AI to help with decision-making process in areas such as credit approval (Aleksandrova et al., 2023), that it has a significant impact on people's lives and their overall well-being. So, thinking about this topic, conclusions are that AI didn't arrive to answer ethical concerns, instead, it came and created different challenges in this field. However, even these ethical concerns don't appear to be an obstacle to the adoption of AI, including in the financial sector, because like other technologies, it is helpful to the society. Nevertheless, unacceptable behaviours in society, have been solved and prevented throughout the times with regulation and this is a field that needs a lot of attention in the next years.

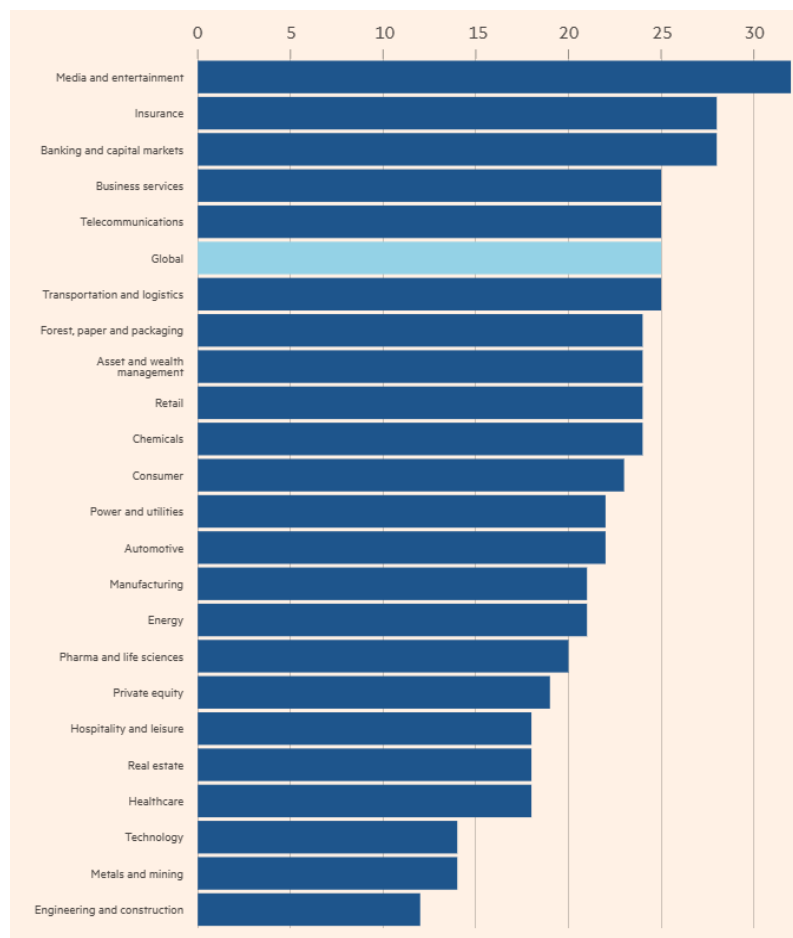
By implementing regulation in this area, we will see that problems such as inequalities and discrimination will decrease. AI systems should be considering the human rights, democratic values and diversity. They should contain appropriate protection to maintain the fair a just society and one way to do this, it is through human intervention in the systems, when it is considered necessary.

Thus, when we study the application of AI in the financial sector, this technology has phenomenal capabilities when used in the right way, such as creating opportunities for financial inclusion with the reduction of costs or giving the works time for a more creative thinking. However, it also has some limitations, and consumers, investors, and professionals, must understand how these systems work to mitigate these risks.

In this chapter, I will bring areas of relevance being the first one: the loss of jobs.

5.2.2 Job losses due to the adoption of AI

Society, including professionals in the financial sector, are beginning to realize that new technologies could replace many jobs and, in the case of AI, not only the less skilled ones which, according to a report published by (Ellingrud et al.,2023), these are the jobs that are most at risk, but they won't be the only ones, people in areas such as accounting and auditing have a higher chance of losing their jobs.



*Figure 3 – CEOs anticipate reducing headcount by 5% or more in 2024 due to Gen AI
Source: Financial Times*

According to **Figure 3**, a study carried out by PWC, (Fleming, 2024), presented in a Financial Times article, shows the forecasts of CEOs, who predict a reduction in headcount due to Gen AI. This figure stands out in professionals linked with the financial sector, such as “banking and Capital Markets”, slightly less, but still

significant, “Asset management”, and with lower percentage, around 17%, “Private Equity”.

Despite these values being significant for employability in the financial sector, the PWC study found that the biggest concern of these CEOs isn't the employability of the sector, nor in the laying off employees, but rather the risks that the implementation of AI will bring to these areas, such as cybersecurity risks or the spread of misinformation. Other important security and safety issues are mentioned in Appendix D. Security and safety issues

Firstly, one of the biggest debates in history, is the topic of jobs, being one these eras, the industrial revolution (Aleksandrova et al., 2023). As stated by Spool, (2023), we should begin by looking at history, which may help us think about the future. In the beginning, people believed industrial revolution would replace human labour, but as we have seen, this didn't happen or happen in a good way for workers and, we saw advances in efficiency, ending workers having more time to focus on activities that required human thought. The article also mentions a significant distinction: when AI is used in the financial sector, it doesn't remove jobs, but rather changes how some jobs are performed. A person could be focused on the client relationship management or in the opportunities of other innovations besides the AI, there will always be a need to check what is being done and ensure that it is delivering the results that the organization and the client want. Persons in the financial sector will always be needed, despite the transformation of some professions within the sector.

In addition, as mentioned, other jobs will be created with the use of AI, which would not make sense to exist otherwise, because the costs for costumers when they use these services would be very expensive. People to analyse if the models are obeying with the regulation, ensuring that there has no bias and discrimination in algorithms and defining the goals for the models are just a few examples of how humans continue to have a key role. For consumers, systems

like Robo-Advisors, could help them following their strategic goals in the market and make it possible to have constant advice, which would otherwise be extremely expensive for users.

The thinking behind the implementation of AI should always be that these are intelligent systems that come up to support human tasks and not “expert systems” that have been created to replace human labour (Russel and Norving, 2009). In the financial industry, especially in the areas of fraud prevention and detection, this becomes clear, AI systems are used to analyse data, find patterns and potential distortions in the information, but it is important that this information be checked by a human expert in the field, to guarantee that the detection is suitable for each organisation.

Not all the countries will respond to this innovation in the same way, although, according to the IMF, 40% of global employment will be faced with AI. However, more developed countries, with face more risks compared to developing countries but also, developed ones could have greater opportunities due to the implementation of AI. The income question also comes to the debate and could affect the gap between employees, the ones who can follow the development of this technology will see an increase in income, but the others will a decrease, augmenting the gap between classes.

Finally, the integration of AI in the financial industry not only affects traditional roles, but also creates new professions. While automation may result in the displacement of positions, it also generates a demand for employment focused on ethical issues and oversight. As the financial sector continues to evolve with technology advances, an equilibrium between AI-driven systems and human-centric ethical governance becomes crucial.

5.2.3 Transparency and explainability

A further significant ethical concern in the relationship between AI and finance, is the lack of transparency and explainability of AI's models. In this context, transparency means that the models should be possible to understand for regulators or stakeholders (Weber et al., 2023). Explainability, on the other hand, slightly above the concept of transparency, means that those who provide these services must be able to explain, for example, their clients need to understand, what models are providing and why they made certain decisions. The concept of transparency and explainability are connected to each other, given that, when talking about explainability, before, we need to have transparency of the models, if the analysts do not understand, cannot explain it to their clients. These concepts are very important when talking about finance and AI because creates trust in both sides, in the AI models and in the financial sector.

