Review Article

Explaining the Unexplainable: A Systematic Review of Explainable AI in Finance

Md Talha Mohsin[1], Nabid Bin Nasim[2]

1. University of Tulsa, United States; 2. University of Dhaka, Bangladesh

Practitioners and researchers trying to strike a balance between accuracy and transparency center Explainable Artificial Intelligence (XAI) at the junction of finance. This paper offers a thorough overview of the changing scene of XAI applications in finance, together with domain-specific implementations, methodological developments, and trend mapping of research. Using bibliometric and content analysis, we find topic clusters, significant research, and most often used explainability strategies in the financial industries. Our results show a substantial dependence on post-hoc interpretability techniques; attention mechanisms, feature importance analysis, and SHAP are the most often used techniques among them. This review stresses the need for multidisciplinary approaches combining financial knowledge with improved explainability paradigms and exposes important shortcomings in present XAI systems.

1. Introduction

The widespread adoption of artificial intelligence (AI) has caused people to reassess their own positions[1]. Thanks to significant developments in processor capacity and advancements in optimization methods, we are seeing widespread acceptance of automated decision-making by AI[2]. Because of these AI systems' popularity, they have also received significant attention in recent popular press[3][4][5][6][7][8]. Through the use of modern computational approaches, decision-making processes in a variety of financial industries have also been revolutionized.

In the field of finance, AI has surprised everyone[9]. Both the large amount of data that is available and the increasing reliance on machine learning (ML) models have impacted the world of finance substantially. Specifically, the growing outflow of data created by consumers, investors, businesses, and governments has helped AI transform the financial sector[10]. Also, the professionals in the finance sector are increasingly fascinated by "alternative data" outside the scope of macroeconomic indicators, securities pricing, and basic business knowledge, which include satellite imagery, news stories, phone records, and social media posts, which require more understanding and have AI implications[11]. AI is not only a phase; rather, it is a presence that is fundamentally changing how companies are handling their assets and finances.

Despite their remarkable success in producing high-performance models, many of these AI and ML techniques have come under fire for their lack of transparency and interpretability. In many scenarios, machine learning models acquire great accuracy, usually at the price of poor explainability[12]. The more complicated these models get, the more often their decision-making processes remain opaque, which creates a barrier for stakeholders trying to understand why and how a particular choice was taken. For applications where sophisticated ML models are part of how stakeholders come to make decisions, giving user-centered explanations is very important[13]. Especially in the financial sector, where decision-makers, including investors, authorities, and financial analysts, depend on not only accurate data but also the justification for these to guarantee informed, responsible, and accountable actions.

 doi.org/10.32388/PUO5CK

Reacting to this problem of interpretability, explainable artificial intelligence (XAI) has become a major focus of study. Although expert Systems researchers have used explanation approaches before[14][15][16][17], the term explainable AI (XAI) was coined by DARPA[18]. Explainability can be thought of as an active feature of a model, referring to any action or process that a model takes to elucidate its internal operations[19]. The quality of "being an explanation" is an interaction rather than a characteristic of assertions[20]. Explanation techniques have advanced significantly as XAI is now acknowledged as a must rather than merely an option[21]. What the user requires, what information they already possess, and most importantly, their objectives, determines what constitutes an explanation. This raises the question of why a particular user needs an explanation, which brings up the function and context of the AI system (software, algorithm, tool). By integrating interpretability restrictions into the model's structure, XAI approaches make the models easier to understand[22].

Although scholars have paid more and more attention to the adoption of AI technologies in a wide spectrum of financial applications in recent years, the existing literature is rather broad and heterogeneous in terms of research questions, level of analysis, and method, making it difficult to draw solid conclusions and to understand which research areas require further investigation[10]. Despite the fact that a large number of machine learning interpretability research and approaches have been established in academia, they hardly ever make up a significant portion of machine learning pipelines and workflows[23]. As such, focusing on its theoretical foundations, approaches, and applications across many financial sectors, this review article attempts to give a thorough understanding of the present situation of XAI in finance. By combining current research, this paper highlights the most significant XAI approaches and their contributions to improve model transparency and interpretability.

We examined 323 papers from journals listed in the Scopus database published between 2015 and 2025 using a bibliometric approach. We applied co-authorship networks, trend analysis, and keyword frequency analysis to understand the current scenarios. Then, based on their significance and influence in the field, we chose thirty papers from this corpus for additional study. We divided these 30 papers into seven streams: risk management, fraud detection, time series forecasting, financial analysis and decision-making, credit evaluation and scoring, financial modeling and prediction, and trading and investment. We conducted citation analysis and co-citation analysis among several bibliometric approaches to understand the current situation. Our evaluation aims to underline the advantages and drawbacks of present XAI approaches as well as the limitations to increase AI system justice, responsibility, and openness.

The remainder of this study is structured as follows: Section 2 discusses background and context, Section 3 outlines the data and methodology, Section 4 presents the findings and discussion, and Section 5 concludes the study.

2. Background and Context

Artificial Intelligence (AI) has gained significant traction in recent years and, with the right application, might meet or surpass expectations in a wide range of application areas[19]. However, by today's standards, it is not advisable to blindly trust the output of AI because of the significant influence of adversarial instances, trustability, and data bias in machine learning[24]. Even when we understand the mathematical underpinnings of machine learning (ML) architectures, it is frequently impossible to gain insight into how the models operate internally[25]; explicit modeling and reasoning techniques are required to clarify how and why a particular outcome was obtained. That's where Explainable Artificial Intelligence (XAI) comes in. Explainability is the ability of an interested stakeholder to understand the major reasons for a model-driven decision[26]. XAI is a component of the third wave of AI[27], a new generation of AI technologies whose goal is to create algorithms that can accurately explain themselves. It entails the capacity of an artificial intelligence system to provide coherent and easily accessible explanations for decisions and actions it makes[28]. XAI makes the end user, who relies on decisions, suggestions, or actions made by an AI system, the focus, as that person must comprehend the reasoning behind the system[18].

A machine learning model's explainability and prediction accuracy are typically inversely correlated[29]; the better the prediction accuracy, the lower the model's explainability. XAI aims to create approaches that, without compromising their prediction accuracy, increase the transparency and interpretability of artificial intelligence and machine learning models. Clear descriptions of model outputs help XAI practitioners and users to understand the underlying processes and justifications for AI-driven judgments. Explainability can be viewed as an active characteristic of a model, denoting any action or procedure taken by a model with the intent of clarifying or detailing its internal functions. Understandability stands out as the most crucial idea in XAI[19], which is closely related to both interpretability and transparency: interpretability gauges how well a human can comprehend a model's choice, whereas transparency describes a model's ability to be intelligible by a human on its own. Understandability and comprehensibility are related in that both depend on the audience's capacity to comprehend the information presented in the model.

In order to cover every facet of XAI, concentrate on providing answers to the questions "What, Why, and How"[19][24][30]. The "What" aims to clarify the current meanings of explainable AI and the significance of elucidating the function of a user. The "Why" gives a summary of the main objectives of XAI research, such as establishing credibility, fulfilling legal obligations, detecting bias, guaranteeing generalization in AI models, and debugging. The "How" section looks at how to achieve explainability prior to the modeling phase, including how to fully comprehend and record the datasets used in modeling. The "How" section examines the methods for attaining explainability before the modeling stage, including techniques for thoroughly understanding and documenting the datasets utilized in modeling **(Fig 1)**. XAI seeks to explain how AI models make decisions in a way that people can grasp, therefore increasing their transparency.

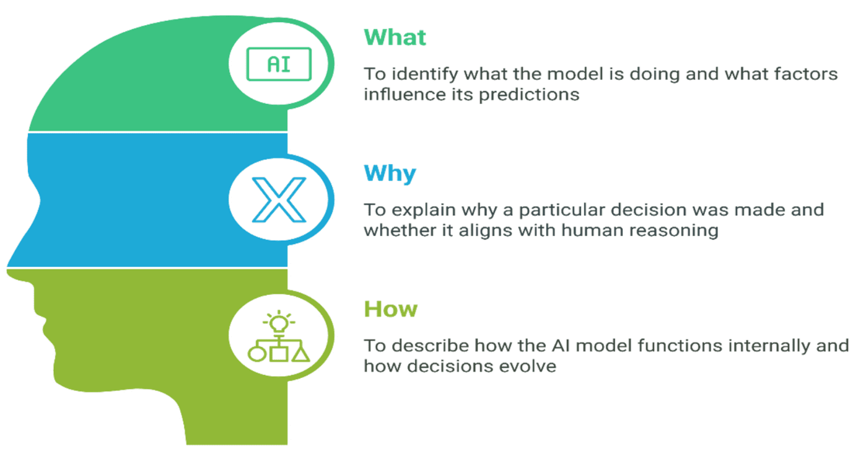

Figure 1. Core aspects of AI interpretability

XAI comprises a wide range of strategies and methodologies that can be broadly classified into two types: post-hoc and ante-hoc explainability. Post-hoc explainability refers to methods used after a model has been trained to provide reasons for its predictions, whereas ante-hoc explainability entails creating intrinsically interpretable models which are transparent by definition and produce interpretable outcomes directly **(Fig 2)**.

