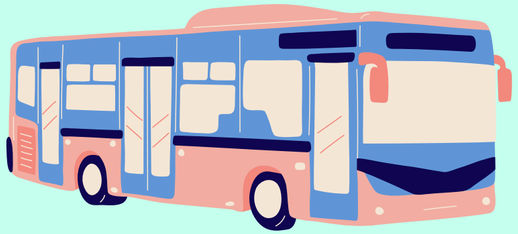


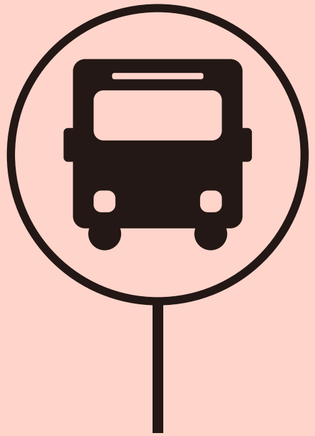
Bus Arrival Time Prediction

Sharipkhanova Dariya BDA-2201

Introduction

- Improving reliability of public transportation.
- Reducing waiting times with accurate bus arrival predictions.
- Integrating real-time GPS, traffic, and weather data.

Goals and Objectives



Goal:

Develop and evaluate models for accurate predictions.

Objectives:

- Integrate various data sources.
- Implement machine learning models.
- Evaluate model performance.

Methods and metrics

Data Sources:

- GTFS schedules
- GPS data
- Traffic & weather data

Models:

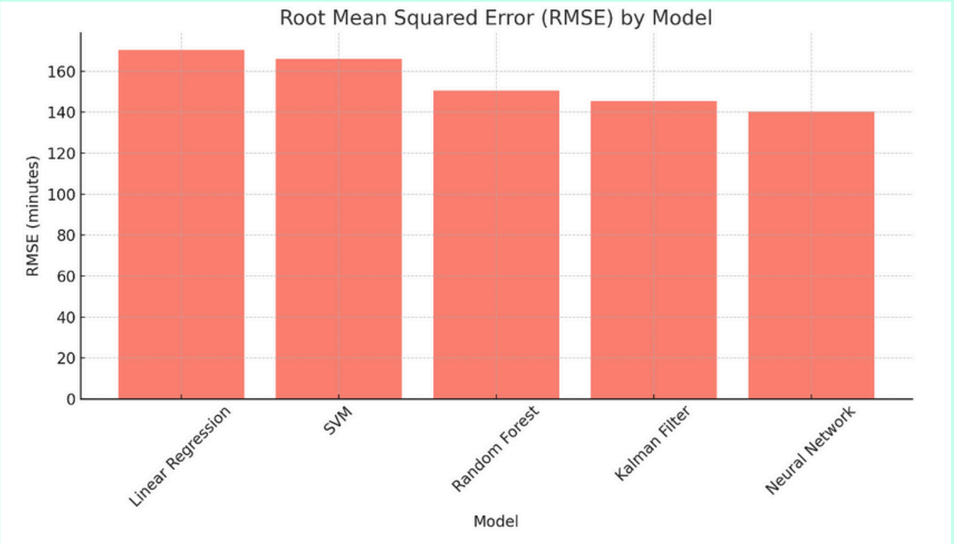
- Linear Regression
- SVM
- Random Forest
- Kalman Filter

Evaluation Metrics

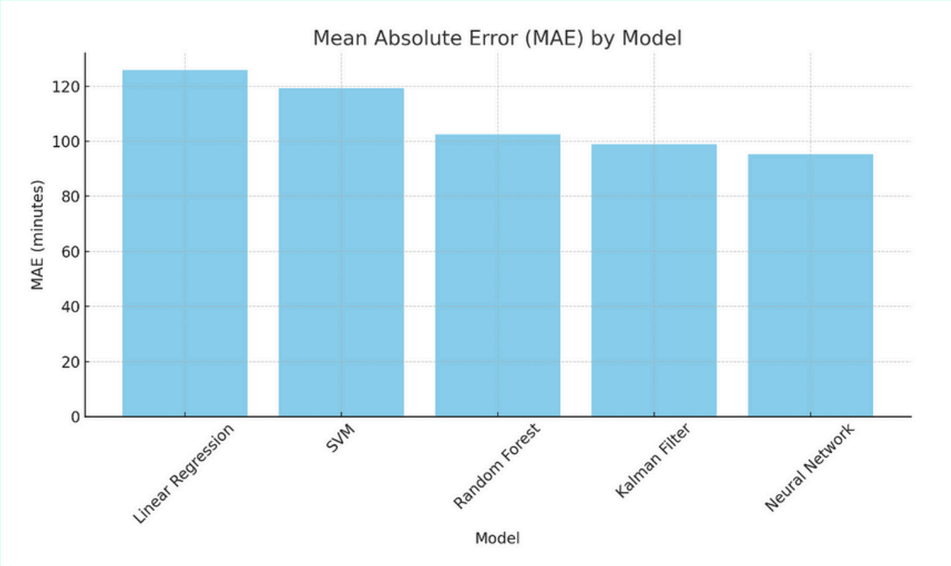
- Mean Absolute Error (MAE): Measures average prediction error.
- Root Mean Squared Error (RMSE): Penalizes larger errors.

Results

- Neural Network achieved the lowest error.
- Kalman Filter improved sequential estimations.
- Random Forest captured non-linear patterns effectively.



RMSE by Model



MAE by Model

Discussion and Conclusion

- Neural Networks provide the most accurate predictions.
- Kalman Filtering improves real-time accuracy.
- Future work:
  - Hyperparameter tuning
  - Real-time deployment
  - Weather and traffic data integration

References

1. https://www.sciencedirect.com/science/article/abs/pii/S0968090X11000155
2. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=34e59070faa4472fb94e3d3413c897f5ad409b11
3. https://www.kaggle.com/datasets/charvibannur/gtfs-traffic-prediction-dataset
4. https://www.communitytransit.org/open-data
5. https://docs.2gis.com/ru/api/navigation/directions/overview
6. https://pro.arcgis.com/ru/pro-app/3.1/help/analysis/networks/historical-traffic.htm
7. https://yandex.ru/dev/weather/doc/ru/concepts/history?utm_source=chatgpt.com

Thanks for your attention!