As already mentioned in other chapters, AI has the nature of back-boxes models, being mandatory to find a balance between the explainability and the model performance. As an illustration, in Germany, it isn't allowed to use models that are complete black-boxes – meaning that there is no possible or at least easy audit of how the system functions (OECD, 2021). These measures are needed in a highly regulated industry to avoid ethical issues, such as, discrimination and to make the fairest decisions for persons and organisations concerned. On the other hand, the comprehension and transparency of the AI's models aren't a sufficient condition to ensure that the model is a good one, other factors such as a robustness, that is, capable of dealing with unpredictable situations, data privacy, and most importantly, an open explanation of the risks underlying the models, which risks are they exposed to, must be taken into account as well.

These two factors have serious implications in terms of ethics, if they are not fulfilled. Insider trading is a good example of this, it consists in having private information, and using it to buy or sell stocks, making profits without any or very

low risk, however this makes some questions about the ethics behind it and opinions differ from investor to investor (Werhane, 1989). When it comes to transparency, if regulators don't know where the model is getting information, how they know if the information is public or private, the model is not transparent.

Moreover, in portfolio management in the lack of these factors AI models will take to the same, behaviours like deciding what are the assets that benefit some organizations, or allocating some capital in regions that the investors don't want to be exposed, the results are that the investor is incurring in unwanted risks, this is the case of political risks, or just some industries that the investor don't want to be exposed, despite of the profitability that the model could bring. Also, this could bring the portfolio to have returns lower than what they are expecting.

We should have transparent models, with transparent variables, however this is not easy to deal with. Financial institutions try to create their own models and use them to have competitive advantages, otherwise, the need to protect this advantage emphasises the reluctance to disclose extensive explanations of their models, as doing so may potentially risk their competitive advantage in the market.

As a result, we need to overcome this problem and here comes the use of different methods of XAI (Weber et al., 2023). The models and algorithms are black-boxes, but the first thing to this concept of XAI began to be studied a long time ago, almost when the topic of AI arrived, since the beginning of the research into AI, one of the concerns of scholars has always been to explain these methods (Xu et al., 2019).

XAI comes to overcome this problem of non-transparency and explainability of the models of AI. According to IBM⁵, XAI has 3 main techniques, prediction accuracy, traceability, and decision understanding. The first one, prediction

⁵ IBM – “What is explainable AI” – [What is explainable AI? | IBM](#)

accuracy, means that for a given model to be credible, it needs to be accurate in their predictions and to measure the accuracy of the model, a technique such as, *Local Interpretable Model-Agnostic Explanations* (LIME), make certain parts of the models transparent and interpretable. This creates trust in an algorithm, turning models into a white box. Afterwards, traceability, related to a method of Deep Learning Important Features (DeepLIFT), leads us another way to meet XAI, where is an attempt to limit the paths that these systems can follow. In the specific case of DeepLIFT, it is used to understand how decisions are made by deep neural networks and it shows how each neuron is interconnected with others. The last technique mentioned by IBM is decision understanding, mostly related to the human part of the systems and it is the part that people need to comprehend and be educated to trust these methods. Hence, the team assigned with managing these AI models must have the necessary training to fully understand the rationale behind and procedures guiding the decisions made with these models.

IBM highlights multiple advantages connected with each of those XAI approaches. They help in continuous model evaluation, allowing the detection and measurement of risks, while also evaluating the model's overall performance. In addition, when the model is constantly being monitored, we can always improve the results and consider the different risks presented over the time.

Additional benefits include the development of trust in the models, which is a fundamental element of the financial industry. These strategies improve transparency and explainability leading to an overall rise in confidence.

Furthermore, this simplifies the way for regulatory oversight. Greater transparency in models allows authorities to supervise whether they meet fundamental requirements and uncover any possible concern, such as discriminatory behaviour. This transparency not only helps supervisors work,

but aids consumers and analysts in making informed decisions, for example, in the context of credit approval, analysts understand why they could not approve a credit, but also, consumers comprehend the reason why it has not been approved. Moreover, the use of Knowledge graphs gives the analysts the necessary tools for understanding the superficial factors of the models and their conclusions, being easy to explain the results to the consumer (Saranya & Subhashini, 2023).

To conclude, XAI serves as a crucial tool in improving transparency and explainability within AI's complex models. By giving insights into the decision-making process of AI systems, XAI helps to develop trust and responsible use.

5.2.4 Human control and alignment

Considering human control this is essential for the good functioning of AI in the financial sector. In the financial context, maintaining the right balance between the human control and technology's potential is the key.

Firstly, analysts and investors, must take a progressive learning strategy and be open to learn more, when it comes to adopting AI technologies. The transition should happen one step at a time, with one example being the deployment of AI systems, such as Chatbots and Robo-Advisors, to help people adjust to the changes. These technologies can take over certain client engagements, but it's important to maintain a balance by including the human factor, ensuring an easier and smooth acceptance by clients and analysts and considering that some of these automated tasks are just one step for the outcome.

Moreover, ethics hold a significant role in the field of AI and Financial decisions. While AI models are designed to derive conclusions based on patterns and on the information available, they frequently lack the ability to recognise or account for the emotional components of decision-making. Despite efforts to minimise emotional biases, it is essential to recognise that there are exceptional

cases that should be examined. Humans play an important role in adding empathy, understanding, and ethical judgment to offset possible discrimination issues from an exclusively computational approach. In essence, a collaborative effort between humans and AI is required to manage the ethical challenges inherent in this dynamic sector. Financial industry, nowadays, still relies heavily on human interactions to maintain its credibility.

Although these systems provide numerous benefits to financial institutions, financial industry, nowadays, still relies heavily on trust and in a significant amount of human contact, which cannot be compromised.

Chapter 6: Conclusions

6.1. Discussion of results

In recent years, AI has become more integrated into the financial industry and in risk management, particularly in credit risk management. With this paradigm AI has the potential to completely transform how financial institutions evaluate, manage and monitor financial risks, but it also raises several ethical issues that deserve careful consideration, as was already demonstrated in this master thesis.

The possible influence on employment is one of the main ethical challenges connected to the use of AI models in finance (Tsang, 2023). Certain jobs may be lost because of automation brought driven by AI, particularly those that involve big analyses of data, time consuming analysis, meaning that with this 4.0 revolution, not only the less qualified professionals will suffer, but also the most qualified ones. However, it is essential to understand that this automation can also lead to new employment opportunities and motivate employees to requalify.

While human monitoring of operations driven by AI models is essential to guaranteeing ethical compliance and responsible decision-making, which will enable notable advances in productivity, professionals can profit from a shift to a more strategic and creative positions. Furthermore, to ensure customer trust and regulatory compliance, AI models must be transparent and explicable. The consumers should understand how decisions are made and what criteria the AI models consider. For this reason, it is important to use AI explainability techniques, like XAI, to give comprehensive insights into the inner workings of the models, as well as transparency principles while developing and deploying AI models. This promotes consumer trust and makes regulatory oversight and auditing easier.

A further significant concern in AI-based financial risk management is data protection and privacy. Strict cyber security and data protection procedures must

be put in place by financial institutions to ensure the availability, confidentiality, and integrity of customer data.

Additionally, to safeguard consumer rights and ensure the moral and lawful use of data, it is necessary to implement responsible data governance practices and comply with privacy laws, such as the *General Data Protection Regulation* (GDPR) in the EU, that in 2018, in a horizon way, started to regulate the use of personal data with AI.

Finding the right balance in the regulatory framework between fostering innovation and safeguarding the interests of consumers and financial stability in the key. To design laws that support ethical and responsible innovation, regulators should work with stakeholders in the financial sector, such as banks, financial institutions and technology companies. This can involve creating guidelines for AI systems in financial risk management that are clear, explicable, and accountable. It can also involve implementing supervision mechanisms to ensure compliance with standards, like periodical audits by independent authorities, establish clear ethical standards including, among other variables, model transparency, explainability, justice criteria and submitting periodic reports to supervision bodies.

By weighing up the benefits and challenges of incorporating AI, especially in credit risk management, a complex and dynamic picture of the future of the sector emerges. Although the implementation of AI may initially pose significant challenges, its presence is here to stay and simplify life in the financial field.

