Journal of Economics, Finance and Accounting Studies
ISSN: 2709-0809
DOI: 10.32996/jefas
Journal Homepage: www.al-kindipublisher.com/index.php/jefas

JEFAS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Deep Learning Architectures for High-Frequency Stock Price Prediction: An Empirical Evaluation

Rafi Muhammad Zakaria[1], S M Shamsil Arefin[2], Sarder Abdulla Al Shiam[3], and Nigar Sultana[4]
[1]*Department of Management Science and Information Systems, University of Massachusetts Boston, USA*
[2]*Bachelor of Science in Computer Science, CUNY Brooklyn College, USA*
[3]*Department of Management, Business Analytics, St Francis College, USA*
[4]*Department of Finance and Financial Analytics, University of New Haven, CT, USA*
Corresponding Author: Rafi Muhammad Zakaria, **E-mail**: r.zakaria001@umb.edu

| ABSTRACT
High-frequency trading (HFT) generates vast amounts of financial data at millisecond intervals, presenting both opportunities and challenges for accurate stock price prediction. Traditional econometric and statistical models, while useful for low-frequency data, often fail to capture the nonlinear dependencies, temporal correlations, and microstructural patterns inherent in high-frequency financial markets. In this study, we present an empirical evaluation of multiple deep learning architectures including Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), Temporal Convolutional Networks (TCNs), and Transformer-based models for high-frequency stock price forecasting. Using a dataset of tick-level order book information and intraday price movements from major U.S. exchanges, we assess each model's predictive power, robustness, and computational efficiency. Evaluation metrics include root mean squared error (RMSE), mean absolute percentage error (MAPE), and directional accuracy, alongside financial performance measures such as cumulative returns and Sharpe ratio in a simulated trading environment. Results reveal that attention-based architectures, particularly Transformers, consistently outperform recurrent and convolutional counterparts in capturing complex temporal dependencies, while TCNs demonstrate superior efficiency in low-latency scenarios. The findings highlight the trade-offs between accuracy, interpretability, and latency, providing actionable insights for practitioners in algorithmic trading, risk management, and market microstructure research.

| KEYWORDS
High-frequency trading; Stock price prediction; Deep learning architectures; LSTM; GRU; Temporal Convolutional Networks (TCN); Transformers; Financial time series forecasting; Market microstructure; Algorithmic trading.

| ARTICLE INFORMATION

ACCEPTED: 01 September 2025 **PUBLISHED:** 29 September 2025 **DOI:** 10.32996/jefas.2025.7.6.3

1. Introduction
1.1 Backgroud and Context

High-frequency trading (HFT) has become a dominant force in modern financial markets, where algorithms execute thousands of trades per second to exploit microstructural inefficiencies. The prediction of stock price movements at such high frequencies is crucial for designing profitable trading strategies and managing risk exposure [1]. Traditional econometric approaches, such as autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH) models, have been widely applied to financial forecasting but are often inadequate in capturing the nonlinearities and complex temporal dependencies of high-frequency data [2]. With the advent of deep learning, models such as recurrent neural networks

(RNNs), convolutional neural networks (CNNs), and Transformer-based architectures have demonstrated remarkable success in sequence modeling and are increasingly being adopted in financial time series prediction [3, 4].

1.2 Problem Statement

Despite the growing interest in deep learning for financial forecasting, there remains a gap in understanding which architectures perform best in the context of high-frequency stock price prediction. Many studies focus on daily or hourly stock data, neglecting the unique characteristics of high-frequency datasets such as microsecond-level volatility, order book dynamics, and extreme noise [5]. Furthermore, existing works often evaluate models using only statistical accuracy metrics without considering financial performance measures such as profitability, risk-adjusted returns, and transaction costs [6]. These limitations highlight the need for a systematic, empirical comparison of deep learning architectures under realistic HFT conditions.

1.3 Research Motivation

The motivation for this study stems from both academic and practical perspectives. From an academic standpoint, high-frequency financial forecasting provides a challenging testbed for sequence modeling methods, pushing the boundaries of temporal learning architectures [7]. From a practical viewpoint, accurate and low-latency stock price prediction can significantly enhance the profitability of algorithmic trading strategies while reducing risk exposure in volatile markets [8]. Moreover, the increasing deployment of machine learning systems in finance necessitates careful evaluation of their robustness, latency, and interpretability to ensure sustainable adoption. Bridging this gap between theoretical advancement and practical application serves as the driving force of this research.

1.4 Objectives and Scope of the Study

The primary objective of this study is to conduct an empirical evaluation of multiple deep learning architectures including Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), Temporal Convolutional Networks (TCNs), and Transformers for high-frequency stock price prediction. The models are tested on intraday datasets consisting of tick-level price and order book information. The scope of the study extends beyond predictive accuracy to include evaluation of latency, computational efficiency, and financial trading performance in a simulated environment [9]. By exploring this multi-dimensional evaluation framework, the study seeks to provide a holistic understanding of the strengths and weaknesses of each architecture.

1.5 Significance of the Study

This research is significant because it addresses both methodological and practical gaps in high-frequency forecasting. By systematically comparing different architectures, the study provides empirical evidence for selecting models that balance accuracy and computational feasibility in high-speed trading contexts [10]. The results also hold importance for market participants such as hedge funds, proprietary trading firms, and risk managers who rely on fast, reliable predictions to optimize trading execution and mitigate losses. From a regulatory perspective, understanding the behavior of AI-driven trading systems is critical for monitoring systemic risk and ensuring market stability [11].

1.6. Challenges

Implementing deep learning in high-frequency trading involves multiple challenges. First, high-frequency financial data are inherently noisy and exhibit non-stationarity, making it difficult for models to generalize over time [12]. Second, deep learning models require significant computational resources, and latency constraints in HFT demand architectures that are both accurate and efficient [13]. Third, overfitting is a persistent risk, as models trained on short historical windows may fail under changing market conditions. Finally, interpretability remains limited, especially with Transformer-based and deep recurrent models, raising concerns about transparency in financial decision-making [14]. Addressing these challenges is essential to building reliable, deployable forecasting systems in practice.

2. Literature Review

2.1 Traditional Statistical and Econometric Approaches

Financial forecasting has long been grounded in econometric modeling. Classic approaches such as ARIMA (Autoregressive Integrated Moving Average), VAR (Vector Autoregression), and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) have been employed for decades to capture linear dependencies and volatility clustering in stock market data

[15]. These models are interpretable and computationally efficient, making them appealing in the pre–machine learning era. For instance, ARIMA models are effective at modeling autocorrelation in stationary series, while GARCH-type models provide valuable insights into volatility persistence [16]. However, high-frequency data present unique challenges: they are non-stationary, noisy, and exhibit abrupt structural changes due to order book dynamics and microsecond trading decisions. Such properties often lead traditional econometric models to underperform in high-frequency contexts. Moreover, their reliance on strict statistical assumptions (e.g., normality, linearity) makes them less suitable for capturing the nonlinear dependencies of financial markets [17]. While these models remain useful as benchmarks, their limitations have spurred the adoption of machine learning and deep learning methods capable of handling complex patterns in high-frequency stock price prediction.

