Machine learning in accounting and finance research: a literature review

Evangelos Liaras[1] · Michail Nerantzidis[1] · Antonios Alexandridis[2]

Abstract

In recent years, scholars in accounting and finance have shown a growing interest in employing machine learning for academic research. This study combines bibliographic coupling and literature review to analyze 575 papers from 93 well-established journals in the field of accounting and finance published between 1996 and 2022, and addresses three interrelated research questions (RQs): RQ1 How is research on the impact of machine learning on accounting and finance developed? RQ2 What is the focus within this corpus of literature? RQ3 What are the future avenues of machine learning in accounting and finance research? We adopt a critical approach to the research foci identified in the literature corpus. Our findings reveal an increased interest in this field since 2015, with the majority of studies focused either on the US market or on a global scale, with a significant increase in publications related to Asian markets during 2020–2022 compared to other regions. We also identify that supervised models are the most frequently applied, in contrast to unsupervised models, which mainly focus on clustering applications or topic extraction through the LDA algorithm, and reinforcement models, which are rarely applied, yield mixed results. Additionally, our bibliographic analysis reveals six clusters, and we discuss key topics, current challenges and opportunities. Finally, we outline machine learning constraints, highlighting common pitfalls, and proposing effective strategies to overcome current barriers and further advance research on this issue.

✉ Evangelos Liaras
  liarasev@o365.uth.gr

  Michail Nerantzidis
  nerantzidismike@uth.gr

  Antonios Alexandridis
  alexandridis@uom.edu.gr

1 Department of Accounting and Finance, University of Thessaly, University Campus "Geopolis" National Road Larissa Trikala, 411 10 Larissa, Greece

2 Department of Accounting and Finance, University of Macedonia, 156 Egnatia Street, 546 36 Thessaloniki, Greece

1 Introduction

The concept of artificial intelligence (AI) and its subset, machine learning (ML), traces its roots back to Turing (1950) and Samuel (1959) respectively. Turing introduced the concept through the "imitation game" to assess a machine's ability to simulate human-like behavior. Samuel, in turn, defined ML as "the field of study that gives computers the ability to learn without being programmed". However, despite its importance, ML lacks a universally accepted definition, leading to a gap in research (Gu et al. 2020). This ambiguity raises questions about the delineation between ML, broader AI, and purely statistical models. Additionally, the classification of a novel algorithm or methodology as part of the ML domain is left to the discretion of the author. Our research adopts Samuel's (1959) definition, including statistical regression models within the domain of ML, while excluding "explicitly programmed" methods associated with AI. In this study, we utilize bibliometric analysis to provide insights into the application of ML within the accounting and finance (A&F) discipline. We aim to explore the development of academic research in this area, shed light on its current focus, and identify potential future avenues of study.

Frequently, scholars encounter a dilemma when selecting between statistical methods and ML approaches. Statistical methods often provide clear interpretations of results, allowing for a better understanding of the relationships between variables. Moreover, statistical methods, are often based on well-defined assumptions, which can help in understanding the limitations of the model, and they are useful for making inferences about populations based on sample data, especially in hypothesis testing and confidence interval estimation. Lastly, they perform well with smaller datasets and limited computational resources. However, if the assumptions underlying the statistical model are violated, it can lead to inaccurate or misleading results. In addition, some statistical methods might struggle to handle complex relationships or high-dimensional data which leads to limited predictive accuracy. Finally, they are less flexible in adapting to various types of data structures or patterns.

On the other hand, ML algorithms can often handle complex patterns and large datasets, resulting in better predictive accuracy. They are characterized by flexibility and can adapt to different data structures and are often more versatile in handling various types of data, including unstructured data like images and text. In addition, many ML algorithms automatically learn relevant features from the data. Finally, machine learning algorithms usually scale efficiently with big data by leveraging parallel processing and distributed computing. However, one of the major critiques against ML models is that they can be challenging to interpret, leading to a lack of transparency in understanding how the model arrives at its decisions. In addition, these models may suffer from overfitting especially when the model complexity is high and the data is limited. Moreover, ML algorithms are computationally expensive and, usually, larger amounts of data is needed compared to statistical methods further increasing the training times. In practice, the choice between statistical methods and ML often depends on the specific problem, the available data, the desired level of interpretability, and the trade-offs between predictive accuracy and understanding the underlying relationships in the data.

ML algorithms are commonly separated into three main categories based on the training methodology; supervise, unsupervised and reinforcement learning. Supervised models require labeled datasets in which training data pairs (x, y) are provided during the training phase. The dependent variable y is used to enhance performance by iteratively adjusting the model parameters, (see, Mohri et al. 2012; Vapnik 2000). For example, in plant

identification, images of plants along their corresponding plant names are presented to the ML model. This type of model is used in both classification and regression problems. On the other hand, unsupervised models identify patterns and hidden features without labeled data. Clustering, association, and dimensionality reduction are common problems addressed through those models. Finally, reinforcement learning algorithms autonomously learn from their environment through trial and error. Data scientists provide reward and penalty rules that are used by the agent of the algorithm to distinguish correct from wrong actions. This type of models is commonly used in gaming industry and robot navigation applications, Russell and Norvig (2020).

Resent ML applications in A&F encapsulate a plethora of traditional subjects, inter alia fraud identification (e.g., Achakzai and Peng 2023; Debener et al. 2023), sentiment extraction from financial corpus (e.g., Blankespoor et al. 2023; Huang et al. 2023), portfolio optimization (e.g., Kaniel et al. 2023; Wang et al. 2024), and bankruptcy predictions (e.g., Cao et al. 2022; Nguyen et al. 2023). Moreover, ML utilization also incorporates the latest trends related to cryptocurrency markets (e.g., Cohen 2023; Han et al. 2024), environmental implications (e.g., Frost et al. 2023; Sautner et al. 2023) and Covid-19 impact (e.g., Chortareas et al. 2024; Yang et al. 2024) indicating a continuous interest in their capabilities.

The above discussion suggests several reasons for undertaking this study. Firstly, empirical studies on ML in A&F have experienced a significant increase in recent years and are widely scattered across various academic journals, making it challenging to obtain a clear picture of this expanding research area. Secondly, to our knowledge, no other study presents a comprehensive literature review on ML in A&F. While a limited number of review studies address this issue to some extent from either an AI or a ML perspective (Gray et al. 2014; Sutton et al. 2016; El-Haj et al. 2019; Weigand 2019; Karolyi and Van Nieuwerburgh 2020; Ahmed et al. 2022; Han et al. 2023), their focus is considerably narrower than the exhaustive coverage offered in this review. For instance, Han et al.'s (2023) review focuses on blockchain applications in accounting, whereas El-Haj et al. (2019) examines the utilization of Computational Linguistics, a subset of AI, in financial disclosure. Thirdly, this review responds to the recent calls[1] for more research in machine learning within A&F.

To address the research gap discussed above, we employ a combination of quantitative techniques, such as bibliometric analysis, and a critical review of all identified research foci within the literature corpus. This approach enables us to offer a comprehensive and interdisciplinary synthesis of knowledge in this field. Specifically, we aim to address three key research questions in this stream of research: RQ1 How is research on the impact of ML on A&F research developed? RQ2 What is the focus within this corpus of literature? RQ3 What are the future avenues of ML in A&F research? We adopt a critical approach to the research foci identified in the corpus of literature by analyzing 575 papers from 93 established quality journals and providing a comprehensive and interdisciplinary synthesized state of knowledge regarding this field of study.

Our results reveal an increased interest in this field since 2015 with the majority of studies focused either on the US market or on a global scale. Publications related to

Asian markets gained momentum as they increased by 950% during 2020–2022. Further, our analysis highlights we show that supervised models are by far the most frequently applied, contrary to unsupervised models, which are mainly focused either on topic extraction through the Latent Dirichlet Allocation (LDA) algorithm or clustering. Additionally, through comprehensive bibliographic analysis, our study identifies six distinct clusters. For each cluster we present the key topics, we examine the current challenges, and discuss the various prospective opportunities. We also examine and propose future avenues of research of ML in A&F. Finally, we analytically present and discuss the various limitations of ML and possible directions for future research to overcome them.

Our research contributes to the relevant literature in several ways. Firstly, to the best of our knowledge, this is the first literature review that investigates both A&F research exclusively under the prism of ML. Secondly, our analysis focuses on established journals included in the 2021 Academic Journal Guide (AJG) in the field of A&F (ranked as 4*, 4, 3, 2 and 1) to ensure that our findings are derived from high-quality academic research and to identify the direction of future research. Third, for each cluster we summarize the research topics, as well as the preferred methodologies and the best performing models for each topic. This summary intends to offer valuable guidance to scholars and early career researchers interested in employing ML in the fields of A&F.

The paper is organized as follows. In Sect. 2 we introduce our methodology. Section 3 discusses the results regarding the three interrelated research questions and adopts a critical approach to the research foci identified in the corpus of literature. In Sect. 4 we present opportunities and challenges of ML. Finally, Sect. 5 outlines the main conclusions and presents the limitations of the paper.

2 Methodology

In this study, we utilize bibliometric analysis[2] to provide insights into the application of ML within the A&F discipline. We aim to explore the development of academic research in this area, shed light on its current focus, and identify potential future avenues of study. Bibliometric serves the dual purpose of mitigating author bias (MacCoun 1998) and efficiently summarizing extensive datasets (Broadus 1987). Given the nascent stage of ML in A&F we combine quantitative with qualitative data by complement bibliometric analysis with literature review for a deeper understanding of our topic (Rialti et al. 2019). This combination of techniques differentiates our study from previous literature research in Artificial Intelligence and ML (Goodell et al. 2021; Ahmed et al. 2022; Ranta et al. 2022) by shedding light on traditional methods, current implications and comparison with ML models for A&F tasks.

Firstly, we conduct our initial research in August 2023 exclusively in the Web of Science (WoS) database as considered the most credible, transparent, and reliable source of information (Modak et al. 2019; Levine-Clark and Gil 2021). In our search keyword selection process, we begin by identifying literature papers that are both pertinent to our study and have been published in high-quality journals rated as 4* (internationally recognized as examples of excellence), 4 (top journals in the field), and 3 (highly regarded journals)

---

[2] We also applied the Dynamic Topic Model as proposed by (Blei & Lafferty 2006). However, due to the sensitivity of the results, we decided to undertake bibliometric analysis.

following the guidelines established by the Chartered Association of Business Schools through the Academic Journal Guide (AJG). Our analysis aims to reveal keywords from titles, author-provided keywords, and publication abstracts that comprehensively address our research topic. In particular, we adopt the ML keywords outlined in the review conducted by Ghoddusi et al. (2019), adjusting our query to align with the objectives of our study. Therefore, our keyword list is composed of 25 keywords related to ML and 13 to A&F. To further expand our research, we also apply stemming and the asterisk "*" regular expression as suffix for generic keywords (e.g., Auditing becomes Audit*) as supported by WoS database. The final query can be found in Appendix.

As a next step, we refine our search results by retrieving any published article until 2022. Additionally, we restrict our search to peer-reviewed and scholarly journals included in the 2021 AJG guide in the field of A&F (ranked as 4*, 4, 3, 2 and 1) to ensure a minimum level of research quality (Chartered Assosiation of Business Schools 2021; Harvey et al. 2010). As a consequence, the initial dataset comprises 3,709 articles, of which 1,004 align with our rating criteria. After eliminating articles that are not pertinent to our research and those that are irretrievable, our final corpus consists of 575 articles.

To answer our three research questions, we employ quantitative and qualitative tools to conduct our analysis. Specifically, for RQ1, we perform preliminary descriptive analysis using the Bibliometrix R software (Aria and Cuccurullo 2017) to gain further insights into our corpus. Furthermore, we examine the evolution and contribution of journals in academic research, and we present the geographic focus trends.

To address our RQ2, we carry out Bibliometric coupling, a method known for its capacity to produce highly accurate clustering results (Boyack and Klavans 2010). This method hinges on the assumption that publications delving into similar topics will share common citations; the greater the number of shared citations, the stronger the connection between the articles, and the higher the likelihood that they address a shared topic. Hence, through bibliographic coupling, we can cluster our results and conduct a comprehensive literature review to uncover the key topics and considerations addressed in these clusters. We have selected VOSviewer (Van Eck and Waltman 2010), a tool widely employed by scholars for literature review and known for its graphical representation of clustering (e.g., Ciampi et al. 2021). Furthermore, to enhance our understanding on cluster topics, we construct wordclouds through bag-of words, a standard technique exercised in literature analysis (Baker et al. 2021).

Lastly, to identify potential future avenues for our RQ3, we employ, through VOSviewer tool, co-word analysis in author specific keywords. Thus, each keyword is distributed through time unfolding trends and patterns. Moreover, our literature review identified key considerations and limitations of ML that could be put under the microscope for future research.

3 Results

3.1 (RQ1) How is research on the impact of ML on A&F developed?

To address our initial research question, we undertake a quantitative analysis of 575 publications to offer an overview of current research. In particular, we identify trends in number of publications per year while pinpointing the top journals within our topic. Moreover, we perform thorough literature review to recognize the geographic data sources and trends of

Table 1 Descriptive statistics

Description	Results
Panel A: Main Data Information	
Timespan	1996–2022
Sources (Journals, Books, etc.)	93
# of Documents	575
Annual Growth Rate %	15.92%
Document Average Age	3.27
Average citations per doc	18.32
# of References	22,518
Panel B: Document Contents	
# of Keywords Plus[1]	1105
# of Author's Keywords	1832
Panel C: Authors	
# of Authors	1495
# of Authors of single-authored docs	61
Co-Authors per Doc	2.87
International co-authorships %	34.43%
Panel D: Document Types	
# of Articles	483
# of Articles; Early Access	69
# of Articles; Proceedings Paper	9
# of Review	12
# of Review; Early Access	2

[1]Keywords added by WoS

selected publications. Finally, we identify the most impactful countries based on publications and citations.

Firstly, the descriptive statistics of selected publications that emerged from Bibliometrix R (Aria and Cuccurullo 2017) are included in Table 1. Specifically, this table is composed of four panels namely "Main Data Information", "Document Contents", "Authors", and "Document Types", each providing a different aspect of our corpus. With an annual growth rate of 15.92% and average document age of 3.27 years, ML research in A&F is a trending topic in its infancy. Additionally, a higher co-authorship average in comparison to previous literature analyses by Gaunt (2014) and Korkeamaki et al. (2018) in A&F publications, may indicate the incorporated complexity of machine learning application in these disciplines. "Document types" panel, illustrates the distribution of publications per document type including early access articles to be published after 2022; an indicative trend of a continuous momentum in this topic.