The analysis conducted for this dissertation indicates that, in the medium term, the financial landscape is going to experience a significant transformation, with the ethical challenges related to AI becoming progressively controlled.

AI brings numerous benefits, from making more accurate forecasts, optimised portfolio management, and personalised financial services (for example, through virtual assistants), reducing operational costs (Nguyen et al., 2023) and analysing

vast amounts of data in real time. Specially in credit risk management, the inclusion of unstructured information sources has the potential to raise the quality of default probability forecasts and expand the scope of inclusion of new consumer profiles. However, we cannot underestimate the ethical and moral challenges that arise when AI is used. The potential loss of jobs, the black box of AI models, and the extra risks for smaller financial institutions are all concerns that need to be considered carefully.

To overcome these issues, therefore, an ethical and responsible approach is necessary, with focus on transparency, and compliance with ethical regulations. Moreover, it is essential that those involved in the financial industry are dedicated to applying and making use of AI in an ethical and sustainable way. Only through safeguarding consumer interests and advancing the integrity and stability of the financial system, AI's benefits could be maximised.

Lastly, it is imperative to note that the field of AI is always changing due to continuous developments in technology and regulations as well as ongoing innovation in new approaches and methods. It is important to define that the time frame covered in this dissertation that is from March 2023, when the research was first proposed, to February 2024, when investigation into the gaps in the existing literature finished. The data and results presented in this study, however, may need to be updated and revised when new laws emerge, and new advancements are made, due to the dynamic and constantly evolving nature of the AI field. Thus, the results presented here must be interpreted within the specific chronological context that has been discussed, keeping in mind that the continuous evolution of the field may require revisions to the conclusions that have been made thus far.

6.2. Contributions

This master thesis focuses on a critical gap found in the already existing research existing in the *Scopus* platform. Through a comprehensive analysis on Scopus, a lack of articles addresses the convergence of three critical and diverse yet complementary domains: Finance, Artificial Intelligence and Ethics. This research tries to fill this gap, by highlighting the convergence of these areas and explaining their combined relevance to the financial sector. Additionally, by offering a study of the interactions between these areas, a list of selected papers were listed, with their contributions and reported restrictions in Appendix E. Selected papers Additionally, as a result research of the current offer in AI tools, a selected list of available AI tools is given in Appendix F: A selected list of Financial related AI Tools to illustrate the already existing use and possibilities that can impact the way both individuals, organizations and finance stakeholders can operate – thus providing material for further research.

6.3. Future Work

At the end this research, it is suggested multiple ways for deeper investigation into AI's impact on finance, reflecting the growing interest among researchers.

- **Regulatory analysis:** Studying how regulations affect the financial sector and how it impacts ethical AI innovation. Surveys for banking institutions could be a good way to do this type of research to see their perception of the laws that are in place.
- **Impacts on financial integration:** Research how the application of AI is influencing global financial integration. Evaluating whether AI promotes or delays integration and whether it makes a bigger or minor social class differences.

- **Long-term expectations and ethical challenges:** It is also critical to investigate the long-term prediction for AI's effects on finance to anticipate potential issues. Also, research in economic stability and address emerging ethical challenges.
- **Evaluation of additional risks:** Operational and legal risks resulting from the widespread implementation of AI should be assessed in addition to the risks typically connected with AI, such as algorithmic bias and data privacy. Developing models and techniques are essential to prevent the increase of these risks.

References

- Ahmed, S., Alshater, M. M., Ammari, A. El, & Hammami, H. (2022). Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance*, 61. <https://doi.org/10.1016/j.ribaf.2022.101646>
- Aleksandrova, A., Ninova, V., & Zhelev, Z. (2023). A Survey on AI Implementation in Finance, (Cyber) Insurance and Financial Controlling. In *Risks* (Vol. 11, Issue 5). MDPI. <https://doi.org/10.3390/risks11050091>
- Aziz, S., & Dowling, M. (2019). AI and machine learning for risk management.
- Boatright, J. R. (2011). Ethics in Finance. In *Finance Ethics: Critical Issues in Theory and Practice* (pp. 1–19). John Wiley and Sons. <https://doi.org/10.1002/9781118266298.ch1>
- Boatright, R., John. (2024). Finance Ethics: critical Issues in Theory and Practice. 4 de março. 10h27. Boatright, R., John. [Online]. Available at https://books.google.pt/books?hl=pt-PT&lr=&id=TC6m9kiHzGgC&oi=fnd&pg=PT16&dq=ethics+in+finance+authors&ots=QWzaVAoQ4W&sig=8NkbVXAkqJivQjge8mCDsE6agg&redir_esc=y#v=onepage&q=ethics%20in%20finance%20authors&f=false Consulted in [4/03/2024].
- Božić, V. (2023). The Role of Artificial Intelligence in Risk Management. <https://doi.org/10.13140/RG.2.2.29886.77126>
- Busetto, L., Wick, W., & Gumbinger, C. (2020). How to use and assess qualitative research methods. In *Neurological Research and Practice* (Vol. 2, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s42466-020-00059-z>
- CEPEJ. (2024). European Charter on the use of artificial intelligence (AI) in judicial systems and their environment – European Commission for the Efficiency of Justice. 12 de março. 17h10. CEPEJ. [Online]. Available at <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment> Consulted in [12/03/2024].

- Choithani, T., Chowdhury, A., Patel, S., Patel, P., Patel, D., & Shah, M. (2022). A Comprehensive Study of Artificial Intelligence and Cybersecurity on Bitcoin, Crypto Currency and Banking System. In *Annals of Data Science*. Springer Science and Business Media Deutschland GmbH.
<https://doi.org/10.1007/s40745-022-00433-5>
- Comissão Europeia. (2024). Proposta de Regulamento do Parlamento Europeu e do Conselho. 22 de março. 10h42. Comissão Europeia. [Online]. Available at <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206> Consulted in [22/03/2024].
- CFA Institute. (2024). Measuring and Managing Market Risk. 27 de fevereiro. 15h40. CFA Institute. [Online]. Available at <https://www.cfainstitute.org/en/membership/professional-development/refresher-readings/measuring-managing-market-risk> Consulted in [27/02/2024].
- Chen, X. Q., Ma, C. Q., Ren, Y. S., Lei, Y. T., Huynh, N. Q. A., & Narayan, S. (2023). Explainable artificial intelligence in finance: A bibliometric review. *Finance Research Letters*, 56.
<https://doi.org/10.1016/j.frl.2023.104145>
- Cisco. (2024). What is Cybersecurity. 3 de março. 13h17. Cisco. [Online]. Available at [What Is Cybersecurity? - Cisco](#). Consulted in [3/03/2024].
- Corporate finance Institute. (2023). Finance Definition. 18 de dezembro. 15h30. Corporate finance Institute. [Online]. Available at <https://corporatefinanceinstitute.com/resources/wealth-management/what-is-finance-definition/> Consulted in [18/12/2023].
- Corporate finance Institute. (2023). Personal Finance – Defined, Guide, Example, Template. 29 de dezembro. 13h00. Corporate finance Institute. [Online]. Available at <https://corporatefinanceinstitute.com/resources/wealth-management/personal-finance/> Consulted in [29/12/2023].
- Corporate finance Institute. (2023). Public Finance. 19 de dezembro. 10h20. Corporate finance Institute. [Online]. Available at <https://corporatefinanceinstitute.com/resources/economics/public-finance/> Consulted in [19/12/2023].

