Assignment 3: Time-Series Data

GROUP 8

This report discusses using Recurrent Neural Networks (RNNs) for time-series data analysis, particularly in the context of weather forecasting. The objective is to investigate different strategies for enhancing the effectiveness of RNN models in predicting weather patterns. The techniques involve modifying the structure of the RNN model, testing various types of recurrent layers like LSTM and GRU, and integrating 1D convolutions with RNN layers. The report details how these methods were put into practice, assesses their performance using validation datasets, and showcases the top-performing models validated on the test set. Ultimately, the report aims to demonstrate the capability of RNNs in managing time-series data and to highlight methods for enhancing their accuracy in weather forecasting.

An overview of time-series data models:

Model	Dense Units	Dropout	Loss	Test MAE
Basic Machine Learning Model	16	No	11.4217	2.64
1D Convolution model	16	No	14.8259	3.04

RNN Model:

RNN Model	Dense Units	Dropout	Loss	Test MAE
Simple RNN Model	16	No	151.2728	9.92
Stacked Simple RNN Model	16	No	151.1893	9.92
Stacked Simple RNN Model	32	No	151.0980	9.90
Stacked Simple RNN Model	64	No	151.1381	9.91

Gated Recurrent Unit:

GRU	Dense Units	Dropout	Loss	Test MAE
Simple GRU	16	No	9.9697	2.49

Long Short-Term Memory:

LSTM	Dense Units	Dropout	Loss	Test MAE
LSTM-Simple	16	No	10.9701	2.58
LSTM - Dropout Regularization	16	Yes	10.557	2.55
LSTM - Stacked setup with 16 units	16	No	10.4460	2.51
LSTM - Stacked setup with 32 units	32	No	11.5477	2.66
LSTM - Stacked setup with 8 units	8	No	10.4114	2.52
LSTM - dropout-regularized, stacked model	8	Yes	11.1375	2.59
Bidirectional LSTM	16	No	10.5141	2.56

Combination:

Combination	Dense Units	Dropout	Loss	Test MAE
1D Convnets and LSTM together	16	No	23.1759	3.87

Using RNNs with Time-Series Data:

- The results show that the basic RNN and stacked Simple RNN models perform poorly in terms of Mean Absolute Error (MAE) on the test set, showing much larger MAE values than other models. This suggests that basic RNNs might not be the best choice for this time-series forecasting task.

- However, GRU and LSTM models exhibit superior performance. In particular, the test Mean Absolute Error (MAE) values for the Simple GRU and Bidirectional LSTM models are the lowest, indicating that they are the best at capturing the temporal patterns found in the data.
- While they are not the best, LSTM models with various configurations (dropout regularization, layered setup with variable units) also exhibit acceptable performance.

Enhancing the Network's Time-Series Data Performance:

- Some LSTM models use dropout regularization to avoid overfitting. Nevertheless, there is no guarantee that LSTM models with dropout regularization would perform better than those without.
- Better performance is not always achieved by stacking more LSTM units in the configuration. Bidirectional LSTM, which gathers data from both previous and future time steps, has promising outcomes in comparison to unidirectional LSTM models.
- For example, the LSTM stacking arrangement with 32 units has a slightly higher test MAE compared to the setup with 16 units.

Various Deep Learning Layers Applied to Time-Series Data:

- Performance is lower when 1D Convolutional Neural Networks (CNNs) and LSTM are combined than when LSTM or GRU models are used alone. This implies that CNNs might not be as successful in extracting pertinent characteristics from the time-series data for this task.

Recommendations:

- Pay attention to models like Simple GRU and Bidirectional LSTM that exhibit strong performance on the test set.
- Try out several architectures and hyperparameters to determine which model performs the best for this dataset.
- To enhance the model's capacity to identify pertinent patterns in the data, think about adding new features or developing new features. It is important to regularly assess models on validation and test sets to make sure that gains in performance extend far beyond the training set.

- Examine more deep learning strategies designed for time-series forecasting, such as hybrid models that combine deep learning and conventional statistical techniques, or attention mechanisms.