The paper from Gray (1996) represents the oldest article in our corpus, and it was published in the "Journal of Financial Economics". Consistent with prior research, our analysis reveals that the interest in ML applications within the A&F domain started to surge in 2015, a trend that remains robust due to technological advancements enabling the execution of these algorithms on conventional personal computers. These findings are presented in Fig. 1. Specifically, we categorize journals with ten or fewer publications as "Other" and exclude 53 unpublished articles from our analysis. Among the
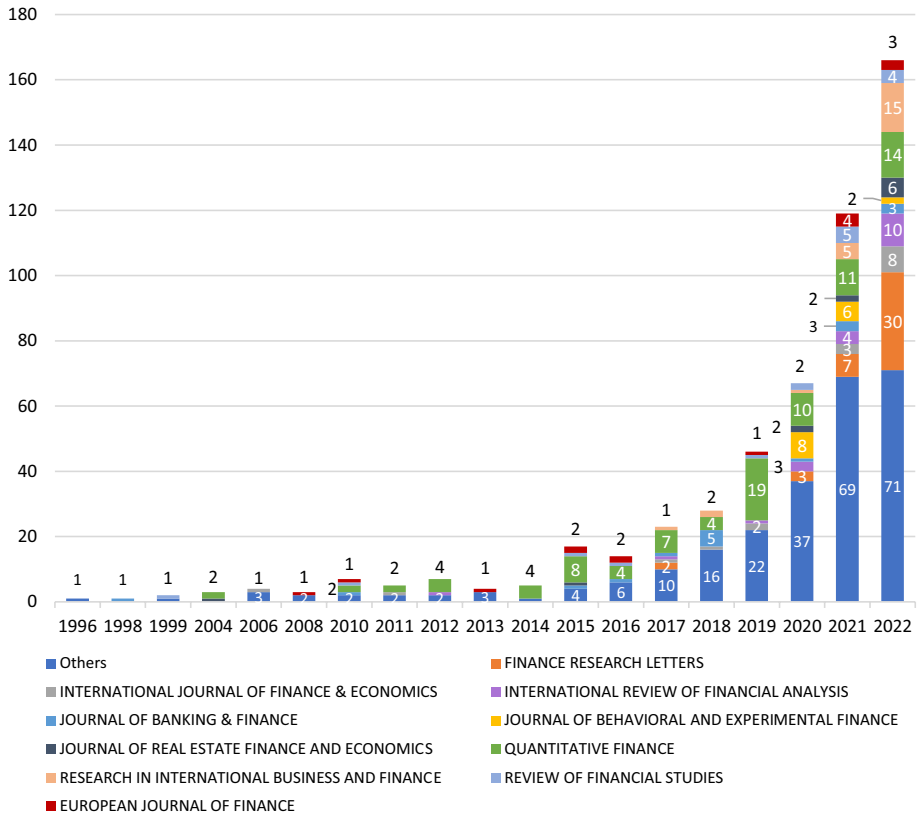
Fig. 1 Number of papers per journal and published year

journals, "Quantitative Finance" stands out by contributing 16.35% of the papers in our corpus, while "Finance Research Letters" published the highest number of articles in 2022, totaling 30. Noticeably, 88.38% of publications included in our corpus relate to finance discipline. This can be attributed to several factors. In practice, accounting captures and provides information (dual role of accounting, valuation and contracting perspective), while finance considers this information to make informed decisions (Baker and Wurgler 2002; Ruch and Taylor 2015). Despite the fact that A&F often co-exist in a single academic unit (Smith and Urquhart 2018), the skills and expertise required for each field differ accordingly. While ML is used in both fields, it seems more prevalent in finance due to its wider range of applications. ML is ideal for handling the vast amount of complex data available in financial markets, enabling real-time decision-making and predictive analysis. On the other hand, accounting frequently entails nuanced tasks, such as providing tax advice or handling matters that require more intuition, where ML may not be as effective. Also, most ML models lack transparency making ML adaptation more difficult for practitioners who often must clearly justify their decision-making process. Finally, maybe it is just the lack of expertise. However, over the past decade, there has been a significant increase in publications within the field of accounting that incorporate ML algorithms. Especially following the publication of Loughran and

McDonald's paper in 2011, there has been an exponential rise in the utilization of text analysis applications in both A&F.

In Table 2, we provide a list of the top 10 journals from a total of 93, based on several key criteria, including the number of articles, total citations, average citations per article, h-index,[3] g-index[4] and journal impact factor as provided by Clarivate. Each journal is compared and ranked based on these six criteria, with at least four of them showing better performance compared to the preceding journal. In the event of a tie, we consider the AJG Ranking as the determinant of the journal's impact.

Observing, through literature review, a considerable concentration of studies within a particular geographic region, it becomes imperative to replicate these experiments in alternative markets or on a global scale. This approach serves two critical purposes: validating results and examining potential variations across different geographic locations. Hence, we implement a geographic data filter, assessing the specific data utilized in each publication. In some cases, we identify datasets encompassing multiple countries, whether from the same continent or across different continents. In the first case we classify those papers to the related continent, while in the latter we label them as "Global". Our results can be found in Table 3 presenting up to the top five regions per continent. Literature that either lacks explicit mention of the dataset's composition or does not utilize any dataset has been excluded from our analysis.

While the majority of publications are tailored to the US markets, almost 30% of papers are focused on global datasets. We further examine this trend in Fig. 2, depicting the evolution of the scientific focus over the years. Please note that early access papers published after 2022 have been removed from our analysis.

With few exceptions, a significant amount of ML applications in North American datasets has been observed since 1996, constituting 38.26% of our dataset population. While there has been a positive upward trend since 2016, the proportion attributed to this group has stabilized in 2022, despite the ongoing increase in total publications. Specifically, publications for Asian markets increased by 975%, European by 243% and Global by 95.83% over the last 3 years. During the same timeframe, we notice only one publication for Oceania and one publication per year for Africa. In summary, although as expected a considerable number of papers still rely on North American data in absolute terms, there is a noticeable increase in applications across the globe, particularly in Asian datasets, reflecting a growing trend.

Lastly, we perform bibliographic analysis through VosViewer (VOSviewer—Visualizing scientific landscapes 2021) to identify the most influential countries through average citations per paper in our corpus. Countries with five or more papers published are presented in Fig. 3. Papers from Austria and the United States are the most cited (with 86 and 33.31 average citations respectively), in contrary to Japan and Thailand being on the other side of this spectrum (2.58 and 4.6 respectively).

Stepping further into our analysis, we calculate the most influential countries in terms of publications by measuring total publications, total citations, and average citations per publication. In particular, the United Stated of America published 27.48% of papers ($N = 158$) achieving 5,295 citations followed by China which accredited with 1,132 citations from 102 publications. Countries with more than 200 citations are illustrated in Fig. 4. Table 4

---

[3] The h-index is defined as the maximum value of h such that the given author has published at least h papers that have each been cited at least h times.

[4] The g-index is the unique largest number such that the top g articles received together at least g^2 citations.

Table 2 Top journals in ML in A&F

Journal	Field	AJG Ranking	TP	TC	C/P	H-Index	G-Index	JIF
Quantitative Finance	Finance	3	94	1682	17.89	19	38	1.3
Finance Research Letters	Finance	2	46	529	11.5	11	22	10.4
Review of Financial Studies	Finance	4*	18	1321	73.39	11	18	8.2
Journal of Financial Economics	Finance	4*	9	1581	175.67	7	9	8.9
Journal of Banking & Finance	Finance	3	18	773	42.94	10	18	3.7
International Review of Financial Analysis	Finance	3	22	299	13.59	8	17	8.2
Journal of Behavioral and Experimental Finance	Finance	1	17	245	14.41	9	15	6.6
Research In International Business and Finance	Finance	2	28	257	9.18	7	15	6.5
Journal of Accounting Research	Accounting	4*	9	727	80.78	5	9	4.4
Journal of Real Estate Finance and Economics	Finance	3	13	209	16.08	6	13	1.9

TP = Total Production; TC = Total Citations; C/P = Citations per Publication; JIF = Journal Impact Factor as per Clarivate

Table 3 ML application in continents and countries

| Continent	Country\Region	Times Referred	Continent Total	Percentage
Africa	South Africa	3	5	0.87%
	Africa	1		
	Tunisia	1		
Asia	China	52	91	15.83%
	India	10		
	South Korea	8		
	Asia	6		
	Taiwan	6		
Europe	Europe	38	78	13.57%
	UK	8		
	France	7		
	Germany	4		
	Italy	3		
Global	Global	171	171	29.74%
North America	US	198	207	36.00%
	Canada	6		
	North America	2		
	Mexico	1		
Oceania	Australia	3	4	0.70%
	Oceania	1		
South America	Colombia	1	1	0.17%
Unknown	N/A	18	18	3.13%
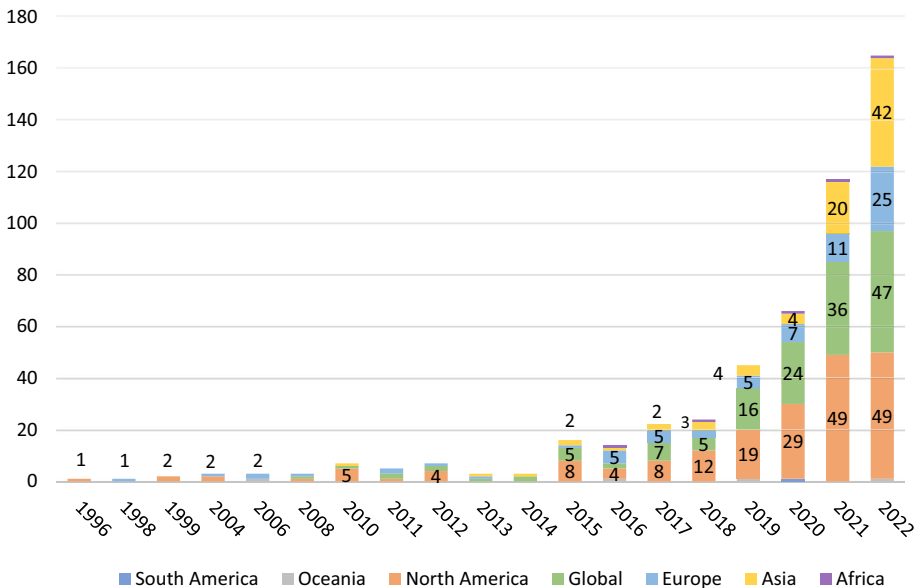
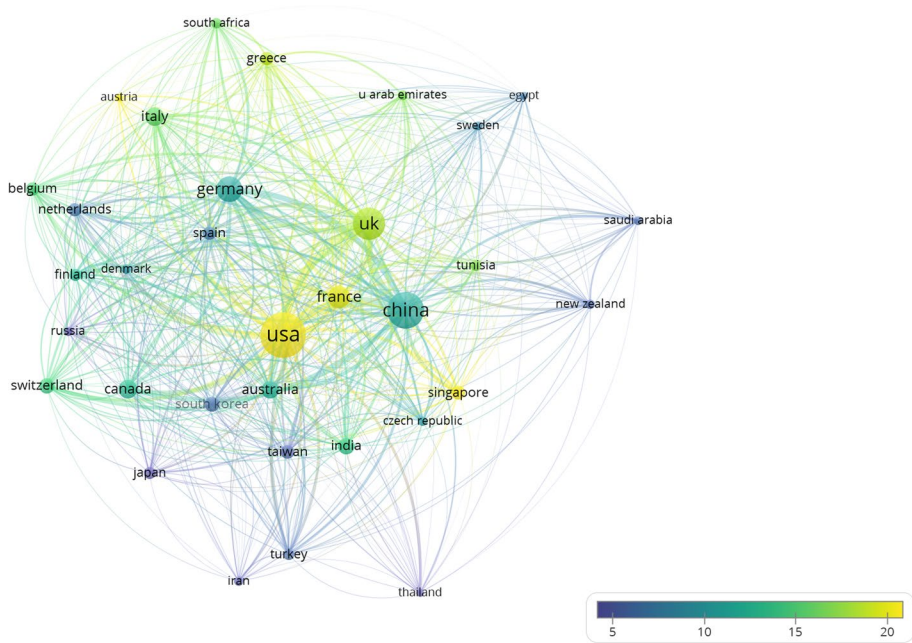
Fig. 2 Geographic dataset trends

Fig. 3 Bibliographic coupling countries

presents the top 10 countries in relation to h-index, g-index and total citations. Each country is compared and ordered by the three criteria having at least better performance in two of them in relation to the proceeding country.

3.2 (RQ2) What is the focus within this corpus of literature?

3.2.1 Bibliographic coupling results

In this section, we provide an in-depth analysis of our corpus of literature by initially performing clustering through bibliographic coupling and combining bag-of words with literature review to identify the key topics for each cluster. Moreover, we extract key considerations and proposed ML models as to assist future research. Lastly, we perform cluster distribution over time to indicate the academic focus and progression.