- Cortella, M. S. (2020). *Filosofia. E nós com isso?* Rio de Janeiro: Vozes.
- Danielsson, J., Macrae, R., & Uthemann, A. (2017). Artificial intelligence, financial risk management and systemic risk. www.riskresearch.org.
- Dickens, C. (1859). *A Tale of Two Cities*.
- Dobson, J. (1997). Ethics in Finance II. In Source: Financial Analysts Journal (Vol. 53, Issue 1).
- Expresso. (2024). Acordo na EU para primeira lei do mundo sobre inteligência artificial. 17 de janeiro. 13h20. Expresso. [Online]. Available at <https://expresso.pt/sociedade/tecnologia/2023-12-08-Acordo-na-UE-para-primeira-lei-do-mundo-sobre-inteligencia-artificial-d639d81a> Consulted in [17/01/2024].
- Federal reserve. (2024). The Fed- Supervisory Policy and Guidance Topics – Credit Risk Management. 26 de fevereiro. 16h00. Federal reserve. [Online]. Available at https://www.federalreserve.gov/supervisionreg/topics/credit_risk.htm Consulted in [26/02/2024].
- Financial Times. (2024). Generative artificial intelligence will lead to job cuts this year, CEOs say. 24 de janeiro. 13h30. Financial Times. [Online]. Available at <https://www.ft.com/content/908e5465-0bc4-4de5-89cd-8d5349645dda> Consulted in [24/01/2024].
- Fisher, J. (2004). *Social Responsibility and Ethics: Clarifying the Concepts*.
- Forbes. (2024). AI Hype Vs. AI Reality For Finance And Accounting. 16 de Janeiro. 21h56. Forbes. [Online]. Available at <https://www.forbes.com/sites/forbesfinancecouncil/2023/11/17/ai-hype-vs-ai-reality-for-finance-and-accounting/> Consulted in [16/01/2024].
- Garg, G., Shamshad, M., Gauhar, N., Tabash, M. I., Hamouri, B., & Daniel, L. N. (2023). A Bibliometric Analysis of Fintech Trends: An Empirical Investigation. In *International Journal of Financial Studies* (Vol. 11, Issue 2). MDPI. <https://doi.org/10.3390/ijfs11020079>
- Gerring, J. (2017). *Qualitative Methods*. <https://doi.org/10.1146/annurev-polisci-092415>

- Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. In *Journal of Behavioral and Experimental Finance* (Vol. 32). Elsevier B.V. <https://doi.org/10.1016/j.jbef.2021.100577>
- Hasan, Z., Vaz, D., Athota, V. S., Désiré, S. S. M., & Pereira, V. (2023). Can artificial intelligence (AI) manage behavioural biases among financial planners? *Journal of Global Information Management*, 31(2). <https://doi.org/10.4018/JGIM.321728>
- IBM. (2024). What is explainable AI?. 9 de fevereiro. 18h23. IBM. [Online]. Available at <https://www.ibm.com/topics/explainable-ai> Consulted in [9/02/2024].
- IMF. (2024). AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity. 2 de março. 15h14. IMF. [Online]. Available at <https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity> Consulted in [2/03/2024].
- Economics Observatory. (2024). How will artificial intelligence affect financial regulation?. 21 de fevereiro. 10h55. Economics Observatory. [Online]. Available at [How will artificial intelligence affect financial - Economics Observatory](https://www.economicsobservatory.com/2024/02/21/how-will-artificial-intelligence-affect-financial-regulation/) Consulted in [21/02/2024].
- Investopedia. (2023). What Does Finance Mean? Its History, Types, and Importance Explained. 18 de dezembro. 14h00. Investopedia. [Online]. Available at <https://www.investopedia.com/terms/f/finance.asp#toc-types-of-finance> Consulted in [18/12/2023].
- K Tsang, E. P. (2023). *AI for Finance*.
- Kong, H. T., Zhang, Q., & Yin, G. G. (2011). A trend-following strategy: Conditions for optimality. *Automatica*, 47(4), 661–667. <https://doi.org/10.1016/j.automatica.2011.01.039>
- Königstorfer, F., & Thalmann, S. (2020). Applications of Artificial Intelligence in commercial banks – A research agenda for behavioral finance. *Journal of Behavioral and Experimental Finance*, 27. <https://doi.org/10.1016/j.jbef.2020.100352>

- Kovvuri, V. R. R., Fu, H., Fan, X., & Seisenberger, M. (2023). Fund performance evaluation with explainable artificial intelligence. *Finance Research Letters*, 58. <https://doi.org/10.1016/j.frl.2023.104419>
- Lee, J. (2020). Access to Finance for Artificial Intelligence Regulation in the Financial Services Industry. *European Business Organization Law Review*, 21(4), 731–757. <https://doi.org/10.1007/s40804-020-00200-0>
- Luca, Kleinberg, Mullainathan (2019). *Inteligência Artificial, Análise de dados e a Nova Era das Máquinas*. 1st ed.
- McKinsey. (2024). Capturing the full value of generative AI in banking. 16 de janeiro. 18h43. McKinsey. [Online]. Available at <https://www.mckinsey.com/industries/financial-services/our-insights/capturing-the-full-value-of-generative-ai-in-banking> Consulted in [16/01/2024].
- McKinsey. (2024). Generative AI's impact on jobs and workflows. 24 de janeiro. 10h48. McKinsey. [Online]. Available at <https://www.mckinsey.com/mgi/our-research/generative-ai-how-will-it-affect-future-jobs-and-workflows> Consulted in [24/01/2024].
- McKinsey. (2024). Seven technologies shaping the future of fintech. 17 de janeiro. 17h21. McKinsey. [Online]. Available at <https://www.mckinsey.com/cn/our-insights/our-insights/seven-technologies-shaping-the-future-of-fintech> Consulted in [17/01/2024].
- Melé, D., Rosanas, J. M., & Fontrodona, J. (2017). Ethics in Finance and Accounting: Editorial Introduction. In *Journal of Business Ethics* (Vol. 140, Issue 4, pp. 609–613). Springer Netherlands. <https://doi.org/10.1007/s10551-016-3328-y>
- Moloi, T., & Marwala, T. (2020). Introduction to artificial intelligence in economics and finance theories. In *Advanced Information and Knowledge Processing* (pp. 1–12). Springer. https://doi.org/10.1007/978-3-030-42962-1_1
- Musleh Al-Sartawi, A. M. A., Hussainey, K., & Razzaque, A. (2022). The role of artificial intelligence in sustainable finance. In *Journal of Sustainable Finance and Investment*. Taylor and Francis Ltd. <https://doi.org/10.1080/20430795.2022.2057405>

- Nasdaq. (2024). AI Trading Software: Here's What You Need To Know. 22 de janeiro. 22h55. Nasdaq. [Online]. Available at <https://www.nasdaq.com/articles/ai-trading-software-heres-what-you-need-to-know> Consulted in [22/01/2024].
- Nguyen, D. K., Sermpinis, G., & Stasinakis, C. (2023). Big data, artificial intelligence and machine learning: A transformative symbiosis in favour of financial technology. *European Financial Management*, 29(2), 517–548. <https://doi.org/10.1111/eufm.12365>
- Observador. (2024). Inovação e setor financeiro: simplificar é a regra. 15 de janeiro. 11h22. Observador. [Online]. Available at <https://observador.pt/2023/11/29/inovacao-e-setor-financeiro-simplificar-e-a-regra/>, Consulted in [31/03/2024].
- Organisation for Economic Co-operation and Development. (2021). OECD business and finance outlook 2020 : sustainable and resilient finance.
- PwC. (2024). What is the prize- What is the real value of AI for your business and how can you capitalize. 1 de março. 14h19. PwC. [Online]. Available at <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf> Consulted in [1/03/2024].
- Rane, N. (2023). Role and Challenges of ChatGPT and Similar Generative Artificial Intelligence in Finance and Accounting. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4603206>
- Roberts, J., & Karras, J. (2019). What Is Blockchain? A PRIMER FOR WHY IT MATTERS FOR ECONOMIC DEVELOPMENT (Vol. 18). www.tipstrategies.com
- Russell, S. & Norvig, P. (2009). *Artificial Intelligence: A modern Approach*. 2nd ed. Pearson.
- San-Jose, L., & Retolaza, J. L. (2018). Ethics in Finance Research: Recommendations from an Academic Experts Delphi Panel. *Journal of Academic Ethics*, 16(1), 19–38. <https://doi.org/10.1007/s10805-017-9293-y>
- Saranya, A., & Subhashini, R. (2023). A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends. In *Decision Analytics Journal* (Vol. 7). Elsevier Inc. <https://doi.org/10.1016/j.dajour.2023.100230>

