SN Bus Econ (2024) 4:23
https://doi.org/10.1007/s43546-023-00618-x

REVIEW

Artificial intelligence in Finance: a comprehensive review through bibliometric and content analysis

Salman Bahoo[1] · Marco Cucculelli[2] · Xhoana Goga[2] · Jasmine Mondolo[2]

Received: 25 April 2023 / Accepted: 13 December 2023 / Published online: 20 January 2024

Abstract
Over the past two decades, artificial intelligence (AI) has experienced rapid development and is being used in a wide range of sectors and activities, including finance. In the meantime, a growing and heterogeneous strand of literature has explored the use of AI in finance. The aim of this study is to provide a comprehensive overview of the existing research on this topic and to identify which research directions need further investigation. Accordingly, using the tools of bibliometric analysis and content analysis, we examined a large number of articles published between 1992 and March 2021. We find that the literature on this topic has expanded considerably since the beginning of the XXI century, covering a variety of countries and different AI applications in finance, amongst which Predictive/forecasting systems, Classification/detection/early warning systems and Big data Analytics/Data mining /Text mining stand out. Furthermore, we show that the selected articles fall into ten main research streams, in which AI is applied to the stock market, trading models, volatility forecasting, portfolio management, performance, risk and default evaluation, cryptocurrencies, derivatives, credit risk in banks, investor sentiment analysis and foreign exchange management, respectively. Future research should seek to address the partially unanswered research questions and improve our understanding of the impact of recent disruptive technological developments on finance.

Keywords Artificial intelligence · Finance · Machine learning · Bibliometric analysis · Content analysis

JEL Classification O33 · G3

Introduction

The first two decades of the twenty-first century have experienced an unprecedented way of technological progress, which has been driven by advances in the development of cutting-edge digital technologies and applications in Artificial Intelligence

Extended author information available on the last page of the article

(AI). Artificial intelligence is a field of computer science that creates intelligent machines capable of performing cognitive tasks, such as reasoning, learning, taking action and speech recognition, which have been traditionally regarded as human tasks (Frankenfield 2021). AI comprises a broad and rapidly growing number of technologies and fields, and is often regarded as a general-purpose technology, namely a technology that becomes pervasive, improves over time and generates complementary innovation (Bresnahan and Trajtenberg 1995). As a result, it is not surprising that there is no consensus on the way AI is defined (Van Roy et al. 2020). An exhaustive definition has been recently proposed by Acemoglu and Restrepo (2020, p.1), who assert that Artificial Intelligence is "(…) the study and development of intelligent (machine) agents, which are machines, software or algorithms that act intelligently by recognising and responding to their environment." Even though it is often difficult to draw precise boundaries, this promising and rapidly evolving field mainly comprises machine learning, deep learning, NLP (natural language processing) platforms, predictive APIs (application programming interface), image recognition and speech recognition (Martinelli et al. 2021).

The term "Artificial intelligence" was first coined by John McCarthy in 1956 during a conference at Dartmouth College to describe "thinking machines" (Buchanan 2019). However, until 2000, the lack of storage capability and low computing power prevented any progress in the field. Accordingly, governments and investors lost their interest and AI fell short of financial support and funding in 1974–1980 and again in 1987–1993. These periods of funding shortage are also known as "AI winters[1]".

However, the most significant development and spread of AI-related technologies is much more recent, and has been prompted by the availability of large unstructured databases, the explosion of computing power, and the rise in venture capital intended to support innovative, technological projects (Ernst et al. 2018). One of the most distinctive The term AI winter first appeared in 1characteristics of AI technologies is that, unlike industrial robots, which need to receive specific instructions, generally provided by a software, before they perform any action, can learn for themselves how to map information about the environment, such as visual and tactile data from a robot's sensors, into instructions sent to the robot's actuators (Raj and Seamans 2019). Additionally, as remarked by Ernst et al. (2018), whilst industrial robots mostly perform manual tasks, AI technologies are able to carry out activities that, until some years ago, were still regarded as typically human, i.e. what Ernst and co-authors label as "mental tasks".

The adoption of AI is likely to have remarkable implications for the subjects adopting them and, more in general, for the economy and the society. In particular, it is expected to contribute to the growth of the global GDP, which, according to a study conducted by Pricewater-house-Coopers (PwC) and published in 2017, is likely to increase by up to 14% by 2030. Moreover, companies adopting AI

[1] The term AI winter first appeared in 1984 as the topic of a public debate at the annual meeting of the American Association of Artificial Intelligence (AAAI). It referred to hype generated by over promises from developers, unrealistically high expectations from end users, and extensive media promotion.

technologies sometimes report better performance (Van Roy et al. 2020). Concerning the geographic dimension of this field, North America and China are the leading investors and are expected to benefit the most from AI-driven economic returns. Europe and emerging markets in Asia and South America will follow, with moderate profits owing to fewer and later investments (PwC 2017). AI is going to affect labour markets as well. The demand for high-skilled employees is expected to increase, whilst the demand for low-skilled jobs is likely to shrink because of automation; the resulting higher unemployment rate, however, is going to be offset by the new job opportunities offered by AI (Ernst et al. 2018; Acemoglu and Restrepo 2020).

AI solutions have been introduced in every major sector of the economy; a sector that is witnessing a profound transformation led by the ongoing technological revolution is the financial one. Financial institutions, which rely heavily on Big Data and process automation, are indeed in a "unique position to lead the adoption of AI" (PwC 2020), which generates several benefits: for instance, it encourages automation of manufacturing processes which in turn enhances efficiency and productivity. Next, since machines are immune to human errors and psychological factors, it ensures accurate and unbiased predictive analytics and trading strategies. AI also fosters business model innovation and radically changes customer relationships by promoting customised digital finance, which, together with the automation of processes, results in better service efficiency and cost-saving (Cucculelli and Recanatini 2022). Furthermore, AI is likely to have substantial implications for financial conduct and prudential supervisors, and it also has the potential to help supervisors identify potential violations and help regulators better anticipate the impact of changes in regulation (Wall 2018). Additionally, complex AI/machine learning algorithms allow Fintech lenders to make fast (almost instantaneous) credit decisions, with benefits for both the lenders and the consumers (Jagtiani and John 2018). Intelligent devices in Finance are used in a number of areas and activities, including fraud detection, algorithmic trading and high-frequency trading, portfolio management, credit decisions based on credit scoring or credit approval models, bankruptcy prediction, risk management, behavioural analyses through sentiment analysis and regulatory compliance.

In recent years, the adoption of AI technologies in a broad range of financial applications has received increasing attention by scholars; however, the extant literature, which is reviewed in the next section, is quite broad and heterogeneous in terms of research questions, country and industry under scrutiny, level of analysis and method, making it difficult to draw robust conclusions and to understand which research areas require further investigation. In the light of these considerations, we conduct an extensive review of the research on the use of AI in Finance thorough which we aim to provide a comprehensive account of the current state of the art and, importantly, to identify a number of research questions that are still (partly) unanswered. This survey may serve as a useful roadmap for researchers who are not experts of this topic and could find it challenging to navigate the extensive and composite research on this subject. In particular, it may represent a useful starting point for future empirical contributions, as it provides an account of the state of the art and of the issues that deserve further investigation. In doing so, this study complements some previous systematic reviews on the topic, such as the ones recently conducted by Hentzen et al. (2022b) and (Biju et al. 2020), which differ from our work in the following main respects: Hentzen and co-authors' study focuses on customer-facing financial services, whilst the valuable contribution of Biju et al. poses particular attention to relevant technical aspects and the assessment of the effectiveness and the predictive capability of machine learning, AI and deep learning mechanisms within the financial sphere; in doing so, it covers an important issue which, however, is out of the scope of our work.

From our review, it emerges that, from the beginning of the XXI century, the literature on this topic has significantly expanded, and has covered a broad variety of countries, as well as several AI applications in finance, amongst which Predictive/forecasting systems, Classification /detection/early warning systems and Big data Analytics/Data mining /Text mining stand out. Additionally, we show that the selected articles can be grouped into ten main research streams, in which AI is applied to the stock market, trading models, volatility forecasting, portfolio management, performance, risk & default evaluation, cryptocurrencies, derivatives, credit risks in banks, investor sentiment analysis and foreign exchange management, respectively.

The balance of this paper is organised as follows: Sect. "Methodology" shortly presents the methodology. Sect. "A detailed account of the literature on AI in Finance" illustrates the main results of the bibliometric analysis and the content analysis. Sect. "Issues that deserve further investigation" draws upon the research streams described in the previous section to pinpoint several potential research avenues. Sect. "Conclusions" concludes. Finally, Appendix 1 clarifies some AI-related terms and definitions that appear several times throughout the paper, whilst Appendix 2 provides more information on some of the articles under scrutiny.

Methodology

To conduct a sound review of the literature on the selected topic, we resort to two well-known and extensively used approaches, namely bibliometric analysis and content analysis. Bibliometric analysis is a popular and rigorous method for exploring and analysing large volumes of scientific data which allows us to unpack the evolutionary nuances of a specific field whilst shedding light on the emerging areas in that field (Donthu et al. 2021). In this study, we perform bibliometric analysis using HistCite, a popular software package developed to support researchers in elaborating and visualising the results of literature searches in the Web of Science platform. Specifically, we employ HistCite to recover the annual number of publications, the number of forward citations (which we use to identify the most influential journals and articles) and the network of co-citations, namely, all the citations received and given by journals belonging to a certain field, which help us identify the major research streams described in Sect. "Identification of the major research streams". After that, to delve into the contents of the most pertinent studies on AI in finance, we resort to traditional content analysis, a research method that provides a systematic and objective means to make valid inferences from verbal, visual, or written data which, in turn, permit to describe and quantify specific phenomena (Downe-Wambolt 1992).

In order to identify the sample of studies on which bibliometric and content analysis were performed, we proceeded as follows. First, we searched for pertinent articles published in English be-tween 1950 and March 2021. Specifically, we scrutinised the “Finance”, “Economics”, “Business Finance” and “Business” sections of the “Web of Science” (WoS) database using the keyword “finance” together with an array of keywords concerning Artificial Intelligence (i.e. “Finance” AND (“Artificial Intelligence” OR “Machine Learning” OR “Deep Learning” OR “Neural Networks*” OR “Natural Language Processing*” OR “Algorithmic Trading*” OR “Artificial Neural Network” OR “Robot*” OR “Automation” OR “Text Mining” OR “Data Mining” OR “Soft Computing” OR “Fuzzy Logic Analysis” OR “Biometrics*” OR “Geotagging” OR “Wearable*” OR “IoT” OR “Internet of Thing*” OR “digitalization” OR “Artificial Neutral Networks” OR “Big Data” OR “Industry 4.0″ OR “Smart products*” OR Cloud Computing” OR “Digital Technologies*”). In doing so, we ended up with 1,218 articles. Next, two researchers independently analysed the title, abstract and content of these papers and kept only those that address the topic under scrutiny in a non-marginal and non-trivial way. This second step reduced the number of eligible papers to 892, which were used to perform the first part of the bibliometric analysis. Finally, we delved into the contents of the previously selected articles and identified 110 contributions which specifically address the adoption and implications in Finance of AI tools focussing on the economic dimension of the topic, and which are employed in the second part of the bibliometric analysis and in the content analysis.

A detailed account of the literature on AI in Finance

In this section, we explore the patterns and trends in the literature on AI in Finance in order to obtain a compact but exhaustive account of the state of the art. Specifically, we identify some relevant bibliographic characteristics using the tools of bibliometric analysis. After that, focussing on a sub-sample of papers, we conduct a preliminary assessment of the selected studies through a content analysis and detect the main AI applications in Finance. Finally, we identify and briefly describe ten major research streams.

Main results of the bibliometric analysis

First, using HistCite and considering the sample of 892 studies, we computed, for each year, the number of publications related to the topic “AI in Finance”. The corresponding publication trend is shown in Fig. 1, which plots both the annual absolute number of sampled papers (bar graph in blue) and the ratio between the latter and the annual overall amount of publications (indexed in Scopus) in the finance area (line graph in orange). We also compute relative numbers to see if the trend emerging from the selected studies is not significantly attributable to a “common trend” (i.e. to the fact that, in the meantime, also the total number of publications in the financial area has significantly increased). It can be noted that both graphs exhibit a

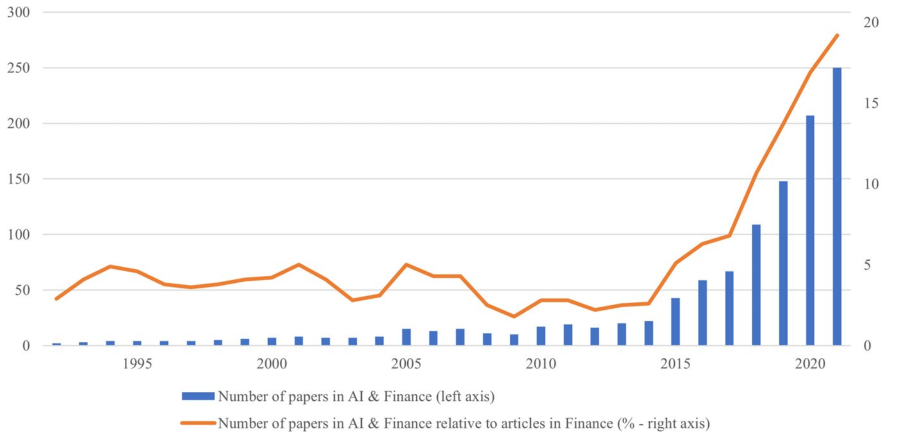

Fig. 1 Publication Trend, 1992–2021

strong upward trend from 2015 onwards; during the most recent years, the pace of growth and the degree of pervasiveness of AI adoption in the financial sphere have indeed remarkably strengthened, and have become the subject of a rapidly growing number of research articles.

After that, focussing on the more pertinent (110) articles, we checked the journals in which these studies were published. Table 1 presents the top-ten list of journals reported in the Academic Journal Guide-ABS List 2020 and ranked on the basis of the total global citation score (TGCS), which captures the number of times an article is cited by other articles that deal with the same topic and are indexed in the WoS database. For each journal, we also report the total number of studies published in that journal. We can notice that the most influential journals in terms of TGCS are the Journal of Finance (with a TGCS equal to 1283) and the Journal of Banking and Finance (with a TGCS of 1253), whilst the journals containing the highest number of articles on the topic are Quantitative Finance (68 articles) and Intelligent Systems in Accounting, Finance and Management (43).

Finally, Fig. 2 provides a visual representation of the citation-based relationships amongst papers starting from the most-cited papers, which we obtained using the Java application CiteSpace.

