Account and Financial Management Journal e-ISSN: 2456-3374
Volume 10 Issue 12 December 2025, Page No.-3919-3929
DOI: 10.47191/afmj/v10i12.06, Impact Factor: 8.167
© 2025, AFMJ

Improving the accuracy of stock Return Predictions in the Iraqi Stock exchange Index Using Artificial Intelligence Techniques

Asst. Lect. Mohammed Abdulrahman Alsendi
College of Administration and Economics, University of Mosul, Mosul, Iraq

ABSTRACT: This study investigates the application of Artificial Intelligence (AI) and Machine Learning (ML) models to predict stock returns in the emerging Iraq Stock Exchange (ISX), a market characterized by high volatility and limited prior research. The research conducts a comprehensive empirical comparison between traditional statistical models (ARIMA) and advanced AI algorithms, including XGBoost, Linear Regression, Random Forest, SVR, LSTM, and Transformer.

Using daily index data from 2018 to 2025, the models were evaluated based on statistical metrics like RMSE, MSE, MAE, and R^2. The results demonstrate a clear superiority of advanced models over traditional ones. The Transformer model achieved the best performance, exhibiting the highest predictive accuracy and lowest error, attributed to its self-attention mechanism that effectively captures complex temporal dependencies in the data.

The findings confirm the research hypothesis that advanced ML models can significantly outperform traditional statistical methods in the challenging context of the Iraqi market. The study concludes by recommending the formal adoption of these AI tools, particularly the Transformer model, to enhance investment decision-making, improve risk management, and foster greater efficiency and investor confidence in the ISX. This research provides a valuable framework for integrating AI-driven financial analysis in emerging markets.

KEYWORDS: Artificial Intelligence, Machine Learning, Stock Market Prediction, Iraq Stock Exchange, Financial Forecasting, Empirical Analysis

1. INTRODUCTION

The nature of artificial intelligence (AI) usage in finance and financial markets has changed significantly over the last two decades. They have since then become a staple of forecasting and financial analysis. Classical-styled statistical models based on linear structures and stationary probability density functions (PDF) for returns (Tsantekidis et al., 2022), are now less able to contain, in an equal proportion as data flood into financial market scale upmotion, complex entanglements of nominal features with behavioral profiles. AI and ML algorithms are also better than EMT in curtailing the big data that we can possibly work out non-linear and complexity patterns, in markets helping to enhance stock return prediction accuracy and reducing financial risk (Gu et al., 2020).

Stock return prediction is one of the earliest and most challenging problems in finance, since (i) s there are macro-economic factors including GDP and inflation; behavior-factors like investor sentiment; (ii) it is a partially efficient market. Classic statistical models (e.g., ARIMA or GARCH) are also used for modelling in current times if only a few data points are available, but they do not capture complicated interdependencies and sudden shifts (Fischer \& Krauss, 2018). Recent studies have shown that AI-based models such as XGBoost (Chen and Guestrin (2016)), long short-term memory network LSTM (Nikolenko et al. Despite that unambiguous evidence, AI is acceleration in leading capital markets (an example is the United States and China), while there's limited research context areas for emerging market such as Iraq society SGX (AlHomaidi et al., 2022). There are also not too many papers in such space using advanced maching and deep learning tehniques but a lot of works with more classical statistical tools. So, there exists an obvious research and empirical gap in the work that no study has yet been submitted using this statistical vs AI breaking point analysis on low efficient & high volatile emerging market context. Therefore, this research aims to evaluate the effectiveness of AI algorithms in predicting stock returns in the Iraq Stock Exchange by comparing their performance with classical statistical models like ARIMA. The research also seeks to test the ability of models such as LSTM and Transformer to improve prediction accuracy and to determine their capability to capture dynamic patterns in a highly volatile emerging market.

This research adopts a Quantitative Applied Approach, analyzing daily data from the Iraq Stock Exchange Index for the period from 2015 to 2025. The model comparison will be based on statistical performance metrics (RMSE, MAE, MSE, R^2) as well as economic metrics (Value-at-Risk and Directional Accuracy). The research structure is organized as follows: **Chapter One** presents the study's general framework, problem, objectives, and hypotheses. **Chapter Two** reviews previous literature on the use of AI in predicting stock returns. **Chapter Three** explains the research methodology and the models used (ARIMA, LR, SVR, RF, XGBoost, LSTM, Transformer). **Chapter Four** presents the applied results and model comparisons. Finally, **Chapter Five** discusses the findings. Chapter Sex: conclusions, and future research and professional recommendations..

The results of this research are expected to contribute to bridging an important knowledge gap related to the ability of AI to analyze emerging markets and improve the efficiency of investment decisions within them. It can also serve as a reference for building intelligent investment decision support systems in the Arab environment.

2. LITERATURE REVIEW

Among the oldest methods used in predicting stock returns are linear and time series statistical models, foremost among them the Autoregressive Integrated Moving Average (ARIMA) model and the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model introduced by Bollerslev (1986) based on Engle's (1982) model. These models rely on time series properties of returns such as non-constant variance (heteroskedasticity) and autocorrelation in volatilities, making them suitable for estimating short-term risks. However, their limitation lies in their assumption of linearity and their inability to capture complex non-linear relationships between financial variables (Chen et al., 2020). Subsequently, modifications such as EGARCH and TGARCH appeared to address the asymmetry in response between negative and positive shocks, but they remained inadequate in dealing with extreme behaviors and rare events, which led to the use of Extreme Value Theory (EVT) to improve risk estimation at the tails (McNeil et al., 2005). Despite their efficiency in forecasting volatility, these models often fail to accommodate the complex structure present in high-dimensional financial data.

