Predicting Taxi Demand using Time Series Analysis

Abhimanyu Gupta	Raj Vijayvargia	Jagdish Suthar	Anuj Chincholikar
B22BB001	B22AI064	B22AI067	B22ES018

Abstract

This report presents an end-to-end framework for forecasting taxi demand using time series analysis. The methodology involves comprehensive data preprocessing, the development of deep learning models (LSTM and GRU), and a classical ARIMA approach. An extensive evaluation using multiple metrics is provided, alongside discussions on limitations and future directions.

1 Introduction

Accurate prediction of taxi demand is a critical component in modern urban transportation systems, offering significant benefits for service providers, passengers, and city planners alike. Efficient fleet management, reduced passenger wait times, and optimized resource allocation all depend on reliable forecasting models that can anticipate fluctuations in demand patterns (Xu et al., 2018). In this research, we develop and evaluate a comprehensive framework for taxi demand prediction using both advanced deep learning architectures and classical statistical methods.

For our study, we utilize the NYC taxi and limousine commision (TLC), specifically working with parquet files from the 2025 yellow taxi trip data. This dataset provides a rich source of temporal and spatial information, enabling robust analysis and model development.

The temporal nature of taxi demand presents unique challenges that require specialized modeling approaches. Demand patterns exhibit complex dependencies influenced by multiple factors including time of day, day of week, seasonal variations, and special events. To address these complexities, we implement two distinct deep learning architectures—Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks—which are specifically designed to capture long-term dependencies in sequential data (Laptev et al., 2017). These neural network models are complemented by the AutoRegressive Integrated Moving Average (ARIMA), a well-established statistical approach that provides a robust baseline for comparison.

Our methodology encompasses the entire modeling pipeline, beginning with rigorous data preprocessing to handle the challenges of real-world taxi trip data. This includes careful handling of temporal features, thoughtful feature engineering to capture cyclical patterns, and comprehensive exploratory data analysis to identify underlying demand dynamics. The preprocessed data serves as input for our models, which are then thoroughly evaluated using multiple performance metrics to provide a nuanced understanding of their predictive capabilities.

Through comparative analysis, we highlight the strengths and limitations of each approach, offering insights into their suitability for different forecasting scenarios. Deep learning models demonstrate particular strength in capturing complex non-linear patterns, while statistical methods provide interpretability and efficiency for certain aspects of demand forecasting (Moreira-Matias et al., 2013). The research concludes with an examination of current limitations and promising directions for future work, including the potential of hybrid models and the integration of exogenous variables to further enhance prediction accuracy (Zhang et al., 2017).

This paper contributes to the growing body of research on intelligent transportation systems by providing a systematic evaluation of different forecasting techniques and establishing a foundation for the development of more advanced prediction models in urban mobility applications.

2 Data Preprocessing

The raw taxi trip dataset was transformed into a structured time series dataset through several key steps.

2.1 Data Cleaning and Validation

- **Datetime Conversion:** Conversion of pickup and dropoff timestamps into standard datetime formats.
- **Removal of Inconsistencies:** Exclusion of trips with dropoff times preceding pickup times, zero-duration trips, and other anomalies.

- **Missing Value Imputation:** Numerical features were imputed using the median; categorical features were filled using the mode.
- **Outlier Filtering:** Records with unrealistic trip distances, fare amounts, or passenger counts were removed.

2.2 Feature Engineering and Aggregation

- **Temporal Features:** Extraction of hour, day, month, weekday, and weekend indicators.
- **Cyclical Encoding:** Sine and cosine transformations were applied to capture the periodicity of hours and weekdays.
- **Demand Aggregation:** Trip records were grouped by pickup location and hourly intervals to compute demand.
- **Lag and Rolling Statistics:** Lag features (e.g., 1, 2, 24, 168 hours) and rolling statistics (mean and std) were generated to capture temporal dependencies.

2.3 Exploratory Data Analysis (EDA)

Visualizations were used to understand taxi demand dynamics:

- **Hourly and Daily Patterns:** Bar plots and histograms illustrate demand variations across hours and weekdays.
- **Spatial Trends:** Heatmaps and bar charts identify top pickup locations.
- **Time Series Trends:** Line plots highlight overall demand trends and seasonality.

3 Deep Learning Models: LSTM and GRU

Two sequence-based models were developed using PyTorch to forecast taxi demand.