Preliminary results of the content analysis

In this paragraph, we shortly illustrate some relevant characteristics of our sub-sample made up of 110 studies, including country and industry coverage, method and underpinning theoretical background. Table 2 comprises the list of countries under scrutiny, and, for each of them, a list of papers that perform their analysis on that country. We can see that our sample exhibits significant geographical heterogeneity, as it covers 74 countries across all continents; however, the most investigated areas are three, that is Europe, the US and China. These results corroborate the fact that the above-mentioned regions are the leaders of the AI-driven financial industry, as

Table 1 Top ten journals publishing articles on the selected topic

Ranking	Journal	No of articles	TGCS	Journal	No of articles	TGCS
1	Journal of Finance	9	1283	Quantitative Finance	68	368
2	Journal of Banking and Finance	28	1256	Intelligent Systems in Accounting, Finance and Management	43	273
3	International Journal of Forecasting	20	521	Journal of Banking and Finance	28	1256
4	Journal of Economic Dynamics and Control	4	377	International Journal of Finance and Economics	21	66
5	Quantitative Finance	68	368	International Journal of Forecasting	20	521
6	Journal of Forecasting	17	275	Computational Economics	17	87
7	Intelligent Systems in Accounting, Finance and Management	43	273	Journal of Forecasting	17	275
8	Accounting Organizations and Society	1	210	European Journal of Finance	16	73
9	Mathematical Finance	11	188	Technological Forecasting and Social Change	15	63
10	Journal Of Business Research	5	182	Pacific-Basin Finance Journal	14	53

[a]Note: ARMA: Autoregressive–moving-average model; ARIMA: Autoregressive integrated moving average; HAR: Heterogeneous autoregressive approach, VAR: vector autoregressive approach

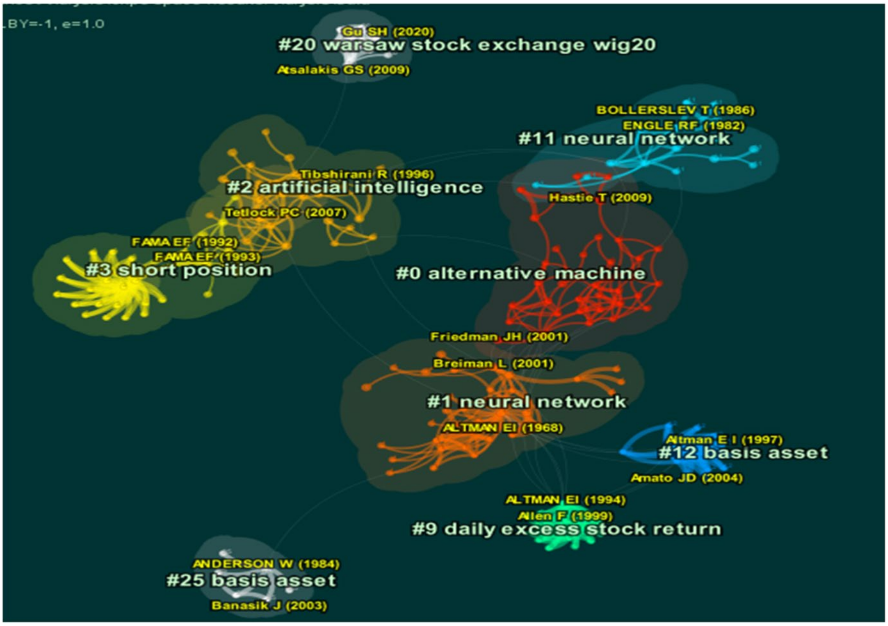

Fig. 2 Citation Mapping and identification of the research streams. Source: authors' elaboration of data from Web of Science; visualisation produced using CiteSpace

suggested by PwC (2017). The United States, in particular, are considered the "early adopters" of AI and are likely to benefit the most from this source of competitive advantage. More lately, emerging countries in Southeast Asia and the Middle East have received growing interest. Finally, a smaller number of papers address underdeveloped regions in Africa and various economies in South America.

The most investigated sectors are reported in Table 3. We can notice that, although it primarily deals with banking and financial services, the extant research has addressed the topic in a vast array of industries. This confirms that the application potential of AI is very broad, and that any industry may benefit from it.

Through our analysis, we also detected the key theories and frameworks applied by researchers in the prior literature. As shown in Table 4, 73 (out of 110) papers explicitly refer to some theoretical framework. Specifically, ten of them (14%) resort to computational learning theory; this theory, which is an extension of statistical learning, provides researchers with a theoretical guide for finding the most suitable learning model for a given problem, and is regarded as one of the most important and most used theories in the field. Specific theories concerning types of neural networks and learning methods are used too, such as the fuzzy set theory, which is mentioned in 8% of the sample, and to a lesser extent, the Naive Bayes theorem, the theory of neural networks, the theory of genetic programming and the TOPSIS analytical framework. Finance theories (e.g. Arbitrage Pricing Theory; Black and Scholes 1973) are jointly employed with portfolio management theories (e.g. modern portfolio theory), and the two of them account together for 21% (15) of the total number of

Table 2 List of countries covered by the literature

Sr. No	Country where study is conducted	Author(s) / Year
Europe		
	Denmark	Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Iceland	Calomiris and Mamaysky 2019; Jain, 2005;
	Austria	Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Italy	Altman et al. 1994; Varetto 1998; Guotai and Abedin 2017; Chen et al. 2013; Deku et al. 2020; Corazza et al. 2021; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Ireland	Deku et al. 2020; Rasekhschaffe and Jones 2019; Jain, 2005;
	France	Chen et al. 2013; Deku et al. 2020; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Mselmi et al. 2017; Holopainen and Sarlin 2017; Jain, 2005;
	Netherlands	Chen et al. 2013; Deku et al. 2020; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Norway	Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Jain, 2005;
	Finland	Caglayan et al. 2020; Calomiris and Mamaysky 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Sweden	Deku et al. 2020; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Holopainen and Sarlin 2017;
	Belgium	(Pompe, and Bilderbeek 2005) Chen et al. 2013; Deku et al. 2020; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; D'Hondt et al. 2020; Holopainen and Sarlin 2017; Jain, 2005;
	Luxembourg	Calomiris and Mamaysky 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Greece	Lahmiri 2016; Guotai and Abedin 2017; Frino et al. 2017; Rodrigues and Stevenson 2013; Chen et al. 2013; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Loukeris and Eleftheriadis 2015; Jain, 2005;
	Switzerland	Deku et al. 2020; Rasekhschaffe and Jones 2019; Jain, 2005;
	UK	Chen et al. 2013; Deku et al. 2020; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Abdou et al. 2021; Kanas 2001; Wanke et al. 2016a, b, c, d; Holopainen and Sarlin 2017; Kim and Kim 2020; Jain, 2005; Sermpinis et al. 2013;

Table 2 (continued)

Sr. No	Country where study is conducted	Author(s) / Year
	Germany	Lahmiri 2016; Chen et al. 2013; Trinkle and Baldwin 2016; Deku et al. 2020; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Xu et al. 2019; Holopainen and Sarlin 2017; Seriev and Germano, 2020; Jain, 2005; Creamer 2012;
	Portugal	Deku et al. 2020; Rasekhschaffe and Jones 2019; Holopainen and Sarlin 2017; Jain, 2005; Calomiris and Mamaysky 2019; Jain, 2005;
	Spain	Cortés et al. 2008; Caglayan et al. 2020; Calomiris and Mamaysky 2019; Holopainen and Sarlin 2017; Jain, 2005;
	Czech Republic	Calomiris and Mamaysky 2019; Jain, 2005;
	Ukraine	Calomiris and Mamaysky 2019; Jain, 2005;
	Romania	Sabău Popa et al. 2021; Jain, 2005;
	Slovenia	Jagric et al. 2011; Jain, 2005;
	Slovakia	Calomiris and Mamaysky 2019; Jain, 2005;
	Poland	Calomiris and Mamaysky 2019; Trinkle and Baldwin 2016; Seriev and Germano, 2020; Jain, 2005;
	Estonia	Caglayan et al. 2020; Calomiris and Mamaysky 2019; Jain, 2005;
	Hungary	Calomiris and Mamaysky 2019; Jain, 2005;
	East Europe	Seriev and Germano, 2020; Jain, 2005;
	Europe (no specif country)	Bucci et al.,2020; Jones and Wang 2019; Kumar et al. 2019; Jones and Wang 2019; Booth et al. 2015; Creamer 2012;
North America		

Table 2 (continued)

Sr. No	Country where study is conducted	Author(s) / Year
	USA	Coats and Fant 1993; Jones et al. 2015; Jones et al. 2017; Butaru et al. 2016; Gepp et al. 2010; Scholtus et al. 2014; Qi and Maddala 1999; Sirignano 2018; Fernandes et al. 2014; Kelejian and Mukerji 2016; Le and Viviani 2018; Wei et al. 2019; Qi 1999; Litzenberger et al. 2012; Chen et al. 2013; Kanas 2001; Vortelinos 2017; Renault 2017; Huang and Kuan 2021; Houlihan and Creamer 2021; Tao et al. 2021; Jain et al. 2021;Calomiris and Mamaysky 2019; Bekiros and Georgoutsos 2008;Heston and Sinha 2017;Rasekhschaffe and Jones 2019; Cao et al. 2022; Kercheval and Zhang 2015; Papadimitriou, Goga, and Agrapetidou, 2020; Soleymani and Vasighi 2020;Abedin et al. 2019; Bucci et al.,2020; Jones and Wang 2019; Creamer and Freund 2010; Booth et al. 2015; Zhao et al. 2018; Jang and Lee 2019; Kim and Kim 2020; Jain, 2005;
	Canada	Jones et al. 2015; Moshiri, and Cameron 2000; Chen et al. 2013; Calomiris and Mamaysky 2019; Jain, 2005;
South America		
	Brazil	Sun and Vasarhelyi 2018; Calomiris and Mamaysky 2019; Jain, 2005;
	Latin America	Jones et al. 2015; Jain, 2005;
	Mexico	Calomiris and Mamaysky 2019; Jain, 2005;
	Peru	Calomiris and Mamaysky 2019; Jain, 2005;
	Argentina	Calomiris and Mamaysky 2019; Jain, 2005;
	Colombia	Calomiris and Mamaysky 2019; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Jain, 2005;
	Bolivia	Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Jain, 2005;
	Chile	Calomiris and Mamaysky 2019; Jain, 2005;
	Latin America	Jones et al. 2015; Jain, 2005;
	Mexico	Calomiris and Mamaysky 2019; Jain, 2005;
Asia		
	India	Calomiris and Mamaysky 2019; Jain, 2005;
	Japan	Lahmiri 2016;
	Thailand	Calomiris and Mamaysky 2019; Jain, 2005;

Table 2 (continued)

Sr. No	Country where study is conducted	Author(s) / Year
	Indonesia	Calomiris and Mamaysky 2019; Wanke et al. 2016a, b, c, d; Jain, 2005;
	Philippines	Calomiris and Mamaysky 2019; Jain, 2005;
	Malaysia	Wanke et al. 2016a, b, c, d ; Calomiris and Mamaysky 2019; Wanke et al. 2016a, b, c, d; Jain, 2005;
	Mauritius	Amelot et al. 2021; Jain, 2005;
	Singapore	Chen et al. 2013; Jain, 2005; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019;
	Pakistan	Wanke et al. 2016a, b, c, d; Jain, 2005;
	Hong Kong	Chen et al. 2013; Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Jain, 2005; Kim and Kim 2020;
	Asia Pacific	Jones et al. 2015; Booth et al. 2015;
	China	Guotai and Abedin 2017; Lu et al. 2013; Jiang and Jones 2018; Zhang et al. 2021; Calomiris and Mamaysky 2019;; Xu and Zhao 2022; Uddin et al. 2020; Yin et al. 2020; Abedin et al. 2019; Gao et al. 2016; Li et al. 2020; Jain, 2005;
	Bangladesh	Wanke et al. 2016a, b, c, d; Jain, 2005;
	South Korea	Calomiris and Mamaysky 2019; Jain, 2005;
	Russia	Calomiris and Mamaysky 2019; Jain, 2005;
	Taiwan	Lu et al. 2013; Jain, 2005; Abedin et al. 2019;
Oceania		
	New Zealand	Calomiris and Mamaysky 2019; Rasekhschaffe and Jones 2019; Jain, 2005;
	Australia	Lahmiri 2016; Guotai and Abedin 2017; Chen et al. 2013; Calomiris and Mamaysky 2019; Xu et al. 2019; Frino et al. 2020; Jain, 2005;
Middle East		
	Iran	Zhang and Feng, 2019; Wanke et al. 2016a, b, c, d;; Jain, 2005;
	Qatar	Wanke et al. 2016a, b, c, d;
	Bahrain	
	UAE	Wanke et al. 2016a, b, c, d; Jain, 2005;

Table 2 (continued)

Sr. No	Country where study is conducted	Author(s) / Year
	Kuwait	Wanke et al. 2016a, b, c, d; Jain, 2005;
	Turkey	Calomiris and Mamaysky 2019; Wanke et al. 2016a, b, c, d; Jain, 2005;
	Israel	Dunis et al. 2013; Deku et al. 2020; Feldman and Gross 2005; Calomiris and Mamaysky 2019; Jain, 2005;
	Saudi Arabia	Wanke et al. 2016a, b, c, d; Jain, 2005;
	Sudan	Wanke et al. 2016a, b, c, d
	Tunis	Wanke et al. 2016a, b, c, d; Jain, 2005;
	Egypt	Abdou et al. 2021; Wanke et al. 2016a, b, c, d; Jain, 2005;
Africa		
	South Africa	Calomiris and Mamaysky 2019; Jain, 2005;
	Nigeria	Calomiris and Mamaysky 2019; Jain, 2005;
	Kenya	Calomiris and Mamaysky 2019; Jain, 2005;
	Ghana	Calomiris and Mamaysky 2019; Jain, 2005;
	Gambia	Wanke et al. 2016a, b, c, d; Jain, 2005;

