Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 4 Issue 2 (2024)

A Bibliometric Analysis of Corporate Bankruptcy Prediction Research: Trends and Insights

Sakshi Sharma[1], Prof. Bhag Singh Bodla[2]
[1]University School of Management, Kurukshetra University, Kurukshetra, sakshi.usm@kuk.ac.in
[2]Kurukshetra University, Kurukshetra, bsbkuk@gmail.com

Abstract
This paper is the outcome of a bibliometric study aimed at providing a worldwide overview of the prediction of corporate bankruptcy. This research paper provides a thorough examination of bibliometric analysis in the field of corporate bankruptcy prediction, specifically emphasising Artificial Intelligence-based models. This study uses rigorous bibliometric methods to identify and analyse the journals with the highest productivity, the papers with the most citations, and the authors with the most citations in this particular field. In addition, the research environment is illuminated by descriptive analysis of authors, countries, journals, and keywords. The study also examines the expansion and patterns in publications, offering significant insights for identifying periods of heightened interest and possible domains for additional investigation. In addition, this study examines country-specific annotations and analyzes individual and country-specific co-authorship to reveal collaboration networks and research contributions in various locations. Co-citation analysis and co-occurrence analysis of journals provide a more profound understanding of how knowledge is spread and the formation of topic clusters within the field. The results of this study have important consequences for research institutions and politicians, providing valuable insights that can lead to future research and influence policy decisions in the field of corporate bankruptcy prediction. It will guide future research and advance the discourse on financial stability and risk management**.**

Keywords: *Corporate Bankruptcy, Artificial Intelligence, financial stability, Risk management*

1 Introduction

When companies are unable to fulfil their financial obligations, they are deemed bankrupt, and this legal status is conferred upon them by their creditors. (Aziz & Dar, 2006) Bankruptcy filing allows corporations and banks to seek forgiveness for unpayable debts. The act of monitoring and documenting administrative errors, deviations from organizational policies, and the general state of well-being often culminate in the initiation of bankruptcy proceedings. The financial crisis has led to a rise in the occurrence of bankruptcies. The occurrence of financial failures can have significant repercussions on several stakeholders, including partners, employees, investors, and society at large, resulting in substantial financial burdens (Altman, 1994). The significance and use of forecasting an organization's financial health are readily apparent. The occurrence of bankruptcy can be mitigated through the provision of timely notification to partners. This elucidates the reason behind the significant amount of attention that the point has garnered in the scholarly literature. (Muslim & Dasril, 2021) The subject of bankruptcy prediction has been the subject of intensive research for a significant duration. The examination of financial ratios, statistics, and attributes is conducted to ascertain the financial health of a company, determining whether it is thriving or susceptible to insolvency (Guha & Veeranjaneyulu, 2019).

(H. Kim et al., 2022) Predicting corporate bankruptcy is a crucial area of study in the field of finance and business analysis. The ability to forecast the financial health of companies and anticipate the risk of bankruptcy has far-reaching implications for investors, creditors, and other stakeholders. By identifying early warning signs and potential indicators of financial distress, predictive models play a vital role in informed decision-making and risk management. (Cao et al., 2022) Corporate bankruptcy signifies a state of financial insolvency where a company is unable to meet its debt obligations. This event can have profound effects on employees, shareholders, suppliers, and the broader economy. Therefore, developing accurate and reliable methods for predicting bankruptcy is essential for mitigating negative consequences and enabling proactive measures. Traditionally, (Soui et al., 2020) bankruptcy prediction relied on financial ratios and accounting-based indicators. However, with the advent of advanced data analytics and machine learning techniques, the landscape of predictive modelling has evolved. These modern methods allow researchers and analysts to process large datasets, uncover intricate patterns, and capture subtle signals that may indicate an impending financial crisis. While the prediction of corporate bankruptcy is a universal challenge, it is influenced by diverse economic, regulatory, and industry-specific factors. Consequently, (H. Kim et al., 2020) focused on developing a comprehensive understanding of bankruptcy prediction within a specific context, aiming to contribute valuable insights to the existing body of knowledge. By concentrating on this critical facet of financial analysis, we seek to enhance our understanding of the dynamics that lead to corporate bankruptcy and the tools available to anticipate

and manage such situations. (Horak et al., 2020) Contemporary economies are constructed upon the fundamental principle of bankruptcy. The concept of limited liability in companies emerged during the 17th century in England as a response to bankruptcy, allowing firms to bear responsibility for their shares while minimizing personal risk. The transition in ownership and management gave rise to fresh concerns, including the divergence of interests between management and equity owners, giving rise to a distinct set of challenges. Consequently, the presence of symmetric information might lead to management and stockholders making unwise judgments. There is a potential risk of investors being misled if they heavily depend on management information. The narrowing of the knowledge gap occurs through the provision of trustworthy models to investors, enabling them to assess the degree of risk.

(A. G. Kim & Yoon, 2021) The concepts of bankruptcy and credit checks have been established for a significant duration. During the 1890s, a prominent method of probability computation emerged, which presented compelling evidence. (Ben Jabeur et al., 2023) Private Banks mostly utilized assessments of enterprises' financial soundness to determine whether to lend them money. This practice had a significant role in the dissemination of ratio analysis. At this juncture, there was an increasing quantity of credit professionals within this structure. The initiation of a discourse on credit risk occurred in 1919 with the release of the Federal Reserve's inaugural report on the analysis of government bank ratios. The prediction of bankruptcy utilizing contemporary methodologies was suggested by the authors (Zelenkov & Volodarskiy, 2021). (du Jardin et al., 2019) conducted research on a total of 706 organizations for five years. Each year, the sample was systematically chosen to deliberately exclude specific segments, resulting in an about equal division between these two groups. The 30 factors that were chosen were categorized into five separate subgroups, each with its unique characteristics. (Vochozka et al., 2020) Various subgroups exhibited heterogeneous financial structures, encompassing factors such as income levels, the prevalence of financially troubled organizations, and the constraints associated with accommodating these proportions. The previously mentioned areas, including net gain, turnover, and analysis thresholds, have now been expanded to encompass more comprehensive standards. Four recommendations were provided in response to his utilization of this methodology for ratio identification.

The primary aim of this study is to perform a detailed bibliometric analysis that comprehensively explores the diverse landscape of research on predicting corporate bankruptcy. The main objective of our study is to conduct a systematic investigation and provide a comprehensive understanding of the historical growth trajectories, authorship dynamics, citation networks, research characteristics, geographical patterns, and evolving techniques within this particular field. This comprehensive analysis aims to offer a detailed and comprehensive understanding of the current landscape of corporate bankruptcy prediction research. Our objective is to identify key trends, influential contributors, and emerging research clusters, as well as to uncover the intricate web of interdisciplinary connections and regional nuances. Moreover, the primary objective of this study is to function as an essential reference for researchers, policymakers, and industry professionals. It aims to provide in-depth insights that can guide future research priorities, foster innovative methodologies, and contribute to the ongoing discussion on financial stability, risk management, and corporate sustainability. This study aims to address this gap by providing a holistic understanding of the subject's development in the Indian corporate landscape, offering valuable insights into its growth dynamics, visibility, and collaborative networks among researchers and institutions.

2 Literature Review

This literature review explores the field of corporate bankruptcy prediction, with a specific focus on employing data analytics. It examines the historical evolution of bankruptcy prediction models, discusses key influencing factors, evaluates methodologies, and sets the stage for a comprehensive bibliometric analysis of research in this critical area.

Using a bibliometric analytic method, Shi & Li (2019) seek to review the body of literature on big data applications in banking. The publications were chosen from 2012 to 2020 and arranged according to the number of citations in the analysis and outcomes. Nobanee et al. (2021) state that bibliometric analysis is a useful technique for conducting a quantitative examination of academic output and addressing current trends in a particular field of study. To evaluate its progress and characterize the research trend over the past five decades, this work intends to explore the use of intelligent strategies in bankruptcy forecasts.

