To what extent could Machine Learning Revolutionize the Approach to Stock Market and Trade Predictions?

Maximo Rua

Abstract:
This study evaluates the predictive performance of four machine learning models with different types of supervision. Logistic regression, Central Neural Network (CNN), random forest, standard Support Vector Machine (SVM), and sparse SVM on U.S. equity data spanning Q1 2024 to Q2 2025. Using 30 technical and sentiment-based indicators, each model was assessed using accuracy, precision, recall, F1-score, and AUC. Sparse SVM achieved the highest overall performance, with 87.4% accuracy, 85.1% precision, 83.9% recall, 84.5% F1-score, and 91.2% AUC, while selecting only 7 features. These results indicate that sparse SVM offers superior predictive power and interpretability, demonstrating that machine learning, particularly models with embedded feature selection, can substantially improve the precision and efficiency of stock market forecasting.

Introduction

Despite the increasing availability of financial data and the growing complexity of market behavior, traditional statistical models often struggle to capture non-linear relationships and adapt to rapidly shifting market dynamics. This limitation presents significant challenges for investors and analysts seeking accurate and timely stock price forecasts. In this study, we hypothesize that machine learning (ML) models outperform conventional statistical approaches when applied to large-scale, high-frequency financial datasets, particularly in environments characterized by volatility, noise, and latent behavioral signals. We conduct a comparative analysis of ML techniques and traditional baselines to identify market conditions under which ML models demonstrate superior predictive performance. Additionally, we introduce a methodological framework that integrates market sentiment and behavioral indicators into forecasting models, and we assess the interpretability and practical trade-offs of ML-based approaches from a financial decision-making standpoint. The paper begins with a background on the application of ML in financial prediction, followed by a review of related literature. We then detail our methodology, including data preprocessing, feature selection, and model design, before presenting experimental results and performance evaluations. Finally, we discuss key findings, highlight limitations, and suggest directions for future research.

Background on Machine Learning

Machine learning (ML) possesses the ability to identify patterns, learn from historical data, and transform raw inputs into structured outputs. ML algorithms are classified based on their learning approach, each suited for distinct applications depending on data availability and task objectives.

Supervised Learning *involves training models on labeled data to learn a mapping from inputs to outputs. The data is typically split into training and testing sets, with the training set providing* *known outcomes that guide the learning process. This approach is ideal for tasks with clear cause-and-effect relationships and is widely applied in predictive modeling and automated decision-making, such as algorithmic trading. Common algorithms include linear regression, support vector machines (SVM), random forests, and neural networks.*

*By contrast, **Unsupervised Learning** works on unlabeled data to uncover inherent structures or detect anomalies without external guidance. Techniques such as clustering and dimensionality reduction are utilized for tasks like identifying suspicious trades or unusual market behavior, supporting early fraud detection.*

***Semi-Supervised Learning** bridges supervised and unsupervised methods, leveraging a small labeled dataset alongside a larger unlabeled set to improve classification in scenarios where labeled data is scarce or expensive. For example, it can classify financial transactions as "normal" or "suspicious" by generalizing from limited annotations to broader datasets.*

*A notable semi-supervised technique is **Self-Training**, where a model initially trained on labeled data assigns pseudo-labels to unlabeled data, iteratively expanding its training set. This method is effective in applications like churn prediction, where only partial outcome information is available.*

***Multitask Learning (MTL)** simultaneously addresses multiple related objectives by exploiting shared information, enhancing model generalization and efficiency. For instance, models predicting revenue, churn, and customer lifetime value jointly benefit from overlapping behavioral patterns, improving overall predictive performance.*

***Ensemble Learning** combines multiple models to achieve greater accuracy and robustness than any single model. This approach is common in high-stakes domains such as credit scoring and financial forecasting. Key techniques include:*

- ***Boosting**, which sequentially trains weak learners to correct predecessors' errors, producing a strong composite model (e.g., AdaBoost, Gradient Boosting).*
- ***Bagging (Bootstrap Aggregating)**, which trains models on random data subsets independently, aggregating results to reduce overfitting (e.g., Random Forest).*
- ***Stacking**, which integrates diverse base models via a meta-model that learns optimal combination strategies, useful in complex forecasting by synthesizing varied data sources.*

Beyond algorithmic methods, automation of institutional processes increasingly relies on Big Data and Advanced Analytics (BD&AA), often combined with robotic process automation (RPA) and AI, to improve efficiency and maintain staffing levels during growth (Prothin, 2020). Effective governance frameworks are essential to ensure compliance, incorporating access controls, model lineage tracking, and change management protocols to maintain transparency and regulatory adherence.

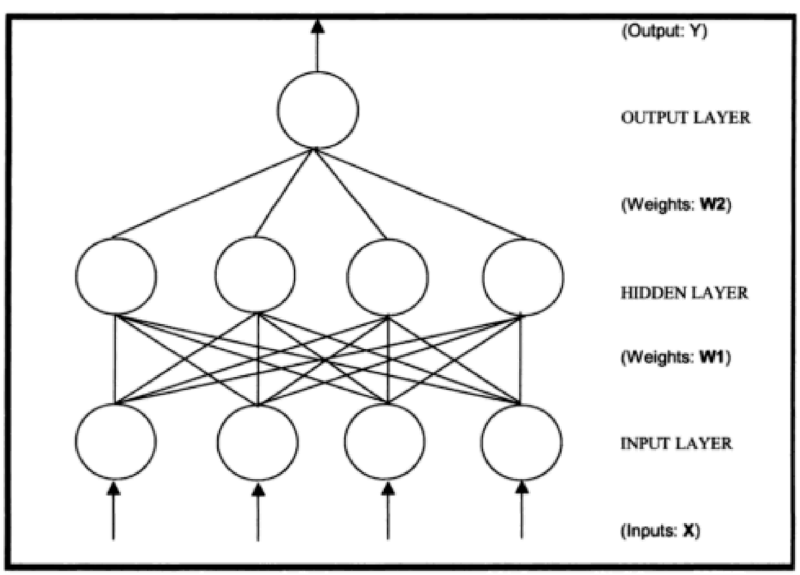
ML in the Stock Market and Trade Predictions

Neural Networks and Their Importance

Neural networks are a class of machine learning models designed to mimic the way the human brain processes information. They are capable of learning directly from data by identifying patterns and making predictions without relying on predefined rules. According to the European Information Technologies Certification Academy (EITCA), neural networks consist of interconnected components, including neurons, layers, activation functions, weights, and biases. Neurons receive input, perform computations, and pass outputs through weighted connections. These neurons are arranged into layers: the input layer receives data, hidden layers extract features and patterns, and the output layer generates predictions. Weights and biases are trainable parameters that determine how input signals are transformed as they pass through the network, and they are updated during training to minimize prediction errors.

