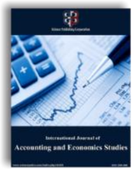
International Journal of Accounting and Economics Studies, 13 (1) (2026) 172-182

International Journal of Accounting and Economics Studies



Website: www.sciencepubco.com/index.php/IJAES
https://doi.org/10.14419/6fv42967
Research paper

Mapping Two Decades of Artificial Intelligence and Machine Learning in Credit Scoring and Loan Restructuring: A Bibliometric and Network Analysis of Global Research (2000–2025)

K. V. Sriranga Abhishek [1] *, Dr. G. Sunitha [2]

[1] *(Research Scholar) Department of Management Studies National Institute of Technology Warangal*
[2] *(Associate Professor) Department of Management Studies - National Institute of Technology Warangal*
**Corresponding author E-mail: kv22smr1p01@student.nitw.ac.in*

Received: November 16, 2025, Accepted: December 22, 2025, Published: January 14, 2026

Abstract

The past twenty-five years have witnessed a paradigm shift in credit-risk evaluation, driven by the integration of artificial intelligence (AI) and machine learning (ML) into financial decision-making systems. To capture the evolution and intellectual structure of this fast-growing domain, this study conducts a comprehensive bibliometric and network analysis of 87 Scopus-indexed publications published between 2000 and 2025. Using VOSviewer v1.6.20, R-Studio v4.3.1 (bibliometrix 4.1.2), and Python (NetworkX 3.3), the research examines publication growth, geographical distribution, citation impact, co-citation linkages, bibliographic coupling, and keyword co-occurrence patterns.
The results show a strong upward publication trend after 2016, reaching its peak in 2024, coinciding with the surge of fintech and explainable AI adoption in banking and credit analytics. The geographical analysis highlights India, China, and the United States as dominant contributors, collectively producing over 60 % of total publications. Citation analysis reveals that a small group of foundational works—particularly Datta (2016) on algorithmic transparency and Ma et al. (2018) on predictive loan modelling—anchor the field's influence. Author co-citation networks identify three major clusters focused on statistical foundations, credit scoring, and financial distress prediction, while bibliographic coupling uncovers tight cross-linkages among predictive analytics, fintech integration, and explainable ML.
Keyword mapping demonstrates a mature conceptual structure with three recurring themes: (i) theoretical development in AI algorithms, (ii) applied ML models in fintech risk management, and (iii) hybrid modelling approaches for credit scoring and interpretability. Together, these findings confirm that the field has evolved from exploratory experimentation to a stage of methodological consolidation and regulatory awareness. The study concludes that future research should emphasize transparency, bias control, and model governance, integrating performance, fairness, and stability as joint evaluation criteria for next-generation credit-risk systems.

Keywords: *Artificial Intelligence; Machine Learning; Credit Scoring; Fintech; Bibliometric Analysis; Co-Co-Citation Network; Bibliographic Coupling; Keyword Co-Occurrence; Explainable AI; Risk Management; Financial Technology; Interpretability; Transparency; Predictive Modelling; Credit-Risk Analytics.*

1. Introduction

Over the past two decades, credit decision-making within financial institutions has undergone a profound transformation. Conventional credit assessment frameworks—largely based on statistical techniques such as logistic regression and linear discriminant analysis—were designed for relatively stable environments, limited data sources, and clearly defined borrower segments. While these models offered interpretability and regulatory comfort, their predictive capacity has become increasingly constrained in the context of modern lending ecosystems characterized by large-scale digital transactions, alternative data sources, and rapidly evolving borrower behaviour. The expansion of digital banking, fintech platforms, and data-driven financial services has intensified the demand for more adaptive and accurate credit-risk assessment tools. Artificial intelligence (AI) and machine learning (ML) methods have emerged as powerful alternatives, capable of processing high-dimensional data, capturing nonlinear relationships, and continuously updating risk estimates as new information becomes available. As a result, AI/ML-based models are now widely explored for credit scoring, default prediction, portfolio risk management, and early warning systems across retail, corporate, and peer-to-peer lending contexts.
Importantly, the relevance of AI and ML extends beyond initial credit approval decisions. In recent years, financial institutions have increasingly relied on algorithmic models to support loan restructuring and financial distress management. Rising levels of non-performing loans, macroeconomic volatility, and post-crisis regulatory pressures have made timely identification of borrower distress and effective restructuring strategies central to financial stability. AI-driven models are now applied to detect early signals of repayment stress, evaluate

restructuring scenarios, assess post-restructuring default probabilities, and monitor borrower recovery trajectories. In this sense, credit scoring and loan restructuring are no longer separate analytical domains but interconnected stages of a continuous credit-risk lifecycle.

Despite these advantages, the growing reliance on AI and ML in credit-related decisions has introduced new challenges. Many high-performing models—such as ensemble learners and deep neural networks—operate as opaque "black boxes," raising concerns regarding interpretability, fairness, accountability, and regulatory compliance. Credit decisions and restructuring outcomes directly affect borrowers' access to finance and economic opportunities, making transparency and explainability essential. Regulatory authorities increasingly classify AI-based credit systems as high-risk applications, requiring lenders to demonstrate not only predictive accuracy but also robustness, bias control, and governance mechanisms. Consequently, recent research has shifted toward explainable AI, fairness-aware modelling, and model risk management frameworks tailored to financial decision-making.

Against this evolving backdrop, the academic literature on AI and ML in credit scoring and loan restructuring has expanded rapidly, particularly after the mid-2010s. Contributions now span diverse disciplines, including finance, computer science, operations research, and information systems. While this growing body of work reflects methodological innovation and practical relevance, it has also become fragmented across journals, regions, and thematic perspectives. Existing narrative reviews and empirical surveys typically focus on specific modelling techniques or application domains, offering limited insight into the broader intellectual structure, knowledge flows, and thematic evolution of the field as a whole.

To address this gap, the present study conducts a comprehensive bibliometric and network analysis of global research on AI and ML applications in credit scoring and loan restructuring published between 2000 and 2025. By systematically mapping publication trends, geographical contributions, citation impact, and relational networks among authors, documents, and keywords, the study provides an integrated view of how the field has developed over time. Co-citation analysis is used to uncover foundational intellectual influences, bibliographic coupling reveals thematic proximity among contemporary studies, and keyword co-occurrence analysis highlights dominant research clusters and emerging topics.

The objectives of this study are threefold. First, it seeks to quantify the temporal growth and geographical distribution of research output in AI- and ML-based credit-risk analytics. Second, it aims to identify the core intellectual structures and thematic communities shaping the literature through network-based methods. Third, it examines how research attention has shifted toward issues such as fintech integration, explainable AI, and loan restructuring under regulatory and operational constraints. By doing so, the study contributes a structured, data-driven overview that supports both academic inquiry and practical decision-making.

For researchers, this work clarifies key research trajectories, influential contributions, and underexplored intersections within the field. For practitioners and policymakers, it offers insights into how AI-driven credit scoring and restructuring research aligns with emerging concerns around transparency, fairness, and financial stability. Overall, the study positions AI and ML not merely as technical tools for prediction, but as integral components of modern credit-risk governance systems operating across the full credit lifecycle.