Halder (2016) analyzed the literature on machine learning (ML) and artificial intelligence (AI) in the finance sector. The author gathered 348 articles from journals indexed in the Scopus database that were published in the years 2011 through 2021 using a bibliometric technique. The data was analyzed using many programs (RStudio, VOSviewer, and Excel), and the most active actors in terms of nations, institutions, sources, documents, and authors were shown. (Almeida, 2023) To identify academic domains that have made substantial research contributions, we summarized 711 bibliometric publications

on AI & its sub-sets and BDA that were published in various fields. From the Scopus Q1 and Q2 journal databases, we selected bibliometric review articles from publications between 2012 and 2022. The Scopus database produced 711 publications from 59 nations and periodicals across various fields, with an average of 17.9 citations every year. (Fakhar et al., 2023) Interfered the subject organization of AI and ML research in finance from 1986 to April 2021 using co-citation and bibliometric-coupling studies. the author further identified three overarching categories of finance scholarship that are roughly identical for both types of analysis by identifying nine (co-citation) and eight (bibliometric coupling) specialized clusters of finance that employ AI and ML. Zapata & Mukhopadhyay (2022) examined the predictive performance of 16 classifiers based on a large sample of US corporate bankruptcies, ranging from the most limited classifiers (such as logit, probit, and linear discriminant analysis) to more sophisticated methods like neural networks, support vector machines (SVMs), and 'new age' statistical learning models like generalized boosting, AdaBoost, and random forests.

To choose the most acceptable cases of bankruptcy, Alam et al. (2021) have recommended an optimization technique of cluster-based sampling utilizing the fuzzy c-means algorithm to the classifier. To address the issue of the current data imbalance, this method was investigated with the assistance of a clustering method and a GA-based artificial neural network. Grosu et al. (2023) compared and contrasted statistical, rough computing, and hybridized methodologies. The Greek industrial bank ETEVA's financial bankruptcy data set is used for the comparison analysis. In conclusion, hybridized computing approaches still offer greater accuracy 94.1% than rough computing techniques, whereas rough computing techniques still offer better accuracy 88.2% than statistical techniques. Mirzakhanyan (2005) provided a bibliometric analysis of various forms of risk and bankruptcy, intending to contribute to academic understanding in this field. The author employed bibliometric tools in R and VOS viewer, according to the key bibliometric laws. Between 1995 and 2023, the author examined 7163 relevant scholarly papers extracted from the WOS database. Li & Faff (2019) consisted of bibliometric and theme modelling studies of corporate bankruptcy. It demonstrates that despite possessing 55-year-old literature, research on corporate insolvency has shown an exponential increase in the recent decade. In this study, the authors have examined 1254 papers to uncover knowledge structures and the advancement of bankruptcy models from 1968 to 2022.

The review of the literature indicates that a notable research gap exists in the realm of predicting corporate bankruptcy in the Indian context. Despite a lot of literature on this topic globally, there is a distinct absence of bibliometric analyses specifically examining the trends, citation impact, and research networks within the Indian context.

3 Research Methodology & Descriptive data analysis

3.1 Research Methodology

This study evaluates the research conducted for predicting corporate bankruptcy in selected Indian companies by extracting data from the Web of Science and Scopus databases. The advanced search was conducted using the terms ‘Bankruptcy prediction’, ‘Data analytics’ and ‘Machine learning’ for 2010 to 2023 and in English (Languages).

Data Collection:

The search strategy yielded 346 documents. Out of a total of 346 documents, 228 have been indexed in the Web of Science database, while 118 are listed in Scopus. In this paper, we identify scholarly databases, repositories, and other sources that provide open access to predict the corporate bankruptcy of selected Indian companies. Access these sources to retrieve metadata and full-text articles authored by Indian researchers. For the analysis, data was gathered for a specific time frame i.e., 2010-2023.

Utilized VOS viewer: A visualization tool developed by Leiden University's Centre for Science and Technology Studies, to generate overlay visualizations that depict bibliographic coupling among documents, co-authorship between organizations, and co-authorship among authors.

Framework for Anatomizing the Data

The analysis in this article was performed through the most relevant components, that is author analysis, source analysis, country analysis and theme analysis. At first, description analysis was performed from 2010 to 2023 on the corpus of 346 research documents by taking the top 10 authors, sources, countries and themes using MS Excel. Network clustering and visualization were executed using the bibliometric software ‘VOS Viewer 1.6.18. Figure 1 presents the research framework which contains extraction and analysis phases for bibliometric visualization.