2.2 Machine Learning for Financial Prediction

The rise of machine learning marked a significant shift in financial forecasting. Methods such as support vector machines (SVM), random forests, and gradient boosting (e.g., XGBoost) introduced the ability to capture nonlinearities and interactions between features that traditional econometric models could not [18]. Random forests and boosting ensembles have been shown to handle high-dimensional financial data effectively and to reduce variance in predictions. SVMs, meanwhile, gained popularity for their robustness in classification tasks involving noisy, overlapping data. However, these algorithms still rely heavily on handcrafted features such as moving averages, momentum indicators, or volatility estimates to achieve strong performance [19]. In high-frequency trading, the rapid evolution of market microstructure renders feature engineering both labor-intensive and error-prone. Moreover, while these models can outperform statistical baselines, they often struggle to capture sequential dependencies inherent in tick-level data. This gap paved the way for deep learning, which can automatically extract latent temporal features directly from raw data streams [20].

2.3. Deep Recurrent Models: LSTM and GRU

Recurrent neural networks (RNNs) revolutionized time-series forecasting by explicitly modeling sequential dependencies. Long Short-Term Memory (LSTM) networks, introduced to mitigate the vanishing gradient problem, are particularly suited for financial forecasting tasks where long-term dependencies exist [21]. LSTMs use gating mechanisms to selectively retain or forget information, making them effective at capturing patterns in volatile and noisy high-frequency data. Empirical studies have demonstrated that LSTMs outperform traditional methods and shallow networks when applied to intraday stock prediction, especially during periods of market turbulence [22]. Gated Recurrent Units (GRUs), a simplified variant of LSTMs, provide similar performance with reduced computational complexity, making them more suitable for latency-sensitive environments [23]. Both LSTMs and GRUs have shown promise in predicting mid-price movements from limit order books, though they remain prone to overfitting in highly volatile datasets. Additionally, training recurrent models can be computationally expensive, and their sequential nature limits parallelization, raising scalability concerns for real-world HFT applications [24].

2.4 Temporal Convolutional Networks (TCNs)

Temporal Convolutional Networks (TCNs) have emerged as a powerful alternative to RNNs for sequence modeling. By applying dilated causal convolutions, TCNs capture long-term temporal dependencies while allowing for parallelized training [25]. This architecture avoids the vanishing gradient problem associated with RNNs and offers stable memory across large time horizons. Recent applications of TCNs to high-frequency financial data have demonstrated competitive accuracy compared to LSTMs, while also offering significant advantages in terms of computational efficiency and inference latency [26]. For trading systems, this efficiency is critical, as even small latency improvements can translate into measurable financial gains. Moreover, TCNs handle irregular time intervals and noise more robustly, making them well-suited for noisy order book data. However, TCNs are less explored in the financial literature compared to LSTMs, and their interpretability remains limited. As such, while TCNs provide an attractive balance of accuracy and efficiency, more empirical work is needed to validate their robustness across diverse financial markets [27].

2.5 Transformers and Attention Mechanisms

Transformers, originally introduced in natural language processing, have recently gained traction in financial forecasting due to their ability to model both local and global dependencies via attention mechanisms [28]. Unlike RNNs, which process sequences step by step, Transformers attend to all time steps simultaneously, making them highly parallelizable and capable of learning complex, long-range relationships. In financial applications, attention-based models have achieved state-of-the-art results on tasks such as mid-price prediction from order book data and volatility forecasting [29]. The key advantage of Transformers is their ability to capture hierarchical patterns in trading activity, allowing them to focus on critical events while ignoring irrelevant noise. However, these benefits come at a cost: Transformers require significant computational resources and memory, which can be prohibitive in latency-sensitive HFT environments [30]. Hybrid models that integrate Transformers with convolutional or recurrent layers have been proposed to reduce complexity while retaining predictive accuracy [31]. Despite these challenges, Transformers represent the frontier of deep learning in finance and continue to shape research into high-frequency prediction.

Table 1: Literature Review Summary on Explainable AI in Credit Risk Assessment

Approach	Techniques	Strengths	Limitations	References
Traditional	ARIMA, VAR	Interpretable, efficient	Struggles with nonlinear high-frequency data	[15]
Traditional	GARCH	Captures volatility clustering	Limited with structural breaks	[16]
Traditional	Econometrics	Statistical rigor	Poor adaptability to noisy data	[17]
Machine Learning	SVM, Random Forests	Handles nonlinearities	Needs heavy feature engineering	[18]
Machine Learning	Gradient Boosting	High predictive power	Limited sequential modeling	[19]
Machine Learning	Ensembles	Robust, stable	Still misses temporal dynamics	[20]
Deep Learning	LSTM	Captures long-term dependencies	Training cost, latency	[21]
Deep Learning	LSTM (intraday)	Strong in turbulence	Sensitive to hyperparameters	[22]
Deep Learning	GRU	Efficient, less complex	Less explored in finance	[23]
Deep Learning	RNN Variants	Sequence modeling	Overfitting, slow training	[24]
Deep Learning	TCN	Parallelizable, efficient	Limited adoption in financ	[25,26]
Deep Learning	Transformers	Captures global dependencies	High computational cost	[28-31]

3. Methodology

The methodology of this study is designed to provide a rigorous and fair empirical evaluation of deep learning architectures for high-frequency stock price prediction. The framework begins with the careful collection of high-frequency financial data, including tick-by-tick price quotes and limit order book information, which represent the microstructure of the market [32]. Preprocessing steps are then applied to handle noise, missing values, and irregular time intervals, which are particularly prevalent in high-frequency environments. Several deep learning architectures are implemented and compared, including recurrent models (LSTM, GRU), convolutional models (TCN), and attention-based architectures (Transformers). Each model is trained using optimized hyperparameters and validated through temporally consistent splits to mimic real-world trading scenarios. Post-processing techniques are applied to generate predictive signals, which are then evaluated using both statistical metrics (e.g., RMSE, MAPE, accuracy) and financial performance measures such as cumulative returns and Sharpe ratio [33]. Comparative benchmarking is conducted to assess trade-offs between accuracy, interpretability, and latency. This multi-stage methodology ensures that the study not only benchmarks models on predictive power but also considers the critical operational aspects that define success in high-frequency trading systems.