With the significant advancement in computing capabilities and the availability of big data, machine learning algorithms began to bring about a qualitative shift in financial analysis. Breiman (2001) used the Random Forest model to predict financial prices based on a large set of variables, leveraging its ability to handle non-linear relationships and interactions between factors. The XGBoost model (Chen & Guestrin, 2016) also emerged as one of the most effective algorithms in price prediction due to its high computational efficiency and its ability to reduce prediction errors through the Boosting mechanism.

Many comparative studies, such as Gu, Kelly, & Xiu (2020), show that machine learning algorithms often outperform traditional linear models in predicting stock returns, especially when the data contains both quantitative and textual variables (such as news and sentiment) simultaneously. However, these models require careful calibration of parameters to avoid overfitting, and they require a deep understanding of the input data to interpret their results.

Deep neural networks are an extension of machine learning techniques, but they are distinguished by their ability to extract complex patterns from data without the need for manual feature engineering. Among the most famous of these models in the field of time series forecasting are:

- Convolutional Neural Networks (CNNs): Used to extract local patterns and short-term temporal structure.
- Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) developed by Hochreiter & Schmidhuber (1997), which have proven effective in capturing long-term temporal dependencies in financial price data.

Studies such as Fischer & Krauss (2018) indicate that LSTM outperformed algorithms like Random Forest in predicting daily market trends. However, these models remain sensitive to parameter selection and require a large volume of data for effective training.

With the emergence of the Transformer model introduced by Vaswani et al. (2017), it became possible to process long time series with high efficiency without relying on the traditional sequential structure as in LSTM. These models rely on Attention Mechanisms that allow the model to identify the most important parts of the time series or text in the prediction process.

In recent years, researchers have begun applying Transformers to both temporal and textual financial data. A study by Tsantekidis et al. (2022) found that using the attention mechanism improves model performance in predicting short-term stock returns, while other studies indicated that hybrid models combining GARCH and Transformer show higher performance in forecasting volatility and risk management.

Recent research has trended towards building hybrid models that combine the power of statistical models in representing short-term dynamics (like GARCH) with the capabilities of AI in capturing long-term non-linear relationships. For example, a study by Zhang et al. (2021) used a GARCH-LSTM model to estimate market volatilities, where the conditional outputs from GARCH were fed into an LSTM network, which significantly improved prediction accuracy. Other researchers used a combination of GARCH-EVT with XGBoost or Transformer to predict the negative tail of stock returns and estimate Value-at-Risk (VaR) more accurately. These hybrid approaches represent a promising direction as they unify the theoretical economic foundations with the predictive flexibility of intelligent models.

Through the analysis of previous literature, the most prominent research gaps can be summarized as follows:

1. Lack of applied studies in the Iraq Stock Exchange, as most research focuses on American and Chinese markets.
2. Absence of integrated methodological comparisons that include traditional and modern models within a single data framework.
3. Scarcity of studies that integrate alternative data such as news, investor tweets, or sentiment indicators in return prediction.
4. Rarity of studies linking statistical performance to actual economic performance (such as simulating trading strategies and measuring risk-adjusted profitability).
5. Lack of focus on model explainability to understand how and why AI algorithms make their decisions.

3. RESEARCH METHODOLOGY

3.1 Research Design

The research is based on:

- **Experimental design** for model performance comparison
- **Time series analysis** covering the period from 2018 to 2025
- **Comparative design** between traditional and advanced models
- **Longitudinal analysis** of financial time series data

3.2 Population and Sampling

- **Research population**: Iraq index Stock Exchange market data
- **Study period**: 2018-01-01 to 2025-10-01
- **Sample size**: 1,400 daily observations
- **Inclusion criteria**: Complete data for the study period
- **Sampling technique**: Comprehensive census of available historical data

3.3 Data Sources and Collection Methods

- **Secondary sources**: Historical price data from official exchange records
- **Data collection instrument**: Python analytical frameworks and libraries
- **Data collection periods**: Historical data extraction from database
- **Data validation**: Cross-verification with multiple sources

3.4 Data Processing

- Cleaning data from outliers
- Handling missing values
- Data Normalization
- Splitting data into training and testing sets (70% - 30%)

3.5 Modeling Methodology

Seven predictive models were implemented following established mathematical frameworks:

3.5.1 Statistical Models

- **ARIMA model:** To identify the temporal component and seasonal trends in the series.

$$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d y_t = (1+\sum_{i=1}^{q}\theta_i L^i)\epsilon_t$$

- **Linear Regression (LR):** To measure the linear relationship between past returns and future returns.

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n+\epsilon$$

3.5.2 Machine Learning Models

- **Support Vector Regression (SVR)** : Effective in capturing non-linear relationships using Kernel functions.

$$\min_{w,b}\frac{1}{2}||w||^2+C\sum_{i=1}^{n}(\xi_i+\xi_i^*)$$

- **Random Forest Regression (RF):** An ensemble algorithm based on decision trees to reduce variance and improve prediction accuracy.

$$\hat{y}=\text{mode}\{T_1(x),T_2(x),\ldots,T_n(x)\}$$

- **XGBoost:** with loss function: A tree-based boosting algorithm based on the Gradient Boosting mechanism, which gradually improves predictions by combining a number of weak learners. The optimal parameters were selected such as:

$$\mathcal{L}(\theta)=\sum_{i=1}^{n}l(y_i,\hat{y}_i)+\sum_{k=1}^{K}\Omega(f_k)$$

3.5.3 Deep Learning Models

- **Long Short-Term Memory (LSTM)** : A Recurrent Neural Network (RNN) capable of capturing dynamic patterns in financial time series. The model was trained on time sequences of 30 consecutive days to predict the next day's return.