Bibliographic coupling analysis reveals six clusters by linking the citing documents based on the number of papers cited together. VOSviewer (VOSviewer—Visualizing scientific landscapes 2021) constructs distance-based maps in which the smaller the distance between two items, the stronger their relation (Van Eck and Waltman 2010). In Fig. 5 we present the relative map where each node represents a paper and its color indicates the cluster assigned. The node size indicates the times a paper is cited. The produced map enhances our motivation to review all papers in order to identify the characteristics of each cluster; common citations between groups and overlapping districts indicate a somewhat strong correlation between some of the papers in different clusters. Natural Language Processing (NLP) analysis techniques namely Bag-Of-Words and n-grams are employed on keywords, titles and abstracts of publications solidifying cluster results. Specifically,
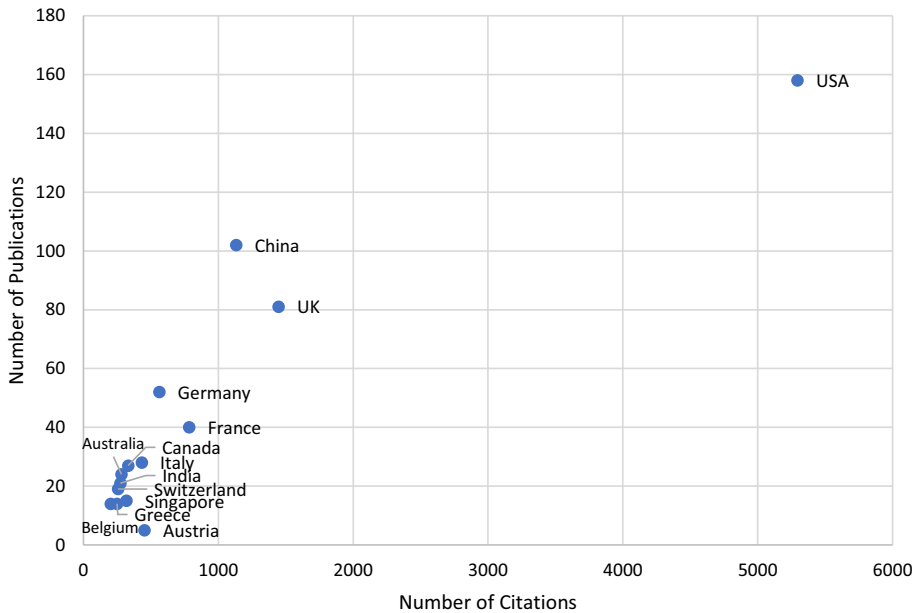
Fig. 4 Most influential countries (> 200 times cited)

Table 4 Most influential countries

Rank	Country	TP	TC	H-Index	G-Index
1	USA	158	5295	35	70
2	UK	81	1447	17	35
3	China	102	1132	15	29
4	France	40	783	15	27
5	Germany	52	562	11	21
6	Italy	28	432	9	20
7	Canada	27	331	8	17
8	Australia	24	282	10	16
9	Singapore	15	319	8	15
10	India	21	274	7	16

TP = Total Production; TC = Total Citations

for each created cluster, we consolidated information collected from keywords, titles and abstracts of included publications, and applied stemming process. Finally, we calculated the occurrence of each word or n-gram withing the cluster to construct our wordclouds. Words that are common across all clusters, such as "machine" and "learning", are excluded from this analysis. The full corpus is available as online supplementary material.[5]

---

[5] Available at https://github.com/Liarasev/Machine-Learning-in-Accounting-and-Finance-Research-A-Literature-Review
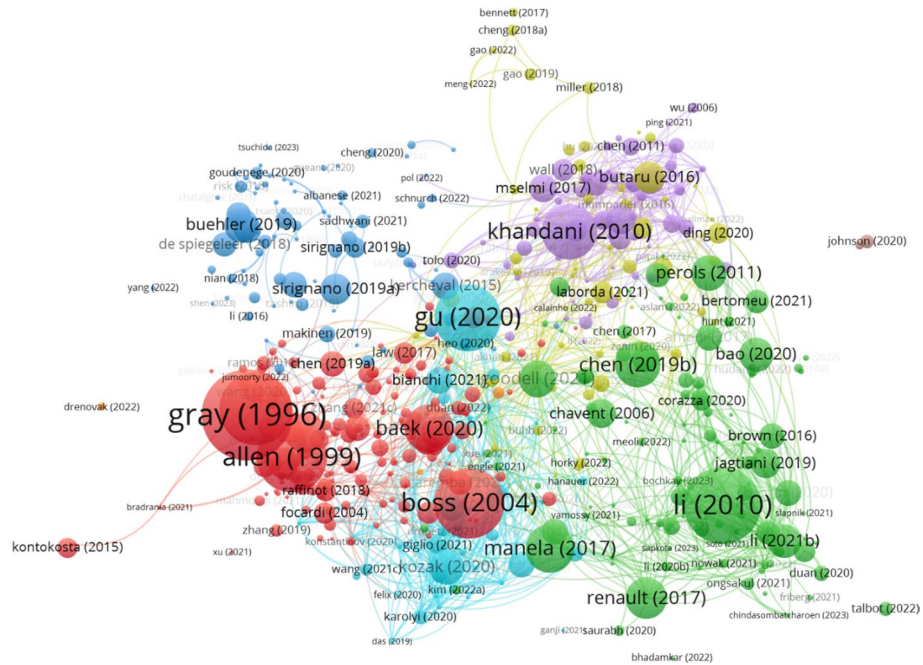
Fig. 5 Bibliographic coupling map

3.2.2 Red cluster: markets and time-series forecasts

The first cluster consists of 172 publications with a total of 4,149 citations. Our analysis reveals that the most common bi-grams in this cluster are "Neural Network" (64 occurrences), "Time Series" (61 occurrences), "Stock Market" (50 occurrences), "Covid-19" (50 occurrences) and "Exchange Rate" (33 occurrences). As presented in Fig. 6 panel A, this extensive cluster encompasses topics related to market volatility, portfolio creation, trading behavior, cryptocurrency, and forex markets, all of which are associated with "time series" analysis. The dynamic, noisy, and non-linear nature of financial time series forecasting makes it a complex endeavor (Karathanasopoulos et al. 2015).

Starting with volatility subcluster, the primary goal is to minimize tolerated risk and maximize gains. Traditional GARCH and Stochastic Volatility models are not suitable in the current high-frequency data environment (Liu et al. 2018). Furthermore, the GARCH model is prone to explosive conditional variance, which has implications for volatility forecasting (Gray 1996). Moreover, the Markowitz mean–variance portfolio model (Markowitz 1952), ignores transaction costs and is more prone to estimation error than minimum variance models (Clarke et al. 2011). Thus, the topic of discussion in this subcluster is the comparison of econometric models with ML models that have the ability to process multi-dimensional and non-linear data. For example, the results of Hu and Tsoukalas (1999) indicate that realized volatility approximates stock volatility through a non-linear approach in which neural networks outperform the GARCH model. The choice of model selection for forecasting depends on the time frame, with long-term volatility forecasts favoring ML models over the econometric GARCH for Forex

Fig. 6 Wordclouds of six clusters. *Note* The font size indicates the times each word occurs withing the cluster

and the Chinese CSI 300 index, while errors remain identical for short-term forecasts (Zhai et al. 2020). Combining Neural Networks with MC-GARCH for high-frequency data processing is another option that shows promising results, achieving higher accuracy compared to standalone MC-GARCH. Similarly, the Heterogeneous Autoregressive model (HAR), yields exceptional results on high-frequency oil price prediction (Gkillas et al. 2020), but underperforms over Neural Networks and tree-based ML algorithms for stock volatility forecast (Christensen et al. 2022). Lastly, the combination of ML models for volatility forecasting is also under the microscope. For example, Qiu et al. (2020) propose the combination of the HAR model with random forest for forecasting the price volatility of 100 Exchange traded funds (ETF).

Portfolio creation is the second subcluster of this group. Traditional approximations of optimal asset allocation, assessing the risk/benefit ratio, include the Black-Litterman (Black and Litterman 1992) and the Markowitz models (Markowitz 1952). The first model suffers from high transactional costs and does not account for stock-specific views, while the latter is sensitive to assumptions. Rebalancing models and the naïve 1/*N* rule, which involves

equally investing in *N* assets, often perform better than the Markowitz model (Mulvey et al. 2001; DeMiguel et al. 2007). Thus, creating portfolios of as independent assets as possible and rebalancing only when a certain risk threshold is exceeded is a promising alternative (Liu et al. 2015; Li et al. 2016). Supervised, unsupervised, and reinforcement models are proposed as replacements or enhancements to the econometric ones. Pyo and Lee (2018) recognize low-risk anomalies and the outperformance of low-risk portfolios compared to high-risk ones. They experiment with ML algorithms and the GARCH model to forecast volatility initially. Subsequently, they integrate these forecasts with the Black-Litterman model for portfolio construction purposes. The combination of the Markowitz model with Neural Networks (Bradrania et al. 2021) or with Convolutional Neural Networks (CNN) and reinforcement learning (Aboussalah et al. 2021), seems to exceed expectations. In contrast, the poor performance of most reinforcement learning algorithms is attributed to noisy and non-stationary financial environments (Aboussalah et al. 2021).

Identification of key aspects of trading behavior, strategies, and price forecasting is the topic of the third subcluster. Starting with the identification of informed trading, current models are computationally intensive in choice of initial parameters requiring up to months of effort (Gan et al. 2015). The unsupervised Hierarchical Agglomerative clustering model is a better alternative in terms of speed and accuracy (Gan et al. 2015; Lin et al. 2021). In the pursuit of automated trading, Genetic Algorithm is able to learn trading strategies and apply them accordingly however after inclusion of transaction costs, their return is similar to a buy and hold strategy (Allen and Karjalainen 1999). One solution is to import risk factors into the model and apply boosting algorithms that avoid unnecessary costly trades by relying exclusively on "strong" signals for decision making (Creamer and Freund 2010). In financial prediction, current econometric and statistical models like ARMA, ARIMA, VAR and GARCH perform well with linear data but struggle when this assumption is violated (Wu et al. 2019). The comparison of ARMA models with Support Vector Machines (SVM) for indices prediction enhances the above statement (Karathanasopoulos et al. 2015). We also find that Neural Networks has been used with success stock prices prediction (Chen and Ge 2019; Zhang et al. 2021) however Genetic Programming could be a better alternative (Dunis et al. 2013).

Crypto-currencies provide an intriguing avenue for investigating market efficiency in high-frequency trading by training ML algorithms to simulate investor actions (Manahov and Urquhart 2021) or predicting prices on a daily timeframe by incorporating significant macroeconomic variables (Liu et al. 2021). Unsupervised models through clustering can also be employed to examine market efficiency and behavior by identifying bubbles (El Montasser et al. 2022). Moreover, the realm of price prediction remains a central topic in academic research. In Aggarwal et al, (2020) it is shown that the accuracy of the SVM model fluctuates based on the selected forecasting period. Notably, forecasting the bitcoin price for the fifth day in the future yields more accurate results than predicting the 15th or 30th day. Lastly, ML models and technical analysis play a role in identifying trends in cryptocurrency prices. However, the absence of observed abnormal returns suggests that cryptocurrency markets may be more analogous to traditional financial markets (Anghel 2021).

The fourth subcluster comprises nine publications that assess the impact of the Covid-19 pandemic on financial markets and aim to identify the key features that minimize risk. Specifically, social media and Covid-related news improve market impact predictability, with varying impacts on Gulf Cooperation Council countries (Al-Maadid et al. 2022). Through Hierarchical clustering, interconnectedness of markets is examined both before and after the financial crisis. Tightly connected countries tend to strengthen their interconnectedness

in contrast to less strongly connected ones that tend to move closer to another cluster after the crisis (León et al. 2017). At an industry level, examining the correlation between positive and negative news related to COVID-19 and US stock prices can serve as an indicator of systemic risk (Baek et al. 2020). Zaremba et al. (2021) investigate 67 markets using multiple factors and observe that low unemployment rates, conservative investment policies, and undervalued companies tend to be more protected from global pandemics. Lastly, when comparing simple linear regression with ML regression models, it is observed that Support Vector Regression (SVR) and Random Forest achieved better accuracy in correlating the Covid-19 death rate with stock market performance in India (Behera et al. 2022).

In summary, this cluster identifies the importance of ML for time-series forecasting as traditional models are more suitable for linear data. However, a combination of econometric with ML models could be beneficial as previous research indicates. The supervised models SVM, Neural Networks and tree-based models are quite common while we find Hierarchical clustering to be the choice for explaining the relationship between entities such as trading patterns.

3.2.3 Green cluster: textual analysis

This cluster consists of 115 publications accredited with 2695 citations. Our analysis reveals two prominent ML models: "Neural Network" (25 occurrences) and "Support Vector Machine" (13 occurrences). The most common bi-grams are "Textual Analysis" (43), "Social Media" (26 occurrences), "Fraud Detection" (26 occurrences), "Annual Reports" (25 occurrences) and "Data Mining" (25 occurrences). Notable tri-grams include "Corporate Social Responsibility" (14 occurrences), "Natural Language Processing" (11 occurrences), "Accounting Information Systems" (9 occurrences) and "Insurance Fraud Detection" (9 occurrences). As Fig. 6 panel B indicates, the key distinction of this cluster is its focus on experimentation with textual data, enabling the extraction of sentiment and the identification of the key corporate insights, particularly in the context of fraud detection.

A common approach for sentiment extraction involves word classification, categorizing words as positive, neutral, negative, or other categories to calculate the sentiment of each sentence and extend it to the entire document. This approach often relies on dictionaries and lexicons, which can help remove subjectivity and reduce the researcher's efforts (Loughran and McDonald 2016). However, word lists may not always be readily available, and issues related to homographs can arise (Loughran and McDonald 2016). Furthermore, domain-specific dictionaries may not perform effectively in different contexts (Bochkay et al. 2019). In contrast, ML models can be trained to discover text features and assign unique sentiment weight to individual words via manual classification of the training and verification sets. In our corpus, social media provide the textual data required for stock price prediction using ML techniques (Renault 2017; Chun et al. 2020; Vamossy 2021) This data allows for the assessment of investor sentiment and beliefs regarding specific firms at given times. Forcasting trading trends is also achievable by identifing topics from news articles via an unsupervised ML model known as LDA (Han and Kim 2021a). Commonly identified ML models include Naïve Bayes (Li 2010; Slapnik and Lončarski 2021), SVM (Liu et al. 2021) and Neural Networks (Chun et al. 2020; Saurabh and Dey 2020; Azimi and Agrawal 2021).