- Svoboda, Dr. A. (2023). The Impact of Artificial Intelligence on the Banking Industry. *Journal of Banking and Finance Management*, 4(1), 7–13. <https://doi.org/10.22259/2642-9144.0401002>
- Théate, T., & Ernst, D. (2021). An application of deep reinforcement learning to algorithmic trading. *Expert Systems with Applications*, 173. <https://doi.org/10.1016/j.eswa.2021.114632>
- The Economic Times. (2024). Trading Tools & Tech: Harnessing power of innovation for successful trading. 23 de janeiro. 12h32. The Economic Times. [Online]. Available a <https://economictimes.indiatimes.com/markets/stocks/news/trading-tools-tech-harnessing-power-of-innovation-for-successfultrading/articleshow/100727636.cms?from=mdr> Consulted in [23/01/2024].
- Upstart. (2024). Personal loans with low fixed rates. 26 de fevereiro. 15h50. Upstart. [Online]. Available at [Upstart Powered Loans: Personal, Car Refinance & Consolidation](#) Consulted in [26/02/2024].
- Weber, P., Carl, K. V., & Hinz, O. (2023). Applications of Explainable Artificial Intelligence in Finance—a systematic review of Finance, Information Systems, and Computer Science literature. *Management Review Quarterly*. <https://doi.org/10.1007/s11301-023-00320-0>
- Werhane, P. H. (1989). The Ethics of Insider Trading.
- World Economic Forum. (2024). The Fourth Industrial Revolution: what it means and how to respond. 2 de março. 8h36. World Economic Forum. [Online]. Available at [The Fourth Industrial Revolution: what it means and how to respond | World Economic Forum \(weforum.org\)](#) Consulted in [2/03/2024].
- Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11839 LNAI, 563–574. https://doi.org/10.1007/978-3-030-32236-6_51

Appendix

Appendix A. Trading

A reported released by OECD (2021), notes that significant progress has been made in the methods used in trading throughout time. Between 1987 and around 2007, there were three main ways to trade. First, there's a chance that the market may oscillate between two distinct stages: a bull market, in which investor "go long" and a bear market, in which investors "go short". This way, traders can choose contra trend or the trend-following (TF) strategy (Kong et al., 2011).

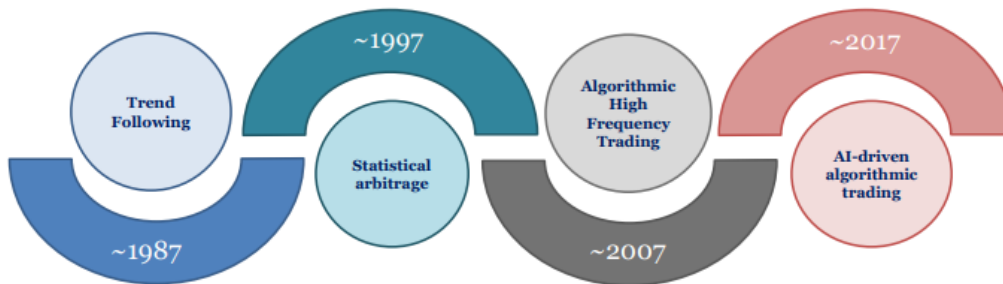


Figure 4 - Trading and AI Historical Evolution
Source: OECD (2021)

Then, statistical arbitrage came, and it can be characterized as highly automated approach that uses probability analysis to build trading strategies and carry out market operations. Stochastic algorithm-based algorithms are used to search for market inefficiencies, or situations in which asset prices do not reflect their true value, and then take advantage of these opportunities.

A different approach, around the year 2007, gained prominence in the financial markets, even though had started to be considered in the mid-1990s, was algorithmic high-frequency trading. This was a reaction to the financial sector's high degree of automation and technological innovation, together with the growing number of procedures being completed using computers.

From the year 2017 forward, AI-driven algorithmic trading strategies, have influenced the future of trading, providing solutions that give benefits including improved liquidity management and the capacity to carry out large orders with little effect on the market. This is accomplished for example, by optimizing size and length and quickly adjusting to market conditions. AI emerges as a technology that can make forecasts and suggests trading strategies without requiring human intervention (OECD, 2021).

NASDAQ⁶ states that two groups can be identified: beginner traders, who employ technology that is easy to use and just has a few basic features, and in the other way, active traders, who use more sophisticated algorithms to help analyse markets and carry out complex strategies. As Moloi & Marwala (2020) points out, another important idea in AI is that, at this point, doesn't provide intelligence, rather it provides an essential component of intelligence, which is prediction. So, when we use back-testing analysis, it lacks when it comes to a critical component: converting positive returns in real-time. This emphasises the advantages of real-time market analysis and prediction, as demonstrated by techniques such as walk-forward experiments.

In the sphere of trading, the blending of speed and certainty in decision-making emerges as another key point. A noteworthy example, as presented in the book (K Tsang, 2023), emphasizes this significance: a trader accurately predicts an asset's price increase. However, neglecting to take immediately act on this insight, allows other investors to enter the market, driving up the asset's price.

One of the simplest versions, without the use of AI, is algorithmic trading and in this method, conditions and strategies are established and then when these criteria are met, the program acts on the trader's behalf, taking advantage of the computer's speed (K Tsang, 2023). However, the increasing prominence of this

⁶ NASDAQ – [AI Trading Software: Here's What You Need To Know | Nasdaq](#)

strategy in the field may have a significant impact on investors' earnings and risk management of their assets. This is because, ML systems can learn from historical data, which could impact the future investment decisions. This is useful in trading decisions such as whether an investor should buy, sell, or just keep a certain asset. Also, allows for the development of strategies as well as the identification of more relevant chances for traders.

Over the years, a wide range of research has been carried out in an effort to identify the factors and information sources that are most useful for forecasting the values and patterns of financial assets. Understanding how information will affect asset prices is crucial, and this can be done by using NLP (natural language processing) algorithms to include new data sources, like tweets and publications from social media.

In a globalized world, it is unavoidable that assets are interrelated, posing greater risks for investors in these assets, but also carrying out new opportunities.

The introduction of AI acts as a tool to support risk management by establishing loss limits and implementing hedging strategies to reduce the impact of market volatility. It also improves forecasting skills, as already mentioned, which have been shown a high level of accuracy in predicting complex trends and anticipating market movements. This is due to their ability to analyse massive amounts of data, like textual data, historical prices, trading volumes and understand it and also the continuous monitoring of the market, as noted by NASDAQ. An example of this, as stated in (OECD, 2021) is risk management through the analysis performed by an AI system, evaluating whether a trader should keep or close a position depending on their future objectives. Despite frequent volatility in financial markets, these AI systems are unaffected, they show the potential to learn autonomously, decreasing the requirement for human interaction and periodic training, which is often necessary with alternative approaches. This advantage of continuous learning

has proven to be one of the techniques' key distinguishing features when compared to other systematic trading systems that rely on human interaction. In this circumstance, time can potentially result in significant financial costs reductions, and with these methods there is an increase in efficiency, because the algorithms are able to adapt independent of the situations, to the changing of the market without the need of human intervention. Moreover, it reduces the human biases, most of the times, allowing better decisions to be made, and, lastly, the availability of this systems is twenty for hours.

This is only possible due to systems such as deep learning (DL) methods, mentioned by Théate & Ernst (2021) as a technique that is based on deep neural networks, models that try to simulate the human brain through mathematical models. Also, this paper mentioned another technique, deep reinforcement learning (DRL), that consists in a set of approaches focus on the learning process of an intelligent agent. In this approach, the agent interacts sequentially with an unknown environment, working to enhance its cumulative rewards, using methods to expand knowledge obtained through interaction with the environment.

The NADASQ article discusses other advantage of AI in trading, customisation, which means that each investor has their own aims and specifies their own level of risk, ending up using different methods with tailored trading strategies and alerts for each client.