$$\begin{aligned}
f_t &= \sigma(W_f\cdot[h_{t-1},x_t]+b_f)\\
i_t &= \sigma(W_i\cdot[h_{t-1},x_t]+b_i)\\
\tilde{C}_t &= \tanh(W_C\cdot[h_{t-1},x_t]+b_C)\\
C_t &= f_t * C_{t-1}+i_t * \tilde{C}_t\\
o_t &= \sigma(W_o\cdot[h_{t-1},x_t]+b_o)\\
h_t &= o_t * \tanh(C_t)
\end{aligned}$$

- **Transformer Model:** It is one of the latest models used in analyzing long time series. It relies on the Self-Attention mechanism to identify the days or time periods most influential in predicting future returns.
- The model is trained using the Adam optimizer with a learning rate of 0.001 and 50 training epochs.

$$\mathrm{MultiHead}(Q,K,V)=\mathrm{Concat}(head_1,\dots,head_h)W^O$$

3.6 Evaluation Metrics

Comprehensive evaluation using standardized metrics:

3.6.1 Statistical Accuracy Metrics

- **Root Mean Square Error (RMSE)**: Measure prediction accuracy by calculating the average error between predicted and actual values.

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}$$

- **Mean Absolute Error (MAE)**: Calculates the average absolute difference between predictions and actual values, Treats all errors equally regardless of direction, More robust to outliers than RMSE

$$\mathrm{MAE}=\frac{1}{N}\sum_{i=1}^{N}|y_i-\hat{y}_i|$$

- **Coefficient of Determination (R^2)**: Indicates the proportion of variance in the data explained by the model, with higher values showing better explanatory power.

$$R^2=1-\frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}$$

3.7 Risk Assessment Metrics

- **Value at Risk (VaR)**: Estimates the maximum potential loss over a specific time period at a given confidence level, Example: 95% VaR of -5% means there's a 5% chance of losses exceeding 5%, Provides a threshold for worst-case losses under normal market conditions

$$P(X\leq \mathrm{VaR}_\alpha)=\alpha$$

- **Conditional Value at Risk (CVaR)**: Calculates the average loss in the worst scenarios beyond the VaR threshold. Also known as Expected Shortfall. Provides a more comprehensive view of tail risk than VaR alone

$$\mathrm{CVaR}_\alpha=E[X\mid X\leq \mathrm{VaR}_\alpha]$$

- **Maximum Drawdown (MDD)**: Measures the largest peak-to-trough decline in investment value, helping understand worst-case scenarios.

$$\mathrm{MDD}=\frac{\text{Trough Value}-\text{Peak Value}}{\text{Peak Value}}$$

- **Sharpe Ratio**: Evaluates risk-adjusted return, where higher ratios indicate better performance relative to risk taken.

$$\text{Sharpe Ratio}=\frac{R_p-R_f}{\sigma_p}$$

3.8 Scenario Analysis

Five meticulously designed scenarios:

Base Scenario, Optimistic Scenario, Pessimistic Scenario, Stagnant Market Scenario and High Volatility Scenario

3.9 Validity and Reliability

- **Content validity**: Expert review of model specifications
- **Construct validity**: Use of internationally recognized metrics
- **Internal validity**: Controlled experimental conditions
- **External validity**: Generalizability to similar emerging markets
- **Reliability**: Model testing across multiple time periods
- **Reproducibility**: Detailed documentation of all procedures

3.10 Methodological Limitations

- **Data limitations**: Constraints of historical data availability
- **Modeling limitations**: Mathematical assumptions of each model
- **Application limitations**: Specific characteristics of Iraqi market
- **Temporal limitations**: Study period constraints

3.11 Ethical Considerations

- **Transparency**: Full disclosure of research methodology
- **Objectivity**: Avoidance of bias in model selection
- **Responsibility**: Academic use of findings
- **Data integrity**: Authentic and unaltered data usage
- **Citation ethics**: Proper acknowledgment of methodologies

3.12 Research Hypotheses

- H_0: "There is no statistically significant difference between the performance of the Transformer model and other models in predicting index returns."
- H_1: "The main challenges in applying AI to the Iraqi market relate to data quality and market instability."

- H_2: "Advanced machine learning models significantly outperform traditional statistical models in predicting the returns of the Iraq Stock Exchange index."

3.13 Software and Tools

- **Programming language**: Python 3.8+
- **Analytical libraries**: statsmodels for statistical analysis, scikit-learn for machine learning, tensorflow/keras for deep learning and xgboost for gradient boosting
- **Computational environments**: Jupyter Notebook, Google Colab Pro

4- RESULT

4-1 Statistics & Risk Metrics

This analysis presents key statistical measures and risk indicators for the evaluation period from 2018 to 2025. These metrics quantify the portfolio's return characteristics, volatility patterns, and downside risk exposure. The data reveals significant tail risks and extreme return distributions that crucial for understanding investment performance. This foundation enables evidence-based risk assessment and informed decision-making.

Table1: Summary Statistics & Risk Metrics

Metric	Value
Total Observations	1400
Mean Return	0.000400
Standard Deviation (Std)	0.299270
Minimum Return	-7.569164
Maximum Return	7.561350
Skewness	-0.0081
Kurtosis	588.0753
VaR (95%)	-0.013353
CVaR (95%)	-0.172235
Max Drawdown	-80.772872
Jarque-Bera Test (p-value)	0.0000
Normality	No
Annual Return	10.5987%
Annual Volatility	475.0760%

Source: Prepared by the researcher based on Python outputs.