Table 3 List of sectors covered by the literature

Sector	Author(s) / Year
Aerospace, airline, aircraft	Kelejian and Mukerji 2016; Zhang et al. 2021; Reber 2014; Kanas 2001;
Agriculture, Hunting and forestry fishing	Cortés et al. 2008; Jones and Wang 2019;
Agriculture Machinery	Kelejian and Mukerji 2016;
Automotive industry, Vehicle Manufacturing, Repair of vehicles	Kelejian and Mukerji 2016; Cortés et al. 2008; Zhang et al. 2021;
Banking /financial services	Khandani et al. 2010; Butaru et al. 2016; Lahmiri 2016; Kim and Kim 2014; Sun and Vasarhelyi 2018; Dunis et al. 2013; Sirignano 2018; Feldman and Gross 2005; Fernandes et al. 2014; Wanke et al. 2016a, b, c, d; Guotai and Abedin 2017; Frino et al. 2017; Le and Viviani 2018; Wei et al. 2019; Cortés et al. 2008; Jagric, Jagric, and Kracun, 2011; Trinkle and Baldwin 2016; Culkin and Das 2017; Law and Shawe-Taylor 2017; Vortelinos 2017; Renault 2017; Jiang and Jones 2018; Zhang et al. 2021; Deku et al. 2020; Caglayan et al. 2020; Calomiris and Mamaysky 2019; Reber 2014; Kumar et al. 2019; Cao, Liu, Zhai, et al., 2020;; Xu and Zhao 2022; Papadimitriou, Goga, and Agrapetidou, 2020; Tao et al. 2021; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Kanas 2001; Loukeris and Eleftheriadis 2015; Abedin et al. 2019; Xu et al. 2019; Wanke et al. 2016a, b, c, d ; Jones and Wang 2019; Episcopos, Pericli and Hu, 1998; Funahashi 2020; Lu and Ohta 2003; Holopainen and Sarlin 2017; Zhao et al. 2018; Guotai and Abedin 2017; Hentzen et al. 2022a; IBM Cloud Education 2020; Petukhina et al. 2020; PwC 2017; PwC 2018; Tao et al. 2021; Yang et al. 1999; Zheng et al. 2019
Business services	Uddin et al. 2020;
Raw Materials	Kim and Kim 2014;
Commercial and service industry and/or general machinery	Varetto 1998; Kelejian and Mukerji 2016;
Construction	Altman et al. 1994; Varetto 1998; Cortés et al. 2008;Sabau, Popa et al., 2021; Reber 2014; Uddin et al. 2020; Kanas 2001; Jones and Wang 2019;
Consumer goods	Kim and Kim 2014; Kelejian and Mukerji 2016; Kanas 2001;
Commodities	Yang et al. 1999; Fernandes et al. 2014; Kelejian and Mukerji 2016; Trinkle and Baldwin 2016; Zhang et al. 2021; Li et al. 2020;
Computer and peripheral equipment	Kelejian and Mukerji 2016;
Communication	Kelejian and Mukerji 2016; Cortés et al. 2008; Jones and Wang 2019;
Cryptocurrency	Pichl and Kaizoji 2017; Burggraf 2021; Petukhina et al. 2020;
Education	Cortés et al. 2008;

Table 3 (continued)

Sector	Author(s) / Year
Electronics Equipment and Manufacturing industry	Reber 2014; Kelejian and Mukerji 2016;
Electronics	Kelejian and Mukerji 2016;
Energy and utilities	Jones et al. 2017; Kim and Kim 2014; Jiang and Jones 2018; Sabau, Popa et al., 2021; Zhang et al. 2021; Cortés et al. 2008; Jones et al. 2017; Reber 2014; Kelejian and Mukerji 2016; Li et al. 2020;
Extractive industry	Sabău Popa et al. 2021;
FinTech	Jones et al. 2017; Kelejian and Mukerji 2016; Cortés et al. 2008; Tao et al. 2021;
Food, Tobacco, Beverages	Jones et al. 2017; Zhang et al. 2021; Cortés et al. 2008; Kanas 2001; Reber 2014;
Footwear	Kanas 2001;
Health Care	Kelejian and Mukerji 2016; Kim and Kim 2014; Cortés et al. 2008; Jones et al. 2017; Reber 2014; Kanas 2001;
Gold	Law and Shawe-Taylor 2017
Name of Industry	Author(s) / Year
Heating Industry	Kelejian and Mukerji 2016; Pompe, and Bilderbeek 2005;
Household goods	Jones et al. 2017;
Information services	Uddin et al. 2020;
IT industry	Jones et al. 2017; Uddin et al. 2020; Kanas 2001; Varetto 1998; D'Hondt et al. 2020; Creamer 2012; Creamer and Freund 2010;
Manufacturing (of woods, textile, leather products)	Sabau, Popa et al., 2021; Cortés et al. 2008; Reber 2014; Jones and Wang 2019;
Manufacture of Chemical, Plastics, Rubber	Coats and Fant 1993; Gepp et al. 2010; Cortés et al. 2008; Reber 2014; Kanas 2001;
Manufacture of electrical and optical equipment	Cortés et al. 2008;
Medical equipment and supplies	Kelejian and Mukerji 2016; Cortés et al. 2008;
Metal	Li et al. 2020;
Mining industry	Kelejian and Mukerji 2016; Rodrigues and Stevenson 2013; Zhang et al. 2021; Jones and Wang 2019;
Paper, paper products, publishing, printing	Cortés et al. 2008;
Pharmaceutical and medicine	Kelejian and Mukerji 2016; Cortés et al. 2008; Zhang et al. 2021; Reber 2014; Kanas 2001;
Power and automation Technology	Kelejian and Mukerji 2016;
Petroleum, Nuclear fuel	Law and Shawe-Taylor 2017; Kanas 2001;
Restaurants. Hotel, tourism and personal services	Cortés et al. 2008;Sabau, Popa et al., 2021; Reber 2014; Uddin et al. 2020;
Wholesale and Retail	Jones et al. 2017; Cortés et al. 2008; Jiang and Jones 2018; Sabau, Popa et al., 2021; Reber 2014; Uddin et al. 2020; Kanas 2001; Jones and Wang 2019;

Table 3 (continued)

Sector	Author(s) / Year
Public administration and defence	Altman et al. 1994; Jones et al. 2015; Jones et al. 2017; Gepp et al. 2010; Cortés et al. 2008; Reber 2014; Jones and Wang 2019;
Real estate, renting and business activities	Cortés et al. 2008; Chen et al. 2013; Zhang et al. 2021; Uddin et al. 2020; Jones and Wang 2019;
Robotics /automation industry	Cortés et al. 2008;
Hygiene products	Jones et al. 2017;
Social media platforms	Houlihan and Creamer 2021;; Xu and Zhao 2022;
Software Engeerning	
Soap, cleaning compound andtoilet preparation	Kelejian and Mukerji 2016;
Technology company	Cortés et al. 2008; Kim and Kim 2014;
Trading	
Telecommunication (service and manufacturing, companies)	Kim and Kim 2014; Jones et al. 2017; Zhang et al. 2021; Reber 2014; Heston and Sinha 2017; Kanas 2001;
Transportation and storage	Dunis et al. 2013; Scholtus et al. 2014; Reboredo et al. 2012; Sabau, Popa et al., 2021; Cortés et al. 2008; Reber 2014; Uddin et al. 2020; Jones and Wang 2019;
Vehicle manufacturing	Jones et al. 2017;
Professional Scientific and technical activities	Sabău Popa et al. 2021;
Warehousing	Uddin et al. 2020;

papers. Finally, bankruptcy theories support business failure forecasts, whilst other theoretical underpinnings concern mathematical and probability concepts.

The content analysis also provides information on the main types of companies under scrutiny. Table 5 indicates that 30 articles (out of 110) focus on large companies listed on stock exchanges, whilst only 16 studies cover small and medium enterprises. Similarly, trading and digital platforms are examined in 16 papers that deal with derivatives and cryptocurrencies.

Furthermore, Table 6 summarises the key methods applied in the literature, which are divided by category (note that all the papers employ more than one method). Looking at the table, we see that machine learning and artificial neural networks are the most popular ones (they are employed in 41 and 51 articles, respectively). The majority of the papers resort to different approaches to compare their results with those obtained through autoregressive and regression models or conventional statistics, which are used as the benchmark; therefore, there may be some overlaps. Nevertheless, we notice that support vector machine and random forest are the most widespread machine learning methods. On the other hand, the use of artificial neural networks (ANNs) is highly fragmented. Backpropagation, Recurrent, and Feed-Forward NNs are considered basic neural nets and are commonly employed. Advanced NNs, such as Higher-Order Neural network (HONN) and Long Short-Term Memory Networks (LSTM), are more performing than their standard version but also much more complicated to apply. These methods are usually compared to autoregressive

Table 4 Theories and frameworks employed in the literature

Theories/Frameworks	No of Articles	Author(s) / Year
Statistical Learning Theory/ Computational Learning Theory	10	Qi 1999; Rodrigues and Stevenson 2013; Law and Shawe-Taylor 2017; Xu et al. 2019; Episcopos, Pericli, and Hu., 1998; Chabou et al., 2014; Jones et al. 2017; Lahmiri 2016; Reboredo et al. 2012; Le and Viviani 2018;
Finance theories (Arbitrage Pricing Theory, Efficient Market Theory, Black and Scholes' theory)	10	Qi and Maddala 1999; Lu et al. 2013; Caglayan et al. 2020; Moshiri and Cameron 2000; Kim and Kim 2020; Litzenberger et al. 2012; Fernandes et al. 2014; Culkin and Das 2017; Chen and Wan 2021; Lu and Ohta 2003;
Fuzzy set theory	6	Trinkle and Baldwin 2016; Huang and Guo 2021; Xu et al. 2019; Jiang and Jones 2018; Lahmiri and Bekiros 2019; Uddin et al. 2020;
Modern Portfolio Theory	5	Loukeris and Eleftheriadis 2015; Soleymani and Vasighi 2020; Zhao et al. 2018; Petukhina et al. 2020;
Naives Bayes ' theorem (Information Criterion, decision-making)	5	Lahmiri 2016; Law and Shawe-Taylor 2017; Jones et al. 2017; Moshiri and Cameron 2000; Jagric, Jagric, and Kracun, 2011; Yang et al. 1999; Gepp et al. 2010;
Econometric Theory	4	Reboredo et al. 2012; Parot et al. 2019; Bucci 2020;
Theory of Neural networks	4	Altman et al. 1994; Wanke et al. 2016a, b, c, d; Qi 1999; Sariev and Germano 2020;
Framework of Hasbrouck	3	Hendershott et al. 2011; Frino et al. 2017;
Probability theories (Dempster–Shafer (D–S) evidence theory)	3	Gepp et al. 2010; Coats and Fant 1993; Jiang and Jones 2018;
Bankruptcy theory / Business failure theory	2	Varetto 1998; Cortés et al. 2008;
Random matrix theory	2	Soleymani and Vasighi 2020; D'Hondt et al. 2020;
Signal detection theory	2	Varetto 1998; Mselmi et al. 2017;
Theory of intraday patterns	2	Fernandes et al. 2014; Litzenberger et al. 2012;
Entropy theory	2	Lu et al. 2013; Heston and Sinha 2017;
Markov decision-making process	1	Dunis et al. 2013;
Agency theory	1	Cao et al. 2022;
Behavioural consistency theory	1	Kamiya et al. 2018;
Theory of power-law distribution (financial markets)	1	Booth et al. 2015;
Conventional valuation theory	1	Jiang and Jones 2018;
Cox–Ross–Rubinstein framework	1	Reber 2014;
Decision theory	1	Law and Shawe-Taylor 2017;

Table 4 (continued)

Theories/Frameworks	No of Articles	Author(s) / Year
Economic theory	1	Bucci 2020; Wei et al. 2019;
Economic theories of 'Matching and managerial talent'	1	Jiang and Jones 2018;
Elder financial abuse: conceptual framework	1	Kumar et al. 2019;
Forecast combinations framework	1	Rasekhschaffe and Jones 2019;
Gradient Theory	1	Culkin and Das 2017;
Graph theory	1	Burggraf 2021;
Individual theory	1	Cao et al. 2022;
KPCA theory (Kernel principal component analysis)	1	Amelot et al. 2021;
Limit order book Dynamics (theoretical model)	1	Sirignano 2018;
Managerial signalling theory	1	Cao et al. 2022;
Preference theory	1	Guotai and Abedin 2017;
Risk parity approach	1	Burggraf 2021;
Sentiment theory	1	Heston and Sinha 2017; Yin et al. 2020;
Stochastic optimal Theory	1	Chen and Ge 2021;
Theory of Genetic programming	1	Dunis et al. 2013;
Time-varying risk premium theory	1	Bekiros and Georgoutsos 2008;
TOPSIS Analytical framework	1	Wanke et al. 2016a, b, c, d
Grey system Theory	1	Chen et al. 2013;
Generalizability Theory	1	Varetto 1998; Feldman and Gross 2005;
Transactions on Information Theory	1	Reboredo et al. 2012;
Transaction Cost Theory	1	Feldman and Gross 2005;

Table 5 Types of firms studied in the literature

Type/Nature of Company/Firm	No of Articles	Author (s) / Year
Based on size		
Large Companies (listed)	30	Khandani et al. 2010; Hendershott et al. 2011; Kim and Kim 2014; Sirignano 2018; Feldman and Gross 2005; Jones et al. 2015,Qi and Maddala 1999; Fernandes et al. 2014; Kelejian and Mukerji 2016; Qi and Maddala 1999; Qi 1999; Cortés et al. 2008; Litzenberger et al. 2012; Lu et al. 2013; Rodrigues and Stevenson 2013;Vortelinos 2017; Renault 2017; Sabau, Popa et al., 2021; Zhang et al. 2021; Jain et al. 2021; Kamiya et al. 2018; Bekiros and Georgoutsos 2008; Heston and Sinha 2017; Rasekh-schaffe and Jones 2019; Cao et al. 2022; Abdou et al. 2021; Soleymani and Vasighi 2020; Kanas 2001; Seriev and Germano, 2020; Butaru et al. 2016;
Small Medium Enterprises (small cap) (SMEs)	16	Altman et al. 1994; Pompe and Bilderbeek 2005; Cortés et al. 2008; Litzenberger et al. 2012; Rodrigues and Stevenson 2013; Chen et al. 2013; Tao et al. 2021; Corazza et al. 2021; Jain et al. 2021; Kamiya et al. 2018; Heston and Sinha 2017; Rasekhschaffe and Jones 2019; Mselmi et al. 2017; Jones and Wang 2019; Jones and Wang 2019; Seriev and Germano, 2020
Trading/ Digital platforms/ Lending online platform (Stocks, crypto, derivatives, loans)	16	Kelejian and Mukerji 2016; Cortés et al. 2008; Litzenberger et al. 2012; Trinkle and Baldwin 2016; Tao et al. 2021; Amelot et al. 2021; Jain et al. 2021; D'Hondt et al. 2020; Gao et al. 2016; Frino et al. 2020; Funahashi 2020; Lu, and Ohta 2003; Creamer and Freund 2010; Creamer 2012; Tashiro et al. 2019; Caglayan et al. 2020;
Banks / Financial Institution (Large and small)	9	Sun and Vasarhelyi 2018; Sirignano 2018; Feldman and Gross 2005; Frino et al. 2017; Wei et al. 2019; Kumar et al. 2019;; Xu and Zhao 2022; Wanke et al. 2016a, b, c, d; Papadimitriou, Goga and Agrapetidou, 2020;
Micro enterprises	3	Gepp et al. 2010; Uddin et al. 2020; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021;
National Banks	1	Wei et al. 2019;
Venture capitals	1	Reber 2014;

Note: Types of firms analysed. Only 76 articles out of 110 define firms' size. 34 articles do not provide information on companies' size for privacy reasons

