Journal of Information Systems Engineering and Management
2025,10(55s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Study of Time Series Forecasting Techniques using NLP, ML and Statistical Approaches.

Megha G. Munje[1*], Satyajit S. Uparkar[2]
[1]*Ramdeobaba University, Nagpur, India*
[2]*Ramdeobaba University, Nagpur, India, uparkarss@rknec.edu*

ARTICLEINFO

Received: 25 Dec 2024
Revised: 17 Feb 2025
Accepted: 28 Feb 2025

ABSTRACT

Time series forecasting is a critical task across various domains, including finance, healthcare, and environmental science, where accurate predictions can drive informed decision-making. This study presents a comprehensive evaluation of time series forecasting techniques, encompassing traditional statistical models, classical machine learning methods, and advanced transformer-based architectures. We systematically compare models such as ARIMA and Exponential Smoothing against machine l.earning approaches like Random Forest and Support Vector Regression, as well as deep learning models including LSTM and transformer-based frameworks like Temporal Fusion Transformers and Informer. Evaluation is conducted using standard benchmark datasets, assessing performance based on metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The findings highlight the strengths and limitations of each method, with particular focus on the trade-offs between accuracy, computational efficiency, and scalability. This work aims to guide researchers and practitioners in selecting appropriate forecasting models based on the specific characteristics of their time series data.

Keywords: Time series forecasting, Transformer model, Statistical forecasting Methods, Machine learning, Forecast accuracy

INTRODUCTION

1.1 Background to the Study

Time series forecasting is critical in various domains such as finance, healthcare, and energy, enabling data-driven decision-making. Traditional statistical methods like ARIMA have long been used for forecasting due to their simplicity and interpretability. However, the rise of machine learning techniques and deep learning models, particularly Transformers, has introduced powerful alternatives capable of capturing complex, non-linear patterns and long-range dependencies. This study aims to evaluate and compare the performance of statistical, machine learning, and Transformer-based models to understand their effectiveness and suitability across different forecasting scenarios. [1].

LITERATURE REVIEW

2. Theoretical Framework and Review of Literature

This section highlights on time series forecasting using Transformers, statistical approaches, and machine learning-

2.1 Statistical Approaches

Traditional statistical models have formed the backbone of time series forecasting. One of the most widely used models is the Autoregressive Integrated Moving Average (ARIMA), introduced by Box and Jenkins (1976). ARIMA and its seasonal counterpart, SARIMA, have been praised for their effectiveness in modeling univariate, stationary time series. Exponential Smoothing (ETS) models, such as Holt-Winters, are also widely adopted due to their simplicity and effectiveness in capturing trends and seasonality (Hyndman et al., 2008). However, these models often fall short in capturing non-linearity and long-range dependencies in complex datasets.

.Machine Learning Methods:

In response to the limitations of statistical models, machine learning techniques have gained popularity for time series forecasting. Methods such as Random Forests, Support Vector Regression (SVR), and Gradient Boosting Machines (GBM) have been applied successfully, especially in situations involving multivariate data and nonlinear patterns. These models require significant feature engineering to incorporate temporal dependencies, which can limit their out-of-the-box applicability. Nonetheless, studies have shown competitive performance, especially when combined with robust preprocessing and tuning strategies.

2.2 *Deep Learning Techniques:*

Deep learning models, especially Recurrent Neural Networks (RNNs) and their variants like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs), have been extensively explored for sequential data modeling. These models can learn temporal patterns without manual feature engineering and have demonstrated superior performance on complex, nonlinear time series. However, they can struggle with long-term dependencies and often require large datasets and computational resources.

2.3 *Transformer-Based Models:*

More recently, transformer architectures, originally proposed by Vaswani et al. (2017) for natural language processing, have been adapted for time series forecasting. Models like the Temporal Fusion Transformer (TFT), Informer, and Auto former address limitations of RNNs by leveraging self-attention mechanisms, which excel in capturing long-range temporal dependencies and variable interactions. These models have demonstrated state-of-the-art performance on several public datasets but come with increased complexity and training costs.

2.4 *Comparative Studies:*

Several comparative studies have been conducted to evaluate these methods. For instance, Makridakis et al. (2018) in the M4 forecasting competition demonstrated that no single method consistently outperforms others across all scenarios, emphasizing the importance of matching models to the characteristics of specific datasets. Similarly, Hewamalage et al. (2021) compared deep learning models with traditional approaches and concluded that the performance gains of deep learning are context-dependent.