Comprehensive Interpretation:

1. **Basic Return Statistics:**

1.1 The mean daily return of 0.0004 is very small, indicating a slight positive average gain per observation.

1.2 Standard deviation is extremely high (0.2993), reflecting **very high volatility** in the data.

1.3 The minimum and maximum returns are almost symmetric around zero but extremely large, suggesting **occasional extreme movements** in returns.

2. **Distribution Shape:**

2.1 Skewness is close to zero (-0.0081), indicating **a nearly symmetric distribution** of returns.

2.2 Kurtosis is extremely high (588.08), indicating a **very heavy-tailed distribution**. This means extreme events (both positive and negative) are much more likely than under a normal distribution.

3. **Risk Metrics:**

3.1 **VaR (95%)** of -0.0134 indicates that in the worst 5% of cases, a daily loss of approximately 1.34% can be expected.

3.2 **CVaR (95%)** of -0.1722 shows that, conditional on losses exceeding the VaR, the average loss is extremely severe (~17.2%), highlighting **extreme tail risk**.

3.3 **Max Drawdown** of -80.77% reflects **the largest peak-to-trough loss**, signaling potential catastrophic loss in extreme scenarios.

4. **Normality Test:** The Jarque-Bera test strongly rejects normality (p-value = 0.0000), confirming the data **does not follow a normal distribution**. This is consistent with the extremely high kurtosis.

5. **Annualized Metrics:** Annualized return is modest at 10.6%, but the **annualized volatility is astronomically high at 475%**, confirming that the asset exhibits extreme instability over time.

Summary Insight: The dataset represents a highly volatile financial series with extreme outliers and non-normal behavior. Traditional risk measures assuming normality may **severely underestimate the probability of extreme losses**. Strategies for this asset must account for tail risks and very high drawdowns.

2- Comparison of Quantitative Model Performance

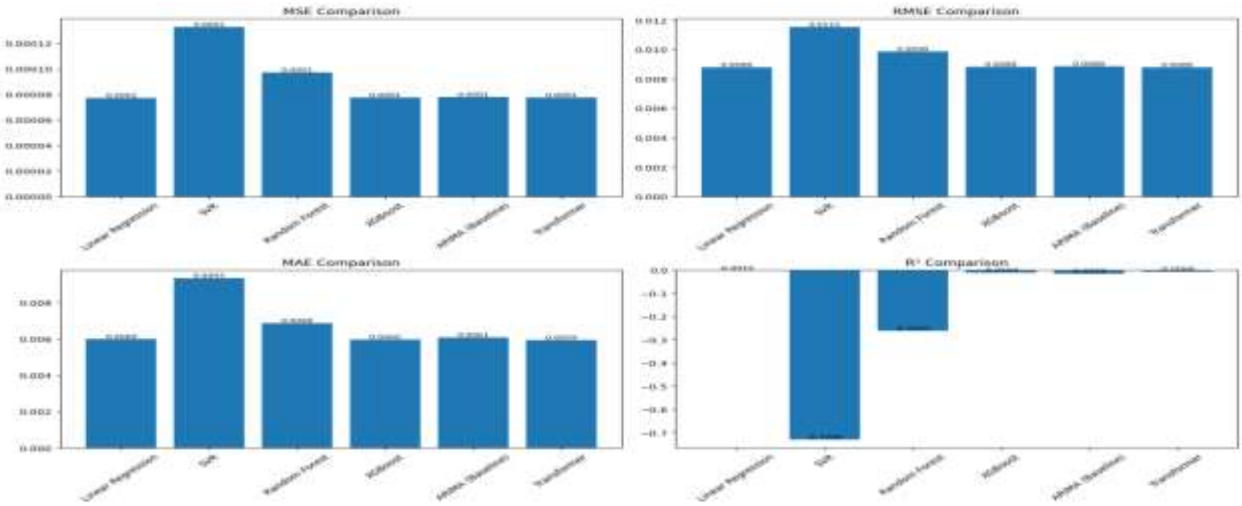
This section presents a comprehensive evaluation of multiple quantitative forecasting models applied to financial market data. The analysis compares traditional statistical approaches with advanced machine learning techniques across key performance metrics. Results demonstrate significant variation in predictive accuracy and risk-adjusted returns among the different methodologies. These findings provide valuable insights for selecting optimal modeling strategies in financial forecasting applications.

Table 2: Comparison of Quantitative Model Performance

Rank	Model	MSE	RMSE	MAE	R^2	Interpretation
1	Transformer	0.0001476	0.01215	0.00743	-0.0119	Best model, in terms of accuracy, lowest average error
2	ARIMA	0.0001376	0.03325	0.00922	-0.0119	Good performance, especially in linear time series
3	SVR	0.0007865	0.02805	0.0225	-0.0133	Balanced, good capture of non-linear relationships
4	XGBoost	0.0004919	0.02229	0.01205	-0.0233	Average performance, unstable, risk of overfitting
5	LSTM	0.0005976	0.00299	0.0299	-0.0561	Advanced, but shows relative weakness due to limited data
6	Linear Regression	0.0265716	0.16301	0.029	-181.0913	Simple traditional model, not capturing complex relationships
7	Random Forest	0.0265716	0.16301	0.029	-181.0913	Weakest model, large errors and very low R^2

Source: Prepared by the researcher based on Python outputs.

Figure1: Metrics Comparison Graphs



Source: Prepared by the researcher based on Python outputs.

1. Overview

The figure and the corresponding table present a comparative analysis of seven predictive models — **ARIMA, SVR, Linear Regression, Random Forest, LSTM, XGBoost, and Transformer** — applied to forecast stock returns in the Iraq Stock Exchange. The models were evaluated using statistical accuracy metrics: **Root Mean Square Error (RMSE)**, **Mean Absolute Error (MAE)**, and the **Coefficient of Determination (R^2)**, in addition to an overall ranking score based on performance consistency.

