sustainability

MDPI

Article

Mapping the Role of Artificial Intelligence and Machine Learning in Advancing Sustainable Banking

Alina Georgiana Manta [1,*], Claudia Gherțescu [1], Roxana Maria Bădîrcea [1], Liviu Florin Manta [2], Jenica Popescu [1] and Mihail Olaru [1]

[1] Faculty of Economics and Business Administration, University of Craiova, 200585 Craiova, Romania; claudiaghertescu@gmail.com (C.G.); roxana.badircea@edu.ucv.ro (R.M.B.); jenica.popescu@edu.ucv.ro (J.P.); mihai.olaru@noarimgroup.ro (M.O.)

[2] Faculty of Automation, Computers and Electronics, University of Craiova, 200585 Craiova, Romania; florin.manta@edu.ucv.ro

* Correspondence: alina.manta@edu.ucv.ro

Abstract

The convergence of artificial intelligence (AI), machine learning (ML), blockchain, and big data analytics is transforming the governance, sustainability, and resilience of modern banking ecosystems. This study provides a multivariate bibliometric analysis using Principal Component Analysis (PCA) of research indexed in Scopus and Web of Science to explore how decentralized digital infrastructures and AI-driven analytical capabilities contribute to sustainable financial development, transparent governance, and climate-resilient digital societies. Findings indicate a rapid increase in interdisciplinary work integrating Distributed Ledger Technology (DLT) with large-scale data processing, federated learning, privacy-preserving computation, and intelligent automation—tools that can enhance financial inclusion, regulatory integrity, and environmental risk management. Keyword network analyses reveal blockchain's growing role in improving data provenance, security, and trust—key governance dimensions for sustainable and resilient financial systems—while AI/ML and big data analytics dominate research on predictive intelligence, ESG-related risk modeling, customer well-being analytics, and real-time decision support for sustainable finance. Comparative analyses show distinct emphases: Web of Science highlights decentralized architectures, consensus mechanisms, and smart contracts relevant to transparent financial governance, whereas Scopus emphasizes customer-centered analytics, natural language processing, and high-throughput data environments supporting inclusive and equitable financial services. Patterns of global collaboration demonstrate strong internationalization, with Europe, China, and the United States emerging as key hubs in shaping sustainable and digitally resilient banking infrastructures. By mapping intellectual, technological, and collaborative structures, this study clarifies how decentralized intelligence—enabled by the fusion of AI/ML, blockchain, and big data—supports secure, scalable, and sustainability-driven financial ecosystems. The results identify critical research pathways for strengthening financial governance, enhancing climate and social resilience, and advancing digital transformation, which contributes to more inclusive, equitable, and sustainable societies.

Keywords: artificial intelligence; machine learning; digitalization; sustainable banking

Check for updates

Academic Editor: Bin Ji

Received: 27 November 2025
Revised: 29 December 2025
Accepted: 5 January 2026
Published: 7 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The use of artificial intelligence (AI) and machine learning (ML) in the financial sector is continuously expanding and transforming, having a profound impact on industry and society [1,2]. From traditional financial institutions such as investment and retail banks or hedge fund management firms such as JPMorgan Chase to new players in financial technology such as Revolut, thus AI and ML are widely used to optimize operations and improve customer service [3].

For example, JPMorgan Chase has deployed the Contract Intelligence (COiN) platform, an AI-powered system that automates the analysis of legal documents. COiN can process and extract critical data from complex contracts in seconds, saving an estimated 360,000 h of manual labor annually and significantly reducing the risk of human error [4]. In the FinTech sector, Revolut uses machine learning algorithms to detect fraudulent behavior and protect customers from scams. The company has launched an advanced fraud detection feature that uses machine learning to identify suspicious transactions in real time and prevent financial losses [5].

This digital transformation trend is also supported by the rapid expansion of the global AI market in the banking sector. Thus, the AI market size has been estimated at USD 19.90 billion in 2023 and increased to USD 26.23 billion in 2024, and is projected to reach around USD 315.50 billion by 2033, registering a compound annual growth rate (CAGR) of 31.83% from 2024 to 2033 [6]. This expansion is fueled by the accelerated digitalization and modernization of the banking sector, as well as the increasing adoption of advanced technologies by financial institutions [7–10].

Artificial intelligence (AI) and machine learning (ML) have begun to play an increasingly important role in the banking sector, transforming it and significantly improving efficiency and decision-making [11–13]. AI, which refers to the use of advanced technologies to enable systems to learn, process and make autonomous decisions, has been integrated into numerous banking operations, such as the use of machine learning algorithms for fraud detection, the automation of credit-granting processes through intelligent scoring or the deployment of virtual assistants to improve customer interaction [14].

AI-driven technologies enable banks to streamline operations, reduce costs, and enhance risk management strategies. For instance, AI-powered chatbots provide 24/7 customer support, handling routine inquiries and allowing human agents to focus on complex issues. Additionally, AI enhances regulatory compliance by automating the monitoring of transactions to detect money laundering activities, ensuring adherence to financial regulations. In investment banking, AI-powered trading algorithms analyze market trends in real time, executing trades at optimal moments to maximize returns and minimize risks. These applications demonstrate how AI is revolutionizing banking by making operations more secure, efficient, and customer-centric [7,10].

In this context, machine learning (ML), a branch of AI, has become an essential tool, helping banks to analyze large volumes of data, identify patterns and predict financial market fluctuations [15,16]. Thus, ML is actively used for fraud prevention by detecting suspicious transactions in real time based on classification and anomaly recognition algorithms, for assessing customer creditworthiness through predictive models that analyze financial history and payment behavior, and for personalizing financial offers through recommendation systems that use clustering to tailor banking products to individual customer needs [17–20].

Over the decades, the use of artificial intelligence (AI) and machine learning (ML) technologies in the banking sector has undergone a profound transformation. Early initiatives in the 1980s, such as expert systems used for financial advice, laid the foundation for modern digitization, and the introduction of the FICO score in 1989 marked a turning point

in the standardization of credit risk assessment in financial institutions around the world. The evolution of these technologies aligns with the directions highlighted in the recent literature, including the analysis by the Congressional Research Service, which highlights how AI has matured from rule-based models to advanced systems capable of supporting critical processes such as fraud prevention, risk assessment, and financial decision automation [21]. At the same time, modern research, such as the study by Islam et al. [22], highlights the catalytic role of global contexts, for example, the COVID-19 pandemic, in accelerating the adoption of AI and ML in key areas, including financial services, facilitating the transition to integrated digital ecosystems oriented towards the concept of Society 5.0.

Although the use of artificial intelligence (AI) and machine learning (ML) techniques offers substantial opportunities for increasing efficiency and innovation in the banking sector, the literature also highlights the existence of significant risks that require increased attention and appropriate management mechanisms. Durongkadej et al. [23] demonstrate that AI-related incidents can directly affect the performance and reputation of financial institutions, leading to operational volatility and increased exposure to technological risks. Similarly, the analysis carried out by the European Central Bank highlights that the rapid adoption of AI can introduce systemic vulnerabilities, including risks related to opaque models, technological dependence, and potential malfunctions that can threaten macroeconomic financial stability [24]. At the same time, the study conducted by Naveed et al. [25] on large language models indicates a number of specific risks, such as the generation of erroneous content, algorithmic bias, and exposure to cyber attacks, elements that have a direct impact on the security of automated banking processes [25].

First, the collection and use of large amounts of sensitive customer data raises major privacy and security concerns [26]. AI algorithms depend on access to such data to function efficiently, and any security breach could have serious consequences, including financial losses and reputational damage for banking institutions. In this context, compliance with data protection regulations, such as the European Union's General Data Protection Regulation (GDPR), becomes essential [27]. Furthermore, the European Union has initiated the development of specific regulations for AI aimed at setting clear boundaries for the use of these technologies [28].

The aim of this paper is to conduct a detailed bibliometric analysis of Artificial Intelligence (AI) and Machine Learning (ML) concepts in banking. Through this analysis, we aim to identify research trends, key developments and models used in the field. We will also examine existing gaps in previous research and highlight future research directions, proposing new approaches and insights to deepen our understanding of the impact of these technologies on efficiency and innovation in the banking sector. A bibliometric analysis is particularly relevant at this stage, given the rapid expansion and fragmentation of the literature on AI and ML in banking, which makes it necessary to systematically map research streams, identify underexplored areas, and assess how current contributions differ from or extend existing reviews. To this end, we propose a series of research questions that will be addressed throughout this paper:

RQ1: How has the publication of academic articles on the use of artificial intelligence and machine learning in the banking sector evolved according to data from Scopus and Web of Science databases?

RQ2: What are the main emerging research directions on the use of artificial intelligence and machine learning in the banking sector, as identified through keyword network analysis?

RQ3: Who are the authors with the highest scientific contributions in the field of artificial intelligence and machine learning applied to the banking sector, according to publications indexed in Scopus and Web of Science?

RQ4: Which research institutions have had the greatest impact on the development of artificial intelligence and machine learning research in the banking sector?

RQ5: How is research on artificial intelligence and machine learning in the banking sector geographically distributed and which countries have the most intense scientific activity in this field?

RQ6: Which academic journals publish the most influential research on artificial intelligence and machine learning in banking, by number of citations?

For a more comprehensive analysis, a multivariate bibliometric analysis using Principal Component Analysis (PCA) will be conducted. Principal Component Analysis (PCA) is employed in this study as a multivariate bibliometric technique to identify latent thematic structures within the keyword co-occurrence matrix. PCA, as a form of factor analysis, operates on correlations within a single dataset and does not synthesize effect sizes across empirical studies. This will provide an integrated perspective on the impact and emerging trends of artificial intelligence and machine learning in the banking sector.

Therefore, the study makes significant contributions to understanding the research landscape of Artificial Intelligence (AI) and Machine Learning (ML) in banking by integrating factor-analytic and bibliometric methodologies, providing a nuanced analysis that bridges quantitative statistical rigor with networked insights. A key contribution lies in its ability to synthesize a large and fragmented body of literature, identifying dominant themes like predictive analytics, customer interaction, and credit risk management, while also revealing emerging areas such as fintech and blockchain. This dual-layered approach not only validates existing findings but also uncovers latent trends and underexplored niches, offering a richer perspective than either method could achieve independently.

The paper is delimited into six sections: Section 2 reviews the relevant literature; Section 3 explains the methodology and data; Section 4 focuses on interpreting the results; Section 5 discusses the results; and Section 6 concludes with the study's conclusions, practitioners' implications, future research, and limitations.

2. Literature Review

2.1. Artificial Intelligence in Banking

Milana and Ashta [29], Manta et al. [30] and Fares et al. [31] conducted bibliometric and systematic literature reviews on the use of artificial intelligence in finance, emphasizing its role in improving efficiency, reducing risk and driving innovation in financial services. These studies emphasized both the potential positive and negative economic effects of AI on the financial sector, highlighting the need for further research to understand these impacts.

Likewise, authors Doumpos et al. [32] provide a detailed review of recent research applied in the banking industry using operations research and artificial intelligence methods, addressing topics such as bank efficiency, risk assessment, bank performance, mergers and acquisitions, bank regulation, customer surveys and fintech, and propose future research directions based on their findings. Along the same lines, Jena et al. [33] used the literature review to identify emerging themes in financial engineering research, including AI and IoT applications, highlighting the growing influence of AI and ML in this field.

Another study [34] shows that people are more willing to use AI-based banking services when they trust these technologies and find them useful. The study's results show that factors such as age, digital experience, and attitude toward technology influence how customers interact with AI-based banking systems, and these links have been confirmed by specific trust measurement tools. Haddad's [35] study explores the impact of artificial intelligence technologies on accounting information system (AIS) excellence in Jordanian banks, based on a survey of 278 respondents from 13 commercial banks. The results indicated the need to enhance the use of artificial intelligence in banks to increase their

excellence, recommending that banking administrations support Jordanian knowledge systems to help senior management access relevant information for decision-making.

The study [15] showed that the intention to adopt artificial intelligence in banks is significantly influenced by factors such as knowledge about AI, subjective norms, perceived risk, perceived usefulness, attitude towards AI and awareness of AI. The results obtained were valid and applicable in a broader context, confirmed by the significance of the model at a level of $p < 0.01$, indicated by F-statistics.

The results of another study [36] using the feasible generalized least squares and generalized method of moments techniques show that technological innovation based on artificial intelligence has a positive impact on banks' return on assets, and the collaboration between AI and economic growth contributes to improved financial performance. The analysis also found that, in the long run, prolonged exposure to innovation can lead to lower financial performance, and non-performing loans negatively affect financial performance, while regulatory capital and economic growth have a positive effect.

2.2. Machine Learning in Banking

Numerous studies address how machine learning (ML) techniques can be used in the financial sector, particularly for systemic risk analysis, risk assessment, and maintaining financial stability [37–42]. The literature highlights the evolution of these methods and the diversification of their applications in financial supervision, banking risk management, and market dynamics interpretation. Existing studies confirm that ML contributes to improving the analytical tools used by financial institutions, providing more advanced means for detecting vulnerabilities and strengthening the resilience of the financial system.

According to the study by Martin Leo, Sharma, and Maddulety [43], the use of machine learning in banking risk management is analyzed through a review of existing literature, highlighting already implemented applications and identifying areas with under-explored potential.

Kou et al. [42] highlighted the potential of ML methods to improve both the accuracy and efficiency of financial systemic risk analysis. At the same time, they emphasized the need to address data quality and data integration challenges and proposed future directions for research. Likewise, Alessi and Savona [37] noted the superior predictive performance of ML approaches in ensuring financial stability. However, they signaled important challenges related to interpretability and accountability, highlighting the need to develop ML models that are both interpretable and accountable.

The study by Lagasio, Pampurini, Pezzola, and Quaranta [44] explores the use of machine learning algorithms to identify the main determinants of bank failures in the euro area over the period 2018–2020. A notable contribution is the application of a graphical neural network and the use of a balanced dataset through advanced oversampling methods, thus highlighting new insights in bank risk analysis. While, the study by Adamyk et al. [45] analyzes the factors influencing trust in Ukrainian banks using World Values Survey data and machine learning algorithms such as Random Forest and XGBoost. Results show that variables such as age, financial satisfaction, income, and overall trust have a significant impact, supported by visualizations and model accuracy estimates.

The study by Beutel, List, and von Schweinitz [46] compares different early warning models for banking crises and finds that while machine learning methods perform well on existing data, traditional logit models are more accurate when applied to new data, identifying factors such as credit expansion and external imbalances as the main signs of banking crises. Thus, the study by authors Alessa et al. [47] presented a framework that combines machine learning and big data analytics for bank marketing, demonstrating

that using PySpark with ML libraries provides faster time to value compared to other ML algorithms, and is ideal for big data applications due to its distributed computing power.

ML in the banking sector is not only used for fraud detection, but also for the classification and performance evaluation of banks, contributing to early risk identification and improved bank management. For example, the study by Meitei et al. [48] applies machine learning techniques to classify banks in India as "strong" or "weak" based on indicators such as return on equity (ROE). The results show that all five models used (Naïve Bayes, SVM, kNN, random forest and neural networks) predicted accurately, highlighting non-performing assets, inflation and exchange rate as key factors.

2.3. The Synergistic Impact of Artificial Intelligence and Machine Learning on Banking Transformation

Artificial intelligence (AI) and machine learning (ML) have demonstrated significant influence over traditional econometric models, generating considerable research interest. These technologies have been successfully applied in various financial domains, such as algorithmic trading [49], asset and derivatives pricing [50], automation [51], financial modeling [52], fraud detection [53], loan management and insurance underwriting [54], financial risk prevention [55], risk management [56], sentiment analysis [57], and trade regulation [58].

The expanding literature on artificial intelligence (AI) and machine learning (ML) in finance has generated important academic reviews, especially in the context of banks. For example, Omarova [58] investigates research on predictive analytics and text mining in finance, addressing the use of AI and ML for analyzing financial data. De Prado et al. [59] review studies of credit risk and bankruptcy in financial institutions, noting a growing trend in financial research to adopt hybrid models combining traditional techniques with AI and neural networks. West and Bhattacharya [60] provide a detailed review of the literature on financial fraud detection in banks, categorizing it based on the types of fraud and the AI and ML algorithms used to identify them. Königstorfer and Thalmann [61] analyze the benefits and challenges of implementing AI in commercial banks, highlighting their applicability in improving banking services. In addition, Ciampi et al. [62] evaluate studies on the use of AI and ML for implicit prediction in SMEs, advocating the adoption of these technologies in banks to improve financial forecasting.

Thus, Nenad Milojević and Srdjan Redzepagic [63] emphasize that artificial intelligence and machine learning are playing an increasingly important role in the financial sector, having a significant impact on bank risk management, especially after the global financial crisis. Their research suggests that careful and well-prepared implementation of these technologies can significantly improve the management of credit, market, liquidity and operational risks, providing effective solutions to current economic and financial challenges.

Another study by Mytnyk et al. [64] addresses the use of artificial intelligence for bank fraud detection, an increasingly relevant issue in the context of the transition to online operations due to the COVID-19 pandemic. The authors develop machine learning models to identify fraudulent transactions and propose data preprocessing techniques, demonstrating that the logistic regression algorithm gives the best results, with an AUC value of about 0.946.

2.4. Artificial Intelligence and Digital Technologies in Sustainable Banking

Recent literature in the financial field highlights an increasingly clear intersection between artificial intelligence, blockchain, and banking sustainability, particularly with regard to climate risk assessment and management. In most studies, these technologies no longer appear as experimental tools but as elements integrated into the analysis and monitoring processes of banking institutions. The first mention of DLT (Distributed Ledger

Technology) appears in relation to increasing the transparency and verifiability of green financial flows, but the literature places particular emphasis on the role of AI and ML in modeling environmental risks.

Hasan et al. [65] show, using advanced semantic analysis techniques applied to an extensive FinTech-sustainability corpus, that machine learning models are already being used to anticipate climate risks, classify green financing according to ESG criteria, and identify discrepancies in climate reporting data. Their results indicate a substantial gain in the quality of predictions regarding the exposure of bank portfolios to climate factors, which reduces delays in economic capital adjustment. In addition, Kozar and Paduszyńska [66] highlight that ML algorithms, including neural networks and ensemble techniques, contribute to the development of much more stable ESG scores, while blockchain facilitates the traceability of funded projects and the confirmation of the authenticity of sustainability indicators.

From an empirical perspective, Li and Chen [67] analyze a large sample of banks and show that the adoption of AI- and blockchain-based digital tools optimizes green lending performance. AI models allow for a more accurate selection of projects with a positive environmental impact, while digital monitoring reduces the risk of misreporting. The same direction is confirmed by Yuan et al. [68], who show that the integration of smart technologies into banking processes strengthens both the financial performance and the climate sustainability profile of institutions through a more consistent assessment of transition and physical risks.

An important contribution comes from the literature on climate risk. Yao and Yang [69] propose a new index for measuring climate risk, demonstrating econometrically that the development of digital components reduces the volatility of bank exposures to climate shocks. In parallel, the synthesis by Tian et al. [70] highlights the use of ML models (such as LSTM, XGBoost, and convolutional neural networks) in climate risk forecasting, transition scenario simulation, and climate impact estimation on financial assets. These applications are implemented by banks in modern stress testing exercises, showing that AI/ML are becoming central components in financial resilience analysis.