Our second subcluster is focused on the detection of misreporting in financial statements, which has repercussions for both investors and employees and contributes to uncertainty in financial markets. Challenges in this group are the ratio of the nonfraud firms to

fraud ones, which affects ML classification, and to find the right attributes as often are noisy due to the attempt to mask the financial statements to be as similar as possible to the ones found in nonfraud firms (Perols 2011). Addressing class imbalance often involves under-sampling, aiming for a 1:4 ratio between fraud and non-fraud instances. However, it's crucial to evaluate additional metrics such as the F-measure or area under the curve (AUC) to comprehensively assess model performance. (Papík and Papíková 2022). Different sets of variables are proposed as fraud detection indicators including raw financial data instead of ratios (Bao et al. 2020), combinations of ratios, raw data and dummy variables (Perols 2011), inclusion of non-accounting variables such as governance, capital markets, and auditing (Bertomeu et al. 2020) and even textual analysis (Chen et al. 2017; Brown et al. 2020; Zhang et al. 2022). Regarding the latter category, an NLP method called Term Frequency-Inverse Document Frequency (TF-IDF) can identify the most important fraudulent accounting narratives in annual reports, enabling Queen Genetic Algorithm and SVM models for the classification of fraud and non-fraud financial statements (Chen et al. 2017). While the "Bag of Words" technique assesses word relevance based solely on the frequency of words within a document, TF-IDF considers not only the frequency of words in a document but also considers how frequently those words appear across the entire collection of documents. A more sophisticated method known as "word embedding" takes into account the sentence structure and creates multidimensional vector representations for each word, allowing for similar representations of synonym words. SVM combined with "word embedding", achieved an accuracy of 77% in fraud detection for Chinese firms (Zhang et al. 2022).

In summary, textual analysis can be conducted through a range of techniques, from simple word counting to vector representation of words. ML can complement both approaches for solving regression, classification, and clustering problems, given their capability to handle multidimensional data. Among unsupervised models, LDA is widely used for topic extraction. In our corpus, neural networks and SVM supervised models are frequently employed and have shown promising results.

3.2.4 Blue cluster: options and limit order trading

The third cluster consists of 89 publications with a total of 904 citations. Our textual analysis indicates keywords related to both ML and traditional models namely "Deep Learning" (69 occurrences), "Neural Networks" (37 occurrences), "Monte Carlo" (20 occurrences) and "Long Short-term Memory" (11 occurrences). The most common topic-related *n*-grams are "Option Pricing" (24 occurrences), "Limit Order Book" (20 occurrences) and "High Dimensional" (20 occurrences), while Fig. 6 panel C illustrates the most frequently repeated keywords in this cluster, including "Price" and "Option".

European, American, and Bermudan option pricing and hedging are scrutinized due to limitations of commonly used traditional techniques. For European options, the Black and Scholes (1973) and the Heston (1993) models are among the most frequently applied methods for identifying underlying prices. However, both models are parametric and rely on certain assumptions that can impact accuracy. For instance, the Black–Scholes model assumes constant volatility, leading to inaccurate prices (Funahashi 2020; Nian et al. 2021). Slow Monte Carlo method, another common choice for option pricing, increase accuracy but at the expense of speed (Horvath et al. 2021), which can be problematic in high-volatility markets. In contrast, ML models, specifically Neural Networks in our corpus, are trained directly on market data, avoiding the misspecification issues that parametric models can

suffer from (Nian et al. 2021). Modeling American options is even more complex as they can be exercised at any time during the contract's life. The regression-based Monte Carlo approach is gaining popularity; however, the objectivity of variable selection in regression-based methods can be compromised, particularly in the case of high-dimensional options (Hu and Zastawniak 2020). Research suggests either combining Monte Carlo with ML or replacing it. For example, in Goudenège et al. (2020), the combination of trees and Monte Carlo methods for pricing American options results in a computation time decrease without sacrificing accuracy. On the contrary, De Spiegeleer et al. (2018) advocate for a new ML model that achieves faster execution speeds compared to Monte Carlo, even if it involves sacrificing some accuracy within acceptable thresholds. In our corpus, Neural Networks are the most commonly applied ML family, including CNN, Long Short-term Memory (LSTM) and recursive neural networks (RNN) (Jang and Lee 2019; Wei et al. 2020; Zhang and Huang 2021).

In the context of limit order trading, the evaluation of the immediate stock pricing movement relies on the asks and bids order flows in high-frequency trading environments. One common approximation involves statistical modeling; however, these methods often require assumptions and intensive computations, making them less suitable for existing trading environments (Bouchaud et al. 2002). Modeling in this context is further complicated by the high-dimensionality of a limit order book, which includes multiple price levels (Sirignano 2018). At the same time, information provided by order book is often overlooked by human observers due to their short-lived nature (Kercheval and Zhang 2015). A deep learning solution has the capacity to generalize and identify relationships between order flow and market prices in a non-parametric manner that can be applied to different stocks (Sirignano and Cont 2019). Neural Networks, for instance, can determine whether price changes result from successful or cancelled orders by considering the order series (Tashiro et al. 2019). However, to train machine models effectively, it is necessary to include multiple levels of limit orders, accounting for factors such as order size (number of stocks), price, and precise placement time. In high-frequency trading markets, data is abundant, which can make model training a time-consuming process. To address this, training is often carried out on GPUs rather than CPUs, as graphic cards can efficiently parallelize the training process across thousands units (Sirignano and Cont 2019). Except for Neural Networks that dominate this group, Random Forest have been used to identify high order cancellation rate impact on the market (McInish et al. 2019) or SVM to predict mid-sized movement (Kercheval and Zhang 2015).

In summary, ML is a good alternative to option pricing and limit order trading for mainly two reasons. Firstly, the current traditional models are prone to assumptions in contrary to ML models which learn by facts. Secondly, due to the high-dimensional nature of this application, current models can be slow to produce a reliable outcome in a highly volatile environment. Lastly, Neural Networks are recognized both for their demand and their ability to handle substantial amount of data; therefore, this characteristic could explain both the researchers' preference for neural networks as well as their success in this cluster.

3.2.5 Yellow cluster: risk management

This cluster comprises 76 articles credited with 632 citations, with the most prolific publication (Butaru et al. 2016) cited 81 times. Our analysis unveils two ML models, "Random Forest" (19 occurrences) and "Neural Networks" (14 occurrences). The most common topic related bi-grams are "Credit Risk" (16 occurrences) and "Risk Management" (15

occurrences), while tri-grams "Telematics Car Driving" (10 occurrences) and "Loss Given Default" (9 occurrences) conclude the focus of this cluster.

Academic research offers a plethora of indicators for achieving financial distress prediction through early warning systems. These indicators encompass consumer credit risk (Butaru et al. 2016; Zanin 2020), volatility (Laborda and Olmo 2021), transaction and connectivity network construction (Akbari et al. 2021; Laborda and Olmo 2021) in combination with traditional financial indicators. Robust modeling requires a substantial amount of data, while a crisis is a rare event thus creating data imbalance. To address this, the creation of synthetic data (Zanin 2020) and application of Synthetic Minority Oversampling Technique (SMOTE) (Lee 2020) are proposed to mitigate imbalances during the training phase of supervised ML models. In the context of short time horizons for credit card delinquency prediction, empirical comparisons have favored Random Forests and decision trees over logistic regression when evaluating and selecting models (Butaru et al. 2016). This preference for ML models over regression models is attributed to their superior accuracy, driven by their ability to handle non-linear data (Colak et al. 2020; Amini et al. 2021).

Risk management is of utmost importance for the insurance and lending industries, and publications on this topic are included in this cluster. In the insurance industry, the Random Forest model has demonstrated the capability to replace human predictions of future claim payments, yielding superior estimates (Ding et al. 2020). While the LASSO model can be used for the same task, it should be avoided when working with small datasets, as it can impact prediction accuracy (Devriendt et al. 2021). In addition, Neural Networks can process complex telematics car driving data to measure risk scores and identify driving styles and patterns, offering viable tools for insurance firms (Gao et al. 2022; Meng et al. 2022). Research on discrimination in the lending industry suggests that Random Forest is either able to capture structural relationships or uncover "identities" of minorities leading to a lower acceptance ratio for mortgage loans (Fuster et al. 2021). Loan officers may consider both 'hard' and 'soft' data, but the delinquency rate on approved loans by gradient boosting models is 33% lower compared to the decisions made by human experts.

Our literature review reveals an additional subcluster focused on the implementation of ML within the real estate domain. ML models are compared to linear regression models as an alternatives for tasks such as house pricing and rent estimation (Deppner and Cajias 2022), commercial pricing (Calainho et al. 2022) and renovation premiums (Mamre and Sommervoll 2022). The conclusion from those publications is that models like Random Forest, XGboost and Bagging outperform linear regression models. Calainho et al. (2022) attribute this performance to the combination of processing non-linear data with the non-parametric nature of those models.

In conclusion for this cluster, the linear nature of simple regression models emerges as a significant factor driving the adoption of tree-based ML algorithms. Simultaneously, the structure of the data appears to play a crucial role in model selection. Additionally, when dealing with unbalanced data, various techniques should be applied in classification problems.

3.2.6 Purple cluster: bankruptcy prediction, credit risk

This cluster encompasses 70 articles cited 1,094 times, with the most prolific publication (Khandani et al. 2010) being cited 266 times. Our analysis reveals the prominence of three ML models: "Neural Networks" (35 occurrences), "Support Vector Machine" (27 occurrences) and "Random Forest" (23 occurrences). The most common topic-related bi-grams

are "Bankruptcy Prediction" (32 occurrences), "Financial Ratios" (31 occurrences), "Credit Risk" (27 occurrences), "Banking Crisis" (25 occurrences) and "Early Warning" (25 occurrences).

In bankruptcy prediction, the focus is on the identification of the correct set of attributes able to forecast imminent delinquency but a variety of methods and financial features are selected in publications. In (Mselmi et al. 2017) several ML models are compared against the statistical logit model to predict financial distress in 212 French firms. The combination of SVM with partial least squares achieves a forecast accuracy of 94.28% for a two-year forecast compared to the 92.86% of the SVM model. SVM outperforms logit model in the default prediction of German firms as well, supplied by eight predictors (Chen et al. 2011). Lahmiri and Bekiros (2019) explore the use of qualitative, rather than quantitative, data and highlight that statistical model assumptions, such as multivariate normality, are often violated.

In credit risk subcluster, consumer behavior is treated as an important factor for predicting financial distress. ML models are capable of forecasting delinquency between 3 and 12 months in advance, even when a small training population is provided (Khandani et al. 2010). Moreover, the ElasticNet regression model can enhance the performance of ML models by identifying the most significant features for credit score classification (Xu et al. 2019). SVM is a common choice for credit analysis (Yu et al. 2020; Ala'raj et al. 2018), however in case of small datasets, logistic regression may be a better alternative (Ala'raj et al. 2018).

In the banking sector, rating agencies frequently misclassify banks, highlighting the need for improved predictions (Viswanathan et al. 2020). Traditional econometric tools for banking crisis assume that individual factors can explain their occurrence. In contrast, Duttagupta and Cashin (2011) through binary classification tree propose that a combination of factors must occur. Le and Viviani (2018) achieve small accuracy improvements using ML over traditional statistical models by measuring 31 different ratios from banking financial statements. In contrary, Beutel et al. (2019) promote logistic regression over ML models as the latter underperform in out-of-sample data in comparison to former. This observation, however, may be subject to overfitting, as a small dataset applied. Lastly, Viswanathan et al. (2020) classify 44 Indian banks through unsupervised K-means based on their credit risk using financial statements and ratios. They employ LDA algorithm for topic extraction to explain the clusters, and turn to supervised models such as Classification and Regression Tree (CART) and Random Forest for predicting credit ratings by comparing result to rating agencies. In light of these varying results, it is clear that consensus on the most important features for predicting banking crises has yet to be reached, necessitating further research.

Most publications within this group are referring to classification problems, therefore supervised models are chosen. SVM, Neural Networks and treed based models are commonly found however there is an indication that for small datasets, simpler statistical models may be more suitable.

3.2.7 Cyan cluster: asset pricing

This cluster consists of 43 publications accredited with 878 citations. Word analysis, indicated four prevailing bi-grams which are "Asset Pricing" (33 occurrences), "Cross Section" (25), "Stock Returns" 22 occurrences) and "Stock Market" (19 occurrences). As Fig. 6 panel F illustrates, this group relates to asset pricing, the Holy Grail for investors and financial institutes given that an accurate estimation of the fair price

allows for investment opportunities while minimizing risk. Many models have been proposed since the Capital Asset Pricing model (CAPM) was first introduced, as previous research revealed empirical failures (Karolyi and Van Nieuwerburgh 2020) and a growing number of anomalies incorporated into newer models (Geertsema and Lu 2020). The three-factor model (Fama and French 1993) was replaced by, among others, the four-factor model (Hou et al. 2015) and the five-factor model (Fama and French 2015). However, those models fail to encapsulate the full spectrum of cases as the problem is high-dimensional and, therefore, a larger number of characteristics is needed (Kozak et al. 2020). Although hundreds of estimators have been proposed for both cross-section and time series problems, often highly correlated and investigated under a linear prism (Weigand 2019; Gu et al. 2020), regression models require the incorporation of a priori knowledge of multiple predictors.

One proposed solution, is to combine a plethora of predictors, thereby generating new variables that effectively mitigate dimensionality and correlation issues (Fang et al. 2020). Statistical tools have been used for factor selection but not for the construction of new ones. Fang et al. (2020), propose a ML approach combining multiple Neural Networks with a "prior knowledge" feature to create and select the best features for the prediction model. Azevedo and Hoegner (2022) demonstrate the significance of non-linearity and high dimensionality. They achieved nearly 2% monthly returns using the Gradient Boosting model, outperforming both linear and traditional models such as the CAPM, the four-factor (Hou et al. 2015), the three-factor (Fama and French 1993) and the five-factor (Fama and French 2015) models. Non-linearities seems to be an important factor for predicting abnormal bond returns thus ML may be useful in prediction of asset price movements (Bianchi et al. 2020). Embracing market anomalies as possible predictors of excess returns has also been studied with promising results (Kozak et al. 2020; Dong et al. 2021). Geertsema and Lu (2020), through unsupervised Hierarchical Clustering, clustered anomalies and tested 41 factors to identify the ones that can explain all of them concluding with nine factors that scored higher in the Sharp ratio.

In most publications, supervised models are employed for both regression and classification problems with Neural Networks being the most common choice. The results of the aforementioned studies indicate an outperformance of ML over simple regression and other traditional models, due to their ability to handle high dimensional non-linear data (Table 5).

3.2.8 Cluster distribution per year

To gain more insights on the academic focus on ML in A&F, we conduct clustering distribution per year. Early access publications and articles not included in the main clusters are omitted. During the first years of experimentation with ML, we find that time-series topic was of sole interest in academic research. Bankruptcy prediction and textual analysis have also been introduced since 2006 while interest in risk management sparked after a decade. In 2021 all clusters increased significantly while in 2022 we find a slowdown for option and limit order trading with 11.5% decrease and a modest 14.28% increase for asset pricing in comparison to other groups. Noticeably, during the last 3 years risk management had 520% increase, textual analysis 300% and time-series forecasts 245% (Table 6).