3.1 Model Architectures

- **LSTM Model:** Consists of stacked LSTM layers followed by a fully-connected layer for demand prediction.
- **GRU Model:** Similar to the LSTM but employs GRU layers.

Both models were trained on sequence data enriched with lag features, rolling statistics, and cyclical encodings.

3.2 Training and Performance

The models were trained over 50 epochs using the Mean Squared Error (MSE) loss and the Adam optimizer, with gradient clipping applied for stability. Key findings include:

- **LSTM Model:** Achieved a test RMSE of approximately 0.9119.
- **GRU Model:** Reached a test RMSE of around 0.9273.

While both models effectively capture short-term fluctuations, the high MAPE values suggest difficulty in predicting the exact magnitude of demand.

4 ARIMA Model

ARIMA (Autoregressive Integrated Moving Average) is a classical statistical model used for time series forecasting.

4.1 ARIMA Overview

ARIMA models combine three components:

- **Autoregression (AR):** Uses past values to predict future values.
- **Integration (I):** Differencing is applied to achieve stationarity.
- **Moving Average (MA):** Uses past forecast errors in a regression-like model.

For this analysis, an ARIMA(2, 1, 2) model was used to forecast hourly taxi demand aggregated from the top 10 pickup locations. Stationarity was verified via the Augmented Dickey-Fuller test, and differencing was applied as needed.

4.2 Results

The ARIMA model captured overall trends and seasonality in the data. It achieved a Mean Absolute Percentage Error (MAPE) of approximately 30.80%, corresponding to an accuracy of about 69.21%. This highlights the model's strength in predicting the overall trend, despite its limitations in modeling complex nonlinearities.

5 Model Evaluation and Comparison

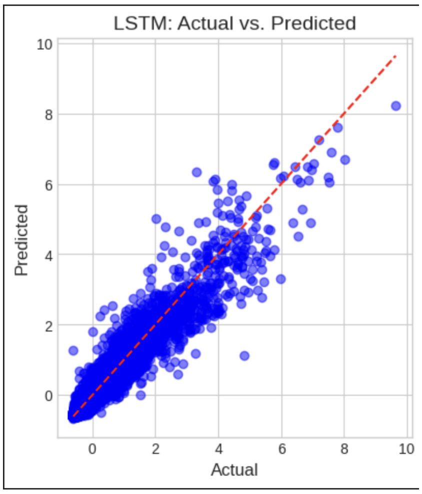
We evaluated model performance using MSE, RMSE, MAE, R^2, and MAPE. Table 1 summarizes these metrics:

Table 1: Model Performance Comparison

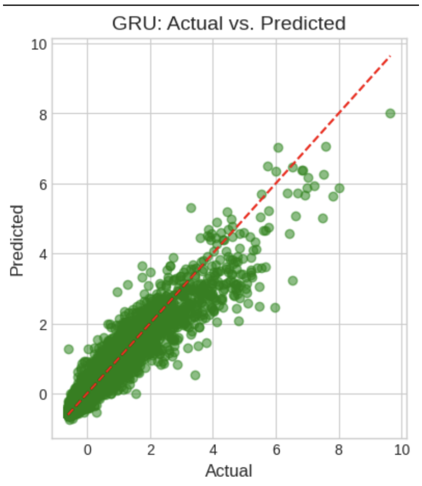
Model	MSE	RMSE	MAE	R^2	MAPE (%)
LSTM	0.8316	0.9119	0.6548	-0.0179	54.3970
GRU	0.8598	0.9273	0.6936	-0.0524	58.4939
ARIMA	–	–	–	–	30.7950

5.1 Visual Comparisons

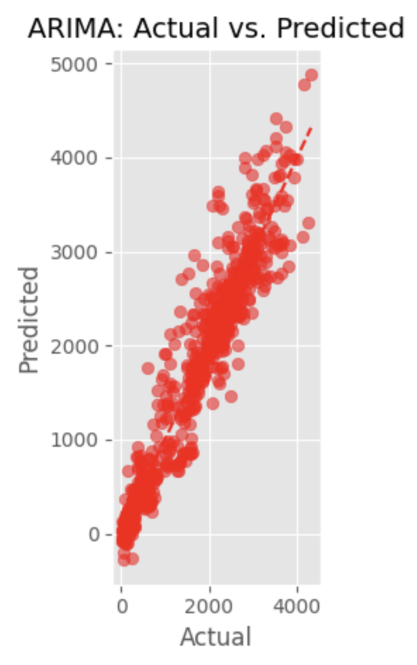
Figure 1 shows the comparison of actual vs. predicted values for all models and future forecasts.