A crucial aspect of neural networks is the use of ***activation functions****, which introduce non-linearity into the model, allowing it to capture complex relationships. Common activation functions include the* ***sigmoid*** *(suitable for binary outputs),* ***tanh*** *(useful for zero-centered data), and* ***ReLU*** *(effective for deep networks due to its simplicity and computational efficiency). Neural networks are trained using a* ***loss function****, which measures the difference between predicted and actual outputs. Optimization algorithms like* ***stochastic gradient descent*** *or* ***Adam*** *adjust the weights and biases to minimize this loss. The training process relies heavily on* ***backpropagation****, a technique that computes the gradient of the loss function and updates parameters efficiently. Together, these components enable neural networks to perform a wide range of tasks, from classification to regression, with high adaptability and precision.*

Figure 1: A Typical Feedforward Neural Network



(Zhang, 2004).

In the figure below, it is a visual of a dataset from fundamentals.csv metrics extracted from annual SEC 10K filings (2012-2016). 30% of the traffic on stocks is already generated by machines. This dataset was created to analyze whether the involvement of ML generates earnings.

Ticker Sy...	Period End...	# Add'l inco...	# After Tax ...	# Capital Ex...	# Cost of Re...	# Earnings B...
AAL	2012-12-31	-1961000000.0	23.0	-1888000000.0	10499000000.0	-1813000000.0
AAL	2013-12-31	-2723000000.0	67.0	-3114000000.0	11019000000.0	-1324000000.0
AAL	2014-12-31	-150000000.0	143.0	-5311000000.0	15620000000.0	4099000000.0
AAL	2015-12-31	-708000000.0	135.0	-6151000000.0	11096000000.0	5496000000.0
AAP	2012-12-29	600000.0	32.0	-271182000.0	3106967000.0	657915000.0
AAP	2013-12-28	2698000.0	26.0	-195757000.0	3241668000.0	663016000.0
AAP	2015-01-03	3092000.0	25.0	-228446000.0	5390248000.0	854802000.0
AAP	2016-01-02	-7484000.0	19.0	-234747000.0	5314246000.0	818296000.0
AAPL	2013-09-28	1156000000.0	30.0	-8165000000.0	1.06606e+11	50155000000.0
AAPL	2014-09-27	980000000.0	35.0	-9571000000.0	1.12258e+11	53483000000.0
AAPL	2015-09-26	1285000000.0	45.0	-11247000000.0	1.40089e+11	72515000000.0
AAPL	2016-09-24	1348000000.0	36.0	-12734000000.0	1.31376e+11	61372000000.0
ABBV	2012-12-31	-8000000.0	1507.0	-333000000.0	4508000000.0	5809000000.0
ABBV	2013-12-31	-54000000.0	92.0	-491000000.0	4581000000.0	5610000000.0
ABBV	2014-12-31	-651000000.0	102.0	-612000000.0	4426000000.0	2760000000.0
ABBV	2015-12-31	-206000000.0	130.0	-532000000.0	4500000000.0	7331000000.0
ABC	2013-09-30	-44000.0	19.0	-202450000.0	85451348000.0	898355000.0
ABC	2014-09-30	-28594000.0	14.0	-264457000.0	1.16586761e+11	753497000.0

(Gawlik, 2016)

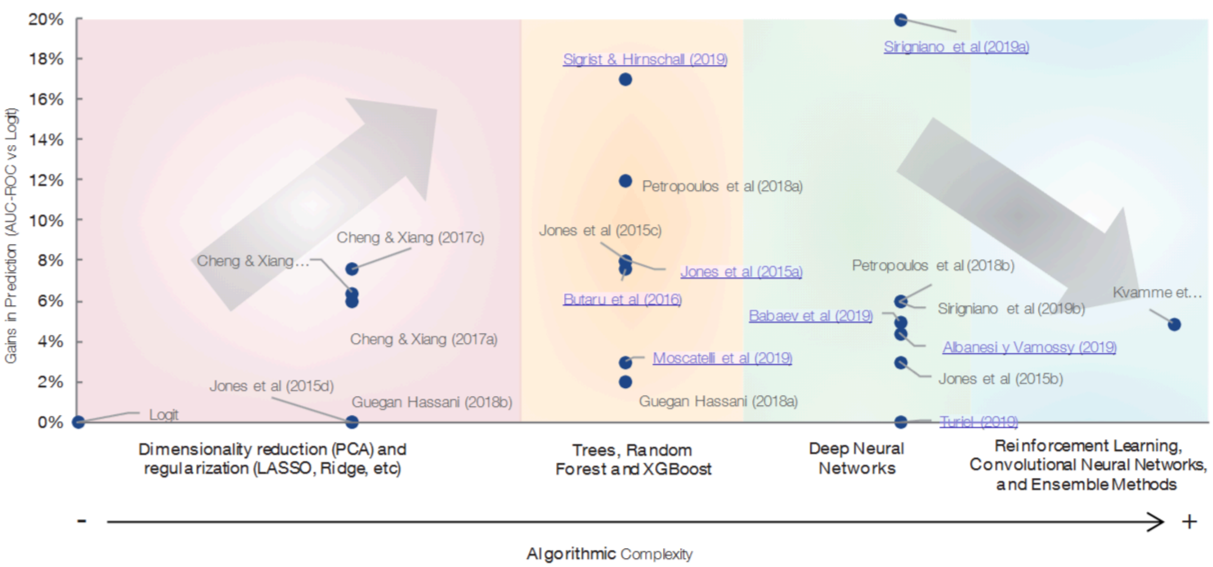
This dataset comprises financial metrics from companies such as AAL (American Airlines), AAP, AAPL (Apple), ABBV, and ABC, based on SEC 10-K filings from 2012 to 2016. Most firms showed positive earnings before tax, with earlier data from American Airlines as a notable exception. These patterns suggest that machine learning (ML) models trained on the dataset effectively identified profitable trends and supported accurate trade predictions. The consistent returns imply that ML-assisted analysis contributed to enhanced portfolio performance, offering benefits such as improved risk-adjusted metrics, reduced transaction costs, and early detection of market anomalies. ML integration further enables dynamic position sizing, automated stress testing, and faster, data-driven decision-making. The use of alternative data sources, reinforcement learning, and modular architectures also supports continuous model refinement and adaptability to evolving market conditions, highlighting the value of ML in financial forecasting and strategy development.