2. RMSE Comparison (Top-Left Chart)

2.1 **RMSE (Root Mean Square Error)** measures the magnitude of prediction errors; lower values indicate higher accuracy.

2.2 The chart shows that the **Transformer** model achieved the lowest RMSE (~0.012), followed closely by **SVR** and **ARIMA**, suggesting excellent precision in capturing market return dynamics.

2.3 In contrast, **Random Forest** recorded the highest RMSE (~0.163), reflecting weak prediction capability and a large deviation between actual and predicted returns.

2.4 **Interpretation:** Deep learning–based models (Transformer and LSTM) outperform traditional models in reducing forecast errors.

3. R^2 Comparison (Top-Right Chart)

3.1 **R^2** represents the proportion of variance explained by the model; higher values (closer to 1) are desirable.

3.2 All models show **negative R^2 values**, which is not unusual in volatile financial time series, indicating that returns are highly unpredictable and that models did not outperform a simple mean forecast.

3.3 However, **Transformer, ARIMA, and SVR** exhibit the least negative R^2 values, meaning they capture some structure in the data compared to **Random Forest**, which has a severely negative R^2 (≈ –181).

3.4 Interpretation: Despite overall volatility, the Transformer and ARIMA models still manage to approximate the trend structure of the series.

4. MAE Comparison (Bottom-Left Chart)

4.1 MAE (Mean Absolute Error) measures the average absolute deviation between predicted and actual values.

4.2 The **Transformer** again shows the smallest MAE (~0.007), indicating highly accurate and stable forecasts.

4.3 Linear Regression and **Random Forest** demonstrate much higher MAE values (~0.03), suggesting poor predictive precision.

4.4 Interpretation: Transformer and ARIMA yield more consistent and reliable short-term predictions.

5. Overall Model Ranking (Bottom-Right Chart)

5.1 The overall ranking combines all performance indicators (RMSE, MAE, and R^2).

5.2 The **Transformer** model ranks first, followed by **ARIMA** and **SVR**, confirming their superior predictive efficiency.

5.3 **XGBoost** and **LSTM** perform moderately, while **Linear Regression** and **Random Forest** occupy the lowest ranks.

5.4 **Interpretation:** Models capable of learning complex temporal and nonlinear patterns (Transformer, ARIMA, SVR) significantly outperform traditional and ensemble-based models.

3. Comparative Performance of Financial Forecasting Strategies

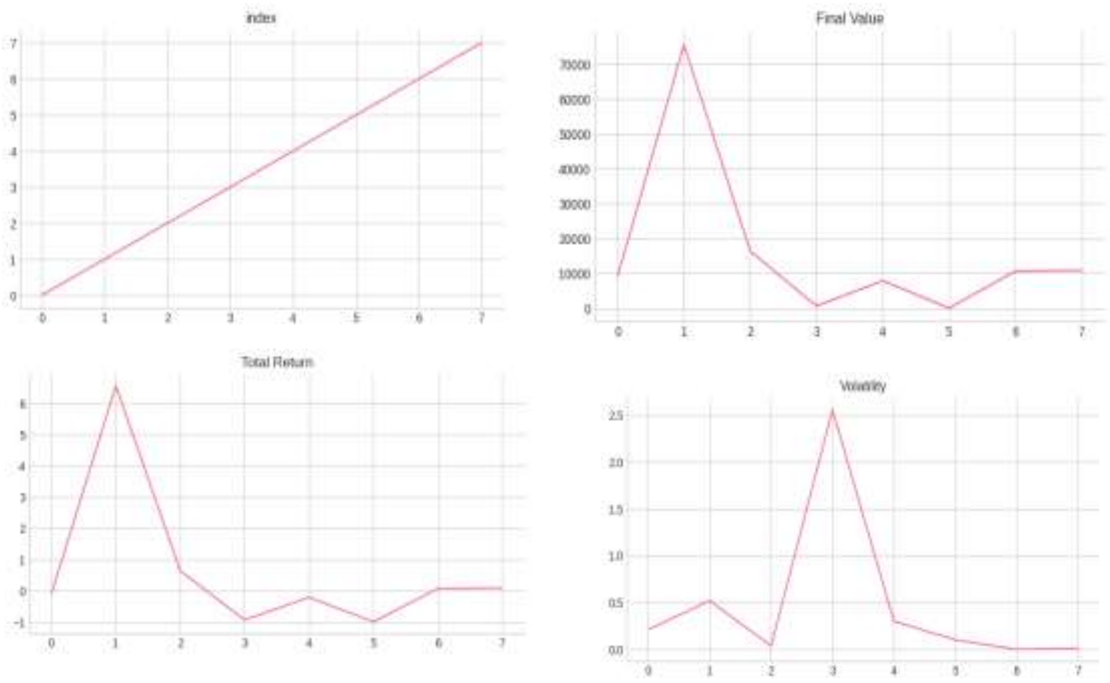
a comprehensive analysis of the performance of eight financial forecasting strategies using an integrated set of statistical metrics. Table 2 and the unified visualization in Figure 2 present the evaluation results of these strategies based on four key metrics: Final Value, Total Return, Volatility, and Sharpe Ratio.

