PressAcademia
www.pressacademia.org/conferences

GBRC
ISBN 978-625-97685-1-9

PressAcademia Procedia YEAR 2025 VOLUME 21

11th Global Business Research Congress, July 3, 2025

A COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS ON NETWORK TRAFFIC FORECASTING

DOI: 10.17261/Pressacademia.2025.1992
PAP- V.21-2025(11)-p.54-58

Buse Dilan Uslan[1], Ferhan Cebi[2]
[1]Istanbul Technical University, Department of Management Engineering, Istanbul, Turkiye.
uslan15@itu.edu.tr , ORCID: 0009-0007-9470-690X
[2]Istanbul Technical University, Department of Management Engineering, Istanbul, Turkiye.
cebife@itu.edu.tr , ORCID: 0000-0003-3100-3020

To cite this document
Uslan, B.D., Cebi, F., (2025). An analysis of unequal exchange in Turkiye's international trade. PressAcademia Procedia (PAP), 21, 54-58.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2025.1992

ABSTRACT

Purpose- The purpose of this study is to evaluate and compare the performance of four different time series forecasting models applied to mobile network traffic data, a domain characterized by high variability and complex seasonal patterns. Accurate forecasting of mobile capacity needs in the telecommunications sector is of great importance for providing uninterrupted and high-quality service. Since each network is unique, it is necessary to build a model that best predicts the seasonal traffic changes of the network.
Methodology- This research utilizes a comparative approach by implementing SARIMA, Prophet, LSTM, and a novel hybrid Prophet-LSTM model on monthly mobile traffic data from a telecommunications operator. The models were evaluated based on standard error metrics including MAE, MSE, RMSE and R^2 score.
Findings- The hybrid model leverages Prophet's trend-seasonality decomposition with LSTM's capability to learn nonlinear residual dynamics. The analysis reveals that the hybrid Prophet-LSTM model significantly outperforms the standalone SARIMA, LSTM, and Prophet models in terms of forecasting accuracy, flexibility, and adaptability. While SARIMA was limited in capturing complex, long-term trends and nonlinear fluctuations, LSTM required extensive hyperparameter tuning and was sensitive to data structure. Prophet proved to be effective in handling trend and seasonality with minimal parameter tuning, making it particularly suitable for cyclic patterns commonly observed in the telecommunications sector. However, the hybrid model's ability to leverage Prophet's decomposition strengths along with LSTM's temporal learning capacity enabled it to deliver the most robust predictions with the lowest error rates.
Conclusion- This model offers direct practical applications in network capacity planning, financial forecasting, and resource optimization processes. Moreover, it can be adapted for use in other sectors such as energy, transportation, and finance that rely heavily on time series data. Based on these findings, the hybrid Prophet-LSTM model is recommended for mobile traffic forecasting tasks involving both seasonal and nonlinear dynamics. Future studies may incorporate real-time streaming data and external factors to further improve predictive performance and real-world applicability.

Keywords: Network traffic prediction, SARIMA, Prophet, LSTM, hybrid model
JEL Codes: C32, C53, L96

1. INTRODUCTION

The rapid increase in mobile data usage, driven by the widespread adoption of smartphones, video streaming, and 5G technologies, has made efficient capacity planning a vital requirement for telecommunications operators. As user behavior becomes more dynamic and network infrastructures more complex, accurate forecasting of network traffic has emerged as a critical challenge in the industry. Poor forecasting can result in either underutilization of resources or service disruptions, both of which can significantly affect customer satisfaction and operational costs.

Time series forecasting models play a key role in predicting future traffic trends and enabling proactive network planning. However, the selection of an appropriate model is not straightforward, as mobile traffic data typically exhibits non-linear, seasonal, and sometimes abrupt behavioral shifts. Traditional statistical methods such as SARIMA are often limited in capturing such complexities. However, at the expense of greater model complexity and training time, machine learning-based models such as Long Short-Term Memory (LSTM) networks are more flexible and capable to learn complex time relationships.

In recent years, hybrid modeling approaches have gained traction, aiming to combine the strengths of different paradigms to improve forecasting performance. One such approach is the Prophet-LSTM hybrid model, which integrates Prophet's strength in trend and seasonality decomposition with LSTM's ability to learn residual dynamics. This hybridization seeks to produce more accurate and robust forecasts, particularly in domains like telecommunications where traffic patterns are both seasonal and irregular.