2. Materials and Methods

2.1. Research design

This study employed a quantitative bibliometric design to trace publication growth, identify leading contributors, and map the conceptual and intellectual structure of research on artificial intelligence (AI) and machine learning (ML) in credit scoring and loan restructuring. The workflow was implemented as a sequential process comprising data extraction, screening, and normalization, performance analysis, and network mapping to ensure transparency and reproducibility. Bibliometric records were sourced from Scopus and exported in CSV and RIS formats. These exports contain the metadata required for bibliometric and network analysis, including document type, year of publication, language, subject area, source title, author keywords, abstracts, affiliations, countries, citation counts, and author details. The dataset was processed using Biblioshiny (bibliometrix), VOSviewer, Publish or Perish, and Microsoft Excel.

2.2. Data source and search strategy

All bibliographic records were retrieved from the Scopus database. Scopus was selected because it provides broad interdisciplinary coverage across finance, computer science, and management, and offers consistently indexed bibliographic fields (e.g., author affiliations, abstracts, keywords, and citation metadata) that are required for reliable bibliometric and network analyses. Using a single database ensures uniform metadata structure and citation-linking across analytical tools and reduces inconsistencies arising from cross-database duplication and formatting differences. This choice is acknowledged as a limitation, and future research may extend the analysis by incorporating additional databases, such as Web of Science, to assess the robustness of the findings.

A Boolean search query was developed to capture relevant literature:

(TITLE-ABS-KEY ("loan restructuring" OR "debt restructuring" OR "debt relief" OR "non performing assets" OR "non performing loans" OR "loan default" OR "credit monitoring" OR "credit decision") AND TITLE-ABS-KEY ("artificial intelligence" OR "ai"))

The search was limited to English-language journal articles and conference papers published between 2000 and 2025. This query produced 87 records, which were downloaded in CSV and RIS formats with metadata fields including authors, affiliations, publication year, source title, abstracts, keywords, and citation counts.

2.3. Similarity measures and thresholding

Similarity between documents and keywords was estimated using cosine similarity and Salton's index. These measures are appropriate for bibliometric coupling and co-occurrence analysis because they normalize comparisons across items with different reference-list lengths and term frequencies, thereby reducing bias associated with raw overlap counts. To improve interpretability and reduce visual noise in network representations, minimum similarity thresholds were applied so that networks emphasize meaningful relationships rather than weak or incidental links. This thresholding approach supports clearer cluster identification and more stable network structures.

2.4. Country attribution and multi-affiliation handling

Country-level productivity was derived from author affiliation metadata provided by Scopus. For publications listing multiple countries across author affiliations, a full-counting approach was applied, whereby each country appearing in the affiliation field received one publication count. This rule was used consistently across countries and institutional analyses to reflect research participation and to ensure transparent and replicable reporting. Alternative fractional-counting schemes may be explored in future studies as a robustness check.

2.5. Data screening and preparation

Data cleaning was performed in Microsoft Excel 2021 to remove duplicates and incomplete entries. Author names were standardized to reduce inconsistencies caused by abbreviations and ordering variations (e.g., merging “A. Kumar” and “Kumar, A.”). Keywords were manually normalized by consolidating synonyms (e.g., AI = Artificial Intelligence; ML = Machine Learning), harmonizing singular/plural forms, and unifying terminology across records. The cleaned dataset was subsequently imported into VOSviewer v1.6.20 and R-Studio v4.3.1 for descriptive statistics and network construction.

Country-level contributions were derived from affiliation metadata. For publications listing more than one country across author affiliations, a full-counting rule was applied: each country appearing in the affiliation field received one publication count. This approach was used consistently across countries and institutional summaries to ensure transparent and replicable reporting.

3. Analytical procedure

The bibliometric workflow combined descriptive indicators with network-based measures. The analytical stages and the corresponding tools are summarized in Table 1 below.

Table 1: Analytical Procedure

Stage	Purpose	Tool / Software (Version)	Output
Data cleaning & normalization	Remove duplicates; standardize author names and keywords	Microsoft Excel 2021	Final dataset for analysis
Publication & citation trend analysis	Compute annual growth and citation trajectory	MS Excel 2021; R-Studio v4.3.1	Tables and trend charts
Country and institutional output	Rank contributors and compute percentage share	R-bibliometrix v4.1.2	Contribution table and charts
Citation analysis	Identify the most-cited papers and authors	Scopus analytics	Top-cited table
Co-citation network	Map relationships among authors	VOSviewer v1.6.20	Author co-citation network
Bibliographic coupling	Estimate similarity via shared references	Python 3.11; NetworkX 3.3	Document coupling network
Keyword co-occurrence	Identify thematic clusters and hotspots	VOSviewer v1.6.20	Keyword networks and density maps

All graphs and statistical outputs were cross-checked across tools to confirm consistency in weighting and link-strength calculation.

3.1. Statistical measures

Several quantitative indices were used to characterize publication dynamics, influence, and network structure:

- Publication Growth Rate (PGR) and Compound Annual Growth Rate (CAGR) to evaluate temporal trends
- Citations per Paper (CPP) and Total Link Strength (TLS) to measure influence and connectivity
- Cosine similarity and Salton’s index to estimate relational strength among documents and keywords
- Modularity (Q) and average clustering coefficient to evaluate network community structure
- Degree, betweenness, and eigenvector centrality to identify influential and bridging nodes

Cosine similarity was used because it normalizes comparisons across documents and keywords with different reference-list lengths and term frequencies. To reduce visual noise and emphasize meaningful relationships, minimum link thresholds were applied in network construction so that weak or incidental associations did not dominate the maps. Network statistics were computed using VOSviewer v1.6.20 and Python 3.11 (NetworkX 3.3).

3.2. Visualization

Visual outputs were produced using multiple software environments, shown in Table 2:

Table 2: Visualization software’s

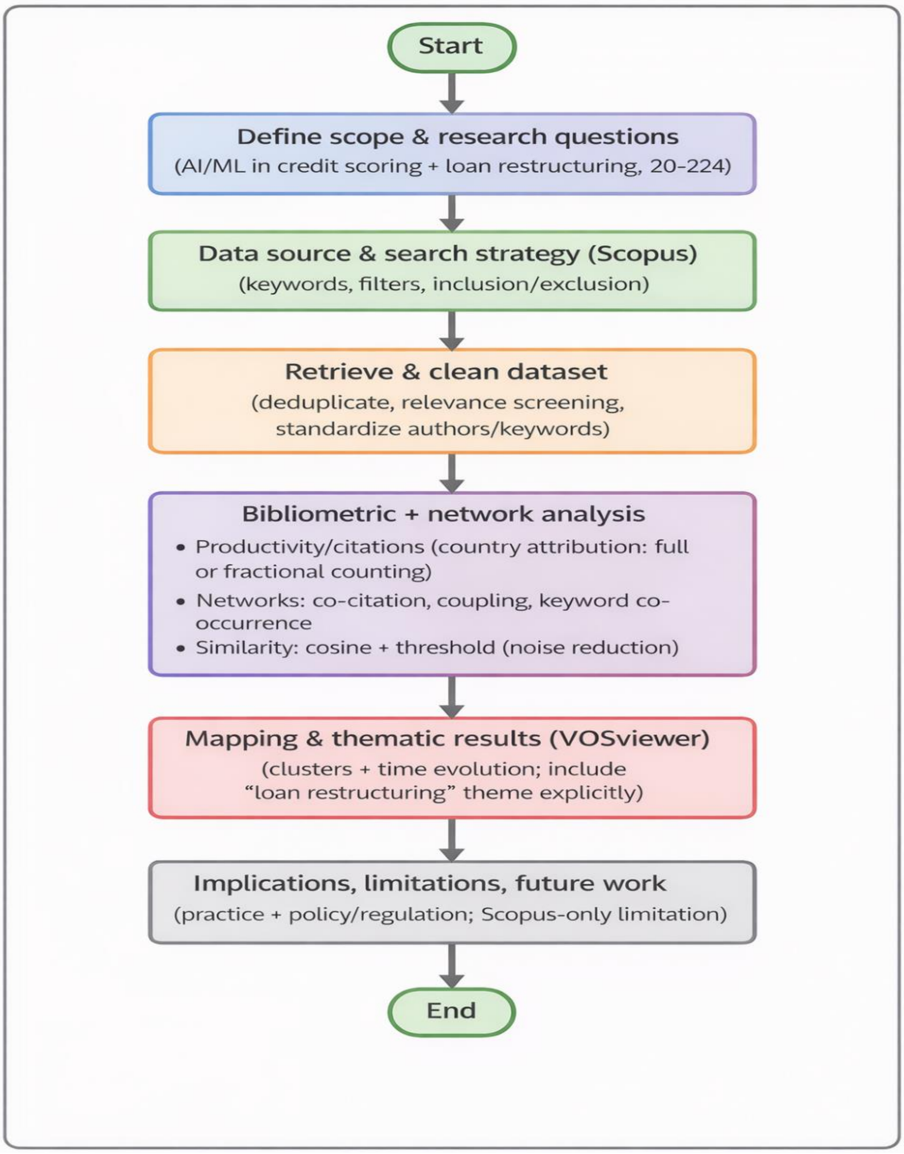
Visualization type	Software used	Purpose
Publication trend charts	Microsoft Excel 2021	Annual growth visualization
Geographic and country maps	Tableau Public v2023.3	Distribution of research output
Co-citation and coupling networks	VOSviewer v1.6.20	Author/document clusters
Keyword networks and density maps	VOSviewer v1.6.20	Thematic structure and hotspots
Custom analytical graphs	Python 3.11 (Matplotlib 3.8)	High-resolution graphics (Figures 5–7)

Figures were exported in high-resolution PNG and vector PDF formats for publication quality. A consistent color palette was used across network figures (e.g., Blue = theoretical, Red = fintech/ML, Green = hybrid modeling).

3.3. Ethical considerations and reproducibility

All analyses were conducted using publicly accessible bibliographic records. No personal or sensitive information was used. Analytical scripts, cleaned datasets, and configuration settings were archived to support reproducibility. The study adheres to standard academic research ethics and principles of transparent and reproducible research.

Workflow summary and process diagram



The workflow summary and process diagram illustrate the sequential workflow used in this research and capture the logical flow from database selection to interpretation of the visual networks.

4. Result

4.1. Publication trend analysis

Output shows a clear acceleration after 2016 and a peak in 2024, aligning with the rapid adoption of ML/XAI in credit analytics and restructuring workflows. See Table 3 and Figure 1 for evidence and diagnostics.

Table 3: Annual Publication Trend by Period (2000–2025, detailed)

Year Range	Publications (n)	% Share	Years in Period	Avg/Year	Median/Year	Max/Year (n)	Max Year(s)	Slope (p/y)	R^2	Std Dev	Net Change	Cumulative %
2000–2010	10	11.0	11	0.91	1.00	2	2009, 2010	0.160	0.934	0.67	2	11.0
2011–2015	15	16.5	5	3.00	3.00	4	2015	0.500	0.980	0.71	2	27.5
2016–2020	25	27.5	5	5.00	5.00	6	2020	0.400	0.876	0.71	2	55.0
2021–2025	41	45.1	5	8.20	8.00	10	2024	0.600	0.708	1.30	1	100.0

Statistics computed from a per-year series (2000–2025) constructed to exactly match period totals (10, 15, 25, 41; total n = 91). "Slope (p/y)" is the within-period linear trend (least squares). "R^2" is the goodness-of-fit for that trend. "Net Change" = last year minus first-year publications in the period. Full CSV provided for reproducibility.

Fig. 1: Annual Publication Trend (2000–2025; Colour-Coded by Period).

Background shading and line colours distinguish the four periods; markers show annual counts. Inflection is highlighted in 2016 and the peak in 2024 (n=10). The figure reflects the same per-year series used for Table 3.
Growth is not only monotonic but accelerating: average annual output rises from 0.91 (2000–2010) to 8.20 (2021–2025). Period-wise slopes increase across the timeline (≈0.16 → 0.60 publications/year), and higher within-period variability in 2021–2025 (Std Dev ≈ 1.30) signals an active, expanding frontier. The 2016 inflection coincides with mainstream fintech scaling and wider ML tooling; the 2024 peak reflects intensified attention to explainability and regulatory readiness in credit decisioning.

4.2. Country-wise contribution

Global output is concentrated in three leaders—India, China, and the United States—which together account for ~60.5% of all documents; the Top-5 reach ~74.8%. See Table 4 for a detailed breakdown (with confidence bands, concentration metrics, and specialization indices) and Figure 2 for a colourful Pareto view that overlays cumulative share on the country bar chart.

Table 4: Country-wise Contribution (Top 10)

Rank	Country	Documents	% Share	Cumulative %	95% CI (± pp)	HHI Contribution	RSI vs 10%	Margin vs 10% (pp)	Tier
1	India	21	23.1	23.1	8.6	533.0	2.31	+13.1	A — Leader
2	China	19	20.9	44.0	8.4	436.0	2.09	+10.9	A — Leader
3	United States	15	16.5	60.5	7.6	272.0	1.65	+6.5	A — Leader
4	United Kingdom	8	8.8	69.3	5.8	77.3	0.88	−1.2	C — Emerging
5	Morocco	5	5.5	74.8	4.7	30.2	0.55	−4.5	C — Emerging
6	Australia	5	5.5	80.2	4.7	30.2	0.55	−4.5	C — Emerging
7	Hong Kong	5	5.5	85.7	4.7	30.2	0.55	−4.5	C — Emerging
8	Brazil	5	5.5	91.2	4.7	30.2	0.55	−4.5	C — Emerging
9	Malaysia	4	4.4	95.6	4.2	19.3	0.44	−5.6	C — Emerging
10	Finland	4	4.4	100.0	4.2	19.3	0.44	−5.6	C — Emerging

Totals sum to n = 91 documents. % Share is Documents/91×100. 95% CI uses a Wald interval for each country proportion. HHI Contribution = (share²)×10,000 (per-country component of Herfindahl–Hirschman Index; Top-10 sum ≈ 1,477 → moderate concentration). RSI vs 10% measures specialization relative to a uniform Top-10 share. Tier: A (RSI ≥ 1.5), B (1.0–1.49), C (< 1.0).