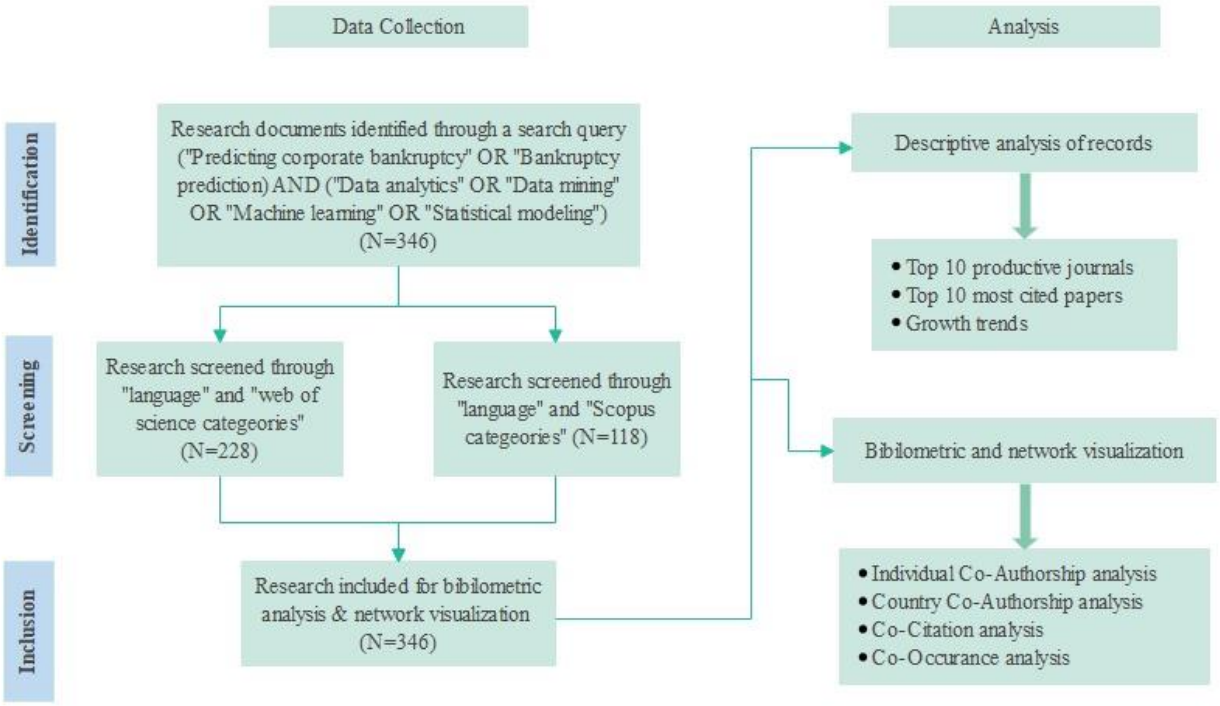


Figure 1 Research Framework

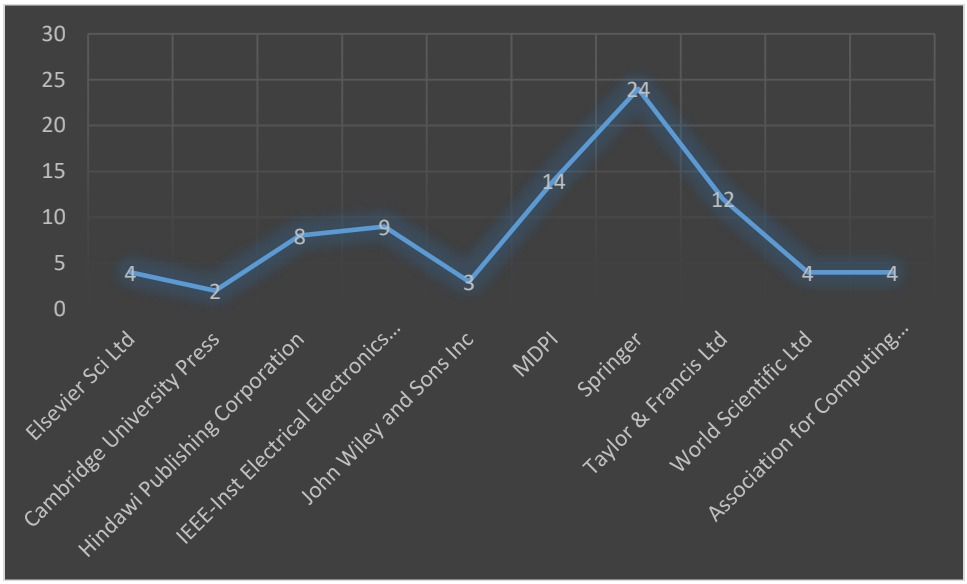
3.2 Top 10 most productive journals and total citations

The below table presents a snapshot of publication and citation metrics for ten journals based on their publication and citation statistics, Springer emerges as the front-runner, securing the top position with a substantial 24 publications and an impressive 1637 citations. This achievement underscores its significant scholarly influence and widespread recognition in the academic community. Following closely is IEEE-Inst Electrical Electronics Engineers Inc., taking the second spot with nine publications and an impressive 726 citations, a testament to its substantial impact in the field of electrical and electronics engineering. Taylor & Francis Ltd claims the third position with 12 publications and 785 citations, highlighting its considerable scholarly reach and influence. Meanwhile, MDPI secures the fourth spot with 14 publications and 583 citations, indicating prolific contributions to the realm of research. Hindawi Publishing Corporation, despite a comparatively lower publication count of eight, ranks fifth with 311 citations, showcasing commendable impact relative to its output. John Wiley and Sons Inc., Elsevier Sci Ltd, World Scientific Ltd, and the Association for Computing Machinery also make valuable contributions, each with their distinctive scholarly footprint. This ranking serves as a valuable resource for researchers seeking journals with substantial publication records and citations in their respective fields.

Table 1. Top 10 most productive journals and total citations

S. no	Name of the Journal	Number of publications	Total Citations
1	Elsevier Sci Ltd	4	250
2	Cambridge University Press	2	45
3	Hindawi Publishing Corporation	8	311
4	IEEE-Inst Electrical Electronics Engineers Inc	9	726
5	John Wiley and Sons Inc	3	36
6	MDPI	14	583
7	Springer	24	1637

8	Taylor & Francis Ltd	12	785
9	World Scientific Ltd	4	171
10	Association for Computing Machinery	4	568



3.3 Top 10 most-cited papers on intelligent techniques-based models

Table 2 presents a collection of research works authored by various scholars in the field of finance, data mining, and machine learning, each contributing unique perspectives to the domain of financial prediction and decision support systems. Barboza, Kimura, and Altman's 2017 study on "Machine learning models and bankruptcy prediction" delves into the application of machine learning techniques to predict bankruptcy, offering valuable insights into risk assessment. Amani and Fadlalla's 2017 work provides a comprehensive review of data mining applications in accounting, organizing the vast literature in this area. Kuiziniene, Krilavicius, Damasevicius, and Maskeliunas (2022) contribute a systematic review that investigates the utilization of artificial intelligence methods for financial distress identification, reflecting the ongoing evolution of predictive technologies in finance.

Lin, Hu, and Tsai's 2012 survey on "Machine learning in financial crisis prediction" serves as a valuable resource for understanding the role of machine learning in crisis anticipation. Zhang, Liu, Heidari, Wang, Chen, and Wang (2021) introduce augmented kernel extreme learning models for bankruptcy prediction, enhancing algorithmic capabilities in this domain. Espejo, Ventura, and Herrera's 2010 survey focuses on the application of genetic programming to classification, expanding the horizons of machine learning techniques. Kim and Kang's 2010 work proposes an ensemble approach using neural networks for bankruptcy prediction, contributing to the development of predictive models. Doumpos and Zopounidis (2011) delve into preference disaggregation and statistical learning, offering insights into multicriteria decision support systems. Liang, Lu, Tsai, and Shih's 2016 study comprehensively investigates the use of financial ratios and corporate governance indicators in bankruptcy prediction. Finally, Lin, Hu, and Tsai (2012) present another survey on "Machine Learning in Financial Crisis Prediction," reaffirming the significance of this topic in the financial domain. These works collectively contribute to the evolving landscape of financial prediction, decision support, and risk assessment, offering valuable knowledge and methodologies for researchers and practitioners alike.

Table 2. Top 10 most-cited papers on intelligent techniques-based models

S.no	Author Name	Title	Year	Citation
1	Barboza F.; Kimura H.; Altman E.	Machine learning models and bankruptcy prediction	2017	374
2	Amani, FA; Fadlalla, AM	Data mining applications in accounting: A review of the literature and organizing framework	2017	262
3	Kuiziniene, D; Krilavicius, T; Damasevicius, R; Maskeliunas, R	Systematic Review of Financial Distress Identification Using Artificial Intelligence Methods	2022	218
4	Lin W.-Y.; Hu Y.-H.; Tsai C.-F.	Machine learning in financial crisis prediction: A survey	2012	216
5	Zhang Y.; Liu R.; Heidari A.A.; Wang X.; Chen Y.; Wang M.; Chen H.	Towards augmented kernel extreme learning models for bankruptcy prediction: Algorithmic behaviour and comprehensive analysis	2021	187
6	Espejo, PG; Ventura, S; Herrera, F	A Survey on the Application of Genetic Programming to Classification	2010	182
7	Kim M.-J.; Kang D.-K.	Ensemble with neural networks for bankruptcy prediction	2010	181
8	Doumpos, M; Zopounidis, C	Preference disaggregation and statistical learning for multi-criteria decision support: A review	2011	175
9	Liang D.; Lu C.-C.; Tsai C.-F.; Shih G.-A.	Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study	2016	173
10	Lin, WY; Hu, YH; Tsai, CF	Machine Learning in Financial Crisis Prediction: A Survey	2012	171

4 Bibliometric and network analysis results and findings

Publication communication has witnessed remarkable growth in the field of corporate bankruptcy research, providing researchers with unrestricted access to scientific literature and promoting collaboration in addressing the global environmental crisis (Morrison, H., Borges, L., Zhao, X., Kakou, T. L., & Shanbhoug, 2022). The growth trends in scholarly communication on corporate bankruptcy highlight the increasing availability and impact of open-access research in this critical area. One of the significant growth trends in scholarly communication on predicting corporate bankruptcy is the rising number of publications. As awareness and concern for bankruptcy have grown, so has the production of open-access research in this field. Researchers and institutions are recognizing the importance of making predicting corporate bankruptcy research freely accessible to maximize its reach and impact. The increasing number of publications serves as evidence of the expanding knowledge base and the commitment of the scientific community to open-access principles (Burright, 2006).