Briefly, AI has advanced significantly, over the years, bringing to the financial market, algorithmic trading strategies, taking advantage of machine learning methods, risk management, the optimization of the investments, predictive analysis, all of this in a real-time analysis of the data, decreasing the likelihood of market opportunities going unnoticed. The ability to learn twenty-four hours a day, without human interaction, makes the application of AI in trading increasingly powerful.

Appendix B. Portfolio management

Regarding portfolio management, over time there have been many advances that have multiple goals, ranging from lowering personal costs to simply trying to improve the consumer experience and speed up the processing of new information. One application of these developments is the use of robo-advisors to assist with investment decisions and manage a portfolio (Nguyen et al., 2023).

Also, in this context, it is useful to use AI for simulations of different future scenarios, it is a much quicker way to simulate returns or risks when the portfolio is exposed to different market conditions (OECD, 2021). The OECD report mentions the benefit of AI to analyse large amounts of data in shorter and real times, from different sources to see if the strategy that have been followed is the best one or if the model could suggest another one more profitable.

One of the challenges faced by investors, when they're trying to optimize their portfolios or making investment decisions are cognitive biases. These biases, according to Hasan et al. (2023), are described as emotions and feelings that make rational financial decision-making complex, since it is difficult to distinguish the rational from the emotional. Furthermore, due to the complexity and non-linearity of evaluating the performance of equity funds, it was required to explore options for this evaluation using ML models (Kovvuri et al., 2023). The authors highlight the advantages of these models over traditional methods, since they take advantage of non-linearity, for example, through the Shapley Additive Explanations (SHAP) model, an Explainable AI (XAI) model based on Shapley values.

Appendix C. Challenges

As it has been presented throughout this report, AI brings many challenges for the world. According to (OECD,2021), AI should bring advantages for people and drive inclusive growth, sustainable development and well-being.

The expression “garbage in, garbage out” emphasizes the importance of input quality in deciding the outcome. In the context of AI, using insufficient or untrustworthy data sources may compromise the results of the model and in the financial world information is a key factor (OECD, 2021). In 2018, the European Commission, started to develop a European Ethical Charter⁷ on the use of AI in judicial systems and it was based on five fundamental principles that lead to the mentioned challenges:

- Principle of respect fundamental rights, ensuring that when AI systems are applied, they do not violate the human fundamental rights.
- Principle of non-discrimination, AI systems should not have biases like gender or race preferences.
- Principle of quality and security, where the data used in AI systems should be secure and from reliable sources. A multi-disciplinary management is recommended.
- Principle of transparency, impartiality and fairness. AI models should consider these three factors.
- Principle of “under user control”, meaning that the user is making informed decisions and know how to control them.

Although these are mainly for judicial systems, they are a good starting point for the challenges presented by the implementation of AI systems in finance and the ethical considerations that arise from that.

⁷ [CEPEJ European Ethical Charter on the use of artificial intelligence \(AI\) in judicial systems and their environment - European Commission for the Efficiency of Justice \(CEPEJ\) \(coe.int\)](#)

Costs of Implementation

Developing and deploying AI systems requires significant financial costs of implementation, generating entry barriers. Businesses with less financial resources may find it too expensive to acquire cutting-edge technology, hire skilled workers and maintain the infrastructures.

There are several huge players, such as investment funds (the ones leading the implementation of AI models), that are capable to support this entry costs, but the small institutions have lots of difficult to compete with them, in the implementation of this systems. This large funds, companies and banks, are increasing their competitive advantage and creating a barrier for the entry of small players, limiting competition (OECD, 2021). The ability of large funds to invest extensively in AI research and development widens the gap, since they can benefit from economies of scale and attract top people to drive their innovation. While huge organisations invest a lot of money in technological progress, they also raise issues about diversity and inclusion in the market. Smaller companies, startups, and institutions in emerging regions may struggle to stay up, leading to a potential digital gap.

Addressing these cost-related entry barriers is essential for building a more inclusive and competitive AI ecosystem, ensuring that the benefits of AI innovation are available to wider range of businesses, with the help of supervisors that need to promote a fair competition in the markets.

Liquidity Challenges

Market liquidity risks should not be ignored. It refers to how easy, quickly and efficiently, is to buy or sell a position, without having significant barriers such as costs or risks.

The search for developed AI models offers benefits such as cost reduction for small companies, but also presents a liquidity and volatility challenge.

Financial institutions which do not develop their own models and just buy them from another entity, there is a risk of a one-way market, increasing the liquidity risks. As these models will respond in a very similar way for signals. A big number of participants in the market executing the same strategy at the same time could make a liquidity pressure in the market. In these cases, if there are big transactions executed at the same time, also the volatility of the market will increase and at a time of more stress, this could increase risks (OECD, 2021).

The intervention of regulators in this is essential for the minimization of risks, like as mentioned before, the models being sold must meet the minimal transparency criteria, which should be defined by regulators. Financial institutions and supervisors should also analyse problem of bias. Organisations can buy already made models, but they must modify them to their goals and ensure that they understand what they are buying.

Appendix D. Security and safety issues

AI is bringing in a new era in finance, with implication in multiple fields. It is capable to give lots of benefits but at the same time many challenges arise in the regulatory environment. Due to the black-box nature of AI algorithms, it turns difficult to comprehend how it works, arising some problems of security, data privacy and cybersecurity. These factors create a barrier to the implementation of this systems in financial institutions, which have strict requirements to fulfil. When using AI for personalised product suggestions, it is necessary to save consumer data over time. However, given the confidential nature of this information, there is a risk that it will be revealed to the public, if a cyber-attack occurs, for example.

As an illustration, in the context of banks, customers will provide confidential information to financial institutions. As time goes by, systems store this data about client actions, such as unemployment, payment delays, preferences for investing in some companies or sectors, the risk profile of the investor, helping to improve the models and decisions for future clients with less information available.

As mentioned by Sekhar et al. (2023) in the context of credit risk decisions, financial institutions should guarantee that the decisions they are talking are not biased and the models used are transparent. There is a very residual regulation in this context and the ones that have been developed are very uncertain. In this report, an alert is issued for regulators to collaborate with management teams and study the implication of this implementation in transparency, bias and ethics concerns.

Many regulators and supervisors as well as analysts lack the knowledge and skills required to efficiently manage this new technology. To ensure a harmonious interaction between finance and AI, well-defined rules with clear principles are required. This ensures that both the technological sector and the

financial industry should operate together to define the better and clear regulations.

According to Quinn et al. (2023) this regulation should involve organisations such as banks, brokerages and investment funds to prevent financial crimes, giving that it has a crucial role in promoting economic stability, protecting financial institutions and their consumers from unethical behaviour. Otherwise, if there are not appropriate regulations, the probability of financial crises will increase and the credibility in the sector will decrease, becoming a concern.

Cybersecurity⁸ is the act of protecting systems, networks, and programs from digital threats. These attacks usually involve the unauthorised access, modification or destruction of critical information, extorting money or just disrupt the usual functioning of organisations. These attacks are not new and not only came with the use of AI, but companies committee also entire departments to addressing these challenges, sometimes due to the high complexity of these issues, organisations often delay the use of these new technologies. When analysing the role of AI in this field, we found two opposite positions that, if correctly used can become complementary. Firstly, the idea is that AI has the capability of analysing large amounts of data and can track data daily, being more informed when decisions are made, however, this motivates the use of AI to check if companies are using and obeying to the regulation. On the other hand, there is a chance that these systems will make mistakes and be misled in their analysis.

As mentioned before, the XAI is, another time, a fundamental tool for overcoming these problems, according to the Economics Observatory. Also, the training of regulators in the understanding of AI has mentioned as a needed factor, helping them testing the outputs of the models.