 doi.org/10.32388/PUO5CK

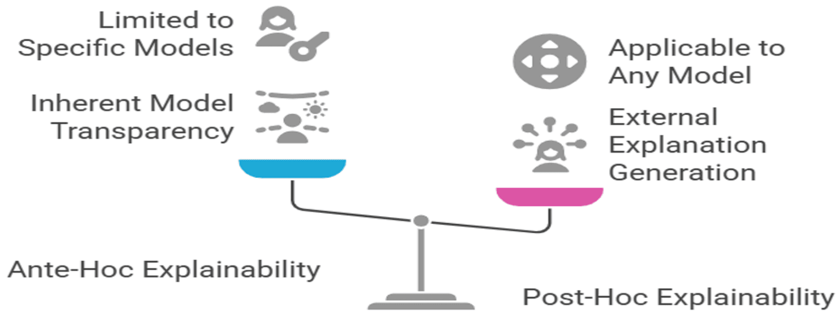

Figure 2. Ante-Hoc vs. Post-Hoc Explainability

Also, there are two types of machine learning models based on their understandability: white-box models and black-box models. These models are a way to group AI models that affect the need for XAI methods. Then, XAI techniques are used to figure out what these models mean. The difference between white-box and black-box models is directly linked to the difference between post-hoc and ante-hoc explainability. Ante-hoc explainability fits well with white-box models because these models are meant to be interpretable and give information about how they make decisions without needing to be interpreted further. These are better when trust and openness are very important. On the one hand, there are "black-box" models, such as deep learning[31] and ensembles[32][33][34]. Especially those based on deep learning are frequently thought of as "black-box" models[35][24][18][36]. These black-box models need post-hoc explainability tools to help us understand how they make decisions when they aren't being clear. (**Fig 3)** depicts the decision-making process when selecting the appropriate model type for XAI.

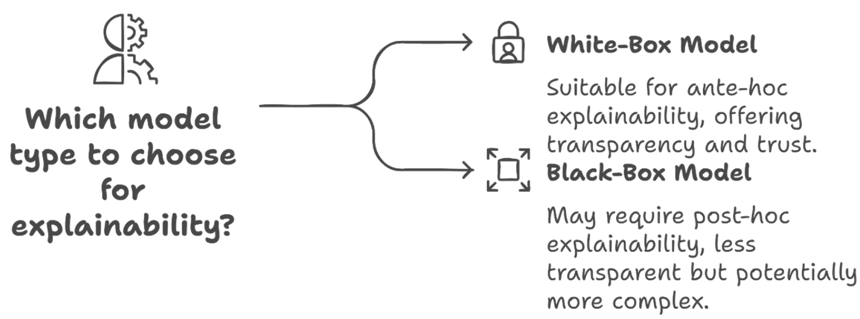

Figure 3. Choosing Between White-Box and Black-Box Models for Explainability

XAI helps release the power of machine learning algorithms to challenge the black-box character[37]. White box models are transparent and very simple to comprehend, while black box models are opaque[36]. Since they enable improved quality assurance of black box ML models, XAI tools are a significant addition to the data science toolkit, especially to understand the characteristics of the data collection and domain expertise helps them to complement other facets of quality assurance, including several approaches of model performance testing[38].

Although modern machine learning models are more easily available than ever, it has been difficult to design and implement systems that support real-world financial applications[39]. This is mostly due to their lack of transparency and explainability—both

doi.org/10.32388/PUO5CK

of which are fundamental for establishing trustworthy technology. In the financial domain, justification for the choice of action is just as important as the result itself. XAI's ability to advance justice and lower bias in artificial intelligence systems is among its most significant advantages for the financial industry. For example, XAI techniques can help find the factors underlying the discrimination if a credit score model is unfairly excluding particular demographic groups, thereby enabling corrective action. Maintaining ethical AI principles depends on this, especially in lending where biased models might cause financial exclusion. By giving financial analysts, investors, and officials better knowledge of how artificial intelligence models operate, XAI also helps with decision-making. XAI technology can find economic indicators, market data, or historical trends driving model forecasts, whether used for stock price prediction or portfolio management. This not only boosts model confidence but also lets financial decision-makers act with certainty based on these realizations. XAI can assist in stock price forecasting, for instance, by dissecting the pertinent elements—such as earnings reports, market sentiment, or macroeconomic data—thereby enabling an investor to understand why a given model forecasts an increase in stock price.

Apart from the complexity of financial models, the demand for XAI in finance results from rising expectations for openness from authorities and the public overall. Financial institutions—including banks, asset managers, and insurance firms—are expected more and more to justify their AI-driven decisions to stakeholders. Explainable artificial intelligence models—which include specifics or justifications to make the operation of artificial intelligence clear or simple—are essential.

3. Data and Methodology

3.1. Data Collection

We systematically searched the Scopus database using a well-chosen set of keywords combining finance and Explainable AI (XAI) to identify the relevant papers. We then carefully examined their titles, abstracts, and keywords to ensure they matched our screening criteria. To ensure the inclusion of recent developments in XAI and its applications in finance, we limited the search for this study to English-language publications published between 2015 and the present.

3.2. Identification of Keywords

The keyword selection procedure for this study was intended to capture the most relevant material on the intersection of Explainable AI (XAI) and finance while minimizing irrelevant results **(Table 1)**. The keyword identification involved a thorough investigation of AI and machine learning (ML) glossaries, dictionaries, and academic literature to discover fundamental terminologies used in the area. XAI-related research regularly uses terminology like "explainable AI," "interpretable machine learning," "post-hoc explanations," "transparency," and "model interpretability." Key terms pertinent to the financial domain were also selected, such as "finance," "financial systems," "fraud detection," "credit scoring," "portfolio optimization," and "stock price prediction." This methodical methodology enabled a targeted retrieval procedure to obtain papers relevant to the research topic.

Domain	Keywords
XAI	"explainable AI" OR XAI OR "interpretable machine learning" OR interpretability OR "explainable artificial intelligence" OR "model interpretability" OR "post hoc explanations" OR "feature attribution" OR transparency OR explainability OR "Attention Mechanisms" OR "Attention Mechanism" OR SHAP OR LIME OR "Integrated Gradients" OR "Counterfactual Explanations" OR "Saliency Maps" OR "Feature Importance" OR "Partial Dependence Plots" OR "Individual Conditional Expectation" OR "Explainable Neural Networks"
Finance	finance OR stock OR "stock price prediction" OR "stock price forecasting" OR "fraud detection" OR "credit scoring" OR loan OR "risk management" OR "risk assessment" OR "Credit risk assessment" OR "portfolio optimization" OR "portfolio management" OR "algorithmic trading" OR "wealth management" OR "financial forecasting" OR ESG OR "investment analysis" OR "personal finance" OR cryptocurrency OR "financial systems" OR "financial technology" OR fintech OR "Regulatory Compliance" OR "Sentiment Analysis for Financial Markets" OR "financial markets" OR "banking" OR "insurance" OR "asset management" OR "derivatives pricing"

Table 1. Query Terms

3.3. Refinement and Search Optimization

The initial keyword query returned 6,086 papers. Given the vast dataset, we used a multi-stage refining approach to improve specificity. First, a strong exclusion criterion was applied to eliminate papers that were outside the scope of the inquiry. Articles were removed if they addressed finance without XAI, discussed XAI but lacked financial applications, or covered AI in finance but omitted explainability aspects.

Secondly, we assessed each of these documents by title, abstract, and keywords. In cases where the relevance was uncertain, further sections such as the introduction and conclusion were reviewed. This stage resulted in the removal of 5,763 publications that did not fit the research objectives. Thirdly, only papers with at least 50 Scopus citations were retained, ensuring the inclusion of well-known and influential studies. This restricted the dataset to 323 articles for further refinement.

In the final screening phase, full-text assessments of these 323 publications were conducted to assess their methodological contributions, application to financial decision-making, and alignment with XAI frameworks. Finally, 30 papers were identified as meeting the inclusion criteria and study objectives in their entirety.