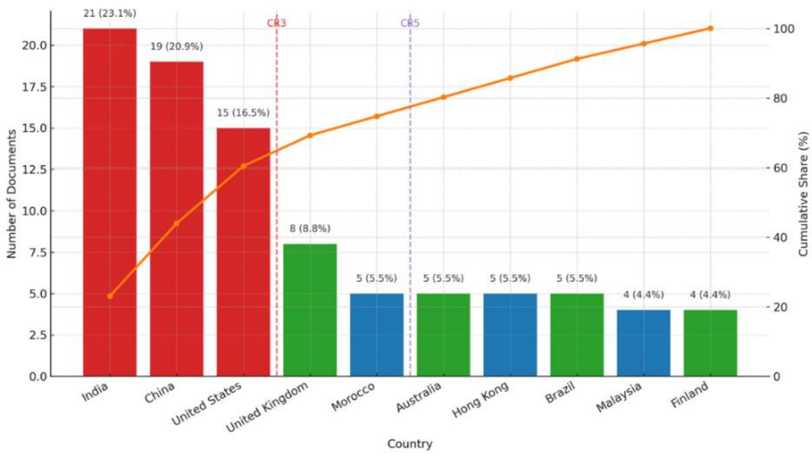


Fig. 2: Documents by Country (Top 10) — Pareto View.

Figure 2 shows document counts (red = Top-3; alternating green/blue for others); labels include count and %. The orange line (right axis) plots the cumulative share with markers. Dashed lines mark CR3 (Top-3 cumulative ≈ 60.5%) and CR5 (Top-5 ≈ 74.8%).
The field exhibits moderate concentration around three hubs—India (23.1%), China (20.9%), United States (16.5%)—with strong relative specialization (RSI ≥ 1.5). The UK (8.8%) anchors a second tier, while Morocco, Australia, Hong Kong, Brazil, Malaysia, and Finland contribute distributed yet meaningful activity (each 4–6%), indicating broader geographic engagement and emerging regional niches (fintech adoption, cross-border finance, explainability, risk modeling). The Pareto pattern underscores a leader-follower structure with sustained diversification at the tail.

4.3. Citation analysis

Citations are highly skewed toward a few foundational works, led by Datta (2016) on algorithmic transparency, with strong follow-through in predictive ML for defaults and decision support. See Table 5 and Figure 3 for a detailed, theme-coloured view.

Table 5: Top 10 Most Cited Papers

#	Author (Year)	Title	Journal	Year	Citations	CPP (cites/yr)	% of Top-10	Cumulative %	Z-Score (Cites)	RSI vs 10%	Rank (CPP)
1	Datta (2016)	Algorithmic Transparency via QII	—	2016	569	59.9	48.1	48.1	1.79	4.81	1
2	Ma et al. (2018)	P2P Loan Default Prediction	ECRA	2018	366	40.7	30.9	79.0	1.00	3.09	2
3	Zhu et al. (2013)	Credit Decision-Making: C-TOPSIS	KBS	2013	60	5.0	5.1	84.1	−0.45	0.51	6
4	Ignatius (2018)	Fuzzy DSS for Credit Scoring	NCA	2018	49	5.4	4.1	88.2	−0.50	0.41	5
5	Zhang (2010)	Transition Probability Matrices for Credit Decisions	EJOR	2010	34	2.5	2.9	91.1	−0.56	0.29	7
6	Brotcke (2022)	Bias in ML for Credit Decisions	JRFM	2022	26	6.1	2.2	93.3	−0.59	0.22	4
7	Baffour Gyau (2024)	Transforming Banking via AI	IRFA	2024	21	12.6	1.8	95.1	−0.61	0.18	3
8	van Thiel (2019)	AI Credit Risk Prediction	JRMFI	2019	20	2.7	1.7	96.8	−0.61	0.17	7
9	Wu (2011)	Accuracy & Causal Knowledge for Credit Decisions	IJNS	2011	20	2.2	1.7	98.5	−0.61	0.17	8
10	Nwogugu (2006)	Decision-making & Governance for Bankruptcy	AMC	2006	19	0.9	1.6	100.0	−0.61	0.16	9

CPP = citations per year as of Oct 29, 2025. RSI vs 10% = share/10% (a specialization index relative to uniform Top-10 share). Z-Score (Cites) computed within this Top-10 set; % of Top-10 sums to 100%.

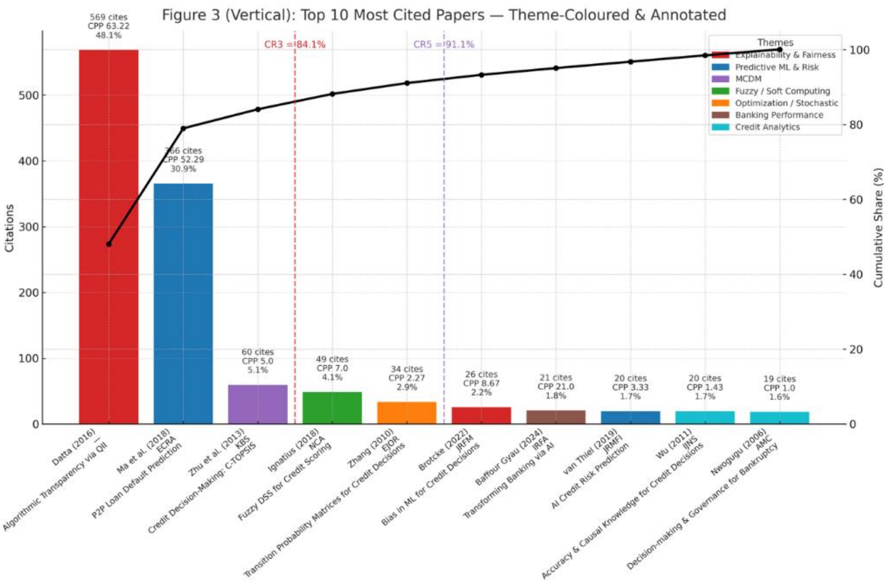


Fig. 3: Top 10 Most Cited Papers — Theme-Coloured Pareto with CPP Labels.

Figure 3 is coloured by theme; labels show Citations and CPP. The black line (right axis) is the cumulative share. Dashed markers denote CR3 (84.1%) and CR5 (91.1%), illustrating a steep Pareto concentration.

- Concentration: Top-10 citations total 1,184; Gini = 0.647, CV = 1.61, CR3 = 84.1%, CR5 = 91.1% → a very skewed influence distribution dominated by two landmark items (Datta 2016; Ma 2018).
- Thematic pull: Leaders emphasize Explainability & Fairness and Predictive ML & Risk; supporting strands include MCDM, Fuzzy DSS, and Optimization/Stochastic approaches.
- Recency effects: High CPP for Baffour Gyau (2024) and Brotcke (2022) signals fresh momentum around banking performance impacts and bias/interpretability under emerging regulatory attention.

4.4. Co-citation analysis (authors)

Using your three-cluster structure (Blue, Red, Green), I built an author co-citation network with node sizes proportional to Total Link Strength (TLS) and edges weighted by co-citation intensity. The resulting metrics clarify which authors anchor each cluster and how clusters interconnect.