This paper presents a comparative analysis of four forecasting models—SARIMA, Prophet, LSTM, and a Prophet-LSTM hybrid—applied to real-world monthly mobile traffic data. By evaluating their performance using multiple error metrics, this study aims to identify the most effective modeling approach for telecommunications traffic forecasting. In doing so, it contributes both to the academic literature and to the practical needs of the telecommunications industry.

The remainder of the paper is organized as follows: Section 2 provides a review of related literature. Section 3 describes the dataset and methodology used. Section 4 presents the experimental results and analysis. Section 5 concludes the paper and outlines directions for future work.

2. LITERATURE REVIEW

Forecasting network traffic is a vital task for ensuring the efficiency, quality of service, and sustainability of operations in the telecommunications sector. As mobile data volumes increase due to 5G adoption and digitalization, accurate traffic prediction models are needed to support resource planning and proactive management.

Over the years, a wide range of time series forecasting models have been proposed and evaluated. Traditional statistical models such as ARIMA have been widely used for modeling linear time-dependent data. However, they often fall short in handling non-linear dynamics and complex seasonal behaviors. Prophet, developed by Facebook, has gained popularity due to its ability to model trend and seasonality with minimal tuning, as shown in the works of Subashini et al. (2019) and Cembaluk et al. (2022). While Prophet provides reliable predictions in many scenarios, some studies point out its limitations in handling high-frequency fluctuations without appropriate parameter optimization.

Deep learning-based models, particularly Long Short-Term Memory (LSTM) networks, have shown great promise in capturing long-term dependencies and non-linear patterns in sequential data (Katwal et al., 2023; Prajam et al., 2022). However, LSTM models require substantial training data and computational resources, and may be sensitive to hyperparameter configurations.

To address the limitations of single-model approaches, several studies have proposed hybrid models combining statistical and machine learning methods, Madan and Mangipudi (2018) introduced a DWT–ARIMA–RNN hybrid model, where wavelet transformation was used to decompose noise and complexity, followed by statistical and deep learning techniques for improved accuracy.

Shi et al. (2021) explored optimization-based hybrid models, integrating techniques such as Particle Swarm Optimization (PSO) and Variational Mode Decomposition (VMD) to enhance model precision. These decomposition-based frameworks improve forecasting by modeling each component individually and combining them. In a more recent example, Zaraket et al. (2024) proposed the "Hyper-Flophet" model, combining Prophet and LSTM to leverage both trend decomposition and sequential learning. Their results demonstrated that hybrid approaches generally outperform single models in complex traffic scenarios.

To support this idea, Karthika et al. (2017) proposed a hybrid method consisting of ARIMA and SVM models for short-term electricity load forecasting (STLF). The results showed that the ARIMA–SVM hybrid model (MAPE: 4.15%) achieved higher accuracy than using only ARIMA (MAPE: 5.16%) or only SVM (MAPE: 4.97%). Furthermore, it was noted that the hybrid model, combined with outlier correction, further reduced the error rate. This study makes a strong contribution to the idea of combining time series models with artificial intelligence methods to increase forecast accuracy.

Arslan (2022) introduces a hybrid time series forecasting model that combines Prophet and a stacked bidirectional LSTM to improve prediction accuracy. The approach involves decomposing the data to isolate and remove seasonality, training the LSTM on deseasonalized data for efficiency, and then reintegrating the seasonal component for final predictions. Evaluated on a real-world dataset of monthly energy consumption from seven countries, the hybrid model outperformed or matched the accuracy of several traditional and advanced models.

Zhang (2003) proposes a hybrid time series forecasting model that combines the linear ARIMA model with the nonlinear artificial neural network (ANN) to leverage their complementary strengths. While ARIMA captures linear patterns and ANNs handle nonlinear relationships, the hybrid model is designed to address complex time series with both components. Empirical results on three real datasets show that the combined model outperforms ARIMA and ANN used individually.