4. Findings/Discussion

4.1. General Information

Our initial dataset contained 6,086 research papers taken from the Scopus database. The search criteria concentrated on the junction of Explainable AI (XAI) and finance. Given the large extent of the dataset, a thorough selection and filtering method produced 323 papers. **(Fig 4)** illustrates the keyword co-occurrence network of the principal research topics and their interrelations to these papers. This network emphasizes the most prevalent keywords, showcasing prominent subfields and emerging trends in the literature. Each of the keywords is represented by a circle, and the size of the circle denotes the weight of the keyword. The network is segmented into different clusters, underscoring the thematic domains in the literature.

doi.org/10.32388/PUO5CK

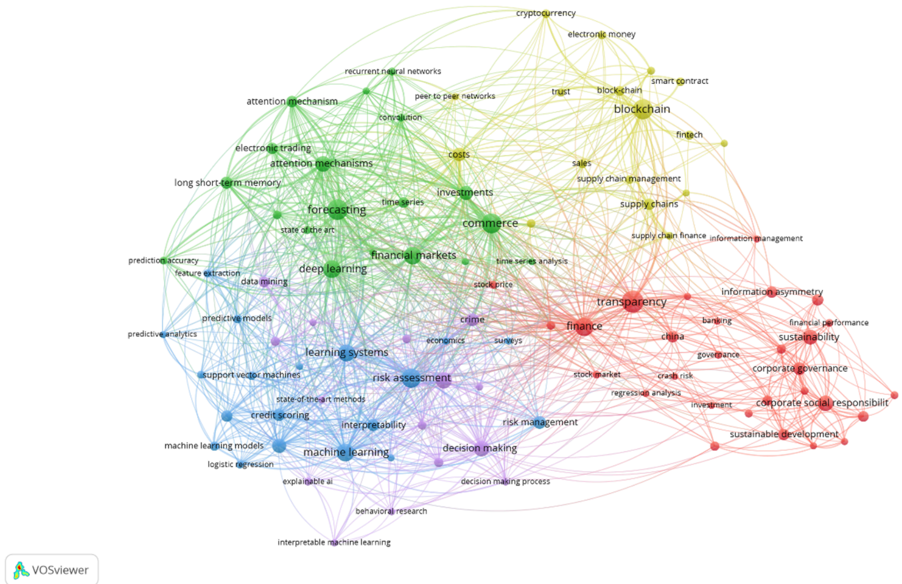

Figure 4. Keyword Co-occurrence Network

(Fig 5) denotes the co-authorship network, visualizing a structural overview of collaboration trends in the field of Explainable AI (XAI) in finance, demonstrating the extent to which researchers collaborate on similar subjects. The network is composed of comparatively small, unrelated clusters, which suggest that research in this field is still relatively niche and dispersed.

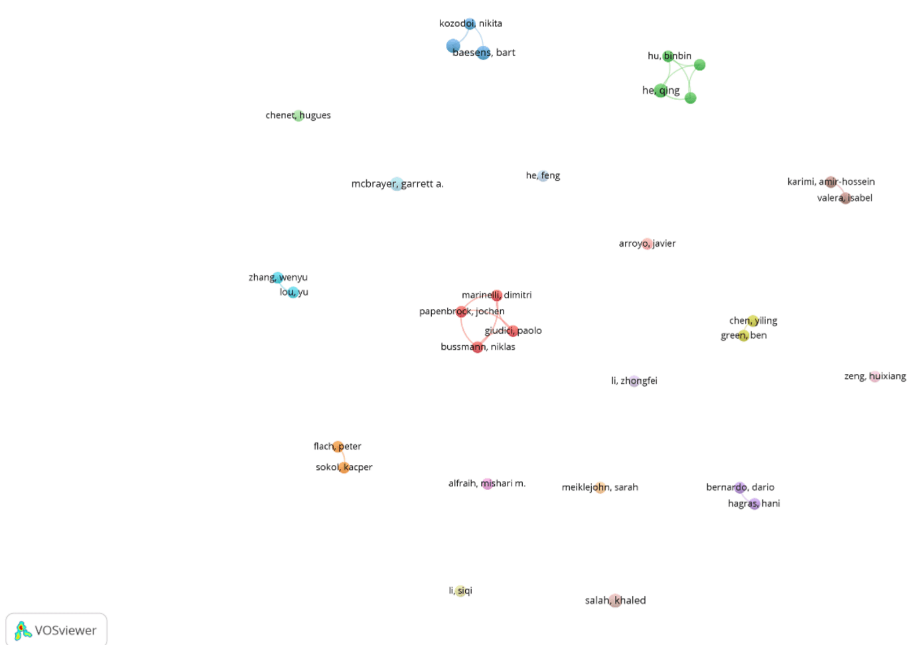

Figure 5. Co-authorship Network

4.2. Citation Analysis

After further refinement through methodological contribution, journal repute, and citation numbers, the collection was reduced to 30 papers, those judged most pertinent and influential for our research. We split them into seven groups of the financial domain: risk management, fraud detection, time series forecasting, financial analysis and decision-making, credit evaluation and scoring, financial modeling and prediction, and trading and investment **(Table 2)**.

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction	[40]	2017	Time Series Forecasting		Dual-Stage Attention-Based RNN (DA-RNN), LSTM, GRU	Input Attention, Temporal Attention	RMSE, MAE, MAPE	arXiv preprint
A Boosted Decision Tree Approach Using Bayesian Hyper-parameter Optimization for Credit Scoring	[41]	2017	Credit Evaluation and Scoring	Credit scoring, Boosted decision tree, Bayesian hyper-parameter optimization	Extreme Gradient Boosting (XGBoost), Bayesian Optimization	Feature Importance Scores, Decision Chart	Accuracy, Error Rate, AUC-H Measure, Brier Score	Expert Systems With Applications
APATE: A Novel Approach for Automated Credit Card Transaction Fraud Detection using Network-Based Extensions	[42]	2015	Fraud Detection	Credit card transaction fraud, network analysis, bipartite graphs, supervised learning	Logistic Regression, Neural Networks, Random Forest	Feature Importance Analysis	AUC, Accuracy, Balanced Accuracy	Decision Support Systems
Stock closing price prediction based on sentiment analysis and LSTM	[43]	2020	Trading and Investment	Stock market prediction, Long short-term memory, Attention mechanism, Empirical mode decomposition	LSTM, CNN	Attention Mechanism	MAE, RMSE, MAPE, R^2, Granger Causality	Neural Computing and Applications

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
Explainable Machine Learning in Credit Risk Management	[26]	2021	Risk Management	Credit risk management, Explainable AI, Financial technologies, Similarity networks	XGBoost, Logistic Regression	Shapley Values, Correlation Network Models, TreeSHAP	AUC, Misclassification Rate, Receiver Operating Characteristics (ROC)	Computational Economics
Machine Learning for Credit Scoring: Improving Logistic Regression with Non-Linear Decision-Tree Effects	[44]	2022	Credit Evaluation and Scoring	Risk management, Credit scoring, Machine learning, Interpretability, Econo-metrics.	Penalised Logistic Tree Regression (PLTR), Random Forest, Support Vector Machine (SVM), Neural Networks (NN)	Feature Engineering with Short-Depth Decision Trees, Adaptive Lasso	AUC, Brier Score, Kolmogorov-Smirnov Statistic (KS), Percentage of Correct Classification (PCC), Partial Gini Index (PGI)	European Journal of Operational Research
Multiobjective Evolution of Fuzzy Rough Neural Network via Distributed Parallelism for Stock Prediction	[45]	2020	Trading and Investment	Distributed parallelism, evolutionary neural network, fuzzy rough neural network (FRNN), multiobjective evolution, stock price prediction	Fuzzy Rough Neural Network (FRNN), Multiobjective Evolutionary Algorithm (MOEA)	Interval type-2 fuzzy neurons, IF-THEN rules-based interpretability	Prediction Precision (MAE, RMSE), Network Simplicity	IEEE Transactions on Fuzzy Systems
Financial time series forecasting with multi-modality graph neural network	[46]	2022	Time Series Forecasting	Graph neural network, Graph attention, Deep learning, Quantitative investment	Multi-modality Graph Neural Network (MAGNN)	Two-phase Attention Mechanism (Inner-Modality and Inter-Modality Attention)	Micro-F1, Macro-F1, Weighted-F1, Accumulated Return, Daily Return, Sharpe Ratio	Pattern Recognition
Disparate Interactions:	[47]	2019	Risk Management	Fairness, risk assessment,	Gradient Boosted Trees	Algorithm-in-the-Loop	Area Under Curve (AUC), False Positive	FAT* ’19 Conference on