With regard to blockchain, the literature focuses mainly on its role in the transparency of green finance. The study by Christodoulou et al. [71] proposes a distributed ledger-based architecture for monitoring green bonds, using smart contracts to automatically verify climate indicators. This approach minimizes the risk of greenwashing and strengthens the credibility of environmental reporting, an important aspect for banking institutions involved in financing climate-impact projects. A similar perspective is presented in the recent literature [72] on banking sustainability, where blockchain and AI are analyzed as complementary technologies: AI interprets data, and DLT (Distributed Ledger Technology) ensures their integrity and verifiability.

2.5. Identified Research Gaps in the Literature

While the existing literature emphasizes the benefits of *artificial intelligence* in streamlining banking operations, risk management and financial services innovation, there are some important gaps. First, most studies focus on the technological and economic aspects without analyzing in depth the ethical and regulatory implications required for the responsible use of AI in banking. Second, there is a lack of comparative studies across different regions or economies, making it difficult to generalize conclusions. Also, existing studies do not adequately address the impact of AI on employment and customer interaction with financial institutions, which are key to understanding the digital transformation of banks. In addition, the relationship between AI and financial sustainability remains under-explored, requiring further investigation to assess the extent to which emerging technologies can contribute to a greener and more sustainable

banking system. Moreover, current studies rarely explore how AI-based analytical tools can support climate-risk modeling, green credit allocation or ESG scoring, despite these becoming essential pillars of sustainable banking.

Despite significant progress in the application of *machine learning* techniques in the banking sector, the current literature has several gaps. A major challenge is the lack of studies comparing the performance of ML models in different contexts, such as developed versus emerging economies. Also, the interpretability and transparency of ML algorithms remain critical issues that are insufficiently addressed, which may affect their adoption in banking decision-making processes. In addition, most existing research focuses on the application of ML in risk assessment and fraud detection, while areas such as personalization of banking services and asset management optimization are less explored. Another limitation of current studies is the lack of long-term evaluations of the impact of ML implementation on bank performance, which could influence future digitalization strategies. Furthermore, ML applications related to environmental performance prediction, climate stress testing, carbon-intensive asset risk assessment and automated detection of greenwashing are examined only superficially or not at all, despite their growing relevance in regulatory and sustainability frameworks.

The literature on the combined use of *AI and ML* in the banking sector is burgeoning, but some significant gaps remain. First, there is a need for more detailed studies on the integration of these technologies into traditional econometric models, as most current research focuses on one-off implementations without providing a comprehensive view of the long-term impact. Second, most studies analyze AI and ML separately, without examining the synergies between them and how their interaction can optimize banking processes. Also, the impact of AI and ML on financial stability and systemic risks is insufficiently investigated, which limits the ability of regulators to adopt effective policies. Another under-explored area is the integration of AI and ML into banks' sustainability strategies, in particular in relation to green project finance and climate risk assessment. There is also a notable scarcity of studies evaluating how AI and ML architectures can be linked to sustainability taxonomies, climate disclosure standards (e.g., CSRD, TCFD) or transition plan modeling, which limits understanding of their practical role in supporting regulatory compliance and environmental resilience.

These gaps underline the need for future research that explores more deeply the implications of AI and ML in banking, given the complexity and rapidity of digital transformations in this field.

Although interest in the application of AI and ML in the banking sector is growing, the number of dedicated studies remains low, and existing research does not yet provide a solid bibliometric analysis that systematically compares information from the Web of Science and Scopus databases. This absence limits a coherent understanding of the field's evolution, thematic trends, and the main authors and institutions contributing to its development. In addition, the literature does not sufficiently map how AI and ML connect with banking sustainability objectives, such as climate resilience, green finance efficiency, ESG criteria integration, or environmental risk reduction, which remain fragmented and insufficiently analyzed. Therefore, an integrated approach is needed to clarify both the directions of research and the contributions of these technologies to the sustainable transformation of the banking system.

3. Materials and Methods

3.1. Bibliometric Analysis Through the Bibliometrix Lens

In this study, the bibliometric analysis was performed using Bibliometrix within RStudio Version: 2024.12.1+563, an integrated development environment for the R language,

specialized in statistical computing and graphical visualization. RStudio is available as both a desktop application and a server version, allowing browser-based access for remote data processing. Through this environment, Bibliometrix facilitates the import, cleaning and analysis of bibliographic datasets, providing advanced functionalities for thematic mapping, co-citation analysis and identification of collaborative networks [73].

This methodology is a tool for investigating the structure and dynamics of academic domains, facilitating the customization of bibliometric indicators and the generation of relevant visualizations for the interpretation of results [74].

For the bibliometric analysis, we used two of the most important academic databases: Web of Science and Scopus. These platforms were chosen because of their relevance and extensive coverage in the field of scientific research, providing access to a considerable number of articles published in prestigious journals. The dual analysis based on these sources allows a more comprehensive assessment of academic trends and contributions in the field, reducing the risk of missing relevant studies.

The bibliometric analysis process was structured in several successive stages, aiming to ensure methodological rigor and reproducibility of results:

1. Defining the search strategy and collecting data: In the first stage, relevant literature was identified by querying the Web of Science Core Collection and Scopus databases. The search strategy was built based on the expression "artificial intelligence" AND "machine learning" AND "banking," applied in the title, abstract, and keywords fields. This procedure generated an initial set of 687 records from Web of Science and 561 records from Scopus.
2. Filtering by domain and thematic delimitation of the corpus: In order to restrict the analysis to literature relevant from a disciplinary point of view, a domain filter was applied, selecting only publications in the economic and financial area (economics, finance, business, and management). Following this thematic filtering stage, the analysis corpus was reduced to 380 documents in Web of Science and 109 documents in Scopus (Figure 1). No time constraints were applied, as all publications available at the time of data extraction were included.
3. Criteria for inclusion and eligibility of documents: The bibliometric analysis included several types of documents, without imposing additional restrictions on the typology of publications. Thus, articles, reviews, book chapters, conference papers, and editorial materials were included. No duplicate records were identified in the final sets analyzed separately for each database, and all documents resulting from thematic filtering were considered eligible for analysis.
4. Data processing and construction of the bibliometric matrix: The bibliographic data were downloaded in BibTeX format and imported into the RStudio environment using the Bibliometrix package. The dataset was structured as a bibliometric matrix, including information on keyword frequency, co-occurrence relationships, co-citation links between authors and sources, as well as data on institutional and geographical collaborations.
5. Bibliometric analysis and interpretation of results: The analysis aimed to identify thematic and relational structures in the literature by examining keyword co-occurrence networks, author co-citation networks, and collaboration patterns between institutions and countries. The results obtained highlighted the main research directions and the evolution of studies on the application of artificial intelligence and machine learning techniques in the banking sector.

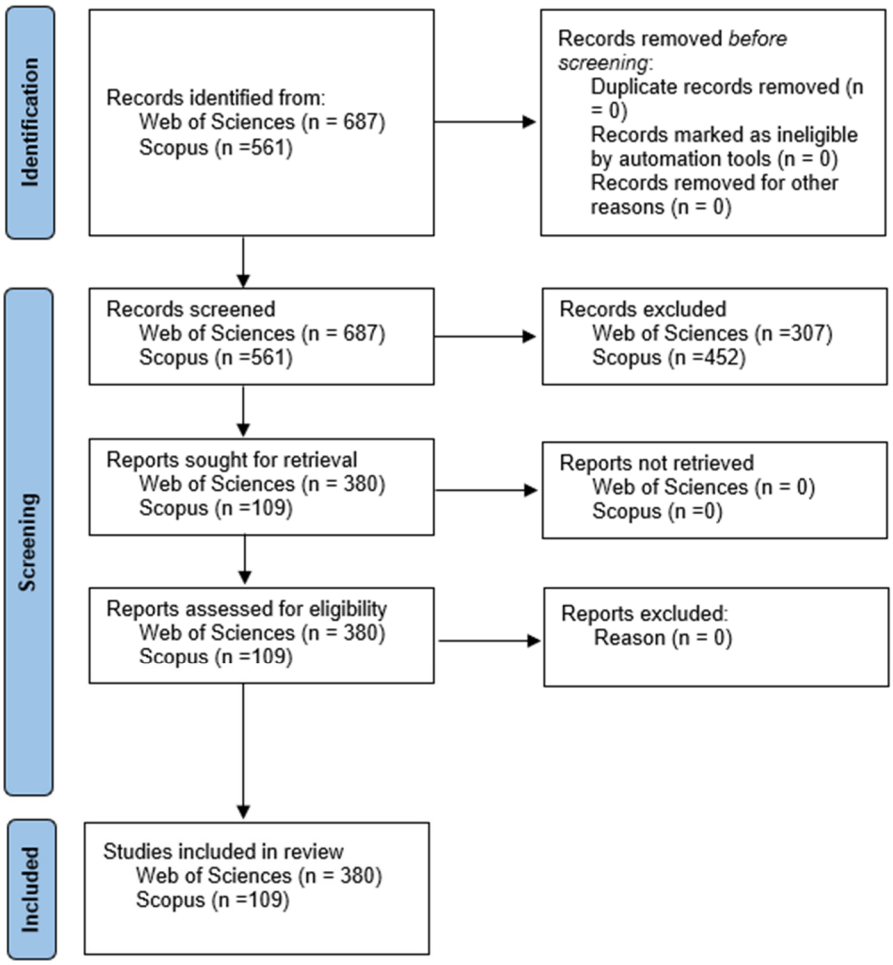


Figure 1. PRISMA flow diagram of the bibliometric search and screening process in Web of Science and Scopus. Source: own processing.

3.2. Complementing the Bibliometric Analysis Through Principal Component Analysis (PCA) Approach

The quantitative procedure applied in this study constitutes a multivariate bibliometric analysis based on Principal Component Analysis (PCA). PCA belongs to the family of factor-analytic techniques and operates by examining the correlation structure among variables within a single dataset, with the goal of reducing dimensionality and identifying latent components. Given that the present study applies PCA to keyword co-occurrence metrics extracted from Scopus and Web of Science, the procedure aligns methodologically with established practices in multivariate bibliometric mapping.

Moreover, factor analysis identifies latent dimensions underlying observed variables by analyzing shared variance. This method is particularly effective in reducing high-dimensional data, allowing researchers to extract dominant themes and relationships [75]. Through factor analysis, AI and ML research can pinpoint core and niche topics, such as "machine learning" and "credit risk,", respectively, within the broader research landscape.

Additionally, factor analysis can group related keywords, such as "classification" and "support vector machines," into a single component, highlighting their shared methodological foundation. Similarly, topics like "deep learning" and "big data" may form another component, reflecting advanced computational techniques in banking research. This method enables researchers to not only identify trends but also assess the relative importance of various themes.

The process typically involves several key steps:

1. Data Aggregation and Preparation: The dataset includes variables derived from bibliometric analyses, such as keyword occurrences, total link strength, and thematic clusters. For instance, keywords like "artificial intelligence" and "machine learning" are treated as variables, with their frequency and connectivity forming the data matrix.
2. Correlation and Covariance Matrix: A strong correlation between "machine learning" and "classification" may indicate their frequent co-occurrence in the same research contexts [76].
3. Principal Component Extraction: Factor analysis method extracts components using eigenvalues and eigenvectors. For example, in the present factor analysis, the first principal component (PC1) accounted for 77.98% of the variance, signifying the dominance of foundational themes like AI and ML.
4. Factor Loadings and Dimensional Reduction: Factor loadings indicate the contribution of each variable to a component. High loadings for "machine learning" and "artificial intelligence" on PC1 confirm their centrality, while lower loadings for "credit risk" or "bankruptcy prediction" on PC2 highlight their niche status.
5. Visualization: Tools such as scree plots, biplots, and scores plots facilitate the interpretation of results. For instance, the scree plot confirms that the first two components sufficiently explain the variance, while biplots visualize the clustering of keywords and their relationships.

This, the integration of bibliometric analysis with factor analysis provides a multidimensional understanding of AI and ML research in banking. Bibliometric methods map co-occurrence networks and collaboration patterns, while factor analysis quantifies the contribution of each theme. For instance, bibliometric analysis highlights the prominence of "artificial intelligence" and "machine learning" in co-occurrence networks, while factor analysis confirms their statistical dominance as primary components [77].

This integration is particularly evident in identifying niche themes. While bibliometric clustering groups keywords like "deep learning" and "credit risk," factor analysis quantifies their contribution to PC2, confirming their secondary but significant role. Together, these methodologies provide a comprehensive picture, balancing networked perspectives with statistical rigor.

4. Results

This section addresses RQ1 by analyzing the evolution of academic articles published to date, aiming to highlight the trends and developmental stages of research on the use of artificial intelligence and machine learning in the banking sector. Analyzing the number of publications registered in the Web of Science (WOS) and Scopus databases during the years 2013–2025, a general trend of significant increase in the number of articles in both WOS and Scopus are observed, indicating a growing interest in the topics analyzed in this study.

In the WOS database (Figure 2), there is a steady increase in the number of publications, with a significant jump since 2020, when the number of published articles reached 28. This is followed by a rapid increase in 2021 and 2022, when publications reached 46 and 56, respectively, reflecting a growing research interest in this sector. In 2023 and 2024, the number of published articles continued to increase, reaching 63 and 79, respectively, suggesting an expansion of the thematic focus and deepening research. This upward trend can be correlated with the rapid development of emerging technologies and their applicability in various economic domains, including the banking sector.

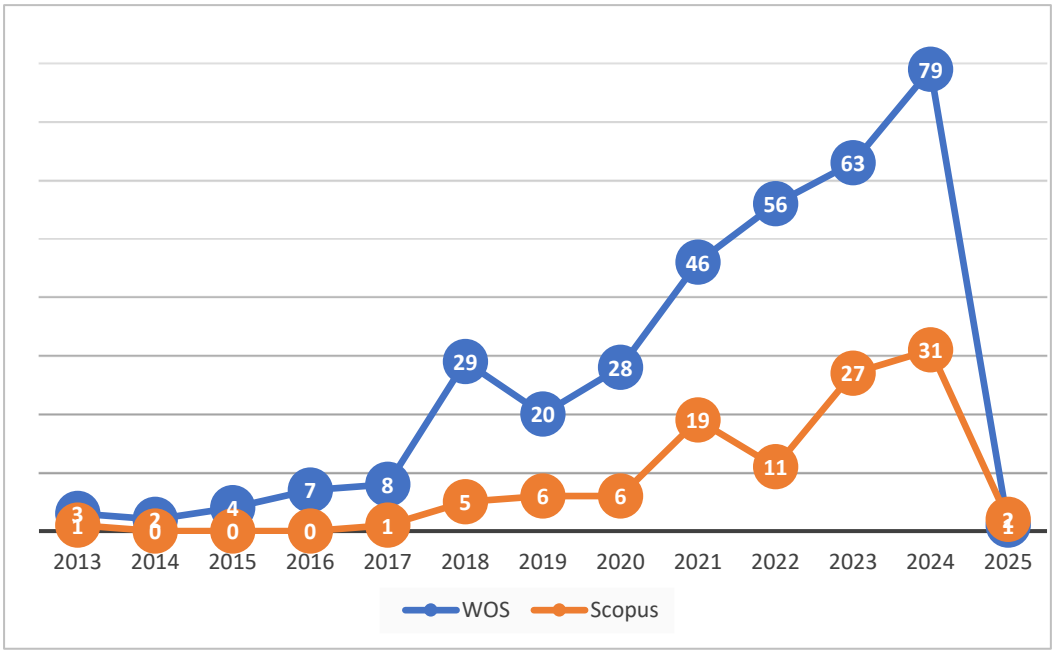


Figure 2. Publications on artificial intelligence and machine learning in banking. Source: Web of Science and Scopus, 2025.

Regarding the Scopus database (Figure 2), the number of publications follows a similar trend, but with lower values compared to WOS, indicating a somewhat narrower coverage of the areas covered. For example, in 2017 and 2018, the number of articles published in Scopus is significantly lower compared to WOS (1 and 5 articles, respectively), but from 2019 onwards, a gradual increase is observed, reaching 6 articles in 2020 and 19 articles in 2021. The increase continues in 2022 and 2023, with 11 and 27 publications, suggesting a wider expansion of the topic in academic research.

This topic is not so widely researched in the early years of the analyzed period, which can be explained by the fact that the integration of artificial intelligence and machine learning in the banking sector is a relatively recent trend, and the applications of these technologies have only started to be explored more intensively in recent years. Thus, interest in the topic has grown significantly as technologies have evolved and started to be deployed more frequently in industry.

4.1. Co-Compete Network Analysis of Keywords

This section addresses RQ2 by examining the emerging trends in research on the use of artificial intelligence and machine learning in the banking sector, as identified through the analysis of keyword networks from relevant scholarly works. The co-occurrence map of the keywords "artificial intelligence" and "machine learning in banking" illustrates the relationships and frequency of connections between relevant terms, highlighting key research trends and emerging areas of interest in the application of these technologies within the banking sector.

The keyword map (Figure 3) generated from the Web of Science database underlines the central importance of the terms "artificial intelligence" and "machine learning" in the context of banking, as they are connected to both practical aspects of the field, such as performance evaluation and credit risk analysis (purple cluster), and advanced predictive methods, such as prediction, models and classification (red cluster). This highlights the use of these technologies for decision automation and process optimization.

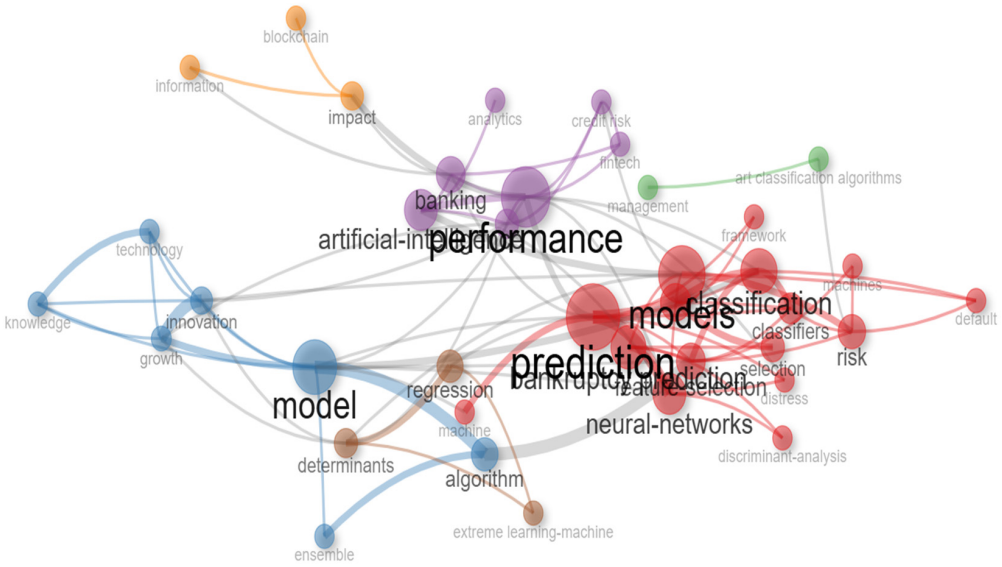


Figure 3. Keyword co-occurrence network in Web of Science database. Source: own processing in Bibliometrix.

The purple cluster indicates a direct application in credit risk management, performance analysis and the use of data-driven analytics in the banking sector. In parallel, the red cluster shows the relevance of predictive methods, where neural networks, classification and extreme machine learning play a key role in fraud detection, customer classification and bankruptcy prediction.