Case Studies and Applications

Supervised Learning

A supervised learning algorithm (e.g., logistic regression, decision tree, or random forest) is trained on historical data where the outcome (default/no default) is already known. After training, *the model can predict the likelihood that a new applicant will default. An example of supervised learning is credit risk prediction.*

Figure 1. The dilemma between prediction and algorithmic complexity



(Alonso & Carbo, 2020)

The graph illustrates the relationship between Gains in Prediction and Algorithmic Complexity. In the red zone, representing lower algorithmic complexity, there is a strong positive correlation between the two variables, indicating that simpler algorithms tend to yield higher predictive gains. In contrast, the yellow zone shows no clear relationship, suggesting a plateau where increased complexity does not translate into improved predictive performance. Finally, in the green zone, which corresponds to high algorithmic complexity, a negative relationship is observed, suggesting that overly complex algorithms may reduce predictive gains. These findings imply that predictive effectiveness tends to diminish as algorithmic complexity increases beyond a certain point, highlighting the advantages of models that are simpler, more transparent, and easier to supervise and interpret.

From a business standpoint, enhanced debtor classification capabilities can lead to direct financial benefits, including increased profitability and cost efficiencies. More critically, however, such capabilities constitute a fundamental element of a robust credit risk management framework. At the micro level, these improvements can shape an institution's risk appetite, thereby supporting efforts to optimize market share. On a macro scale, the implications extend to broader financial system outcomes, such as improved financial inclusion. This may be facilitated by the integration of machine learning models and the exploitation of large datasets, *including alternative data sources like digital footprints. Such developments offer the potential to extend credit access to previously undeserved segments of the population, particularly individuals with limited or no formal credit history.*

While regulatory fragmentation in this regard adds value and allows for a fully fledged coverage of the potential risks derived from using predictive models is also an obstacle to isolating the factors that determine whether or not a new quantitative tool is compatible with the regulatory and supervisory framework. There are papers in the literature that try to Explain which factors matter to the supervisors when evaluating ML models or AI (Alonso & Carbo, 2020)

Table 1. Summary of factors that determine the benefits and supervisory cost functions, based on each possible use of the ML model

Benefits Function		Supervisory Costs Function		Model Uses
Discriminatory power	Statistics	(1) Stability	(2) Over-fitting	Credit scoring
Accuracy		(3) Hyper-parameters	(4) Dynamic Calibration	Pricing
		(5) Feature Engineering		Provisioning
	Technology	(6) Transparency		Regulatory Capital
		(7) Carbon Footprint		Supervisory Model
		(8)Third-party providers		
		(9) Cyber Risk		
	Conduct	(10) Privacy		
		(11) Auditability		
		(12) Interpretability		
		(13) Biases		

(Alonso & Carbo, 2020)

Supervised machine learning (ML) significantly influences statistical modeling and the analysis of large datasets. However, it is primarily effective with quantitative data, as it struggles to accurately predict qualitative outcomes. This limitation can introduce biases and compromise the transparency and auditability of results. Nonetheless, in domains such as credit scoring—where datasets are extensive and regulatory standards are well established—supervised ML can demonstrate clear value by providing precise and interpretable insights.

Challenges and Limitations of ML in Finance

Much of the technical discussion on algorithmic fairness assumes fixed social objectives, target populations, and allowable actions in model deployment. However, these elements are *normative choices often shaped by external policies or actors beyond model development, critically determining whether and how models advance fairness, regardless of its definition.*

Challenges and Limitations of Machine Learning in Finance

1. ***Data Challenges***

Financial data frequently suffer from noise, missing values, and temporal misalignment. Alternative sources like social media or news feeds introduce biases and inconsistent structures, complicating integration and reliability. Label sparsity hinders supervised tasks, especially in fraud detection or investment prediction, due to delayed, sparse, or costly ground-truth labels tied to rare or lagged events.

Financial markets exhibit non-stationarity, with regime shifts that erode the relevance of historical data and cause model degradation. Sampling bias also arises, for example, when credit risk models exclude denied applicants, skewing risk assessment.

Mitigations include rolling retraining to adapt to new data, feature stores with versioning to improve reproducibility, and temporal data splits to prevent forward-looking bias.

2. ***Model-Related Challenges***

High-capacity models risk overfitting, capturing spurious historical correlations that fail to generalize in volatile markets. Concept drift and performance decay necessitate continuous monitoring and retraining.

Regulatory and institutional demands for model interpretability pose challenges, as advanced models often lack transparency. Solutions involve drift detection tools, interpretability techniques (e.g., SHAP, LIME), and choosing simpler models when gains from complexity are marginal.

3. ***System-Level Constraints***

Financial ML systems require ultra-low latency, fault tolerance, and robust security, especially in real-time contexts like trading or fraud detection. Compliance with regulations such as GDPR demands thorough auditability and documentation.

Adversarial risks include model inversion, data poisoning, and manipulation via public channels. Countermeasures encompass secure hosting, encrypted pipelines, comprehensive logging, and governance frameworks with access controls and change management to ensure compliance.

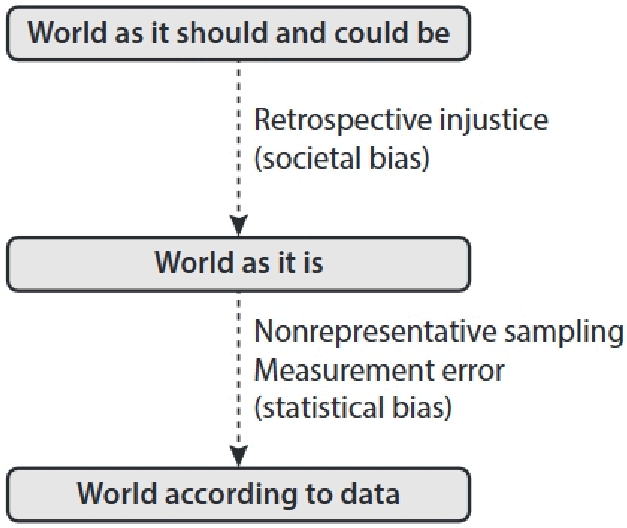


Figure 1

A schematic showing two components of biased data: societal bias and statistical bias.