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments				behavioral experiment, Mechanical Turk		Framework, Fairness Analysis	Rate, Brier Score	Fairness, Accountability, and Transparency
Credit Risk Analysis Using Machine and Deep Learning Models	[48]	2018	Risk Management	Credit risk, financial regulation, data science, Big Data, deep learning	Elastic Net, Random Forest, Gradient Boosting Machine (GBM), Neural Network (Deep Learning)	Feature Importance Analysis, ROC Analysis	AUC (Area Under Curve), Root Mean Square Error (RMSE), Gini Index	Risks
Temporal Attention-Augmented Bilinear Network for Financial Time-Series Data Analysis	[49]	2018	Time Series Forecasting	Bilinear projection, feedforward neural network, financial data analysis, temporal attention, time-series prediction.	Bilinear Network, Temporal Attention Mechanism	Attention Mechanism for Temporal Interpretability	Micro-F1, Macro-F1, Weighted-F1, Computational Complexity, Training Time	IEEE Transactions on Neural Networks and Learning Systems
A Benchmark of Machine Learning Approaches for Credit Score Prediction	[50]	2021	Credit Evaluation and Scoring	Credit score prediction, Benchmark, Supervised learning, Machine learning, Explainable artificial intelligence	Logistic Regression, Random Forest, Multi-layer Perceptron (MLP)	LIME, SHAP, Anchors, BEEF, LORE	AUC, Sensitivity (TPR), Specificity (TNR), Accuracy (ACC), G-Mean, False Positive Rate	Expert Systems with Applications
[51]		2023	Financial Analysis and	Deep learning, large language model, transfer	BERT, FinBERT, Convolutional	Feature Importance Analysis,	Accuracy, Precision, Recall, F1 Score, AUC	Contemporary Accounting Research

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
			Decision-Making	learning, interpretable machine learning, sentiment classification, environment, social, and governance (ESG)	Neural Network (CNN), Long Short-Term Memory (LSTM), Support Vector Machine (SVM), Random Forest (RF)	Interpretable Machine Learning		
A Compact Evolutionary Interval-Valued Fuzzy Rule-Based Classification System for the Modeling and Prediction of Real-World Financial Applications With Imbalanced Data	[52]	2014	Financial Modeling and Prediction	Evolutionary algorithms, financial applications, interval-valued fuzzy rule-based classification systems, intervalvalued fuzzy sets (IVFSs).	Interval-Valued Fuzzy Rule-Based Classification System (IVTURSFARC-HD), Evolutionary Algorithm	Interpretability through Fuzzy Rules, Rule Weight Rescaling Method	Geometric Mean (GM), True Positive Rate (TPR), True Negative Rate (TNR), Accuracy	IEEE Transactions on Fuzzy Systems
Hierarchical Multi-Scale Gaussian Transformer for Stock Movement Prediction	[53]	2020	Trading and Investment	Transformer, Multi-Scale Gaussian Prior, Orthogonal Regularization, Trading Gap Splitter, Stock Movement Prediction	Transformer, Multi-Scale Gaussian Transformer	Attention Mechanism for Hierarchical Interpretability	Accuracy, Matthews Correlation Coefficient (MCC)	International Joint Conference on Artificial Intelligence (IJCAI-20)
An Explainable AI Decision-Support System to	[2]	2020	Credit Evaluation and Scoring	Explainable artificial intelligence, Interpretable	Gradient Boosting Models (GBM),	SHAP, ICE	Loan Approval Rates, Model Fairness	Expert Systems with Applications

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
Automate Loan Underwriting				machine learning, Loan underwriting, Evidential reasoning, Belief-rule-base Automated decision making	Random Forest			
Deep Learning for Detecting Financial Statement Fraud	[54]	2020	Fraud Detection	Fraud detection, Financial statements, Deep learning, Text analytics	Hierarchical Attention Network (HAN), LSTM, Random Forest, XGBoost, GPT-2	Attention Mechanism, LIME (Local Interpretable Model-Agnostic Explanations)	AUC, Sensitivity, Specificity, F1-Score, F2-Score, Accuracy	Decision Support Systems
Stock market index prediction using deep Transformer model	[55]	2022	Trading and Investment	Deep learning, Transformer, Stock index prediction	Transformer, CNN, RNN, LSTM	Attention Mechanism	Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), Sharpe Ratio, Volatility, Max Drawdown	Expert Systems With Applications
Exploring the Attention Mechanism in LSTM-based Hong Kong Stock Price Movement Prediction	[56]	2019	Trading and Investment	LSTM, Stock price, Prediction	Long Short-Term Memory (LSTM), Attention-based LSTM	Attention Mechanism	Accuracy, Precision, Recall, F1 Score	Quantitative Finance
A Deep Learning Approach for Credit Scoring of Peer-to-Peer Lending Using	[57]	2018	Credit Evaluation and Scoring	P2P lending, credit scoring, machine learning, deep learning, LSTM,	LSTM, Bi-directional LSTM (BLSTM), Attention Mechanism	Attention Mechanism for Interpretability	ROC Curve, AUC (Area Under Curve), Kolmogorov-Smirnov (KS) Statistic	IEEE Access

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
Attention Mechanism LSTM				attention mechanism				
Financial Defaulter Detection on Online Credit Payment via Multi-view Attributed Heterogeneous Information Network	[58]	2020	Fraud Detection	Financial defaulter detection, Multi-view attributed heterogeneous information network, Meta-path encoder	Meta-path based Path Encoder, Neural Networks, Attention Mechanism, Softmax Classifier	Attention Mechanism for Node, Link, and Meta-path Level Interpretability	AUC (Area Under the Curve), Recall@Precision=0.1	The Web Conference (WWW) 2020
Multiobjective Evolutionary Optimization of Type-2 Fuzzy Rule-Based Systems for Financial Data Classification	[59]	2017	Financial Analysis and Decision-Making	Financial datasets, multiobjective evolutionary fuzzy systems, type-2 fuzzy rule-based classifiers, unbalanced datasets.	Type-2 Fuzzy Rule-Based Classifier, Multiobjective Evolutionary Algorithm (MOEA), C4.5 Decision Tree, FURIA	Rule-Based Interpretability, Scaled Dominance Approach	Accuracy, Complexity (Number of Rules), True Positive Rate (TPR), False Positive Rate (FPR)	IEEE Transactions on Fuzzy Systems
Financial Defaulter Detection on Online Credit Payment via Multi-view Attributed Heterogeneous Information Network	[58]	2020	Fraud Detection	Cross-country study, national culture, individualism, stock price crash risk	Meta-path-based Path Encoder, Attention Mechanism, Multi-view Learning	Attention Mechanism for Interpretability	AUC (Area Under Curve), Recall@Precision=0.1	Proceedings of The Web Conference (WWW)
Explainable AI in Fintech Risk Management	[60]	2020	Risk Management	Credit risk management, explainable AI, financial technologies, peer to peer lending,	Logistic Regression, XGBoost	SHAP	Accuracy, ROC-AUC	Frontiers in Artificial Intelligence

doi.org/10.32388/PUO5CK

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
				logistic regression, predictive models				
AlphaStock: A Buying-Winners-and-Selling-Losers Investment Strategy using Interpretable Deep Reinforcement Attention Networks	[61]	2019	Trading and Investment	Quantitative Trading, Portfolio Management	Reinforcement Learning, Deep Attention Networks, Long Short-Term Memory (LSTM)	Sensitivity Analysis, Cross-Asset Attention Mechanism	Sharpe Ratio, Annualized Percentage Rate (APR), Maximum Drawdown (MDD), Calmar Ratio (CR), Downside Deviation Ratio (DDR)	KDD '19 - ACM SIGKDD Conference
Forecasting Daily Stock Trend Using Multi-Filter Feature Selection and Deep Learning	[62]	2021	Trading and Investment	Stock trend prediction, Feature selection, Deep learning, Machine learning	Logistic Regression, Support Vector Machine (SVM), Random Forest, Deep Generative Models, Variational AutoEncoder (VAE), Attention Mechanism	Attention Mechanism	Prediction Accuracy, Classification Accuracy, Sensitivity, Specificity	Expert Systems with Applications
Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment	[63]	2015	Risk Management	Ensemble strategies, Credit scoring, Decision forests, Diversity, Gradient boosting, Random forests	Correlated-Adjusted Decision Forest (CADF)	SHAP	Accuracy, RMSE, Precision	Annals of Operations Research

doi.org/10.32388/PUO5CK

Paper Name	Author(s)	Year of Publication	Category	Keywords	ML Algorithm	XAI Methods	Evaluation Metrics	Journal Name
Transformer-based Attention Network for Stock Movement Prediction	[64]	2022	Trading and Investment	Stock movement prediction, Deep learning, Transformer, Attention	Transformer, LSTM, Multihead Attention, Temporal Attention	Visualization of Attention Scores (Multihead and Temporal Attention Mechanisms)	Accuracy, Matthews Correlation Coefficient (MCC)	Expert Systems with Applications
Accurate Multivariate Stock Movement Prediction via Data-Axis Transformer with Multi-Level Contexts	[65]	2021	Trading and Investment	Stock movement prediction, transformers, attention mechanism	Data-axis Transformer with Multi-Level Contexts (DTML)	Attention Map, Temporal Attention	Accuracy, Matthews Correlation Coefficient (MCC), Investment Simulation (Annualized Return)	ACM SIGKDD Conference on Knowledge Discovery and Data Mining
Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment. A correlated-adjusted decision forest proposal	[63]	2015	Fraud Detection	Ensemble strategies, Credit scoring, Decision forests, Diversity, Gradient boosting, Random forests	Gradient Boosting, Random Forests, Decision Trees	Rule-based reasoning, Decision Trees for Interpretability	Accuracy Rate, Type I & II Error, AUC, Statistical Tests	Expert Systems with Applications

Table 2. Taxonomy of XAI Applications

(Fig 6) shows the keyword analysis, which offers substantial insights into the prevailing issues and research focal areas at the intersection of finance and Explainable AI (XAI). The Fig highlights the critical aspects of learning, explainability, prediction, credit, stock, finance, and risk management. The increasing emphasis on interpretability, particularly in complex financial models, underscores the importance of fuzzy logic, explainability, and attention mechanisms.

