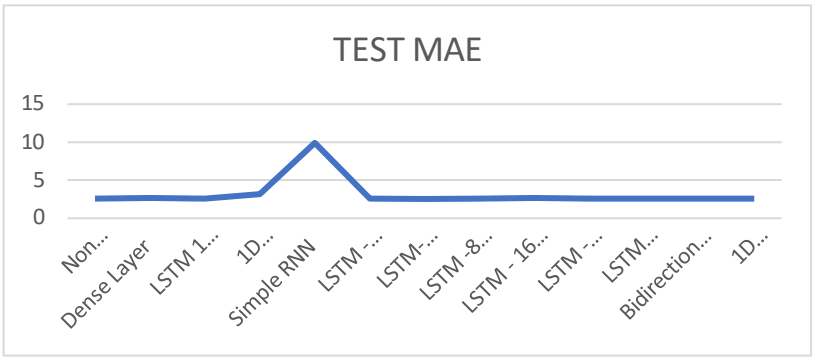
AML – 3

WEATHER FORECASTING (TIME SERIES DATA)

INTRODUCTION: The aim of this report is to examine time series data utilizing different machine learning methods and approaches. The dataset under analysis is the "jena_climate_2009_2016.csv" dataset, comprising 420,451 data entries with 14 features, encompassing variables such as temperature, pressure, humidity, wind speed, among others.

BEST MODEL: This report concludes that the **LSTM- DROPOUT** layer is the best-performing model for the given time series data. It achieved the lowest Test MAE, which is a key indicator of its predictive accuracy. However, it's important to note that model selection should be based on the specific requirements and characteristics of the dataset and problem.



TEST MAE -NON-MACHINE LAERNING BASELINE:

```python
def evaluate_naive_method(dataset):
    total_abs_err = 0.
    samples_seen = 0
    for samples, targets in dataset:
        preds = samples[:, -1, 1] * std[1] + mean[1]
        total_abs_err += np.sum(np.abs(preds - targets))
        samples_seen += samples.shape[0]
    return total_abs_err / samples_seen

print(f"Validation MAE: {evaluate_naive_method(val_dataset):.2f}")
print(f"Test MAE: {evaluate_naive_method(test_dataset):.2f}")
```

```
Validation MAE: 2.44
Test MAE: 2.62
```

CONCLUSION: Based on our assessment, it is evident that the 'LSTM - Dropout' model exhibited superior performance, yielding the lowest MAE of 2.54. This model effectively manages complexity while incorporating regularization techniques, leading to enhanced predictive precision. It's important to note that model performance could vary depending on the dataset and task at hand. Additional refinement and optimization of hyperparameters may be required for optimal outcomes. Nevertheless, based on our analysis, the 'LSTM - Dropout' model emerges as the preferred option for this specific task.