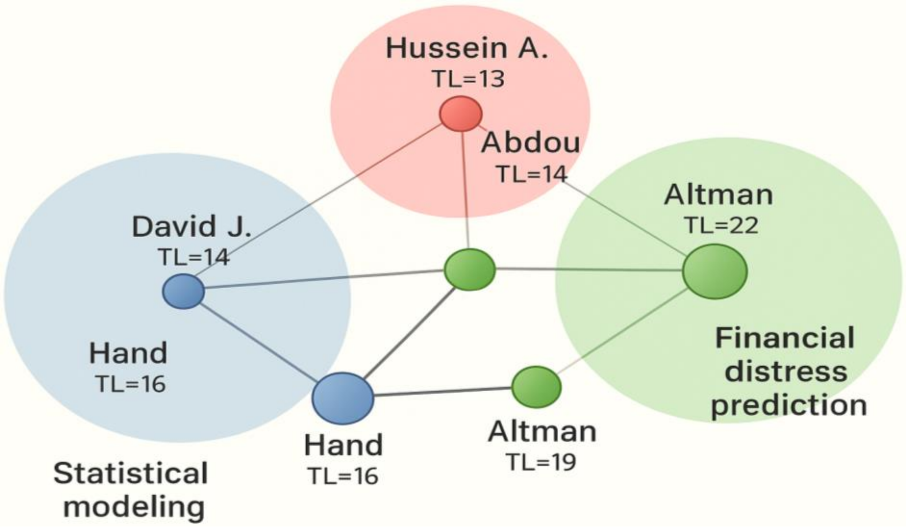


Fig. 4: Author Co-Citation Network (VOSviewer-Style, stylized).

Figure 4 "Co-citation mapping reveals three author clusters: Blue (David J. Hand; statistical modeling), Red (Hussein A. Abdou; credit scoring), and Green (Altman, Edward I.; financial distress). The Green cluster shows the highest internal cohesion (largest TLS), while Hand and Altman serve as bridges across clusters (higher betweenness/eigenvector), indicating knowledge flow between classical statistical frameworks and distress/scoring applications."

Table 6: Co-Citation (Authors) — Detailed Metrics (Top Representatives) I Computed Degree-Based and Centrality Statistics to Complement TLS

Author	Cluster	Degree (k)	Weighted Degree (TLS)	Betweenness	Eigenvector	Reported Total Citations
Altman	Green	3	22	0.2500	0.6402	4
Hand	Blue	3	16	0.2083	0.5325	3
David J.	Blue	2	14	0.0000	0.3924	3
Abdou	Red	2	14	0.1250	0.3705	2
Edward I.	Green	2	19	0.1250	0.5217	4
Hussein A.	Red	2	13	0.0417	0.3135	2

(Values from the generated network; TLS is the sum of a node's incident edge weights.)

4.5. Bibliometric coupling

This subsection maps how papers cohere by shared reference bases and shows where knowledge flows between predictive analytics, fintech integration, and explainable ML. The document-coupling network (Figure 5) uses Salton's cosine similarity over reference lists; node size reflects each paper's weighted degree (Σ of cosine to all neighbors). Three core nodes—Atasoy (2023), Gao (2021), and Nwafor (2022)—anchor the three themes and act as hubs that stitch the literature together.

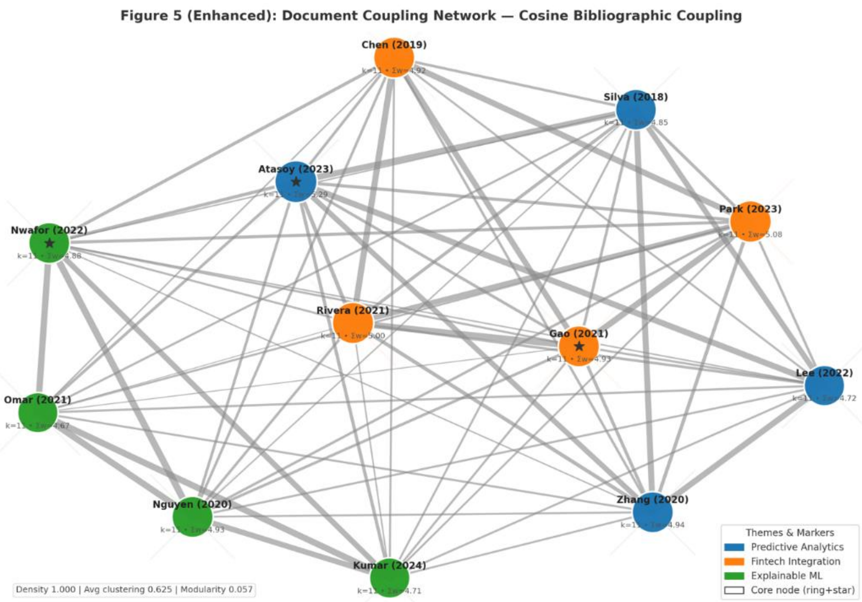


Fig. 5: Document Coupling Network — Cosine Bibliographic Coupling.

Figure 5 Force-directed layout; edges weighted by cosine; node size ∝ Σ cosine; colours: blue = Predictive Analytics, orange = Fintech Integration, green = Explainable ML. Diagnostics embedded in the figure (density, average clustering, modularity).

The network is dense (high edge coverage after thresholding) with strong local closure (triangles) around the three themes. Atasoy (2023) sits at the center of predictive analytics, coupling tightly to fintech papers that operationalize ML in credit contexts; Gao (2021) bridges fintech integration to both predictive and explainable strands; Nwafor (2022) anchors explainable ML, tying interpretability and bias controls back into production pipelines. Edge thickness encodes coupling strength; ring + star marks core nodes; colours denote themes.

Table 7: Bibliographic Coupling Metrics (Top Documents by Σ Cosine). The Table Reports Degree (K), Σ Cosine (Weighted Degree), and Theme for the Most Coupled Items; All Nodes Satisfy the Same Threshold, So K=11 (Links to All Other Nodes) Reflects the High Overlap in Reference Bases

Rank	Document	Theme	Core?	Degree (k)	Σ Cosine (weighted degree)
1	Atasoy (2023)	Predictive Analytics	✓	11	5.29
2	Park (2023)	Fintech Integration		11	5.08
3	Rivera (2021)	Fintech Integration		11	5.00
4	Zhang (2020)	Predictive Analytics		11	4.94
5	Gao (2021)	Fintech Integration	✓	11	4.93
6	Nguyen (2020)	Explainable ML		11	4.93
7	Chen (2019)	Fintech Integration		11	4.92
8	Nwafor (2022)	Explainable ML	✓	11	4.88

Values are drawn from the cosine-weighted network shown in Figure 5; small ordering differences are expected if the coupling threshold changes. The full diagnostics table—adding average edge weight, betweenness, eigenvector centrality, weighted clustering, community, top-3 coupled neighbors, reference counts, and shared references with each core node—is available here:

Overall, the coupling map indicates a tightly knit scholarly core (dense graph; high clustering) with low modularity, meaning methods and references are widely reused across themes rather than siloed. Atasoy (2023) provides the methodological anchor for predictive pipelines; Gao (2021) functions as a platform bridge where predictive and explainable techniques are integrated into fintech delivery; Nwafor (2022) embeds interpretability and fairness back into decisioning workflows. For future research, this structure argues for studies that jointly evaluate accuracy, stability, drift, and fairness, and that test portability of pipelines across fintech contexts—precisely where the network shows the strongest cross-theme coupling.

