**Assignment 3: Time-Series Data**

**Models Evaluation:**

* A common-sense non-machine-learning baseline method calculates MAE in both a validation dataset and a test dataset. The model gave test MAE: 2.62 and validation MAE: 2.44 it means the predictions made by this common-sense method are off by 2.44 on the validation dataset and 2.62 on the test dataset.
* We trained Densely connected neural network model, which consists of a flattening layer, a dense layer with 16 units, and a ReLU activation function. RMSprop as the optimizer and Mean Squared Error as the loss function were used to train the model across ten epochs. MAE was tracked as a performance indicator. After 10 epochs, the model produced training validation MSE 11.90 and MAE of 2.72, indicating that the model can fit the training data and generalize to validation data. Following training, the model was evaluated on a test dataset which gave test MAE of 2.68. The evaluation on the test dataset suggested a competitive level of performance with a test MAE of 2.68.
* we have also evaluated the performance of a convolutional neural network. Convolutional layers with max-pooling and a global average pooling layer were used in the model's design, followed by a single-unit dense output layer for regression. Each of the three convolutional layers in this model has 8 units, while the final dense output layer has 1 unit. On the training dataset the model was trained for 10 epochs. The mean absolute error of the test was around 3.10. The model exhibits its ability to learn and adapt to data by reducing MAE of 2.10 on the validation dataset. In comparison to the previous densely connected model, this CNN model achieves a slightly higher test MAE of approximately 3.10. When compared to densely connected model, it has a lower test MAE of 2.68 which is typically indicative of better prediction accuracy.
* The first model is made up of two SimpleRNN layers of 16 units each, both of which are programmed to return sequences. The optimizer used is RMSprop, and the loss function is Mean Squared Error. The training procedure consists of ten epochs. The stacked SimpleRNN model's test MAE is 2.58 suggesting that its predictions on the test data were off by 2.58 units on the same scale as the target variable on average.
* The second model has several SimpleRNN layers. It consists of three SimpleRNN layers with increasing unit counts 30, 62, and 90. The model was trained for ten epochs on the train dataset before being tested on the validation dataset. The stacked SimpleRNN model's test MAE is around 2.61
* The third model has three layers with different unit sizes 64, 58, and 49 units and a final output layer with a SimpleRNN with 16 units. During training it went through 10 epochs of RMSprop optimization to reduce Mean Squared Error and was assessed using Mean Absolute Error. The model's performance was evaluated using its test MAE, which is 2.65 units.
* Among the three provided SimpleRNN models the best performing model in terms of Test MAE is the first one with two SimpleRNN layers of 16 units each which has a test MAE of 2.58. The other two models, which have more complex architectures with varying numbers of units. The best performing model is the one that achieves the lowest test MAE, as it indicates better generalization to unseen data. In this case, the first model with two SimpleRNN layers is the best choice due to its slightly lower test MAE which indicates a decent trade-off between model complexity and prediction performance.
* The model includes an LSTM layer with 32 units and recurrent dropout set to 0.25. It is trained over a period of ten epochs, which has RMSprop as the optimizer and MSE as the loss function. The LSTM model's test MAE is similarly 2.53.

Among all the models the LSTM model is the best performing model with the lowest Test MAE of 2.53. This indicates that, on average, its predictions on the test data were off by 2.53 units in the same scale as the target variable. This model outperforms the others in terms of predicted accuracy. The lower Test MAE of the LSTM model shows that it generalizes effectively to new data and able to efficiently capture and model the underlying patterns in data,leading in more accurate predictions on new, previously unseen data points. The LSTM model's architecture and hyperparameters such as the number of units and usage of recurrent dropout were fine-tuned to deliver the optimum performance for this task.