ADVANCED MACHINE LEARNING

(BA – 64061)

ASSIGNMENT – 3

TIME SERIES DATA

DIVYA CHANDRASEKARAN

ID: 811284790

OBJECTIVES

1. To compare various models for time series data analysis, including common-sense methods, basic machine learning models, and advanced neural network architectures.

2. To identify the best-performing model that accurately captures long-term dependencies and patterns within sequential data.

3. To investigate the impact of the vanishing gradient problem on simple Recurrent Neural Networks (RNNs) and evaluate solutions through advanced RNN variants such as LSTM and GRU.

4. To explore the effectiveness of combining convolutional models with RNNs in handling time series data.

PROBLEM STATEMENT

Time series data analysis poses unique challenges due to the necessity of capturing temporal dependencies and patterns within sequential data. Traditional machine learning models and simple neural networks often fail to retain the temporal context, leading to suboptimal predictions. The vanishing gradient problem further exacerbates the issue, particularly in simple RNNs, preventing the effective modeling of long-term dependencies. This study seeks to identify and validate more sophisticated models, such as LSTM and GRU, which are designed to overcome these limitations and provide superior performance in time series data analysis.

INTRODUCTION

In this report, I have attempted to discover the different machine learning approaches for time-series modeling and evaluate their performance using Mean Absolute Error (MAE) as a metric.

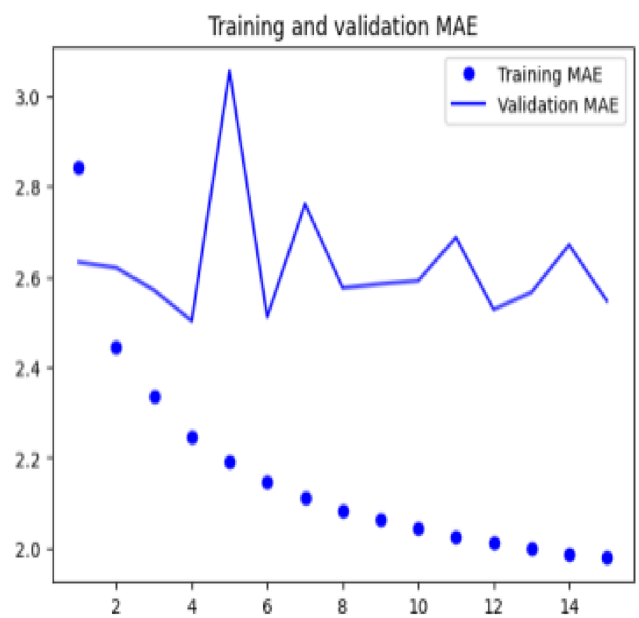
I conducted a comparison of several model configurations. Initially, I implemented a basic machine learning example to establish a baseline. Next, I integrated Long Short-Term Memory (LSTM) layers. LSTM networks are a subtype of recurrent neural networks designed to grasp order dependencies in sequence prediction tasks. Following that, I employed a blend of 1D convolutional layers (1D convnets) and recurrent neural network (RNN) layers, specifically LSTM, to capture both spatial and temporal patterns simultaneously. Lastly, I fine-tuned the number of units in each recurrent layer to optimize the model's architecture.

Structure of the report:

- **<u>Simple Machine Learning Illustration:</u>**

  - The test dataset produced an MAE outcome of 8.85.

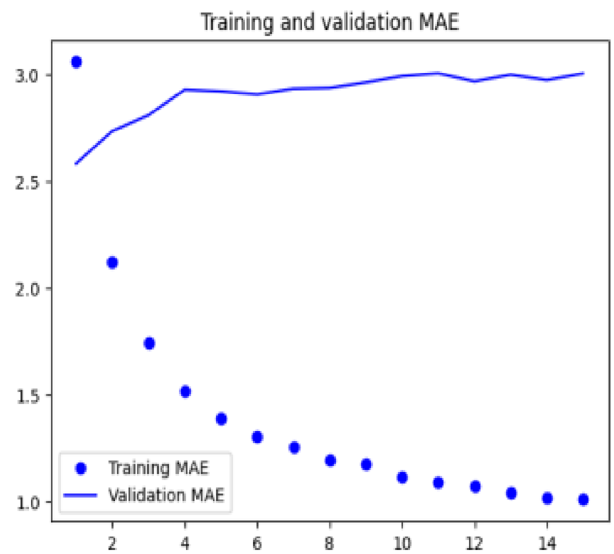
  - This model served as a reference point for my subsequent endeavors in the time-series forecasting task, excluding the utilization of advanced methodologies tailored to sequential data. The somewhat elevated MAE indicates potential shortcomings in capturing temporal patterns and dependencies.