4.1 Descriptive Analysis of Authors, Countries, Journals and Keywords

Numerous quantitative data on document trends, citations, authors, collaboration, and keywords are presented in a descriptive analysis. The document analysis is shown in Table 3.

Among 346 research documents, the total 346 documents are multi-authored documents published in 10 journals. These are published by 5112 authors where 5112 authors published the articles with mutual collaboration.

Table 3 Descriptive Analysis of Documents Related to Research Theme

Description	Results

Total documents	346
Period	2010-2023
Web of Science database	265
Scopus database	118
Authors	1254
Number of journals	10
Total Number of publications	84
Total Citations	5112
Documents per author	3
Authors per document	5
Co-authors per documents	658

Table 3 provides an overview of research results encompassing a period from 2010 to 2023, drawing from a total of 383 documents sourced from both the Web of Science and Scopus databases. A diverse group of 1,254 authors contributed to these documents, resulting in an average of 3 documents per author. Collaboration was a key theme, with an average of 5 authors per document, highlighting the interdisciplinary nature of the research. The extensive network of collaboration is evident in the remarkable figure of 658 co-authors per document. This comprehensive research effort spanned across 10 different journals, resulting in a total of 84 publications. The collective impact of these publications is underscored by a total of 5,112 citations, emphasizing their significance within the academic community. This table provides valuable insights into the scale and impact of a multi-faceted research endeavour over the specified timeframe.

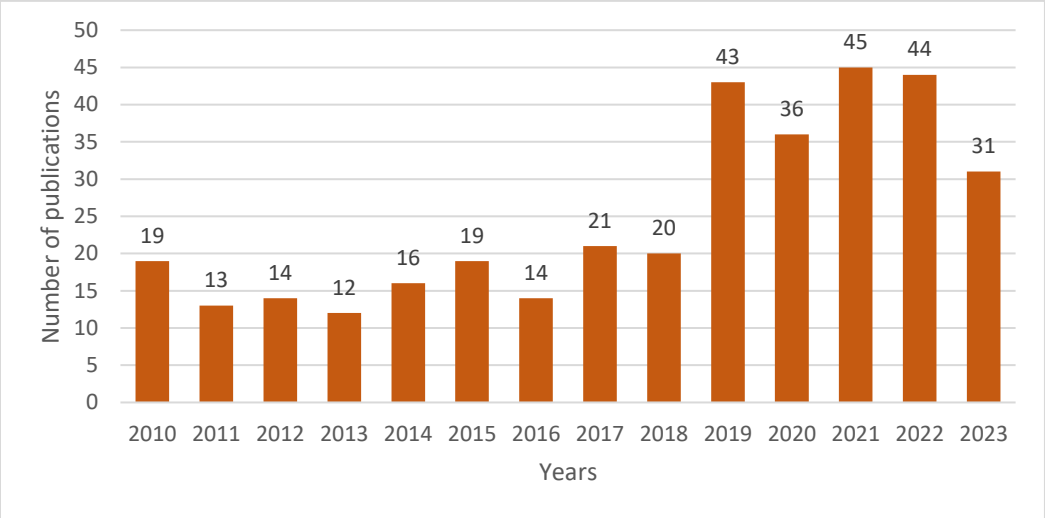
Table 4: Growth Trends in publications on bankruptcy

Year	No. of Publication	% of Pub. (out of Total publications)	Citation References	Citation Impact
2010	19	4.97%	200	0.095
2011	13	3.40%	136	0.09558824
2012	14	3.66%	370	0.03783784
2013	12	3.14%	36	0.33333333
2014	16	4.18%	385	0.04155844
2015	19	4.97%	180	0.10555556
2016	14	3.66%	260	0.05384615
2017	21	5.49%	750	0.028
2018	20	5.23%	166	0.12048193
2019	43	11.25%	487	0.08829569
2020	36	9.42%	251	0.14342629
2021	45	11.78%	338	0.13313609
2022	44	11.51%	36	1.22222222
2023	31	8.11%	48	0.64583333

This table offers a comprehensive overview of scholarly publications from 2010 to 2023, detailing the number of publications, the percentage they constitute out of the total publications, citation references, and their resulting citation impact. Over this period, the number of publications fluctuated, with the highest count in 2021 at 45 and the lowest in 2013 at 12. The percentage of these publications concerning the total varies accordingly, reaching a peak in 2021 at 11.78%.

Citation references also vary significantly, with a notable peak in 2017 at 750 and a low point in 2013 at 36. The citation impact demonstrates fluctuations as well, with the highest impact in 2022 at 1.22 and the lowest in 2013 at 0.33. This data allows for the observation of trends in scholarly output, the relative contribution of these publications to the total, their

citation patterns, and their impact on the academic community, providing valuable insights for researchers and institutions alike.



4.2 Country-wise Annotation

Table 5. Top 10 Countries Based on Total Citations.

COUNTRY	NO. of citations	% of TC
China	**250**	**4.89**
Canada	**45**	**0.88**
Spain	**311**	**6.08**
USA	**726**	**14.2**
Austria	**36**	**0.7**
Italy	**583**	**11.4**
South Korea	**1637**	**32.022**
South Africa	**785**	**15.35**
INDIA	**171**	**3.34**
Saudi Arabia	**568**	**11.11**
Total	**5112**	**100**

Table 5 presents a breakdown of the total number of citations (NO. of citations) and the corresponding percentage of the total citations (% of TC) for a selection of countries within a research context, contributing to a cumulative total of 5,112 citations. Notably, South Korea stands out with a substantial 32.02% of the total citations, reflecting a significant influence in the field. The United States also makes a substantial impact, with 14.2% of the total citations. Italy and South Africa follow closely with 11.4% and 15.35%, respectively in Figure 2. China, Saudi Arabia, and Spain each contribute a notable share of the citations, underscoring their significance in this research landscape. Furthermore, Canada, Austria, India, and Estonia play more modest yet still important roles in the overall citation distribution. This table offers valuable insights into the global reach and influence of research, showcasing the varying contributions of these nations to the body of knowledge in the field.

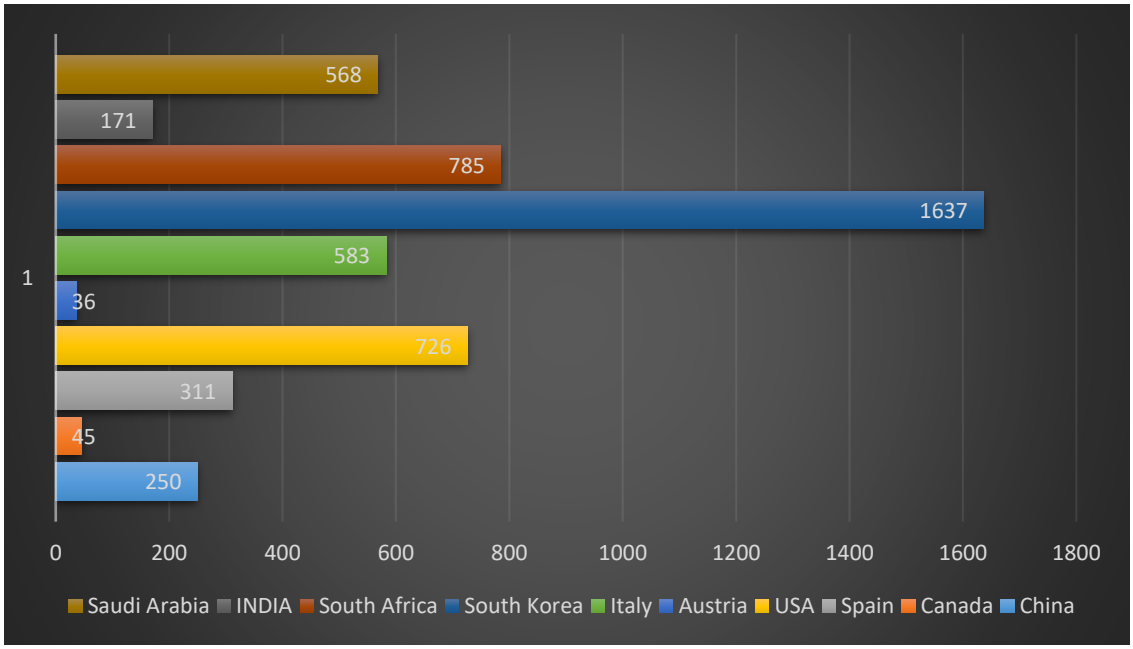


Figure 2 Top 10 Countries Based on Total Publication Production.