Table 6 Main methods applied in the literature

Method	n. of articles	Author (s) / Years
Machine learning	41	Khandani et al. 2010; Varetto 1998; Jones et al. 2017; Galeshchuk and Mukherjee 2017; Butaru et al. 2016; Lahmiri 2016; Kercheval and Zhang 2015; Kim, and Kim 2014; Sun and Vasarhelyi 2018; Gepp et al. 2010; Dunis et al. 2013; Feldman and Gross 2005; Reboredo et al. 2012; Le and Viviani 2018; Cortés et al. 2008; Butaru et al. 2016; Law and Shawe-Taylor 2017; Renault 2017; Jiang and Jones 2018; Lahmiri and Bekiros 2019; Burggraf 2021; Huang and Guo 2021; Deku et al. 2020; Houlihan and Creamer 2021; Amelot et al. 2021; Caglayan et al. 2020; Kumar et al. 2019; Rasekhschaffe and Jones 2019; Cao et al. 2022; Papadimitriou, Goga and Agrapetidou, 2020); Soleymani and Vasighi 2020; Uddin et al. 2020; Abedin et al. 2019; Mselmi et al. 2017; D'Hondt, De Winne, Ghysels et al., 2020; Jones and Wang 2019; Creamer and Freund 2010; Creamer 2012; Booth et al. 2015; Holopainen and Sarlin 2017; Jang and Lee 2019;; Xu and Zhao 2022;
Support Vector Machine (SVM)	19	Jones et al. 2015; Jones et al. 2017; Galeshchuk and Mukherjee 2017; Lahmiri 2016; Kercheval and Zhang 2015; Reboredo et al. 2012; Le and Viviani 2018; Law and Shawe-Taylor 2017; Huang and Guo 2021; Houlihan and Creamer 2021; Kumar et al. 2019; Rasekhschaffe and Jones 2019; Cao et al. 2022; Papadimitriou, Goga and Agrapetidou, 2020; Abedin et al. 2019; Mselmi et al. 2017; Booth et al. 2015; Holopainen and Sarlin 2017; Jang and Lee 2019; Hamdi and Aloui 2015;
Random forest	8	Jones et al. 2015; Jones et al. 2017; Butaru et al. 2016; Deku et al. 2020; Kumar et al. 2019; Uddin et al. 2020; D'Hondt et al. 2020; Booth et al. 2015;
Naïve Bayes	5	Lahmiri 2016; Kim and Kim 2014; Sun and Vasarhelyi 2018; Deku et al. 2020;; Xu and Zhao 2022;
Adaboost	5	Jones et al. 2015; Jones et al. 2017; Cortés et al. 2008; Rasekhschaffe and Jones 2019; Creamer 2012;
Least absolute shrinkage and selection operator (LASSO)	3	Caglayan et al. 2020; Cao et al. 2022; Holopainen and Sarlin 2017;
CART	3	Khandani et al. 2010; Gepp et al. 2010;
Decision trees	3	Butaru et al. 2016; Sun and Vasarhelyi 2018; Gepp et al. 2010;
Generalised Boosting	2	Jones et al. 2015; Jones et al. 2017;
Genetic programming	2	Dunis et al. 2013; Feldman and Gross 2005;
Logiboost	1	Creamer and Freund 2010;
TreeNet	1	Jones and Wang 2019;

Table 6 (continued)

Method	n. of articles	Author (s) / Years
Genetic algorithm	1	Varetto 1998;
Gradient boosted regression tree	1	Rasekhschaffe and Jones 2019;
Artificial Neutral Network (ANN)	51	Altman et al. 1994; Jones et al. 2015; Trippi and DeSieno 1992; Yang et al. 1999; Lahmiri 2016; Dunis et al. 2013; Wanke et al. 2016a, b, c, d; Guotai and Abedin 2017; Qi 1999; Nag and Mitra 2002; Pompe and Bilderbeek 2005; Rodrigues and Stevenson 2013; Galeshchuk and Mukherjee 2017; Sun and Vasarhelyi 2018; Qi and Maddala 1999; Reboredo et al. 2012; Le and Viviani 2018; Parot et al. 2019; Chen et al. 2013; Pichl and Kaizoji 2017; Vortelinos 2017; Lahmiri and Bekiros 2019; Sabau, Popa et al., 2021; Zhang et al. 2021; Chen and Wan 2021; Corazza et al. 2021; Amelot et al. 2021; Bekiros and Georgoutsos 2008; Dunis et al. 2013; Dunis et al. 2010; Sermpinis et al. 2013; Reber 2014; Heston and Sinha 2017; Rasekhschaffe and Jones 2019; Abdou et al. 2021; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Kanas 2001; Loukeris and Eleftheriadis 2015; Abedin et al. 2019; Mselmi et al. 2017; D'Hondt et al. 2020; Bucci 2020; Funahashi 2020; Lu and Ohta 2003; Booth et al. 2015; Holopainen and Sarlin 2017; Zhao et al. 2018; Jang and Lee 2019; Sariev and Germano 2020; Jagric, Jagric, and Kracun, 2011; Hamdi and Aloui 2015;
Multilayer perceptron (MLP)	9	Dunis et al. 2013; Dunis et al. 2010; Sermpinis et al. 2013; Reber 2014; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Kanas 2001; Loukeris and Eleftheriadis 2015; Abedin et al. 2019; Zhao et al. 2018;
Backpropagation Neural Network (BPNN)	8	Qi and Maddala 1999; Lahmiri and Bekiros 2019; Moshiri and Cameron 2000; Pichl and Kaizoji 2017; Yang et al. 1999; Sermpinis et al. 2013; Hamdi and Aloui 2015; Amelot et al. 2021;
Recurrent Neural Network (RNN)	5	Dunis et al. 2013; Dunis et al. 2010; Sermpinis et al. 2013; Zhao et al. 2018;
Cascade-correlation Neural network (CASCOR)	3	Altman et al. 1994; Coats and Fant 1993; Reber 2014;
Higher order Neural network (HONN)	3	Dunis et al. 2013; Dunis et al. 2010; Sermpinis et al. 2013;
Long Short-Term Memory Networks (LSTM)	2	Zhang et al. 2021; Bucci 2020;
Radial basis function Network (RBFN)	2	Lahmiri and Bekiros 2019; Episcopos, Pericli, and Hu, 1998;
Principal component combining/ analysis	2	Vortelinos 2017; Amelot et al. 2021;
Feed-forward neural network	1	Hamdi and Aloui 2015;

Table 6 (continued)

Method	n. of articles	Author (s) / Years
Generative Bayesian Neural network (GBNN)	1	Jang and Lee 2019;
NARX-Neural Network	1	; Xu and Zhao 2022; Amelot et al. 2021;
Fixed geometry neural networks (FGNN)	1	Nag and Mitra 2002;
Genetic algorithm neural networks (GANN)	1	Nag and Mitra 2002;
Psi sigma Neural Network	1	Dunis et al. 2010;
Probabilistic Neural Network	1	Yang et al. 1999;
Leaning vector quantization Neural network (LVQ)	1	Jagric, Jagric, and Kracun, 2011;
Deep learning (Deep neural networks / deep convolution neural network CNN)	12	Galeshchuk and Mukherjee 2017; Tashiro et al. 2019; Sun and Vasarhelyi 2018; Sirignano 2018; Dixon et al. 2017; Culkin and Das 2017; Cao et al. 2022; Kim and Kim 2020; Chen and Ge 2021; Lahmiri and Bekiros 2019; Abdou et al. 2021; Lahmiri and Bekiros 2019;
Hybrid Methods	3	Mselmi et al. 2017; Wanke et al. 2016a, b, c, d ;
Multi criteria decision-making (TOPSIS) combined with Neural Network	2	Wanke et al. 2016a, b, c, d ;
Partial least squares regression (PLS + Support Vector Machine)	1	Mselmi et al. 2017
REGRESSION (Panel regression, linear regression, multivariate regression, cross-sectional, OLS, multivariate adaptive regression splines MARS)	9	Qi 1999; Reber 2014; Scholtus et al. 2014; Qi and Maddala 1999; Calomiris and Mamaysky 2019; Jain et al. 2021; Kim and Kim 2014; Hendershott et al. 2011; Kamiya et al. 2018; D'Hondt et al. 2020; Jain et al. 2021; Kim and Kim 2014; Hendershott et al. 2011;
Logit	17	Jones et al. 2015; Jones et al. 2017; Jagric, Jagric, and Kracun, 2011; Butaru et al. 2016; Sun and Vasarhelyi 2018; Le and Viviani 2018; Lu et al. 2013; Rodrigues and Stevenson 2013; Deku et al. 2020; Kamiya et al. 2018; Kumar et al. 2019; Cao et al. 2022; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Mselmi et al. 2017; Episcopos, Pericli, and Hu, 1998; Creamer 2012; Holopainen and Sarlin 2017;

Table 6 (continued)

Method	n. of articles	Author (s) / Years
Autoregressive models (Arma, Arima, Garch, Har, Var)[a]	14	Zhao et al. 2018; Galeshchuk and Mukherjee 2017; Moshiri and Cameron 2000; Amelot et al. 2021; Herdershott, Jones, and Menkveld, 2011; Chabou et al., 2014; Frino et al. 2017; Parot et al. 2019; Calomiris and Mamaysky 2019; Reboredo et al. 2012; Vortelinos 2017; Fernandes et al. 2014; Bucci 2020; Jones et al. 2015;
Linear Discriminant Analysis (LDA)	7	Altman et al. 1994; Jones et al. 2017; Varetto 1998; Lahmiri 2016; Le and Viviani 2018; Cortés et al. 2008; Holopainen and Sarlin 2017;
Probit	3	Jones et al. 2015; Jones et al. 2017; Lahmiri 2016;
Nonlinear autoregressive exogenous model (NARX)	2	Amelot et al. 2021; Bucci 2020;
Multiple discriminant analysis (MDA)	2	Coats and Fant 1993; Pompe and Bilderbeek 2005;
Exponential smoothing (ETS)	1	Galeshchuk and Mukherjee 2017;
Generalised additive model (GAM)	2	Jones et al. 2015; Petukhina et al. 2020;
Other		
Spatial model	2	Litzenberger et al. 2012; Kelejian and Mukerji 2016;
Data mining/Text mining/Text analysis	6	Wei et al. 2019; Lu et al. 2013; Huang and Kuan 2021; Li et al. 2020; Trinkle and Baldwin 2016; Gepp, 2018; Yin et al. 2020;
Sentiment analysis	2	Huang and Kuan 2021; Houlihan and Creamer 2021;
Natural Language Processing	1	Calomiris and Mamaysky 2019;
Asset pricing models	1	Tao et al. 2021;
Image processing	1	Kamiya, 2018;
Grey relational analysis	1	Chen et al. 2013;
Random walk	2	Qi and Maddala 1999; Reboredo et al. 2012;
Type of Method/Paper		
Qualitative Paper	3	Huang and Kuan 2021; Wei et al. 2019; Xu et al. 2019;

Table 6 (continued)

Method	n. of articles	Author (s) / Years
Quantitative Paper	103	Hendershott et al. 2011; Calomiris and Mamaysky 2019; Altman et al. 1994; Chabou et al., 2014; Khandani et al. 2010; Coats and Fant 1993; Jones et al. 2015; Jones et al. 2017; Galeshchuk and Mukherjee 2017; Butaru et al. 2016; Varetto 1998; Trippi and DeSieno 1992; Jain, 2005; Lahmiri 2016; Kim and Kim 2014; Kercheval and Zhang 2015; Sun and Vasarhelyi 2018; Gepp et al. 2010; Dunis et al. 2013; Dunis et al. 2010; Yang et al. 1999; Sermpinis et al. 2013; Mirmirani and Li, 2004; Scholtus et al. 2014; Sirignano 2018; Qi and Maddala 1999; Feldman and Gross 2005; Reboredo et al. 2012; Fernandes et al. 2014; Wanke et al. 2016a, b, c, d ; Kelejian and Mukerji 2016; Guotai and Abedin 2017; Frino et al. 2017; Le and Viviani 2018; Parot et al. 2019; LeBaron, Arthur, and Palmer, 1999; Qi 1999; Moshiri and Cameron 2000; Nag and Mitra 2002; Pompe and Bilderbeek 2005; Cortés et al. 2008; Jagric, Jagric, and Kracun, 2011; Lu et al. 2013; Rodrigues and Stevenson 2013; Chen et al. 2013; Dixon et al. 2017; Culkin and Das 2017; Law and Shawe-Taylor 2017; Pichl and Kaizoji 2017; Vortelinos 2017; Renault 2017; Dubey et al. 2017; Jiang and Jones 2018; Sabau, Popa et al., 2021; Zhang et al. 2021; Chen and Ge 2021; Deku et al. 2020; Burggraf 2021; Huang and Guo 2021; Chen and Wan, 2020; Petukhina et al. 2020; Houlihan and Creamer, 2019; Tao et al. 2021; Corazza et al. 2021; Amelot et al. 2021; Jain et al. 2021; Caglayan et al. 2020; Calomiris and Mamaysky 2019; Kamiya et al. 2018; Bekiros and Georgoutsos 2008; Reber 2014; Kumar et al. 2019; Heston and Sinha 2017; Rasekhschaffe and Jones 2019; Cao et al. 2022; Abdou et al. 2021;; Xu and Zhao 2022; Papadimitriou, Goga and Agrapetidou, 2020; Soleymani and Vasighi 2020; Uddin et al. 2020; Yin et al. 2020; Durango-Gutiérrez, Lara-Rubio, and Navarro-Galera, 2021; Kanas 2001; Loukeris and Eleftheriadis 2015; Abedin et al. 2019; Mselmi et al. 2017; D'Hondt et al. 2020; Bucci 2020; Gao et al. 2016; Frino et al. 2020; Wanke et al. 2016a, b, c, d ; Jones and Wang 2019; Episcopos, Pericli, and Hu, 1998; Funahashi 2020; Lu and Ohta 2003; Creamer and Freund 2010; Creamer 2012; Booth et al. 2015; Holopainen and Sarlin 2017; Zhao et al. 2018; Jang and Lee 2019; Tashiro et al. 2019; Sariev and Germano 2020; Kim and Kim 2020; Li et al. 2020;
Literature Review	4	Litzenberger et al. 2012; Trinkle and Baldwin 2016; Lahmiri and Bekiros 2019; Hamdi and Aloui 2015;

models and regressions, such as ARMA, ARIMA, and GARCH. Finally, we observe that almost all the sampled papers are quantitative, whilst only three of them are qualitative and four of them consist in literature reviews.

A taxonomy of AI applications in Finance

After scrutinising some relevant features of the papers, we make a step forward and outline a taxonomy of AI applications used in Finance and tackled by previous literature. The main uses of AI in Finance and the papers that address each of them are summarised in Table 7.

Many research papers (39 out of 110) employ AI as a predictive instrument for forecasting stock prices, performance and volatility. In 23 papers, AI is employed in classification problems and warning systems to detect credit risk and frauds, as well as to monitor firm or bank performance. The former use of AI permits to classify firms into two categories based on qualitative and quantitative data; for example, we may have distressed or non-distressed, viable–nonviable, bankrupt–non-bankrupt, or financially healthy–not healthy, good–bad, and fraud–not fraud. Warning systems follow a similar principle: after analysing customers' financial behaviour and classifying potential fraud issues in bank accounts, alert models signal to the bank unusual transactions. Additionally, we see that 14 articles employ text mining and data mining language recognition, i.e. natural language processing, as well as sentiment analysis. This may be the starting point of AI-driven behavioural analysis in Finance. Amongst others, trading models and algorithmic trading are further popular aspects of AI widely analysed in the literature. Moreover, interest in Robo-advisory is growing in the asset investment field. Finally, less studied AI applications concern the modelling capability of algorithms and traditional machine learning and neural networks.