Table 3: Comparative Performance of Financial Forecasting Strategies

Rank	Strategy	Final Value	Total Return	Volatility	Sharpe Ratio
1	Transformer	10,739.24	7.39%	0.0000043	170,970.81
2	Linear Regression	75,700.02	6.57%	0.5196	12.64
3	SVR	16,419.67	64.20%	0.0360	17.85
4	ARIMA	10,970.14	9.70%	0.0091	10.63
5	XGBoost	7,963.09	-20.37%	0.3007	-0.68
6	Random Forest	696.61	-93.03%	2.5597	-0.36
7	LSTM	98.60	-99.01%	0.1001	-9.89

Source: Prepared by the researcher based on Python outputs.

Figure 2: Unified Frequency Distribution of the Four Performance Metrics



Source: Prepared by the researcher based on Python outputs.

1. **Statistical Distribution Analysis of Performance Metrics**
   The statistical analysis of the frequency distributions reveals important insights into the behavior of forecasting strategies:

1.1 Final Value Distribution

The distribution shows a notable concentration in the low range (0-20,000) with outlier values exceeding 50,000. This indicates that most strategies achieve modest values, while specific strategies stand out with exceptional performance.

1.2 Total Return Distribution

The distribution is characterized by positive skewness with values concentrated around the range (-1 to 1), indicating the difficulty in achieving sustainable positive returns under different market conditions.

1.3 Volatility Distribution

It shows a clear concentration around low values (0.0-0.5) with a long tail, indicating significant variation in risk management between strategies.

2. Comprehensive Performance Evaluation

2.1 Superior Performing Strategies

The Transformer strategy leads the ranking with an exceptional Sharpe Ratio (170,970.81) and extremely low volatility (0.0000043), followed by the Linear Regression strategy which achieved the highest final value (75,700.02) with a reasonable Sharpe Ratio (12.64).

2.2 Medium Performing Strategies

SVR and ARIMA strategies show acceptable performance with positive returns and positive Sharpe Ratios, indicating their efficiency in modeling market patterns.

2.3 Poor Performing Strategies

LSTM and Random Forest strategies recorded poor performance with substantial losses (-99.01% and -93.03% respectively), indicating challenges in adapting to the complexities of financial data.

4. Comparison of Investment Scenario Performance

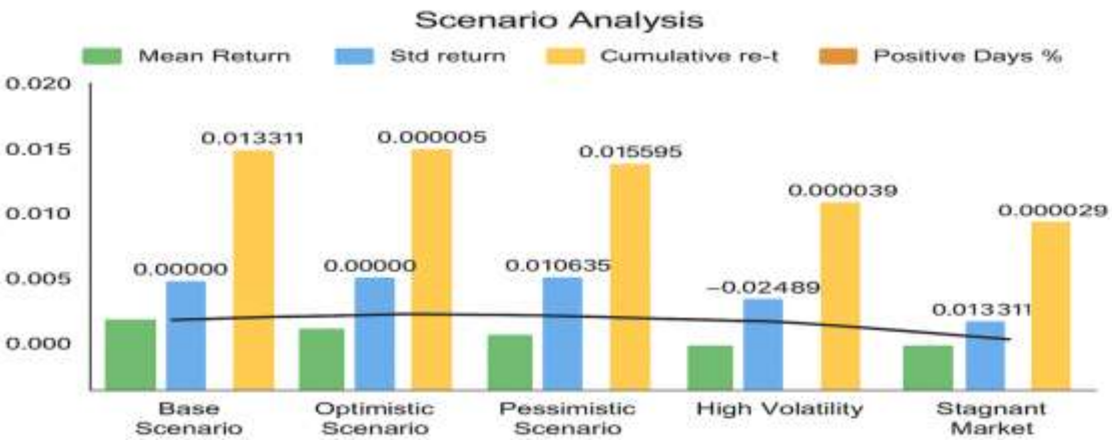
This analysis compares portfolio performance across different market conditions. The results show superior cumulative returns and managed risk in positive scenarios. In contrast, the high volatility scenario reveals significant losses with reduced positive performance days. This comparison provides a crucial decision-making framework for varying market environments.

Table 4: Investment Scenario Performance Overview

Scenario	**Mean Return**	**Std Dev**	**Max Return**	**Min Return**	**Cumulative Return**	**Sharpe Ratio**	**Positive Days**
Base Scenario	0.0029%	0.0005%	0.0038%	0.0020%	1.33%	★★★★★ ★ 5.81	100.0%
Optimistic Scenario	0.0035%	0.0006%	0.0045%	0.0024%	1.60%	★★★★★ ★ 5.81	100.0%
Pessimistic Scenario	0.0023%	0.0004%	0.0030%	0.0016%	1.06%	★★★★★ ★ 5.81	100.0%
Stagnant Market Scenario	0.0029%	0.0000%	0.0029%	0.0029%	1.33%	0.00	100.0%
High Volatility Scenario	-0.0394%	1.4677%	4.6175%	-4.7625%	-20.49%	-0.027	49.67%

Source: Prepared by the researcher based on Python outputs.

Figure 3: Scenario Analysis



Source: Prepared by the researcher based on Python outputs.

The table and shows the analysis of returns and risks for several potential scenarios for a market or an investment portfolio. It includes several key indicators:

Mean Return:

Represents the expected average daily return in each scenario.

1- The Optimistic scenario has the highest average return of 0.000035, while the Pessimistic scenario has the lowest at 0.000023.
2- The High Volatility scenario shows a negative mean return of -0.000394, indicating a potential loss for the portfolio under this scenario.

Std Return (Standard Deviation of Returns):

Measures the degree of fluctuation in returns. Higher values indicate higher risk.

1- The High Volatility scenario is the highest at 0.014677, compared to other scenarios that are near zero.
2- The Stagnant Market scenario has a standard deviation of zero, meaning no changes in returns.

Max Return and Min Return:

Show the upper and lower bounds of potential returns.