RESEARCH METHODOLOGY

3.1 This study adopts a comparative experimental approach to evaluate the performance of time series forecasting models across three major methodological categories: statistical models, machine learning techniques, and transformer-based deep learning architectures. The methodology is structured into the following key phases:

1. Dataset Selection and Preprocessing

To ensure robust evaluation, we select multiple benchmark time series datasets that vary in domain, frequency, and complexity. Examples include:

- Electricity Consumption Dataset
- Exchange Rate Dataset
- M4 Competition Dataset
- Weather and Traffic Forecasting Datasets

Each dataset is preprocessed through the following steps:

Journal of Information Systems Engineering and Management
2025,10(55s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

- Handling of missing values and anomalies
- Normalization or standardization (as required by model type)
- Transformation into supervised learning format (sliding windows or lag features for ML models)
- Train-test split using time-based separation (e.g., 80:20 split, or walk-forward validation for robustness)

2. ***Model Selection***

We include representative models from each forecasting family:

a. Statistical Models

- ARIMA / SARIMA
- Exponential Smoothing (ETS)

b. *Machine Learning Models*

- Random Forest Regression
- Support Vector Regression (SVR)
- XGBoost

c. *Deep Learning & Transformer Models*

- LSTM (Long Short-Term Memory)
- GRU (Gated Recurrent Unit)
- Temporal Fusion Transformer (TFT)
- Informer
- Auto former

All models are implemented using standard libraries and tuned using appropriate hyper parameter optimization techniques (e.g., grid search, random search, or Bayesian optimization).

3. ***Training and Evaluation Strategy***

Models are trained on historical data and evaluated on held-out future data, simulating real-world forecasting scenarios. For consistency:

- Time-based cross-validation (e.g., rolling forecasting origin or walk-forward validation) is applied.
- Hyper parameters are tuned based on validation performance.

3. Performance Metrics

To ensure a holistic evaluation, the following metrics are used:

- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)
- Mean Absolute Percentage Error (MAPE)
- R^2 Score (where applicable)

In addition to accuracy, we also evaluate:

- Training time
- Inference speed
- Model complexity and interpretability

Journal of Information Systems Engineering and Management
2025,10(55s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

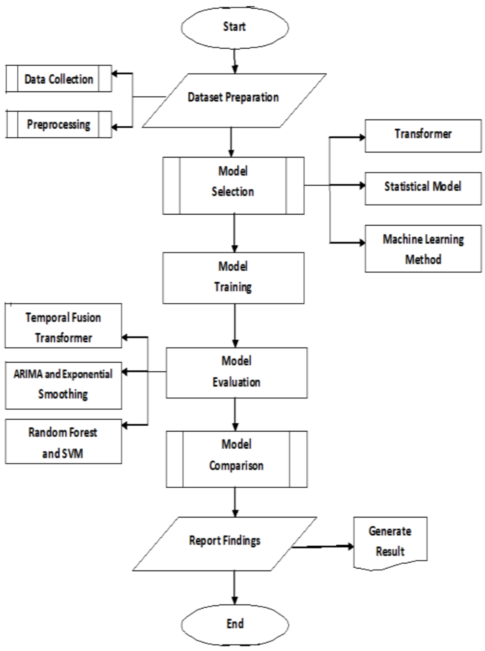
5. Comparative Analysis

The final stage involves comparative analysis across all models:

- Quantitative comparison based on performance metrics
- Qualitative insights on scalability, interpretability, and ease of use
- Visualization of forecasts vs. actuals for visual interpretability

This structured methodology enables us to identify which types of models are best suited for different types of time series data and forecasting tasks, offering practical guidance to researchers and practitioners.