Identification of the major research streams

Drawing upon the co-citation analysis mentioned in Sect. "Methodology", we detected ten main research streams: (1) AI and the stock market; (2) AI and Trading Models; (3) AI and Volatility Forecasting; (4) AI and Portfolio Management; (5) AI and Performance, Risk, and Default Valuation; (6) AI and Bitcoin, Cryptocurrencies; (7) AI and Derivatives; (8) AI and Credit Risk in Banks; (9) AI and Investor Sentiments Analysis; (10) AI and Foreign Exchange Management. Some research streams can be further divided into sub-streams as they deal with various aspects of the same main topic. In this section, we provide a compact account for each of the aforementioned research streams. More detailed information on some of the papers fuelling them is provided in Appendix 2.

Stream 01: AI and the stock market

The stream "AI and the Stock Market" comprises two sub-streams, namely algorithmic trading and stock market, and AI and stock price prediction. The first sub-stream deals with the impact of algorithmic trading (AT) on financial markets. In

Table 7 Main uses of AI in finance addressed by the literature

AI applications	n. of articles	Authors(s) / Years
Predictive/ forecasting systems	39	Jones et al. 2017; Yang et al. 1999; Sun and Vasarhelyi 2018; Gepp et al. 2010; Dunis et al. 2013; Qi, and Maddala 1999; Reboredo et al. 2012; Fernandes et al. 2014; Wanke et al. 2016a, b, c, d ; Le and Viviani 2018; Parot et al. 2019; Moshiri and Cameron 2000; Nag and Mitra 2002; Rodrigues and Stevenson 2013; Chen et al. 2013; Trinkle and Baldwin 2016; Dixon et al. 2017; Law and Shawe-Taylor 2017; Pichl and Kaizoji 2017; Vortelinos 2017; Lahmiri and Bekiros 2019; Sabău Popa et al. 2021; Zhang et al. 2021; Houlihan and Creamer 2021; Caglayan et al. 2020; Bekiros and Georgoutsos 2008; Dunis et al. 2010; Sermpinis et al. 2013; Heston and Sinha 2017; Loukeris and Eleftheriadis 2015; Abedin et al. 2019; Bucci 2020; Jones and Wang 2019; Episcopos, Pericli and Hu, 1998; Booth et al. 2015; Tashiro et al. 2019; Kim and Kim 2014; Papadimitriou, Goga, and Agrapetidou, 2020
Classification / detection / early warning systems	23	Altman et al. 1994; Coats and Fant 1993; Khandani et al. 2010; Jones et al. 2017; Jones et al. 2015; Butaru et al. 2016; Varetto 1998; Feldman and Gross 2005; Jagric, Jagric and Kracun, 2011; Lu et al. 2013; Jiang and Jones 2018; Huang and Guo 2021; Deku et al. 2020; Corazza et al. 2021; Kumar et al. 2019; Durango-Gutiérrez, Lara-Rubio and Navarro-Galera, 2021; Loukeris and Eleftheriadis 2015; Mselmi et al. 2017; Holopainen and Sarlin 2017; Renault 2017; Le and Viviani 2018; Lahmiri 2016; Xu et al. 2019
Big data Analytics / Data mining / Text mining	14	Houlihan and Creamer 2021; Huang and Kuan 2021; Abdou et al., 2020; Kanas 2001; Durango-Gutiérrez, Lara-Rubio and Navarro-Galera, 2021; Wanke et al. 2016a, b, c, d ; Lu and Ohta 2003; Li et al. 2020; Kamiya et al. 2018; Renault 2017; Heston and Sinha 2017;; Xu and Zhao 2022; Yin et al. 2020; Xu et al. 2019;
Algorithmic trading/ Trading models	12	Hendershott et al. 2011; Chaboud et al. 2014; Scholtus et al. 2014; Kelejian and Mukerji 2016; Frino et al. 2017; Kelejian and Mukerji 2016; Litzenberger et al. 2012; Petukhina et al. 2020; Jain et al. 2021; Gao et al. 2016; Frino et al. 2020; Creamer 2012;
Natural Language processing/ sentiment analysis	9	Kim and Kim 2014; Wei et al. 2019; Calomiris and Mamaysky 2019; Heston and Sinha 2017; Renault 2017; Heston and Sinha 2017;; Xu and Zhao 2022; Yin et al. 2020; Houlihan and Creamer 2021;
Artificial Neural Networks	8	Reber 2014; Funahashi 2020; Zhao et al. 2018; Jang and Lee 2019; Sariev and Germano 2020; Loukeris and Eleftheriadis 2015; Heston and Sinha 2017; Dunis et al. 2010;
Robo-advisory	7	Trippi and DeSieno 1992; Rodrigues and Stevenson 2013; Petukhina et al. 2020; Tao et al. 2021; Loukeris and Eleftheriadis 2015; D'Hondt et al. 2020; Creamer 2012; Creamer and Freund 2010;

Table 7 (continued)

AI applications	n. of articles	Authors(s) / Years
Modelling	6	Fernandes et al. 2014; Guotai and Abedin 2017; Chen and Wan 2021; Amelot et al. 2021; Dunis et al. 2010; Funahashi 2020;
Machine learning	5	Rasekhschaffe and Jones 2019; Kercheval and Zhang 2015; Soleymani and Vasighi 2020; Burggraf 2021; Xu et al. 2019
Deep Learning	5	Culkin and Das 2017; Dixon et al. 2017; Kim and Kim 2020; Chen and Ge 2021; Galeshchuk and Mukherjee 2017
Digitalization / digital technology	1	Lu and Ohta 2003

this regard, Herdershott et al. (2011) argue that AT increases market liquidity by reducing spreads, adverse selection, and trade-related price discovery. This results in a lowered cost of equity for listed firms in the medium–long term, especially in emerging markets (Litzenberger et al. 2012). As opposed to human traders, algorithmic trading adjusts faster to information and generates higher profits around news announcements thanks to better market timing ability and rapid executions (Frino et al. 2017). Even though high-frequency trading (a subset of algorithmic trading) has sometimes increased volatility related to news or fundamentals, and transmitted it within and across industries, AT has overall reduced return volatility variance and improved market efficiency (Kelejian and Mukerji 2016; Litzenberger et al. 2012).

The second sub-stream investigates the use of neural networks and traditional methods to forecast stock prices and asset performance. ANNs are preferred to linear models because they capture the non-linear relationships between stock returns and fundamentals and are more sensitive to changes in variables relationships (Kanas 2001; Qi 1999). Dixon et al. (2017) argue that deep neural networks have strong predictive power, with an accuracy rate equal to 68%. Also, Zhang et al. (2021) propose a model, the Long Short-Term Memory Networks (LSTM), that outperforms all classical ANNs in terms of prediction accuracy and rational time cost, especially when various proxies of online investor attention (such as the internet search volume) are considered.

Stream 02: AI and trading models

From the review of the literature represented by this stream, it emerges that neural networks and machine learning algorithms are used to build intelligent automated trading systems. To give some examples, Creamer and Freund (2010) create a machine learning-based model that analyses stock price series and then selects the best-performing assets by suggesting a short or long position. The model is also equipped with a risk management overlayer preventing the transaction when the trading strategy is not profitable. Similarly, Creamer (2012) uses the above-mentioned logic in high-frequency trading futures: the model selects the most profitable and less risky futures by sending a long or short recommendation. To construct an efficient trading model, Trippi and DeSieno (1992) combine several neural networks into a single decision rule system that outperforms the single neural networks; Kercheval and Zhang (2015) use a supervised learning method (i.e. multi-class SVM) that automatically predicts mid-price movements in high-frequency limit order books by classifying them in low-stationary-up; these predictions are embedded in trading strategies and yield positive payoffs with controlled risk.

Stream 03: AI and volatility forecasting

The third stream deals with AI and the forecasting of volatility. The volatility index (VIX) from Chicago Board Options Exchange (CBOE) is a measure of market sentiment and expectations. Forecasting volatility is not a simple task because of its very persistent nature (Fernandes et al. 2014). According to Fernandes and co-authors, the VIX is negatively related to the SandP500 index return and positively related to its volume. The heterogeneous autoregressive (HAR) model yields the best predictive results as opposed to classical neural networks (Fernandes et al. 2014; Vortelinos 2017). Modern neural networks, such as LSTM and NARX (nonlinear autoregressive exogenous network), also qualify as valid alternatives (Bucci 2020). Another promising class of neural networks is the higher-order neural network (HONN) used to forecast the 21-day-ahead realised volatility of FTSE100 futures. Thanks to its ability to capture higher-order correlations within the dataset, HONN shows remarkable performance in terms of statistical accuracy and trading efficiency over multi-layer perceptron (MLP) and the recurrent neural network (RNN) (Sermpinis et al. 2013).

Stream 04: AI and portfolio management

This research stream analyses the use of AI in portfolio selection. As an illustration, Soleymani and Vasighi (2020) consider a clustering approach paired with VaR analysis to improve asset allocation: they group the least risky and more profitable stocks and allocate them in the portfolio. More elaborate asset allocation designs incorporate a bankruptcy detection model and an advanced utility performance system: before adding the stock to the portfolio, the sophisticated neural network estimates the default probability of the company and asset's contribution to the optimal portfolio (Loukeris and Eleftheriadis 2015). Index-tracking powered by deep learning technology minimises tracking error and generates positive performance (Kim and Kim 2020). The asymmetric copula method for returns dependence estimates further promotes the portfolio optimization process (Zhao et al. 2018). To sum up, all papers show that AI-based prediction models improve the portfolio selection process by accurately forecasting stock returns (Zhao et al. 2018).

Stream 05: AI and performance, risk, default valuation

This research stream comprises three sub-streams, namely AI and Corporate Performance, Risk and Default Valuation; AI and Real Estate Investment Performance, Risk, and Default Valuation; AI and Banks Performance, Risk and Default Valuation.

The first sub-stream examines corporate financial conditions to predict financially distressed companies (Altman et al. 1994). As an illustration, Jones et al. (2017) and Gepp et al. (2010) determine the probability of corporate default. Sabău Popa et al. (2021) predict business performance based on a composite financial index. The findings of the aforementioned papers confirm that AI-powered classifiers are extremely accurate and easy to interpret, hence, superior to classic linear models. A quite interesting paper surveys the relationship between face masculinity traits in CEOs and firm riskiness through image processing (Kamiya et al. 2018). The results reveal that firms lead by masculine-faced CEO have higher risk and leverage ratios and are more frequent acquirers in MandA operations.

The second sub-stream focuses on mortgage and loan default prediction (Feldman and Gross 2005; Episcopos, Pericli, and Hu, 1998). For instance, Chen et al. (2013) evaluate real estate investment returns by forecasting the REIT index; they show that the industrial production index, the lending rate, the dividend yield and the stock index influence real estate investments. All the forecasting techniques adopted (i.e. supervised machine learning and ANNs) outperform linear models in terms of efficiency and precision.

The third sub-stream deals with banks' performance. In contradiction with past research, a text mining study argues that the most important risk factors in banking are non-financial, i.e. regulation, strategy and management operation. However, the findings from text analysis are limited to what is disclosed in the papers (Wei et al. 2019). A highly performing NN-based study on the Malaysian and Islamic banking sector asserts that negative cost structure, cultural aspects and regulatory barriers (i.e. low competition) lead to inefficient banks compared to the U.S., which, on the contrary, are more resilient, healthier and well regulated (Wanke et al. 2016a, b, c, d; Papadimitriou et al. 2020).

Stream 06: AI and cryptocurrencies

Although algorithms and AI advisors are gaining ground, human traders still dominate the cryptocurrency market (Petukhina et al. 2021). For this reason, substantial arbitrage opportunities are available in the Bitcoin market, especially for USD–CNY and EUR–CNY currency pairs (Pichl and Kaizoji 2017). Concerning daily realised volatility, the HAR model delivers good results. Likewise, the feed-forward neural network effectively approximates the daily logarithmic returns of BTCUSD and the shape of their distribution (Pichl and Kaizoji 2017).

Additionally, the Hierarchical Risk Parity (HRP) approach, an asset allocation method based on machine learning, represents a powerful risk management tool able to manage the high volatility characterising Bitcoin prices, thereby helping cryptocurrency investors (Burggraf 2021).

Stream 07: AI and derivatives

ANNs and machine learning models are accurate predictors in pricing financial derivatives. Jang and Lee (2019) propose a machine learning model that outperforms traditional American option pricing models: the generative Bayesian NN; Culkin and Das (2017) use a feed-forward deep NN to reproduce Black and Scholes' option pricing formula with a high accuracy rate. Similarly, Chen and Wan (2021) suggest a deep NN for American option and deltas pricing in high dimensions. Funahashi (2020), on the contrary, rejects deep learning for option pricing due to the instability of the prices, and introduces a new hybrid method that combines ANNs and asymptotic expansion (AE). This model does not directly predict the option price but measures instead, the difference between the target (i.e. derivative price) and its approximation. As a result, the ANN becomes faster, more accurate and "lighter" in terms of layers and training data volume. This innovative method mimics a human learning process when one learns about a new object by recognising its differences from a similar and familiar item (Funahashi 2020).

Stream 08: AI and credit risk in banks

The research stream labelled “AI and Credit Risk in Banks”[2] includes the following sub-streams: AI and Bank Credit Risk; AI and Consumer Credit Risk and Default; AI and Financial Fraud detection/ Early Warning System; AI and Credit Scoring Models.

The first sub-stream addresses bank failure prediction. Machine learning and ANNs significantly outperform statistical approaches, although they lack transparency (Le and Viviani 2018). To overcome this limitation, Durango-Gutiérrez et al. (2021) combine traditional methods (i.e. logistic regression) with AI (i.e. Multiple layer perceptron -MLP), thus gaining valuable insights on explanatory variables. With the scope of preventing further global financial crises, the banking industry relies on financial decision support systems (FDSSs), which are strongly improved by AI-based models (Abedin et al. 2019).

The second sub-stream compares classic and advanced consumer credit risk models. Supervised learning tools, such as SVM, random forest, and advanced decision trees architectures, are powerful predictors of credit card delinquency: some of them can predict credit events up to 12 months in advance (Lahmiri 2016; Khandani et al. 2010; Butaru et al. 2016). Jagric et al. (2011) propose a learning vector quantization (LVQ) NN that better deals with categorical variables, achieving an excellent classification rate (i.e. default, non-default). Such methods overcome logit-based approaches and result in cost savings ranging from 6% up to 25% of total losses (Khadani et al. 2010).