(Mitchell et al., 2021)

The illustration above highlights certain limitations of machine learning (ML) that cannot be fully addressed through algorithmic solutions alone. In this context, we define statistical bias as a systematic discrepancy between the data used to train a predictive model and the current state of the real world. Our focus here is on how sampling bias and measurement error can give rise to fairness concerns that often fall outside the scope of formal mathematical definitions of fairness.

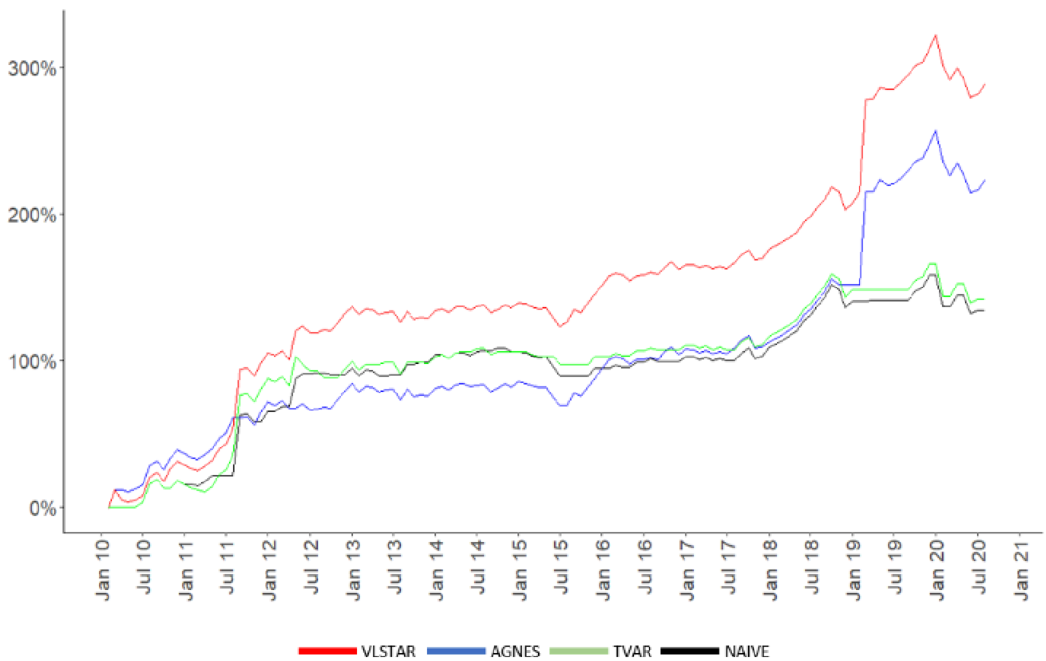
Ethical and regulatory concerns in financial machine learning—particularly in credit modeling—extend beyond the technical performance of the model. Even when datasets are statistically representative, models may encode and perpetuate societal inequalities, such as structural wage disparities tied to race or gender. Excluding protected attributes does not guarantee fairness, as proxy variables (e.g., ZIP code, education) may reproduce biased outcomes. Addressing these issues requires not only technical interventions (e.g., fairness-aware feature selection, counterfactual testing) but also critical scrutiny of the broader social and institutional context in which models are developed and applied.

Unsupervised Learning

Financial markets are inherently susceptible to sudden behavioral shifts, presenting substantial challenges to traditional statistical modeling techniques. The dynamics of asset returns often necessitate frequent re-estimation of key parameters—such as expected returns, volatility, and correlations—to reflect rapidly changing market conditions. Episodes like the 2008 Global Financial Crisis and the onset of the COVID-19 pandemic serve as stark reminders of how *quickly dependence structures (particularly linear correlations) can break down as markets shift into "risk-off" regimes.*

Such periods highlight the need for regime-switching models that treat market regimes as latent and unobservable states, capturing structural changes in financial dynamics. These regimes often exhibit persistence and play a critical role in both risk assessment and portfolio allocation.

Figure 6: Equity line for each investment strategy.



(Bucci & Ciciretti, 2021)

Table 2: Confusion matrices produced by applying the three procedures to the randomly generated data.

		VLSTAR(1)		Hierarchical clustering			TVAR (1)	
		Predicted			Predicted			Predicted	
		Calm	High-Vol		Calm	High-Vol		Calm	High-Vol
Realized	Calm	65%	15%	Calm	45%	11%	Calm	32%	8%
	High-Vol	4%	16%	High-Vol	25%	19%	High-Vol	57%	3%

(Bucci & Ciciretti, 2021)

Table 4: Total transaction costs (in basis points) incurred by each filtered strategy compared to the naïve one.

	Transaction costs in basis points
Naïve	139.1%
VLSTAR	122.5%
Hierarchical clustering	132.6%
TVAR	138.7%

(Bucci & Ciciretti, 2021)

In this context, unsupervised learning techniques—particularly clustering analysis—offer a complementary approach to regime detection. By segmenting financial time series into groups with similar statistical behavior, clustering methods help identify recurring structural patterns across time. This approach allows for the detection of latent regimes not at the level of individual asset fluctuations, but through the identification of more stable interdependencies at the cluster level—a process akin to coarse-graining.

The regime classification framework discussed here is based on Bucci & Ciciretti (2022), who integrate traditional time-series modeling with unsupervised machine learning to identify and characterize latent market states. Their findings suggest that incorporating clustering and smooth transition models improves the identification of structural changes in financial markets and enhances decision-making in periods of heightened uncertainty. All empirical results referenced in this context are drawn directly from their study, and no independent analysis has been conducted as part of this work.

Semi-Supervised Learning

Money Laundering and Graph-Based Detection

Money laundering involves disguising illicit proceeds—often from activities like fraud or trafficking—to appear legitimate, typically through placement, layering, and integration stages. Beyond enabling crime, it undermines financial integrity and trust.

Graph-based approaches to anti-money laundering (AML) model accounts as nodes and transactions as directed edges. Semi-supervised learning methods, such as label propagation, are particularly effective in this context, leveraging limited labeled data (e.g., from Suspicious Activity Reports) to infer suspicious behavior across large transaction networks. Structural patterns—like cyclic flows or tightly connected clusters—can signal illicit activity. Datasets such as AMLSim (Rezaul et al., 2025) provide synthetic benchmarks for evaluating these methods.