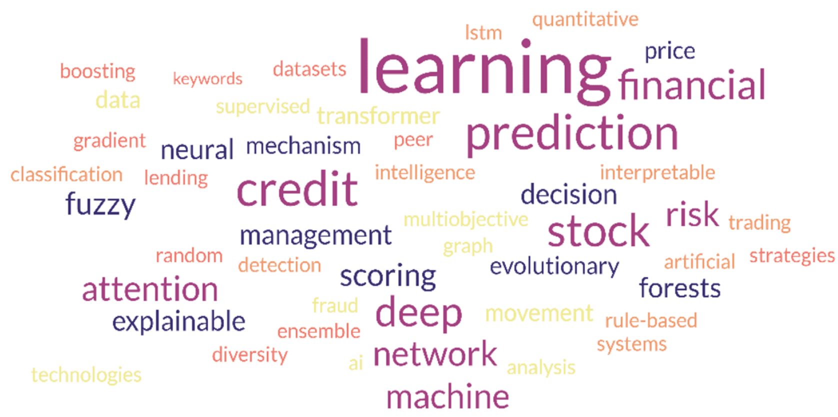

Figure 6. Keyword Analysis

The bar chart in (**Fig 7**) depicts the distribution of papers published each year. The analysis of publication trends indicates that there has been a lot more research on Explainable AI (XAI) in finance after 2019. The interest rate peaked in 2019-2020 and remained high in the years following.

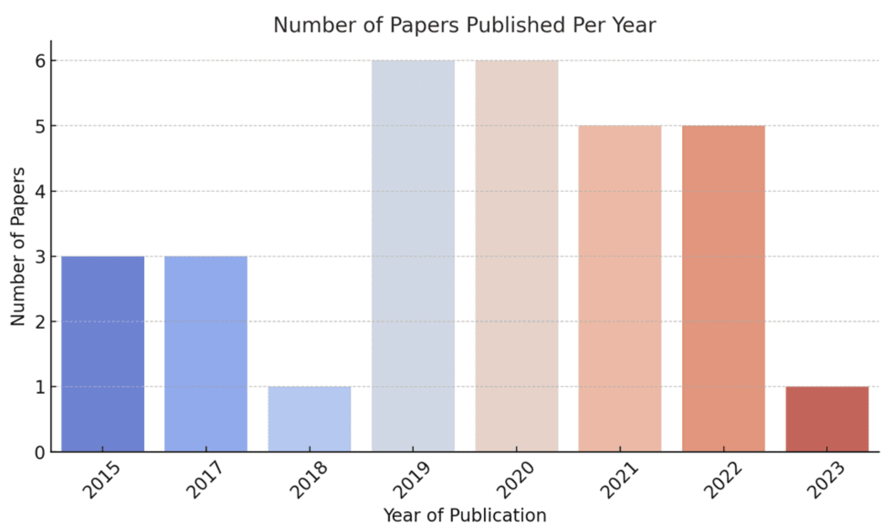

Figure 7. Annual Trend of Published Papers

The expansion of research across diverse areas of finance indicates that a lot of attention is being paid to critical areas. Trading and Investment lead the category, followed by Credit Evaluation, Fraud Detection, and Risk Management **(Fig 8)**.

doi.org/10.32388/PUO5CK

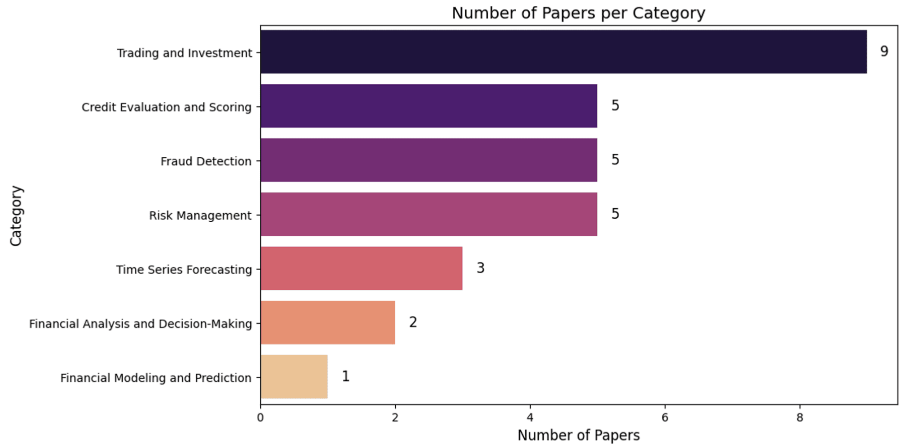

Figure 8. Distribution of Research Papers Across Financial Domains

The leading journal in the distribution of research publications is Expert Systems with Applications, followed by IEEE Transactions on Fuzzy Systems and Decision Support Systems. Top conferences such as IJCAI, KDD, and WWW highlight the increasing scholarly and commercial attention on XAI-driven financial applications **(Fig 9)**.

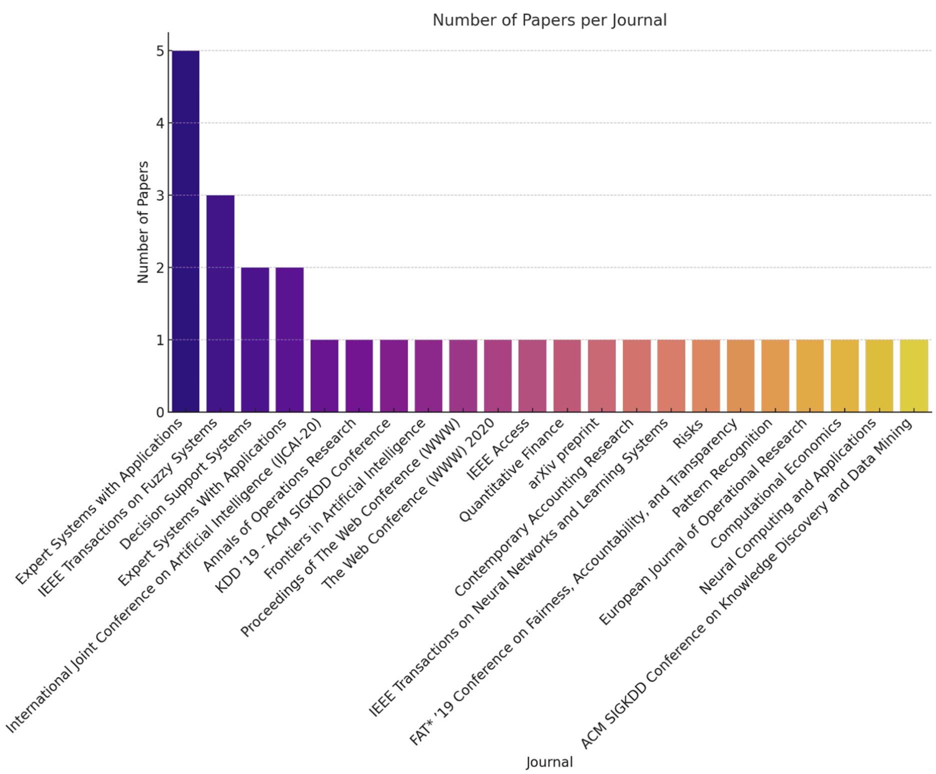

Figure 9. Journal and Conference Distribution

As in the case of the usage of XAI methods, attention mechanisms, SHAP, and feature importance analysis emerge as the most frequently employed techniques. The trend also shows that the acceptance of XAI approaches is the increasing dependence on

 doi.org/10.32388/PUO5CK

attention-based models and SHAP explanations **(Fig 10).** The adoption of feature attribution techniques such as SHAP and LIME has acquired momentum in credit scoring and fraud detection. Furthermore, complementing sophisticated black-box models in high-frequency trading and algorithmic decision-making are post-hoc explainability methods.