4.3 Individual and Country Co-Authorship Analysis

Individual and country co-authorship analysis, further enriched by the integration of powerful tools like VOS Viewer, Scopus, and Web of Science, offers a comprehensive and insightful approach to dissecting collaborative research endeavours. VOS Viewer, known for its advanced visualization capabilities, empowers researchers to explore intricate networks of co-authored publications, providing dynamic visual representations of academic collaboration patterns. This visual clarity aids in deciphering the complex relationships between individual authors, institutions, and research themes. Simultaneously, the inclusion of Scopus and Web of Science databases ensures a comprehensive analysis, incorporating a vast array of scholarly journals and research publications. By harnessing the data from these renowned platforms, scholars gain access to an extensive repository of academic work, facilitating a deeper understanding of global research networks and the influential role of countries within this ecosystem. In summary, the amalgamation of VOS Viewer, Scopus, and Web of Science into individual and country co-authorship analysis not only provides invaluable insights into collaborative research but also equips researchers and decision-makers with robust tools to visualize, interpret, and make informed decisions in academia, research policy, and resource allocation.

4.3.1 Individual Co-Authorship Analysis

Co-authorship analysis, conducted by examining a dataset of 221 authors using VOS Viewer in conjunction with the Web of Science database, offers a rich and nuanced perspective on collaborative research within a specific domain. The dataset of 221 authors represents a diverse array of contributors, each bringing their unique expertise and perspectives to the research landscape. Through the lens of VOS Viewer, researchers can delve deep into the intricacies of these collaborations, visualizing the connections and relationships between these 221 individuals. This analysis not only unveils the extent of co-authorship but also provides insights into the clustering of researchers, identification of prolific collaborators, and the formation of research communities within the field. By incorporating the comprehensive data from the Web of Science, this analysis ensures that the study encompasses a wide array of scholarly publications and research contributions, making it a valuable tool for understanding collaborative dynamics and knowledge dissemination within the academic and scientific community.

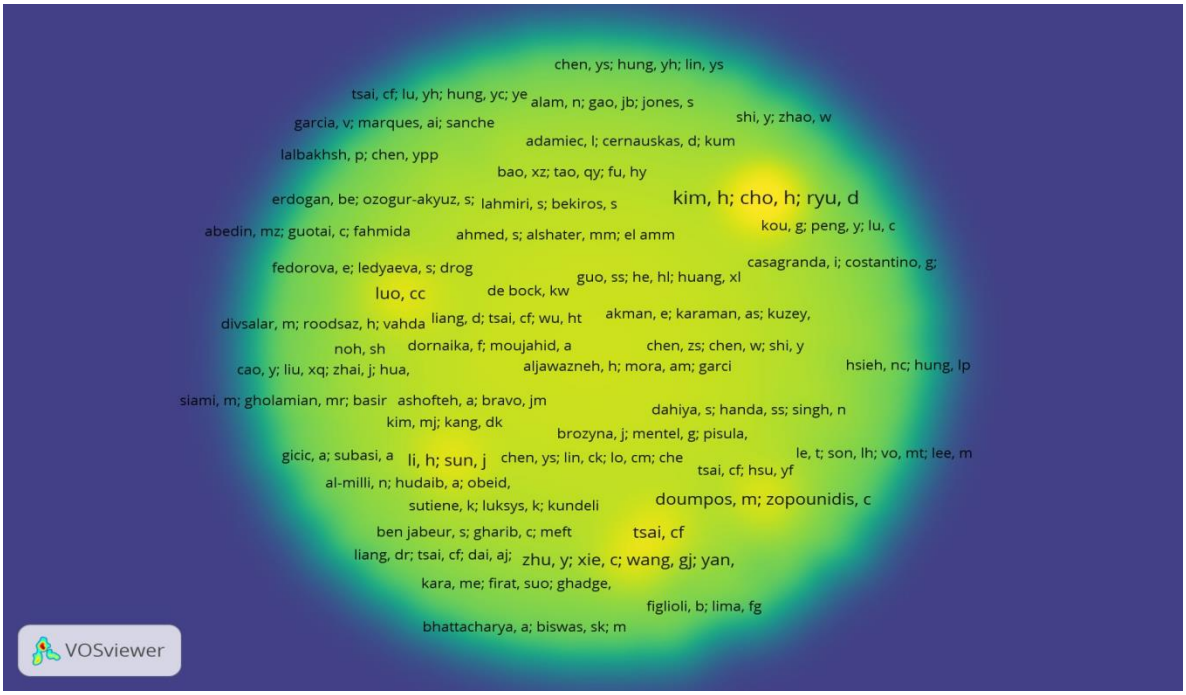


Figure 3 Individual co-authorship for the Web of Science

Co-authorship analysis, conducted through the examination of a dataset featuring 115 authors using VOS Viewer within the context of the Journal of Scopus, provides a comprehensive exploration of collaborative research dynamics within this specific academic journal. The dataset comprising 115 authors reflects a diverse and interconnected group of researchers who have collectively contributed to the scholarly discourse within the journal. By employing VOS Viewer, researchers gain the ability to visually represent and analyze the intricate web of collaborations among these authors. This analysis not only reveals the extent of co-authorship within the Journal of Scopus but also allows for the identification of research clusters, influential authors, and emerging trends within the journal's domain. Leveraging the extensive data available in the Scopus database ensures that this analysis encompasses a wide spectrum of scholarly publications, providing valuable insights into the collaborative dynamics and knowledge dissemination processes within the context of this particular journal.