Table 5 Cluster topics

Cluster	Cluster focus	Major topics	TC	Most cited articles Title	Authors	Year	TC	CPY
Red	Markets and time-series fore-casting	Covid-19, Crypto, Asset Allocation/portfolio creation, automated trading/trading behavior, volatility, Forex	4149	Modeling the conditional distribution of interest rates as a regime-switching process	Gray, SF	1996	720	26.67
				Stock market overreaction to bad news in good times: A rational expectations equilibrium model	Veronesi, P	1999	447	18.63
				Network topology of the interbank market	Boss, M; Elsinger, H; Summer, M; Thurner, S	2004	439	23.11
				Using genetic algorithms to find technical trading rules	Allen, F; Karjalainen, R	1999	428	17.83
				COVID-19 and stock market volatility: An industry level analysis	Baek, S; Mohanty, SK; Glambosky, M	2020	196	65.33
Green	Textual analysis	Textual Analysis, Sentiment Trading, Financial Misreporting	2695	The Information Content of Forward-Looking Statements in Corporate Filings-A Naive Bayesian Machine Learning Approach	Li, Feng	2010	501	38.54
				News implied volatility and disaster concerns	Manela, Asaf; Moreira, Alan	2017	198	33.00
				How Valuable Is FinTech Innovation?	Chen, Mark A.; Wu, Qinxi; Yang, Baozhong	2019	174	43.50
				Intraday online investor sentiment and return patterns in the US stock market	Renault, Thomas	2017	152	25.33
				Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms	Perols, Johan	2011	108	9.00

Table 5 (continued)

Cluster	Cluster focus	Major topics	TC	Most cited articles		Year	TC	CPY
				Title	Authors			
Blue	Options, Limit Order Trading	Limit Order Trading, Options	904	Deep hedging	Buehler, H.; Gonon, L.; Teichmann, J.; Wood, B	2019	95	23.75
				Universal features of price formation in financial markets: perspectives from deep learning	Sirignano, Justin; Cont, Rama	2019	78	19.50
				Modelling high-frequency limit order book dynamics with support vector machines	Kercheval, Alec N.; Zhang, Yuan	2015	66	8.25
				Application of data clustering and machine learning in variable annuity valuation	Gan, Guojun	2013	64	6.40
				Quant GANs: deep generation of financial time series	Wiese, Magnus; Knobloch, Robert; Korn, Ralf; Kretschmer, Peter	2020	54	18.00

Table 5 (continued)

| Cluster | Cluster Focus | TC | Major topics | Most cited articles | | | | |
				Title	Authors	Year	TC	CPY
Yellow	Risk Management	632	Risk Management, Insurance, Real Estate	Risk and risk management in the credit card industry	Butaru, Florentin; Chen, Qingqing; Clark, Brian; Das, Sanmay; Lo, Andrew W.; Siddique, Akhtar	2016	81	11.57
				Volatility spillover between economic sectors in financial crisis prediction: Evidence spanning the great financial crisis and Covid-19 pandemic	Laborda, Ricardo; Olmo, Jose	2021	39	19.50
				Machine learning improves accounting estimates: evidence from insurance payments	Ding, Kexing; Lev, Baruch; Peng, Xuan; Sun, Ting; Vasarhelyi, Miklos A	2020	38	12.67
				Predictably Unequal? The Effects of Machine Learning on Credit Markets	Fuster, Andreas; Goldsmith-Pinkham, Paul; Ramadorai, Tarun; Walther, Ansgar	2022	35	35.00
				Subjectivity in sovereign credit ratings	De Moor, Lieven; Luitel, Prabesh; Sercu, Piet; Vanpee, Rosanne	2018	29	5.80

Table 5 (continued)

| Cluster | Cluster Focus | Major topics | TC | Most cited articles | | | | |
				Title	Authors	Year	TC	CPY
Purple	Bankruptcy Prediction, Credit Risk	Bankruptcy Prediction, Credit Risk, Banking Crysis, Early Warning Systems	1094	Consumer credit-risk models via machine-learning algorithms	Khandani, Amir E.; Kim, Adlar J.; Lo, Andrew W	2010	266	20.46
				Identifying excessive credit growth and leverage	Alessi, Lucia; Detken, Carsten	2018	77	15.40
				Financial distress prediction: The case of French small and medium-sized firms	Mselmi, Nada; Lahiani, Amine; Hamza, Taher	2017	74	12.33
				Anatomy of banking crises in developing and emerging market countries	Duttagupta, Rupa; Cashin, Paul	2011	62	5.17
				Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios	Le, Hong Hanh; Viviani, Jean-Laurent	2018	46	9.20
Cyan	Asset Pricing	Asset Pricing	878	Empirical Asset Pricing via Machine Learning	Gu, Shihao; Kelly, Bryan; Xiu, Dacheng	2020	363	121.00
				Shrinking the cross-section	Kozak, Serhiy; Nagel, Stefan; Santosh, Shrihari	2020	104	34.67
				Bond Risk Premiums with Machine Learning	Bianchi, Daniele; Buchner, Matthias; Tamoni, Andrea	2021	59	29.50
				Machine Learning for Stock Selection	Rasekhschaffe, Keywan Christian; Jones, Robert C	2019	52	13.00
				Machine learning in the Chinese stock market	Leippold, Markus; Wang, Qian; Zhou, Wenyu	2022	47	47.00

Table 6 Cluster evolution over time

	Cluster					
	Markets and time-series forecasts	Textual analysis	Options, limit order trading	Bankruptcy prediction, credit risk	Asset pricing	Risk management
1996–2004	6					
2006–2010	7	3	1	3		
2011–2013	10	1	2	2	1	
2014–2016	17	6	6	4	1	2
2017–2019	35	13	12	20	5	12
2020–2022	87	76	60	33	32	55
TP	162	99	81	62	39	69
TC	4136	2668	900	1079	874	622
C/P	25.53	26.95	11.11	17.40	22.41	9.01
Rank (TP)	1	2	3	5	6	4
Rank (TC)	1	2	4	3	5	6
Rank (C/P)	2	1	5	4	3	6

TP = Total Production; TC = Total Citations; C/P = Citations per Publication

3.2.9 Cluster conclusions

The six clusters identified the plethora of problems where ML can be applied, either in conjunction with traditional models or independently. Their ability to process high dimensional, non-linear data are among the top factors for promoting them as good candidates in A&F discipline. Supervised models are the most frequently applied indicating that regression and classification problems types are the most common. We also find 72 clustering applications while reinforcement learning is barely employed. In Table 7, cluster key points are summarized. Lastly, our initial analysis indicated 2 smaller clusters composed of 4 and 2 papers respectively that we refrain from analyzing due to their size.

3.3 (RQ3) What are the future avenues of ML in A&F research?

Foremost, the analysis of clustering distribution over time provides valuable insights into research trends that elaborate on addressing our third research question. To further enhance our understanding of future avenues, we employ co-word analysis through VOSviewer. This method facilitates the visualization of interconnection between common words and importantly their average year of occurrence. For our purpose, following similar studies (e.g., Burton et al. 2020; Rojas-Lamorena et al. 2022) we select author specified keywords identified at least four times in our corpus thus focusing on the most important topics. As Fig. 7 depicts, each keyword is assigned a color indicating the average year of occurrence; darker color palette denotes years prior to 2019 while vibrant colors relate to last couple of years. Our results reveal both ML algorithms application and topics progression that we analyze under three periods. Given the number of publications in the early years, we observe that only since 2017 author-specified keywords occur at least four times.

Table 7 Summary of cluster challenges and solutions

| Cluster	Proposed algorithms	Key challenges	Proposed solutions
Markets & Time-Series Forecasting	Neural Networks SVM Tree-Based Models Genetic Algorithm Hierarchical Clustering	High Frequency High Dimensionality Non-Linearity	Combination of ML with Traditional Models non-parametric ML
Textual Analysis	Neural Networks SVM LDA	Textual cleanup	Bag-Of-Words Word Embedding Dictionary Creation
Options, Limit Order Trading	Neural Networks	Traditional Model Limitations High Frequency High Dimensionality	Combination of ML with Monte Carlo Supervised ML
Risk Management	Random Forest Neural Networks	Data Imbalances Non-Linearity	Synthetic Data Usage of SMOTE Supervised non-linear models
Bankruptcy Prediction, Credit Risk	Neural Networks SVM Random Forest	Small Dataset No Current Predictor Consensus	ML Models Able to be Trained with Small Datasets
Asset Pricing	Neural Networks	No Current Predictor Consensus High Dimensionality Non-Linearity	Supervised non-linear models Combination of Predictors
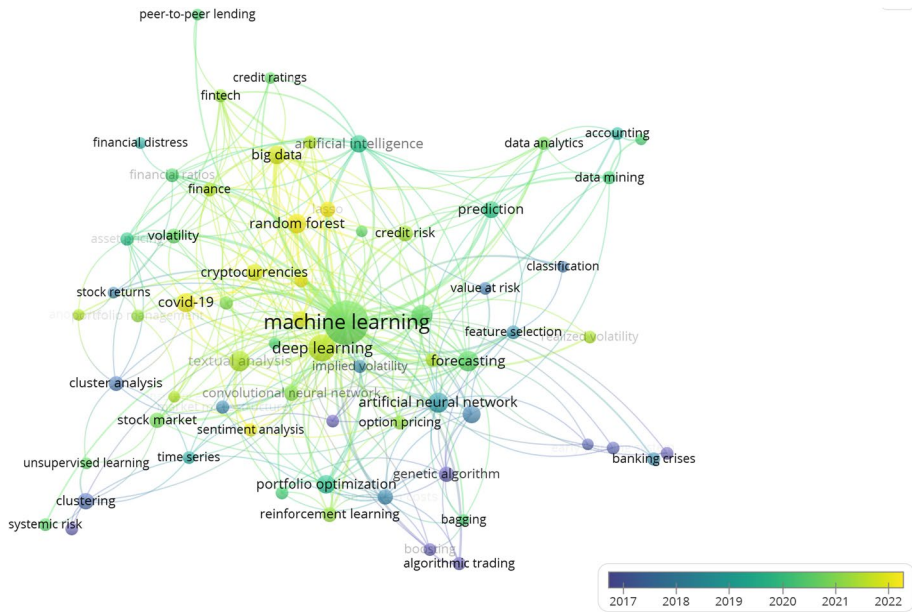
Fig. 7 Co-Word Analysis of author specified keywords

Starting with the first period, until 2018, the key focus of research can be categorized under market returns and crisis prediction. Topic keywords during this period are algorithmic trading (Allen and Karjalainen 1999; Creamer 2012; Cont and Kukanov 2016), stock returns (Constantinou et al. 2006), exchange rate (Amat et al. 2018), implied volatility (Bekiros and Georgoutsos 2008; Manela and Moreira 2017), early warning systems (Joy et al. 2016; Alessi and Detken 2018), and banking crises (Duttagupta and Cashin 2011; Alessi and Detken 2018). Moreover, we notice a plethora of clustering and classification applications during this timeframe where Genetic Programming (Payne and Tresl 2014; Karathanasopoulos et al. 2015), Neural Networks (Fioramanti 2008; Chen et al. 2013) and boosting algorithms (Creamer and Freund 2010; Creamer 2015) were widely adopted.

Between 2019 and 2020, publications related to time-series forecasting (Wiese et al. 2020), asset pricing (Calomiris and Mamaysky 2019; Weigand 2019) and data mining (Anouze and Bou-Hamad 2019) were introduced. Volatility and risk remain prominent themes in the literature, as indicated by keywords such as credit ratings (Abedin et al. 2019; Xu et al. 2019), systemic risks (Arakelian et al. 2019; Dungey et al. 2020), peer-to-peer lending (Jagtiani and Lemieux 2019; Zanin 2020) and financial distress (Gkillas et al. 2020; Samitas et al. 2020). Regarding ML keywords, Neural Networks (Mäkinen et al. 2019; Sun 2019) and bagging algorithms (Pace and Hayunga 2019) are proposed while, although limited, reinforcement learning emerged (Buehler et al. 2019; Wang and Zhou 2020).

During the last couple of years, research has shifted in topics driven by technological breakthroughs. Cryptocurrency markets (Anghel 2021; El Montasser et al. 2022), big data (Xue et al. 2021; Obaid and Pukthuanthong 2022), fintech (Han and Kim 2021b), textual analysis (Ongsakul et al. 2021; Aziz et al. 2022) and sentiment extraction (Liu et al. 2021; Obaid and Pukthuanthong 2022) are advocates of this trend. Publications related to Covid-19 pandemic and its impact on financial markets (Guo et al. 2021) and banks response

(Talbot and Ordonez-Ponce 2022) is also a main consideration of academic research during last two years. Traditional themes are also explored such as portfolio management (Mahmoudi et al. 2021; Pun and Wang 2021), realized volatility (Engle et al. 2021; Lu et al. 2022), and option pricing (Chataigner et al. 2021; Bayer et al. 2022). In addition to Neural Networks and deep learning (Jiang et al. 2022), recent research expanded by incorporating additional ML models such as Random Forest (Laborda and Olmo 2021; Ghosh et al. 2022), SVM (Petridis et al. 2022) and LASSO (Devriendt et al. 2021; Shahzad et al. 2022).

From the co-words analysis, we observe that volatility, risk management, and price forecasting are predominant topics throughout our selected timeframe, suggesting the potential continuation of this trend in the near future. Neural Networks are expected to maintain a central position in research; however, we anticipate the emergence of specialized algorithms tailored for specific tasks, departing from the prevalent broader-purpose algorithms. Our observation is based on the increased number of publications on journals such as Quantitative Finance which excel in this field e.g. (Horvath et al. 2021; Kim et al. 2022). Furthermore, our analysis in RQ1 indicates a shift in the application of ML methodologies from the US market towards the Asian and global markets, indicating an evolving landscape for research focus.

As shown above, ML has multiple advantages over traditional methods. However, there is still a debate on how it can be adapted to current strict regulations on A&F and how it can be implemented given the lack of literacy on data analysis of the existing workforce in those fields. We believe that further research will be conducted to enhance the traceability of decision-making of the algorithms as well as to identify the actions needed to support new or updated regulatory requirements. We argue that practitioners in A&F should be more technology inclined and be able to work alongside advanced automation tools to enhance decision-making capabilities.