The third group discusses the role of AI in early warning systems. On a retail level, advanced random forests accurately detect credit card fraud based on customer financial behaviour and spending pattern, and then flag it for investigation (Kumar et al. 2019). Similarly, Coats and Fant (1993) build a NN alert model for distressed firms that outperforms linear techniques. On a macroeconomic level, systemic risk monitoring models enhanced by AI technologies, i.e. k-nearest neighbours and sophisticated NNs, support macroprudential strategies and send alerts in case of global unusual financial activities (Holopainen, and Sarlin 2017; Huang and Guo 2021). However, these methods are still work-in-progress.

The last group studies intelligent credit scoring models, with machine learning systems, Adaboost and random forest delivering the best forecasts for credit rating changes. These models are robust to outliers, missing values and overfitting, and require minimal data intervention (Jones et al. 2015). As an illustration, combining data mining and machine learning, Xu et al. (2019) build a highly sophisticated model that selects the most important predictors and eliminates noisy variables, before performing the task.

[2] Since credit risk in the banking industry remarkably differs from credit risk in firms, the two of them are treated separately.

Stream 09: AI and investor sentiment analysis

Investor sentiment has become increasingly important in stock prediction. For this purpose, sentiment analysis extracts investor sentiment from social media platforms (e.g. StockTwits, Yahoo-finance, eastmoney.com) through natural language processing and data mining techniques, and classifies it into negative or positive (Yin et al. 2020). The resulting sentiment is regarded either as a risk factor in asset pricing models, an input to forecast asset price direction, or an intraday stock index return (Houlihan and Creamer 2021; Renault 2017). In this respect, Yin et al. (2020) find that investor sentiment has a positive correlation with stock liquidity, especially in slowing markets; additionally, sensitivity to liquidity conditions tends to be higher for firms with larger size and a higher book-to-market ratio, and especially those operating in weakly regulated markets. As for predictions, daily news usually predicts stock returns for few days, whereas weekly news predicts returns for longer period, from one month to one quarter. This generates a return effect on stock prices, as much of the delayed response to news occurs around major events in company life, specifically earnings announcement, thus making investor sentiment a very important variable in assessing the impact of AI in financial markets. (Heston and Sinha 2017).

Stream 10: AI and foreign exchange management

The last stream addresses AI and the management of foreign exchange. Cost-effective trading or hedging activities in this market require accurate exchange rate forecasts (Galeshchuk and Mukherjee 2017). In this regard, the HONN model significantly outperforms traditional neural networks (i.e. multi-layer perceptron, recurrent NNs, Psi sigma-models) in forecasting and trading the EUR/USD currency pair using ECB daily fixing series as input data (Dunis et al. 2010). On the contrary, Galeshchuk and Mukherjee (2017) consider these methods as unable to predict the direction of change in the forex rates and, therefore, ineffective at supporting profitable trading. For this reason, they apply a deep NN (Convolution NNs) to forecast three main exchange rates (i.e. EUR/USD, GBP/USD, and JPY/USD). The model performs remarkably better than time series models (e.g. ARIMA: Autoregressive integrated moving average) and machine learning classifiers. To sum up, from this research stream it emerges that AI-based models, such as NARX and the above-mentioned techniques, achieve better prediction performance than statistical or time series models, as remarked by Amelot et al. (2021).

Issues that deserve further investigation

As shown in Sect. "A detailed account of the literature on AI in Finance", the literature on Artificial Intelligence in Finance is vast and rapidly growing as technological progress advances. There are, however, some aspects of this subject that are unexplored yet or that require further investigation. In this section, we further scrutinise, through content analysis, the papers published between 2015 and 2021 (as we want

Table 8 Research questions for future research

Research streams	Research questions	Authors (s) / Year
AI and Stock Market	Which AI-based technique (e.g. ML, clustering algorithms) is the best for Stock market prediction?	Law, and Shawe-Taylor (2017)
	Which kind of order book information best improves the accuracy of AI-based models for stock market prediction?	Tashiro, et al. (2019)
	How does policy and regulation impact Algorithmic trading?	Litzenberger et al. (2012)
	What effect have market cycles on the accuracy of intelligent stock price prediction models? Can it be leveraged to improve the model's performance?	Booth et al. (2015)
AI and Trading Models	How do Robo advisors perform during major unexpected financial crisis such as COVID-19?	Tao et al. (2021)
	Can limit order books data embedded in AI-based techniques boost trading models accuracy?	Sirignano (2018)
AI and Volatility Forecasting	Do more elaborated neural network architectures enhance realised volatility prediction? What are the benefits and results of using NNs multivariate time series in forecasting realised volatility?	Bucci (2020)
	Which AI optimising algorithms most improve index-tracking portfolio strategy?	Kim and Kim (2020)
AI and Portfolio Management	Which machine learning approach (e.g. fuzzy clustering) best improves portfolio construction?	Soleymani and Vasighi (2020)
	How can deep learning techniques contribute to volatility forecasting for portfolio selection?	Chen and Ge (2021)

Table 8 (continued)

Research streams	Research questions	Authors (s) / Year
AI and Performance, Risk, and Default Valuation	How would multiple classifiers based on AI technology perform compared to binary classifiers in predicting corporate bankruptcy, bond default, corporate mergers, reconstructions, and takeovers? What are the benefits of combining sophisticated data mining techniques with experts' opinion in corporate default forecasts? What are possible solutions for transforming and manipulating missing data in AI predictive models?	Jones et al. (2017)
	What impact have corporate credit ratings and social media data on the accuracy of AI-powered risk predictors?	Uddin et al. (2020)
	Which AI tools help overcome ANNs limitations (e.g. overfitting, black box)?	Sariev and Germano (2020)
AI and Bitcoin, Cryptocurrency	Which AI techniques are best for the optimization of a cryptocurrency portfolio?	Burggraf (2021)
	What are future developments in the crypto market in terms of AI-based trading methods and blockchain? What impact has regulation and blockchain on crypto markets and AI models performance?	Petukhina et al. (2020)
AI and Derivatives	What are potential deployments and results of text-based input data and sentiment analysis in option pricing?	Jang and Lee (2019)
	What are the best designs of AI models that minimise computational cost?	Chen and Wan (2021)
	Are there further human learning paths to be implemented in AI technology?	Funahashi (2020)

Table 8 (continued)

Research streams	Research questions	Authors (s) / Year
AI and Credit Risk in Banks	What type of data (e.g. bank market data) best improves the result of bank default forecasting models?	Le and Viviani (2018)
	What methods reduce AI training speed and enhance classification accuracy?	Kumar et al. (2019)
	How can early warning models be further simplified to be widely implemented? Which AI technique is best for combining visual data or visual interfaces with systemic risk measurement to "visualise" and interact with future risk scenarios?	Holopainen and Sarlin (2017)
AI and Investor Sentiments Analysis	Can the combination of both textual data and market data improve AI predictive models in specific sectors and industries?	Houlihan and Creamer (2021)
	How do diverse types of news and "social" data impact financial markets? How does the market process that information?	Heston and Sinha (2017)
	Which AI model best captures the impact of social networks sites' sentiment (SNS) on individual stock for portfolio management? Does the increasing role of "influencers" in Finance (e.g. investor advisors, expert analysts) affect market returns and how can AI technology use it for financial forecasts?	Xu and Zhao (2020)
	What are the strategies to simplify and make machine learning leaner and faster? Which AI-based trading strategy best performs in the forex market during a financial crisis?	Galeshchuk and Mukherjee (2017)
AI and Foreign Exchange Management	Which AI model based on advanced time series (e.g. genetic algorithm (GA), hybrid genetic algorithm optimised long short-term memory, ETS models or APGARCH or hybrid ANN Gravitational models) is most performing in foreign exchange rates or stock market forecasting?	Amelot et al. (2021)

to focus on the most recent research directions) in order to define a potential research agenda. Hence, for each of the ten research streams presented in Sect. "Identification of the major research streams", we report a number of research questions that were put forward over time and are still at least partly unaddressed. The complete list of research questions is enclosed in Table 8.

AI and the stock market

This research stream focuses on algorithmic trading (AT) and stock price prediction. Future research in the field could analyse more deeply alternative AI-based market predictors (e.g. clustering algorithms and similar learning methods) and draw up a regime clustering algorithm in order to get a clearer view of the potential applications and benefits of clustering methodologies (Law, and Shawe-Taylor 2017). In this regard, Litzenberger et al. (2012) and Booth et al. (2015) recommend broadening the study to market cycles and regulation policies that may affect AI models' performance in stock prediction and algorithmic trading, respectively.[3] Furthermore, forecasting models should be evaluated with deeper order book information, which may lead to a higher prediction accuracy of stock prices (Tashiro et al. 2019).

AI and trading models

This research stream builds on the application of AI in trading models. Robo advisors are the evolution of basic trading models: they are easily accessible, cost-effective, profitable for investors and, unlike human traders, immune to behavioural biases. Robo advisory, however, is a recent phenomenon and needs further performance evaluations, especially in periods of financial distress, such as the post-COVID-19 one (Tao et al. 2021), or in the case of the so-called "Black swan" events. Conversely, trading models based on spatial neural networks (an advanced ANN) outperform all statistical techniques in modelling limit order books and suggest an extensive interpretation of the joint distribution of the best bid and best ask. Given the versatility of such a method, forthcoming research should resort to it with the aim of understanding whether neural networks with more order book information (i.e. order flow history) lead to better trading performance (Sirignano 2018).

AI and volatility forecasting

As previously mentioned, volatility forecasting is a challenging task. Although recent studies report solid results in the field (see Sermpinis et al. 2013; Vortelinos 2017), future work could deploy more elaborated recurrent NNs by modifying the activation function of the processing units composing the ANNs, or by adding hidden layers and then evaluate their performance (Bucci 2020). Since univariate time

[3] As this issue has not been addressed in the latest papers, we include these two papers although their year of publication lies outside the established range period.

series are commonly used for realised volatility prediction, it would be interesting to also inquire about the performance of multivariate time series.

AI and portfolio management

This research stream examines the use of AI in portfolio selection strategies. Past studies have developed AI models that are capable of replicating the performance of stock indexes (known as index tracking strategy) and constructing efficient portfolios with no human intervention. In this regard, Kim and Kim (2020) suggest focussing on optimising AI algorithms to boost index-tracking performance. Soleymani and Vasighi (2020) recognise the importance of clustering algorithms in portfolio management and propose a clustering approach powered by a membership function, also known as fuzzy clustering, to further improve the selection of less risky and most profitable assets. For this reason, analysis of asset volatility through deep learning should be embedded in portfolio selection models (Chen and Ge 2021).

AI and performance, risk, default valuation

Bankruptcy and performance prediction models rely on binary classifiers that only provide two outcomes, e.g. risky–not risky, default–not default, good–bad performance. These methods may be restrictive as sometimes there is not a clear distinction between the two categories (Jones et al. 2017). Therefore, prospective research might focus on multiple outcome domains and extend the research area to other contexts, such as bond default prediction, corporate mergers, reconstructions, takeovers, and credit rating changes (Jones et al. 2017). Corporate credit ratings and social media data should be included as independent predictors in credit risk forecasts to evaluate their impact on the accuracy of risk-predicting models (Uddin et al. 2020). Moreover, it is worth evaluating the benefits of a combined human–machine approach, where analysts contribute to variables' selection alongside data mining techniques (Jones et al. 2017). Forthcoming studies should also address black box and over-fitting biases (Sariev and Germano 2020), as well as provide solutions for the manipulation and transformation of missing input data relevant to the model (Jones et al. 2017).

AI and cryptocurrencies

The use of AI in the cryptocurrency market is in its infancy, and so are the policies regulating it. As the digital currency industry has become increasingly important in the financial world, future research should study the impact of regulations and blockchain progress on the performance of AI techniques applied in this field (Petukhina et al., 2021). Cryptocurrencies, and especially Bitcoins, are extensively used in financial portfolios. Hence, new AI approaches should be developed in order to optimise cryptocurrency portfolios (Burggraf 2021).

AI and derivatives

This research stream examines derivative pricing models based on AI. A valuable research area that should be further explored concerns the incorporation of text-based input data, such as tweets, blogs, and comments, for option price prediction (Jang and Lee 2019). Since derivative pricing is an utterly complicated task, Chen and Wan (2021) suggest studying advanced AI designs that minimise computational costs. Funahashi (2020) recognises a typical human learning process (i.e. recognition by differences) and applies it to the model, significantly simplifying the pricing problem. In the light of these considerations, prospective research may also investigate other human learning and reasoning paths that can improve AI reasoning skills.

AI and credit risk in banks

Bank default prediction models often rely solely on accounting information from banks' financial statements. To enhance default forecast, future work should consider market data as well (Le and Viviani 2018). Credit risk includes bank account fraud and financial systemic risk. Fraud detection based on AI needs further experiments in terms of training speed and classification accuracy (Kumar et al. 2019). Early warning models, on the other hand, should be more sensitive to systemic risk. For this reason, subsequent studies ought to provide a common platform for modelling systemic risk and visualisation techniques enabling interaction with both model parameters and visual interfaces (Holopainen and Sarlin 2017).

AI and investor sentiment analysis

Sentiment analysis builds on text-based data from social networks and news to identify investor sentiment and use it as a predictor of asset prices. Forthcoming research may analyse the effect of investor sentiment on specific sectors (Houlihan and Creamer 2021), as well as the impact of diverse types of news on financial markets (Heston and Sinha 2017). This is important for understanding how markets process information. In this respect, Xu and Zhao (2022) propose a deeper analysis of how social networks' sentiment affects individual stock returns. They also believe that the activity of financial influencers, such as financial analysts or investment advisors, potentially affects market returns and needs to be considered in financial forecasts or portfolio management.

AI and foreign exchange management

This research stream investigates the application of AI models to the Forex market. Deep networks, in particular, efficiently predict the direction of change in forex rates thanks to their ability to "learn" abstract features (i.e. moving averages) through hidden layers. Future work should study whether these abstract features can be inferred from the model and used as valid input data to simplify the deep network structure

(Galeshchuk and Mukherjee 2017). Moreover, the performance of foreign exchange trading models should be assessed in financial distressed times. Further research may also compare the predictive performance of advanced times series models, such as genetic algorithms and hybrid NNs, for forex trading purposes (Amelot et al. 2021).

Conclusions

Despite its recent advent, Artificial Intelligence has revolutionised the entire financial system, thanks to advanced computer science and Big Data Analytics and the increasing outflow of data generated by consumers, investors, business, and governments' activities. Therefore, it is not surprising that a growing strand of literature has examined the uses, benefits and potential of AI applications in Finance. This paper aims to provide an accurate account of the state of the art, and, in doing so, it would represent a useful guide for readers interested in this topic and, above all, the starting point for future research. To this purpose, we collected a large number of articles published in journals indexed in Web of Science (WoS), and then resorted to both bibliometric analysis and content analysis. In particular, we inspected several features of the papers under study, identified the main AI applications in Finance and highlighted ten major research streams. From this extensive review, it emerges that AI can be regarded as an excellent market predictor and contributes to market stability by minimising information asymmetry and volatility; this results in profitable investing systems and accurate performance evaluations. Additionally, in the risk management area, AI aids with bankruptcy and credit risk prediction in both corporate and financial institutions; fraud detection and early warning models monitor the whole financial system and raise expectations for future artificial market surveillance. This suggests that global financial crises or unexpected financial turmoil will be likely to be anticipated and prevented.