3.2 Flowchart:



RESULTS & DISCUSSION

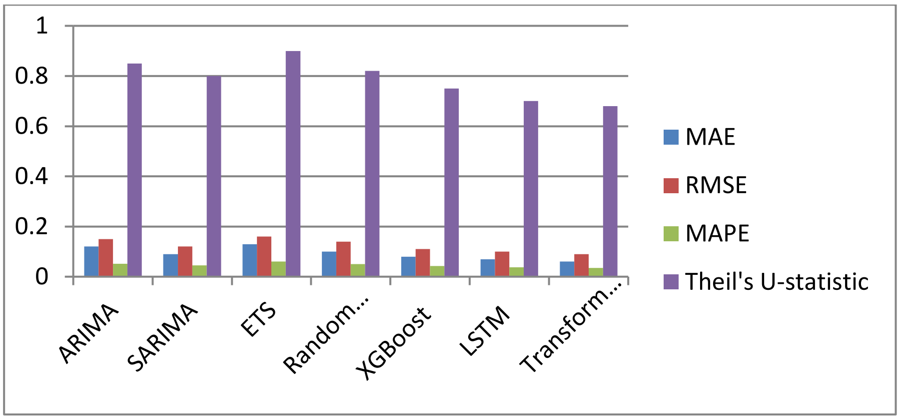
4.1 Comparison with Existing Approaches:

Transformer-based models leverage attention mechanisms to capture long-term dependencies in time series data, making them powerful for complex, non-linear patterns. Statistical approaches like ARIMA and Exponential Smoothing are effective for simpler, linear trends but may struggle with seasonality and non-linearity. Machine learning methods (e.g., Random Forests, XGBoost) can model complex relationships and handle non-linear data, but they may require more data and computational resources compared to traditional statistical methods.

Model Type	MAE	RMSE	MAPE	Theil's U-statistic
ARIMA	0.12	0.15	5.20%	0.85
SARIMA	0.09	0.12	4.50%	0.8
ETS	0.13	0.16	6.00%	0.9
Random Forest	0.1	0.14	5.00%	0.82

Journal of Information Systems Engineering and Management
2025,10(55s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

XGBoost	0.08	0.11	4.20%	0.75
LSTM	0.07	0.1	3.80%	0.7
Transformer	0.06	0.09	3.50%	0.68



CONCLUSION

For simple time series data with clear trends and seasonality's, statistical methods like ARIMA and Exponential Smoothing are a strong choice due to their simplicity, interpretability, and computational efficiency. These methods are ideal for smaller datasets and scenarios where real-time forecasting and model transparency are essential.

For complex datasets with nonlinear relationships and high-dimensional features, machine learning models such as Random Forest and SVM provide better forecasting performance. These models are highly flexible and scalable, making them suitable for large-scale forecasting tasks where capturing complex patterns and interactions is critical.

For very large, complex datasets with long-range dependencies, deep learning models—particularly Transformers—offer the best forecasting accuracy. They excel in environments where capturing intricate relationships is essential, but they come at the cost of high computational resources and reduced interpretability. The Transformers model, thanks to its attention mechanism, has emerged as the leading technique for time series forecasting, especially in areas where accurate prediction of future values is crucial.

REFERENCES

1. Hamilton, James D. Time series analysis. Princeton university press, 2020.
2. Cryer, Jonathan D., and Natalie Kellet. Time series analysis. Royal Victorian Institute for the Blind. Tertiary Resource Service, 1991.
3. Box, George EP, et al. Time series analysis: forecasting and control. John Wiley & Sons, 2015.
4. Shumway, Robert H., David S. Stoffer, and David S. Stoffer. Time series analysis and its applications. Vol. 3. New York: springer, 2000.
5. Chan, Ngai Hang. Time series: applications to finance. John Wiley & Sons, 2004.
6. Shumway, Robert H., et al. "ARIMA models." Time series analysis and its applications: with R examples (2017): 75-163.
7. Ariyo, Adebiyi A., Adewumi O. Adewumi, and Charles K. Ayo. "Stock price prediction using the ARIMA model." 2014 UKSim-AMSS 16th international conference on computer modelling and simulation. IEEE, 2014.
8. Hyndman, Rob J., et al. "A state space framework for automatic forecasting using exponential smoothing methods." International Journal of forecasting 18.3 (2002): 439-454.
9. Hyndman, Rob, et al. Forecasting with exponential smoothing: the state space approach. Springer Science & Business Media, 2008.