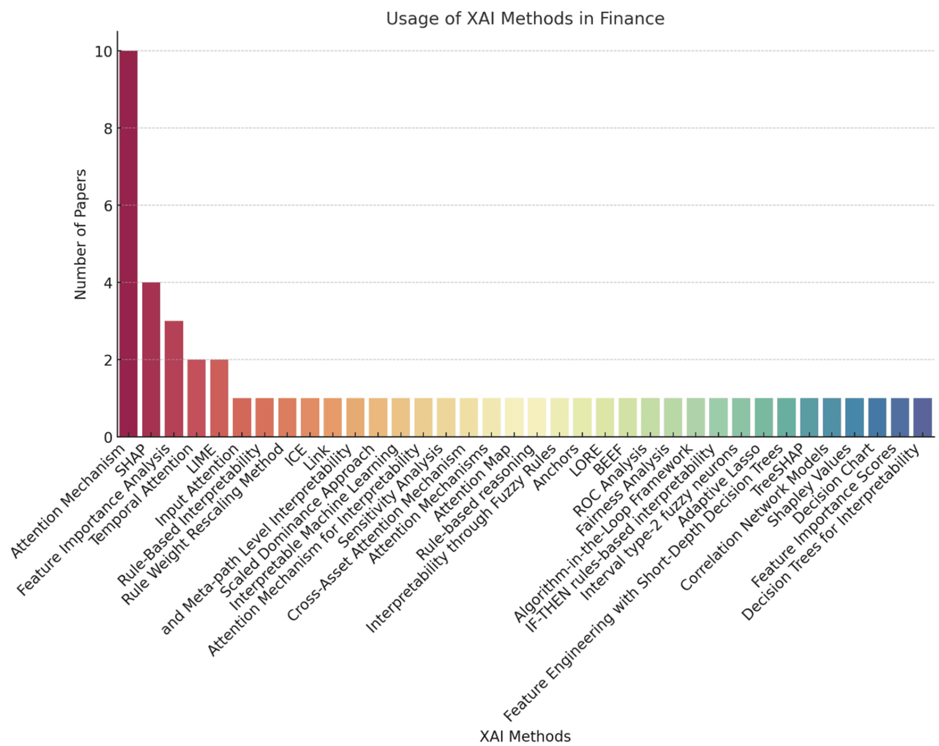

Figure 10. Adoption of XAI Techniques

The co-occurrence analysis of XAI and ML algorithms in financial research demonstrates distinct preferences for explainability techniques based on machine learning model complexity **(Fig 11).** SHAP, attention mechanisms, and feature importance analysis emerge as the most commonly used algorithms, especially when combined with XGBoost, LSTM, CNNs, and Transformers. A notable trend is the strong correlation between post-hoc interpretability techniques, such as SHAP and LIME, and tree-based models like XGBoost and Random Forest, whereas attention mechanisms and temporal interpretability techniques are more commonly used in conjunction with deep learning models, such as LSTMs and Transformers. The heatmap also shows a shift toward hybrid interpretability approaches, which combine different XAI techniques to improve model transparency. The inclusion of decision charts, correlation network models, and sensitivity analysis alongside feature attribution methods indicates a desire to create a multi-layered interpretability framework that balances local and global explanations. The findings show a divergent trajectory in explainability strategies, with traditional financial models relying on feature attribution and rule-based reasoning, and deep learning-driven financial systems requiring layer-wise attention mechanisms and neural network transparency tools.

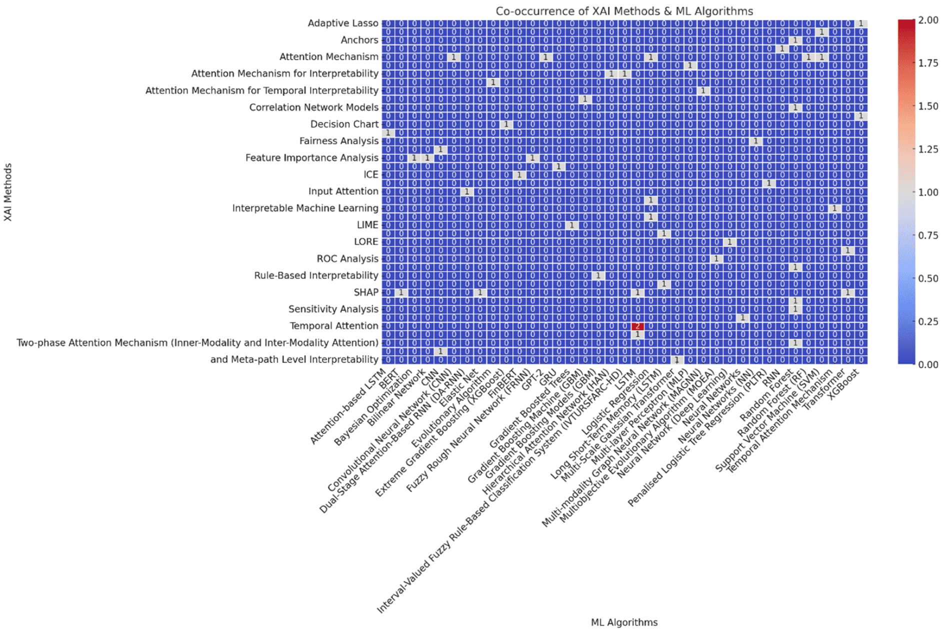

Figure 11. Co-occurrence Analysis of XAI Methods and ML Algorithms

Of the chosen studies, the trade and investment category accounts for 26.5%, the credit evaluation and scoring category accounts for 23.5%, and the fraud detection category accounts for 20.6% of the studies. The concentration on investment, risk reduction, and decision transparency shown by the high presence of these categories points to their increasing importance and dominance. Many of these studies use machine learning models—often combining explainability techniques such as SHAP, LIME, and Feature Importance Scores to increase openness. Particularly in financial time series forecasting, stock price prediction, and portfolio optimization, a great deal of research centers on trading and investment. To improve prediction accuracy, researchers use LSTM, Transformer-based architectures, and Reinforcement Learning models. In this field, attention methods are extensively applied to highlight the primary financial factors affecting market movements, hence enhancing model interpretability. Several publications also provide multi-objective optimization methods that balance explainability with forecast accuracy, therefore enabling financial analysts to better understand AI-driven trading strategies.

With many studies using network-based analysis, anomaly detection, and deep learning methods to identify fraudulent transactions, fraud detection is still a major focus of research. These studies seek to minimize false positives—a recurring difficulty in the financial industry—while simultaneously increasing fraud detection accuracy. Often used approaches include hierarchical attention networks for deep learning-based fraud identification, random forests for transaction analysis, and graph-based fraud detection models. These models depend heavily on feature attribution and counterfactual explanations, which offer clear justifications for identified fraudulent activities. Comprising 14.7% of the chosen papers, research in risk management investigates the function of explainable artificial intelligence in credit risk assessment, financial regulation, and decision-making transparency. Commonly utilized approaches include Shapley values, correlation network models, and counterfactual explanations. In banking and finance, where responsibility and fairness in AI decision-making are vital, these studies show the increasing requirement for interpretable artificial intelligence models. The variety of methodological techniques used in different financial subfields emphasizes how rapidly hybrid machine learning and deep learning models are being accepted. Underlining the industry's need for

doi.org/10.32388/PUO5CK

open, interpretable, and regulation-compliant AI solutions, the most often employed explainability methodologies are SHAP, LIME, Attention Mechanisms, and Feature Attribution Scores.

Our analysis of the publication trends shows a notable rise in research post-2018 in line with the acceptance of explainability in finance. Extensively referenced are deep learning-based explainability methods for trading and investment, especially Transformer-based models and LSTM architectures used for portfolio optimization and stock price prediction. Many studies propose multi-head attention processes, which offer a more thorough understanding of the main market factors affecting investment choices. These studies show how urgently interpretable artificial intelligence-driven trading techniques that match financial expert expectations are needed. Particularly, studies using graph-based and anomaly detection algorithms have attracted a lot of attention in fraud detection research. Many of these studies center on financial transaction fraud detection using deep learning explainability methods to increase model transparency. By means of saliency maps and hierarchical attention processes, financial analysts can grasp the reasons behind specific transaction flagging as fraudulent, thereby lowering false positives and raising detection accuracy. Research on risk management and regulatory compliance is also becoming more popular; papers on fairness in AI-driven financial systems are driving more citations for both. Research including causality-driven risk assessment systems, fairness-aware artificial intelligence models, and counterfactual explanations shows a rising focus on guaranteeing ethical and fair financial AI models. Combining decision forests with correlation-adjusted models to improve interpretability while maintaining good predictive accuracy, several articles provide ensemble solutions for reducing financial risk.

There are several limitations that are evident across the reviewed studies as well.[40][41][50] It suffers from limited real-time application. The need for further exploration of the model behavior is required in[42]. Dataset constraints present another significant limitation in[43][44][55][56][58][63]. Significantly higher computational costs pose a constraint in studies such as[45][46][49][51][52][54][59][61][42][48], where complex model architectures lead to significant processing overheads as well as limiting scalability.