⁸ [What Is Cybersecurity? – Cisco](#)

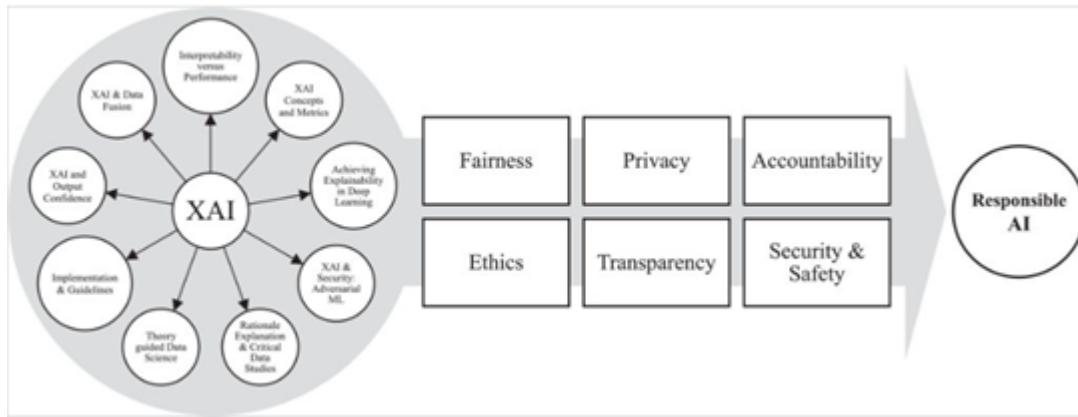


Figure 5 - Taxonomy of responsible AI

Source: How will artificial intelligence affect financial regulation?

As XAI could try to help in this problem, also the use of technologies such as antivirus, firewall, *security incident and event management* (SIEM), data loss protection and introduction detection system, are presented as technologies that would help financial institutions to prevent from online attacks (Choithani et al., 2022).

The Blockchain, associated in the finance world to the Bitcoin, is another alternative technology characterised by fragmentation, tenacity, confidentiality and testability (Choithani et al., 2022) and comes to guarantee data security, most of the times known as a distributed ledger, that is a system programmed to protect against hackers, prohibit unauthorized changes to records, reduce the risk of fraud and the best part is that all of this occur without an intermediary. This is difficult to change and hack because the information is distributed by blocks and not in just one server. All the information is stored and there is no possibility to change the transactions that have been done before and they are available for all in this network, it is a chain. However, despite having multiple blocks, they are linked to each other, in a chain (Roberts & Karras, 2019). The way blockchain works is presented in a simple way in the **Figure 6**.

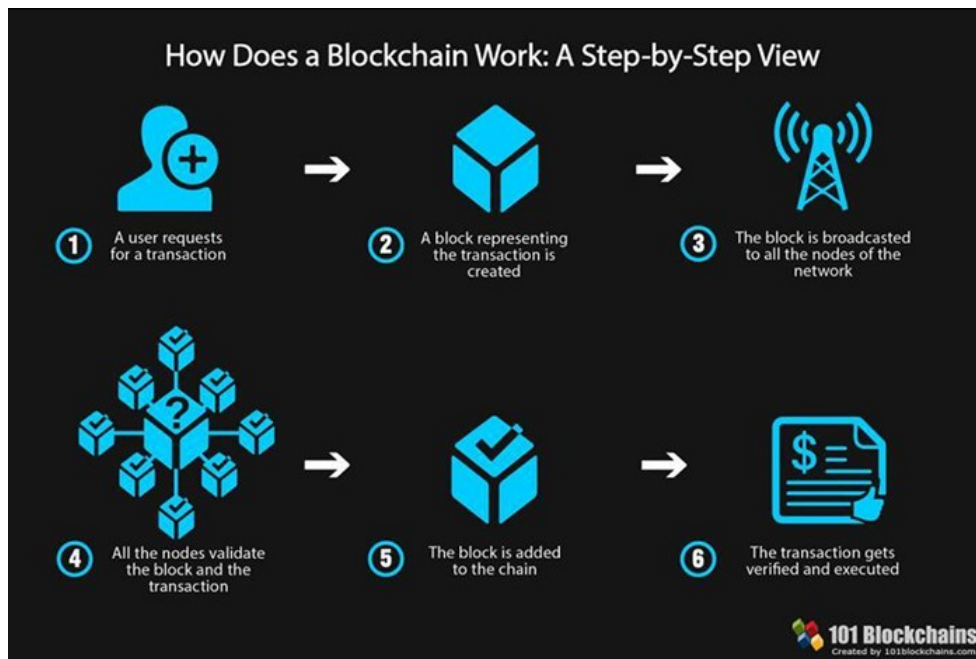


Figure 6 – How a blockchain works.

Source: A Comprehensive Study of Artificial Intelligence and cybersecurity on Bitcoin, Crypto Currency and Banking System

Despite its significance in the financial industry, there are numerous challenges that limit its implementation. For example, the vulnerability of the system in cases of selfish mining, where a small group of miners can deceive the system even if they don't have much power. Another problem is privacy leakage, meaning that the users don't need to reveal their identity, it is an option to prove the identity, transforming this technology in a challenge to the financial sector, where all the information in these contexts should be clearly expressed.

In the pursuit of ethical and fair technology, as already mentioned in other chapters, for the safety of the models it is also crucial to ensure that models not only avoid discrimination but also provide inclusive security for those in less favourable circumstances, such as lower incomes or coming from developing countries. It is necessary not to rule out complex situations, as one of the limitations of these models is the lack of sensitivity to specific details. By developing an approach that values equity and recognises the diversity of

experiences, we can aspire to a more responsible and ethical use of AI promoting more just and ethical societies for all.

This is where regulation appear, technology is transforming the financial sector, but there is an urgent need to handle the difficulties that arise, ensuring that regulatory frameworks promote both innovation and consumer safety. In the banking sector, financial institutions that use AI models could increase their minimum requirements of capital to mitigate the risks like systematic ones applying these methods the models should have a kill switch, that is a method of switch of models when there a period of stress, regulators obey the financial institutions to switch off their models and humans need to intervene and process the transactions without the use of these technologies. Although, this could mitigate many risks, it could also increase the operational risks (OECD, 2021). Models should be tested in situation of stress and a suggestion made by OECD is the creation of a supervisor only for testing models and approve them to be used by financial institutions.

In conclusion, the first step towards making these systems more secure and reliable is recognising their limitations as well as the benefits (Luca et al., 2019).

Appendix E. Selected papers

The following table presents 18 selected papers for which were extracted their contributions and the constraints, concerning the researched topics of AI and their potential impact in Finances. The papers were published between 2020 and 2023.

Author	Contributions	Constraints
(Chen, Ma, Ren, Lei, Huynh, & Narayan, 2023)	Applying a bibliometric approach, the paper presents the notion of <i>Explainable Artificial Intelligence (XAI)</i> , a branch of AI. XAI aims to improve the transparency and comprehensibility, mitigating risks, and fostering ethical business practices of AI models and algorithms for human comprehension. By identifying important subjects, significant writers, and research gaps, this makes it possible to evaluate the state of the field's research in both quantitative and qualitative terms.	The bibliometric analysis that this paper performs is one of its limitations. Although it enthralls the basic patterns, such as the number of publications and citations, it needs to be enhanced with a qualitative approach of the themes discussed.
(Aleksandrova, Ninova, & Zhelev, 2023)	Based on the application of AI in the financial system, in sectors such as insurance and financial control, the article discusses the advantages of the application, such as speeding up the development of procedures and services, as well as the drawbacks of using AI in these areas, which raises ethical concerns. For possible future research, the paper explores multiple areas of interdisciplinarity between finance and AI.	This article uses a "theoretical desk study" approach and is based on a small number of academic research articles. Given the topic's interdisciplinary nature, the paper is unable to handle each problematic area and it is impossible to foresee what topics could be covered in the future.
(Nguyen, Sermpinis, & Stasinakis, 2023)	The paper is a valuable resource in comprehending the complex relationship between big data, ML, and AI in the financial sector. It explains how these three fundamental parts of data science	s/d