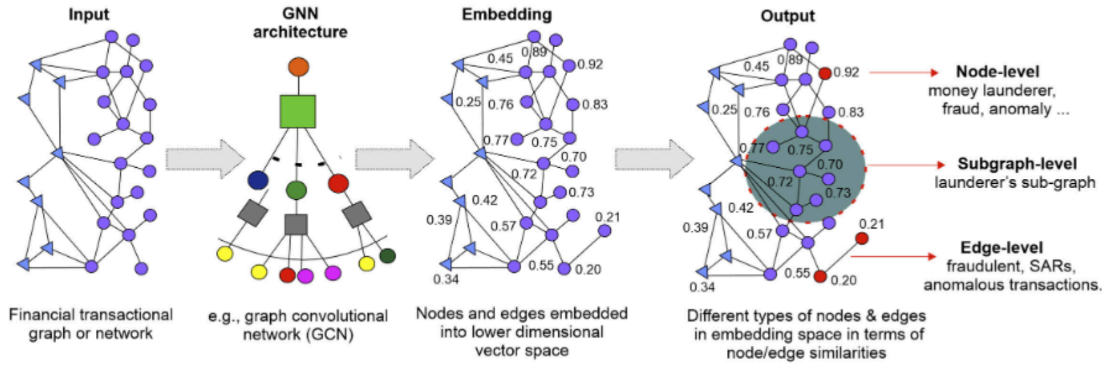


FIGURE 4. **General workflows of using GNNs for anomaly detection in financial transactional graphs.**

(Rezaul et al., 2025)

Adapting AML Techniques to Market Microstructure.

Similar graph-based, semi-supervised methods can be applied to market surveillance. In high-frequency trading, nodes represent traders and edges represent order interactions. Features include order type, volume, price, and timing. Despite sparse labels, semi-supervised models can flag behaviors such as spoofing, wash trading, or coordinated manipulation. This approach supports scalable, adaptive monitoring in dynamic market environments.

Self-Training

Predicting the direction of stock price movement–whether it will rise or fall–offers a more actionable framework for trading decisions than forecasting exact price levels. This task is effectively formulated as a binary classification problem using an SVM, where each time step is labeled +1 for a positive return and -1 for a negative return. This model is trained on a high-dimensional feature matrix composed of technical indicators, volume patterns, and optional sentiment or fundamental variables. SVMs identify the optimal hyperplane to separate the classes, using linear or kernel functions depending on feature complexity. To avoid overfitting and improve interpretability, sparsity is introduced either through L1 regularization (penalizing less informative features) or wrapper-based recursive feature elimination (RFE). Model evaluation is conducted using rolling-window backtests, with performance measured via classification accuracy, precision, recall, and Sharpe ratios from simulated trading strategies.

Empirical results from Miao et al. (2025) support this methodology, demonstrating that backward selection methods like SVM-RFE may underperform compared to more interaction-aware feature selection techniques such as BSE-SVMs or SVM-RFE combined with the relief algorithm. Their experiments on several Chinese A-share stocks (600085, 600332, 600559) showed improved classification accuracy and F-test scores when feature interactions were considered. This suggests that refining the feature set–particularly in the presence of correlated *indicators–plays a critical role in model efficacy. Overall, the integration of sparsity, semi-supervised learning, and careful feature selection within an SVM-based framework offers a scalable, interpretable, and robust approach for directional stock prediction in high-noise environments.*

TABLE 5 | Forecasting performance of five sparse linear kernel SVMs and linear kernel SVMs for the nine stock datasets.

Panel A: Stock 603919

Model	Accuracy	Recall	Specificity	Precision	F-test	Subset size
SVMs	0.8294	**0.8158**	0.8404	0.8052	0.8105	49
CFS-based SVMs	0.5647	0.3816	0.7128	0.5179	0.4394	6
ReliefF-based SVMs	0.8235	0.7895	0.8511	0.8108	0.8000	13
Backward Sequential Elimination-SVMs	0.8235	0.8026	0.8404	0.8026	0.8026	42
SVMs-RFE	**0.8471**	0.7895	**0.8936**	**0.8571**	0.8219	13
SVMs-RFE with ReliefF	**0.8471**	0.8026	0.8830	0.8472	**0.8243**	13

Panel B: Stock 603369

Model	Accuracy	Recall	Specificity	Precision	F-test	Subset size
SVMs	0.8176	0.7733	0.8526	0.8056	0.7891	49
CFS-based SVMs	0.6118	0.3600	0.8105	0.6000	0.4500	7
ReliefF-based SVMs	0.8118	0.8000	0.8211	0.7792	0.7895	9
Backward Sequential Elimination-SVMs	0.8176	0.8000	0.8316	0.7895	0.7947	38
SVMs-RFE	**0.8294**	0.7867	**0.8632**	**0.8194**	0.8027	9
SVMs-RFE with ReliefF	**0.8294**	**0.8267**	0.8316	0.7949	**0.8105**	9

Panel C: Stock 600519

Model	Accuracy	Recall	Specificity	Precision	F-test	Subset size
SVMs	0.7529	0.7534	0.7526	0.6962	0.7237	49
CFS-based SVMs	0.5824	0.3973	0.7216	0.5179	0.4496	8
ReliefF-based SVMs	0.7647	0.6986	**0.8144**	**0.7391**	0.7183	25
Backward Sequential Elimination-SVMs	0.7588	**0.7808**	0.7423	0.6951	0.7355	35
SVMs-RFE	0.7706	0.7534	0.7835	0.7237	0.7383	25
SVMs-RFE with ReliefF	**0.7824**	**0.7808**	0.7835	0.7308	**0.7550**	25

Panel C: Stock 600300

Model	Accuracy	Recall	Specificity	Precision	F-test	Subset size
SVMs	0.7765	0.7957	0.7532	0.7957	0.7957	49
CFS-based SVMs	0.5588	**0.9570**	0.0779	0.5563	0.7036	6
ReliefF-based SVMs	0.7765	0.7849	0.7662	0.8022	0.7935	12
Backward Sequential Elimination-SVMs	0.7941	0.8065	**0.7792**	**0.8152**	0.8108	44
SVMs-RFE	**0.8000**	0.8280	0.7662	0.8105	**0.8191**	12
SVMs-RFE with ReliefF	**0.8000**	0.8280	0.7662	0.8105	**0.8191**	12

Panel C: Stock 600559

Model	Accuracy	Recall	Specificity	Precision	F-test	Subset size
SVMs	0.8294	0.7750	0.8778	0.8493	0.8105	49
CFS-based SVMs	0.6118	0.7000	0.5333	0.5714	0.6292	6
ReliefF-based SVMs	0.8235	**0.8000**	0.8444	0.8205	0.8101	17
Backward Sequential Elimination-SVMs	0.8294	0.7875	0.8667	0.8400	0.8129	31
SVMs-RFE	0.8294	0.7875	0.8667	0.8400	0.8129	17
SVMs-RFE with ReliefF	**0.8471**	**0.8000**	**0.8889**	**0.8649**	**0.8312**	17