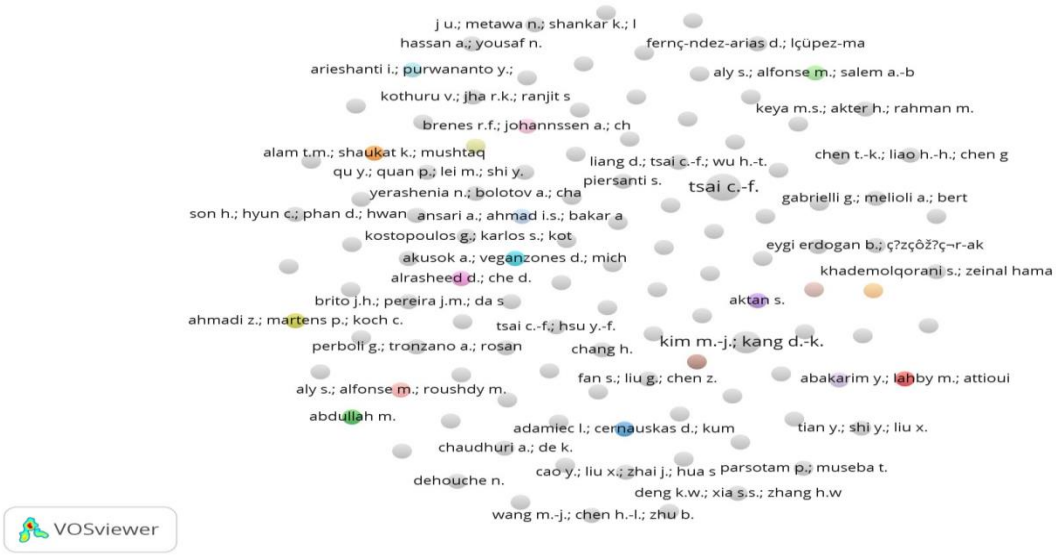


Figure 4 Individual co-authorship for Scopus

4.3.2. Country co-authorship analysis

Co-authorship analysis for Web of Science and Scopus publications, conducted on a country-wise basis, is a methodical examination of collaborative research endeavours within specific nations. This approach delves into academic papers and their authors to uncover intricate networks, interdisciplinary trends, and the global impact of a country's scientific contributions. By scrutinizing co-authorship patterns and affiliations, it provides valuable insights into the extent of international research collaboration, highlighting opportunities for fostering cross-border partnerships. Additionally, it offers a quantitative assessment of research productivity, showcasing the quantity and quality of a country's scholarly output and its role in the broader academic landscape. Furthermore, this analysis serves as a compass for policymakers and funding agencies, aiding in resource allocation, strategic decision-making, and the cultivation of a vibrant research ecosystem. Ultimately, co-authorship analysis in Web of Science and Scopus publications on a country-wise basis plays a pivotal role in enhancing a nation's global research presence, fostering innovation, and advancing the frontiers of knowledge.

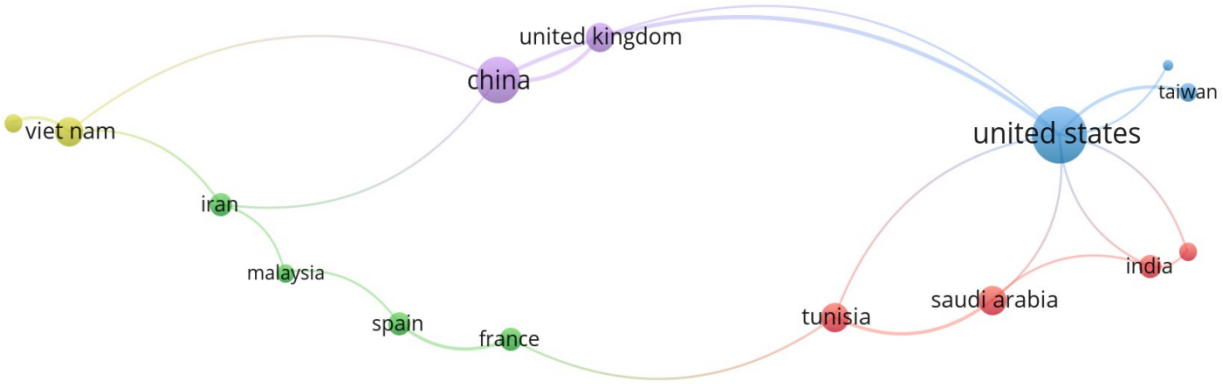


Figure 5 Co-Authorship Analysis for Scopus

Figure 6 Co-Authorship Analysis for Web of Science

4.4 Journal co-citation analysis

Table 6 presents a snapshot of the co-citation frequencies for a selection of academic journals across diverse subject areas. These frequencies signify the extent to which each journal is cited in conjunction with others, reflecting their prominence and influence within their respective fields. Notably, ELSEVIER's "Applied Soft Computing" stands out with an impressive co-citation frequency of 716, underscoring its pivotal role in the domain of soft computing. "Scientific Programming" by HINDAWI LTD also demonstrates its significance, boasting a co-citation frequency of 149. Meanwhile, journals such as "Applied Artificial Intelligence" by Taylor and Francis Inc. (with a frequency of 290) and "Computational Economics" by Springer New York LLC (with a frequency of 331) reaffirm their importance in their respective areas of study. These co-citation frequencies offer valuable insights into the academic landscape, highlighting the journals that play pivotal roles in shaping scholarly discourse and knowledge dissemination across various disciplines (Adakawa & Harinarayana, 2022) (Kumar, 2015).

Table 6. Top 15 Journals (According to Co-Citation Frequency)

Name of journal	Subject	Co-Citation frequency
ELSEVIER	APPLIED SOFT COMPUTING	716
HINDAWI LTD	SCIENTIFIC PROGRAMMING	149
Taylor and Francis Inc.	APPLIED ARTIFICIAL INTELLIGENCE	290
SPRINGER	MACHINE LEARNING	88
Oxford University Press	COMPUTER JOURNAL	62
SPRINGER	FINANCIAL INNOVATION	55
Springer New York LLC	COMPUTATIONAL ECONOMICS	331
PERGAMON-ELSEVIER SCIENCE LTD	COMPUTERS & OPERATIONS RESEARCH	155
WILEY	DECISION SCIENCES	46
ELSEVIER	DECISION SUPPORT SYSTEMS	374
TAYLOR & FRANCIS LTD	ENTERPRISE INFORMATION SYSTEMS	85
PERGAMON-ELSEVIER SCIENCE LTD	ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE	236
WILEY	EUROPEAN FINANCIAL MANAGEMENT	62
UNIV NAC COLOMBIA, FAC INGENIERIA	INGENIERIA E INVESTIGACION	118

4.5 Co-Occurrence Analysis

Co-occurrence analysis for Web of Science and Scopus publications, conducted on a country-wise basis, is a systematic approach to gaining profound insights into the thematic focus and research trends within specific nations. This method involves the meticulous examination of keywords or phrases used in academic papers to unveil prevalent subjects, emerging areas of interest, and the unique research landscape of each country (Björk, 2017). By quantifying the frequency and relevance of these keywords, researchers can identify the core topics driving scientific inquiry in a particular nation. Such analysis is instrumental in assessing a country's contributions to various fields, highlighting its strengths and areas of specialization, and pinpointing potential opportunities for collaboration and innovation (Esh & Ghosh, 2021). Policymakers, funding agencies, and institutions can leverage these findings to inform strategic decisions, allocate resources effectively, and promote research areas of national importance. Ultimately, keyword occurrence analysis in Web of Science and Scopus publications on a country-wise basis empowers nations to understand their research identity, foster interdisciplinary cooperation, and navigate the ever-evolving landscape of science and scholarship.