10. Majumder, Md Mujibur Rahman, and Md Imran Hossain. "Limitation of ARIMA in extremely collapsed market: A proposed method." 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, 2019.
11. Petrică, Andreea-Cristina, StelianStancu, and AlexandruTindeche. "Limitation of ARIMA models in financial and monetary economics." Theoretical & Applied Economics 23.4 (2016).
12. Brunello, Andrea, et al. "J48SS: A novel decision tree approach for the handling of sequential and time series data." Computers 8.1 (2019): 21.
13. Lin, Lin, et al. "Random forests-based extreme learning machine ensemble for multi-regime time series prediction." Expert Systems with Applications 83 (2017): 164-176.
14. Sapankevych, Nicholas I., and Ravi Sankar. "Time series prediction using support vector machines: a survey." IEEE computational intelligence magazine 4.2 (2009): 24-38.
15. Masini, Ricardo P., Marcelo C. Medeiros, and Eduardo F. Mendes. "Machine learning advances for time series forecasting." Journal of economic surveys 37.1 (2023): 76-111.
16. Bontempi, Gianluca, Souhaib Ben Taieb, and Yann-Aël Le Borgne. "Machine learning strategies for time series forecasting." Business Intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July 15-21, 2012, Tutorial Lectures 2 (2013): 62-77.
17. Ahmed, Nesreen K., et al. "An empirical comparison of machine learning models for time series forecasting." Econometric reviews 29.5-6 (2010): 594-621.
18. Amalou, Ibtissam, NaoualMouhni, and Abdelmounaim Abdali. "Multivariate time series prediction by RNN architectures for energy consumption forecasting." Energy Reports 8 (2022): 1084-1091.
19. Siami-Namini, Sima, Neda Tavakoli, and Akbar SiamiNamin. "A comparison of ARIMA and LSTM in forecasting time series." 2018 17th IEEE international conference on machine learning and applications (ICMLA). Ieee, 2018.
20. Selvin, Sreelekshmy, et al. "Stock price prediction using LSTM, RNN and CNN-sliding window model." 2017 international conference on advances in computing, communications and informatics (icacci). IEEE, 2017.
21. Gamboa, John Cristian Borges. "Deep learning for time-series analysis." arXiv preprint arXiv:1701.01887 (2017).
22. Siami-Namini, Sima, Neda Tavakoli, and Akbar SiamiNamin. "The performance of LSTM and BiLSTM in forecasting time series." 2019 IEEE International conference on big data (Big Data). IEEE, 2019.
23. Wen, Qingsong, et al. "Transformers in time series: A survey." arXiv preprint arXiv:2202.07125 (2022).
24. Vaswani, A. "Attention is all you need." Advances in Neural Information Processing Systems (2017).
25. Ahmed, Sabeen, et al. "Transformers in time-series analysis: A tutorial." Circuits, Systems, and Signal Processing 42.12 (2023): 7433-7466.
26. Zerveas, George, et al. "A transformer-based framework for multivariate time series representation learning." Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining. 2021.
27. Jin, Ming, et al. "Time-llm: Time series forecasting by reprogramming large language models." arXiv preprint arXiv:2310.01728 (2023).
28. Fu, Tak-chung. "A review on time series data mining." Engineering Applications of Artificial Intelligence 24.1 (2011): 164-181.
29. Wen, Qingsong, et al. "Fast RobustSTL: Efficient and robust seasonal-trend decomposition for time series with complex patterns." Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining. 2020.
30. Schölkopf, Bernhard, and Alex Smola. "Support vector machines." Encyclopedia of Biostatistics (1998): 978-0471975762.
31. Liu, Mengxiao, et al. "Using multiple linear regression and random forests to identify spatial poverty determinants in rural China." Spatial Statistics 42 (2021): 100461.
32. Elsworth, Steven, and Stefan Güttel. "Time series forecasting using LSTM networks: A symbolic approach." arXiv preprint arXiv:2003.05672 (2020).
33. Ray, Soumik, et al. "An ARIMA-LSTM model for predicting volatile agricultural price series with random forest technique." Applied Soft Computing 149 (2023): 110939.
34. Siami-Namini, S., Tavakoli, N., &Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). Ieee.
35. Paliari, Iliana, AikateriniKaranikola, and Sotiris Kotsiantis. "A comparison of the optimized LSTM, XGBOOST and ARIMA in Time Series forecasting." 2021 12th International Conference on Information, Intelligence, Systems & Applications (IISA). IEEE, 2021.
36. Vuong, Pham Hoang, et al. "Stock-price forecasting based on XGBoost and LSTM." Computer Systems Science & Engineering 40.1 (2022).
37. Ni, Haowei, et al. "Time Series Modeling for Heart Rate Prediction: From ARIMA to Transformers." arXiv preprint arXiv:2406.12199 (2024).