In terms of innovation, the blue cluster, which includes terms such as "innovation", "technology" and "growth", reflects an emerging synergy between AI and technological developments in banking. Likewise, the orange cluster, with a focus on 'blockchain' and 'impact', suggests the integration of blockchain technologies within AI-enabled solutions, supporting security and efficiency.

On the other hand, the keyword map (Figure 4) generated from the Scopus map database is structured around central nodes and related subtopics, reflecting the areas of interest and emerging research directions. Dominant concepts, such as machine learning and artificial intelligence, occupy central positions in the network and are connected with most sub-themes. This centrality emphasizes the importance of these technologies in current research, both as fundamental methods and as solutions applicable in various contexts.

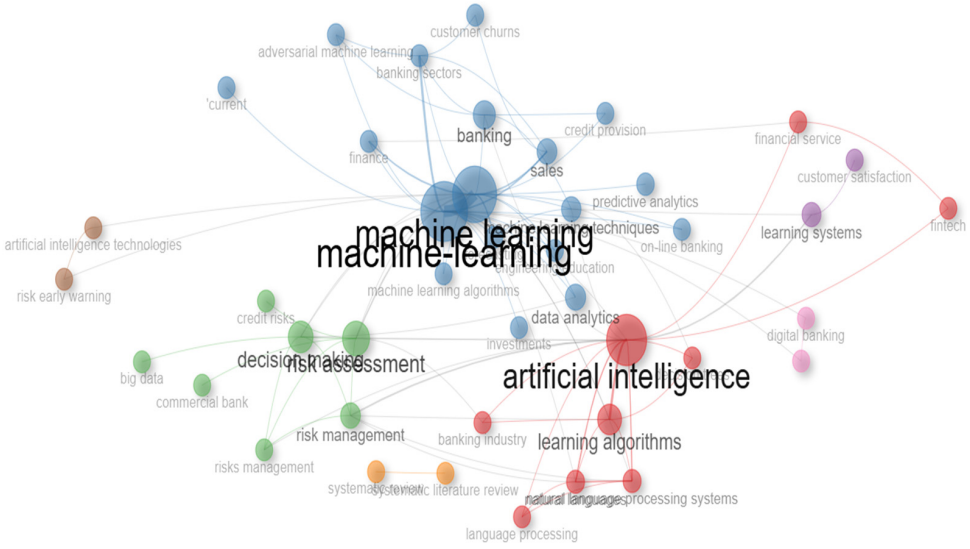


Figure 4. Keyword co-occurrence network in Scopus database. Source: own processing in Bibliometrix.

The network also highlights the existence of distinct thematic clusters, each with a specific role in understanding and applying these concepts. For example, the blue cluster, associated with the term 'banking', connects sub-themes such as 'predictive analytics', 'on-line banking' and 'customer churns', suggesting that machine learning plays a key role in improving banking processes and making predictions about customer behavior. Meanwhile, the red cluster, centered on "artificial intelligence", integrates subtopics such as "learning algorithms", "language processing" and "financial service", indicating the extensive use of artificial intelligence algorithms in financial services and natural language processing.

Another example is the green cluster, which contains concepts such as "risk management", "credit risks" and "decision-making", demonstrating the application of big data and artificial intelligence methods in risk assessment and decision support in the banking sector. In addition, the brown cluster highlights themes such as "artificial intelligence technologies" and "risk early warning", highlighting the use of these technologies in early warning systems to facilitate effective risk management.

In addition to analyzing the thematic clusters, the map also reveals strong links between the core concepts. The relationships between "machine learning" and "banking" are particularly robust, suggesting the intensive use of machine learning algorithms for optimizing banking processes. Similarly, the connection between "artificial intelligence" and "learning systems" reflects continued progress in the development of adaptive intelligent systems capable of responding effectively to complex challenges.

At the same time, the emerging trends highlighted by this network are particularly relevant for research. Sub-themes such as "customer satisfaction" and "digital banking" illustrate a growing interest in improving the customer experience through digital technologies. Also, the presence of the term "systematic literature review" connected with the various sub-themes suggests that researchers are constantly striving to analyze the existing literature in order to identify research gaps and opportunities.

And for a detailed exploration of the relationships between keywords, the analysis is based on the use of factor maps, which allow the identification of thematic clusters and significant connections. These maps provide an in-depth insight into emerging research trends and areas, facilitating an understanding of the distribution and interdependence of concepts.

The bibliometric factor analysis performed with keywords from the Web of Science database reveals the relational structure between relevant concepts, using semantic similarities to highlight the main research directions in the field of artificial intelligence and machine learning in banking. The two main dimensions, Dimension 1 (20.49%) and Dimension 2 (17.85%), together explain approximately 38.34% of the total variance in the data, providing a solid basis for thematic interpretation.

The clustering of the map (Figure 5) terms suggests the existence of four main regions, each with relevant thematic specificities. In the top left region, we identify terms such as knowledge, growth and innovation, which reflect theoretical and conceptual concerns, suggesting studies of technological progress and innovation in the financial industry. This region is strongly connected to the bottom left area, which includes terms such as ensemble algorithms, feature selection and framework. This connection indicates a focus on the development and refinement of technical machine learning methodologies to support progress in this area.

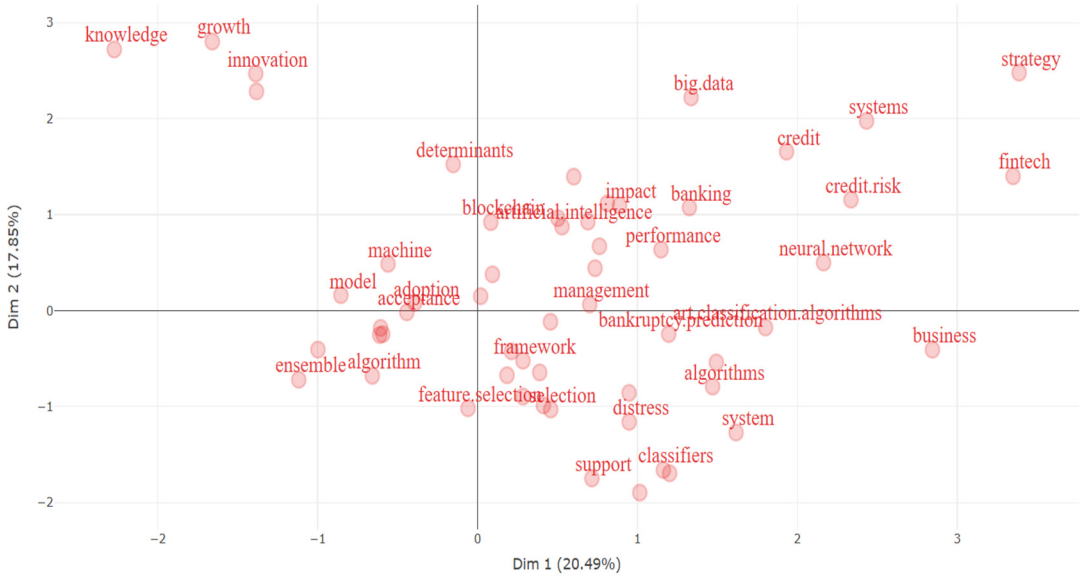


Figure 5. Factorial analysis of keywords from the Web of Science Database. Source: own processing in Bibliometrix.

In parallel, the top right area includes terms such as big data, systems, credit risk, fintech and strategy, which emphasize the applicability of new technologies in financial data analysis, risk management and the definition of banking strategies. This cluster clearly shows how machine learning and artificial intelligence technologies are being used to solve operational and risk analytics challenges in banking, reinforcing the links between technological and application aspects.

At the same time, the bottom right cluster highlights direct applications of machine learning algorithms in finance. Terms such as bankruptcy prediction, neural networks and classification algorithms suggest a focus on the development of advanced predictive models used to detect critical situations such as bankruptcy or risk classification of customers and banking operations. This region confirms the financial industry's orientation towards automation and efficiency through the use of predictive technologies.

A particularly interesting element is the relatively central positioning of the terms artificial intelligence and machine learning, which emphasizes them as connecting points between the different thematic clusters. This position emphasizes the cross-cutting nature of these technologies, which integrate conceptual and methodological research as well as practical applications. Their centrality underlines the fact that AI and machine learning represent a common basis for all the themes explored in this review.

Beyond these observations, the relationships identified in the map highlight some important trends. For example, terms associated with theoretical underpinnings, such as knowledge and growth, indicate concerns about the accumulation and capitalization of knowledge, while the presence of terms such as blockchain and impact reflect interest in emerging technologies that, together with artificial intelligence, could contribute to innovations in the security and efficiency of banking operations. Also, the connections between credit risk, performance and analytics suggest that financial risk and performance measurement remain central research themes, highlighting the practical value of predictive technologies.

The map in Figure 6, generated from the Scopus database shows two dimensions. Which explain a cumulative 22.38% of the total variance, with 14.62% attributed to Dimension 1 and 7.76% to Dimension 2, suggesting a relatively broad distribution of the analyzed themes. Dimension 1 seems to reflect an axis of emerging technologies and processes used in banking and artificial intelligence. Themes associated with this dimension, such

as "digital banking", "predictive analytics" and "blockchain", highlight the innovative application of these technologies in the financial industry. In contrast, Dimension 2 can be interpreted as an operational and behavioral impact axis, including concepts such as "credit risks", "behavioral research" and "accuracy assessment", which highlight decision-making and performance evaluation aspects of AI-based applications.

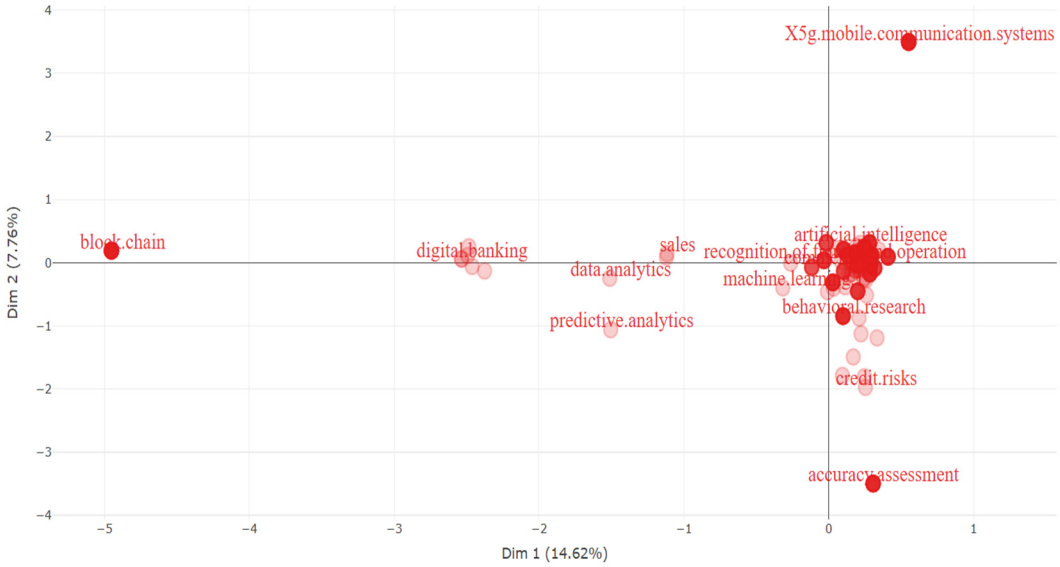


Figure 6. Factorial analysis of keywords from the Scopus. Source: own processing in Bibliometrix.

The thematic distribution is characterized by a number of significant observations. Some concepts, such as "blockchain" and "5G mobile communication systems", are located at the periphery of the diagram, indicating a relatively isolated positioning with respect to the central themes. This suggests that these emerging topics are not yet fully integrated into the prevailing body of literature. In contrast, terms such as 'artificial intelligence', 'machine learning' and 'credit risks' are concentrated in the core area, reflecting a strong link between these themes and their relevance in existing research. This central concentration emphasizes the importance of collaboration between advanced technologies and financial practices in the current literature.

The thematic interactions provide a deeper understanding of how these concepts are interrelated. The proximity between the terms 'predictive analytics' and 'digital banking' points to the application of predictive analytics for process optimization in digital banking, suggesting a synergy between these fields. Also, the connection between "credit risks" and "behavioral research" highlights the growing interest in the use of artificial intelligence to analyze customer behavior and manage associated risks.

The emerging trends identified in the map indicate new directions for research and development. The emergence of the concept of "5G mobile communication systems" reveals a growing interest in the integration of advanced mobile communications in financial services, offering new opportunities for the digitization of the sector. Similarly, blockchain technology is highlighted as a distinct topic, suggesting high potential in revolutionizing banking processes. These trends underline the dynamism and rapid evolution of research in this area, highlighting the opportunities offered by new technologies in reshaping the financial sector.

4.2. Authors' Co-Citation Network

This section addresses RQ3 by identifying the authors with the highest number of published works in the field of artificial intelligence and machine learning applied to the banking sector, based on data from relevant academic sources. The analysis of authors

from the Web of Science database highlights the key contributors in the field of artificial intelligence and machine learning applied to banking, emphasizing their collaboration networks and academic impact.

Figure 7 highlights the relevance of authors according to the number of papers published, providing a clear perspective on their contribution to the field. Zhang Y stands out as the author with the most publications, having a total of four papers, while Cheng D, Krabichler T, Li J, Sharma M and Wang X each have three papers published. Also, authors such as Balland PA, Basalkevych O, Beheshti A and Benatallah B contributed two papers each, showing a relatively even distribution of publications among the relevant authors. This distribution suggests a balanced collaboration in the field without an excessive concentration of publications around a small number of researchers.

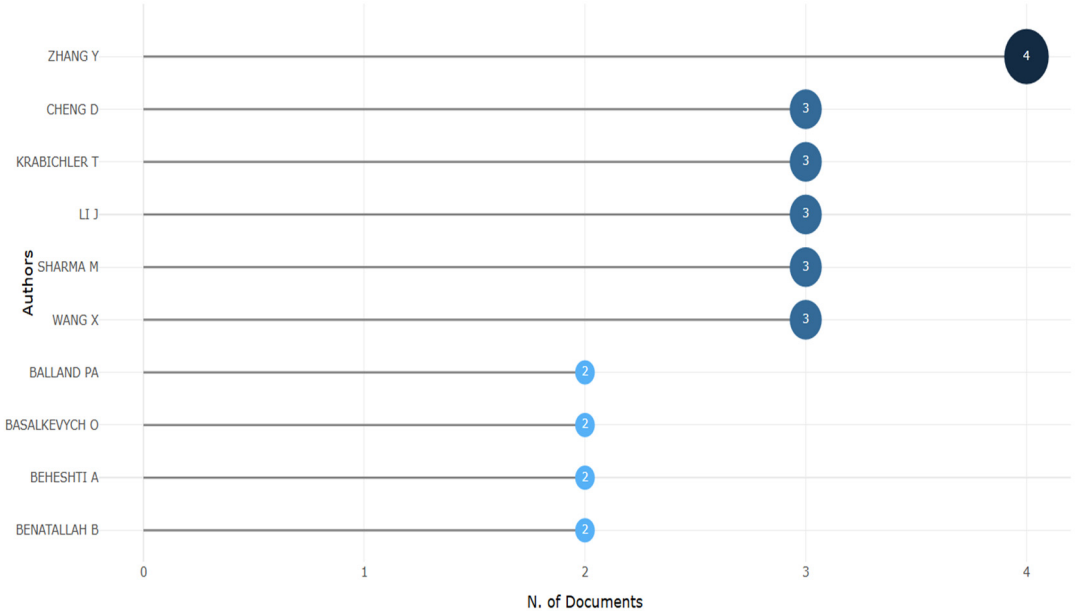


Figure 7. Most relevant authors in Web of Science database. Source: own processing in Bibliometrix.

And Figure 8, is a complement to Figure 7, which reflects the output of the authors over time, highlighting their significant contributions in the field of artificial intelligence (AI) and machine learning (ML) applied to the banking sector. The authors' contributions are distributed over a relevant period, highlighting emerging trends and their role in advancing knowledge in the field.

Among the authors with recurring contributions are Zhang Y, Cheng D, Krabichler T, Li J, and Sharma M, whose papers are distributed consistently between 2020 and 2024. This continuity suggests sustained commitment to the topic and active involvement in applied AI/ML research in banking. Wang X also stands out with a significant volume of publications in a short time span, indicating an intensification of research activity and a possible focus on innovative aspects.

In contrast, authors such as Basalkevych O, Beheshti A and Benatallah B had punctuated contributions concentrated in specific periods. This pattern may reflect temporary involvement or research focused on limited projects, but with potentially significant impact.

The temporal evolution of academic output shows a sharp increase after 2020, coinciding with the acceleration of digitization and the increasing demand for advanced technologies in the banking sector. This trend highlights the growing relevance of AI/ML in optimizing banking processes, analyzing risks and improving customer service.

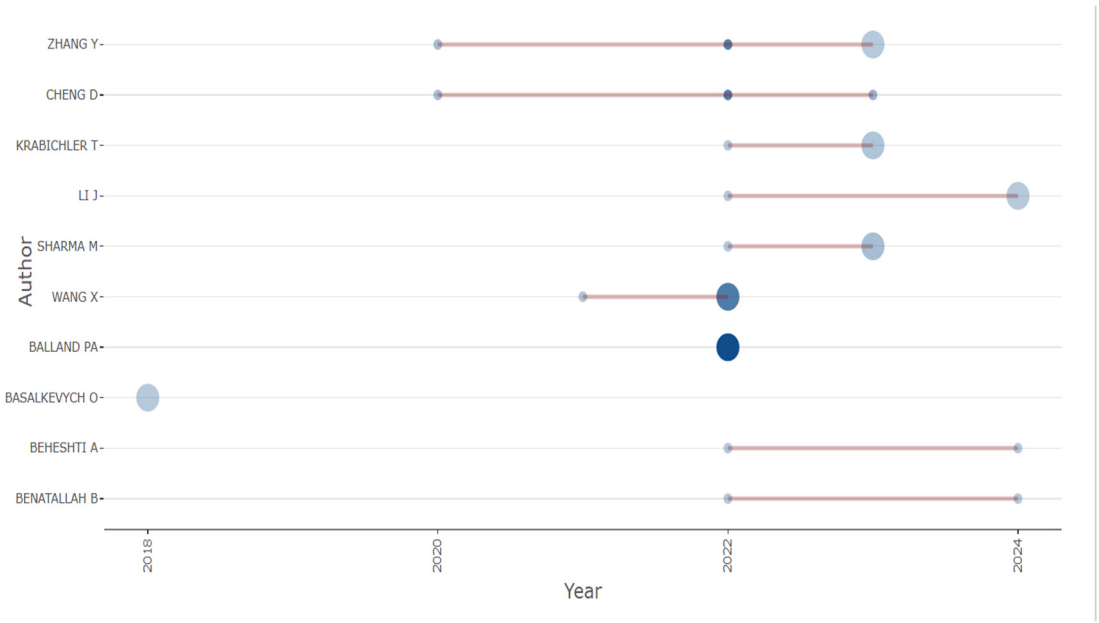


Figure 8. Authors' rate of production over time in Web of Science database. Source: own processing in Bibliometrix.

According to Figure 9, made with keywords from the Scopus database, each author included in the graph is associated with two papers, the maximum value identified, which emphasizes a relatively equal contribution among the researchers analyzed. The most prominent authors include Ceron BM, Chen K, Gupta S, Gupta S, Irfan M, Kumar P, Mehrotra A, Monge M, Taneja S and Abd El-Aal MF, each of whom have an equal number of papers attributed.