Aladag et. al. (2009) proposes a hybrid forecasting model that combines ARIMA with an Elman Recurrent Neural Network (ERNN), aiming to improve upon Zhang's earlier ARIMA-FNN hybrid approach. Since ERNNs include a context layer, they offer better forecasting accuracy than feedforward neural networks (FNNs). The proposed model was tested on the Canadian lynx dataset and showed improved performance compared to other methods.

Khashei (2011) proposes a hybrid forecasting method that combines a traditional time series model with a probabilistic neural network (PNN) to enhance prediction accuracy. The approach modifies the forecasts of the base model by analyzing and classifying the residual trends using the PNN, alongside optimizing the step length through mathematical programming. Tested on three real-world datasets, the hybrid model outperforms the standalone time series model, demonstrating its potential as a more accurate alternative for forecasting tasks.

In related domains, hybrid models such as Prophet-LSTM have also been employed in web traffic (Kong et al., 2021), energy demand (Bashir et al., 2022), and oil production forecasting (Ning et al., 2022). These studies collectively support the idea that hybrid modeling techniques can effectively bridge the gap between interpretability and predictive power.

Inspired by these findings, this study proposes a Prophet-LSTM hybrid model tailored for mobile network traffic forecasting. By leveraging Prophet's ability to handle trend-seasonality and LSTM's strength in modeling residuals, the model aims to deliver accurate, robust, and scalable forecasts. In contrast to previous works, this study applies the hybrid model to real operational data from a telecom operator, aiming to offer both theoretical insights and practical contributions for time series forecasting in large-scale networks.

3. DATA AND METHODOLOGY

The dataset used in this study consists of mobile network traffic records obtained from a telecommunications operator. Since the telecommunications sector is highly confidential, the dataset was created with the guidance of experts in the area. A synthetic dataset spanning from January 2023 to December 2024, including 24 months of historical downlink and uplink traffic data aggregated at the monthly level, was created withthe following steps: the average of actual data was taken, yearly seasonality was added, daily seasonality was added (Turkish customers tend to produce more data during the summer time and on the weekends)

The dataset was splitted into test and training subsets with the ratio of 30% and 70% simultaneously. Each of the models was trained with the same data and made the predictions on the same time span.

For SARIMA, Grid Search was used for hyperparameter tuning. For Prophet, weekly and yearly seasonality were activated. For LSTM, MinMaxScaler was used and hyperparameter tuning process was applied. For hybrid model, first Prophet was applied then the residuals were trained by LSTM and the combination of these two predictions was the final result.

Then, MAE, MSE, RMSE and R^2 score were calculated for each model. These calculations were made by the following formulas:

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

MSE = $\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

RMSE = $\sqrt{MSE}$

R^2 Score = $1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

Lower MAE, MSE, RMSE and R^2 score closer to 1 would give the best model performance.

4. FINDINGS

Performance of the models are given for comparison. MAE, MSE, RMSE and R^2 score were compared to find the best model. Each figure shows the test set and the predictions. Figure 1 shows that although SARIMA's error rates are interpreted as low, its R^2 score of 17.6% is well below expectations. A single metric is not sufficient to evaluate a model's performance, and it requires consideration from many different perspectives.

Figure 1: SARIMA Model Performance

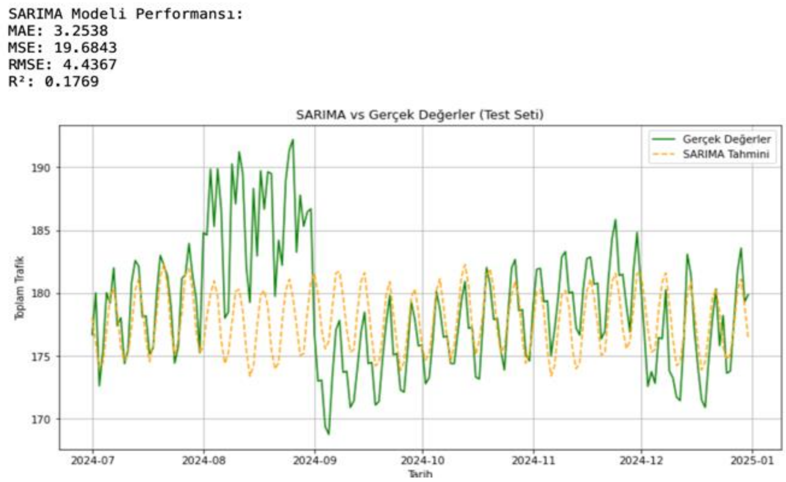


Figure 2 shows that the established model is quite successful since the MAE is close to zero and the R2 score is 76%.