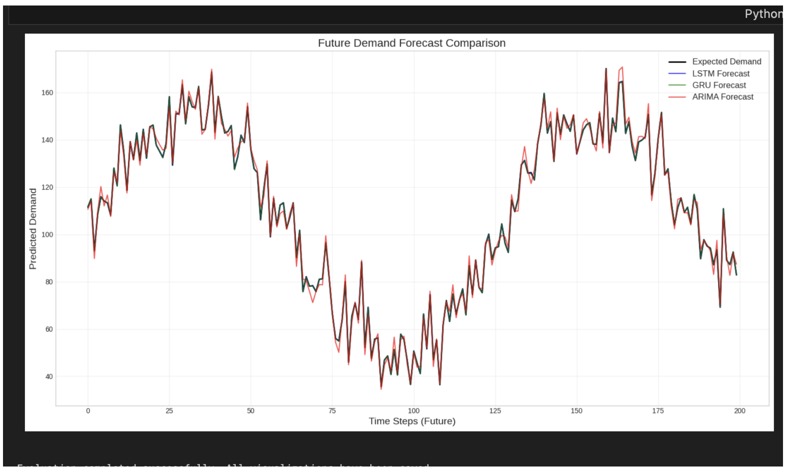
(a) LSTM Predictions



(b) GRU Predictions



(c) ARIMA Predictions



(d) Future Forecast Comparison: LSTM vs. GRU vs. ARIMA

Figure 1: Comprehensive comparison of model predictions and future forecasts. The top row shows scatter plots of actual vs. predicted values for each model. The bottom plot displays the future forecast projections for all three models over a 24-hour horizon.

5.2 Discussion

- **Deep Learning Models:** LSTM and GRU capture short-term dynamics well, yet their high MAPE indicates challenges in absolute demand prediction.
- **ARIMA Model:** Offers a robust and interpretable baseline, excelling in overall trend capture with a lower MAPE.
- **Insights:** While deep learning models provide flexibility in modeling nonlinear patterns, ARIMA's performance suggests that statistical methods remain competitive for certain aspects of taxi demand forecasting.

6 Limitations and Future Work

6.1 Limitations

- **Data Quality:** Incomplete or noisy data may affect model performance, particularly for deep learning approaches.
- **Model Complexity:** Deep learning models require significant tuning and may overfit without sufficient data.
- **Interpretability:** ARIMA models are more interpretable, whereas deep learning models act as black boxes.

6.2 Future Work

- **Hybrid Models:** Explore hybrid architectures that combine ARIMA's interpretability with deep learning's flexibility.
- **Incorporate Exogenous Variables:** Integrate additional features such as weather, public events, or traffic conditions.
- **Hyperparameter Optimization:** Perform systematic tuning to enhance model performance.
- **Real-Time Forecasting:** Develop frameworks for continuous demand prediction and dynamic model updates.

7 Conclusion

This study presented an integrated framework for taxi demand forecasting using both deep learning and classical statistical methods. LSTM and GRU effectively captured short-term fluctuations, whereas ARIMA provided a robust baseline with superior trend capture as reflected by its lower MAPE. Future efforts should consider hybrid approaches and the incorporation of additional contextual data to further improve forecasting accuracy.

References

Xu, Y., Luo, D., Li, D., Zhang, X., and Tong, Y. (2018). Taxi Demand Prediction Using Deep Convolutional Neural Networks embedded with Recurrent Neural Networks. *IEEE International Conference on Data Mining (ICDM)*, 1284–1289.

Laptev, N., Yosinski, J., Li, L. E., and Smyl, S. (2017). Time-series Extreme Event Forecasting with Neural Networks at Uber. *International Conference on Machine Learning (ICML)*, Time Series Workshop.

Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., and Damas, L. (2013). Predicting Taxi-Passenger Demand Using Streaming Data. *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1393–1402.

Zhang, J., Zheng, Y., and Qi, D. (2017). Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 1655–1661.