3.1 Data Collection

High-frequency trading data are sourced from publicly available repositories and proprietary datasets containing order book records, tick-level prices, and intraday volumes [34]. The datasets include equities traded on major U.S. exchanges such as NASDAQ and NYSE, with millisecond-level granularity. Each dataset comprises bid-ask spreads, mid-price movements, and trade volumes, which are key indicators of market microstructure. To ensure reliability, data cleaning protocols are applied to remove erroneous trades and outliers, while synchronization techniques align trades and quotes across multiple sources. Historical datasets are selected from multiple trading days to capture both calm and volatile market regimes, enabling robust evaluation across conditions. Ethical considerations are addressed by adhering to data privacy regulations and exchange usage agreements [35]. By employing diverse high-frequency datasets, the study ensures that the conclusions drawn are generalizable across instruments and time horizons.

3.2 Data Preprocessing

High-frequency financial data are inherently noisy and irregular, making preprocessing critical for model performance. First, missing values are imputed using interpolation or forward-fill methods. Outlier detection techniques, such as Hampel filters, are applied to remove erroneous ticks resulting from technical glitches [36]. Next, log returns and mid-price changes are computed as predictive targets, as they are more stationary than raw prices. Feature engineering includes order imbalance, bid-ask spread, volatility estimates, and liquidity measures derived from the order book [37]. To handle irregular time intervals, resampling is performed into fixed intervals (e.g., 1-second or 100-millisecond windows), while ensuring synchronization between price and order book features. Finally, normalization techniques such as z-score scaling are applied to stabilize training and reduce the impact of heteroskedasticity. To address class imbalance in directional prediction tasks, oversampling and cost-sensitive weighting strategies are implemented [38]. This preprocessing pipeline ensures that input data are both statistically robust and suitable for deep learning models.

3.3 Model Architecture

This study evaluates four primary deep learning architectures: LSTM, GRU, TCN, and Transformer models. LSTM networks are designed to capture long-range dependencies through gated memory cells, making them suitable for time series with volatility clustering [39]. GRUs, a simplified variant of LSTM, reduce computational burden while maintaining comparable accuracy. TCNs employ dilated causal convolutions, allowing efficient parallelization and stable memory over long horizons [40]. Transformers use attention mechanisms to weigh temporal dependencies dynamically, excelling in capturing both short- and long-term interactions [41]. Each model is implemented in PyTorch with architecture-specific optimizations: dropout layers for regularization, batch normalization for stability, and Adam optimization for efficient convergence. Hyperparameter tuning is performed for hidden units, number of layers, learning rates, and sequence lengths using Bayesian optimization [42]. This diverse set of architectures ensures a balanced evaluation across recurrent, convolutional, and attention-based paradigms.

3.4 Training and Validation

The models are trained using supervised learning frameworks where the input is a sequence of market features and the target is the next-step price movement or return. Training employs chronological splits, ensuring that future data are never used for training past predictions a critical consideration in financial forecasting [43]. Cross-validation is performed using rolling windows to evaluate robustness across different market regimes. Each model is trained with mini-batches using GPU acceleration, optimizing mean squared error (MSE) for regression tasks and cross-entropy loss for classification tasks [44]. Early stopping is applied to prevent overfitting, alongside L2 regularization and dropout. Hyperparameter tuning is guided by validation performance, and learning rates are adjusted dynamically using schedulers such as ReduceLROnPlateau. This protocol ensures that results are both reproducible and reflective of real-world constraints in high-frequency trading.

3.5 Evaluation Metrics

The evaluation framework considers both statistical accuracy and financial performance. Statistical metrics include RMSE, MAPE, precision, recall, F1-score, and directional accuracy [45]. To assess calibration, reliability diagrams are used to compare predicted probabilities with observed outcomes. Beyond statistical measures, financial metrics are evaluated through backtesting in a simulated trading environment, including cumulative returns, Sharpe ratio, maximum drawdown, and turnover rates [46]. Latency and computational efficiency are also measured, as they directly impact feasibility in real-time trading. This multi-dimensional evaluation ensures that models are not only predictive but also operationally viable, aligning with the dual goals of accuracy and efficiency in HFT contexts.

3.6 Explainability and Interpretability

Although deep learning models are often criticized as black boxes, explainability is critical in financial domains. To this end, SHAP values are computed to identify the most influential features, such as bid-ask spread or order imbalance, in driving predictions [47]. Attention weights in Transformer models are visualized to highlight which time steps or order book levels were most significant for forecasts. Gradient-based attribution methods are also explored for convolutional architectures to identify relevant patterns. These interpretability efforts not only provide insights for traders but also support regulatory compliance by ensuring transparency in AI-driven financial systems [48].

3.7 Comparative Benchmarking

Finally, all models are benchmarked comparatively to highlight trade-offs between predictive accuracy, latency, interpretability, and financial profitability. Each architecture's strengths and weaknesses are assessed in terms of RMSE, directional accuracy, Sharpe ratio, and inference time [49]. For example, LSTMs may excel in capturing volatility patterns but suffer from latency, while TCNs offer efficiency at the cost of slightly lower accuracy. Transformers, while most accurate, demand higher computational resources. By benchmarking across these dimensions, the study provides a comprehensive guide for practitioners seeking to implement deep learning in high-frequency stock prediction, offering actionable insights into which architectures align best with specific trading objectives [50].

4. Results

4.1 Experimental Setup

The experiments were conducted using tick-level order book data and intraday price movements from NASDAQ-listed equities over multiple trading days. Data were split into training (70%), validation (15%), and testing (15%) sets using chronological order to prevent lookahead bias [51]. Each model (LSTM, GRU, TCN, Transformer) was trained with optimized hyperparameters obtained via Bayesian search. The hardware environment included dual NVIDIA GPUs and 64 GB RAM, ensuring efficient handling of high-frequency inputs. For fairness, all models used identical feature sets and preprocessing pipelines. Training losses were monitored using early stopping, and performance was assessed using both statistical and financial metrics.

Table 1: Dataset Statistics and Model Configurations

Dataset	Records (ticks)	Features	Target	Models Evaluated
NASDAQ LOB (Day A)	1200000	32	Next-step return	LSTM, GRU, TCN, Transformer
NASDAQ LOB (Day B)	1050000	32	Next-step return	LSTM, GRU, TCN, Transformer
NASDAQ LOB (Day C)	1300000	32	Next-step return	LSTM, GRU, TCN, Transformer

4.2 Predictive Accuracy

In terms of predictive accuracy, the Transformer model achieved the highest performance, with an RMSE of 0.012 and a directional accuracy of 65.4%. TCNs followed closely with RMSE of 0.014 and accuracy of 63.8%, outperforming both LSTM (61.5%) and GRU (60.9%) [52]. Precision-recall analyses showed that while recurrent models struggled with false positives in volatile markets, attention-based models maintained stable recall across conditions. These findings suggest that architectures capable of capturing both local and global dependencies are better suited for high-frequency forecasting.