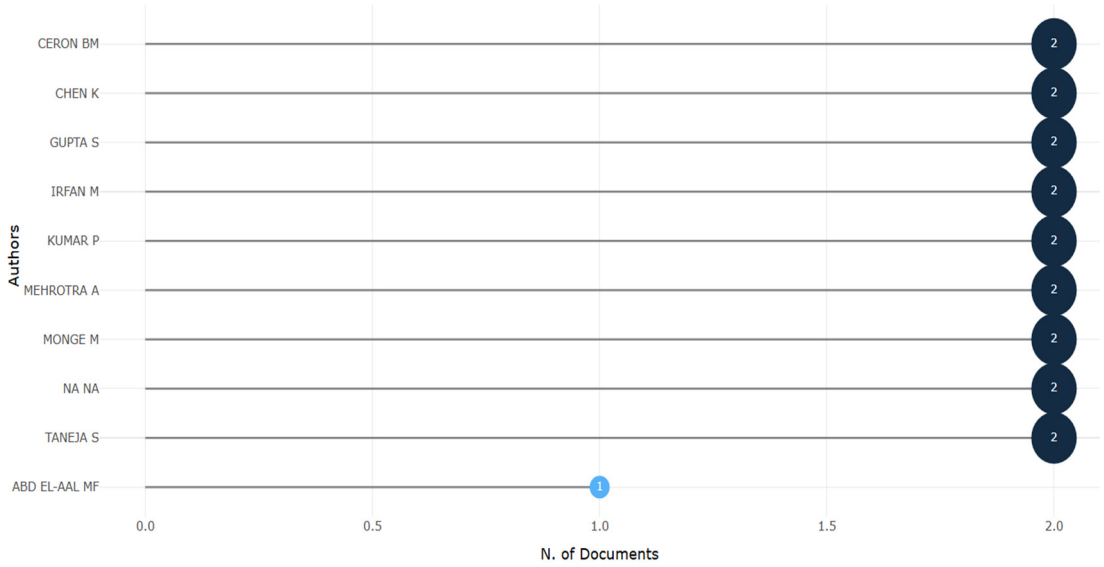


Figure 9. Most relevant authors in Scopus database. Source: own processing in Bibliometrix.

The even distribution of publications across authors, with no clear leader dominating the field, suggests a collaborative nature of research in this thematic space. The lack of a marked concentration of scholarly output in a small number of researchers may reflect the diversity of contributions and approaches in the field analyzed. At the same time, the graph also reveals the presence of one author associated with a single paper, identified as NA NA, which may indicate incomplete or anonymous data.

Figure 10 reflects authors' production over time, in Scopus database, gives an insight into the evolution of the scientific output of different authors over the years, reflecting the diversity of contributions and the dynamics of academic research in this field. The analysis highlights the constant involvement of authors such as Mehrotra A. and Kumar P., who demonstrate a sustained and consistent activity over several years.

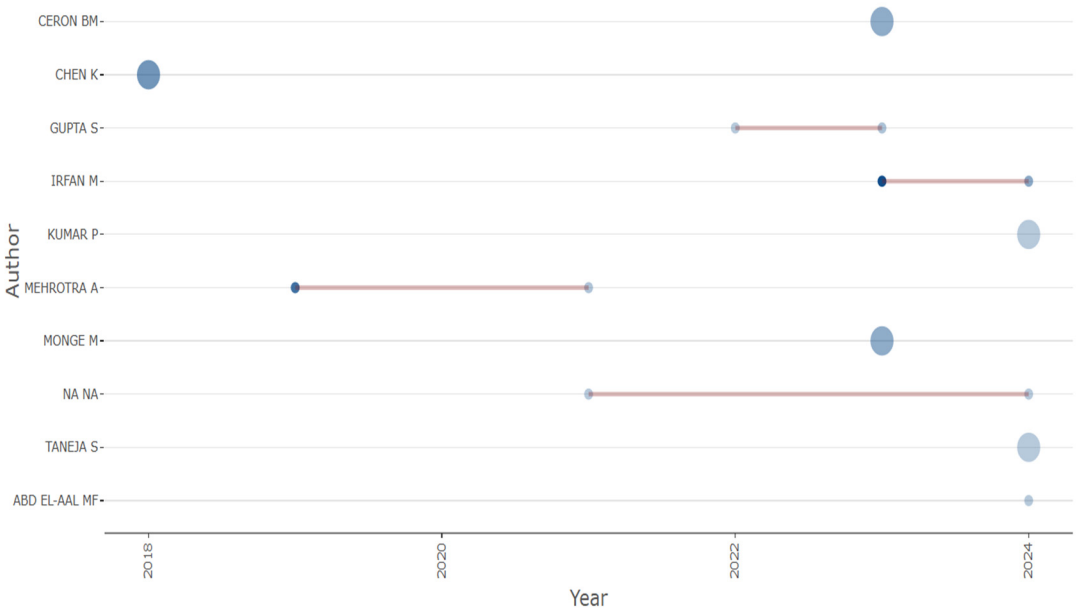


Figure 10. Authors' rate of production over time in Scopus database. Source: own processing in Bibliometrix.

This continuity emphasizes their long-term commitment and their essential contribution in advancing the field. In contrast, other authors, such as Chen K., stand out with productions that are limited in time but significant in volume, indicating a concentrated academic impact in the years of activity.

On the other hand, temporal analysis highlights variations in scholarly output. Some publications are notably concentrated in the recent range 2023–2024, suggesting an increased interest in emerging themes and topical issues. Authors such as Irfan M. and Gupta S. are distinguished by an intense pace of publications in short periods, suggesting either the conduct of a specific research project or a sharp increase in scholarly engagement within a delimited interval. This trend highlights particularities in the individual work of the authors, but also the influence of contextual factors on the pace of research.

Temporal diversity becomes evident in the case of authors with recent contributions, such as Taneja S. and Abd El-Aal MF. who seem to focus on contemporary or emerging research directions. Their involvement reflects the dynamism of the field of study and the tendency towards innovation in the context of current themes. At the same time, the presence of an anonymous author in the data, identified by the 'NA NA' marker, suggests possible inaccuracies or errors.

4.3. Collaborative Institutional Analysis of Co-Authors

This section explores RQ4 by examining the primary research institutions that have made significant contributions to studies on artificial intelligence and machine learning in the banking sector, as highlighted in the academic sources analyzed. The collaborative institutional analysis of co-authors reveals the key partnerships between institutions, highlighting the geographical and academic networks that drive research in artificial intelligence and machine learning within the banking sector.

Figure 11 provides a detailed perspective on the evolution of key institutions' contributions to scientific research over the period 2015–2025, highlighting their dynamics and progress in terms of the number of published articles in Web of Science database. The analysis of this graph highlights the general trends, the performances of the top institutions, the regional contributions and the thematic alignment between them.

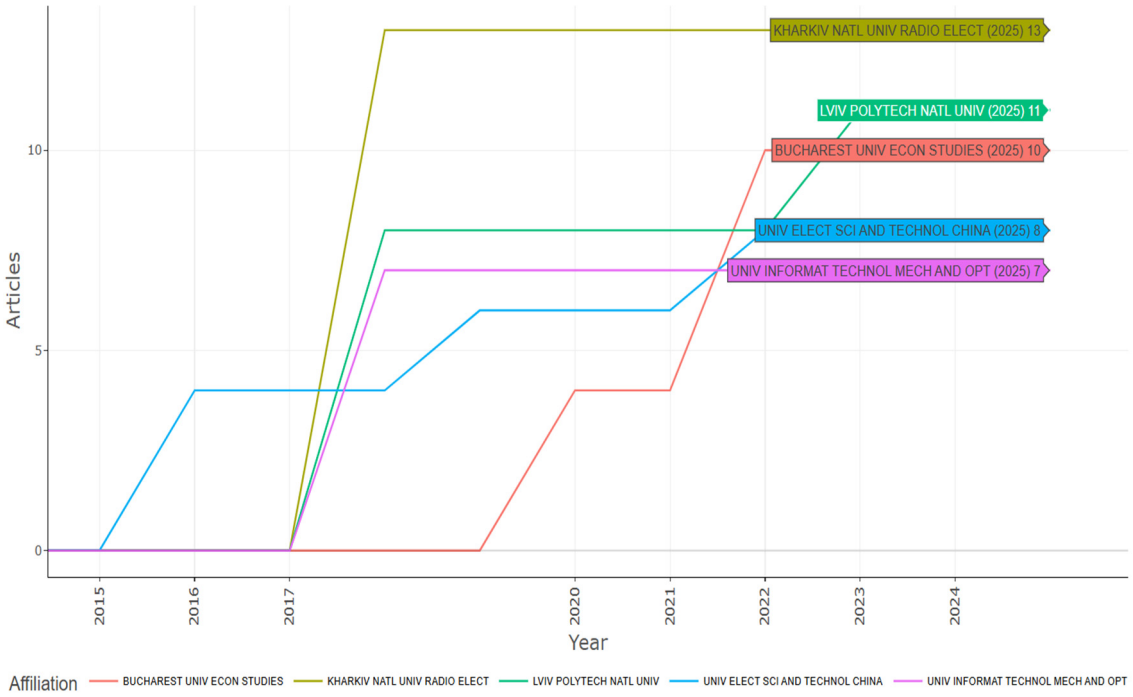


Figure 11. Affiliations' rate of production over time in Web of Science database. Source: own processing in Bibliometrix.

First, a general upward trend is observed, characterized by a sustained increase in the number of articles published by all the institutions analyzed, especially since 2017. This intensification reflects an increase in academic interest in areas such as artificial intelligence, machine learning and their financial applications. The period 2017–2025 thus marks a critical stage of research expansion, being characterized by a thematic diversification and a significant increase in the number of publications.

Among the top institutions, Kharkiv National University of Radio Electronics stands out as the leader, with a total of 13 articles published by 2025. The accelerated growth of its scientific output after 2017 highlights both a major involvement in research and a position of academic leadership in the field. Similarly, Lviv National Polytechnic University, with 11 articles in 2025, is on an upward trajectory, marking a solid contribution to the advancement of scientific knowledge. Bucharest University of Economic Studies also records a significant leap with 2020, reaching a total of 10 articles in 2025, indicating an intensified focus on topics related to digitalization and the economic impact of new technologies.

In addition, institutions such as the University of Electronic Science and Technology of China and the University of Information Technology, Mechanics, and Optics are notable for their steady growth, publishing 8 and 7 articles, respectively, by 2025. This trend highlights the active involvement of Asian and European research centers in the development and application of emerging technologies. At the same time, the synchronized growth of contributions from these institutions indicates a possibility of indirect collaboration or thematic alignment, which reflects a common interest in digital technologies and their impact on the global economy.

The institutional analysis from the Scopus database identifies the leading organizations contributing to research on artificial intelligence and machine learning in banking, showcasing their influence and collaboration networks within the field.

One of the notable trends is the clear rise of Amity University (Figure 12), which, compared to 2019, has demonstrated a steady and rapid increase in scientific production. This institution has come to dominate the academic landscape of the analyzed field, reaching a maximum of six articles published in 2024.

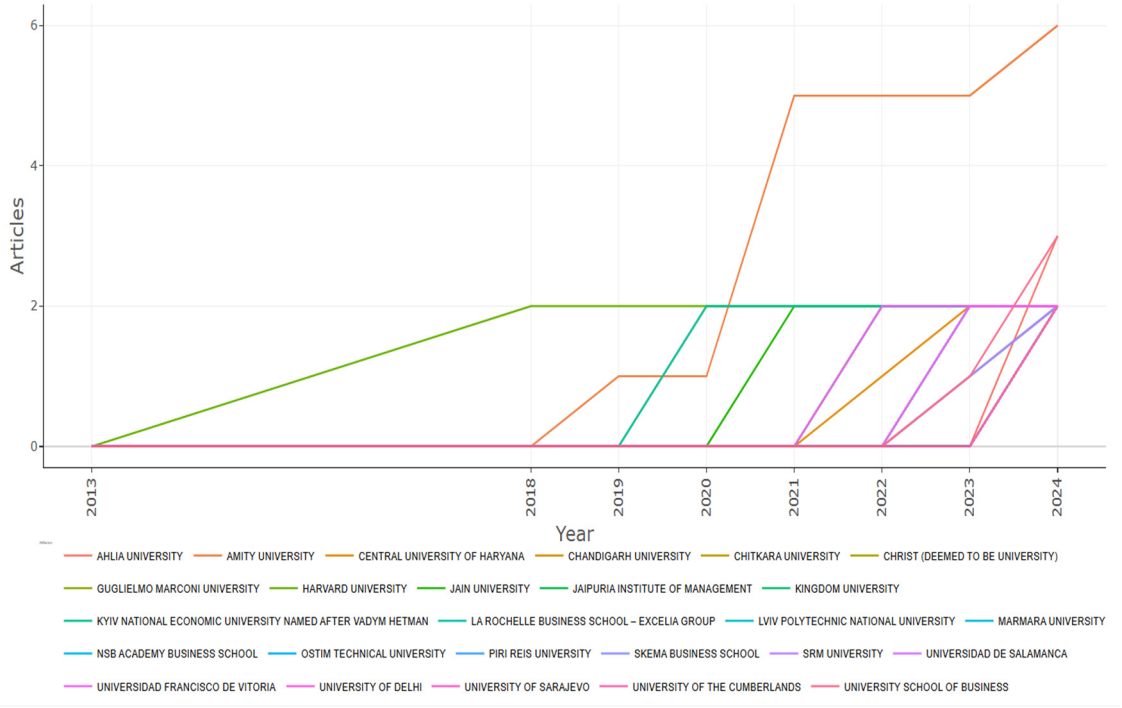


Figure 12. Affiliations' rate of production over time in Scopus database. Source: own processing in Bibliometrix.

On the other hand, Ahlia University stands out with a slower but steady growth, which suggests a stable and relevant presence in the academic landscape. This evolution is also supported by other institutions such as Guglielmo Marconi University and the University School of Business, which have continued to contribute with a moderate number of publications. These increases highlight a diversified expansion in which several institutions find an important role in the development of the analyzed field.

At the same time, the analysis reveals significant differences in the pace of contribution between institutions. Harvard University, University and University School of Business Presentations intermittent, albeit modest, contributions, which may reflect either a focus on specific research niches or the prioritization of other academic fields. This temporal variability adds a level of complexity to the academic landscape, suggesting that methodological and thematic diversity is a characteristic element of this analysis.

Also of interest is the diversified international affiliation. Institutions such as the University of Delhi and the Universidad Francisco de Vitoria demonstrate an increasing commitment to research in the field studied, indicating increasing internationalization and cross-border collaboration in the production of scientific works. This diversity highlights the global impact of the research topic and its ability to attract the interest of a wide range of institutions.

An important peculiarity is the significant increase in contributions in the period 2020–2024, which coincides with the global pandemic context. These recent dynamics

suggest an intensification of interest in the research topics analyzed, probably driven by the transformations and challenges brought to this activity. Institutions have found opportunities for academic exploration within new paradigms, and this has generated a visible increase in scientific production. Thus, it analyses a complex picture of the evolution and distribution of academic contributions, highlighting both divergences and convergences in the global academic landscape.

4.4. Country-Level Research Analysis and Collaboration

This section addresses RQ5 by analyzing the geographical distribution of research contributions and identifying the most active countries in advancing the field of artificial intelligence and machine learning within the banking sector, as revealed by the bibliometric data.

Figure 13, the presentation of the world map of collaboration countries from the Web of Science database. Thus, it provides a comprehensive representation of international networks of research collaboration, highlighting the connections between different countries and underlining the dynamics of global academic partnerships. This visualization suggests the existence of well-established networks, reflecting knowledge exchanges and collaborations between institutions in different regions of the world.

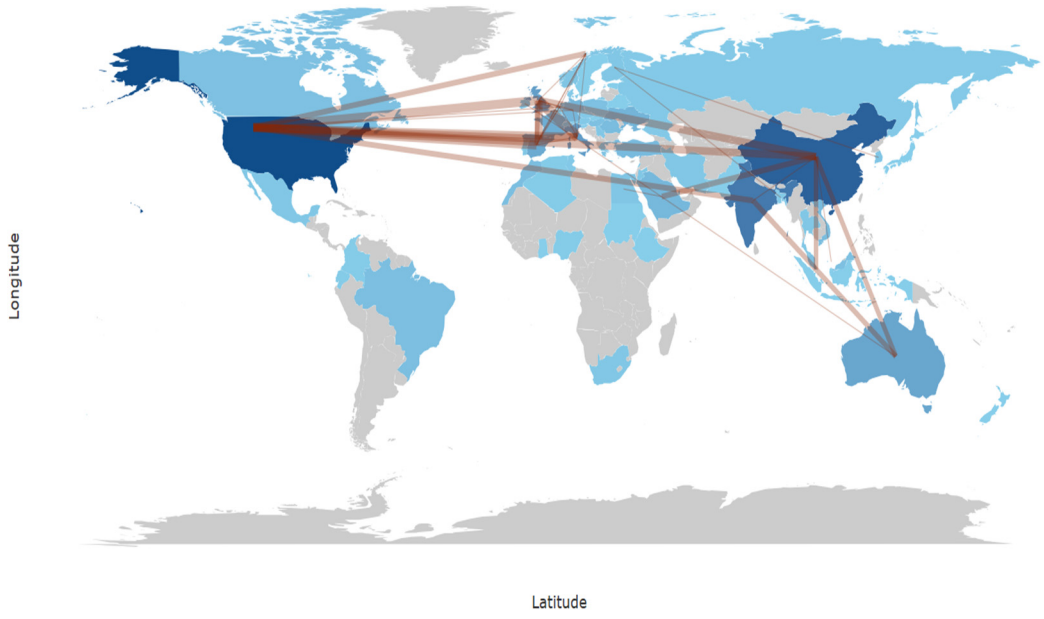


Figure 13. World map of countries' collaboration (Web of Science database). Source: own processing in Bibliometrix.

First, global collaboration centers are dominated by countries such as the United States, China, and Western European countries. They act as central nodes in the global network, highlighting their dominant position in scientific production and transnational research initiatives. Also, the Asia-Pacific regions, including Australia, play a key role by connecting institutions in the southern hemisphere with those in other parts of the world, underlining the integrative nature of scientific research.

On the other hand, the dynamics of connections show that the strongest collaborative relationships are represented by thick links between the United States and countries in Europe and Asia. These well-established partnerships reflect the frequent exchange of knowledge and resources between the most advanced research centers. At the same time, there are signs of emerging collaborations between Central and Eastern European countries, such as Romania, and research centers in Asia and North America. These new connections

highlight the diversification of academic networks and the increasing contribution of developing countries to global research efforts.

Regional and global interactions also highlight an intensification of collaborations between European countries, especially within the European Union. This reflects regional integration initiatives, which facilitate the development of joint research and the exchange of good practices. Similarly, China has an increasingly visible presence in the international collaboration landscape, with extensive links to Europe, North America and Australia. This underlines its rise as a key player in global research, but also its involvement in addressing global challenges.