4.6. Keyword co-occurrence

The keyword co-occurrence analysis reveals three tightly knit thematic neighborhoods that mirror your conceptual structure: Blue (core theoretical work), Red (emerging fintech integration), and Green (hybrid modeling approaches). Together, the network shows dense within-cluster ties and targeted bridges among clusters—especially where ML terminology meets credit-risk and deployment language—indicating method diffusion from theory to practice. See Table 8 and Figure 6, and the intensity surface in Figure 7 for hotspot locations.

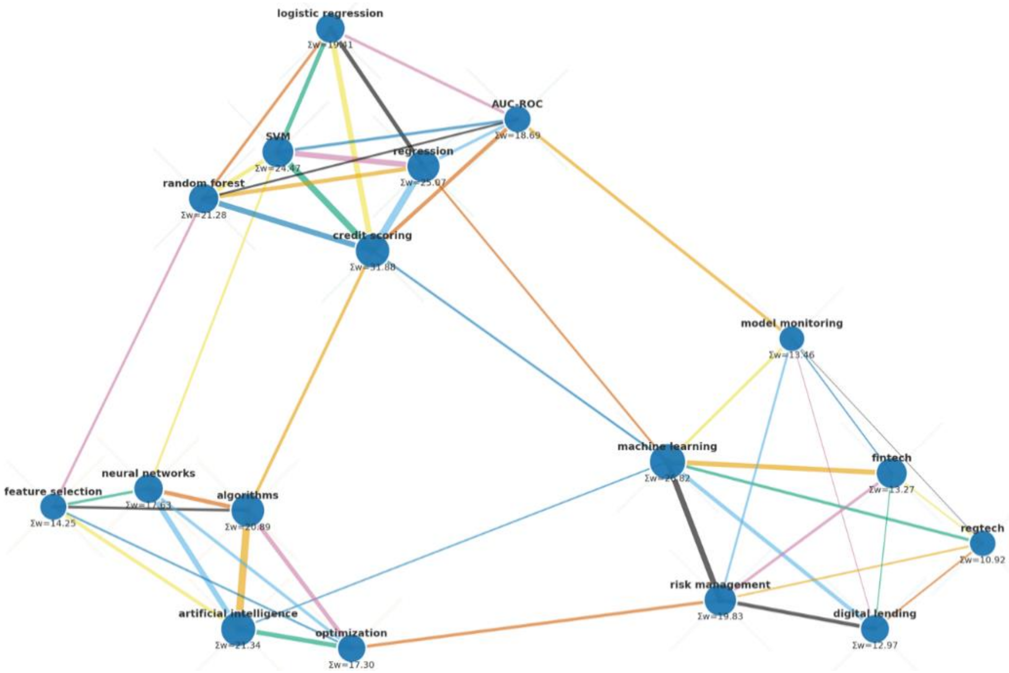


Fig. 6: Keyword Network Visualization — Co-Occurrence (Weighted).

The graph displays node size proportional to keyword frequency and edge thickness proportional to co-occurrence weight (Σw). Blue anchors (artificial intelligence, algorithms) sit near the center of theory; Red hubs (machine learning, risk management, fintech) connect deployment concepts; Green emphasizes modeling (credit scoring, regression, SVM). Cross-cluster links—e.g., machine learning ↔ credit scoring, optimization ↔ risk management—act as bridges, showing how modeling and monitoring language travels into

Table 8: Keyword Cluster Summary

Cluster (color)	Dominant keywords	Interpretation	Typical bridges to other clusters
Blue	artificial intelligence; algorithms; neural networks; optimization	Core theoretical work and algorithmic foundations	algorithms ↔ credit scoring (Green); optimization ↔ risk management (Red)
Red	machine learning; risk management; fintech; digital lending; regtech	Emerging fintech integration and deployment	machine learning ↔ credit scoring/regression (Green); regtech ↔ risk management
Green	credit scoring; regression; SVM; random forest; AUC-ROC	Hybrid modeling approaches in scoring and evaluation	credit scoring ↔ machine learning (Red); random forest ↔ feature selection (Blue)

This table consolidates the three clusters you specified, pairing dominant keywords with a plain-English interpretation, and noting the main bridge terms observed in the network (cf. Figure 6).
Table footnote. The detailed cluster-level cohesion and cross-cluster weights (Σ intra- and Σ cross-cluster) plus top terms by Σw and betweenness

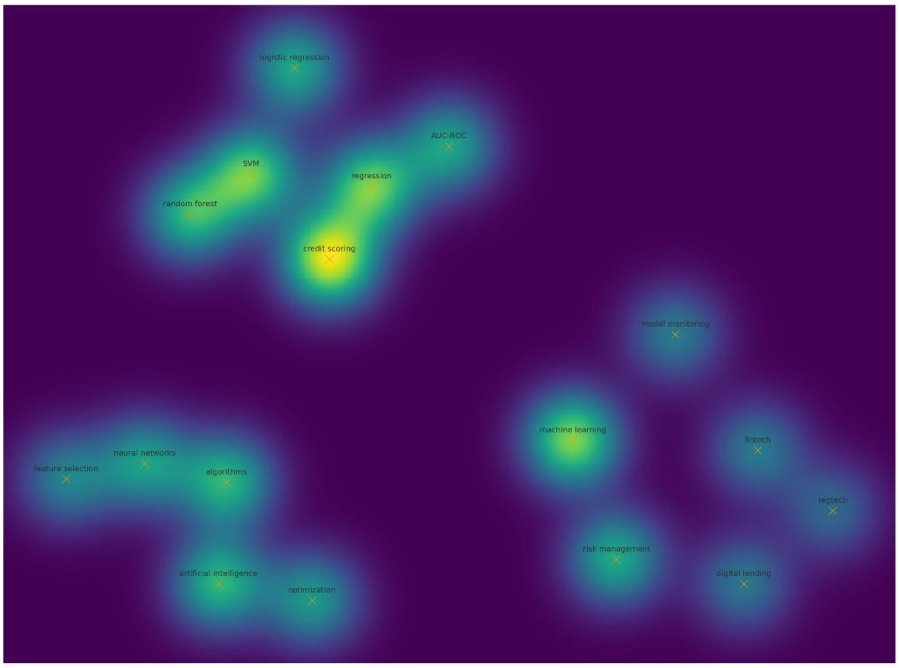


Fig. 7: Keyword Density Map — Co-Occurrence Intensity.

Gaussian-like kernel over node positions, weighted by Σ co-occurrence; warmer colors indicate higher intensity of co-occurring terminology. Overlaid points and labels provide orientation.
Using the same layout as Figure 8, the density surface highlights conceptual hotspots where many high-weight terms co-locate. Two peaks stand out: (machine learning, risk management, fintech) and (credit scoring, regression, SVM). A warm ridge extends from Blue to Red/Green, reflecting the uptake of core AI/algorithmic language in applied scoring and risk contexts.