4 Challenges and opportunities in advancing applications of ML

As we have seen in the previous sections, ML models have gained significant popularity in the recent years. However, researchers should be familiar with model's shortcomings before employing them in any application, whether in the fields of A&F or any other field. In addition, there are also some common pitfalls that researchers should avoid. In this section we address the above issues as well as the most common approaches to overcome or alleviate them. Understanding these constraints is crucial for informed and effective utilisation of these models in practical scenarios in both A&F. In addition, we present various ideas for future research in these areas.

4.1 Can we enhance the transparency of ML approach?

Enhancing transparency in ML models is a crucial concern, especially in fields like A&F, where interpretability is essential for regulatory compliance and risk assessment. While the term "Black Box" is commonly used in ML, it's worth noting that certain ML algorithms, such as Wavelet Networks (considered "grey boxes") and Genetic Programming (considered "white boxes"), offer a degree of transparency.

Recently, various methods to improve model interpretability have been proposed, such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic

Explanations), and surrogate models. In our literature review, SHAP stands out as the most frequently employed approach. It is based on cooperative game theory to allocate contribution of each feature to the prediction (Lundberg and Lee 2017). LIME is used for instance-level interpretations, and it uses a simple model to approximate a complex one (Ribeiro et al. 2016). A global surrogate model is a simple model that approximates a black-box model. Other common approaches dedicated to CNNs are Features Visualization, that uses neuron activation maximization to visualize learned features, and Network Dissection (Bau et al. 2017).

Future research could focus on developing novel techniques for model explainability and visualization tailored specifically to A&F applications. This might involve creating domain-specific interpretability tools that financial analysts and auditors can use effectively (Bertomeu 2020). A clear and transparent decision process for financial managers, accountants and auditors will further enhance the adaptation and acceptance of ML by practitioners.

4.2 Can we conduct any statistical inference and hypothesis test based on ML?

The majority of the papers that employ ML methods focus on point forecasts while statistical inference and prediction intervals are overlooked. However, financial managers are also interested in prediction intervals, statistical inference, and hypothesis testing. Traditional statistical methods often rely on assumptions like normality, linearity, and independence that may not hold in ML models, (Alexandridis and Zapranis 2014). However, recent research has shown ways to conduct hypothesis tests and statistical inference with ML. Techniques like permutation tests and bootstrap methods can be applied to assess the significance of model features or compare different models (Efron and Tibshirani 1994; Zapranis and Refenes 1999). To our knowledge, the only approach to a complete statistical framework is presented in Alexandridis and Zapranis (2013) and Alexandridis and Zapranis (2014). The authors provide a complete model identification framework for a class of Neural Networks called Wavelet Networks, but the methodology is applicable to every family of Neural Networks. More precisely, a model selection procedure as well as a statistical variable significance and a statistical variable selection framework are presented. The methodology is based on a Sensitivity Based Pruning criterion and bootstrap techniques. Furthermore, the authors provide a framework for prediction intervals based on the Bagging and Balancing techniques. The drawback of the previous algorithms is that are based on bootstrapping techniques and hence they are time consuming. Future work could focus on refining these statistical techniques to make them more applicable to ML in A&F. Additionally, exploring the theoretical underpinnings of statistical properties of ML models could provide further insights.

4.3 How ML can help decision making in social science?

ML can be applied to data analysis and pattern recognition, identifying relationships between variables to gain deeper insights into human behaviour, societal trends, and interactions, thereby informing more robust decision-making (Lv et al. 2020). Similarly, ML can be used in predictive modelling in fields such as economic forecasting, crime prediction, and disease outbreak modelling, guiding policymakers in making informed decisions (Kleinberg et al. 2015). Text and sentiment analysis can be employed to analyse social media posts, surveys, reviews, and other textual data to understand public perceptions,

attitudes, and sentiments toward various social issues, policies, products, and services, ultimately shaping decision-making strategies (Shrestha et al. 2021). Personalisation and recommender systems can be used in areas like education (suggesting courses), healthcare (recommending treatment plans), and policy-making (tailoring interventions to specific groups) (Sarker 2021).

ML can also contribute to casual inference to identify how one variable can affect another, to network analysis to identify influential nodes community structures, and to information flow patterns, aiding the understanding of social dynamics and communication patterns and policy analysis and simulation to assess potential outcomes before implementing policies, helping to make more informed decisions and reduce unintended consequences (Sarker 2021). In the areas of healthcare and social services sector, ML aids in resources allocation by predicting demand, identifying high-risk populations, and suggesting personalised interventions (Hoffman and Podgurski 2019). Finally, ML could be employed in ethics and bias detection to help identify biases in datasets and models, promoting fairness and ethical considerations in decision-making processes. It can also help in identifying potential discriminatory outcomes of certain policies or interventions (Di Maggio et al. 2022).

Successful application of ML in social science, requires careful consideration of ethical and privacy issues, as well as transparency and interpretability in order to avoid any biases. As it shown in Di Maggio et al. (2022) minorities were more likely to either be denied credit or be granted credit on unfavourable terms while Hoffman and Podgurski (2019) reports similar algorithmic discrimination issues in health care. Collaboration between domain experts and data scientists is crucial to ensure that ML techniques are applied in a responsible and meaningful way in social science research and decision-making.

Future research should aim to improve model accuracy and robustness in social science applications. Furthermore, it should aim to develop ethical and responsible AI practices to address potential biases in data. Finally, should aim to foster interdisciplinary collaborations between ML experts and social scientists to ensure the relevance and validity of models.

4.4 How can we control the overfitting and over-parameterization issues?

One of the most crucial steps in ML is to identify the correct topology of the model. For example, in Neural Networks a desired architecture should contain as few hidden units (HUs) (or neurons) as necessary while at the same time it should explain as much variability of the training data as possible. A network with less HUs than needed would not be able to learn the underlying function while selecting more HUs than needed will result to an over-fitted model.

The usual approaches proposed in the literature are the early stopping, regularization, Bayesian regularization, L1 and L2 regularization, brute force pruning and irrelevant connection elimination and are commonly used to mitigate these issues. Some other prevention practices include feature pruning by eliminating irrelevant or non-significant dimensions for selected dataset, embedding additional distinct cases in training set and data augmentation in which models separate important features from noise. Some ML models include prevention mechanisms such as Boosting and Bagging models, as well as random node dropouts for Neural Networks.

The previous methods do not use an optimal architecture of a model. A very large model is used and then various methods were developed to avoid over-fitting. Smaller networks

usually are faster to train and need less computational power to build (Reed 1993). Detection tools take into account the accuracy difference between the training and the validation sample. The training stops when the error in the validation sample increases. Other methods include bootstrapping or *k*-fold cross-validation techniques to improve the generalisation of the model. Others propose ad-hoc rules like the observations to parameters ratio should be over a specific number, e.g. five. This is similar to usually requiring around 30 observations for linear models. Zapranis and Refenes (1999) and Alexandridis and Zapranis (2013) propose the Minimum Prediction Risk criterion for the optimal selection of neurons in a Neural Network.

Future research can explore novel regularization techniques tailored to A&F data, which often have specific characteristics such as time-series dependencies, imbalanced classes, or high dimensionality.

4.5 To what extent is ML sensitive to training data? Is there any robustness?

ML models can be sensitive to training data especially in cases of small or unrepresentative datasets. This sensitivity is not confined to any specific model type; it applies universally, whether the models are linear or nonlinear, parametric or nonparametric, when the sample used for model training fails to adequately represent the overall population. Moreover, in classification problems, test set population ideally should be chosen to be evenly split between groups, otherwise careful consideration is required in order to establish the appropriate cut-off point for the classifier to ensure robust and accurate classification.

Another common issue in the literature where ML is applied in A&F is the omission of the data pre-processing step. Trend and periodicities should be removed from the data, and appropriate techniques to treat outliers should be applied. Finally, as in the case of statistical models, ML models are also affected by structural breaks in the data.

Researchers can assess robustness by performing sensitivity analyses, using techniques like dropout, and applying adversarial testing. Another approach is to create bootstrapped versions of the training sample, train a different model on each sample and then use an amalgamation of the predictions. Finally, adaptive models have been applied to update online the architecture of ML models to account for structural breaks or jumps in the data or any other change in the data generating process (e.g., Cao and Tay 2003; Lin et al. 2006).

Achieving complete robustness in all situations is often a challenging and ongoing process. Addressing issues such as data quality, bias, and data drift can help make ML models more resilient to variations in input data. Future work could involve developing techniques for training models that are inherently more robust to variations in training data distribution and exploring methods for model robustness evaluation.

4.6 How researchers fine-tune the hyperparameters in A&F data?

Fine-tuning hyperparameters in ML models is often done using techniques like grid search, random search, or Bayesian optimization. Usually, these approaches are coupled with sampling techniques like bootstrap or cross-validation. However, domain-specific knowledge is crucial in determining appropriate parameter settings.

Future research could focus on automated hyperparameter tuning methods tailored to A&F datasets. Additionally, developing domain-specific guidelines for hyperparameter tuning can help researchers and practitioners navigate this process effectively.

5 Conclusions

ML offers several advantages over traditional methods currently employed in A&F, including the extraction of features, pattern recognition, processing of high-dimensional data and handling nonlinearity. Our paper sheds light on the current state of research and applications in ML while also suggesting new paths for further investigation. Through both literature review and bibliographic coupling, we explored three research questions.

In our first research question, we identified a surge in this research field since 2015 which continues up to date. ML applications in finance constitute 88.38% of our corpus, while 16.35% of papers are published in the Quantitative Finance journal. Among the 67 countries associated with our research, Austria and the United States of America emerged as the most cited in contrast to Japan and Thailand that rank at the bottom of the list.

In the second research question, we constructed six clusters through bibliographic coupling and analyzed them using a Bag-of-words technique and literature review. This leads us to conclusions about current challenges, key ML algorithms proposed, and the evolution that this new technology is bringing about. The strongest assets of the new models lie in their ability to handle multi-dimensionality, non-linearity and multiple sources of information such as images and textual analysis. Neural networks, SVM, and tree based algorithms proved to be effective in a plethora of applications as long as enough training data are at hand; otherwise out of sample predictions may exhibit lower accuracy than traditional models. The most commonly employed models are supervised, while unsupervised models are predominantly used for clustering and topic extraction using the LDA algorithm. Notably, in the past three years, there has been extensive exploration in topics associated with risk management, textual analysis, and time-series forecasts.

To address our third research question, we have conducted a co-word analysis on author keywords, revealing the exploration of volatility, risk management, and price forecasting since the early stages of academic research in ML within the A&F discipline. Furthermore, we tried to identify the future direction of research in the fields of ML and A&F. Our results indicate that the trend in the above topics will continue while we expect the development of more advanced methodologies that are also more tailored to specific applications in the area of A&F. Finally, we examine analytically the limitations of ML algorithms. We present the most common approaches proposed in the literature to alleviate these issues, where possible, as well as directions for future research in these areas.

Appendix: "Search Query"

ALL = (("Machine Learning" or "Supervised Learning" or "Unsupervised Learning" or "Nearest Neighbor" or "Clustering" or "Support Vector Machine" or "Random Forest" or "Classification Tree" or "Deep Learning" or "Convolutional Neural Network" or "Artificial Neural Network" or "Recurrent Neural Network" or "Long-Short Term Memory" or "Ensemble Methods" or "Radial Basis Function Network" or "Kernel-based Extreme Learning" or "Feed-Forward Deep Network" or "Genetic Algorithm"

or "Particle Swarm Optimization" or "Agent-Based Algorithmic Learning" or "Wavelet-Based Neural Networks" or "Ensemble Empirical Mode Decomposition" or "Data Fluctuation Network" or "Soft Computing" or "Simulated-Based Neural Network") and ("financ*" or "capital" or "leverage" or "accounting" or "accountancy" or "ledger" or "irs" or "tax" or "asset" or "audit*" or "stock" or "payroll" or "payable")).

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Ethics approval was not required for this systematic review.

Competing interests The authors state that there is no competing interests.

References

Abedin MZ et al (2019) Topological applications of multilayer perceptrons and support vector machines in financial decision support systems. Int J Financial Econ 24(1):474–507

Aboussalah AM, Xu Z, Lee C-G (2021) What is the value of the cross-sectional approach to deep reinforcement learning? Quant Finance 22(6):1091–1111

Achakzai MA, Peng J (2023) Detecting financial statement fraud using Dynamic Ensemble Machine Learning. Int Rev Financial Anal 89:102827

Aggarwal D, Chandrasekaran S, Annamalai B (2020) A complete empirical ensemble mode decomposition and support vector machine-based approach to predict Bitcoin prices. J Behav Exp Finance 27:100335

Ahmed S, Alshater MM, Ammari AE, Hammami H (2022) Artificial intelligence and machine learning in finance: a bibliometric review. Res Int Bus Finance 61:101646

Akbari A, Ng L, Solnik B (2021) Drivers of economic and financial integration: a machine learning approach. J Empir Financ 61:82–102

Ala'raj, M., Abbod, M. & Radi, M. (2018) The applicability of credit scoring models in emerging economies: an evidence from Jordan. Int J Islam Middle East Financ Manag 11(4):608–630

Alessi L, Detken C (2018) Identifying excessive credit growth and leverage. J Financ Stab 35:215–225

Alexandridis AK, Zapranis AD (2013) Wavelet neural networks: a practical guide. Neural Netw 42:1–27

Alexandridis AK, Zapranis AD (2014) Wavelet networks: methodologies and applications in financial engineering, classification and chaos. Wiley, New Jersey

Allen F, Karjalainen R (1999) Using genetic algorithms to find technical trading rules. J Financial Econ 51(2):245–271

Al-Maadid A, Alhazbi S, Al-Thelaya K (2022) Using machine learning to analyze the impact of coronavirus pandemic news on the stock markets in GCC countries. Res Int Bus Finance 61:101667

Amat C, Michalski T, Stoltz G (2018) Fundamentals and exchange rate forecastability with simple machine learning methods. J Int Money Financ 88:1–24

Amini S, Elmore R, Öztekin Ö, Strauss J (2021) Can machines learn capital structure dynamics? J Corp Finance 70:102073

