UPSKILLS DATA SCIENCE AND MACHINE LEARNING INTERNSHIP

WEEK - 5

KETAN JHANWAR

GOVT. ENGINEERING COLLEGE, AJMER

UPSKILLS DATA SCIENCE AND MACHINE LEARNING INTERNSHIP

WEEK - 5

I would like to provide you with a progress report for my fifth week in the Upskills UCT Machine Learning and Data Science Internship. The following points highlight the key aspects of my activities and experiences:

Project Overview:

The Smart City Traffic Pattern ML project aims to analyze and predict traffic patterns in a smart city environment using machine learning techniques. By understanding and predicting traffic patterns, we can optimize traffic flow, improve transportation efficiency, and enhance overall urban mobility. This report provides an overview of the problem statement and discusses potential algorithms that can be employed in the project.

Problem Statement:

You are working with the government to transform your city into a smart city. The vision is to convert it into a digital and intelligent city to improve the efficiency of services for the citizens. One of the problems faced by the government is traffic. You are a data scientist working to manage the traffic of the city better and to provide input on infrastructure planning for the future. The government wants to implement a robust traffic system for the city by being prepared for traffic peaks. They want to understand the traffic patterns of the four junctions of the city. Traffic patterns on holidays, as well as on various other occasions during the year, differ from normal working days. This is important to take into account for your forecasting.

Now we done the basic study of the PS and our dataset and evaluated the following facts about the given dataset and what we have to submit in the final project report. So, we will work accordingly.

Data Dictionary

Variable	Description
ID	Unique ID
DateTime	Hourly Datetime Variable
Junction	Junction Type
Vehicles	Number of Vehicles (Target)

sample_submission.csv

Column Name	Description
ID	Unique ID
Vehicles	Number of Vehicles (Target)

Progress Summary: During Week 5, we focused on implementing the Keras modeling approach to build a Recurrent Neural Network (RNN) for traffic forecasting. The key activities and achievements are outlined below:

- **Keras Modeling:**
- Leveraged the Keras library to design and construct the RNN architecture for traffic forecasting.

- Keras provides a user-friendly and powerful interface for creating deep learning models, making it ideal for our project's needs.
- **Root Mean Squared Error (RMSE) as Cost Function:**
- Chose RMSE as the cost function (also known as the loss function) for our traffic forecasting RNN model.
- RMSE is a common metric used in regression tasks, and it measures the difference between predicted and actual values, providing insight into the model's accuracy.
- **Initialization of RNN:**
- Initialized the RNN with the necessary parameters and configurations to create a basic structure.
- Selected the LSTM (Long Short-Term Memory) architecture for the RNN as it is well-suited for sequential data, such as time series.
- **Adding Input Layer and LSTM Layer:**
- Set up the input layer to accept the preprocessed and reshaped traffic data as input to the RNN model.
- Added LSTM layers to the RNN to enable memory and sequential learning, which is crucial for time series forecasting tasks.
- **Fitting the RNN to Training Set and Validating the Model:**
- Used the training set generated in Week 4 to train the RNN model.
- Defined the number of epochs (training iterations) and batch size to control the training process.

- Monitored the model's performance on the validation set using RMSE as the evaluation metric.
- Continued to iterate and fine-tune the model based on the validation results to achieve optimal performance.
- **Next Steps:** Moving forward, the following tasks will be undertaken in Week 6:
- **Model Evaluation and Fine-Tuning:**
- Evaluate the trained RNN model on the testing (or validation) set to assess its performance in real-world traffic forecasting scenarios.
- Fine-tune the model's hyperparameters and architecture based on the evaluation results to optimize its accuracy.
- **Integration with Real-Time Traffic Monitoring:**
- Integrate the validated RNN model with the real-time traffic monitoring system.
- Develop mechanisms to update the model with the latest traffic data, enabling adaptive and dynamic traffic predictions.
- **Collaboration with Infrastructure Planning Teams:**
- Continue collaborating with infrastructure planning teams and stakeholders.
- Provide them with updated traffic forecasts and insights to support informed infrastructure planning decisions.
- **Continuous Evaluation and Improvement:**
- Regularly evaluate the traffic management strategies implemented based on real-time data and feedback from citizens and stakeholders.
- Identify areas for improvement and make necessary adjustments to optimize traffic flow, reduce congestion, and enhance overall citizen satisfaction.

- **Challenges and Risks:**
- The choice of hyperparameters in the RNN model can significantly impact its performance, and fine-tuning them requires careful consideration.
- Ensuring the RNN model's generalization ability to handle varying traffic conditions and unexpected events is a potential challenge.
- Integrating the RNN model with the real-time monitoring system might require additional data processing and synchronization.

- **Conclusion:** The fifth week of our traffic management project marked significant progress in building an effective traffic forecasting model using Keras with Long Short-Term Memory (LSTM) architecture. By incorporating lag features, scaling the input data, and leveraging the RMSE as the cost function, we have improved the model's ability to capture temporal dependencies and make accurate traffic predictions.
- The addition of lag features allowed the LSTM model to consider historical traffic patterns, contributing to more accurate forecasts. Proper feature scaling with Min-Max Scaler and Standard Scaler ensured that all features were on a similar scale, preventing any bias in the model training and ensuring convergence.
- Furthermore, we successfully split the dataset into training and testing sets, adhering to the temporal order of the data. This allowed us to effectively evaluate the model's performance and make fine-tuning adjustments during training.
- In the subsequent phase, we designed and implemented the LSTM neural network with appropriate parameters and hyperparameters. By adding input layers and

LSTM layers, we enabled the model to capture temporal dependencies and patterns in the sequential traffic data. The choice of RMSE as the cost function for training was well-suited to the traffic forecasting task, penalizing larger prediction errors more significantly.

- By fitting the RNN to the training set and validating the model on the testing (or validation) set, we assessed the LSTM model's performance and iteratively fine-tuned it to achieve better forecasting accuracy. This validation process allowed us to gauge the model's generalization capabilities and make necessary improvements.
- Looking ahead to Week 6, we plan to conduct further hyperparameter tuning to optimize the LSTM model's architecture and enhance traffic forecasting accuracy. The integration of the validated LSTM model with the real-time traffic monitoring system will enable us to provide up-to-date traffic forecasts and adapt to changing traffic patterns dynamically.
- Moreover, we will continue collaborating closely with infrastructure planning teams and stakeholders, sharing accurate traffic forecasts and insights. This alignment of traffic management strategies with the model's predictions will guide future infrastructure decisions and foster the development of a smart city with efficient traffic management.
- As we move forward, continuous evaluation and improvement will remain crucial in ensuring the effectiveness of the implemented traffic management strategies. We will closely monitor the model's performance in real-time and gather citizen

feedback to assess the impact of our initiatives on overall traffic flow and congestion reduction.

- Overall, the progress made in Week 5 sets a solid foundation for the subsequent stages of the project, paving the way for a comprehensive traffic management system that aligns with the goals of transforming our city into a smart and efficient urban environment.

Thanks and Regards

Ketan Jhanwar