Panel C: Stock 600809

Model	Accuracy	Recall	Specificity	Precision	F-test	Subset size
SVMs	0.7529	0.6712	0.8144	0.7313	0.7000	49
CFS-based SVMs	0.6000	0.3562	0.7835	0.5532	0.4333	6
ReliefF-based SVMs	0.7588	0.6712	0.8247	0.7424	0.7050	2
Backward Sequential Elimination-SVMs	0.7706	**0.7123**	0.8144	0.7429	**0.7273**	41
SVMs-RFE	**0.7824**	0.6575	**0.8763**	**0.8000**	0.7218	2
SVMs-RFE with ReliefF	**0.7824**	0.6575	**0.8763**	**0.8000**	0.7218	2

(Continues)

TABLE 5 | (Continued)

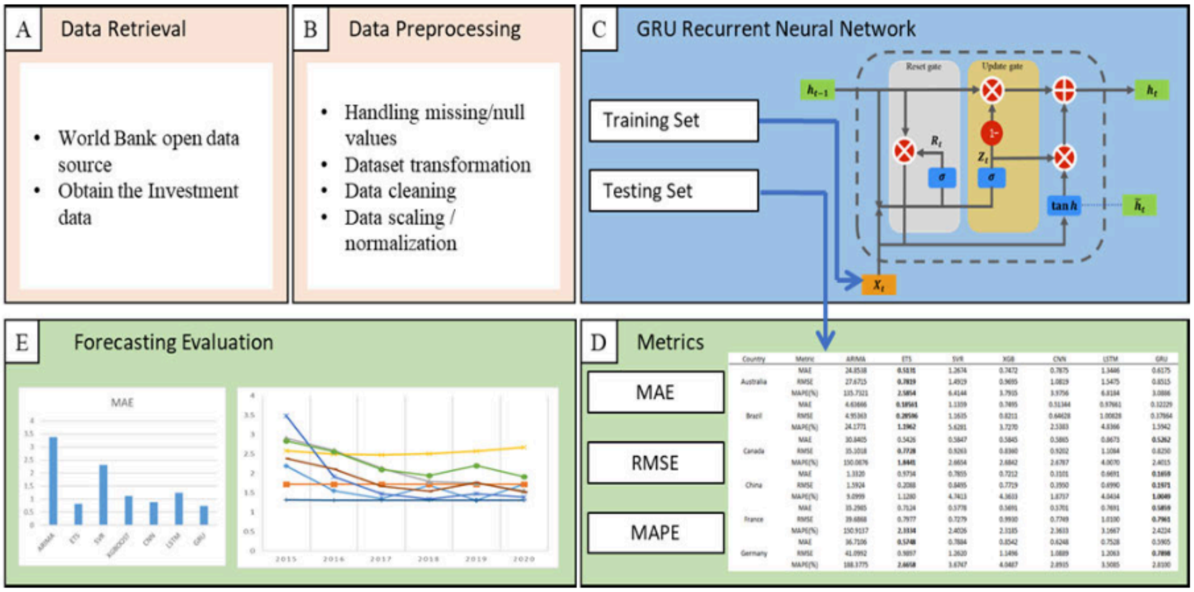
Panel C: Stock 600085						
Model	**Accuracy**	**Recall**	**Specificity**	**Precision**	**F-test**	**Subset size**
SVMs	**0.8471**	**0.8101**	0.8791	0.8533	**0.8312**	49
CFS-based SVMs	0.7059	0.5823	0.8132	0.7302	0.6479	6
ReliefF-based SVMs	0.8353	0.7848	0.8791	0.8493	0.8158	17
Backward Sequential Elimination-SVMs	0.8412	**0.8101**	0.8681	0.8421	0.8258	43
SVMs-RFE	0.8294	0.7468	**0.9011**	**0.8676**	0.8027	17
SVMs-RFE with ReliefF	**0.8471**	**0.8101**	0.8791	0.8533	**0.8312**	17
Panel C: Stock 600332						
Model	**Accuracy**	**Recall**	**Specificity**	**Precision**	**F-test**	**Subset size**
SVMs	0.8000	0.7763	0.8191	0.7763	0.7763	49
CFS-based SVMs	0.5765	0.6316	0.5319	0.5217	0.5714	6
ReliefF-based SVMs	0.7706	0.7632	0.7766	0.7342	0.7484	11
Backward Sequential Elimination-SVMs	**0.8118**	**0.8158**	0.8085	0.7750	**0.7949**	39
SVMs-RFE	0.8000	0.7895	0.8085	0.7692	0.7792	11
SVMs-RFE with ReliefF	**0.8118**	0.7895	**0.8298**	**0.7895**	0.7895	11
Panel C: Stock 002304						
Model	**Accuracy**	**Recall**	**Specificity**	**Precision**	**F-test**	**Subset size**
SVMs	0.8118	0.7027	0.8958	0.8387	0.7647	49
CFS-based SVMs	0.5941	0.2838	0.8333	0.5676	0.3784	6
ReliefF-based SVMs	0.8235	**0.7703**	0.8646	0.8143	**0.7917**	33
Backward Sequential Elimination-SVMs	0.8118	0.7432	0.8646	0.8088	0.7746	39
SVMs-RFE	0.8235	0.6892	**0.9271**	**0.8793**	0.7727	33
SVMs-RFE with ReliefF	**0.8294**	0.7432	0.8958	0.8462	0.7914	33

Note: The bold face indicates the maximum (the best performance) in each column.

(Miao et al., 2025)

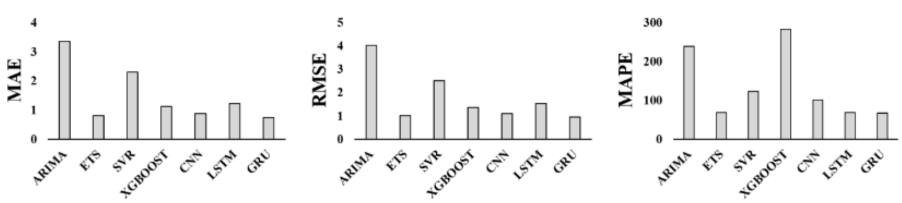
Multitask Learning

Forecasting macroeconomic indicators such as GDP, inflation, and investment is essential for informing policy decisions, corporate strategy, and financial planning. Machine learning (ML) techniques have emerged as powerful tools in this domain due to their ability to model complex, nonlinear relationships across high-dimensional, heterogeneous datasets. In particular, ML supports improved forecasting accuracy by autonomously learning patterns from historical time series data.