The global dimension of research, as highlighted by the map, shows that it is a joint effort of the international community. Frequent transcontinental interactions highlight the fact that global issues such as digitalization, climate change or sustainability require cooperation between countries. These networks facilitate the exchange of knowledge and technologies, contributing to the development of integrated solutions to contemporary challenges. In the future, strengthening international partnerships and extending them to less represented regions such as South America and Africa could further diversify the landscape of scientific collaboration. The involvement of these new actors in global networks could bring new and innovative perspectives, contributing to the development of more equitable and sustainable solutions. Thus, the map reflects the importance of continued international cooperation, which supports scientific progress and responds to global needs.

Figure 14 illustrates the distribution of publications by country of corresponding authors in the Web of Science database, specifically analyzing collaboration at national (SCP—Single Country Publication) and international (MCP—Multiple Country Publication) levels. The distribution highlights both the general trends of global scientific production and the collaboration strategies preferred by different countries.

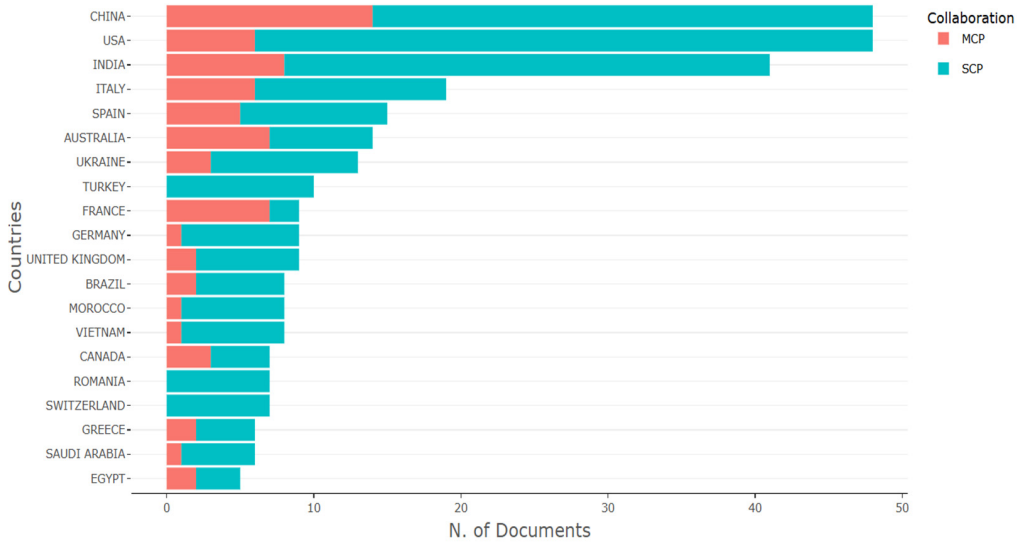


Figure 14. Corresponding authors' countries in Web of Science database. Source: own processing in Bibliometrix.

First of all, countries such as China, the USA and India stand out for their very high number of publications, which indicates a high level of involvement in scientific research. This strong presence demonstrates the capacity of these states to produce new knowledge and to contribute significantly to their fields of interest. Moreover, publications from these countries are mainly of the SCP (Single Country Publication) type, reflecting a majority

focus on domestic research. This approach can be attributed to the significant resources available at national level, which allow for autonomous research.

Secondly, the distribution of MCP (Multiple Country Publication) publications, which highlight international collaborations, is more visible in the case of European countries such as France, Germany and Switzerland. This underlines a strong orientation towards global cooperation networks and active involvement in transnational projects. Europe clearly tends to promote international collaborations more than other regions, a phenomenon that can be explained by the common research policies promoted by the European Union and by the open working cultures in academic fields.

On the other hand, countries such as Romania, Morocco, Vietnam or Egypt contribute with a relatively low number of publications, which can be attributed to more limited resources or research priorities that focus on local needs. In these countries, publications are predominantly SCP type, suggesting a lower participation in global initiatives, but an internal development adapted to the specific problems of each country.

A notable aspect of the analysis is the leading role of the USA and China, not only in volume, but also in involvement in international collaborations. Although SCP is dominant for both nations, they present a significant number of MCP publications, reflecting their influence in global research networks. This balance between domestic research and external collaborations highlights diverse strategies to capitalize on resources and strengthen their position on the global scientific scene.

In terms of geographical distribution, it is evident that Asian regions, such as China, India or Vietnam, favor national research, which reflects either a need to develop their own solutions or a reduced dependence on international collaborations. In contrast, European countries present a different profile, focused more on cross-border collaborations, a phenomenon that indicates a high degree of integration into the global scientific community.

The geographical distribution of international collaborations in the Scopus database also shows networks of academic partnerships between different countries. The intensity of collaboration is represented by the color blue, where darker shades indicate a more significant contribution to the analyzed academic literature, and the lines drawn between countries suggest direct links, indicating the flows of collaboration. Among the main observations, the significant contributions from India stand out, which (Figure 15), due to the intensity of the dark blue, suggests a central role of this country in the production of scientific literature related to the analyzed topics.

The map also highlights active global collaboration networks, especially between India and countries in North America, such as the United States, in Europe, including the United Kingdom and other European states, as well as in Southeast Asia, which emphasizes the transcontinental involvement in research and the globalization of the studied topics. In parallel, countries marked in lighter shades of blue, such as those in Africa, Oceania and South America, reflect a lower level of collaboration, but their inclusion in the map highlights global participation, even at a lower intensity. Emerging trends suggest a concentration of collaborations around major hubs, such as India, the United States and Europe, indicating that these regions are the main drivers of academic research in the selected field.

Figure 16 provides the distribution of published papers, classified by country of corresponding authors in the Scopus database. Thus, India stands out significantly for the high number of papers, around 15, most of which are produced in domestic collaboration (SCP), indicating a high capacity for national academic production. Similarly, China and the United Kingdom stand out for a balanced mix of MCP and SCP, highlighting both active international collaboration and their own research capacity. The United States and Australia also contribute a significant number of papers, most of them in the MCP framework, reflecting an active

involvement in international collaboration networks. On the other hand, countries such as Bahrain, Spain and Turkey, with a lower number of publications, present a balance between SCP and MCP, indicating a diversity in collaboration styles. In terms of small but diverse contributions, countries such as Greece, Brazil, Egypt, Italy and Hungary have a more modest presence but are actively participating in international and national collaborations, suggesting a potential for development in the field under review. Overall, the general trends show a preponderance of SCP in many countries, suggesting a local focus of research; however, countries that combine SCP and MCP, such as China and the United Kingdom, highlight the importance of international collaborations for advancing the field.

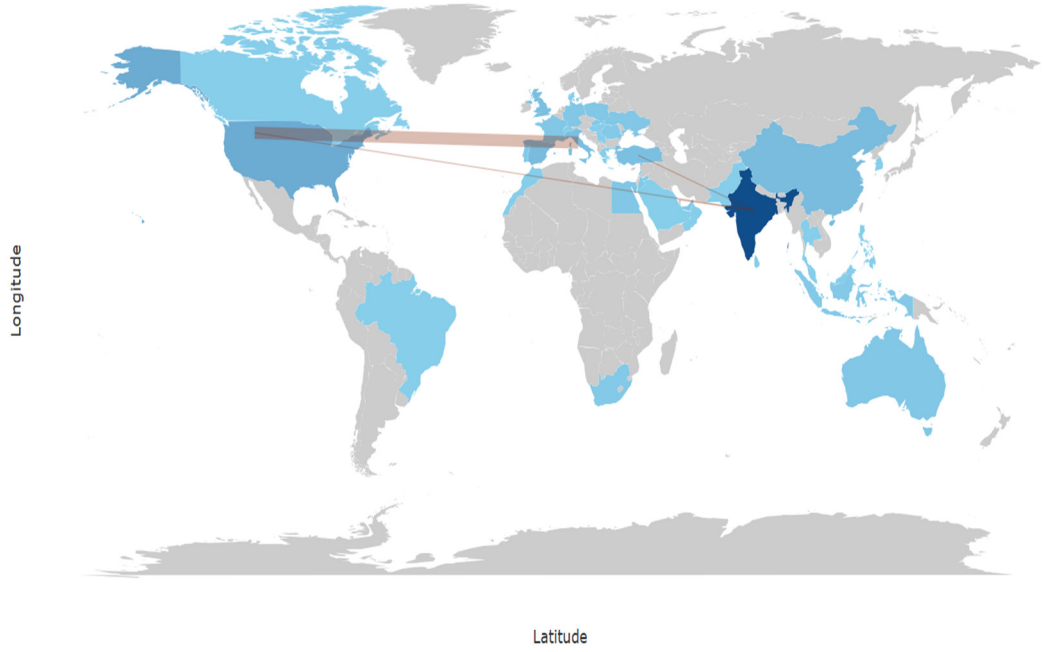


Figure 15. World map of countries' collaboration (Scopus database). Source: own processing in Bibliometrix.

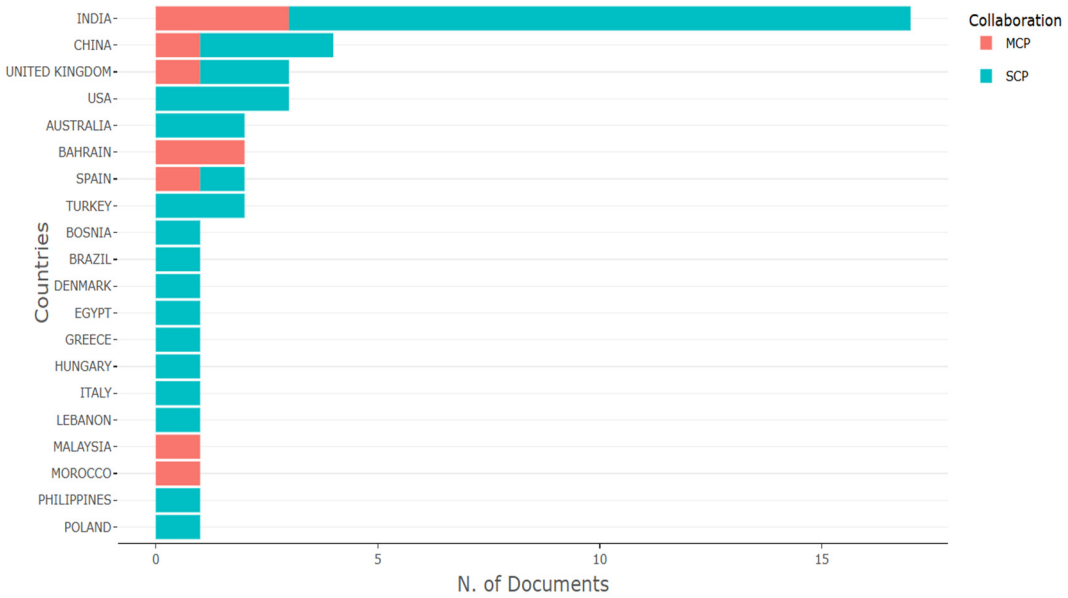


Figure 16. Corresponding authors' countries in Scopus database. Source: own processing in Bibliometrix.

4.5. Analysis of Specialized Journals

This section addresses RQ6 by examining the academic journals that serve as key platforms for disseminating the most highly cited research on artificial intelligence and machine learning in the banking sector, highlighting their role in shaping the field's scholarly discourse.

Figure 17 highlights the distribution of the most frequently cited sources within a corpus of scientific literature from the Web of Science database, revealing the importance of these sources in the local context of the study. The horizontal axis represents the number of citations, which indicates the impact and relevance of each source analyzed. The analysis provides significant insights into the preferences of the academic community and the fundamental sources used in the field of research.

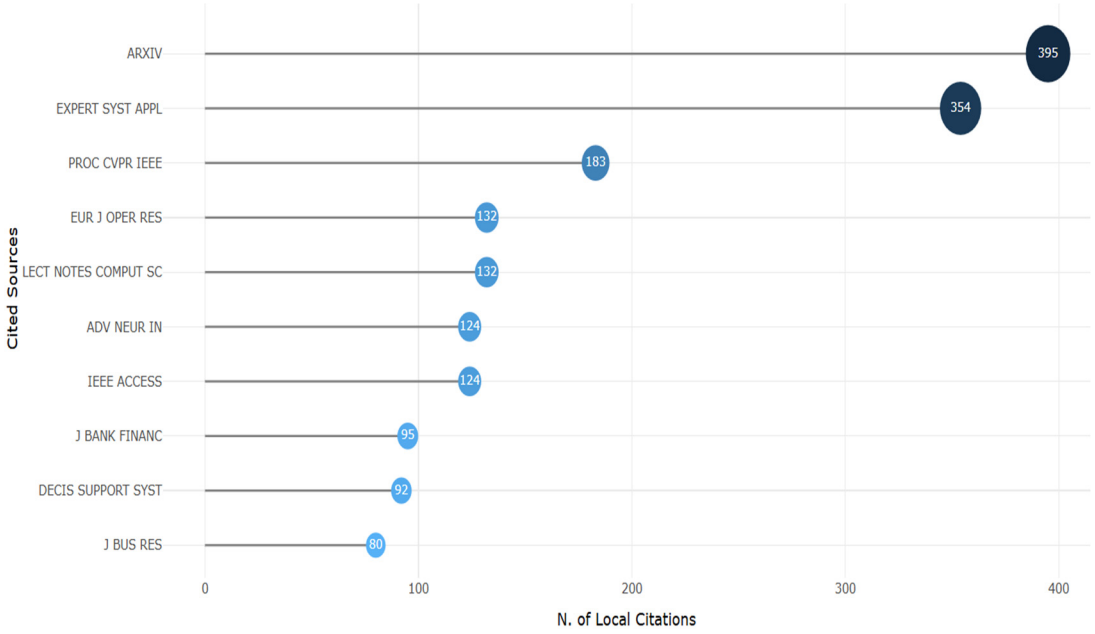


Figure 17. Most locally cited sources in Web of Science database. Source: own processing in Bibliometrix.

First, ARXIV and Expert Systems with Applications stand out as the sources with the greatest local influence. With an impressive number of 395 citations, ARXIV ranks first, demonstrating its popularity as an open platform for articles and preprints. Its free accessibility and interdisciplinary approach make ARXIV a preferred source for the latest research in data science, artificial intelligence and related fields. In second place, Expert Systems with Applications, with 354 citations, confirms its influence in applied studies related to expert systems and industrial uses of artificial intelligence.

In addition, other significant sources in the technical and operational community are noteworthy. Proceedings of CVPR IEEE, with 183 citations, reflect considerable interest in advanced research in image processing, highlighting its relevance in areas such as visual recognition and machine learning. Similarly, Lecture Notes in Computer Science and the European Journal of Operational Research, each with 132 citations, indicate high relevance for computational and analytical research. While the former supports fundamental studies in computer science, the latter emphasizes the application of AI technologies to decision-making and operations optimization.

Mid-level sources also play an important role within the scientific community. Advanced Neural Information Processing Systems (Adv Neur In) and IEEE Access, each with 124 citations, are widely used for technical and innovative approaches in deep learning, but have a smaller impact compared to the top sources. This suggests that, although their influence is consistent, they are more niche and addressed to a specialized audience.

Regarding niche sources, the Journal of Banking and Finance (95 citations) and Decision Support Systems (92 citations) highlight their importance for the financial-banking fields and the economic applications of artificial intelligence. These publications play a key role in exploring the use of AI technologies for decision support and financial analysis. The list is completed by the Journal of Business Research, with 80 citations, which highlights its contribution to studying the impact of artificial intelligence on the economy and organizational behavior.

4.6. Factor Analysis Approach Through Principal Component Analysis (PCA)

The results of the factor analysis, complemented by the bibliometric insights, underline the significant role of Artificial Intelligence (AI) and Machine Learning (ML) in banking research. These analyses reveal a comprehensive structure of research themes, interconnections, and collaborations within the domain, providing a detailed understanding of the field's dynamics.

The keyword analysis confirms the dominance of "machine learning" and "artificial intelligence", which hold the highest occurrences and total link strength. These terms underscore their pivotal role in modern banking research, particularly in areas such as risk management, process optimization, and customer interaction. The bibliometric findings corroborate this, with both databases (Web of Science and Scopus) emphasizing the centrality of these terms. However, the bibliometric analysis additionally highlights differences in thematic focus, where Web of Science leans towards emerging technologies like blockchain, while Scopus prioritizes applications in customer interaction and natural language processing.

The keyword clusters reveal a diverse thematic landscape. While Cluster 1 consolidates the foundational methodologies of AI and ML, other clusters, such as those centered around "banking", "classification", and "credit risk", illustrate applied research areas. Clusters such as big data and deep learning emphasize advanced computational techniques, while prediction and performance suggest a focus on practical outcomes in financial forecasting and operational efficiency. These findings align with bibliometric insights, where institutional collaborations and international research reflect an adaptive response to global challenges, such as the COVID-19 pandemic.

The Principal Component Analysis (PCA) results enhance the keyword analysis by identifying the variance explained by the main components. The scree plot demonstrates that the first component (PC1) captures approximately 78% of the variance, underscoring its significance in summarizing the dataset's primary trends. The orthonormal loadings suggest that variables like C1 (machine learning) and C2 (artificial intelligence) strongly influence PC1, while C3 (more niche topics) contributes less significantly. These results align with the bibliometric conclusion that AI and ML are central themes, with emerging areas like blockchain or natural language processing contributing as secondary components.

Therefore, the PCA was conducted on a reduced keyword matrix derived from co-occurrence analysis using VOSviewerversion 1.6.18. The initial list contained 54 keywords, which were then filtered to include only those with a minimum frequency ≥ 10 and total link strength ≥ 50, leading to the selection of 16 variables. These thresholds were determined based on best practices in co-word analysis to ensure thematic relevance and reduce noise [78].

The PCA was performed in EViews 12, using the correlation matrix approach with unrotated factor solutions. Prior to factor extraction, Kaiser-Meyer-Olkin (KMO) [79] and Bartlett's Test of Sphericity [80] were applied to confirm sampling adequacy and factorability. The KMO measure was 0.631—deemed "mediocre" yet acceptable per Kaiser's criteria—while Bartlett's test was significant ($p < 0.001$), justifying the use of PCA.

Three components were extracted, but interpretation was restricted to those with eigenvalues > 1.0, following the Kaiser criterion, and validated using scree plot inflection points. The first component accounted for 77.98% of the variance and reflects a dense cluster of foundational terms such as "machine learning," "artificial intelligence," and "classification." The second component explained 21.85%, dominated by emergent, high-impact topics like "blockchain," "deep learning," and "prediction." The third component contributed a negligible 0.18% and was not retained in interpretation.

The factor loadings revealed clear thematic separations, allowing for the identification of core versus peripheral research fields. This process advances beyond simple mapping and enables nuanced detection of latent structures within the bibliometric space [76].

The bibliometric analysis highlights variations in collaboration networks. While Web of Science reflects a concentration of research efforts around a few key authors, Scopus suggests a more diverse and collective approach. This distinction is important, as it reveals how different publication ecosystems foster distinct research dynamics. Moreover, the factor analysis confirms the prominence of international collaborations, particularly between the USA, China, and Europe, reinforcing the global relevance of banking research driven by AI and ML. The emphasis on keywords like "credit risk", "bankruptcy prediction", and "support vector machines" reflects the increasing application of AI and ML in addressing financial risks and enhancing predictive analytics. These findings complement the bibliometric insights on the growing relevance of customer interaction tools and optimization techniques, suggesting a convergence of theoretical advancements and practical implementations.