All in all, judging from the rapid widespread of AI applications in the financial sphere and across a large variety of countries, and, more in general, based on the growth rate exhibited by technological progress over time, we expect that the use of AI tools will further expand, both geographically, across sectors and across financial areas. Hence, firms that still struggle with coping with the latest wave of technological change should be aware of that, and try to overcome this burden in order to reap the potential benefits associated with the adoption of AI and remain competitive. In the light of these considerations, policymakers should motivate companies, especially those that have not adopted yet, or have just begun to introduce AI applications, to catch up, for instance by providing funding or training courses aimed to strengthen the complex skills required by employees dealing with these sophisticated systems and languages.

This study presents some limitations. For instance, it tackles a significant range of interrelated topics (in particular, the main financial areas affected by AI which have been the main object of past research), and then presents a concise description for each of them; other studies may decide to focus on only one or a couple of subjects and provide a more in-depth account of the chosen one(s). Also, we are aware that technological change has been progressing at an unprecedented fast and growing pace; even though we considered a significantly long time-frame and a relevant amount of studies have been released in the first two decades of the XXI century, we are aware that further advancements have been made from 2021 (the last year included in the time frame used to the select our sample); for instance, in the last few years, AI experts, policymakers, and also a growing number of scholars have been debating the potential and risks of AI-related devices, such as chatGBT and the broader and more elusive "metaverse" (see for instance Mondal et al. 2023 and Calzada 2023, for an overview). Hence, future contributions may advance our understanding of the implications of these latest developments for finance and other important fields, such as education and health.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43546-023-00618-x.

Author contributions Conceptualization: MC and SB. Methodology: SB. Investigation: MC, XG, SB. Writing: Marco Cucculelli, Xhoana Goga, Salman Bahoo and JM. Writing – Review and Editing: JM. Supervision: MC**.** Project Administration: MC.

Funding Open access funding provided by Università Politecnica delle Marche within the CRUI-CARE Agreement. This study has not received specific funding. We are granted with research funds by our institution which would allow us to cover the publication costs.

Data availability Full data are available from authors upon request.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

References

Abdou HA, Ellelly NN, Elamer AA, Hussainey K, Yazdifar H (2021) Corporate governance and earnings management Nexus: evidence from the UK and Egypt using neural networks. Int J Financ Econ 26(4):6281–6311. https://doi.org/10.1002/ijfe.2120

Abedin MZ, Guotai C, Moula F, Azad AS, Khan MS (2019) Topological applications of multilayer perceptrons and support vector machines in financial decision support systems. Int J Financ Econ 24(1):474–507. https://doi.org/10.1002/ijfe.1675

Acemoglu D, Restrepo P (2020) The wrong kind of AI? Artificial intelligence and the future of labor demand. Cambr J Reg Econ Soc, Cambr Pol Econ Soc 13(1):25–35

Altman EI, Marco G, Varetto F (1994) Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience). J Bank Finance 18(3):505–529. https://doi.org/10.1016/0378-4266(94)90007-8

Amelot LM, Subadar Agathee U, Sunecher Y (2021) Time series modelling, narx neural network and HYBRID kpca–svr approach to forecast the foreign exchange market in Mauritius. Afr J Econ Manag Stud 12(1):18–54. https://doi.org/10.1108/ajems-04-2019-0161

Bekiros SD, Georgoutsos DA (2008) Non-linear dynamics in financial asset returns: The predictive power of the CBOE volatility index. Eur J Fin 14(5):397–408. https://doi.org/10.1080/13518470802042203

Biju AKVN, Thomas AS, Thasneem J (2020) Examining the research taxonomy of artificial intelligence, deep learning & machine learning in the financial sphere—a bibliometric analysis. Qual Quant Online First. https://doi.org/10.1007/s11135-023-01673-0

Black F, Scholes M (1973) The pricing of Options and corporate liabilities. J Pol Econ 81(3):637–654

Booth A, Gerding E, McGroarty F (2015) Performance-weighted ensembles of random forests for predicting price impact. Quant Finance 15(11):1823–1835. https://doi.org/10.1080/14697688.2014.983539

Bresnahan TF, Trajtenberg M (1995) General purpose technologies 'Engines of growth'? J Econom 65(1):83–108. https://doi.org/10.1016/0304-4076(94)01598-T

Bucci A (2020) Realized volatility forecasting with neural networks. J Financ Econom 3:502–531. https://doi.org/10.1093/jjfinec/nbaa008

Buchanan, B. G. (2019). Artificial intelligence in finance - Alan Turing Institute. https://www.turing.ac.uk/sites/default/files/2019-04/artificial_intelligence_in_finance_-_turing_report_0.pdf.

Burggraf T (2021) Beyond risk parity – a machine learning-based hierarchical risk parity approach on cryptocurrencies. Finance Res Lett 38:101523. https://doi.org/10.1016/j.frl.2020.101523

Butaru F, Chen Q, Clark B, Das S, Lo AW, Siddique A (2016) Risk and risk management in the credit card industry. J Bank Finance 72:218–239. https://doi.org/10.1016/j.jbankfin.2016.07.015

Caglayan M, Pham T, Talavera O, Xiong X (2020) Asset mispricing in peer-to-peer loan secondary markets. J Corp Finan 65:101769. https://doi.org/10.1016/j.jcorpfin.2020.101769

Calomiris CW, Mamaysky H (2019) How news and its context drive risk and returns around the world. J Financ Econ 133(2):299–336. https://doi.org/10.1016/j.jfineco.2018.11.009

Calzada I (2023) Disruptive technologies for e-diasporas: blockchain, DAOs, data cooperatives, metaverse, and ChatGPT. Futures 154:103258. https://doi.org/10.1016/j.futures.2023.103258

Cao Y, Liu X, Zhai J, Hua S (2022) A Two-stage Bayesian network model for corporate bankruptcy prediction. Int J Financ Econ 27(1):455–472. https://doi.org/10.1002/ijfe.2162

Chaboud AP, Chiquoine B, Hjalmarsson E, Vega C (2014) Rise of the machines: Algorithmic trading in the foreign exchange market. J Financ 69(5):2045–2084. https://doi.org/10.1111/jofi.12186

Chen S, Ge L (2021) A learning-based strategy for portfolio selection. Int Rev Econ Financ 71:936–942. https://doi.org/10.1016/j.iref.2020.07.010

Chen Y, Wan JW (2021) Deep neural network framework based on backward stochastic differential equations for pricing and hedging American options in high dimensions. Quant Finance 21(1):45–67. https://doi.org/10.1080/14697688.2020.1788219

Chen J, Chang T, Ho C, Diaz JF (2013) Grey relational analysis and neural Network forecasting of reit returns. Quantitative Finance 14(11):2033–2044. https://doi.org/10.1080/14697688.2013.816765

Coats PK, Fant LF (1993) Recognizing financial distress patterns using a neural network tool. Financ Manage 22(3):142. https://doi.org/10.2307/3665934

Corazza M, De March D, Di Tollo G (2021) Design of adaptive Elman networks for credit risk assessment. Quantitative Finance 21(2):323–340. https://doi.org/10.1080/14697688.2020.1778175

Cortés EA, Martínez MG, Rubio NG (2008) FIAMM return persistence analysis and the determinants of the fees charged. Span J Finance Account Revis Esp De Financ Y Contab 37(137):13–32. https://doi.org/10.1080/02102412.2008.10779637

Creamer G (2012) Model calibration and automated trading agent for euro futures. Quant Finance 12(4):531–545. https://doi.org/10.1080/14697688.2012.664921

Creamer G, Freund Y (2010) Automated trading with boosting and expert weighting. Quant Finance 10(4):401–420. https://doi.org/10.1080/14697680903104113

Cucculelli M, Recanatini M (2022) Distributed Ledger technology systems in securities post-trading services. Evid Eur Global Syst Banks Eur J Finance 28(2):195–218. https://doi.org/10.1080/1351847X.2021.1921002

Culkin R, Das SR (2017) Machine learning in finance: The case of deep learning for option pricing. J Invest Management 15(4):92–100

D'Hondt C, De Winne R, Ghysels E, Raymond S (2020) Artificial intelligence alter egos: Who might benefit from robo-investing? J Empir Financ 59:278–299. https://doi.org/10.1016/j.jempfin.2020.10.002

Deku SY, Kara A, Semeyutin A (2020) The predictive strength of mbs yield spreads during asset bubbles. Rev Quant Financ Acc 56(1):111–142. https://doi.org/10.1007/s11156-020-00888-8

Dixon M, Klabjan D, Bang JH (2017) Classification-based financial markets prediction using deep neural networks. Algorithmic Finance 6(3–4):67–77. https://doi.org/10.3233/af-170176

Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2021) How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 133:285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Downe-Wamboldt B (1992) Content analysis: method, applications, and issues. Health Care Women Int 13(3):313–321. https://doi.org/10.1080/07399339209516006

Dubey RK, Chauhan Y, Syamala SR (2017) Evidence of algorithmic trading from Indian equity Market: Interpreting the transaction velocity element of financialization. Res Int Bus Financ 42:31–38. https://doi.org/10.1016/j.ribaf.2017.05.014

Dunis CL, Laws J, Sermpinis G (2010) Modelling and trading the EUR/USD exchange rate at the ECB fixing. Eur J Finance 16(6):541–560. https://doi.org/10.1080/13518470903037771

Dunis CL, Laws J, Karathanasopoulos A (2013) Gp algorithm versus hybrid and mixed neural networks. Eur J Finance 19(3):180–205. https://doi.org/10.1080/1351847x.2012.679740

Durango-Gutiérrez MP, Lara-Rubio J, Navarro-Galera A (2021) Analysis of default risk in microfinance institutions under the Basel Iii framework. Int J Financ Econ. https://doi.org/10.1002/ijfe.2475

Episcopos A, Pericli A, Hu J (1998) Commercial mortgage default: A comparison of logit with radial basis function networks. J Real Estate Finance Econ 17(2):163–178

Ernst, E., Merola, R., and Samaan, D. (2018). The economics of artificial intelligence: Implications for the future of work. ILO Futur Work Res Paper Ser No. 5.

Feldman D, Gross S (2005) Mortgage default: classification trees analysis. J Real Estate Finance Econ 30(4):369–396. https://doi.org/10.1007/s11146-005-7013-7

Fernandes M, Medeiros MC, Scharth M (2014) Modeling and predicting the CBOE market volatility index. J Bank Finance 40:1–10. https://doi.org/10.1016/j.jbankfin.2013.11.004

Frankenfield, J. (2021). How Artificial Intelligence Works. Retrieved June 11, 2021, from https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp

Frino A, Prodromou T, Wang GH, Westerholm PJ, Zheng H (2017) An empirical analysis of algorithmic trading around earnings announcements. Pac Basin Financ J 45:34–51. https://doi.org/10.1016/j.pacfin.2016.05.008

Frino A, Garcia M, Zhou Z (2020) Impact of algorithmic trading on speed of adjustment to new information: Evidence from interest rate derivatives. J Futur Mark 40(5):749–760. https://doi.org/10.1002/fut.22104

Funahashi H (2020) Artificial neural network for option pricing with and without asymptotic correction. Quant Finance 21(4):575–592. https://doi.org/10.1080/14697688.2020.1812702

Galeshchuk S, Mukherjee S (2017) Deep networks for predicting direction of change in foreign exchange rates. Intell Syst Account Finance Manage 24(4):100–110. https://doi.org/10.1002/isaf.1404

Gao M, Liu Y, Wu W (2016) Fat-finger trade and market quality: the first evidence from China. J Futur Mark 36(10):1014–1025. https://doi.org/10.1002/fut.21771

Gepp A, Kumar K, Bhattacharya S (2010) Business failure prediction using decision trees. J Forecast 29(6):536–555. https://doi.org/10.1002/for.1153

Guotai C, Abedin MZ (2017) Modeling credit approval data with neural networks: an experimental investigation and optimization. J Bus Econ Manag 18(2):224–240. https://doi.org/10.3846/16111699.2017.1280844

Hamdi M, Aloui C (2015) Forecasting crude oil price using artificial neural networks: a literature survey. Econ Bull 35(2):1339–1359

Hendershott T, Jones CM, Menkveld AJ (2011) Does algorithmic trading improve liquidity? J Financ 66(1):1–33. https://doi.org/10.1111/j.1540-6261.2010.01624.x

Hentzen JK, Hoffmann A, Dolan R, Pala E (2022a) Artificial intelligence in customer-facing financial services: a systematic literature review and agenda for future research. Int J Bank Market 40(6):1299–1336. https://doi.org/10.1108/IJBM-09-2021-0417
Hentzen JK, Hoffmann AOI, Dolan RM (2022b) Which consumers are more likely to adopt a retirement app and how does it explain mobile technology-enabled retirement engagement? Int J Consum Stud 46:368–390. https://doi.org/10.1111/ijcs.12685
Heston SL, Sinha NR (2017) News vs sentiment: predicting stock returns from news stories. Financial Anal J 73(3):67–83. https://doi.org/10.2469/faj.v73.n3.3
Holopainen M, Sarlin P (2017) Toward robust early-warning models: a horse race, ensembles and model uncertainty. Quant Finance 17(12):1933–1963. https://doi.org/10.1080/14697688.2017.1357972
Houlihan P, Creamer GG (2021) Leveraging social media to predict continuation and reversal in asset prices. Comput Econ 57(2):433–453. https://doi.org/10.1007/s10614-019-09932-9
Huang X, Guo F (2021) A kernel fuzzy twin SVM model for early warning systems of extreme financial risks. Int J Financ Econ 26(1):1459–1468. https://doi.org/10.1002/ijfe.1858
Huang Y, Kuan C (2021) Economic prediction with the fomc minutes: an application of text mining. Int Rev Econ Financ 71:751–761. https://doi.org/10.1016/j.iref.2020.09.020
IBM Cloud Education. (2020). What are Neural Networks? Retrieved May 10, 2021, from https://www.ibm.com/cloud/learn/neural-networks
Jagric T, Jagric V, Kracun D (2011) Does non-linearity matter in retail credit risk modeling? Czech J Econ Finance Faculty Soc Sci 61(4):384–402
Jagtiani J, Kose J (2018) Fintech: the impact on consumers and regulatory responses. J Econ Bus 100:1–6. https://doi.org/10.1016/j.jeconbus.2018.11.002
Jain A, Jain C, Khanapure RB (2021) Do algorithmic traders improve liquidity when information asymmetry is high? Q J Financ 11(01):1–32. https://doi.org/10.1142/s2010139220500159
Jang H, Lee J (2019) Generative Bayesian neural network model for risk-neutral pricing of American index options. Quant Finance 19(4):587–603. https://doi.org/10.1080/14697688.2018.1490807
Jiang Y, Jones S (2018) Corporate distress prediction in China: a machine learning approach. Account Finance 58(4):1063–1109. https://doi.org/10.1111/acfi.12432
Jones S, Wang T (2019) Predicting private company failure: a multi-class analysis. J Int Finan Markets Inst Money 61:161–188. https://doi.org/10.1016/j.intfin.2019.03.004
Jones S, Johnstone D, Wilson R (2015) An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. J Bank Finance 56:72–85. https://doi.org/10.1016/j.jbankfin.2015.02.006
Jones S, Johnstone D, Wilson R (2017) Predicting corporate bankruptcy: an evaluation of alternative statistical frameworks. J Bus Financ Acc 44(1–2):3–34. https://doi.org/10.1111/jbfa.12218
Kamiya S, Kim YH, Park S (2018) The face of risk: Ceo facial masculinity and firm risk. Eur Financ Manag 25(2):239–270. https://doi.org/10.1111/eufm.12175
Kanas A (2001) Neural network linear forecasts for stock returns. Int J Financ Econ 6(3):245–254. https://doi.org/10.1002/ijfe.156
Kelejian HH, Mukerji P (2016) Does high frequency algorithmic trading matter for non-at investors? Res Int Bus Financ 37:78–92. https://doi.org/10.1016/j.ribaf.2015.10.014
Kercheval AN, Zhang Y (2015) Modelling high-frequency limit order book dynamics with support vector machines. Quant Finance 15(8):1315–1329. https://doi.org/10.1080/14697688.2015.1032546
Khandani AE, Kim AJ, Lo AW (2010) Consumer credit-risk models via machine-learning algorithms. J Bank Finance 34(11):2767–2787. https://doi.org/10.1016/j.jbankfin.2010.06.001
Kim S, Kim D (2014) Investor sentiment from internet message postings and the predictability of stock returns. J Econ Behav Organ 107:708–729. https://doi.org/10.1016/j.jebo.2014.04.015
Kim S, Kim S (2020) Index tracking through deep latent representation learning. Quant Finance 20(4):639–652. https://doi.org/10.1080/14697688.2019.1683599
Kumar G, Muckley CB, Pham L, Ryan D (2019) Can alert models for fraud protect the elderly clients of a financial institution? Eur J Finance 25(17):1683–1707. https://doi.org/10.1080/1351847x.2018.1552603
Lahmiri S (2016) Features selection, data mining and financial risk classification: a comparative study. Intell Syst Account Finance Managed 23(4):265–275. https://doi.org/10.1002/isaf.1395