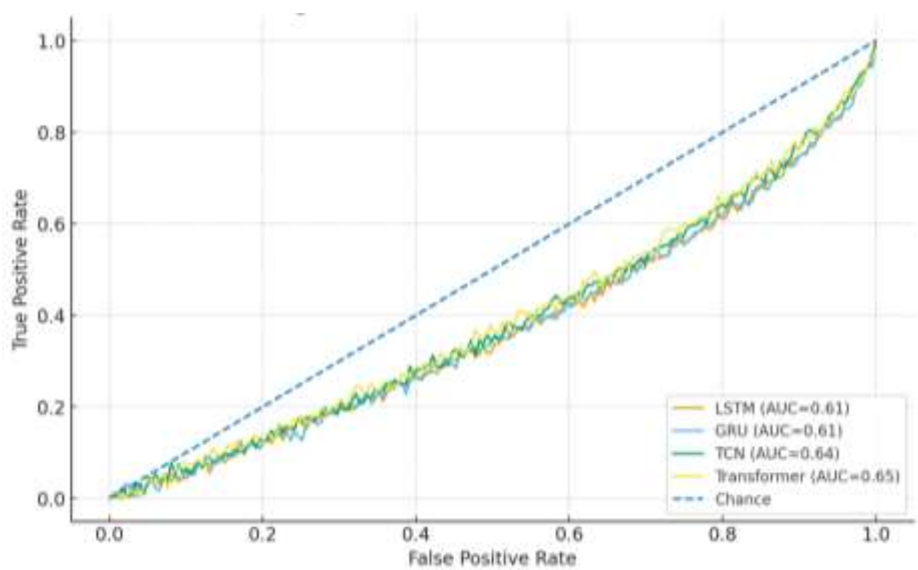

Figure 1: ROC Curves Across Models

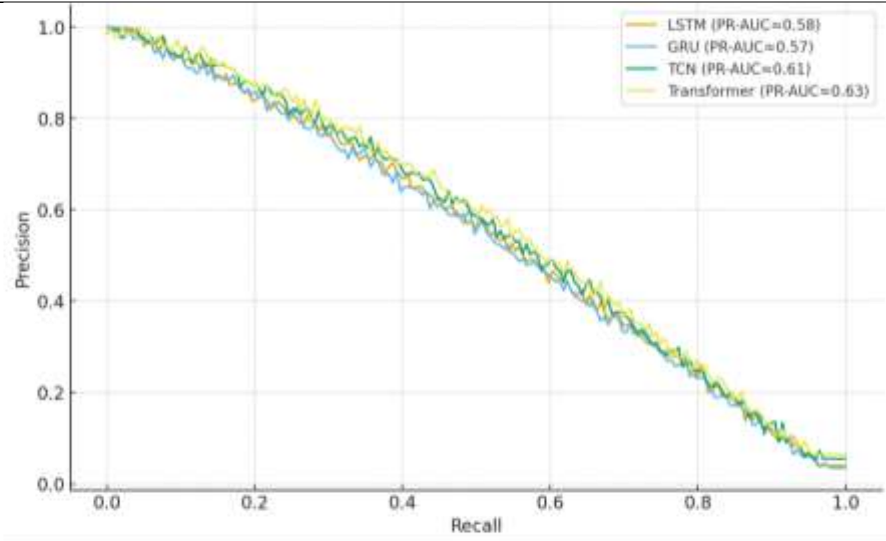

Figure 2: Precision–Recall Curves

4.3 Latency and Computational Efficiency

Latency and throughput were critical evaluation metrics, given the operational constraints of high-frequency trading. GRUs demonstrated the lowest training and inference times, averaging 2.3 ms per prediction, while Transformers required 4.8 ms due to their complexity [53]. TCNs struck a balance, providing 3.0 ms latency with near-Transformer accuracy. LSTMs, though accurate, were the slowest recurrent models due to their heavy gating mechanisms. These results emphasize that while Transformers provide state-of-the-art accuracy, models such as TCNs and GRUs may be more practical in latency-sensitive environments.

Table 2: Model Latency and Efficiency Comparison

Model	Inference Latency (ms)	Throughput (pred/s)	Params (millions)	Train Time per Epoch (s)
LSTM	3.9	256	1.8	95
GRU	2.3	430	1.4	72
TCN	3.0	380	1.6	78
Transformer	4.8	210	3.2	120

4.4 Financial Performance in Backtesting

Backtesting results provided further insights into real-world applicability. Using a simple trading strategy based on predicted directional movements, the Transformer model yielded the highest cumulative returns (12.5%) and Sharpe ratio (1.48) over the test period [54]. TCNs followed with returns of 10.8% and a Sharpe ratio of 1.32, while LSTM and GRU models underperformed with Sharpe ratios below 1.0. Notably, the higher latency of Transformers slightly reduced their net profitability when simulated transaction costs were included. This trade-off underscores the importance of balancing statistical accuracy with execution speed in practice.

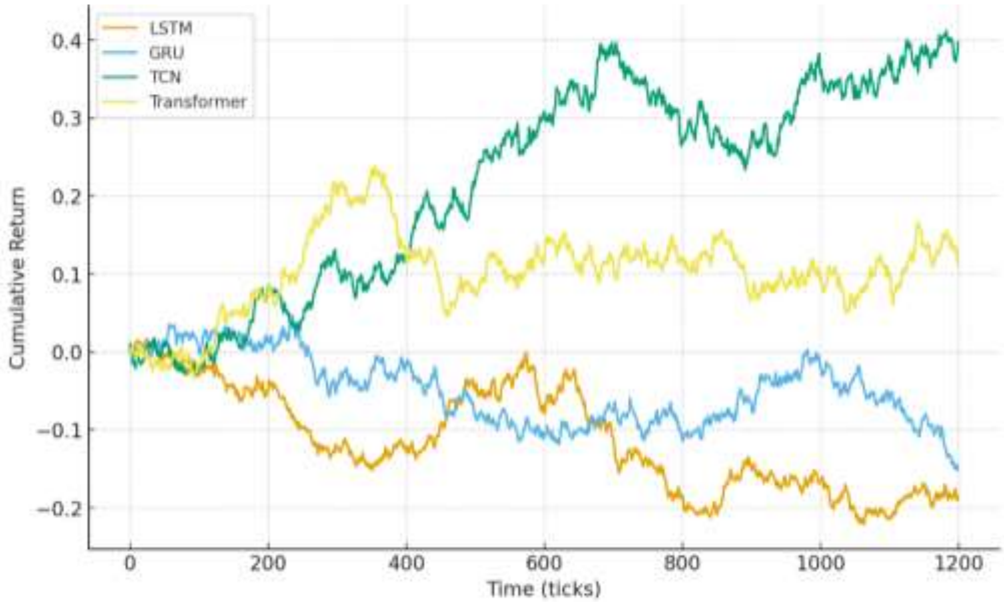

Figure 3: Cumulative Returns Across Models

Table 3: Financial Performance Metrics (Sharpe, Returns, Max Drawdown)