Figure 2: Prophet Model Performance

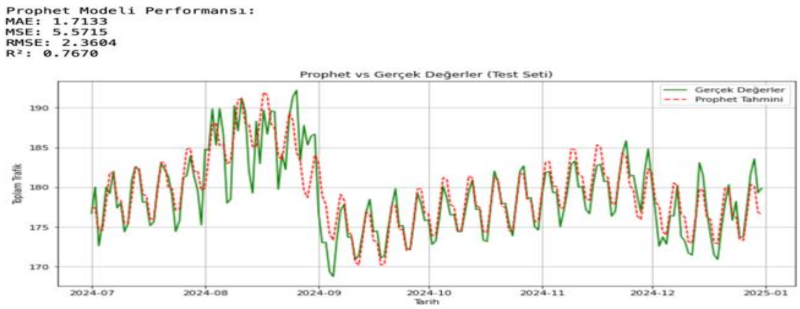


Figure 3 shows the performance of the LSTM model is lower than the Prophet model, but higher than SARIMA. It was able to explain 40% of the data.

Figure 3: LSTM Model Performance

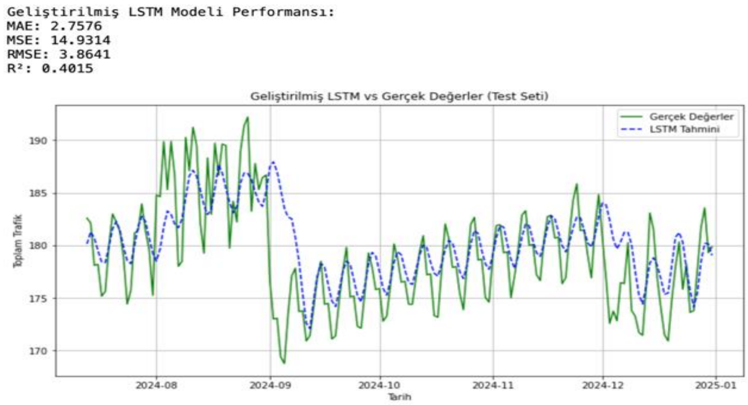
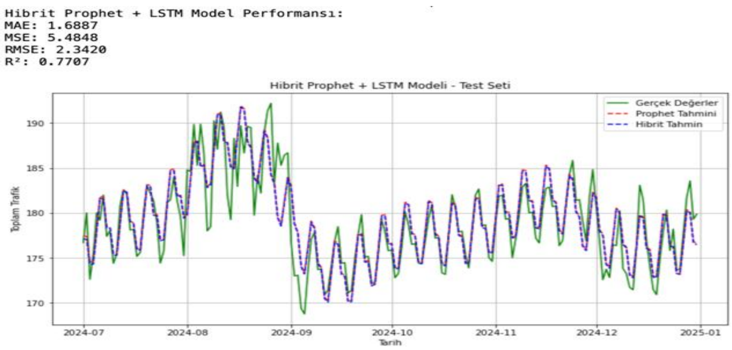


Figure 4 shows that the calculated error rates were quite low. The Prophet model explained 76.7% of the data, while the hybrid model explained 77%. This percentage improvement is anticipated to be even greater in larger, more realistic data sets.

Figure 4: Hybrid Model Performance



The summary of the performance metrics are given in Table 1.

Table 1: Comparison of performance metrics

	MAE	MSE	RMSE	R^2 Score
SARIMA	3.25	19.6	4.43	17%
LSTM	2.75	14.9	3.86	40%
Prophet	1.71	5.6	2.36	76.7%
Hybrid model	1.68	5.4	2.34	77%

5. CONCLUSION

Hybrid models are very promising for sectors dealing with time series data. It is seen that measuring the model performance with a real and longer data set would give a better outcome. External factors such as special days, holidays, match fixtures and campaign periods can be added to the model in order to obtain a real-life pattern. More precise and detailed predictions of models can be obtained by increasing data frequency to hourly level. This type of time series problems and hybrid model approaches can easily be adapted to other sectors working with time series data such as energy, transportation, e-commerce and finance.

