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 the realm of analyzing time series data, a comparative study was conducted involving a diverse array of models, starting from baseline methodologies grounded in intuitive approaches to sophisticated neural network structures. The inception model, based on rudimentary common-sense techniques, established a preliminary benchmark with a Mean Absolute Error (MAE) of **2.62**. Following this, an attempt with a basic machine learning construct featuring a dense layer slightly elevated the **MAE to 2.70**, primarily due to its limitation in preserving the temporal sequence of data. The exploration continued with a convolutional model, which failed to adapt to the unique demands of time series data, as it uniformly processed data segments, thus disrupting the sequential integrity post-pooling.

The study pivoted towards Recurrent Neural Networks (RNNs), celebrated for their adeptness at integrating historical data points into current evaluations, uncovering inherent sequential patterns and dependencies. Despite the theoretical advantages of Simple RNNs, they displayed underwhelming performance across all tested models, attributed to the pervasive "vanishing gradient problem," which compromises their training efficacy. This predicament prompted the exploration of more sophisticated RNN variants, such as Long Short-Term Memory (LSTM) and

Gated Recurrent Unit (GRU) models, with GRU models emerging as the most effective in encapsulating long-range sequential dependencies in a computationally efficient manner.

LSTM architectures were extensively tested, with variations in the number of units and the incorporation of recurrent dropout and bidirectional techniques, to mitigate overfitting and enhance model robustness. These LSTM configurations yielded results surpassing the baseline model, with minor variations in MAE scores among them.

An innovative endeavor to merge a 1D convolutional model with an RNN was also undertaken, resulting in a **higher MAE of 3.79**, indicative of the convolutional approach's inadequacy in preserving sequential data integrity. The comprehensive analysis underscores the inefficacy of simple RNNs for time series analysis due to the vanishing gradient dilemma and their inability to effectively process long-term dependencies. In contrast, advanced RNN frameworks like LSTM and GRU are recommended for their superior capability in overcoming these challenges. While LSTMs are widely acknowledged for their prowess in time series analysis, the findings advocate for a preference towards GRU models for their efficiency and effectiveness. To refine GRU model performance, it is advisable to fine-tune hyperparameters such as stacked layer units, dropout rates, and bidirectional data processing strategies. Moreover, the study advises against combining convolutional methods with RNNs for time series analysis, due to the convolutional models' propensity to disrupt the sequential data flow.

CONCLUSION

The investigation into various models for time series data analysis revealed that advanced RNN architectures, specifically GRU, outperform others, including the baseline common-sense approach, dense layer models, convolutional models, and even LSTMs in some cases. GRU models demonstrated a superior capability in capturing long-range dependencies efficiently, which is crucial for time series data analysis. On the other hand, simple RNNs were found to be impractical for real-world applications due to the vanishing gradient problem. While LSTM models also showed promising results, especially with configurations to combat overfitting and enhance data presentation, GRUs were noted for their computational efficiency and equally effective performance. The attempted combination of 1D convolution with RNNs did not yield improved results, suggesting a misalignment in their operational strengths regarding time series analysis. Based on these findings, it is recommended to prioritize advanced RNN architectures like GRU for effective time series data analysis, while also considering LSTM for specific needs. Adjustments in model configurations, such as units in stacked layers, dropout rates, and bidirectional processing, can further optimize performance.