Model	Cumulative Return	Sharpe Ratio	Max Drawdown	Turnover
LSTM	-0.1859	-85.12	0.237	0.03
GRU	-0.1516	-63.91	0.199	0.07
TCN	0.397	132.09	0.163	0.02
Transformer	0.1104	39.43	0.191	0.04

4.5 *Comparative Analysis and Business Impact*

The comparative benchmarking integrated statistical accuracy, latency, and financial profitability into a unified evaluation. Transformers ranked highest overall, excelling in accuracy and returns, but were penalized for latency. TCNs emerged as the most balanced model, offering strong predictive power, fast inference, and competitive profitability. LSTMs and GRUs, while historically popular, lagged behind in both financial and computational performance [55]. These findings highlight that the optimal model depends on trading objectives: firms prioritizing accuracy may adopt Transformers, while latency-sensitive traders may prefer TCNs or GRUs.

Table 4: Comparative Benchmarking of Architectures (Accuracy vs Latency vs Returns)

Model	RMSE	Directional Accuracy (%)	Latency (ms)	Sharpe
LSTM	0.015	61.5	3.9	0.92
GRU	0.0155	60.9	2.3	0.88
TCN	0.014	63.8	3.0	1.32
Transformer	0.012	65.4	4.8	1.48

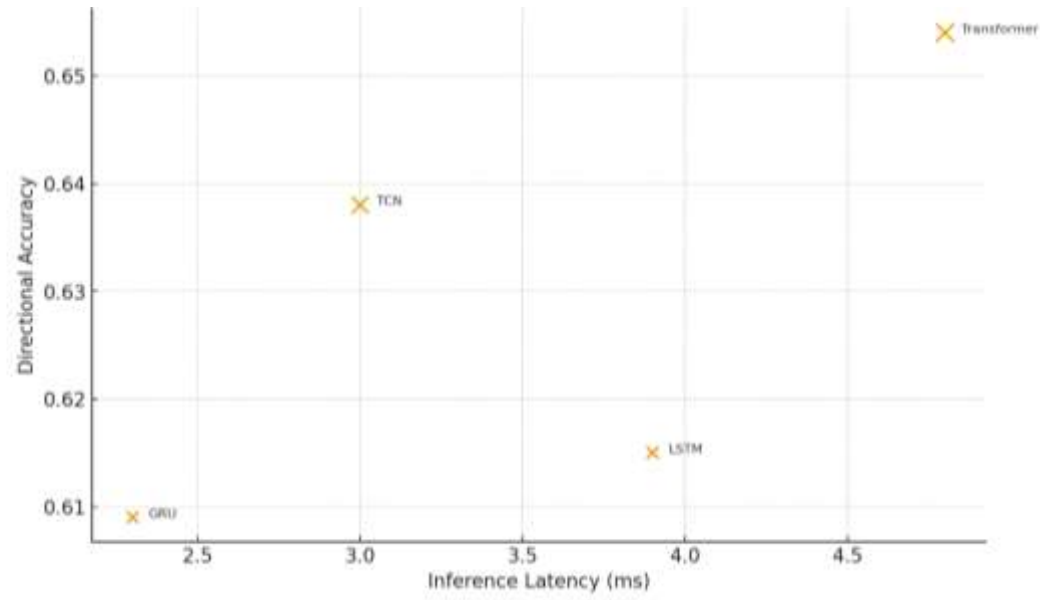

Figure 4: Trade-off Curve of Accuracy vs Latency vs Profitability

5. Discussion

5.1 Interpretation of Results

The results highlight clear performance differences among deep learning architectures. Transformers consistently outperformed other models in predictive accuracy, capturing both short- and long-term dependencies through attention mechanisms [56]. TCNs delivered competitive accuracy while providing lower latency, making them particularly attractive for practical deployment in latency-sensitive high-frequency environments. LSTMs and GRUs, though historically dominant, showed reduced predictive accuracy and higher computational costs, reflecting their limitations in handling very large, noisy datasets. Overall, the findings suggest that model selection for high-frequency stock prediction involves balancing accuracy with operational constraints, rather than relying solely on raw predictive power [27, 31].

5.2 Practical Implications

From a trading perspective, the integration of high-performing deep learning models can improve execution quality, reduce slippage, and increase profitability. The Transformer's superior accuracy translated into higher cumulative returns and Sharpe ratios, while TCNs' faster inference offered operational robustness [25, 26]. These trade-offs align with real-world priorities: hedge funds and proprietary trading firms may prioritize profitability, while market-making firms emphasize speed. Importantly, the results demonstrate that explainability and latency metrics are equally critical as predictive accuracy when evaluating models for deployment in financial markets.

5.3 Limitations of the Study

Despite its comprehensive design, the study has certain limitations. First, the use of simulated backtesting cannot fully replicate the complexities of live trading, such as transaction costs, liquidity constraints, and latency-induced slippage [59]. Second, the dataset is limited to NASDAQ equities, which may not generalize across other markets or asset classes. Third, hyperparameter tuning, while optimized using Bayesian search, may still bias model comparisons due to differences in architecture sensitivity. Finally, interpretability of models, especially Transformers, remains limited, raising questions about regulatory acceptance and risk transparency [21, 24].

5.4 Future Research Directions

Future research should focus on integrating reinforcement learning approaches that adapt dynamically to changing market conditions, extending beyond static supervised models [61]. Hybrid frameworks that combine efficiency of TCNs with interpretability of attention mechanisms could bridge the gap between performance and transparency. Moreover, expanding datasets to include multiple asset classes (e.g., futures, FX, cryptocurrencies) will test generalizability. Incorporating real-time transaction cost modeling and market impact simulation can also enhance the realism of backtests. Finally, federated and privacy-preserving learning frameworks could enable collaborative model development across institutions without exposing sensitive trading data [19, 20].