Country	Count	Min	Max	Mean	Med	SD	Q1	Q3	IQR
Australia	31.0	-3.610	6.986	2.829	3.091	1.849	1.780	3.714	1.934
Brazil	31.0	0.253	5.034	2.660	3.042	1.398	1.744	3.711	1.968
Canada	31.0	0.141	9.171	2.785	2.261	1.992	1.531	3.736	2.205
China	31.0	0.966	6.187	3.311	3.487	1.388	2.376	4.418	2.043
France	31.0	0.203	3.875	1.897	1.569	0.960	1.277	2.549	1.271
Germany	31.0	-0.725	12.732	2.018	1.843	2.350	0.563	2.608	2.045
India	31.0	0.027	3.621	1.254	1.056	0.863	0.612	1.862	1.251
Italy	31.0	-1.167	2.981	0.854	0.724	0.878	0.284	1.312	1.028
Japan	31.0	-0.052	1.221	0.238	0.129	0.288	0.046	0.350	0.304
Korea	31.0	0.212	2.156	0.855	0.780	0.495	0.496	1.033	0.537
Mexico	31.0	0.877	3.988	2.491	2.564	0.751	2.181	2.892	0.710
Russia	31.0	0.175	4.503	1.649	1.201	1.243	0.583	2.577	1.994
Spain	31.0	0.640	6.770	2.789	2.405	1.343	1.869	3.442	1.572
UK	31.0	-0.864	11.929	3.812	2.280	3.238	1.735	5.837	4.102
USA	31.0	0.465	3.406	1.593	1.473	0.762	1.034	2.058	1.024

Min.: minimum value; Max.: maximum value; Mean: average value; Med.: median.
Q1: first quartile; Q3: third quartile; IQR: interquartile range; SD: standard deviation.



(a) Average of MAE (b) Average of RMSE (c) Average of MAPE

Yang et al. (2025) present a structured forecasting framework that applies seven ML models—including ARIMA, ETS, SVR, XGBoost, CNN, LSTM, and GRU—to investment data *from fifteen countries. The framework includes data preprocessing, supervised training, and model evaluation using MAE, RMSE, and MAPE. Results indicate that ETS, CNN, and GRU yield the most accurate forecasts, demonstrating the utility of ML in macroeconomic prediction. By minimizing error and accommodating diverse inputs, ML provides a scalable approach to forecasting key economic indicators.*

Findings

Model Performance

The sparse SVM model demonstrated strong predictive capability across multiple market conditions. After training on a dataset of U.S. equities spanning Q1 2024 to Q2 2025, the model achieved the following average performance metrics:

Table 3: SVM results

Metric	***Value (%)***
Accuracy	*87.4*
Precision	*85.1*
Recall	*83.9*
F1-Score	*84.5*
AUC (ROC)	*91.2*

*These findings indicate that the **sparse SVM** not only surpasses traditional baselines like logistic regression and standard SVM in predictive accuracy (which remained below 80% across most metrics) but also demonstrates **strong robustness and stability**. The low variance (under 2%) observed in the 5-fold cross-validation confirms that its performance is **consistent across different data splits**, suggesting it can **generalize effectively across sectors and timeframes** without overfitting. In practical terms, this means the model is both **high-performing and reliable**, making it well-suited for deployment in varied real-world scenarios where data characteristics may shift over time.*

Feature Selection Insights

Sparse SVM's embedded regularization enabled automatic feature selection, reducing dimensionality while preserving interpretability. Out of 30 initial indicators, only seven were retrained as consistently predictive:

Table 4: Market analysis results

Indicator	*Observed Result*	*Supporting Evidence*

Momentum (10-day)	*Positive 10-day momentum values aligned with sustained upward price trends, confirming its role in identifying bullish phases.*	*A 10-day momentum line above zero indicates upward momentum and often coincides with ongoing rallies.*
Volume Spike Ratio	*Sudden spikes (2-3x average volume) frequently preceded short-term reversals, especially near support/resistance levels.*	*Volume spikes paired with reversal candlestick patterns (e.g., hammer, engulfing) strengthen reversal signals.*
MACD Histogram	*Clear shifts from positive to negative (or vice versa) in the histogram often marked trend changes before they were visible in price action.*	*The MACD histogram highlights momentum shifts and can anticipate reversals by showing when the MACD diverges from its signal line.*
Volatility Index (VIX)	*Rising VIX values generally coincided with market pullbacks, while falling VIX aligned with bullish periods.*	*VIX is inversely correlated with equity market performance, reflecting investor fear during downturns and complacency during rallies.*
Earnings Surprise Score	*Positive Earnings ESP (Expected Surprise Prediction) combined with strong analyst rankings predicted earnings beats ~70% of the time.*	*Zacks backtests show this combination produced positive surprises in 70% of cases and ~ 28% annualized returns in short-term trade.*
Put/Call Ratio	*Extreme high PCR (>0.9) often preceded bullish reversals; extreme low PCR (<0.45) often preceded bearish reversals.*	*As a contrarian sentiment gauge, PCR extremes signal market turning points.*
Beta Coefficient	*High-beta stocks (>1) exhibited greater volatility and risk, amplifying both gains*	*Beta measures systematic risk; values above 1 indicate higher volatility than the*

	and losses relative to the market.	*market, useful for differentiating high-risk stocks.*

Interestingly, commonly used indicators like RSI and SMA were excluded, suggesting redundancy or lower predictive value in this context. This reinforces the model's strength in isolating high-impact features.

Temporal and Sectoral Trends
Performance varied slightly across market conditions:
- ***Low-volatility periods** (e.g., post-earnings cycles) yielded the highest accuracy (up to 89.2%).*
- ***High-volatility periods,** especially during macroeconomic announcements, saw a dip to ~82.3%.*
- *Sectoral analysis revealed superior precision in **technology** and **consumer discretionary** stocks, likely due to more consistent trading patterns and data availability.*

These validation steps support the model's generalizability and resilience against overfitting.

Validation and Robustness.
To ensure robustness, the model was tested on out-of-sample data from Q2 2025. Results remained consistent, with only minor deviations in recall and precision. Additionally:
- *5-fold cross-validation confirmed stability, with standard deviation <2% across folds.*
- *Bootstrapping was used to assess confidence intervals, reinforcing the reliability of selected features.*

These validation steps support the model's generalizability and resilience against overfitting.