The bibliometric analysis also emphasizes the growing role of European and Asian contributions in banking research, possibly driven by increased digital transformation and regulatory innovation in these regions. This trend is mirrored in the clustering of applied topics like "fintech" and "finance", pointing towards region-specific research priorities. Therefore, the integration of factor analysis and bibliometric data provides a comprehensive perspective on the evolving landscape of AI and ML in banking. While the centrality of AI and ML is universally acknowledged, regional and thematic variations suggest a dynamic and multifaceted field. Future research should further explore niche areas, such as blockchain and customer interaction technologies, while fostering balanced international collaborations to maximize the impact of these advancements in banking. This combined approach ensures that theoretical innovation is matched by practical applicability, driving the sustainable transformation of financial services.

In order to prepare the dataset, we have selected the variables C1—occurrences, C2—total link strength and C3—clusters, each corresponding to the most influential keywords. (Table 1). The selection of variables in this table was guided by their ability to reflect the thematic and methodological structure of research in Artificial Intelligence (AI) and Machine Learning (ML) within the banking sector. Each variable represents a keyword that serves as a proxy for a specific research focus or methodological approach, allowing for a comprehensive analysis of the domain. The inclusion of these keywords is based on their frequency of occurrence, their connectivity within the research network, and their clustering patterns, all of which provide critical insights into the centrality, relationships, and substructures within the field.

Table 1. Variable description.

Keywords	C1—Occurrences	C2—Total Link Strength	C3—Clusters
machine learning	153	565	1
artificial intelligence	101	414	1
banking	24	145	4
classification	25	126	2
big data	22	123	9
prediction	20	110	11
performance	16	106	5
models	17	97	13
model	15	90	10
deep learning	25	89	8
bankruptcy prediction	11	88	7
finance	15	86	14
fintech	15	86	4
artificial intelligence	14	80	3
support vector machines	13	75	3
credit risk	13	74	5

Source: own processing.

Occurrences (C1) measure the prominence of each keyword in the dataset, with higher values indicating themes that are extensively studied or foundational, such as "machine learning" and "artificial intelligence." These terms represent core concepts driving innovation and research in banking. In contrast, keywords with lower occurrences, such as "support vector machines" and "bankruptcy prediction," reflect niche or specialized areas of application that contribute to specific challenges or solutions within the sector.

The total link strength (C2) quantifies the degree of co-occurrence between keywords, serving as an indicator of their interconnectedness within the research network. High link strength values suggest that these terms frequently appear together in studies, signifying their role as integrative or complementary elements in the research landscape. For instance, "big data" and "deep learning" show moderate link strengths, highlighting their importance as enablers of advanced analytics in banking.

Clusters (C3) further refine the selection by categorizing keywords into thematic groups based on their co-occurrence patterns. These clusters reveal the subfields or specialized areas within the broader domain, such as financial applications of AI (e.g., "credit risk" and "bankruptcy prediction") or methodological advancements (e.g., "classification" and "support vector machines"). Clustering also demonstrates the interdisciplinary nature of AI and ML research, where techniques and applications converge to address complex problems in banking.

The rigor of this selection process lies in its alignment with bibliometric methodologies, which prioritize variables that balance high relevance with diverse representation. By including keywords that span foundational concepts, methodological innovations, and applied topics, the analysis ensures a multidimensional understanding of the research field. This approach allows for a detailed exploration of both dominant trends and emerging areas, highlighting the interconnected and dynamic nature of AI and ML research in banking.

Furthermore, the decision to employ Principal Component Analysis (PCA) in the analysis of the selected variables is rooted in its methodological rigor and its ability to distill high-dimensional data into a reduced set of uncorrelated components. PCA is

particularly well-suited for analyzing bibliometric data because it identifies underlying patterns and relationships within a complex dataset while preserving the variance structure. This statistical technique aligns with the objectives of the study, which seeks to uncover dominant research themes and their interconnections in the field of Artificial Intelligence (AI) and Machine Learning (ML) in banking.

One of the primary reasons for using PCA is its capacity to address multicollinearity among variables. In bibliometric datasets, keywords often exhibit high correlations due to their frequent co-occurrence in the same studies or thematic areas. For instance, "machine learning" and "artificial intelligence" are strongly correlated as they represent foundational methodologies in banking research. PCA resolves this issue by transforming the original variables into orthogonal components, ensuring that each principal component (PC) represents unique dimensions of the data without redundancy. This transformation enables a clearer interpretation of the relationships among variables.

Additionally, PCA is employed to reduce the complexity of the dataset while retaining most of the original information. Bibliometric data often contain numerous variables, such as occurrences, link strengths, and clusters, which can make direct analysis cumbersome and less insightful. PCA simplifies this complexity by aggregating the variance explained by multiple variables into a few principal components. For example, in this study, the first component (PC1) may capture the variance associated with core themes like "machine learning" and "artificial intelligence," while subsequent components may represent niche areas such as "credit risk" or "deep learning." This dimensionality reduction enhances interpretability and allows researchers to focus on the most influential factors driving the research landscape.

Another justification for PCA lies in its ability to quantify the contribution of each variable to the overall dataset structure. The eigenvalues and eigenvectors computed during PCA indicate the proportion of variance explained by each component and the loading of each variable on these components. This quantification provides empirical evidence for the centrality of dominant themes, such as the high eigenvalue loading of "machine learning" on PC1, validating its importance in the field. Furthermore, the visualization tools associated with PCA, such as biplots and scree plots, facilitate the identification of clusters and outliers, offering additional insights into the thematic organization of the data.

From a methodological perspective, PCA aligns seamlessly with the goals of bibliometric and factor-analytic studies. Its application allows for a systematic and unbiased examination of research trends, ensuring that the analysis captures both the breadth and depth of the field. By uncovering latent structures and reducing dimensionality, PCA not only enhances the rigor of the study but also provides a robust foundation for interpreting the complex interplay of themes, methodologies, and applications within AI and ML research in banking. This approach ensures that the findings are both statistically sound and practically relevant, offering valuable insights into future research and policy development.

Therefore, the results of the PCA are outlined in Table 2 and Figures 18–20. The scree plot (Figure 18) presents the eigenvalues of the components extracted through Principal Component Analysis (PCA). The first eigenvalue is 2.339, explaining 77.98% of the total variance, while the second eigenvalue is 0.655 (21.85% of the variance), and the third is minimal at 0.005 (0.18% of the variance). This steep drop in eigenvalues after the first component demonstrates that the first principal component (PC1) captures most of the variability in the data. The cumulative proportion reaching 99.82% after the second component suggests that the dataset can effectively be reduced to two principal components without significant loss of information.

Table 2. Principal Component Analysis.

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.339319	1.683889	0.7798	2.339319	0.7798
2	0.655430	0.650180	0.2185	2.994750	0.9982
3	0.005250	---	0.0018	3.000000	1.0000
Eigenvectors (loadings):					
Variable	PC1	PC2	PC3		
C1	0.629750	0.325985	−0.705087		
C2	0.631475	0.313776	0.709072		
C3	−0.452387	0.891784	0.008251		
Ordinary correlations:					
	C1	C2	C3		
C1	1.000000				
C2	0.994698	1.000000			
C3	−0.475942	−0.484842	1.000000		

Source: own processing.

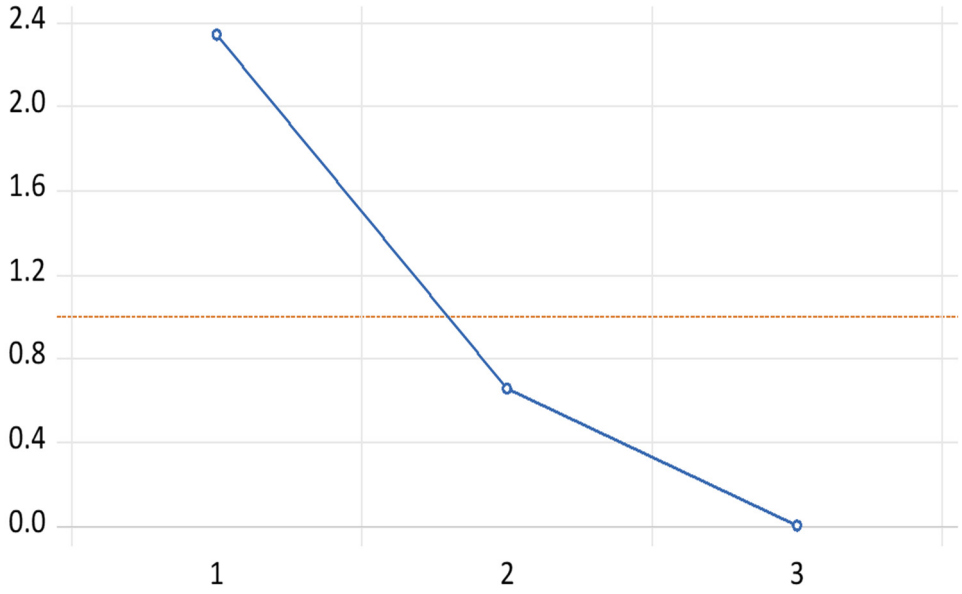
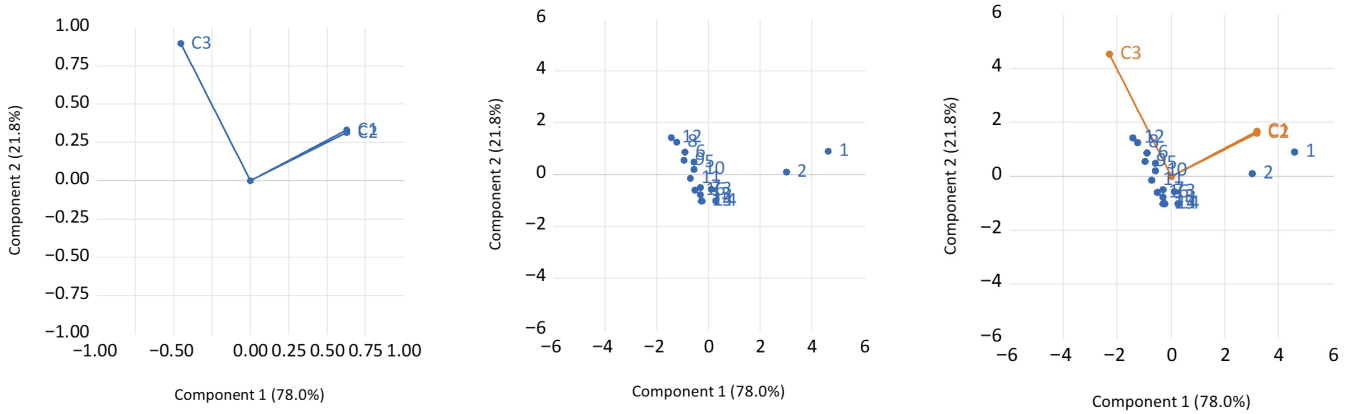


Figure 18. Scree Plot (Ordered Eigenvalues). Source: own processing.



(**A**) (**B**) (**C**)

Figure 19. (**A**)—plot of orthonormal loadings; (**B**)—plot of orthonormal scores; (**C**)—biplot of orthonormal loadings and scores. Source: own processing.

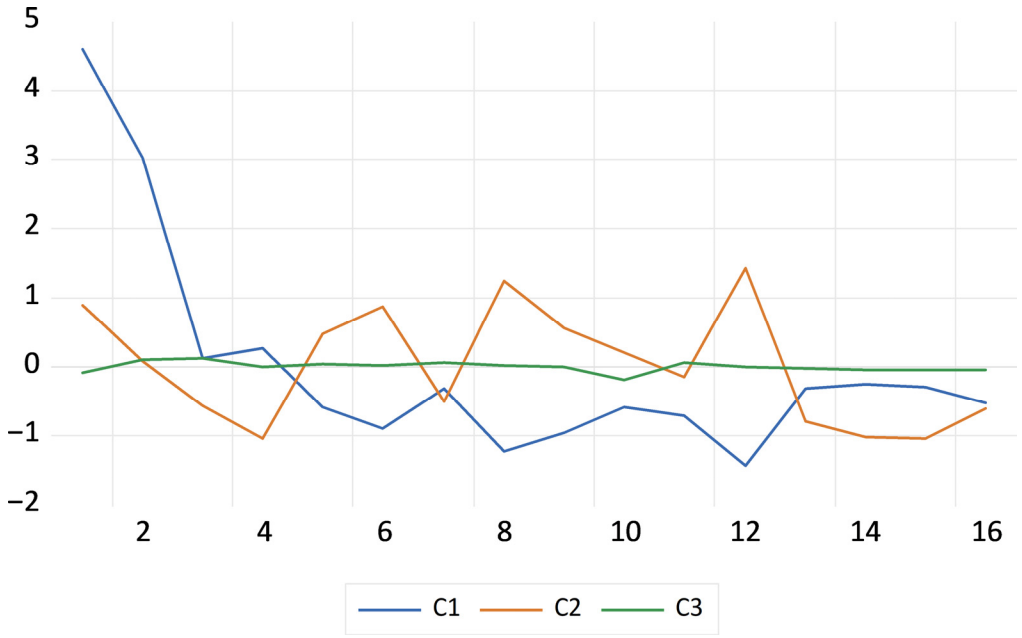


Figure 20. Variability Plot. Source: own processing.

The orthonormal loadings plot (Figure 19A) shows how variables C1, C2, and C3 are distributed across the first two components. C1 (representing "machine learning") and C2 (representing "artificial intelligence") have high positive loadings on PC1, indicating their strong influence on the primary dimension of variance. C3 (representing niche or secondary keywords) has a strong loading on PC2 and negative contributions to PC1. This plot highlights the centrality of AI and ML (captured by C1 and C2) as driving forces in the research dataset. The secondary influence of C3 underscores its role in capturing emerging or specialized areas (e.g., "deep learning" or "credit risk").

The scores plot (Figure 19B) maps individual data points (observations) onto the two primary components, PC1 and PC2. Most observations cluster around the origin, indicating a concentration of research efforts on common or overlapping themes related to AI and ML in banking. However, a few outliers (e.g., points labeled 1 and 2) are situated far from the cluster, representing unique or niche studies within the dataset. This clustering reinforces the notion that AI and ML dominate the research landscape, while the outliers reflect specialized studies focusing on unique methodologies or applications.

The biplot (Figure 19C) combines the loadings and scores, visualizing the relationship between variables (C1, C2, C3) and observations. C1 and C2 point strongly towards PC1, reflecting their dominance in the dataset. C3, pointing along PC2, aligns with studies focusing on specialized themes, as indicated by the outlier data points. The interpretation is that the research field is heavily influenced by core themes like AI and ML, while secondary areas (e.g., "classification" or "support vector machines") complement these dominant themes.

The variability plot (fluctuation of C1, C2, C3) (Figure 20) over observations reveals contrasting trends. C1 and C2 display significant variability, reflecting their dominant but diverse application across studies. C3, by contrast, remains relatively stable, indicating its limited but consistent role in capturing niche aspects of the dataset.

This stability supports the conclusion that C3 (e.g., topics like "credit risk" or "bankruptcy prediction") represents a less variable but specialized focus within the broader field of AI and ML.

Thus, the PCA results confirm that PC1 (with an eigenvalue of 2.339) captures the largest portion of variability (77.98%), followed by PC2 (21.85%), and PC3 with negligible

contribution. This aligns with the scree plot, indicating that most research themes can be reduced to two dimensions without significant loss of information. The eigenvector loadings show that C1 and C2 are positively correlated with PC1, while C3 aligns more with PC2. This division indicates that the primary component captures foundational themes (AI and ML), while the secondary component represents niche or emerging areas.

Factor analysis conducted through Principal Component Analysis (PCA) [70] serves as an important complement to bibliometric analysis performed in Bibliometrix, providing a more rigorous and multidimensional understanding of the research landscape. While bibliometric analysis offers a descriptive and relational overview by identifying trends, thematic clusters, and co-occurrence networks, PCA enhances these insights by introducing a quantitative framework to uncover latent structures, quantify contributions, and address the complexity of the dataset. The integration of these methodologies allows for a deeper exploration of both dominant themes and emerging areas in the field of Artificial Intelligence (AI) and Machine Learning (ML) in banking.

Bibliometric analysis lays the groundwork by organizing data into networks that reveal the frequency and interconnections of keywords, clusters, and research collaborations. However, its descriptive nature limits its ability to quantify the relationships between variables or to prioritize themes based on their explanatory power. This is where PCA becomes indispensable, as it identifies the principal dimensions of variance within the dataset, effectively reducing its complexity while retaining its most significant patterns. By transforming correlated variables into orthogonal components, PCA eliminates redundancy and ensures that each principal component represents a unique aspect of the research landscape.

Moreover, PCA not only validates but also refines the findings of bibliometric analysis. For instance, bibliometric tools might identify "machine learning" and "artificial intelligence" as central themes based on their high frequency and strong co-occurrence. PCA corroborates this centrality by demonstrating that these variables dominate the variance explained by the first principal component, providing empirical evidence for their foundational role. Furthermore, PCA's capacity to quantify the variance associated with secondary components highlights emerging but less prominent themes, such as "deep learning" or "credit risk," offering a nuanced understanding of niche areas that might otherwise be overlooked.

The integration of PCA and bibliometric analysis also facilitates a more systematic visualization of the research structure. While bibliometric clustering maps thematic groupings, PCA enhances this by organizing variables and observations into a hierarchy of components that reflects their statistical significance. This synergy not only clarifies the thematic organization but also enables the identification of outliers or unique contributions, providing actionable insights for future research directions.

Thus, factor analysis through PCA [72] enriches bibliometric analysis by adding statistical rigor and revealing hidden dimensions within the data. This complementary relationship ensures a comprehensive perspective that balances descriptive and quantitative insights, making it possible to capture both overarching trends and subtle thematic variations. Such an integrated approach is particularly valuable in the dynamic and complex field of AI and ML in banking, where a multidimensional understanding of research themes and their interconnections is essential for advancing knowledge and informing evidence-based policy decisions.

When integrated with the bibliometric analysis, these results paint a comprehensive picture of the research field:

- The dominance of AI and ML as central themes is evident both in PCA and bibliometric analyses, supported by high occurrences and strong linkages in keyword networks.

- The PCA's secondary components align with bibliometric findings that highlight emerging areas like customer interaction tools in Scopus or blockchain in Web of Science.
- The internationalization and collaborative nature of research, emphasized in bibliometric analysis, are indirectly reflected in the clustering of data points, showing overlaps and shared themes across studies.

The factor analysis complements the bibliometric findings by quantitatively confirming the centrality of AI and ML while identifying secondary dimensions that represent emerging or specialized research areas. The visualizations and statistical results collectively demonstrate a research field characterized by dominant methodologies and applications, alongside a growing exploration of niche topics. Future research can leverage these insights to explore underrepresented areas or enhance interdisciplinary collaboration.

The integration of bibliometric analysis and factor analysis offers a comprehensive and rigorous exploration of research trends in Artificial Intelligence (AI) and Machine Learning (ML) within the banking sector. By combining these two methodologies, this study bridges quantitative insights from statistical modeling with qualitative and networked understandings from bibliometric tools. This hybrid approach not only confirms the centrality of core themes but also uncovers nuanced patterns that would remain obscured when relying on a single method.

5. Discussion

5.1. Interpretation of Results from a Sustainable Banking Perspective

The keyword network analysis confirms the central role of artificial intelligence and machine learning concepts in banking literature, which are mainly associated with risk management, performance analysis, and predictive models. This thematic structure indicates that current research uses AI and ML primarily as quantitative tools for optimizing banking processes and managing traditional financial risks. The predominant focus on operational efficiency and predictive capability suggests a consolidation of these technologies in the core functions of banking.