REFERENCES

Aladag, C. H., Egrioglu, E., & Kadilar, C. (2009). Forecasting nonlinear time series with a hybrid methodology. Applied Mathematics Letters, 22(9), 1467–1470. https://doi.org/10.1016/j.aml.2009.02.006

Arslan, S. (2022). A hybrid forecasting model using LSTM and Prophet for energy consumption with decomposition of time series data. PeerJ Computer Science, 8, e1001. https://doi.org/10.7717/peerj-cs.1001

Bashir, S., Khan, A., Rehman, A., & Khan, M. (2022). Short-term electric load forecasting using hybrid Prophet–LSTM model. *Energy Reports, 8*, 345–355. https://doi.org/10.1016/j.egyr.2022.03.005

Cembaluk, P., Aniszewski, J., Knapińska, A., & Walkowiak, K. (2022). Forecasting the network traffic with PROPHET. Proceedings of the 3rd Polish Conference on Artificial Intelligence, 215–218.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://otexts.com/fpp2/

Karthika, S., Margaret, V., & Balaraman, K. (2017, August). Hybrid short term load forecasting using ARIMA-SVM. 2017 Innovations in Power and Advanced Computing Technologies (i-PACT) (pp. 1–6). IEEE. https://doi.org/10.1109/IPACT.2017.8245060

Katwal, A., Sharma, P., & Thakur, R. (2023). Comparative study of time series models for website traffic forecasting. Journal of Data Science Applications*,* 5*(2), 44–53.*

Khashei, M., & Bijari, M. (2011). A new class of hybrid models for time series forecasting. Expert Systems With Applications, 39(4), 4344–4357. https://doi.org/10.1016/j.eswa.2011.09.157

Kong, Y. H., Lim, K. Y., & Chin, W. Y. (2021). Forecasting Facebook user engagement using hybrid Prophet and long short-term memory model. International Conference on Digital Transformation and Applications (ICDXA), October 25-26, 2021, Kuala Lumpur, Malaysia. https://www.tarc.edu.my/files/icdxa/proceeding_new/920C21C7-4E28-44B4-9E3A-ECB64CBB13CD.pdf

Madan, R., & Mangipudi, P. S. (2018). Predicting computer network traffic: A time series forecasting approach using DWT, ARIMA and RNN. 2018 Eleventh International Conference on Contemporary Computing (IC3), 1–6. https://doi.org/10.1109/IC3.2018.8530608

Ning, Y., Kazemi, H., & Tahmasebi, P. (2022). A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. Computers & Geosciences, 164, 105126. https://doi.org/10.1016/j.cageo.2022.105126

Prajam, S., Wechtaisong, C., & Khan, A. A. (2022). Applying machine learning approaches for network traffic forecasting. Indian Journal of Computer Science and Engineering, 13(2), 324–330. https://doi.org/10.21817/indjcse/2022/v13i2/221302188

Shi, J., Leau, Y.-B., Li, K., & Obit, J. H. (2021). A comprehensive review on hybrid network traffic prediction model. International Journal of Electrical and Computer Engineering (IJECE), 11(2), 1450–1459. https://doi.org/10.11591/ijece.v11i2.pp1450-1459

Subashini, A., Sandhiya, K., Saranya, S., & Harsha, U. (2019). Forecasting Website Traffic Using Prophet Time Series Model. International Research Journal of Multidisciplinary Technovation, 1(1), 56–63. https://doi.org/10.34256/irjmt1917

Taylor, S. J., & Letham, B. (2017). Forecasting at scale. https://doi.org/10.7287/peerj.preprints.3190v2

Zaraket, K., Harb, H., Bennis, I., Jaber, A., & Abouaissa, A. (2024). Hyper-Flophet: A neural Prophet-based model for traffic flow forecasting in transportation systems. Simulation Modelling Practice and Theory, 134, 102954. https://doi.org/10.1016/j.simpat.2024.102954

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175. https://doi.org/10.1016/s0925-2312(01)00702-0