4.3. Challenges in Implementing XAI in the Financial Domain

Though XAI holds great promise for finance, various obstacles have hampered its acceptance. AI systems should clearly explain their outputs; the way a system interacts with a user should never be taken as though the choice was made by a person rather than a machine[66]. However, developing models that are both accurate and interpretable is challenging, due in great part to the complexity of financial data and the numerous links between variables. Moreover, even if XAI approaches may explain model decisions, there is usually a trade-off between model interpretability and accuracy, since simpler models may offer more direct explanations but at the expense of lowered predictive performance. The ethical and legal ramifications of XAI in finance provide another difficulty, especially in terms of guaranteeing justice, openness, and responsibility in artificial intelligence-driven decision-making. Last but not least, the absence of uniform evaluation systems for XAI in finance poses a major challenge, since practitioners and academics lack common benchmarks to evaluate the success of several approaches. Many of the current artificial intelligence models in financial applications are still seen as "black-box" technology, which makes it challenging for stakeholders to completely trust or understand the decision-making process supporting AI projections. Among the several difficulties this lack of openness generates are ethical ones, questions about regulatory compliance, and doubts on the validity of decisions made by AI-driven companies. Therefore, resolving these challenges and supporting ethical AI deployment in the industry depends on investigating XAI approaches in the framework of finance.

5. Conclusion

In the last few years, AI has achieved notable momentum that, if harnessed appropriately, may deliver the best of expectations in finance. The integration of XAI in finance reflects a basic movement toward more transparent, responsible, and interpretable financial decision-making. In this paper, the junction of XAI and finance has been methodically investigated, underlining important trends, methodological developments, and useful applications of explainability methodologies. Notwithstanding these developments, some issues still exist, especially in the scalability, real-time application, and XAI standardization in financial systems. As the lack of interpretability and auditability of AI and Machine Learning (ML) methods could become a macro-level risk[67], it is imperative to learn how the inherent models work. Researchers, practitioners, and industry working together will help to create stronger, more scalable, and domain-specific explainability solutions from addressing these shortcomings. This improved interpretability and transparency would boost confidence in the models, allowing them to be widely used in real-world applications. As artificial intelligence (AI) continues to transform financial markets, the ability to explain and defend algorithmic models will be important for generating confidence, ensuring fairness, and meeting regulatory obligations. The rise of XAI in finance is more than just a technological advancement; it represents a necessary paradigm shift toward ethical, responsible, and transparent AI-powered financial systems.

References

1. ^*Cao S, Jiang W, Wang J, Yang B (2024). "From Man vs. Machine to Man + Machine: The Art and AI of Stock Analyses." Journal of Financi al Economics. 160:103910. doi:10.1016/j.jfineco.2024.103910.*
2. a, b*Sachan S, Yang J-B, Xu D-L, Benavides DE, Li Y (2020a). "An Explainable AI Decision-Support-System to Automate Loan Underwritin g." Expert Systems with Applications 144:113100. doi:10.1016/j.eswa.2019.113100.*
3. ^*Bornstein AM (2016). "Is Artificial Intelligence Permanently Inscrutable?" Nautilus. Retrieved January 29, 2025. https://nautil.us/is-artif icial-intelligence-permanently-inscrutable-236088/.*
4. ^*Harford T (2014). "Big Data: Are We Making a Big Mistake?" https://www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0.*
5. ^*Hawkins J (2017). "Special Report: Can We Copy the Brain? - What Intelligent Machines Need to Learn from the Neocortex." IEEE Spect rum 54(6):34–71. doi:10.1109/MSPEC.2017.7934229.*
6. ^*Kuang C (2017). "Can A.I. Be Taught to Explain Itself? - The New York Times." Retrieved January 29, 2025 https://www.nytimes.com/20 17/11/21/magazine/can-ai-be-taught-to-explain-itself.html.*
7. ^*Pavlus (2017). "Stop Pretending You Really Know What AI Is and Read This Instead." Quartz. Retrieved January 29, 2025 https://qz.co m/1067123/stop-pretending-you-really-know-what-ai-is-and-read-this-instead.*
8. ^*Mikulak A (2017). "Uncommon Insights Into Common Knowledge." APS Observer 30. Link.*
9. ^*Hidayat M, Defitri SY, Hilman H (2024). "The Impact of Artificial Intelligence (AI) on Financial Management." Management Studies an d Business Journal (PRODUCTIVITY) 1(1):123–29. Link.*
10. a, b*Bahoo S, Cucculelli M, Goga X, Mondolo J (2024). "Artificial Intelligence in Finance: A Comprehensive Review through Bibliometric a nd Content Analysis." SN Business & Economics. 4(2):23. doi:10.1007/s43546-023-00618-x.*
11. ^*Goodell JW, Jabeur SB, Saâdaoui F, Nasir MA (2023). "Explainable Artificial Intelligence Modeling to Forecast Bitcoin Prices." Internatio nal Review of Financial Analysis. 88:102702. doi:10.1016/j.irfa.2023.102702.*

12. ^Giudici P, Raffinetti E (2023). SAFE Artificial Intelligence in Finance. Vol. 56. Elsevier. https://www.sciencedirect.com/science/article/pii/S1544612323004609.
13. ^Zhou Y, Li H, Xiao Z, Qiu J (2023). A User-Centered Explainable Artificial Intelligence Approach for Financial Fraud Detection. Vol. 58. Elsevier. Link.
14. ^Clancey WJ (2014). "Methodology for Building an Intelligent Tutoring System." Pp. 51–83 in Methods and tactics in cognitive science. Psychology Press. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315802619-4/methodology-building-intelligent-tutoring-system-william-clancey.
15. ^Clancey WJ (1986). From GUIDON to NEOMYCIN and HERACLES in Twenty Short Lessons. Vol. 7. https://doi.org/10.1609/aimag.v7i3.548.
16. ^McKeown KR, Swartout WR (1987). "Language Generation and Explanation." Annual Review of Computer Science 2(1):401–49. doi:10.1146/annurev.cs.02.060187.002153.
17. ^Moore JD, Swartout WR (n.d.). "1990-Pointing: A Way Toward Explanation Dialogue." Link.
18. a, b, cGunning D, Aha D (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. Vol. 40. doi:10.1609/aimag.v40i2.2850.
19. a, b, c, dArrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, García S, Gil-López S, Molina D, Benjamins R (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. Vol. 58. Elsevier. https://www.sciencedirect.com/science/article/pii/S1566253519308103?via%3Dihub.
20. ^Hoffman RR, Mueller ST, Klein G, Litman J (2019). "Metrics for Explainable AI: Challenges and Prospects." Link.
21. ^Wazid M, Das AK, Chamola V, Park Y (2022). Uniting Cyber Security and Machine Learning: Advantages, Challenges and Future Research. Vol. 8. Elsevier. Link.
22. ^Rai A (2020). "Explainable AI: From Black Box to Glass Box." Journal of the Academy of Marketing Science 48(1):137–41. doi:10.1007/s11747-019-00710-5.
23. ^Linardatos P, Papastefanopoulos V, Kotsiantis S (2020). "Explainable Ai: A Review of Machine Learning Interpretability Methods." MDPI Vol. 23. Link.
24. a, b, cDas A, Rad P (2020). "Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey." https://arxiv.org/pdf/2006.11371.
25. ^Goebel R, Chander A, Holzinger K, Lecue F, Akata Z, Stumpf S, Kieseberg P, Holzinger A (2018). "Explainable AI: The New 42?" Pp. 295–303 in Machine Learning and Knowledge Extraction. Vol. 11015, Lecture Notes in Computer Science, edited by A. Holzinger, P. Kieseberg, A. M. Tjoa, and E. Weippl. Cham: Springer International Publishing. doi:10.1007/978-3-319-99740-7_21.
26. a, bBussmann N, Giudici P, Marinelli D, Papenbrock J (2021). "Explainable Machine Learning in Credit Risk Management." Computational Economics. 57(1):203–16. doi:10.1007/s10614-020-10042-0.
27. ^Adadi A, Berrada M (2018). Peeking inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). Vol. 6. IEEE. https://ieeexplore.ieee.org/document/8466590.
28. ^Talaat FM, Aljadani A, Badawy M, Elhosseini M (2024). "Toward Interpretable Credit Scoring: Integrating Explainable Artificial Intelligence with Deep Learning for Credit Card Default Prediction." Neural Computing and Applications 36(9):4847–65. doi:10.1007/s00521-023-09232-2.
29. ^Xu F, Uszkoreit H, Du Y, Fan W, Zhao D, Zhu J (2019). "Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges." In: Natural Language Processing and Chinese Computing. Vol. 11839, Lecture Notes in Computer Science, edited by J. Tang, M.-Y. Kan, D. Zhao, S. Li, and H. Zan. Cham: Springer International Publishing. doi:10.1007/978-3-030-32236-6_51.

30. ^Khaleghi B (n.d.). "An Explanation of What, Why, and How of eXplainable AI (XAI). 2020." Link.

31. ^LeCun Y, Bengio Y, Hinton G (2015). "Deep Learning." Nature 521(7553):436–44. doi:10.1038/nature14539.