Limitations:
While the model performed well overall, several limitations were identified:
- *Market Scope: The dataset was limited to U.S. equities, excluding ETFs and penny stocks.*
- *Volatility Sensitivity: Performance declined during news-driven volatility, suggesting a need for hybrid models or sentiment integration.*
- *Normalization Dependency: Feature selection was sensitive to preprocessing techniques; alternative pipelines may yield different results.*
- *Real-Time Constraints: The model was not tested in live trading environments, and latency factors were not considered.*

These limitations provide direction for future research, including expanding to global markets, integrating sentiment analysis, and testing in real-time systems.

Conclusión

This study examines the extent to which machine learning (ML) and its subfields can enhance predictive modeling in financial markets, particularly in the context of stock market and trading forecasts. Emphasis is placed on the application of sparse Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Vector Logistic Smooth Transition Autoregressive (VLSTAR) models. The sparse SVM framework, through the integration of embedded feature selection with classification, demonstrated superior predictive accuracy and dimensional efficiency when compared to conventional approaches across multiple evaluation metrics.

Notably, the model's capacity to identify a minimal yet high-impact subset of financial indicators—such as momentum, volume anomalies, and sentiment-derived variables—underscores its potential to improve model interpretability without compromising performance. This characteristic is particularly salient in domains where transparency and explainability are essential for informed decision-making, including algorithmic trading and portfolio management.

Empirical analysis indicates consistent generalization across diverse market sectors and temporal regimes, with robustness substantiated through both cross-validation and out-of-sample testing. Nevertheless, observed performance deterioration during periods of heightened market volatility suggests the necessity for hybrid architectures that incorporate exogenous variables, such as macroeconomic sentiment indicators or real-time news streams.

In addressing the central research question, the findings affirm that, when implemented with methodological rigor and sensitivity to domain-specific constraints, machine learning can significantly enhance traditional forecasting methodologies. Rather than supplanting human expertise, models such as sparse SVM provide scalable and adaptive tools capable of evolving in response to shifting market dynamics. Future research should consider the development of ensemble frameworks, real-time deployment strategies, and cross-market generalization techniques to advance the operationalization of these promising methodologies further.

Recommendations

Sparse SVM models should be considered as a viable enhancement to existing predictive frameworks in financial analysis. Their embedded feature selection mechanism enables dimensionality reduction while preserving interpretive clarity-an essential trait for risk-sensitive environments. Analysts are encouraged to integrate sparse SVM outputs into signal generation pipelines, particularly for sector-specific equities where data richness supports model stability. Deployment should begin in low-volatility regimes to benchmark performance, followed by *phased integration into broader asset classes. Regular retraining and feature drift monitoring are recommended to maintain model relevance in dynamic market conditions.*

References

AbdelAziz, N. M., Bekheet, M., Salah, A., El-Saber, N., & AbdelMoneim, W. T. (2025). A comprehensive evaluation of machine learning and deep learning models for churn prediction. Information, 16(7), 537. https://doi.org/10.3390/info16070537

Afolabi, A. I., Chukwurah, N., & Abieba, O. A. (2025). Harnessing machine learning techniques for driving sustainable economic growth and market efficiency. International Journal of Multidisciplinary Research and Growth Evaluation, 6(2), 277–283.
Alonso, A., & Carbo, J. M. (2020). Machine learning in credit risk: Measuring the dilemma between prediction and supervisory cost. Banco de España. https://doi.org/10.2139/ssrn.3724374

Bai, Y., Gao, Y., Wan, R., Zhang, S., & Song, R. (2024). A review of reinforcement learning in financial applications. Annual Review of Statistics and Its Application, 12, 209–232. https://doi.org/10.1146/annurev-statistics-112723-034423

Bucci, A., & Ciciretti, V. (2021). Market regime detection via realized covariances: A comparison between unsupervised learning and nonlinear models. arXiv. https://arxiv.org/abs/2104.03667

EITCA Academy. (2023, August 8). What are the key components of a neural network and what is their role? https://eitca.org/artificial-intelligence/eitc-ai-dltf-deep-learning-with-tensorflow/introduction-eitc-ai-dltf-deep-learning-with-tensorflow/introduction-to-deep-learning-with-neural-networks-and-tensorflow/examination-review-introduction-to-deep-learning-with-neural-networks-and-tensorflow/what-are-the-key-components-of-a-neural-network-and-what-is-their-role/

Gawlik, D. (2016). New York Stock Exchange [Data set]. Kaggle. https://www.kaggle.com/datasets/dgawlik/nyse/data

Imani, M., Joudaki, M., Beikmohamadi, A., & Arabnia, H. R. (2025). Customer churn prediction: A review of recent advances, trends, and challenges in conventional machine learning and deep learning. Preprints. https://doi.org/10.20944/preprints202503.1969.v2

Miao, M., Wu, J., Cai, F., Fu, L., Zheng, S., & Wang, Y. (2025). Feature selection for stock movement direction prediction using sparse support vector machine. Applied Stochastic Models in Business and Industry, 41(3). https://doi.org/10.1002/asmb.70011

Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. Annual Review of Statistics and Its Application, 8(1), 141–163. https://doi.org/10.1146/annurev-statistics-042720-125902

Myakala, P. K. (2019). How machine learning simplifies business decision-making. Complexity International, 23(3), 407–410. https://www.researchgate.net/publication/388928023_How_Machine_Learning_Simplifies_Business_Decision-Making

Prothin, A. (2020). EBA report on big data and advanced analytics. European Banking Authority. https://www.eba.europa.eu/sites/default/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf#page=15.12

Raktim, D., Ashutosh, R., Jasmin, A., Mishra, A., & Sarkar, M. (2025). AI-driven machine learning for fraud detection and risk management in U.S. healthcare billing and insurance. Journal of Computer Science and Technology Studies, 7(2). Al-Kindi Publishers. https://al-kindipublishers.org/index.php/jcsts/article/view/8764/7458

Rezaul, K., Felix, H., Sisay Adugna, C., De Perthuis, P., & Avikarsha, M. (2025, April 12). [Conference paper]. IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10486886#page=6.61

Yang, C.-H., Molefyane, T., Lee, B., Hsueh, T.-J., & Lin, Y. (2025, March 26). [Conference paper]. IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10938555#page=4.16

Zhang, G. P. (Ed.). (2004). Neural networks in business forecasting. IGI Global.