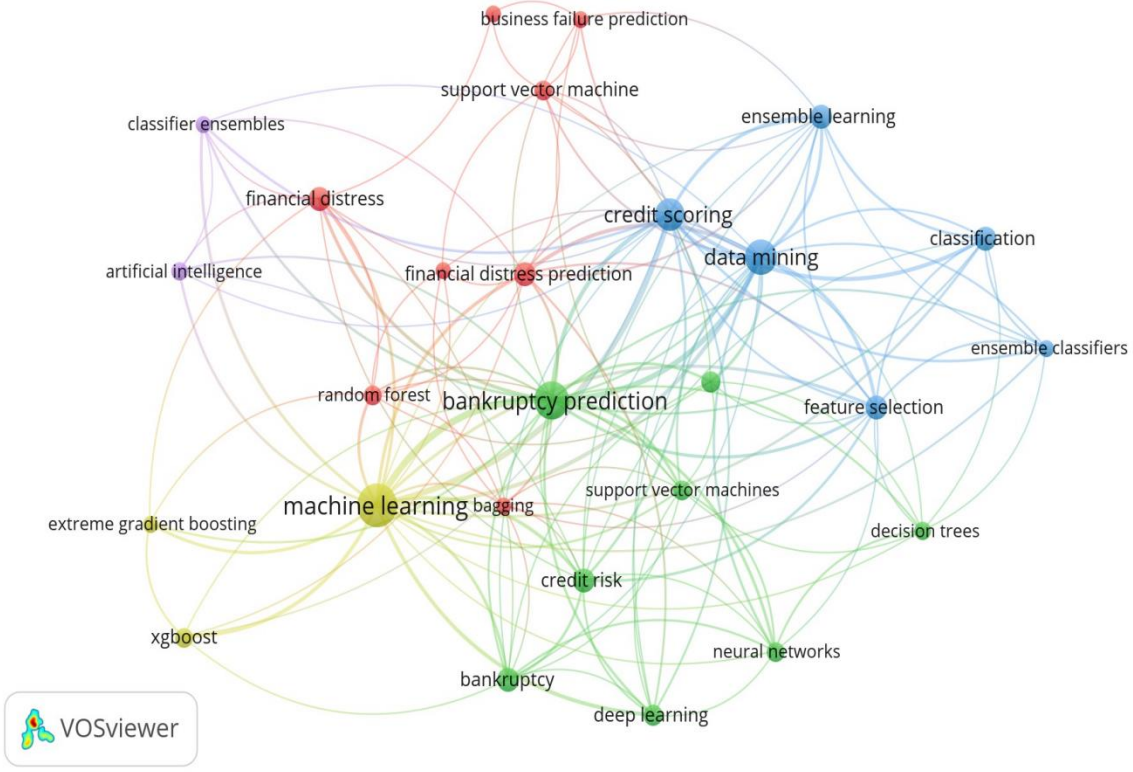


Figure 7 Co-occurrence for Scopus

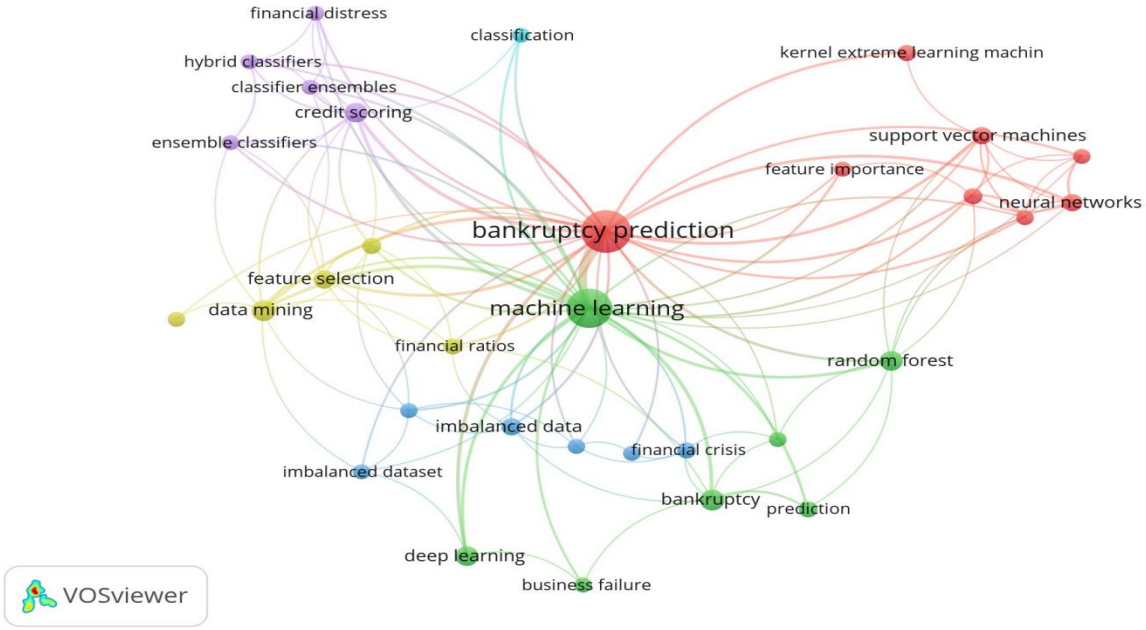


Figure 8 Co-occurrence for Web of Science

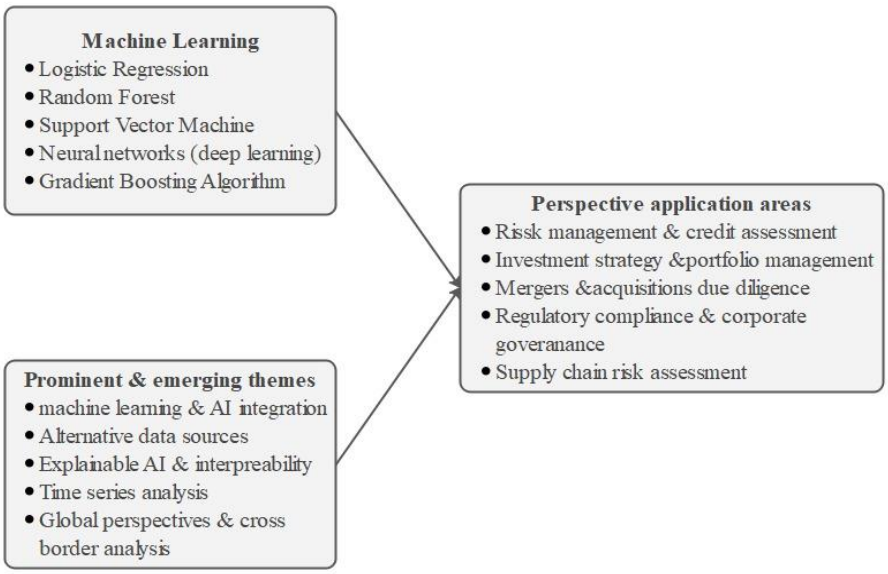


Figure 9 Framework for Emerging Research Themes

The above figure illustrates the conceptual framework for emerging research themes. The framework shows the concept of machine learning techniques, prominent & emerging themes, and perspective application areas, that is, risk management and credit assessment, investment strategy, mergers& acquisitions, regulatory compliance, and supply chain risk assessment. Future research studies can combine these themes to produce the most effective and advanced results for researchers and organisations.

5 Discussion and Conclusion

This study aimed to identify the research gap in predicting corporate bankruptcy among selected Indian companies by conducting a bibliometric analysis of scholarly publications. Key findings include growth trends in publications from 2010 to 2023, which showed a noticeable increase in interest in the topic since 2014, and the relative growth rate and doubling time of scholarly publications on the topic. The study also calculated the Co-Authorship Index (CAI) which provides insights into collaboration patterns among researchers. The analysis of authorship data provided insights into the level of research cooperation within the academic community, which are crucial indicators of collaboration trends and network dynamics. Overlay visualisations of bibliographic coupling with documents, co-authorship with organisations, and co-authorship with authors offered a comprehensive view of the scholarly landscape, helping identify clusters of related papers, research institutions actively contributing to the field, and collaborations among researchers. This approach facilitates knowledge discovery, interdisciplinary research exploration, and the identification of potential research partnerships.

The findings have several implications for researchers, institutions, and policymakers. Researchers can use the growth trends to identify periods of increased interest in the topic and potential areas for further exploration. The relative growth rate and doubling time metrics provide insights into the pace of research development, aiding in planning and resource allocation. The Co-Authorship Index sheds light on collaborative dynamics, fostering opportunities for interdisciplinary research and partnerships.

Limitations and future scope

Limitations of this bibliometric study include potential language bias, as it predominantly focuses on English-language publications, possibly excluding valuable research in other languages. Additionally, it may be subject to publication bias, relying solely on publicly available literature, and potentially overlooking unpublished or proprietary research. The findings are based on data up until a specific cutoff date, possibly missing out on recent developments in the field. Data accuracy and consistency are contingent on the quality of underlying databases, which may have inherent limitations. Furthermore, the study primarily concentrates on research within the scope of corporate bankruptcy prediction, potentially not fully capturing contributions from related fields. Moving forward, future research could aim to incorporate non-English language publications, conduct dynamic real-time analyses, integrate qualitative assessments for deeper insights, explore interdisciplinary studies, and examine the practical application of identified research trends, ultimately advancing the field.