Anghel D-G (2021) A reality check on trading rule performance in the cryptocurrency market: machine learning vs. technical analysis. Finance Res Lett 39:101655

Anon (2021a) Chartered Assosiation of Business Schools. [Online]. Available at: https://charteredabs.org/academic-journal-guide-2021/

Anon (2021b) VOSviewer—visualizing scientific landscapes. [Online]. Available at: https://www.vosviewer.com/

Anouze AL, Bou-Hamad I (2019) Data envelopment analysis and data mining to efficiency estimation and evaluation. Int J Islam Middle East Financ Manag 12(2):169–190

Arakelian V, Dellaportas P, Savona R, Vezzoli M (2019) Sovereign risk zones in Europe during and after the debt crisis. Quant Finance 19(6):961–980

Aria M, Cuccurullo C (2017) Bibliometrix: an R-tool for comprehensive science mapping analysis. J Inf 11(4):959–975

Azevedo V, Hoegner C (2022) Enhancing stock market anomalies with machine learning. Rev Quant Finance Acc 60:195–230

Azimi M, Agrawal A (2021) Is positive sentiment in corporate annual reports informative? Evidence from deep learning. Rev Asset Pricing Stud 11(4):762–805

Aziz S, Dowling M, Hammami H, Piepenbrink A (2022) Machine learning in finance: a topic modeling approach. Eur Financ Manag 28:744–770

Baek S, Mohanty SK, Glambosky M (2020) Covid-19 and stock market volatility: an industry level analysis. Finance Res Lett 37:101748

Baker M, Wurgler J (2002) Market timing and capital structure. J FInance 57(1):1–32

Baker HK, Kumar S, Pattnaik D (2021) Twenty-five years of the Journal of Corporate Finance: a sciento-metric analysis. J Corp Finance 66:101572

Bao Y et al (2020) Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. J Acc Res 58(1):199–235

Bau D, et al (2017) Network dissection: quantifying interpretability of deep visual representations. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 6541–6549

Bayer C, Qiu J, Yao Y (2022) Pricing options under rough volatility with BACKWARD SPDES. SIAM J Financial Math 13(1):179–212

Behera J, Pasayat AK, Behera H (2022) Covid-19 vaccination effect on stock market and death rate in India. Asia-Pac Finan Mark 29(4):651–673

Bekiros SD, Georgoutsos DA (2008) Non-linear dynamics in financial asset returns: the predictive power of the CBOE volatility index. Eur J Finance 14(5):397–408

Bertomeu J (2020) machine learning improves accounting: discussion, implementation and research opportunities. Rev Acc Stud 25(3):1135–1155

Bertomeu J, Cheynel E, Floyd E, Pan W (2020) Using machine learning to detect misstatements. Rev Acc Stud 26(2):468–519

Beutel J, List S, von Schweinitz G (2019) Does machine learning help us predict banking crises? J Financial Stab 45:100693

Bianchi D, Büchner M, Tamoni A (2020) Bond risk premiums with machine learning. Rev Financial Stud 34(2):1046–1089

Black F, Litterman R (1992) Global portfolio optimization. Financ Anal J 48(5):28–43

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

Blankespoor E, Hendricks BE, Miller GS (2023) The pitch: managers' disclosure choice during initial public offering roadshows. Acc Rev 98(2):1–29

Blei DM, Lafferty JD (2006) Dynamic topic models. In: ICML '06: Proceedings of the 23rd international conference on Machine learning, pp 113–120

Bochkay K, Hales J, Chava S (2019) Hyperbole or reality? Investor response to extreme language in earnings conference calls. Acc Rev 95(2):31–60

Bouchaud J-P, Mézard M, Potters M (2002) Statistical properties of stock order books: empirical results and Models. Quant Finance 2(4):251–256

Boyack K, Klavans R (2010) Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the reaserch front most accurately? J Am Soc Inform Sci Technol 61(12):2389–2404

Bradrania R, Pirayesh Neghab D, Shafizadeh M (2021) State-dependent stock selection in index tracking: a machine learning approach. Fin Mark Portf Mgmt 36(1):1–28

Broadus RN (1987) Toward a definition of "bibliometrics." Scientometrics 12:373–379

Brown NC, Crowley RM, Elliott WB (2020) What are you saying? Using topic to detect financial misreporting. J Acc Res 58(1):237–291

Buehler H, Gonon L, Teichmann J, Wood B (2019) Deep hedging. Quant Finance 19(8):1271–1291

Burton B, Kumar S, Pandey N (2020) Twenty-five years of The European Journal of Finance (EJF): a retrospective analysis. Eur J Finance 26(18):1817–1841

Butaru F et al (2016) Risk and risk management in the credit card industry. J Bank Finance 72:218–239

Calainho FD, van de Minne AM, Francke MK (2022) A machine learning approach to price indices: applications in commercial real estate. J Real Estate Finance Econ

Calomiris CW, Mamaysky H (2019) How news and its context drive risk and returns around the world. J Financial Econ 133(2):299–336

Cao L, Tay F (2003) Support vector machine with adaptive parameters in financial time series forecasting. IEEE Trans Neural Netw 14(6):1506–1518

Cao Y, Liu X, Zhai J, Hua S (2022) A two-stage bayesian network model for corporate bankruptcy prediction. Int J Financial Econ 27(1):455–472

Chataigner M et al (2021) Short communication: beyond surrogate modeling: learning the local volatility via shape constraints. SIAM J Financial Math 12(3):SC58–SC69

Chen S, Ge L (2019) Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction. Quant Finance 19(9):1507–1515

Chen S, Härdle WK, Moro RA (2011) Modeling default risk with support vector machines. Quant Finance 11(1):135–154

Chen J-H, Chang T-T, Ho C-R, Diaz JF (2013) Grey relational analysis and neural network forecasting of reit returns. Quant Finance 14(11):2033–2044

Chen Y-J et al (2017) Enhancement of fraud detection for narratives in annual reports. Int J Acc Inf Syst 26:32–45

Chortareas G, Katsafados AG, Pelagidis T, Prassa C (2024) Credit risk modelling within the euro area in the covid-19 period: evidence from an icas framework. Int J Finance Econ

Christensen K, Siggaard M, Veliyev B (2022) A machine learning approach to volatility forecasting. J Financial Econ 21:1680–1727

Chun J, Ahn J, Kim Y, Lee S (2020) Using deep learning to develop a stock price prediction model based on individual investor emotions. J Behav Finance 22(4):480–489

Ciampi F, Giannozzi A, Marzi G, Altman E (2021) Rethinking SME default prediction: a systematic literature review and future perspectives. Scientometrics 126:2141–2188

Clarke R, de Silva H, Thorley S (2011) Minimum-variance portfolio composition. J Portf Manag 37:31–45

Cohen G (2023) Intraday algorithmic trading strategies for cryptocurrencies. Rev Quant Financial Acc 61(1):395–409

Colak G, Fu M, Hasan I (2020) Why are some Chinese firms failing in the US capital markets? A machine learning approach. Pac-Basin Finance J 61:101331

Constantinou E, Georgiades R, Kazandjian A, Kouretas GP (2006) Regime switching and artificial neural network forecasting of the Cyprus Stock Exchange Daily Returns. Int J Financial Econ 11(4):371–383

Cont R, Kukanov A (2016) Optimal order placement in limit order markets. Quant Finance 17(1):21–39

Creamer G (2012) Model calibration and automated trading agent for euro futures. Quant Finance 12(4):531–545

Creamer GG (2015) Can a corporate network and news sentiment improve portfolio optimization using the black–litterman model? Quant Finance 15(2):1405–1416

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

De Spiegeleer J, Madan DB, Reyners S, Schoutens W (2018) Machine learning for quantitative finance: fast derivative pricing, hedging and fitting. Quant Finance 18(10):1635–1643

Debener J, Heinke V, Kriebel J (2023) Detecting insurance fraud using supervised and unsupervised machine learning. J Risk Insur 90(3):743–768

DeMiguel V, Garlappi L, Uppal R (2007) Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Rev Financial Stud 22(5):1915–1953

Deppner J, Cajias M (2022) Accounting for spatial autocorrelation in algorithm-driven hedonic models: a spatial cross-validation approach. J Real Estate Finance Econ 68:235–273

Devriendt S, Antonio K, Reynkens T, Verbelen R (2021) Sparse regression with multi-type regularized feature modeling. Insur Math Econ 96:248–261

Ding K et al (2020) Machine Learning improves accounting estimates: evidence from insurance payments. Rev Acc Stud 25(3):1098–1134

Dong X, Li Y, Rapach DE, Zhou G (2021) Anomalies and the expected market return. J Finance 77(1):639–681

Dungey M, Islam R, Volkov V (2020) Crisis transmission: visualizing vulnerability. Pac Basin Finance J 59:101255

Dunis CL, Laws J, Karathanasopoulos A (2013) GP algorithm versus hybrid and mixed neural networks. Eur J Finance 19(3):180–205

Duttagupta R, Cashin P (2011) Anatomy of banking crises in developing and emerging market countries. J Int Money Finance 30(2):354–376

Efron B, Tibshirani R (1994) An introduction to the bootstrap, 1st edn. Chapman and Hall/CRC, New York

El Montasser G, Charfeddine L, Benhamed A (2022) Covid-19, cryptocurrencies bubbles and digital market efficiency: Sensitivity and similarity analysis. Finance Res Lett 46:102362

El-Haj M et al (2019) In search of meaning: lessons, resources and next steps for computational analysis of financial discourse. J Bus Financ Acc 46(3–4):265–306

Engle RF, Hansen MK, Karagozoglu AK, Lunde A (2021) News and idiosyncratic volatility: the public information processing hypothesis*. J Financ Economet 19(1):1–38

Fama EF, French KR (1993) Common risk factors in the returns on stocks and Bonds. J Financ Econ 33(1):3–56

Fama EF, French KR (2015) A five-factor asset pricing model. J Financ Econ 116(1):1–22

Fang J et al (2020) Neural network-based automatic factor construction. Quant Finance 20(12):2101–2114

Fioramanti M (2008) Predicting sovereign debt crises using artificial neural networks: a comparative approach. J Financial Stab 4(2):149–164

Frost G, Jones S, Yu M (2023) Voluntary carbon reporting prediction: a machine learning approach. Abacus 59(4):1116–1166

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

Fuster A, Goldsmith-Pinkham P, Ramadorai T, Walther A (2021) Predictably unequal? The effects of machine learning on credit markets. J Finance 77(1):5–47

Gan Q, Wei WC, Johnstone D (2015) A faster estimation method for the probability of informed trading using hierarchical agglomerative clustering. Quant Finance 15(11):1805–1821

Gao G, Meng S, Wüthrich MV (2022) What can we learn from Telematics car driving data: a survey. Insur Math Econ 104:185–199

Gaunt C (2014) Accounting and finance: authorship and citation trends. Acc Finance 2(54):441–465

Geertsema P, Lu H (2020) The correlation structure of anomaly strategies. J Bank Finance 119:105934

Ghoddusi H, Creamer GG, Rafizadeh N (2019) Machine learning in energy economics and finance: a review. Energy Econ 81:709–727

Ghosh P, Neufeld A, Sahoo JK (2022) Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. Finance Res Lett 46:102280

Gkillas K, Gupta R, Pierdzioch C (2020) Forecasting realized oil-price volatility: the role of financial stress and asymmetric loss. J Int Money Finance 104:102–137

Goodell JW, Kumar S, Lim WM, Pattnaik D (2021) Artificial Intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. J Behav Exp Finance 32:100577

Goudenège L, Molent A, Zanette A (2020) Machine learning for pricing american options in high-dimensional Markovian and non-Markovian models. Quant Finance 20(4):573–591

Gray SF (1996) Modeling the conditional distribution of interest rates as a regime-switching process. J Financial Econ 42(1):27–62

Gray GL, Chiu V, Liu Q, Li P (2014) The expert systems life cycle in AIS research: What does it mean for future AIS research? Int J Acc Inf Syst 15(4):423–451

Gu S, Kelly B, Xiu D (2020) Empirical asset pricing via machine learning. Rev Financial Stud 33(5):2223–2273

Guo Y, Li P, Li A (2021) Tail risk contagion between international financial markets during COVID-19 pandemic. Int Rev Financial Anal 73:101649

Han JJ, Kim H-J (2021a) Prediction of investor-specific trading trends in South Korean stock markets using a BILSTM prediction model based on sentiment analysis of financial news articles. J Behav Finance 24(4):398–410

Han JJ, Kim H-J (2021b) Stock price prediction using multiple valuation methods based on artificial neural networks for KOSDAQ IPO companies. Investig Anal J 50(1):17–31

Han H et al (2023) Accounting and auditing with blockchain technology and artificial intelligence: a literature review. Int J Acc Inf Syst 48:100598

Han W, et al (2024) The diversification benefits of cryptocurrency factor portfolios: Are they there? Rev Quant Finance Acc

Harvey C, Morris H, Kelly A, Rowlinson M (2010) Academic Journal Quality Guide

Heston SL (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financial Stud 6(2):327–343

Hoffman S, Podgurski A (2019) Artificial intelligence and discrimination in health care. Yale J Health Pol'y L. Ethics 19:1

Horvath B, Muguruza A, Tomas M (2021) Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models. Quant Finance 21(1):11–27

Hou K, Xue C, Zhang L (2015) Digesting anomalies: an investment approach. Rev Financial Stud 28(3):650–705

Hu MY, Tsoukalas C (1999) Combining conditional volatility forecasts using neural networks: an application to the EMS exchange rates. J Int Financial Mark Inst Money 9(4):407–422

Hu W, Zastawniak T (2020) Pricing high-dimensional American options by kernel ridge regression. Quant Finance 20(5):851–865

Huang AH, Wang H, Yang Y (2023) Finbert: a large language model for extracting information from financial text. Contemp Acc Res 40(2):806–841

Jagtiani J, Lemieux C (2019) The roles of Alternative Data and machine learning in fintech lending: evidence from the lendingclub consumer platform. Financial Manag 48:1009–1029

Jang H, Lee J (2019) Generative Bayesian neural network model for risk-neutral pricing of American index options. Quant Finance 19(4):587–603

Jiang K, Du X, Chen Z (2022) Firms' digitalization and stock price crash risk. Int Rev Financial Anal 82:102196

Joy M, Rusnák M, Šmídková K, Vašíček B (2016) Banking and currency crises: differential diagnostics for developed countries. Int J Financial Econ 22(1):44–67