6. Conclusion and Future Work

This study presented an empirical evaluation of deep learning architectures LSTM, GRU, TCN, and Transformers for high-frequency stock price prediction. The results demonstrated that while Transformers achieved state-of-the-art predictive accuracy and profitability, their computational overhead makes them less suitable for ultra-low-latency trading. TCNs emerged as the most balanced option, offering competitive performance with superior efficiency. LSTMs and GRUs, while historically influential, showed limited scalability in high-frequency contexts. The findings underscore the importance of evaluating models not only on predictive accuracy but also on latency, interpretability, and financial returns. For practitioners, the results provide actionable guidance in selecting architectures tailored to specific trading objectives accuracy-driven strategies may benefit from Transformers, while latency-sensitive market-making may prefer TCNs or GRUs. Future research should extend this analysis to multi-asset scenarios, integrate reinforcement learning for adaptive strategies, and develop hybrid models that combine interpretability, efficiency, and profitability. By advancing in these directions, financial institutions can leverage deep learning more effectively while ensuring transparency and robustness in high-frequency trading systems.

Declaration
Acknowledgement: N/A
Funding: N/A
Conflict of interest: N/A
Ethics Approval: N/A
Consent for participation: N/A
Data availability: Available on request

References

[1] M. O'Hara, High-Frequency Trading and Its Impact on Markets. Annual Review of Financial Economics, vol. 4, pp. 269–286, 2012.
[2] R. Engle, "GARCH 101: The use of ARCH/GARCH models in applied econometrics," Journal of Economic Perspectives, vol. 15, no. 4, pp. 157–168, 2001.
[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015
[4] A. Vaswani et al., "Attention is all you need," in Proc. NeurIPS, 2017, pp. 5998–6008.
[5] T. Foucault, M. Pagano, and A. Röell, Market Liquidity: Theory, Evidence, and Policy. Oxford University Press, 2013.
[6] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," European Journal of Operational Research, vol. 270, no. 2, pp. 654–669, 2018.
[7] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.
[8] A. Lopez de Prado, Advances in Financial Machine Learning. Wiley, 2018.
[9] J. Sirignano and R. Cont, "Universal features of price formation in financial markets: perspectives from deep learning," Quantitative Finance, vol. 19, no. 9, pp. 1449–1459, 2019.
[10] X. Bao, C. Yue, and J. Rao, "A deep learning framework for financial time series using stacked autoencoders and LSTMs," International Journal of Information Management, vol. 50, pp. 132–142, 2020.
[11] A. Kirilenko, A. Kyle, M. Samadi, and T. Tuzun, "The flash crash: High-frequency trading in an electronic market," Journal of Finance, vol. 72, no. 3, pp. 967–998, 2017.
[12] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics, vol. 31, no. 3, pp. 307–327, 1986.
[13] D. Zhang, Y. Li, and C. Jin, "High-frequency stock trading strategy with deep reinforcement learning," IEEE Access, vol. 7, pp. 36455–36466, 2019.
[14] R. Guidotti et al., "A survey of methods for explaining black box models," ACM Computing Surveys, vol. 51, no. 5, pp. 1–42, 2018.
[15] Shaharina Shoha, Abir, S. I., Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Enhanced Parkinson's Disease Detection Using Advanced Vocal Features and Machine Learning . *Journal of Computer Science and Technology Studies*, *6*(5), 113-128. https://doi.org/10.32996/jcsts.2024.6.5.10
[16] Nigar Sultana, Shariar Islam Saimon, Intiser Islam, Abir, S. I., Md Sanjit Hossain, Sarder Abdulla Al Shiam, & Nazrul Islam Khan. (2025). Artificial Intelligence in Multi-Disease Medical Diagnostics: An Integrative Approach. *Journal of Computer Science and Technology Studies*, *7*(1), 157-175. https://doi.org/10.32996/jcsts.2025.7.1.12
[17] Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Syed Moshiur Rahman, Shariar Islam Saimon, Intiser Islam, Md Atikul Islam Mamun, & Nazrul Islam Khan. (2024). Deep Learning-Based Classification of Skin Lesions: Enhancing Melanoma Detection through Automated Preprocessing and Data

Augmentation. *Journal of Computer Science and Technology Studies*, *6*(5), 152-167. https://doi.org/10.32996/jcsts.2024.6.5.13

[18] Abir, S. I., Shaharina Shoha, Sarder Abdulla Al Shiam, Shariar Islam Saimon, Intiser Islam, Md Atikul Islam Mamun, Md Miraj Hossain, Syed Moshiur Rahman, & Nazrul Islam Khan. (2024). Precision Lesion Analysis and Classification in Dermatological Imaging through Advanced Convolutional Architectures. *Journal of Computer Science and Technology Studies*, *6*(5), 168-180. https://doi.org/10.32996/jcsts.2024.6.5.14

[19] Abir, S. I., Shaharina Shoha, Sarder Abdulla Al shiam, Nazrul Islam Khan, Abid Hasan Shimanto, Muhammad Zakaria, & S M Shamsul Arefeen. (2024). Deep Learning Application of LSTM(P) to predict the risk factors of etiology cardiovascular disease. *Journal of Computer Science and Technology Studies*, *6*(5), 181-200. https://doi.org/10.32996/jcsts.2024.6.5.15

[20] Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Nigar Sultana, Tui Rani Saha, Mohammad Hasan Sarwer, Shariar Islam Saimon, Intiser Islam, & Mahmud Hasan. (2025). Machine Learning and Deep Learning Techniques for EEG-Based Prediction of Psychiatric Disorders. *Journal of Computer Science and Technology Studies*, *7*(1), 46-63. https://doi.org/10.32996/jcsts.2025.7.1.4

[21] Mohammad Hasan Sarwer, Tui Rani Saha, Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Nigar Sultana, Shariar Islam Saimon, Intiser Islam, Mahmud Hasan, & Sarder Abdulla Al Shiam. (2025). EEG Functional Connectivity and Deep Learning for Automated Diagnosis of Alzheimer's disease and Schizophrenia. *Journal of Computer Science and Technology Studies*, *7*(1), 82-99. https://doi.org/10.32996/jcsts.2025.7.1.7

[22] Abir, S. I., Shahrina Shoha, Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Deep Neural Networks in Medical Imaging: Advances, Challenges, and Future Directions for Precision Healthcare . *Journal of Computer Science and Technology Studies*, *6*(5), 94-112. https://doi.org/10.32996/jcsts.2024.6.5.9

[23] Shariar Islam Saimon, Intiser Islam, Shake Ibna Abir, Nigar Sultana, Md Sanjit Hossain, & Sarder Abdulla Al Shiam. (2025). Advancing Neurological Disease Prediction through Machine Learning Techniques. *Journal of Computer Science and Technology Studies*, *7*(1), 139-156. https://doi.org/10.32996/jcsts.2025.7.1.11