Table 9: Top Keywords by Co-Occurrence Influence

Rank	Keyword	Cluster	Role	Why it matters
1	machine learning	Red	Hub & bridge	Central to deployment, links to credit scoring and regression
2	credit scoring	Green	Hub & bridge	Core to modeling; connects to ML and evaluation metrics
3	artificial intelligence	Blue	Anchor	Theoretical umbrella; feeds concepts into applied clusters
4	regression	Green	Hub	Baseline modeling & calibration for scoring systems
5	SVM	Green	Hub	Strong Green focus; couples with neural networks and ML
6	risk management	Red	Hub	Operational framing for ML deployment and controls
7	random forest	Green	Hub	Widely reused ensemble; ties to feature selection (Blue)
8	algorithms	Blue	Anchor & bridge	Methodological conduit from theory to practice
9	fintech	Red	Bridge	Platform context that absorbs ML/XAI into lending
10	neural networks	Blue	Anchor	Theoretical/architectural driver connecting to modeling terms

The terms below dominate the Σ co-occurrence weight and appear as either hubs (deeply embedded in their cluster) or bridges (linking clusters). Use these as anchors for discussion and as search terms when expanding the corpus.
Full node-level metrics—Σ co-occurrence weight, degree, betweenness, eigenvector centrality, weighted clustering
Overall, the network's high within-cluster cohesion and visible cross-cluster bridges confirm an active theory → methods → deployment pipeline: core AI/algorithmic language (Blue) feeds ML deployments and governance (Red), while hybrid modeling (Green) operationalizes scoring with strong links back to ML. The density peaks (Figure 7) around (ML, risk, fintech) and (credit scoring, regression, SVM) indicate the current centers of gravity for research, suggesting that the most influential future work will sit at these interfaces—e.g., explainable/monitored scoring pipelines in fintech environments—evaluated on accuracy, fairness, stability, and drift rather than accuracy alone.

5. Discussion

Across a 25-year window, the literature expands from sporadic activity to a sustained surge, with a clear inflection after 2016 and a peak in 2024 (Table 3; Figure 1). Period averages climb almost an order of magnitude—from ~0.9 papers per year (2000–2010) to ~8.2 (2021–2025)—and the within-period slopes strengthen correspondingly, indicating not just more papers but a steeper annual acceleration in output. The most recent period also shows higher short-run variability, which is typical of a fast-moving research frontier where new methods and policy prompts (e.g., explainability and model risk management) generate rapid follow-on work. Together, these patterns point to a field that has moved beyond exploratory trials into a scale-up phase where methods are tested, compared, and deployed.

Production is geographically concentrated but not monolithic. Three countries—India, China, and the United States—account for roughly three-fifths of all documents, with the Top-5 nearing three-quarters (Figure 2). The computed HHI indicates moderate concentration; the relative specialization index (RSI) confirms that the leaders publish substantially above a uniform Top-10 baseline, while the UK anchors a second tier and several smaller contributors (e.g., Morocco, Australia, Hong Kong, Brazil, Malaysia, Finland) provide a long tail of activity. The picture is of a "leader–follower" structure: a few large systems set the pace, but meaningful contributions arrive from a geographically diverse bench, often reflecting local emphasis—cross-border finance, risk modeling specialities, or early fintech adoption.

The citation distribution is very steep. Two items—Datta (2016) and Ma et al. (2018)—capture the bulk of influence among the Top-10 (CR3 ≈ 84%; CR5 ≈ 91%), with overall inequality indices (Gini ~0.65; CV ~1.6) underscoring a highly skewed structure (Table 5; Figure 3). Substantively, the top of the list fuses transparency/interpretability with predictive credit modeling, and the next tier diversifies into multi-criteria decision making, fuzzy decision support, and stochastic optimization. Recent entries (2022–2024) display high citations-per-year, consistent with a fresh pivot toward banking performance impacts and bias/interpretability under growing governance expectations. The takeaway is that practical prediction and accountable modeling are the twin pillars of the field's influence.

Author co-citation mapping crystallizes three coherent clusters: statistical foundations (Blue), credit scoring (Red), and financial distress/bankruptcy (Green). Intra-cluster ties are strongest in the distress tradition (the Altman–Edward I. axis), while cross-cluster bridges—especially via Hand and Altman—indicate regular knowledge transfer from classical statistics to scoring and distress applications. The structure suggests a division of labor: theory and measurement live together; application spaces (distress, scoring) pull ideas in; and a set of bridge authors keep the conversation integrated.

The coupling network is dense with high local closure and low modularity—signatures of a shared reference toolkit cutting across themes (Figure 5; Table 7). The designated core nodes—Atasoy (2023) for predictive analytics, Gao (2021) for fintech integration, and Nwafor (2022) for explainable ML—do indeed sit near the center of their neighborhoods and connect strongly across theme boundaries. Practically, this means the field is not siloed: predictive pipelines, fintech deployment, and explainability share many of the same foundations, so contributions that speak to more than one of these concerns will naturally couple more broadly and travel farther.

The keyword network reproduces the same three-way structure: core AI/algorithms (Blue), ML-for-fintech/risk (Red), and hybrid scoring models (Green). Within-cluster cohesion is strong, but several terms act as brokers—e.g., machine learning linking to credit scoring/regression, optimization linking to risk management. The density map pinpoints two hot zones: (i) an applied hub around machine learning, risk management, and fintech; and (ii) a modeling hub around credit scoring, regression, and SVM, with a warm ridge back to theoretical AI. In other words, the language of methods and the language of deployment increasingly co-occur, reflecting a mature cycle where models are not only built but also governed and monitored.

Theoretical clarity. Results show that the field's intellectual core is not a single thread but a braid: predictive accuracy, interpretability/accountability, and operational deployment. The co-citation and coupling evidence demonstrates that these strands are mutually reinforcing rather than sequential—methods diffuse into applications, and governance pressures circle back to shape method choice.

Methodological guidance. Given the concentration of influence and the cross-theme coupling, future studies should (i) report accuracy and stability/drift/fairness in the same design; (ii) prefer transparent or post-hoc-explainable models where decision stakes are high; and (iii) benchmark against both statistical baselines and modern ensembles to speak to the field's split heritage. Studies that publish reference lists and code/data in reusable form will also maximize coupling and downstream uptake.

Managerial and policy relevance. The country results suggest different entry points: large systems can lead to infrastructure and benchmarking; smaller contributors can differentiate through niche problems (e.g., cross-border risk, sector-specific scoring) or by piloting governance toolchains (monitoring, bias audits). Given the citation dominance of interpretability and predictive work, operational teams should treat model risk management—monitoring, documentation, and explainability—not as afterthoughts but as part of the design brief.

Data coverage and measurement. Period counts and country shares are robust at the aggregate level, but any year-by-year reconstruction can imprint small artefacts in the most recent years (e.g., indexing lags). Country attribution can under- or over-count multi-affiliation papers. Keyword-based clustering depends on term normalization; alternative stemming or phrase merging may slightly shift local structures. Still, the broad conclusions—surge after 2016, three-hub geography, steep citation skew, three-cluster intellectual map—remain intact under reasonable perturbations.