Kaniel R, Lin Z, Pelger M, Van Nieuwerburgh S (2023) Machine-learning the skill of mutual fund managers. J Financial Econ 150(1):94–138

Karathanasopoulos A et al (2015) Stock market prediction using evolutionary support vector machines: an application to the ASE20 index. Eur J Finance 22(12):1145–1163

Karolyi GA, Van Nieuwerburgh S (2020) New methods for the cross-section of returns. Rev Financial Stud 33(5):1879–1890

Kercheval AN, Zhang Y (2015) Modelling high-frequency limit order book dynamics with support vector machines. Quant Finance 15(8):1315–1329

Khandani AE, Kim AJ, Lo AW (2010) Consumer credit-risk models via machine-learning algorithms. J Bank Finance 34(11):2767–2787

Kim H, Jun S, Moon K-S (2022) Stock market prediction based on adaptive training algorithm in machine learning. Quant Finance 22(6):1133–1152

Kleinberg J, Ludwig J, Mullainathan S, Obermeyer Z (2015) Prediction policy problems. Am Econ Rev 105(5):491–495

Korkeamäki T, Sihvonen J, Vähämaa S (2018) Evaluating publications across business disciplines: Inferring interdisciplinary "exchange rates" from Intradisciplinary author rankings. J Bus Res 84:220–232

Kozak S, Nagel S, Santosh S (2020) Shrinking the cross-section. J Financial Econ 135(2):271–292

Laborda R, Olmo J (2021) Volatility spillover between economic sectors in financial crisis prediction: evidence spanning the great financial crisis and covid-19 pandemic. Res Int Bus Finance 57:101402

Lahmiri S, Bekiros S (2019) Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design. Quant Finance 19(9):1569–1577

Le HH, Viviani J-L (2018) Predicting bank failure: an improvement by implementing a machine-learning approach to classical financial ratios. Res Int Bus Finance 44:16–25

Lee SC (2020) Addressing imbalanced insurance data through zero-inflated Poisson regression with boosting. ASTIN Bull 51(1):27–55

León C, Kim G-Y, Martínez C, Lee D (2017) Equity markets' clustering and the global financial crisis. Quant Finance 17:1905–1922

Levine-Clark M, Gil E (2021) A new comparative citation analysis: Google Scholar, Microsoft Academic, Scopus, and Web of Science. J Bus Finance Librariansh 26:145–163

Li F (2010) The information content of forward-looking statements in corporate filings—a Naive Bayesian machine learning approach. J Acc Res 48(5):1049–1102

Li T, Chen K, Feng Y, Ying Z (2016) Binary switch portfolio. Quant Finance 17(5):763–780

Lin F, Shieh H, Huang P (2006) Adaptive wavelet neural network control with hysteresis estimation for piezo-positioning mechanism. IEEE Trans Neural Netw 17(2):432–444

Lin E, Kao C-LM, Adityarini NS (2021) Data-driven tree structure for PIN models. Rev Quant Financial Acc 57(2):411–427

Liu H, Mulvey J, Zhao T (2015) A semiparametric graphical modelling approach for large-scale equity selection. Quant Finance 16(7):1053–1067

Liu F, Pantelous AA, von Mettenheim H-J (2018) Forecasting and trading high frequency volatility on large indices. Quant Finance 18(5):737–748

Liu M, Li G, Li J, Zhu X, Yao Y (2021) Forecasting the price of bitcoin using deep learning. Finance Res Lett 40:101755

Liu Q, Wang C, Zhang P, Zheng K (2021) Detecting stock market manipulation via machine learning: evidence from China Securities Regulatory Commission punishment cases. Int Rev Financial Anal 78:10188

Loughran T, McDonald B (2016) Textual analysis in accounting and finance: a survey. J Acc Res 54(4):1187–1230

Lu X, Ma F, Xu J, Zhang Z (2022) Oil futures volatility predictability: new evidence based on machine learning models. Int Rev Financial Anal 83:102299

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 30

Lv Z, Qiao L, Singh AK (2020) Advanced machine learning on cognitive computing for human behavior analysis. IEEE Trans Comput Soc Syst 8(5):1194–1202

MacCoun RJ (1998) Biases in the interpretation and use of reaserch results. Annu Rev Phychol 49:259–287

Di Maggio M, Ratnadiwakara D, Carmichael D (2022) Invisible primes: Fintech lending with alternative data

Mahmoudi A, Hashemi L, Jasemi M, Pope J (2021) A comparison on particle swarm optimization and genetic algorithm performances in deriving the efficient frontier of stocks portfolios based on a mean-lower partial moment model. Int J Financial Econ 26(4):5659–5665

Mäkinen Y, Kanniainen J, Gabbouj M, Iosifidis A (2019) Forecasting jump arrivals in stock prices: new attention-based network architecture using Limit Order Book Data. Quant Finance 19(12):2033–2050

Mamre MO, Sommervoll DE (2022) Coming of age: renovation premiums in housing markets. J Real Estate Finance Econ

Manahov V, Urquhart A (2021) The efficiency of bitcoin: a strongly typed genetic programming approach to Smart Electronic Bitcoin markets. Int Rev Financial Anal 73

Manela A, Moreira A (2017) News implied volatility and disaster concerns. J Financial Econ 123(1):137–162

Markowitz H (1952) Portfolio selection. J Finance 7(1):77–91

McInish TH, Nikolsko-Rzhevska O, Nikolsko-Rzhevskyy A, Panovska I (2019) Fast and slow cancellations and trader behavior. Finance Manag 49(4):973–996

Meng S, Wang H, Shi Y, Gao G (2022) Improving automobile insurance claims frequency prediction with telematics car driving data. ASTIN Bull 52(2):363–391

Modak NM et al (2019) Fifty years of transportation research journals: a bibliometric overview. Transp Res Part A Policy Pract 120:188–223

Mohri M, Rostamizadeh A, Talwalkar A (2012) Foundations of machine learning, 1st edn. The MIT Press, s.l.

Mselmi N, Lahiani A, Hamza T (2017) Financial distress prediction: the case of French small and medium-sized firms. Int Rev Financial Anal 50:67–80

Mulvey JM, Lu N, Sweemer J (2001) Rebalancing strategies for multi-period asset allocation. J Wealth Manag 4(2):51–58

Nguyen HH, Viviani J-L, Ben Jabeur S (2023) Bankruptcy prediction using machine learning and Shapley additive explanations. Rev Quant Finance Acc

Nian K, Coleman TF, Li Y (2021) Learning sequential option hedging models from market data. J Bank Finance 133:106277

Obaid K, Pukthuanthong K (2022) A picture is worth a thousand words: measuring investor sentiment by combining machine learning and photos from news. J Financial Econ 144(1):273–297

Ongsakul V, Chatjuthamard P, Jiraporn P, Chaivisuttangkun S (2021) Corporate integrity and hostile takeover threats: Evidence from Machine Learning and "CEO luck." J Behav Exp Finance 32:100579

Pace RK, Hayunga D (2019) Examining the information content of residuals from hedonic and spatial models using trees and forests. J Real Estate Finance Econ 60(1–2):170–180

Papík M, Papíková L (2022) Detecting accounting fraud in companies reporting under US GAAP through data mining. Int J Acc Inf Syst 45:100559

Payne BC, Tresl J (2014) Hedge fund replication with a genetic algorithm: breeding a usable mousetrap. Quant Finance 15(10):1705–1726

Perols J (2011) Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Auditing J Pract Theory 30(2):19–50

Petridis K, Tampakoudis I, Drogalas G, Kiosses N (2022) A support vector machine model for classification of efficiency: an application to M&A. Res Int Bus Finance 61:101633

Pun CS, Wang L (2021) A cost-effective approach to portfolio construction with range-based risk measures. Quant Finance 21(3):431–447

Pyo S, Lee J (2018) Exploiting the low-risk anomaly using machine learning to enhance the black–litterman framework: evidence from South Korea. Pac Basin Finance J 51:1–12

Qiu Y, Xie T, Yu J, Zhou Q (2020) Forecasting equity index volatility by measuring the linkage among component stocks. J Finance Economet 20:160–186

Ranta M, Ylinen M, Järvenpää M (2022) Machine learning in management accounting research: literature review and pathways for the future. Eur Acc Rev 32(3):607–636

Reed R (1993) Prunning algorithms—a survey. IEEE Trans Neural Netw 4:740–747

Renault T (2017) Intraday online investor sentiment and return patterns in the U.S. stock market. J Bank Finance 84:25–40

Rialti R, Marzi G, Ciappei G, Busso D (2019) Big data and dynamic capabilities: a bibliometric analysis and systematic literature review. Manag Decis 57:2052–2068

Ribeiro MT, Singh S, Guestrin C (2016) Why should I trust you? Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 1135–1144

Rojas-Lamorena ÁJ, Barrio-García S, Alcántara-Pilar J (2022) A review of three decades of academic research on brand equity: a bibliometric approach using co-word analysis and bibliographic coupling. J Bus Res 139:1067–1083

Ruch GW, Taylor G (2015) Accounting conservatism: a review of the literature. J Acc Lit 34(1):17–38

Russell SJ, Norvig P (2020) Artificial intelligence a modern approach, 4th edn. Pearson, Boston

Samitas A, Kampouris E, Kenourgios D (2020) Machine learning as an early warning system to predict financial crisis. Int Rev Financial Anal 71:101507

Samuel A (1959) Some studies in machine learning using the game of checkers. IBM J Res Dev 3:535–554

Sarker IH (2021) Machine Learning: algorithms, real-world applications and research directions. SN Comput Sci 2(3):160

Saurabh S, Dey K (2020) Unraveling the relationship between social moods and the stock market: evidence from the United Kingdom. J Behav Exp Finance 26:100300

Sautner Z, van Lent L, Vilkov G, Zhang R (2023) Firm-Level climate change exposure. J Finance 78(3):1449–1498

Shahzad SJ, Bouri E, Ahmad T, Naeem MA (2022) Extreme tail network analysis of cryptocurrencies and trading strategies. Finance Res Lett 44:102106

Shrestha YR, Krishna V, von Krogh G (2021) Augmenting organizational decision-making with deep learning algorithms: principles, promises, and challenges. J Bus Res 123:588–603

Sirignano JA (2018) Deep learning for limit order books. Quant Finance 19(4):549–570

Sirignano J, Cont R (2019) Universal features of price formation in financial markets: perspectives from deep learning. Quant Finance 19(9):1449–1459

Slapnik U, Lončarski I (2021) On the information content of sovereign credit rating reports: improving the predictability of rating transitions. J Int Financial Mark Inst Money 73:101344

Smith SJ, Urquhart V (2018) Accounting and finance in UK universities: academic labour, shortages and strategies. Br Acc Rev 50(6):588–601

Sun T (2019) Applying deep learning to audit procedures: an illustrative framework. Acc Horiz 33(3):89–109

Sutton S, Holt M, Arnold V (2016) "The reports of my death are greatly exaggerated"—artificial intelligence research in accounting. Int J Acc Inf Syst 22:60–73

Talbot D, Ordonez-Ponce E (2022) Canadian banks' responses to covid-19: a strategic positioning analysis. J Sustain Finance Invest 12(2):423–430

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

Turing A (1950) Computing machinery and intelligence. Mind 236:433–460

Vamossy DF (2021) Investor emotions and earnings announcements. J Behav Exp Finance 30:100474

Van Eck N, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84(2):523–538

Vapnik VN (2000) The nature of statistical learning theory, 2nd edn. Springer Verlag, s.l.

Viswanathan PK, Srinivasan S, Hariharan N (2020) Predicting financial health of banks for investor guidance using machine learning algorithms. J Emerg Mark Finance 19(2):226–261

Wang Z, He T, Ren X, Huynh LD (2024) Robust portfolio strategies based on reference points for personal experience and upward pacesetters. Rev Quant Finance Acc

Wang H, Zhou XY (2020) Continuous-time mean–variance portfolio selection: a reinforcement learning framework. Math Finance 30(4):1273–1308

Wei X et al (2020) An intelligent learning and ENSEMBLING framework for predicting option prices. Emerg Mark Finance Trade 57(15):4237–4260

Weigand A (2019) Machine learning in empirical asset pricing. Fin Mark Portf Mgmt 33(1):93–104

Wiese M, Knobloch R, Korn R, Kretschmer P (2020) Quant Gans: deep generation of financial time series. Quant Finance 20(9):1419–1440

Wu Y-X, Wu Q-B, Zhu J-Q (2019) Improved EEMD-based crude oil price forecasting using LSTM networks. Physica A 516:114–124

Xu D, Zhang X, Feng H (2019) Generalized fuzzy soft sets theory-based novel hybrid ensemble credit scoring model. Int J Finance Econ 24(2):903–921

Xue F, Li X, Zhang T, Hu N (2021) Stock market reactions to the COVID-19 pandemic: the moderating role of corporate big data strategies based on word2vec. Pac Basin Finance J 68:101608

Yang C, Zhang H, Weng F (2024) Effects of COVID-19 vaccination programs on EU carbon price forecasts: evidence from explainable machine learning. Int Rev Financial Anal 91:102953

Yu L et al (2020) A novel dual-weighted fuzzy proximal support vector machine with application to credit risk analysis. Int Rev Financial Anal 71:101577

Zanin L (2020) Combining multiple probability predictions in the presence of class imbalance to discriminate between potential bad and good borrowers in the peer-to-peer lending market. J Behav Exp Finance 25:100272

Zapranis A, Refenes A-PN (1999) Principles of neural model identification, selection and adequacy. Springer-Verlag, s.l.

Zaremba A et al (2021) The quest for multidimensional financial immunity to the COVID-19 pandemic: evidence from international stock markets. J Int Financial Mark Inst Money 71:101284

Zhai J, Cao Y, Liu X (2020) A neural network enhanced volatility component model. Quant Finance 20(5):783–797

Zhang J, Huang W (2021) Option hedging using LSTM-RNN: an empirical analysis. Quant Finance 21(10):1753–1772

Zhang Y, Chu G, Shen D (2021) The role of investor attention in predicting stock prices: the long short-term memory networks perspective. Finance Res Lett 38:101484

Zhang Y, Hu A, Wang J, Zhang Y (2022) Detection of fraud statement based on word vector: evidence from financial companies in China. Finance Res Lett 46:102477