From the perspective of ESG and sustainable banking, however, the results show that these dimensions are poorly integrated into the core of AI and ML-based research. Although AI and ML provide an appropriate analytical framework for assessing climate risks, analyzing the social impact of banking, and improving decision-making governance, these directions appear marginal in the structure of thematic networks. Concepts associated with sustainability and green transition are poorly connected to the dominant clusters, suggesting that sustainable banking is not yet systematically addressed by advanced artificial intelligence tools.

The differences between the Web of Science and Scopus databases reinforce this interpretation. In Web of Science, the strong orientation towards technological innovation and advanced machine learning methods indicates a focus on the development of analytical tools, with a low recurrence of topics related to banking sustainability. In Scopus, although the thematic structure is more diverse and includes applications oriented towards digital services and financial behavior, the directions associated with sustainable banking do not form a distinct and coherent cluster. In conclusion, the results suggest that ESG and sustainable banking are still peripheral in the literature on AI and ML in banking, being insufficiently correlated with the risk and performance models that dominate current research.

The analysis of author and co-citation networks highlights structural differences between the Web of Science and Scopus databases, reflecting distinct levels of intellectual maturity and thematic orientation of AI and ML research in the banking sector. From a sustainable banking perspective, these differences are relevant because they indicate how scientific

knowledge is organized around long-term risks, systemic stability, and the structural transformation of financial intermediation.

In Web of Science, the existence of consolidated intellectual hubs, dominated by authors with consistent output and high visibility in the co-citation network, suggests methodologically stable literature. This stability is associated with the development and refinement of quantitative models used in banking risk management. Such a structure favors analytical consistency and research continuity but may lead to slower integration of emerging dimensions such as climate risks, transition risks, and the assessment of long-term impacts on banking portfolios.

In contrast, the network of authors and co-citations in Scopus is more dispersed and less hierarchical, reflecting a greater diversity of theoretical and methodological perspectives. This structure indicates greater openness to new and interdisciplinary topics, but also a lack of conceptual convergence in the integrated approach to sustainability through AI and ML. The absence of dominant authors or theoretical frameworks suggests that sustainability-related dimensions are still being explored in a fragmented manner, without being consolidated into a unified analytical framework.

A comparative analysis of co-citation networks indicates that, in both databases, the literature continues to be dominated by paradigms oriented towards traditional financial risk and operational efficiency. From a sustainable banking perspective, this suggests that AI and ML are predominantly used to optimize short-term decisions, while their potential for analyzing structural, climate, and sustainability risks is insufficiently integrated into the theoretical core of the field.

An *analysis of institutional collaboration* networks in Web of Science and Scopus highlights relevant differences in the structure and evolution of research on artificial intelligence and machine learning in the banking sector. From the perspective of the long-term sustainability and resilience of the banking system, these networks reflect how scientific expertise is concentrated and disseminated at the institutional level.

In Web of Science, the steady increase in academic output after 2017 indicates a consolidation of research focused on the development of advanced analytical models. This dynamic suggests the existence of stable research centers that contribute to the methodological continuity necessary for assessing structural and systemic risks. The geographical diversification of the institutions involved broadens the basis for analysis to different banking systems, increasing the literature's ability to capture heterogeneous vulnerabilities.

In Scopus, the structure of collaborations is more concentrated, with dominant institutions and rapid increases in output. This configuration reflects a more applied orientation and a focus on specific directions, which may accelerate the exploration of emerging themes but limits the development of integrated analytical frameworks on the long-term stability of the banking sector.

Overall, the analysis shows that institutional networks are expanding and becoming increasingly internationalized, but directions related to sustainability, resilience, and structural transformation of banking activity are not yet consolidated around coherent institutional collaborations. The differences between the two databases reflect distinct models of research organization, with direct implications for how AI and ML can support the long-term stability of the banking system.

The analysis of country-level collaborations highlights structural differences between the Web of Science and Scopus databases in terms of the organization and internationalization of research on artificial intelligence and machine learning applied to the banking sector. From the perspective of the long-term stability and resilience of the banking system, these networks indicate how research capacity and knowledge exchange are concentrated globally.

In Web of Science, the United States, China, and Western European countries act as central nodes of international collaborations, facilitating intense flows of knowledge between advanced research centers. This structure suggests a concentration of expertise in mature economies, which influences the dominant directions of research and the development of analytical models used in banking risk assessment. The presence of emerging connections with countries in Central and Eastern Europe indicates a gradual expansion of networks, but with a still secondary role in global scientific output.

In Scopus, collaborations are more evenly distributed between national and international research, with countries such as India, the United States, China, and the United Kingdom playing a central role. This structure reflects a diversification of contributions and a broader integration of research from emerging economies, without however changing the overall hierarchy of scientific influence.

In both databases, the intensification of research activity in 2020–2024 indicates an acceleration of interest in digitalization and structural transformations in the banking sector. The results show that, although international collaborations are extensive, research remains concentrated around a limited number of advanced economies, with direct implications for the dominant directions of AI and ML development in banking.

5.2. Factor Analysis: Novel Contributions and Implications for Academia and Practitioners

Factor analysis, as applied in this study through Principal Component Analysis (PCA), provides a structured method for reducing multidimensional data into its principal components. By extracting and analyzing eigenvalues, eigenvectors, and correlations, factor analysis identifies latent patterns and hierarchical structures that characterize the research field. The strength of this approach lies in its ability to quantify variance and identify dominant dimensions, such as the foundational importance of AI and ML, while simultaneously isolating niche contributions, such as "credit risk" or "deep learning".

The methodological rigor of factor analysis is evident in the systematic decomposition of variance. The scree plot revealed that the first principal component (PC1) accounts for nearly 78% of the total variance, emphasizing the dominance of key themes. Furthermore, the orthonormal loadings and scores clearly delineate the relationships between keywords and their underlying dimensions. These results provide an empirical foundation for interpreting the prominence of thematic clusters while ensuring objectivity and replicability. Unlike bibliometric tools, which rely on co-occurrence networks, factor analysis adds a robust statistical layer, quantifying the relative influence of variables across dimensions.

While bibliometric analysis maps thematic networks and author collaborations, it lacks the quantitative precision necessary to evaluate hierarchical relationships or measure variance within the data. However, when integrated with factor analysis, as in this study, it enables a more detailed exploration of structural dynamics. For instance, bibliometric results underscore the centrality of "artificial intelligence" and "machine learning," corroborating factor-analytic findings that these themes dominate PC1. Moreover, the bibliometric insight into thematic shifts between databases—where Web of Science emphasizes blockchain and Scopus highlights customer interactions—aligns with the orthonormal scores, which show how PC2 captures these secondary but emerging areas.

This complementarity is particularly evident in the identification of niche clusters. Bibliometric clustering revealed distinct themes such as "fintech" and "deep learning," which factor analysis further validated by showing their limited but focused contributions to PC2. Together, these methodologies capture both the macro trends (core themes like AI and ML) and micro trends (specific applications like "support vector machines" or "bankruptcy prediction"), offering a holistic perspective on the research landscape.

The results provide a layered understanding of the field. The dominance of PC1, as shown in the scree plot and eigenvalues, reflects a strong concentration of research on AI and ML methodologies. These technologies, represented by variables C1 and C2, are not only the most frequently occurring but also the most interconnected, as evidenced by their high eigenvector loadings and correlation coefficients. This centrality underscores their foundational role in advancing predictive analytics, fraud detection, and process optimization within banking.

However, the second principal component (PC2) highlights the diversity within the field, capturing emerging areas that extend beyond traditional AI and ML applications. For example, topics such as "deep learning" and "credit risk," while secondary, represent focused streams of research that address specific challenges in banking. The stability of C3, as seen in the variability plot, reinforces its role in anchoring these niche areas, providing consistent but less dominant contributions.

The orthonormal loadings and scores further reveal how individual studies align with these components. The clustering of observations near the origin indicates a convergence of research around common themes, such as risk management and customer interactions. Conversely, the presence of outliers in the scores plot reflects innovative or specialized studies, such as those exploring blockchain or non-traditional credit evaluation methods. These findings align with bibliometric insights into international collaboration, where regions like Asia and Europe have driven niche innovations, influenced in part by the COVID-19 pandemic.

Overall, the dominant principal component (PC1) highlights well-established themes such as predictive analytics and AI/ML integration in operational banking. However, the second component (PC2), representing niche topics like "credit risk" and "bankruptcy prediction," reveals gaps in integrating ESG-specific variables and climate-related stress testing. Additionally, regional clustering showed that European and Asian institutions focus on fintech and blockchain, but less on sustainability-linked outcomes. These findings suggest underexplored opportunities for applying ML in climate risk modeling and green finance, areas currently underserved in the literature.

The integration of bibliometric and factor-analytic approaches introduces several *novel contributions to the understanding of AI and ML research in banking*. First, the use of factor analysis quantifies the dominance and interrelations of variables, providing empirical validation for bibliometric observations. For example, while bibliometric clustering identifies "machine learning" as a central node, factor analysis assigns it the highest eigenvalue loading, confirming its statistical significance.

Second, the combination of methodologies reveals hidden patterns that would remain inaccessible in isolation. Bibliometric tools, while effective in mapping collaborations and thematic networks, cannot quantify the variance explained by different themes. Conversely, factor analysis, while robust in dimensional reduction, lacks the networked perspective offered by bibliometric tools. Together, they offer a multidimensional view, balancing quantitative precision with qualitative richness.

Finally, this integrated approach enhances the granularity of insights. For instance, bibliometric findings show a shift towards customer-centric applications in Scopus, while factor analysis, through PC2, quantifies the variance associated with these emerging trends. This granularity not only validates the findings but also offers actionable insights for future research and policy.

The combined use of factor analysis and bibliometric methodologies has significant *implications for both academic research and practical applications in banking*. Academically, it establishes a robust framework for future studies, enabling researchers to systematically explore thematic hierarchies and interrelations. Practically, it highlights the critical areas

for innovation, such as the integration of AI with fintech solutions or the use of ML for non-traditional credit assessments. Furthermore, the international collaboration trends revealed by bibliometric analysis suggest opportunities for cross-border partnerships, particularly in addressing global challenges like financial inclusion and cybersecurity.

By integrating bibliometric and factor-analytic methodologies, this study provides a comprehensive and nuanced understanding of the role of AI and ML in banking research. The complementarity of these approaches not only validates core themes but also uncovers emerging trends and niche areas, offering a multidimensional perspective on the research landscape. This hybrid methodology represents a significant advancement in the analysis of scientific literature, demonstrating its potential to inform both academic inquiry and practical innovation in the rapidly evolving field of banking technology. Future studies should continue to leverage these methodologies, exploring their application across other domains and expanding the scope of interdisciplinary collaborations.

5.3. Directions for Future Research

The results of this study show that artificial intelligence and machine learning are currently used primarily to assess traditional financial risks and improve banking performance. For sustainability-oriented banks, a clear direction for research is to analyze how these technologies can be used to identify customers and projects that pose long-term risks, such as exposure to climate change or major economic transformations (Table 3).

Table 3. Future Research Directions on AI and ML in Banking.

No.	Research Direction	Description	Relevance for Sustainable and Resilient Banking
1	Integration of Long-Term Risks into AI Models	Extension of existing credit risk and stress-testing models by incorporating climate-related and structural risk variables.	Enables assessment of portfolio stability beyond traditional short-term financial risks.
2	AI-Supported Decision-Making for Sustainable Banking	Analysis of how AI-generated outputs support banking decisions that consider long-term impacts rather than short-term profitability.	Supports a more stable and responsible banking framework.
3	Cross-Country and Cross-Bank Comparative Studies	Comparative analyses of AI and ML applications across banks from developed and emerging economies using consistent methodologies.	Identifies institutional and infrastructural constraints affecting AI adoption.
4	Human–AI Interaction in Banking Decisions	Investigation of how banking professionals interpret, adjust, or override AI-based recommendations.	Improves understanding of real-world AI deployment and decision robustness.
5	Impact Assessment of AI-Assisted Decisions	Comparison between AI-assisted and traditional decisions at portfolio or institutional level.	Allows evaluation of AI's actual contribution to banking stability.
6	Empirical Studies Based on Real Banking Data	Use of operational banking datasets or collaborations with financial institutions for model validation.	Enhances practical relevance and applicability of academic research.

Source: own processing.

Another important direction is to study how banks can combine automated decisions with sustainable development goals. Future research could examine whether AI models can support lending decisions that focus not only on immediate profit, but also on the long-term stability of bank portfolios and their impact on the environment and society.

Greater attention also needs to be paid to differences between banks and countries. Future studies could compare how banks in developed and emerging economies use AI to support more responsible banking. This would help to understand the real limitations related to data, technology, and institutional capacity.

A practical and easily applicable direction is to analyze how bank employees use AI-generated results. Research can examine whether these results are followed automatically or whether they are adjusted by specialists, especially in decisions that have a long-term impact on the stability of the bank.

Finally, future literature should focus more on studies based on real data from banks, not just simulations or theoretical models. Such research can show more clearly the extent to which artificial intelligence actually contributes to building a more stable, responsible, and sustainable banking sector.

6. Conclusions

The findings of this study reflect a comprehensive and detailed picture of the evolution and research directions on the use of artificial intelligence (AI) and machine learning (ML) in the banking sector, based on the analysis of Web of Science and Scopus databases.

With regard to RQ1 *(How has the publication of academic articles on the use of artificial intelligence and machine learning in the banking sector evolved according to data from Scopus and Web of Science databases?)*, the analysis of the evolution of academic publications shows a rapid and diversified growth of interest in the application of AI and ML in banking, especially in the period 2020–2024. This expansion can be correlated with the recent global challenges, such as the COVID-19 pandemic, which have accentuated the need for digitization and optimization of banking processes.

For RQ2 *(What are the main emerging research directions on the use of artificial intelligence and machine learning in the banking sector, as identified through keyword network analysis?)*, the keyword network analysis shows that *artificial intelligence* and *machine learning* are the core concepts shaping current research in banking, mainly linked to risk management and process optimization. Beyond these central themes, the emerging research directions differ across databases. In the Web of Science, greater attention is given to technological innovations, particularly the integration of blockchain with AI-based solutions. In contrast, Scopus emphasizes AI applications focused on customer interactions, including *natural language processing* and data-driven personalization. Overall, the results indicate a dual research trajectory, combining technological innovation with customer-oriented and operational applications of AI in banking.

In terms of author contributions *(RQ3: Who are the authors with the highest scientific contributions in the field of artificial intelligence and machine learning applied to the banking sector, according to publications indexed in Scopus and Web of Science?)*, Web of Science shows a concentration of collaborations around a small number of active authors, such as Zhang Y and Cheng D, indicating a more focused collaboration. On the other hand, Scopus reflects a more diverse network with a more balanced distribution of contributions, suggesting a collective approach to research.

The analysis of institutional collaborations *(RQ4: Which research institutions have had the greatest impact on the development of artificial intelligence and machine learning research in the banking sector?)* shows a marked internationalization of research, especially in the period 2020–2024, with a significant increase in contributions from Europe and Asia, probably

influenced by global challenges such as the COVID-19 pandemic. This is particularly important as it reflects a trend towards globalization of research in banking and emerging technologies, which can support the development of innovative solutions to global economic and financial problems. International collaborations not only improve the quality and diversity of research but also facilitate knowledge sharing between regions affected by global crises, stimulating adaptation and the implementation of effective strategies in the face of common challenges.

In terms of geographical distribution of research *(RQ5: How is research on artificial intelligence and machine learning in the banking sector geographically distributed and which countries have the most intense scientific activity in this field?)*, both Web of Science and Scopus emphasize the central role of the USA, China and Europe. However, Scopus suggests a greater balance between national and international research, while Web of Science shows an already well-established network of transcontinental collaborations.

Finally, regarding the academic journals with the highest impact *(RQ6: Which academic journals publish the most influential research on artificial intelligence and machine learning in banking, by number of citations?)*, the most influential research on artificial intelligence and machine learning in the banking sector is mainly published in ARXIV and Expert Systems with Applications, which have the highest citation rates and serve as central platforms for disseminating impactful results.

From the analysis, we observe that, even if we did not exclude the duplicates from both databases, we obtained different results. This confirms that the study carried out offers distinct and complementary perspectives on the topic under analysis. Although there is some overlap, the differences identified emphasize the importance of using multiple sources for a broader and more accurate understanding. The results show that each database makes a unique contribution, reflecting the diversity and complexity of the information available. This diversity allows us to gain a more comprehensive picture and to identify aspects that would have been missed using a single source.

From a *novelty perspective*, the study stands out by combining factor analysis through factor-analytic techniques with bibliometric clustering, an integration that has rarely been applied to banking research. Moreover, the use of eigenvalue decomposition to quantify variance explained by core and niche research themes introduces an empirical framework for evaluating thematic centrality. Additionally, the findings extend beyond methodological insights by demonstrating how bibliometric data, such as co-occurrence networks, can be quantitatively analyzed to provide actionable insights into the structure of a rapidly evolving field.

The implications for *policymakers* are profound, as the study underscores the strategic importance of fostering research in AI and ML to enhance banking efficiency, financial inclusion, and fraud detection. Furthermore, it highlights the need for investments in emerging areas like blockchain and customer interaction technologies, emphasizing their potential to transform banking operations. Policymakers can also leverage the study's insights into international collaborations, which reveal the pivotal role of global partnerships in driving innovation. Nevertheless, the study also draws attention to ethical considerations, particularly in the equitable application of AI technologies, suggesting the importance of developing regulatory frameworks that balance innovation with fairness and privacy.

Despite its combination of factor analysis and bibliometric analysis providing a structured overview of the literature, this study has several *limitations*. First, the analysis is based exclusively on the Web of Science and Scopus databases, which may lead to the omission of relevant works indexed in other sources. Second, language restrictions may limit the geographical representativeness of the results.

In addition, bibliometric analysis is based on the co-occurrence of keywords, which captures general thematic structures but does not reflect in detail the full content of the studies analyzed. Finally, the results are not directly correlated with concrete indicators of sustainability or climate performance at the banking level, which limits the assessment of the practical impact of the research. These limitations open up clear directions for future research that integrates more diverse data sources and applied empirical analyses.

In conclusion, the study offers a robust methodological contribution and actionable insights for both academic and policy-making audiences. By addressing its limitations and pursuing the identified future directions, subsequent research can further refine our understanding of AI and ML in banking, ensuring their continued impact on this significant sector.