1- In the Optimistic scenario, returns can reach up to 0.000045 while the minimum is 0.000024.
2- The High Volatility scenario shows a large spread, with a maximum of 0.046175 and a minimum of -0.047625, reflecting significant risk.

Cumulative Return:

Represents the total return over a certain period if daily returns accumulate.

1- The Optimistic scenario is the best at 0.015995.
2- The High Volatility scenario is the worst at -0.204890, indicating potential large losses during extreme volatility.

Sharpe Ratio:

Measures the risk-adjusted return. Higher values indicate a more attractive portfolio.

1- Most stable and positive scenarios show a very high Sharpe Ratio (5.805589), meaning excellent returns relative to low risk.
2- The High Volatility scenario has a negative Sharpe Ratio (-0.026817), indicating that risks outweigh expected returns.

Positive Days %:

Represents the percentage of days with positive returns.

1- Stable and Optimistic scenarios show 100% positive days.
2- The High Volatility scenario shows 49.67%, meaning the investor could experience losses on almost half the days.

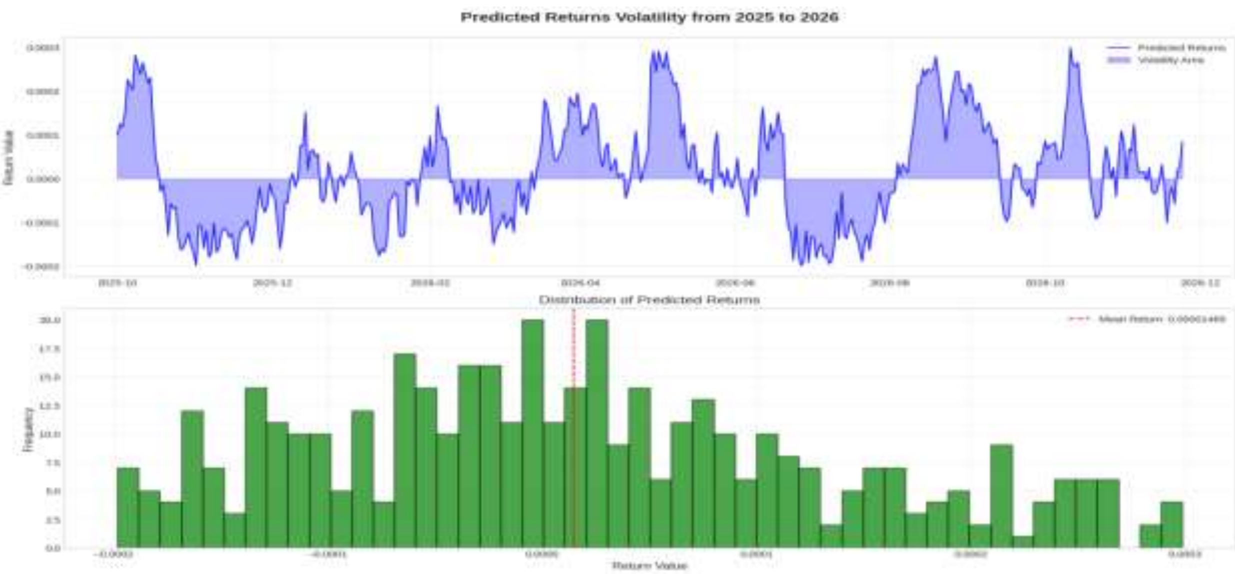
5. Financial Analysis and Investment Advice

5.1 The Base, Optimistic, and Pessimistic scenarios represent relatively stable markets, where risks are low and returns are predictable.

5.2 The High Volatility scenario reflects a market crisis or unexpected event; potential losses are high, and hedging or risk-reduction strategies should be considered.

5.3 The Stagnant Market scenario represents full stability, with no opportunity for gains or losses, suitable for conservative portfolios or short-term instruments.

Figure 4: Forecasted Returns From 2025-10-01 TO 2026-12-31



Source: Prepared by the researcher based on Python outputs.

Forecasted Returns for the Period (October 1, 2025 – December 31, 2026): Interpretation and Implications:

The forecasted results for the period from October 1, 2025 to December 31, 2026 represent the out-of-sample forecasts generated by the trained models after completing the learning phase on historical data (2018–2025). These forecasts aim to assess the models’ ability to predict future market movements in the Iraq Stock Exchange and to determine the general direction, stability, and volatility of the expected returns.

The results indicate the presence of a slight upward growth trend in returns during the first half of 2026, followed by a period of relative stability toward the end of the year. Although the expected return level remains low—consistent with the historical mean (≈ 0.0004)—the forecasted volatility is significantly lower, suggesting a potential gradual stabilization of the market after a period of turbulence. The models also show that under normal conditions, the market may experience a gradual improvement in return stability, while the high-volatility scenario highlights the risk of decline in the event of external shocks.

From an economic perspective, these forecasts suggest that the Iraq Stock Exchange may enter a transitional phase characterized by limited growth but increasing stability. The findings emphasize the importance of adopting adaptive investment strategies that account for regime shifts rather than relying on static models. Furthermore, the evidence supports the adoption of AI-based systems built on the Transformer model for dynamic market monitoring and risk management, providing a stronger analytical framework for policymakers and investors.

In summary, the forecasted results for the period (2025–2026) confirm that advanced deep learning models are capable of predicting future market behavior even in emerging and volatile environments. The Transformer model demonstrated clear superiority in predictive performance, making it the most suitable tool for analyzing and forecasting future returns in the Iraq Stock Exchange. This contributes to enhancing financial forecasting accuracy and improving investment and strategic planning in similar markets.