	<p>are actively shaping the direction of fintech and contributing to the continuing revolution in this field. The article distinguishes by basing its research on academic, professional surveys, and regulatory perspectives. Additionally, it delves into areas of AI application that are still being developed, such as the adoption of robo-advisors for portfolio management.</p>	
(Hasan, Vaz, Athota, Désiré, & Pereira, 2023)	<p>This article delves deeper into the concept of behavioural biases and focus on how AI may help to prevent such distortions. It delves into topics like robo-advisors and their implications in terms of avoiding behavioural biases in the financial system.</p>	s/d
(Li, Sigov, Ratkin, Ivanov, & Ling Li, 2023)	<p>The authors of the article perform a comprehensive review of many ML techniques, including Support Vector Machines (SVM) and K-nearest neighbors (KNN). In addition, the concept of deep learning, a sort of ML that is gaining renown, is introduced. The numerous uses of AI in financial industry, as well as the key challenges and trends related with these applications, are thoroughly examined.</p>	<p>Focus on specific themes and techniques within AI, because of the careful selection of the articles used, with some of them exhibiting limited applicability to the domain of Finance. The scarcity of articles in the field that verify the conclusions through theoretical frameworks is also regarded as a limitation.</p>
(Chen, Wu, & Zhao, 2023)	<p>This article is based on the concept of Generative AI establishing the opportunities, challenges, and applications of this technology. It delves into a case study involving ChatGPT, evaluating its trustworthiness in doing sentiment scoring analysis for firms. Ethical limitations, such as data protection problems associated with the use of these tools, are also discussed.</p>	s/d

(Weber, Carl, & Hinz, 2023)	<p>Built upon the principles of XAI and Post-hoc Explainability, this article explains how these concepts might be applied, through a systematic literature review, transforming the financial sector using AI.</p> <p>Their implementation fosters more transparency, moving away from the black-box mindset, particularly in risk management. The discussion covers regulatory challenges and the conflict arising from AI's black-box nature, along with how XAI might help to overcome these concerns.</p>	<p>Even though the article covers a wide range of sources of information and referenced papers, there remain numerous overlooked sources that could have been investigated. The focus was consciously directed towards constraining the literature review more recent research.</p>
(Zarifis, & Cheng, 2022)	<p>Exploring a trust model in Fintech and Insurtech, the study reveals that consumers have pre-existing beliefs about AI and associated technologies. It provides three levels of transparency regarding AI and related technologies.</p>	<p>Certain limitations listed to both the intrinsic nature of the chosen topic as well as the research methodology used.</p> <p>Furthermore, the geographical scope of this study must be acknowledged as a limitation, as the research was done exclusively within the confines of the European Union (EU).</p>
(Herrmann, & Masawi, 2023)	<p>This paper gives a thorough comprehensive analysis of the use of AI in sectors involving banking, financial services, and insurance.</p> <p>A description of the progress of AI in these fields is also presented.</p>	<p>It relies solely on papers from Scopus as the research platform.</p> <p>The absence of insurance sector representation limits the analysis.</p>
(Ahmed, Alshater, Ammari, & Hammami, 2022)	<p>Through an examination of the existing literature, investigation is carried out across six areas: banking corruption and credit risk assessment, stock price forecasting, portfolio management, volatility and liquidity, commodity price prediction, money laundering detection, personal finance, and lastly, big data analysis.</p>	s/d
(Dziuba, Hrysenko, & Matei, 2022)	<p>The article outlines the several risks involved with the financial system's digitalization and</p>	s/d

	suggests solutions for their mitigation. In addition, it delves into the role of new technologies in the financial services industry.	
(Choithani et al., 2022)	Examining the interaction between cryptocurrencies and AI, this paper goes deep into several AI and ML techniques important to the domain of cryptocurrencies and Bitcoin. Reflecting on AI's critical role in threat identification in the context of cybersecurity, the research additionally looks at the benefits of integrating AI systems in the banking industry.	s/d
(AI-Sartawi, Hussainey, & Razzaque, 2022)	The research project highlights the use of AI in advancing sustainable financing, particularly in influencing long-term investment decisions. Furthermore, it covers numerous challenges and presents solutions to sustainability-related issues in the financial industry.	s/d
(Goodell, Kumar, Lim, & Pattnaik, 2021)	Paper based on a bibliometric review of AI and ML in finance.	The format of a bibliometric review may represent a constraint due to the fact it primarily highlights significant articles, authors, themes, and others in the area without going further into each of them.
(Lee, 2020)	Using Robo-Advisors as an example, the author examines how regulation plays a critical role in boosting the adoption of systems using AI within financial services. The author offers recommendations for using AI to incorporate the general population into the financial markets. Additionally, the debate touches upon proposals for possible benefits for consumers resulting from increased competition in financial markets.	s/d

<p>(Königstorfer & Thalmann, 2020)</p>	<p>Address the impacts, challenges and advantages of AI in commercial banks, especially lending operations and credit risk assessment. The article develops a research agenda, targeting three separate groups: consumers, employees, and other stakeholders.</p>	<p>s/d</p>
<p>(Moloi & Marwala, 2020)</p>	<p>Delineation of distinct AI types, such as ML, computational intelligence, and soft computing, is the focus of this chapter. The discussion broadens to include numerous theories, exploring how they are justified or challenged by the integration of AI.</p>	<p>s/d</p>
<p>(Kovvuri, Fu, Fan, & Seisenberger, 2023)</p>	<p>Given the complexities of evaluating equity funds, this article, combines a ML and XAI to analyse the performance of the G10 fund. The goal is to identify the factors that distinctly influence the fund's performance, taking use of the benefits provided by non-linear models.</p>	<p>Concentrating on a single fund makes generalizing conclusions to other funds more difficult.</p>

Appendix F: A selected list of Financial related AI Tools

The following table presents 14 selected AI tools that are available for use at March 2024. The name, a general description and its Web location is listed. In complement, the associated costs are presented.

Tool name	General description and URL	Associated costs
TradeUI	Increase the quality of your trades and boost your profits quickly with AI. https://tradeui.com/	Stock plan \$45 /month Trader plan \$75 /month
Uptrends AI	An intelligent assistant for investors combining news and social network data for a comprehensive stock market analysis . Ideal for tracking trends and optimising trading decisions. https://www.uptrends.ai/	Free Essentials \$9.99 /month Pro \$49.99 /month
Finta	AI that helps your company raise money and find investors . https://www.trustfinta.com/	e-marketplace: - for founders - for investors
PionexGPT	A cryptocurrency exchange platform offering automatic trading bots. https://www.pionex.com/en/	Depends on services and exchange values
Freshly.ai	Comprehensive AI-powered enterprise reporting provider (GPT-3). https://freshly.ai/discuro-form/	Free
Finchat.io	Use ChatGPT's capabilities to analyze, invest or get answers about a company . https://finchat.io/	Free Plus, \$29 /month Pro, \$79 /month
StockGPT	A financial search engine powered by an AI https://askstockgpt.com/	Free StockGPT Pro \$7 /month Pro \$49 /yearly
Web3 Summary	An AI under ChatGPT that analyzes any crypto project (NFT, Swap, DeFi, Blockchain). https://web3summary.com/	Free
Quadency	Develop effective strategies with an AI bot to follow your directions (crypto) https://quadency.com/	Free
LivePlan	Make your business dream come true with intuitive AI tools. Create a solid business plan and financial forecast without complex spreadsheets. https://www.liveplan.com/	Standard \$20 /month Pro \$40 /month
Fin+	Propel your investments with an AI assistant expert in finance . Get high-quality, accurate	By subscription

	market statistics in seconds. https://www.finplus.ai/	
Dispute AI	Help improve your credit scores with AI in a simple, safe and affordable way. https://disputeai.pro/	Free Premium: \$34.99 /month
Avanzai	An AI co-driver that enables rapid financial analysis by generating Python code from a prompt. https://avanz.ai/	Free
Codenu AI	A no-code platform that allows building AI models without coding and offering various services. https://codenu.ai.versoly.page/	Free