Lahmiri S, Bekiros S (2019) Can machine learning approaches predict corporate bankruptcy? evidence from a qualitative experimental design. Quant Finance 19(9):1569–1577. https://doi.org/10.1080/14697688.2019.1588468

Law T, Shawe-Taylor J (2017) Practical Bayesian support vector regression for financial time series prediction and market condition change detection. Quant Finance 17(9):1403–1416. https://doi.org/10.1080/14697688.2016.1267868

Le HH, Viviani J (2018) Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. Res Int Bus Financ 44:16–25. https://doi.org/10.1016/j.ribaf.2017.07.104

Li J, Li G, Zhu X, Yao Y (2020) Identifying the influential factors of commodity futures prices through a new text mining approach. Quant Finance 20(12):1967–1981. https://doi.org/10.1080/14697688.2020.1814008

Litzenberger R, Castura J, Gorelick R (2012) The impacts of automation and high frequency trading on market quality. Annu Rev Financ Econ 4(1):59–98. https://doi.org/10.1146/annurev-financial-110311-101744

Loukeris N, Eleftheriadis I (2015) Further higher moments in portfolio Selection and a priori detection of bankruptcy, under multi-layer perceptron neural Networks, HYBRID Neuro-genetic MLPs, and the voted perceptron. Int J Financ Econ 20(4):341–361. https://doi.org/10.1002/ijfe.1521

Lu J, Ohta H (2003) A data and digital-contracts driven method for pricing complex derivatives. Quant Finance 3(3):212–219. https://doi.org/10.1088/1469-7688/3/3/307

Lu Y, Shen C, Wei Y (2013) Revisiting early warning signals of corporate credit default using linguistic analysis. Pac Basin Financ J 24:1–21. https://doi.org/10.1016/j.pacfin.2013.02.002

Martinelli A, Mina A, Moggi M (2021) The enabling technologies of industry 4.0: examining the seeds of the fourth industrial revolution. Ind Corp Chang 2021:1–28. https://doi.org/10.1093/icc/dtaa060

Mondal S, Das S, Vrana VG (2023) How to bell the cat? a theoretical review of generative artificial intelligence towards digital disruption in all walks of life. Technologies 11(2):44. https://doi.org/10.3390/technologies11020044

Moshiri S, Cameron N (2000) Neural network versus econometric models in forecasting inflation. J Forecast 19(3):201–217. https://doi.org/10.1002/(sici)1099-131x(200004)19:33.0.co;2-4

Mselmi N, Lahiani A, Hamza T (2017) Financial distress prediction: the case of French small and medium-sized firms. Int Rev Financ Anal 50:67–80. https://doi.org/10.1016/j.irfa.2017.02.004

Nag AK, Mitra A (2002) Forecasting daily foreign exchange rates using genetically optimized neural networks. J Forecast 21(7):501–511. https://doi.org/10.1002/for.838

Papadimitriou T, Goga P, Agrapetidou A (2020) The resilience of the US banking system. Int J Finance Econ. https://doi.org/10.1002/ijfe.2300

Parot A, Michell K, Kristjanpoller WD (2019) Using artificial neural networks to forecast exchange rate, including Var-vecm residual analysis and prediction linear combination. Intell Syst Account Finance Manage 26(1):3–15. https://doi.org/10.1002/isaf.1440

Petukhina AA, Reule RC, Härdle WK (2020) Rise of the machines? intraday high-frequency trading patterns of cryptocurrencies. Eur J Finance 27(1–2):8–30. https://doi.org/10.1080/1351847x.2020.1789684

Petukhina A, Trimborn S, Härdle WK, Elendner H (2021) Investing with cryptocurrencies – evaluating their potential for portfolio allocation strategies. Quant Finance 21(11):1825–1853. https://doi.org/10.1080/14697688.2021.1880023

Pichl L, Kaizoji T (2017) Volatility analysis of bitcoin price time series. Quant Finance Econ 1(4):474–485. https://doi.org/10.3934/qfe.2017.4.474

Pompe PP, Bilderbeek J (2005) The prediction of bankruptcy of small- and medium-sized industrial firms. J Bus Ventur 20(6):847–868. https://doi.org/10.1016/j.jbusvent.2004.07.003

PricewaterhouseCoopers-PwC (2017). PwC's global Artificial Intelligence Study: Sizing the prize. Retrieved May 10, 2021, from https://www.PwC.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.

PricewaterhouseCoopers- PwC (2018). The macroeconomic impact of artificial intelligence. Retrieved May 17, 2021, from https://www.PwC.co.uk/economic-services/assets/macroeconomic-impact-of-ai-technical-report-feb-18.pdf.

PricewaterhouseCoopers- PwC (2020). How mature is AI adoption in financial services? Retrieved May 15, 2021, from https://www.PwC.de/de/future-of-finance/how-mature-is-ai-adoption-in-financial-services.pdf.

Qi M (1999) Nonlinear predictability of stock returns using financial and economic variables. J Bus Econ Stat 17(4):419. https://doi.org/10.2307/1392399
Qi M, Maddala GS (1999) Economic factors and the stock market: a new perspective. J Forecast 18(3):151–166. https://doi.org/10.1002/(sici)1099-131x(199905)18:33.0.co;2-v
Raj M, Seamans R (2019) Primer on artificial intelligence and robotics. J Organ Des 8(1):1–14. https://doi.org/10.1186/s41469-019-0050-0
Rasekhschaffe KC, Jones RC (2019) Machine learning for stock selection. Financ Anal J 75(3):70–88. https://doi.org/10.1080/0015198x.2019.1596678
Reber B (2014) Estimating the risk–return profile of new venture investments using a risk-neutral framework and 'thick' models. Eur J Finance 20(4):341–360. https://doi.org/10.1080/1351847x.2012.708471
Reboredo JC, Matías JM, Garcia-Rubio R (2012) Nonlinearity in forecasting of high-frequency stock returns. Comput Econ 40(3):245–264. https://doi.org/10.1007/s10614-011-9288-5
Renault T (2017) Intraday online investor sentiment and return patterns in the U.S. stock market. J Bank Finance 84:25–40. https://doi.org/10.1016/j.jbankfin.2017.07.002
Rodrigues BD, Stevenson MJ (2013) Takeover prediction using forecast combinations. Int J Forecast 29(4):628–641. https://doi.org/10.1016/j.ijforecast.2013.01.008
Van Roy V, Vertesy D, Damioli G (2020). AI and robotics innovation. In K. F., Zimmermann (ed.), Handbook of Labor, Human Resources and Population Economics (pp. 1–35) Springer Nature
Sabău Popa DC, Popa DN, Bogdan V, Simut R (2021) Composite financial performance index prediction – a neural networks approach. J Bus Econ Manag 22(2):277–296. https://doi.org/10.3846/jbem.2021.14000
Sariev E, Germano G (2020) Bayesian regularized artificial neural networks for the estimation of the probability of default. Quant Finance 20(2):311–328. https://doi.org/10.1080/14697688.2019.1633014
Scholtus M, Van Dijk D, Frijns B (2014) Speed, algorithmic trading, and market quality around macroeconomic news announcements. J Bank Finance 38:89–105. https://doi.org/10.1016/j.jbankfin.2013.09.016
Sermpinis G, Laws J, Dunis CL (2013) Modelling and trading the realised volatility of the ftse100 futures with higher order neural networks. Eur J Finance 19(3):165–179. https://doi.org/10.1080/1351847x.2011.606990
Sirignano JA (2018) Deep learning for limit order books. Quant Finance 19(4):549–570. https://doi.org/10.1080/14697688.2018.1546053
Soleymani F, Vasighi M (2020) Efficient portfolio construction by means OF CVaR and K -means++ CLUSTERING analysis: evidence from the NYSE. Int J Financ Econ. https://doi.org/10.1002/ijfe.2344
Sun T, Vasarhelyi MA (2018) Predicting credit card delinquencies: an application of deep neural networks. Intell Syst Account Finance Manage 25(4):174–189. https://doi.org/10.1002/isaf.1437
Szczepański, M. (2019). Economic impacts of artificial intelligence. Retrieved May 10, 2021, fromhttps://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf
Tao R, Su C, Xiao Y, Dai K, Khalid F (2021) Robo advisors, algorithmic trading and investment management: Wonders of fourth industrial revolution in financial markets. Technol Forecast Soc Chang 163:120421. https://doi.org/10.1016/j.techfore.2020.120421
Tashiro D, Matsushima H, Izumi K, Sakaji H (2019) Encoding of high-frequency order information and prediction of short-term stock price by deep learning. Quant Finance 19(9):1499–1506. https://doi.org/10.1080/14697688.2019.1622314
Trinkle BS, Baldwin AA (2016) Research opportunities for neural networks: the case for credit. Intell Syst Account Finance Manage 23(3):240–254. https://doi.org/10.1002/isaf.1394
Trippi RR, DeSieno D (1992) Trading equity index futures with a neural network. J Portf Manage 19(1):27–33. https://doi.org/10.3905/jpm.1992.409432
Uddin MS, Chi G, Al Janabi MA, Habib T (2020) Leveraging random forest in micro-enterprises credit risk modelling for accuracy and interpretability. Int J Financ Econ. https://doi.org/10.1002/ijfe.2346
Varetto F (1998) Genetic algorithms applications in the analysis of insolvency risk. J Bank Finance 22(10–11):1421–1439. https://doi.org/10.1016/s0378-4266(98)00059-4
Vortelinos DI (2017) Forecasting realized Volatility: HAR against principal components combining, neural networks and GARCH. Res Int Bus Financ 39:824–839. https://doi.org/10.1016/j.ribaf.2015.01.004

Wall LD (2018) Some financial regulatory implications of artificial intelligence. J Econ Bus 100:55–63. https://doi.org/10.1016/j.jeconbus.2018.05.003
Wanke P, Azad MA, Barros C (2016a) Predicting efficiency in Malaysian islamic banks: a two-stage TOPSIS and neural networks approach. Res Int Bus Financ 36:485–498. https://doi.org/10.1016/j.ribaf.2015.10.002
Wanke P, Azad MA, Barros CP, Hassan MK (2016c) Predicting efficiency in Islamic banks: an integrated multicriteria decision Making (MCDM) Approach. J Int Finan Markets Inst Money 45:126–141. https://doi.org/10.1016/j.intfin.2016.07.004
Wei L, Li G, Zhu X, Li J (2019) Discovering bank risk factors from financial statements based on a new semi-supervised text mining algorithm. Account Finance 59(3):1519–1552. https://doi.org/10.1111/acfi.12453
Xu Y, Zhao J (2022) Can sentiments on macroeconomic news explain stock returns? evidence from social network data. Int J Financ Econ 27(2):2073–2088. https://doi.org/10.1002/ijfe.2260
Xu D, Zhang X, Feng H (2019) Generalized fuzzy soft sets theory-based novel hybrid ensemble credit scoring model. Int J Financ Econ 24(2):903–921. https://doi.org/10.1002/ijfe.1698
Yang Z, Platt MB, Platt HD (1999) Probabilistic neural networks in bankruptcy prediction. J Bus Res 44(2):67–74. https://doi.org/10.1016/s0148-2963(97)00242-7
Yin H, Wu X, Kong SX (2020) Daily investor sentiment, order flow imbalance and stock liquidity: Evidence from the Chinese stock market. Int J Financ Econ. https://doi.org/10.1002/ijfe.2402
Zhang Y, Chu G, Shen D (2021) The role of investor attention in predicting stock prices: the long short-term memory networks perspective. Financ Res Lett 38:101484. https://doi.org/10.1016/j.frl.2020.101484
Zhao Y, Stasinakis C, Sermpinis G, Shi Y (2018) Neural network copula portfolio optimization for exchange traded funds. Quant Finance 18(5):761–775. https://doi.org/10.1080/14697688.2017.1414505
Zheng X, Zhu M, Li Q, Chen C, Tan Y (2019) Finbrain: When finance meets ai 2.0. Front Inform Technol Electr Eng 20(7):914–924. https://doi.org/10.1631/fitee.1700822

Authors and Affiliations

Salman Bahoo[1] · Marco Cucculelli[2] · Xhoana Goga[2] · Jasmine Mondolo[2]

✉ Marco Cucculelli
m.cucculelli@univpm.it

Salman Bahoo
sbahoo@edcparis.edu

Xhoana Goga
x.goga@staff.univpm.it

Jasmine Mondolo
j.mondolo@staff.univpm.it

[1] Department of Strategy and Management, EDC Paris Business School, 10074m, Puteaux Cedex, La Défense, 92807 Paris, France

[2] Department of Economics and Social Sciences, Marche Polytechnic University, Piazzale Martelli 8, 60100 Ancona, Italy