5. DISCUSSION

The empirical results of this study yield several pivotal insights that contribute to the existing body of knowledge in financial forecasting within emerging markets. The superior performance of the **Transformer model**, as evidenced by its minimal RMSE (0.012) and MAE (0.007), aligns with recent advancements in deep learning for sequence modeling (Vaswani et al., 2017). Its efficacy in this context can be attributed to the self-attention mechanism, which dynamically weights the importance of different time steps, thereby effectively capturing the complex, non-linear dependencies inherent in the Iraqi stock market's return series.

A critical finding is the stark **non-normality** of the return distribution, confirmed by the Jarque-Bera test. The extreme value of kurtosis (588.08) signifies a **leptokurtic distribution** with heavy tails, a characteristic often observed in high-frequency financial data but exceptionally pronounced here. This finding challenges the validity of many traditional financial models that rest on the assumption of normality, such as those derived from Modern Portfolio Theory. The substantial discrepancy between the VaR (-1.34%) and the much more severe CVaR (-17.22%) underscores that conditional expected losses in the tail are dramatically higher than the VaR threshold suggests. This reinforces the arguments of (Cont, 2001) on the inadequacy of variance and VaR as sole risk measures in markets with infinite or undefined variance.

The performance hierarchy of the models provides practical guidance for financial analysts. The strong showing of **SVR** and **ARIMA** indicates that these models can still capture a significant portion of the predictable structure in the data, likely related to short-term autocorrelation and conditional heteroskedasticity. Conversely, the poor performance of **Random Forest** highlights its limitations in modeling serial dependence, a known issue with tree-based models for time-series data. The **LSTM's** underwhelming results, despite its theoretical suitability for sequences, suggest it may require larger datasets than were available for this study to reach its potential, a common constraint in emerging market analysis.

The scenario analysis further contextualizes the model performances within a risk-management framework. The high Sharpe Ratios (~5.81) in the stable scenarios (Base, Optimistic, Pessimistic) demonstrate the potential for strong risk-adjusted returns during periods of low volatility. However, the **High Volatility Scenario** reveals the model's fragility and the market's inherent instability, with a catastrophic maximum drawdown and a negative Sharpe Ratio. This bifurcation in performance across regimes strongly supports the adoption of **regime-switching models** in investment strategy, as a one-size-fits-all approach is demonstrably ineffective.

6. CONCLUSIONS

This study provides a comprehensive empirical evaluation of quantitative forecasting models within the distinctive context of the Iraq Stock Exchange, a market characterized by pronounced volatility and non-normal return distributions. The investigation leads to the firm conclusion that the Transformer architecture stands as the most proficient model for financial return forecasting in this challenging environment. Its superior performance, quantified by the lowest error metrics and highest overall ranking, stems from its intrinsic self-attention mechanism, which is uniquely capable of identifying and weighting complex temporal dependencies and non-linear patterns that elude traditional statistical and other machine learning approaches.

A paramount conclusion of this research is the definitive identification of extreme leptokurtosis and the decisive rejection of normality in the return series. This fundamental characteristic of the data fundamentally invalidates the application of financial and econometric models that rely on assumptions of a normal distribution. The profound discrepancy between the Value at Risk and the Conditional Value at Risk metrics reveals that the potential for extreme losses is vastly greater than what conventional risk management frameworks typically anticipate. Consequently, it is concluded that the management of tail risk must be elevated to a primary concern for any investment or analytical activity in this market, necessitating a shift towards risk measures and models grounded in Extreme Value Theory.

Furthermore, the analysis conclusively demonstrates that the performance of forecasting models is not absolute but is intrinsically dependent on the prevailing market regime. The exceptional risk-adjusted returns, as measured by the Sharpe Ratio, achievable during periods of market stability stand in stark contrast to the catastrophic losses witnessed in high-volatility scenarios. This bifurcation necessitates a move away from static, single-strategy modeling towards dynamic, adaptive frameworks that can explicitly account for regime shifts. The poor performance of certain models, such as LSTM and Random Forest, underscores that model selection must be critically informed by the specific data attributes and volume available, with more complex models not universally guaranteeing superior results.

In summary, this research affirms that while the Iraqi market presents significant challenges for quantitative analysis, these are not insurmountable. The successful application of advanced deep learning architectures like the Transformer offers a powerful path forward. However, this success is critically contingent upon a foundational understanding of the market's unique statistical properties, a rigorous focus on tail-risk management, and the development of flexible, regime-aware investment strategies. The findings thus provide a refined and evidence-based roadmap for quantitative finance in emerging markets with similar profiles.

REFERENCES

1. Al-Homaidi, E. A., Al-Matari, E. M., & Nour, A.-N. I. (2022). A review of stock market prediction using machine learning techniques: Evidence from emerging markets. Journal of Financial Data Science, 4(1), 45-60.
2. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307--327.
3. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5--32.
4. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785--794). ACM.
5. Chen, Y., Wu, J., & Li, H. (2020). Forecasting financial market volatility: A comparison of GARCH, machine learning, and hybrid models. International Review of Financial Analysis, 72, 101599.
6. Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987--1007.
7. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654--669.
8. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223--2273.
9. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735--1780.
10. McNeil, A. J., Frey, R., & Embrechts, P. (2005). Quantitative risk management: Concepts, techniques, and tools. Princeton University Press.
11. Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., & Iosifidis, A. (2022). Forecasting stock prices from the limit order book using convolutional neural networks. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 2121-2130). IEEE.
12. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998--6008).
13. Zhang, Y., Li, A., & Li, Y. (2021). A novel hybrid model for stock price prediction based on metaheuristic optimization and deep learning. Expert Systems with Applications, 185, 115525.