32. ^Chen T, Guestrin C (2016). "XGBoost: A Scalable Tree Boosting System." Pp. 785–94 in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA: ACM. https://dl.acm.org/doi/10.1145/2939672.2939785.

33. ^Liaw A (2002). "Classification and Regression by randomForest." R News. Link.

34. ^Polikar R (2012). "Ensemble Learning." Pp. 1–34 in Ensemble Machine Learning, edited by C. Zhang and Y. Ma. New York, NY: Springer New York. doi:10.1007/978-1-4419-9326-7_1.

35. ^Lakkaraju H, Kamar E, Caruana R, Leskovec J (2017). "Interpretable & Explorable Approximations of Black Box Models." Link.

36. a, bDwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, Qian B, Wen Z, Shah T, Morgan G, Ranjan R (2023). "Explainable AI (XAI): Core Ideas, Techniques, and Solutions." ACM Computing Surveys. 55(9):1–33. doi:10.1145/3561048.

37. ^Nallakaruppan MK, Balusamy B, Shri ML, Malathi V, Bhattacharyya S (2024). "An Explainable AI Framework for Credit Evaluation and Analysis." Applied Soft Computing 153:111307. Link.

38. ^Bracke P, Datta A, Jung C, Sen S (2019). "Machine Learning Explainability in Finance: An Application to Default Risk Analysis." https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2019/machine-learning-explainability-in-finance-an-application-to-default-risk-analysis.pdf.

39. ^Nallakaruppan MK, Chaturvedi H, Grover V, Balusamy B, Jaraut P, Bahadur J, Meena VP, Hameed IA (2024). "Credit Risk Assessment and Financial Decision Support Using Explainable Artificial Intelligence." MDPI Vol. 12. Link.

40. a, bQin Y, Song D, Chen H, Cheng W, Jiang G, Cottrell G (2017). "A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction." Link.

41. a, bXia Y, Liu C, Li Y, Liu N (2017). A Boosted Decision Tree Approach Using Bayesian Hyper-Parameter Optimization for Credit Scoring. Vol. 78. Elsevier. Link.

42. a, b, cVan Vlasselaer V, Bravo C, Caelen O, Eliassi-Rad T, Akoglu L, Snoeck M, Baesens B (2015). "APATE: A Novel Approach for Automated Credit Card Transaction Fraud Detection Using Network-Based Extensions." Elsevier Vol. 75. Link.

43. a, bJin Z, Yang Y, Liu Y (2020). "Stock Closing Price Prediction Based on Sentiment Analysis and LSTM." Neural Computing and Applications 32(13):9713–29. doi:10.1007/s00521-019-04504-2.

44. a, bDumitrescu E, Hué S, Hurlin C, Tokpavi S (2022). Machine Learning for Credit Scoring: Improving Logistic Regression with Non-Linear Decision-Tree Effects. Vol. 297. Elsevier. https://www.sciencedirect.com/science/article/pii/S0377221721005695?via%3Dihub.

45. a, bCao B, Zhao J, Lv Z, Gu Y, Yang P, Halgamuge SK (2020). Multiobjective Evolution of Fuzzy Rough Neural Network via Distributed Parallelism for Stock Prediction. Vol. 28. IEEE.

46. a, bCheng D, Yang F, Xiang S, Liu J (2022). Financial Time Series Forecasting with Multi-Modality Graph Neural Network. Vol. 121. Elsevier. https://www.sciencedirect.com/science/article/pii/S003132032100399X?via%3Dihub.

47. ^Green B, Chen Y (2019). "Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments." Pp. 90–99 in Proceedings of the Conference on Fairness, Accountability, and Transparency. Atlanta GA USA: ACM. doi:10.1145/3287560.3287563.

48. a, bAddo PM, Guegan D, Hassani B (2018). Credit Risk Analysis Using Machine and Deep Learning Models. Vol. 6. MDPI. https://www.mdpi.com/2227-9091/6/2/38.

49. a, bTran DT, Iosifidis A, Kanniainen J, Gabbouj M (2018). "Temporal Attention-Augmented Bilinear Network for Financial Time-Series Data Analysis." IEEE Vol. 30. Link.

50. a, b *Moscato V, Picariello A, Sperlí G (2021). "A Benchmark of Machine Learning Approaches for Credit Score Prediction." Elsevier Vol. 165. Link.*
51. a, b *Huang AH, Wang H, Yang Y (2023). "FinBERT: A Large Language Model for Extracting Information from Financial Text." Contemporary Accounting Research 40(2):806–41. doi:10.1111/1911-3846.12832.*
52. a, b *Sanz JA, Bernardo D, Herrera F, Bustince H, Hagras H (2014). "A Compact Evolutionary Interval-Valued Fuzzy Rule-Based Classification System for the Modeling and Prediction of Real-World Financial Applications with Imbalanced Data." IEEE Vol. 23. Link.*
53. ^ *Ding Q, Wu S, Sun H, Guo J, Guo J (2020). "Hierarchical Multi-Scale Gaussian Transformer for Stock Movement Prediction." Pp. 4640–46 in IJCAI. https://idea.edu.cn/person/guojian/assets/papers/HMSG-transformer-2020.pdf.*
54. a, b *Craja P, Kim A, Lessmann S (2020). Deep Learning for Detecting Financial Statement Fraud. Vol. 139. Elsevier. https://www.sciencedirect.com/science/article/pii/S0167923620301767?via%3Dihub.*
55. a, b *Wang C, Chen Y, Zhang S, Zhang Q (2022). "Stock Market Index Prediction Using Deep Transformer Model." Elsevier Vol. 208. Link.*
56. a, b *Chen S, Ge L (2019). "Exploring the Attention Mechanism in LSTM-Based Hong Kong Stock Price Movement Prediction." Quantitative Finance. 19(9):1507–15. doi:10.1080/14697688.2019.1622287.*
57. ^ *Wang C, Han D, Liu Q, Luo S (2018). "A Deep Learning Approach for Credit Scoring of Peer-to-Peer Lending Using Attention Mechanism LSTM." IEEE Vol. 7. Link.*
58. a, b, c *Zhong Q, Liu Y, Ao X, Hu B, Feng J, Tang J, He Q (2020). "Financial Defaulter Detection on Online Credit Payment via Multi-View Attributed Heterogeneous Information Network." In: Proceedings of The Web Conference 2020. Taipei Taiwan: ACM. doi:10.1145/3366423.3380801.*
59. a, b *Antonelli M, Bernardo D, Hagras H, Marcelloni F (2017). "Multiobjective Evolutionary Optimization of Type-2 Fuzzy Rule-Based Systems for Financial Data Classification." IEEE Transactions on Fuzzy Systems. 25(2):249–64. doi:10.1109/TFUZZ.2016.2578341.*
60. ^ *Bussmann N, Giudici P, Marinelli D, Papenbrock J (2020). Explainable AI in Fintech Risk Management. Vol. 3. Frontiers Media SA. doi:10.3389/frai.2020.00026.*
61. a, b *Wang J, Zhang Y, Tang K, Wu J, Xiong Z (2019). "AlphaStock: A Buying-Winners-and-Selling-Losers Investment Strategy Using Interpretable Deep Reinforcement Attention Networks." Pp. 1900–1908 in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage AK USA: ACM. Link.*
62. ^ *Haq AU, Zeb A, Lei Z, Zhang D (2021). Forecasting Daily Stock Trend Using Multi-Filter Feature Selection and Deep Learning. Vol. 168. Elsevier. https://www.sciencedirect.com/science/article/pii/S095741742031099X?via%3Dihub.*
63. a, b, c *Florez-Lopez R, Ramon-Jeronimo JM (2015). Enhancing Accuracy and Interpretability of Ensemble Strategies in Credit Risk Assessment. A Correlated-Adjusted Decision Forest Proposal. Vol. 42. Elsevier. https://www.sciencedirect.com/science/article/pii/S0957417415001499.*
64. ^ *Zhang Q, Qin C, Zhang Y, Bao F, Zhang C, Liu P (2022). Transformer-Based Attention Network for Stock Movement Prediction. Vol. 202. Elsevier. Link.*
65. ^ *Yoo J, Soun Y, Park Y, Kang U (2021). "Accurate Multivariate Stock Movement Prediction via Data-Axis Transformer with Multi-Level Contexts." In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. Virtual Event Singapore: ACM. doi:10.1145/3447548.3467297.*
66. ^ *Purificato E, Lorenzo F, Fallucchi F, De Luca EW (2023). "The Use of Responsible Artificial Intelligence Techniques in the Context of Loan Approval Processes." International Journal of Human–Computer Interaction 39(7):1543–62. doi:10.1080/10447318.2022.2081284.*

67. ^*Board FSB Financial Stability (2017). Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications. Financial Stability Board. https://www.fsb.org/uploads/P011117.pdf.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.