- **<u>LSTM Configuration</u>**:

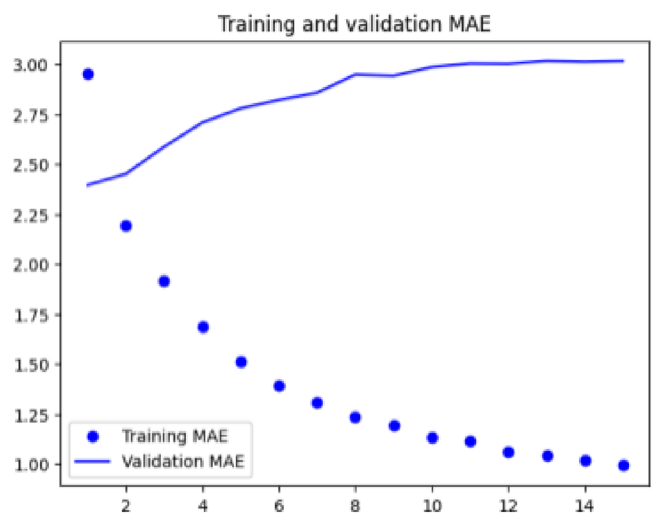
- The test dataset yielded an MAE outcome of 3.12.

- Employing a layered structure of LSTM layers led to significant enhancement compared to the basic machine learning strategy. The incorporation of LSTM layers enabled the model to more effectively grasp long-term dependencies within the time-series data, leading to a reduced MAE.



- **<u>Combinations of 1D Convnets & RNN:</u>**

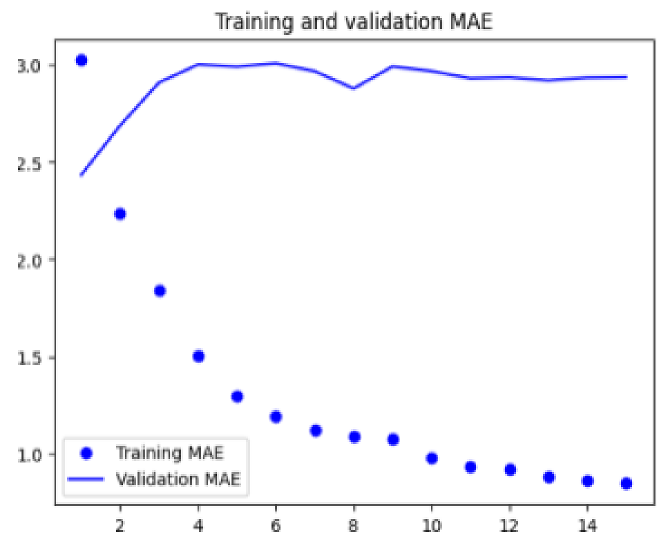
  - The test dataset generated an MAE outcome of 3.11.
  - In this model, I merged 1D convolutional layers (1D convnets) with recurrent neural networks (RNN) layers, particularly LSTM. This amalgamation of layers is likely beneficial in capturing both spatial and temporal patterns within the data, resulting in a comparable MAE to the stacked LSTM approach.



- **<u>Adjusting the number of units in each recurrent layer</u>**:

- The test dataset produced an MAE outcome of 3.08.

- In this model, I attained the lowest MAE among the test models by adjusting the number of units in each recurrent layer within a stacked LSTM configuration. This implies that optimizing the architecture, particularly the number of units, can positively influence model performance.



CONCLUSION

In summary, the outcomes underscore the efficacy of employing sophisticated methodologies like LSTM layering, integrating 1D convnets with RNN, and optimizing the model's structure to enhance the precision of time-series models. Specifically, adjusting the unit counts in each recurrent layer yields the most notable enhancement, underscoring the critical role of model architecture in crafting efficient models for time-series data. Collectively, models incorporating LSTM layers and amalgamations of convolutional and recurrent layers surpass the rudimentary machine learning instances, emphasizing the significance of utilizing advanced architectures for sequential data analysis.