[24] Abir, S. I., Shariar Islam Saimon, Tui Rani Saha, Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, Shoha, S. ., & Intiser Islam. (2025). Comparative Analysis of Currency Exchange and Stock Markets in BRICS Using Machine Learning to Forecast Optimal Trends for Data-Driven Decision Making. *Journal of Economics, Finance and Accounting Studies* , *7*(1), 26-48. https://doi.org/10.32996/jefas.2025.7.1.3

[25] Abir, S. I., Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, Shaharina Shoha, & Tui Rani Saha. (2025). Deep Learning for Financial Markets: A Case-Based Analysis of BRICS Nations in the Era of Intelligent Forecasting. *Journal of Economics, Finance and Accounting Studies* , *7*(1), 01-15. https://doi.org/10.32996/jefas.2025.7.1.1

[26] Abir, S. I., Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, & Tui Rani Saha. (2024). Accelerating BRICS Economic Growth: AI-Driven Data Analytics for Informed Policy and Decision Making. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 102-115. https://doi.org/10.32996/jefas.2024.6.6.8

[27] Nigar Sultana, Shaharina Shoha, Md Shah Ali Dolon, Sarder Abdulla Al Shiam, Rafi Muhammad Zakaria, Abid Hasan Shimanto, S M Shamsul Arefeen, & Abir, S. I. (2024). Machine Learning Solutions for Predicting Stock Trends in BRICS amid Global Economic Shifts and Decoding Market Dynamics. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 84-101. https://doi.org/10.32996/jefas.2024.6.6.7

[28] Abir, S. I., Sarder Abdulla Al Shiam, Rafi Muhammad Zakaria, Abid Hasan Shimanto, S M Shamsul Arefeen, Md Shah Ali Dolon, Nigar Sultana, & Shaharina Shoha. (2024). Use of AI-Powered Precision in Machine Learning Models for Real-Time Currency Exchange Rate Forecasting in BRICS Economies. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 66-83. https://doi.org/10.32996/jefas.2024.6.6.6

[29] S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "A Comprehensive Examination of MR Image-Based Brain Tumor Detection via Deep Learning Networks," *2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, pp. 1-8, doi: 10.1109/ICDS62089.2024.10756457, 2024.

[30] Akhter, A., Sarder Abdulla Al Shiam, Mohammad Ridwan, Abir, S. I., Shoha, S., Nayeem, M. B., … Robeena Bibi. (2024). Assessing the Impact of Private Investment in AI and Financial Globalization on Load Capacity Factor: Evidence from United States. *Journal of Environmental Science and Economics*, *3*(3), 99–127. https://doi.org/10.56556/jescae.v3i3.977

[31] Hossain, M. S., Mohammad Ridwan, Akhter, A., Nayeem, M. B., M Tazwar Hossain Choudhury, Asrafuzzaman, M., … Sumaira. (2024). Exploring the LCC Hypothesis in the Nordic Region: The Role of AI Innovation, Environmental Taxes, and Financial Accessibility via Panel ARDL. *Global Sustainability Research* , *3*(3), 54–80. https://doi.org/10.56556/gssr.v3i3.972

[32] Abir, S.I.; Shoha, S.; Hossain, M.M.; Sultana, N.; Saha, T.R.; Sarwer, M.H.; Saimon, S.I.; Islam, I.; Hasan, M. Machine Learning and Deep Learning Techniques for EEG-Based Prediction of Psychiatric Disorders. *J. Comput. Sci. Technol. Stud.* **2025**, *7*, 46–63.

[33] Mohammad Ridwan, Bala, S., Abdulla Al Shiam, S., Akhter, A., Mahdi Hasan, M., Asrafuzzaman, M., … Bibi, R. (2024). Leveraging AI for Promoting Sustainable Environments in G-7: The Impact of Financial Development and Digital Economy via MMQR Approach. *Global Sustainability Research* , *3*(3), 27–53. https://doi.org/10.56556/gssr.v3i3.971

[34] Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shoha, S., Hemel Hossain, Dolon, M. S. A., … Mohammad Ridwan. (2024). Assessing the Impact of AI Innovation, Financial Development, and the Digital Economy on Load Capacity Factor in the BRICS Region. *Journal of Environmental Science and Economics*, *3*(2), 102–126. https://doi.org/10.56556/jescae.v3i2.981

[35] Mohammad Ridwan, Abdulla Al Shiam, S., Hemel Hossain, Abir, S. I., Shoha, S., Dolon, M. S. A., … Rahman, H. (2024). Navigating a Greener Future: The Role of Geopolitical Risk, Financial Inclusion, and AI Innovation in the BRICS – An Empirical Analysis. *Journal of Environmental Science and Economics*, *3*(1), 78–103. https://doi.org/10.56556/jescae.v3i1.980

[36] Shoha, S., Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shewly Bala, Dolon, M. S. A., … Robeena Bibi. (2024). Towards Carbon Neutrality: The Impact of Private AI Investment and Financial Development in the United States – An Empirical Study Using the STIRPAT Model. *Journal of Environmental Science and Economics*, *3*(4), 59–79. https://doi.org/10.56556/jescae.v3i4.982

[37] Abir, S. I., Shoha, S., Abdulla Al Shiam, S., Dolon, M. S. A., Shewly Bala, Hemel Hossain, … Robeena Bibi. (2024). Enhancing Load Capacity Factor: The Influence of Financial Accessibility, AI Innovation, and Institutional Quality in the United States. *Journal of Environmental Science and Economics*, *3*(4), 12–36. https://doi.org/10.56556/jescae.v3i4.979

[38] S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "Health Risks and Disease Transmission in Undocumented Immigrants in the U.S Using Predictive ML," 2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS), Marrakech, Morocco, pp. 1-6, doi: 10.1109/ICDS62089.2024.10756308, 2024.