References

Alam, T. M., Shaukat, K., Mushtaq, M., Ali, Y., Khushi, M., Luo, S., & Wahab, A. (2021). Corporate Bankruptcy Prediction: An Approach towards Better Corporate World. *Computer Journal*, *64*(11), 1731–1746. https://doi.org/10.1093/comjnl/bxaa056

Almeida, L. (2023). Risk and Bankruptcy Research: Mapping the State of the Art. *Journal of Risk and Financial Management*, *16*(8). https://doi.org/10.3390/jrfm16080361

Altman, E. I. (1994). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, *3*(10), 589–609.

Aziz, A. M., & Dar, H. A. (2006). Predicting corporate bankruptcy: Whither do we stand? *Loughborough University´s Institutional Repository*, *4*(1), 1–52.

Ben Jabeur, S., Stef, N., & Carmona, P. (2023). Bankruptcy Prediction using the XGBoost Algorithm and Variable Importance Feature Engineering. *Computational Economics*, *61*(2), 715–741. https://doi.org/10.1007/s10614-021-10227-1

Björk, B. C. (2017). Gold, green, and black open access. *Learned Publishing*, *30*(2), 173–175. https://doi.org/10.1002/leap.1096

Burright, M. (2006). The Access Principle: The Case for Open Access To Research and Scholarship. In *Issues in Science and Technology Librarianship* (Issue 47). https://doi.org/10.29173/istl2069

Cao, Y., Liu, X., Zhai, J., & Hua, S. (2022). A two-stage Bayesian network model for corporate bankruptcy prediction. *International Journal of Finance and Economics*, *27*(1), 455–472. https://doi.org/10.1002/ijfe.2162

du Jardin, P., Veganzones, D., & Séverin, E. (2019). Forecasting Corporate Bankruptcy Using Accrual-Based Models. *Computational Economics*, *54*(1), 7–43. https://doi.org/10.1007/s10614-017-9681-9

Esh, M., & Ghosh, S. (2021). A Scientometric Introspect of Literary Warrants Published in Annals of Library and Information Studies (ALIS) during the Last Decade. *Library Philosophy and Practice*, *2021*, 1–20.

Fakhar, S., Mohd Khan, F., Tabash, M. I., Ahmad, G., Akhter, J., & Al-Absy, M. S. M. (2023). Financial distress in the banking industry: A bibliometric synthesis and exploration. *Cogent Economics and Finance*, *11*(2). https://doi.org/10.1080/23322039.2023.2253076

Grosu, V., Chelba, A., Melega, A., Botez, D., & Socoliuc, M. (2023). Bibliometric analysis of the literature on evaluation models of the bankruptcy risk. *Strategic Management*, *28*(00), 42–42. https://doi.org/10.5937/straman2200035g

Guha, A., & Veeranjaneyulu, N. (2019). Prediction of bankruptcy using big data analytics based on fuzzy C-means algorithm. *IAES International Journal of Artificial Intelligence*, *8*(2), 168–174. https://doi.org/10.11591/ijai.v8.i2.pp168-174

Horak, J., Vrbka, J., & Suler, P. (2020). Support Vector Machine Methods and Artificial Neural Networks Used for the Development of Bankruptcy Prediction Models and their Comparison. *Journal of Risk and Financial Management*, *13*(3), 60. https://doi.org/10.3390/jrfm13030060

Kim, A. G., & Yoon, S. (2021). Corporate Bankruptcy Prediction with Domain-Adapted BERT. *Proceedings of the 3rd Workshop on Economics and Natural Language Processing, ECONLP 2021*, 26–36. https://doi.org/10.18653/v1/2021.econlp-1.4

Kim, H., Cho, H., & Ryu, D. (2020). Corporate default predictions using machine learning: A literature review. *Sustainability (Switzerland)*, *12*(16), 1–11. https://doi.org/10.3390/SU12166325

Kim, H., Cho, H., & Ryu, D. (2022). Corporate Bankruptcy Prediction Using Machine Learning Methodologies with a Focus on Sequential Data. *Computational Economics*, *59*(3), 1231–1249. https://doi.org/10.1007/s10614-021-10126-

Kumar, S. (2015). Co-authorship networks: a review of the literature. *Aslib Journal of Information Management*, *67*(1), 55–73. https://doi.org/10.1108/AJIM-09-2014-0116

Li, L., & Faff, R. (2019). Predicting corporate bankruptcy: What matters? *International Review of Economics and Finance*, *62*, 1–19. https://doi.org/10.1016/j.iref.2019.02.016

M. Kannadhasan, Halder, A. (2016). *Knowledge Structures, Progression and Emergent Areas of Corporate Bankruptcy: A Bibliometric and Topic Modelling Analyses*. *2*(1), 6–11.

Mirzakhanyan, A. (2005). Economic and social development. *The Armenians: Past and Present in the Making of National Identity*, 196–210. https://doi.org/10.4324/9780203004937

Morrison, H., Borges, L., Zhao, X., Kakou, T. L., & Shanbhoug, A. N. (2022). Change and growth in open access journal publishing and charging trends 2011 – 2021. *Journal of the Association for Information Science and Technology*, *73*(12), 1793–1805.

Muslim, M. A., & Dasril, Y. (2021). Company bankruptcy prediction framework based on the most influential features using XGBoost and stacking ensemble learning. *International Journal of Electrical and Computer Engineering*, *11*(6), 5549–5557. https://doi.org/10.11591/ijece.v11i6.pp5549-5557

Nobanee, H., Dilshad, M. N., Al Dhanhani, M., Al Neyadi, M., Al Qubaisi, S., & Al Shamsi, S. (2021). Big Data Applications the Banking Sector: A Bibliometric Analysis Approach. *SAGE Open*, *11*(4). https://doi.org/10.1177/21582440211067234

Shettar, I., & Hadagali, G. S. (2023). Global Research Publications on Open Data: A Scientometric Analysis. *SSRN Electronic Journal*, *February 2002*, 112–119. https://doi.org/10.2139/ssrn.4344330

Shi, Y., & Li, X. (2019). A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. *Heliyon*, *5*(12), e02997. https://doi.org/10.1016/j.heliyon.2019.e02997

Son, H., Hyun, C., Phan, D., & Hwang, H. J. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications*, *138*, 112816. https://doi.org/https://doi.org/10.1016/j.eswa.2019.07.033

Soui, M., Smiti, S., Mkaouer, M. W., & Ejbali, R. (2020). Bankruptcy Prediction Using Stacked Auto-Encoders. *Applied Artificial Intelligence*, *34*(1), 80–100. https://doi.org/10.1080/08839514.2019.1691849

Vochozka, M., Vrbka, J., & Suler, P. (2020). Bankruptcy or success? The effective prediction of a company's financial development using LSTM. *Sustainability (Switzerland)*, *12*(18). https://doi.org/10.3390/su12187529

Zapata, H. O., & Mukhopadhyay, S. (2022). A Bibliometric Analysis of Machine Learning Econometrics in Asset Pricing. *Journal of Risk and Financial Management*, *15*(11). https://doi.org/10.3390/jrfm15110535

Zelenkov, Y., & Volodarskiy, N. (2021). Bankruptcy prediction on the base of the unbalanced data using a multi-objective selection of classifiers. *Expert Systems with Applications*, *185*, 115559. https://doi.org/https://doi.org/10.1016/j.eswa.2021.115559