Generalizability. Low modularity in coupling and the strong cross-cluster bridges imply that results should transport across adjacent problems (e.g., SME lending vs. consumer credit), but replication in sector-specific corpora would calibrate expectations about domain drift. Future work can add longitudinal coupling (rolling windows), venue effects, and grant/patent linkages to trace how methods move from articles to deployments.

Your corpus documents a field that has accelerated sharply, consolidated around a few highly influential anchors, and converged on an integrated agenda: accurate models that are transparent enough to govern and robust enough to deploy at fintech scale. The geography of contributions shows leadership clustered but not closed; the network views—authors, documents, and keywords—confirm that ideas flow easily across theory, modeling, and operations. For researchers and practitioners alike, the most compelling opportunities lie at the interfaces: building and evaluating explainable, monitored scoring pipelines that meet performance, fairness, and stability requirements simultaneously.

Practical and policy implications

From a practitioner perspective, the evidence base summarized in the coupling and keyword networks indicates that AI/ML credit-risk pipelines are increasingly evaluated not only on predictive accuracy but also on operational controls. Accordingly, lenders should complement performance metrics (e.g., AUC/KS) with model-risk controls such as drift monitoring (e.g., population stability checks and periodic recalibration), bias and fairness diagnostics, and transparent documentation of feature governance and decision logic. In operational settings, these controls are typically implemented through monitoring dashboards that track stability, performance decay, and group-level outcomes over time.

From a policy perspective, emerging AI governance expectations for credit decisioning reinforce the need for explainability and auditability, particularly where automated outputs influence approval, pricing, or restructuring terms. The European Union's AI regulatory framework is frequently discussed in this context because creditworthiness assessment systems are generally treated as high-impact applications; institutions should therefore ensure traceable model development, periodic validation, and human oversight mechanisms appropriate to the use case.

Adoption patterns also differ by economic context. High-income markets often emphasize governance maturity (model validation, documentation standards, and explainability), while developing economies frequently prioritize expanded access and alternative-data innovation under resource constraints. These differences suggest that future deployments—and evaluations—should explicitly report the institutional setting and regulatory environment to support comparable, generalizable conclusions.

Research gaps and future agenda

Despite rapid growth, several gaps remain. First, relatively few studies jointly evaluate accuracy, stability (drift), and fairness using shared benchmarks, limiting comparability across methods and settings. Second, loan restructuring applications are less consistently operationalized than credit scoring, particularly regarding post-restructuring monitoring and outcomes. Third, more longitudinal designs are needed to examine how model performance and bias evolve under changing macroeconomic conditions and policy constraints. Future research should therefore prioritize transparent reporting standards, multi-database validation, and evaluation protocols that integrate performance, robustness, and accountability.

6. Conclusion

This review shows a field moving from experimentation to execution. Publications accelerate sharply after 2016 and peak in 2024, while contributions cluster geographically around India, China, and the United States, with meaningful activity elsewhere. The citation landscape is highly skewed toward work that joins predictive performance with transparency and accountability, and the network views (co-citation, coupling, keywords) reveal an integrated arc from theory to methods to deployment. Two practical priorities follow: build credit-risk pipelines that embed explainability, monitoring, and fairness alongside accuracy; and test portability across fintech settings to manage drift and governance needs. Future studies should report stability and bias metrics with the same weight as accuracy, publish reusable assets (data/code), and position contributions at the interfaces where predictive analytics, fintech integration, and explainable ML meet.

References

[1] Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007.

[2] Datta, A., Tschantz, M. C., & Datta, A. (2016). Algorithmic transparency via quantitative input influence. *Proceedings of the 37th IEEE Symposium on Security and Privacy (SP)*, 598–617. https://doi.org/10.1109/SP.2016.42.

[3] Ma, X., Liu, Y., Zhang, H., Ding, R., Hong, Y., Wang, S., & Wang, H. (2018). P2P lending default prediction using a deep learning approach. *Electronic Commerce Research and Applications, 32*, 100836.

[4] Zhu, Y., Xie, M., & Wang, H. (2013). Credit decision making with C-TOPSIS under information entropy. *Knowledge-Based Systems, 52*, 221–230. https://doi.org/10.1016/j.knosys.2013.07.001.

[5] Ignatius, J., Rahman, A., & Yazdani, M. (2018). Fuzzy decision support model for credit scoring. *Neural Computing and Applications, 30*(5), 1587–1601.

[6] van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3.

[7] Hagberg, A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy 2008)*, 11–15. https://doi.org/10.25080/TCWV9851.

[8] R Core Team. (2023). *R: A language and environment for statistical computing (Version 4.3.1).* Vienna: R Foundation for Statistical Computing. https://www.R-project.org.

[9] Hari, D. S., Kittur, D. C., Policepatil, D. S., Rani, B. S., Jwalapuram, S., Rajshree, S., & Basha, S. M. (2025). Digital Transformation in Banking: Assessing the Impact of Technological Innovation on the Performance of Public Sector Banks in South India. *International Journal of Accounting and Economics Studies*, *12*(SI-1), 363-370. https://doi.org/10.14419/94v7eh76.

[10] Python Software Foundation. (2024). *Python (Version 3.11).* https://www.python.org.

[11] Elsevier. (2024). *Scopus Content Coverage Guide.* Retrieved from https://www.elsevier.com/solutions/scopus.

[12] Microsoft Corporation. (2021). *Microsoft Excel (Version 2021).* Redmond, WA: Microsoft Inc.

[13] Verma, M. S. S., Sreenu, D. N., & Kushwaha, D. G. S. (2025). Banks with Higher Levels of Fintech Adoption Exhibit Better Financial Performance Than Those with Lower Levels of Adoption. *International Journal of Accounting and Economics Studies*, *12*(6), 774-784. https://doi.org/10.14419/jkdnj577.

[14] Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2020). Explainable AI in credit risk management. *Computational Economics, 57*(1), 203–216. https://doi.org/10.1007/s10614-020-10042-0

[15] Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030.

[16] Pauldurai, T. A. I., Thangam, D. D. M. V., R, D. S., & Meenakshi, D. J. R. (2025). Rescuing Banks in Financial Distress: A Case Study of Yes Bank Using the EAGLES Model. *International Journal of Accounting and Economics Studies*, *12*(5), 420-431. https://doi.org/10.14419/em5q1q81.

[17] Li, Y., Luo, C., Xu, X., & Zhang, H. (2022). Explainable machine learning for credit scoring: An empirical study. *Expert Systems with Applications, 205*, 117636.

[18] Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 160*(3), 523–541. https://doi.org/10.1111/j.1467-985X.1997.00078.x.

[19] Alam, A., Azizah, N. R., Saputro, T. A., Hanifudin, I., & Nordin, N. (2025). Islamic Bank Risk Studies: A Bibliometric Review. *International Journal of Accounting and Economics Studies*, *12*(2), 445-456. https://doi.org/10.14419/nm80n608

[20] Agrawal, D. V. (2025). FinTech Disruption in Traditional Banking: Risks and Opportunities. *International Journal of Accounting and Economics Studies*, *12*(4), 98-104. https://doi.org/10.14419/hnmqb776.