Author Contributions: Conceptualization, A.G.M. and C.G.; methodology, A.G.M., C.G. and R.M.B.; software, A.G.M. and C.G.; validation, A.G.M., J.P. and L.F.M.; formal analysis, A.G.M. and C.G.; investigation, C.G. and M.O.; resources, A.G.M.; data curation, J.P. and M.O.; writing—original draft preparation, C.G. and A.G.M.; writing—review and editing, A.G.M. and R.M.B.; visualization, R.M.B. and L.F.M.; supervision, A.G.M. and L.F.M.; project administration, A.G.M. and C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: All data were obtained from the Web of Science Core Collection and Scopus Database. Publicly available datasets were analyzed in this study. This data can be found here: https://0c10qjxkk-y-https-www-webofscience-com.z.e-nformation.ro/wos/woscc/basic-search and here: https://0c109wro9-y-https-www-scopus-com.z.e-nformation.ro/results/results.uri?st1=Artificial+intelligence+and+machine+learning+in+banking&st2=&s=TITLE-ABS-KEY(Artificial+intelligence+and+machine+learning+in+banking)&limit=10&origin=searchbasic&sort=plf-f&src=s&sot=b&sdt=b&sessionSearchId=0738e4d96c169dff733b193639bc1f71 (accessed on 3 December 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Li, X.; Tang, P. Stock Index Prediction Based on Wavelet Transform and FCD-MLGRU. *J. Forecast.* **2020**, *39*, 1229–1237. [CrossRef]
2. Wall, L.D. Some Financial Regulatory Implications of Artificial Intelligence. *J. Econ. Bus.* **2018**, *100*, 55–63. [CrossRef]
3. Gherțescu, C.; Manta, A.G. Fintech Trends and Banking Digitalization: Insights from a Bibliometric Analysis. *Financ.-Chall. Future* **2023**, *24*, 24–36.
4. J.P. Morgan: COiN—A Case Study of AI in Finance. *Superior Data Science*. 2024. Available online: https://superiordatascience.com/jp-morgan-coin-a-case-study-of-ai-in-finance/ (accessed on 20 February 2025).
5. Revolut. Revolut Launches AI Feature to Protect Customers from Card Scams and Break the Scammers' "Spell", 2024. Available online: https://www.revolut.com/en-FR/news/revolut_launches_ai_feature_to_protect_customers_from_card_scams_and_break_the_scammers_spell/ (accessed on 20 February 2025).
6. Zoting, S. Artificial Intelligence (AI) in Banking Market Report by 2033. *Preced. Res.* **2024**. Available online: https://www.precedenceresearch.com/artificial-intelligence-in-banking-market (accessed on 10 December 2024).
7. Columbia University School of Engineering and Applied Science: Artificial Intelligence (AI) vs. Machine Learning, 2023. Available online: https://ai.engineering.columbia.edu/ai-vs-machine-learning/ (accessed on 10 January 2025).
8. Manta, A.G.; Ghertescu, C.; Bădîrcea, R.M.; Manta, L.F.; Popescu, J.; Lăpădat, C.V.M. How Does the Interplay Between Banking Performance, Digitalization, and Renewable Energy Consumption Shape Sustainable Development in European Union Countries? *Energies* **2025**, *18*, 571. [CrossRef]
9. Kothuri, R.K.; Samala, R.K. Artificial intelligence based hybridization for economic power dispatch. *Cogn. Robot.* **2023**, *3*, 218–225. [CrossRef]
10. Faggella, D. Everyday Examples of Artificial Intelligence and Machine Learning. *Emerj.* 2020. Available online: https://emerj.com/everyday-examples-of-ai/ (accessed on 10 January 2025).

11. Flood, M.D.; Jagadish, H.V.; Raschid, L. Big Data Challenges and Opportunities in Financial Stability Monitoring. *Financ. Stab. Rev.* **2016**, *20*, 129–142.
12. Goodell, J.W.; Kumar, S.; Lim, W.M.; Pattnaik, D. Artificial Intelligence and Machine Learning in Finance: Identifying Foundations, Themes, and Research Clusters from Bibliometric Analysis. *J. Behav. Exp. Financ.* **2021**, *32*, 100577. [CrossRef]
13. Siminică, M.I.; Cîrciumaru, D.; Manta, A.G.; Cârstina, S.; Badareu, G.; Gherțescu, C. FinTech, artificial intelligence, and European Union banks: A double-edged sword for performance? *Oeconomia Copernic.* **2025**, *16*, 1099–1176. [CrossRef]
14. Narang, A.; Vashisht, P.; Bajajaj, S.B. Artificial Intelligence in Banking and Finance. *Int. J. Innov. Res. Comput. Sci. Technol.* **2024**, *12*, 130–134. [CrossRef]
15. Polireddi, N.S.A. An Effective Role of Artificial Intelligence and Machine Learning in the Banking Sector. *Meas. Sens.* **2024**, *33*, 101135. [CrossRef]
16. Dewasiri, N.J.; Karunarathnage, S.S.S.N.; Karunarathne, S.M.; Potupitiya Gamaathige, S.A.J.; Rathnasiri, M.S.H. Fusion of Artificial Intelligence and Blockchain in the Banking Industry: Current Application, Adoption, and Future Challenges. In *Transformation for Sustainable Business and Management Practices: Exploring the Spectrum of Industry 5.0*; Saini, A., Garg, V., Eds.; Emerald Publishing Limited: Leeds, UK, 2023; pp. 293–307. [CrossRef]
17. Roseline, J.F.; Naidu, G.; Pandi, V.S.; Alias Rajasree, S.A.; Mageswari, N. Autonomous Credit Card Fraud Detection Using Machine Learning Approach. *Comput. Electr. Eng.* **2022**, *102*, 108132. [CrossRef]
18. Pazarbasioglu, C.; Garcia Mora, A.; Uttamchandani, M.; Natarajan, H.; Feyen, E.; Saal, M. Digital Financial Services. *World Bank Group*. 2020. Available online: https://thedocs.worldbank.org/en/doc/305a39cbb6f35567db78bda6709c5cd8-0430012025/original/World-Bank-DFS-Whitepaper-DigitalFinancialServices.pdf (accessed on 14 October 2025).
19. Manta, L.F.; Manta, A.G.; Gherțescu, C. Decoding Digital Synergies: How Mechatronic Systems and Artificial Intelligence Shape Banking Performance Through Quantile-Driven Method of Moments. *Appl. Sci.* **2025**, *15*, 5282. [CrossRef]
20. Ghertescu, C.; Manta, A.G.; Bădîrcea, R.M.; Manta, L.F. How Does the Digitalization Strategy Affect Bank Efficiency in Industry 4.0? A Bibliometric Analysis. *Systems* **2024**, *12*, 492. [CrossRef]
21. Tierno, P. Artificial Intelligence and Machine Learning in Financial Services (CRS Report No. R47997). *Congr. Res. Serv.* **2024**, *10*, e23492. Available online: https://www.congress.gov/crs-product/R47997 (accessed on 20 October 2025).
22. Islam, A.; Islam, M.; Hossain Uzir, M.U.; Abd Wahab, S.; Abdul Latiff, A.S. The Panorama Between COVID-19 Pandemic and Artificial Intelligence (AI): Can It Be the Catalyst for Society 5.0. *Int. J. Sci. Res. Manag.* **2020**, *8*, 2011–2025.
23. Durongkadej, I.; Hu, W.; Wang, H.E. How Artificial Intelligence Incidents Affect Banks and Financial Services Firms? A Study of Five Firms. *Financ. Res. Lett.* **2024**, *70*, 106279. [CrossRef]
24. Leitner, G.; Singh, J.; van der Kraaij, A.; Zsámboki, B. The Rise of Artificial Intelligence: Benefits and Risks for Financial Stability. *Financial Stability Review*, European Central Bank. 2024. Available online: https://www.ecb.europa.eu/press/financial-stability-publications/fsr/special/html/ecb.fsrart202405_02 58c3ce5246.en.html (accessed on 30 October 2025).
25. Naveed, H.; Khan, A.U.; Qiu, S.; Saqib, M.; Anwar, S.; Usman, M.; Akhtar, N.; Barnes, N.; Mian, A. A comprehensive overview of large language models. *J. LaTeX* **2024**, *16*, 172. [CrossRef]
26. Fundira, M.; Edoun, E.I.; Pradhan, A.; Mbohwa, C. Assessing digital competencies and AI ethics awareness among customers in the banking sector. *Afr. J. Sci. Technol. Innov. Dev.* **2024**, *16*, 792–807. [CrossRef]
27. European Parliament; Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off. J. Eur. Union* **2016**, *119*, 1–88. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679 (accessed on 20 January 2025).
28. Ozili, P.K. Artificial intelligence in central banking: Benefits and risks of AI for central banks. In *Industrial Applications of Big Data, AI, and Blockchain*; IGI Global: New York, NY, USA, 2024. [CrossRef]
29. Milana, C.; Ashta, A. Artificial intelligence techniques in finance and financial markets: A survey of the literature. *Strateg. Change* **2021**, *30*, 189–209. [CrossRef]
30. Manta, A.G.; Bădîrcea, R.M.; Doran, N.M.; Badareu, G.; Ghertescu, C.; Popescu, J. Industry 4.0 Transformation: Analysing the Impact of Artificial Intelligence on the Banking Sector through Bibliometric Trends. *Electronics* **2024**, *13*, 1693. [CrossRef]
31. Fares, O.H.; Butt, I.; Lee, S.H.M. Utilization of artificial intelligence in the banking sector: A systematic literature review. *J. Financ. Serv. Mark.* **2023**, *28*, 835–852. [CrossRef]
32. Doumpos, M.; Zopounidis, C.; Gounopoulos, D.; Platanakis, E.; Zhang, W. Operational research and artificial intelligence methods in banking. *Eur. J. Oper. Res.* **2023**, *306*, 1–16. [CrossRef]
33. Jena, J.R.; Biswal, S.K.; Shrivastava, A.K.; Panigrahi, R.R. A bibliographic overview of financial engineering in the emerging financial market. *Int. J. Syst. Assur. Eng. Manag.* **2023**, *14*, 2048–2065. [CrossRef]
34. Zungu, N.P.; Amegbe, H.; Hanu, C.; Asamoah, E.S. AI-driven self-service for enhanced customer experience outcomes in the banking sector. *Cogent Bus. Manag.* **2025**, *12*, 2450295. [CrossRef]

35. Haddad, H. The effect of artificial intelligence on the AIS excellence in Jordanian banks. *Montenegrin J. Econ.* **2021**, *17*, 155–166. [CrossRef]
36. Gyau, E.B.; Appiah, M.; Gyamfi, B.A.; Achie, T.; Naeem, M.A. Transforming banking: Examining the role of AI technology innovation in boosting banks' financial performance. *Int. Rev. Financ. Anal.* **2024**, *96*, 103700. [CrossRef]
37. Alessi, L.; Savona, R. Machine learning for financial stability. In *Data Science for Economics and Finance: Methodologies and Applications*; Consoli, S., Reforgiato Recupero, D., Saisana, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2021; pp. 65–87. [CrossRef]
38. Bashar, A.; Rabbani, M.R.; Khan, S.; Moh'd Ali, M.A. Data driven finance: A bibliometric review and scientific mapping. In Proceedings of the 2021 International Conference on Data Analytics for Business and Industry, ICDBI, Sakheer, Bahrain, 25–26 October 2021; pp. 161–166. [CrossRef]
39. Heß, V.L.; Damásio, B. Machine learning in banking risk management: Mapping a decade of evolution. *Int. J. Inf. Manag. Data Insights* **2025**, *5*, 100324. [CrossRef]
40. Guerra, P.; Castelli, M. Machine learning applied to banking supervision: A literature review. *Risks* **2021**, *9*, 136. [CrossRef]
41. Herrmann, H.; Masawi, B. Three and a half decades of artificial intelligence in banking, financial services, and insurance: A systematic evolutionary review. *Strateg. Change* **2022**, *31*, 549–569. [CrossRef]
42. Kou, G.; Chao, X.; Peng, Y.; Alsaadi, F.E.; Herrera-Viedma, E. Machine learning methods for systemic risk analysis in financial sectors. *Technol. Econ. Dev. Econ.* **2019**, *25*, 716–742. [CrossRef]
43. Leo, M.; Sharma, S.; Maddulety, K. Machine Learning in Banking Risk Management: A Literature Review. *Risks* **2019**, *7*, 29. [CrossRef]
44. Lagasio, V.; Pampurini, F.; Pezzola, A.; Quaranta, A.G. Assessing bank default determinants via machine learning. *Inf. Sci.* **2022**, *618*, 87–97. [CrossRef]
45. Adamyk, B.; Skirka, A.; Snihur, K.; Adamyk, O. Analysis of trust in Ukrainian banks based on machine learning algorithms. In Proceedings of the 9th International Conference on Advanced Computer Information Technologies (ACIT), Ceske Budejovice, Czech Republic, 5–7 June 2019; pp. 234–239. [CrossRef]
46. Beutel, J.; List, S.; von Schweinitz, G. Does machine learning help us predict banking crises? *J. Financ. Stab.* **2019**, *45*, 100693. [CrossRef]
47. Alessa, N.; Majdua, A.; Alshehri, S.; Alhawiti, M.; Aljohani, R.; Alhakamy, A. Banking in terms of deposit prediction based on machine learning and big data analytics. In Proceedings of the 11th International Conference on Control, Mechatronics and Automation (ICCMA), Grimstad, Norway, 1–3 November 2023; IEEE: New York, NY, USA, 2023; pp. 69–74. [CrossRef]
48. Meitei, A.J.; Arora, P.; Mohapatra, B.B.; Arora, H. Identification of Weak Banks Using Machine Learning Techniques: Evidence from the Indian Banking Sector. *Glob. Bus. Rev.* **2022**, 09721509221113631. [CrossRef]
49. Martínez, R.G.; Román, M.P.; Casado, P.P. Big data algorithmic trading systems based on investors' mood. *J. Behav. Financ.* **2019**, *20*, 227–238. [CrossRef]
50. Houlihan, P.; Creamer, G.G. Leveraging social media to predict continuation and reversal in asset prices. *Comput. Econ.* **2021**, *57*, 433–453. [CrossRef]
51. Kokina, J.; Gilleran, R.; Blanchette, S.; Stoddard, D. Accountant as digital innovator: Roles and competencies in the age of automation. *Account. Horiz.* **2020**, *35*, 153–184. [CrossRef]
52. Chan, T.L.; Hale, N. Pricing European-type, early-exercise and discrete barrier options using an algorithm for the convolution of Legendre series. *Quant. Financ.* **2020**, *20*, 1307–1324.
53. Teng, H.-W.; Lee, M. Estimation procedures of using five alternative machine learning methods for predicting credit card default. *Rev. Pac. Basin Financ. Mark. Policies* **2019**, *22*, 1950021. [CrossRef]
54. Bee, M.; Hambuckers, J.; Trapin, L. Estimating large losses in insurance analytics and operational risk using the g-and-h distribution. *Quant. Financ.* **2021**, *21*, 1207–1221. [CrossRef]
55. Gao, B. The use of machine learning combined with data mining technology in financial risk prevention. *Comput. Econ.* **2021**, *59*, 1385–1405. [CrossRef]
56. Li, Q.; Xu, Z.; Shen, X.; Zhong, J. Predicting business risks of commercial banks based on BP-GA optimized model. *Comput. Econ.* **2021**, *59*, 1423–1441. [CrossRef]
57. Chen, Y.; Xie, Z.; Zhang, W.; Xing, R.; Li, Q. Quantifying the effect of real estate news on Chinese stock movements. *Emerg. Mark. Financ. Trade* **2020**, *57*, 4185–4210.
58. Omarova, S.T. New tech v. new deal: Fintech as a systemic phenomenon. *Yale J. Regul.* **2019**, *36*, 735–793.
59. de Prado, J.W.; de Castro Alcântara, V.; de Melo Carvalho, F.; Vieira, K.C.; Machado, L.K.C.; Tonelli, D.F. Multivariate analysis of credit risk and bankruptcy research data: A bibliometric study involving different knowledge fields (1968–2014). *Scientometrics* **2016**, *106*, 1007–1029. [CrossRef]
60. West, J.; Bhattacharya, M. Intelligent financial fraud detection: A comprehensive review. *Comput. Secur.* **2016**, *57*, 47–66. [CrossRef]
61. Königstorfer, F.; Thalmann, S. Applications of artificial intelligence in commercial bank—A research agenda for behavioral finance. *J. Behav. Exp. Financ.* **2020**, *27*, 100352. [CrossRef]

62. Ciampi, F.; Giannozzi, A.; Marzi, G.; Altman, E.I. Rethinking SME default prediction: A systematic literature review and future perspectives. *Scientometrics* **2021**, *126*, 2141–2188. [CrossRef]
63. Milojević, N.; Redzepagic, S. Prospects of Artificial Intelligence and Machine Learning Application in Banking Risk Management. *J. Cent. Bank. Theory Pract.* **2021**, *10*, 41–57. [CrossRef]
64. Mytnyk, B.; Tkachyk, O.; Shakhovska, N.; Fedushko, S.; Syerov, Y. Application of Artificial Intelligence for Fraudulent Banking Operations Recognition. *Big Data Cogn. Comput.* **2023**, *7*, 93. [CrossRef]
65. Hasan, M.; Hoque, A.; Abedin, M.Z.; Gasbarro, D. FinTech and sustainable development: A systematic thematic analysis using human- and machine-generated processing. *Int. Rev. Financ. Anal.* **2024**, *95*, 103473. [CrossRef]
66. Kozar, Ł.J.; Paduszyńska, M. FinTech and sustainable development: Study focused on identifying green research areas. *Procedia Comput. Sci.* **2024**, *246*, 1762–1769. [CrossRef]
67. Li, Z.; Chen, P. Sustainable finance meets FinTech: Amplifying green credit's benefits for banks. *Sustainability* **2024**, *16*, 7901. [CrossRef]
68. Yuan, X. Integrating FinTech, CSR, and green finance: Impacts on financial and environmental performance in China. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 1072. [CrossRef]
69. Yao, J.; Yang, C. Financial technology and climate risks in the financial market. *Int. Rev. Financ. Anal.* **2025**, *99*, 103920. [CrossRef]
70. Tian, Y.; Wen, H.; Guo, K. Machine learning applications in climate finance: An overview. *Res. Int. Bus. Financ.* **2025**, *79*, 103063. [CrossRef]
71. Christodoulou, P.; Psillaki, M.; Sklias, G.; Chatzichristofis, S.A. A blockchain-based framework for effective monitoring of EU green bonds. *Financ. Res. Lett.* **2023**, *58*, 104397. [CrossRef]
72. Manu, F.M.; Rajendran, N. Sustainable banking in the digital age: Insights from FinTech, blockchain, and green finance research. *Int. J. Res. Commer. Manag. Stud.* **2025**, *7*, 529–558. [CrossRef]
73. Rusydiana, A.; Slamet, A.; As-Salafiyah, A.; Djamaluddin, Y.S.; Marlina, L. Fiqh on Finance: A Scientometric Analysis using Bibliometrix. *Libr. Philos. Pract.* **2021**, 5436. Available online: https://digitalcommons.unl.edu/libphilprac/5436 (accessed on 12 September 2025).
74. Aria, M.; Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [CrossRef]
75. Fabrigar, L.R.; Wegener, D.T.; MacCallum, R.C.; Strahan, E.J. Evaluating the use of exploratory factor analysis in psychological research. *Psychol. Methods* **2020**, *25*, 375–389. [CrossRef]
76. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*, 8th ed.; Cengage Learning: Boston, MA, USA, 2019.
77. Bornmann, L.; Leydesdorff, L. Scientometrics in a changing research landscape: Bibliometrics has become an integral part of research quality evaluation and has been changing the practice of research. *EMBO Rep.* **2014**, *15*, 1228–1232. [CrossRef]
78. Van Eck, N.J.; Waltman, L. Software Survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef]
79. Kaiser, H.F. An index of factorial simplicity. *Psychometrika* **1974**, *39*, 31–36. [CrossRef]
80. Bartlett, M.S. Tests of significance in factor analysis. *Br. J. Psychol.* **1950**, *3*, 77–85. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.