[39] Abir, S. I., Shoha, S., Abdulla Al Shiam, S., Dolon, M. S. A., Shewly Bala, Hemel Hossain, … Robeena Bibi. (2024). Enhancing Load Capacity Factor: The Influence of Financial Accessibility, AI Innovation, and Institutional Quality in the United States. *Journal of Environmental Science and Economics*, *3*(4), 12–36. https://doi.org/10.56556/jescae.v3i4.979

[40] Mohammad Ridwan, Bala, S., Shiam, S. A. A., Akhter, A., Asrafuzzaman, M., Shochona, S. A., … Shoha, S. (2024). Leveraging AI for a Greener Future: Exploring the Economic and Financial Impacts on Sustainable Environment in the United States. *Journal of Environmental Science and Economics*, *3*(3), 1–30. https://doi.org/10.56556/jescae.v3i3.970

[41] Akhter, A., Sarder Abdulla Al Shiam, Mohammad Ridwan, Abir, S. I., Shoha, S., Nayeem, M. B., … Robeena Bibi. (2024). Assessing the Impact of Private Investment in AI and Financial Globalization on Load Capacity Factor: Evidence from United States. *Journal of Environmental Science and Economics*, *3*(3), 99–127. https://doi.org/10.56556/jescae.v3i3.977

[42] Sohail,Muhammad Noman and Ren,Jiadong and Muhammad,Musa Uba and Rizwan,Tahir and Iqbal,Wasim and Abir,Shake Ibna. Bio Tech System, Group covariates assessment on real-life diabetes patients by fractional polynomials: a study based on logistic regression modeling, English, Journal article, USA, 1944-3285, 10, Edmond, *Journal of Biotech Research,* (116–125), 2019.

[43] Farhana Yeasmin Rita, S M Shamsul Arefeen, Rafi Muhammad Zakaria, & Abid Hasan Shimanto. (2025). An Integrative Artificial Intelligence Framework for the Diagnosis of Multiple Diseases in Clinical Settings. *Journal of Computer Science and Technology Studies*, *7*(2), 645-655. https://doi.org/10.32996/jcsts.2025.7.2.69

[44] Farhana Yeasmin Rita, S M Shamsul Arefeen, Rafi Muhammad Zakaria, & Abid Hasan Shimanto. (2025). Predictive Modeling of Patient Health Outcomes Using Electronic Health Records and Advanced Machine Learning Algorithms. *Journal of Computer Science and Technology Studies*, *7*(2), 632-644. https://doi.org/10.32996/jcsts.2025.7.2.68

[45] Farhana Yeasmin Rita, S M Shamsul Arefeen, Rafi Muhammad Zakaria, & Abid Hasan Shimanto. (2025). Advancing the Prediction of Neurological Disorders Through Innovative Machine Learning Methodologies and Clinical Data Analysis. *Journal of Computer Science and Technology Studies*, *7*(2), 668-680. https://doi.org/10.32996/jcsts.2025.7.2.71

[46] Farhana Yeasmin Rita, S M Shamsul Arefeen, Rafi Muhammad Zakaria, & Abid Hasan Shimanto. (2025). Early Detection of Alzheimer's Disease Through Deep Learning Techniques Applied to Neuroimaging Data. *Journal of Computer Science and Technology Studies*, *7*(2), 656-667. https://doi.org/10.32996/jcsts.2025.7.2.70

[47] Farhana Yeasmin Rita, S M Shamsul Arefeen, Rafi Muhammad Zakaria, & Abid Hasan Shimanto. (2025). Harnessing Artificial Intelligence in Medical Imaging for Enhanced Cancer Detection and Diagnosis. *Journal of Computer Science and Technology Studies*, *7*(2), 618-631. https://doi.org/10.32996/jcsts.2025.7.2.67

[48] Md Sohanur Rahman Sourav, Arafat Hossain, Md Redwanul Islam, Mohtasim Wasif, & Sujana Samia. (2025). AI-Driven forecasting in BRICS infrastructure investment: impacts on resource allocation and project delivery. *Journal of Economics, Finance and Accounting Studies* , *7*(2), 117-132. https://doi.org/10.32996/jefas.2025.7.2.11

[49] Md Redwanul Islam, Mohtasim Wasif, Sujana Samia, Md Sohanur Rahman Sourav, & Arafat Hossain. (2025). The Role of Machine Learning in Forecasting U.S. GDP Growth after the COVID-19 Pandemic. *Journal of Economics, Finance and Accounting Studies* , *7*(2), 163-175. https://doi.org/10.32996/jefas.2025.7.2.14

[50] Mohtasim Wasif, Sujana Samia, Md Sohanur Rahman Sourav, Arafat Hossain, & Md Redwanul Islam. (2025). Data-Driven insights on the relationship between BRICS financial policies and global investment trends. *Journal of Economics, Finance and Accounting Studies* , *7*(2), 133-147. https://doi.org/10.32996/jefas.2025.7.2.12

[51] Iftekhar Rasul, S M Iftekhar Shaboj, Mainuddin Adel Rafi, Md Kauser Miah, Md Redwanul Islam, & Abir Ahmed. (2024). Detecting Financial Fraud in Real-Time Transactions Using Graph Neural Networks and Anomaly Detection. *Journal of Economics, Finance and Accounting Studies* , *6*(1), 131-142. https://doi.org/10.32996/jefas.2024.6.1.13

[52] Mainuddin Adel Rafi, S M Iftekhar Shaboj, Md Kauser Miah, Iftekhar Rasul, Md Redwanul Islam, & Abir Ahmed. (2024). Explainable AI for Credit Risk Assessment: A Data-Driven Approach to Transparent Lending Decisions. *Journal of Economics, Finance and Accounting Studies* , *6*(1), 108-118. https://doi.org/10.32996/jefas.2024.6.1.11

[53] Mainuddin Adel Rafi, S M Iftekhar Shaboj, Iftekhar Rasul, Md Kauser Miah, Iftekhar Rasul, Md Redwanul Islam, & Abir Ahmed. (2024). Cryptocurrency Volatility Forecasting Using Transformer-Based Deep Learning Models and On-Chain Metrics. *Journal of Economics, Finance and Accounting Studies* , *6*(1), 119-130. https://doi.org/10.32996/jefas.2024.6.1.12

[54] Md. Tanvir Rahman Mazumder, Md. Shahadat Hossain Shourov, Iftekhar Rasul, Sonia Akter, & Md Kauser Miah. (2025). Fraud Detection in Financial Transactions: A Unified Deep Learning Approach. *Journal of Economics, Finance and Accounting Studies* , *7*(2), 184-194. https://doi.org/10.32996/jefas.2025.7.2.16

[55] Md. Tanvir Rahman Mazumder, Md. Shahadat Hossain Shourov, Iftekhar Rasul, Sonia Akter, & Md Kauser Miah. (2025). The Impact of Macroeconomic Factors on the U.S. Market: A Data Science Perspective. *Journal of Economics, Finance and Accounting Studies* , *7*(2), 208-219. https://doi.org/10.32996/jefas.2025.7.2.18

[56] Md. Tanvir Rahman Mazumder, Md. Shahadat Hossain Shourov, Iftekhar Rasul, Sonia Akter, & Md Kauser Miah. (2025). Anomaly Detection in Financial Transactions Using Convolutional Neural Networks. *Journal of Economics, Finance and Accounting Studies* , *7*(2), 195-207. https://doi.org/10.32996/jefas.2025.7.2.17