GROUP WORK PROJECT 1
GROUP NUMBER: 3356

MScFE 622: Stochastic Modeling

FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
Ncube Trymore	South Africa	trymorencube@yahoo.com	
Sanchit Jain	India	jsanchit1103@gmail.com	
Trin Chavalittumrong	Thailand	mrtrin@gmail.com	

Statement of integrity: By typing the names of all group members in the text boxes below, you confirm that the assignment submitted is original work produced by the group (excluding any non-contributing members identified with an "X" above).	
Team member 1	**Ncube Trymore**
Team member 2	**Sanchit Jain**
Team member 3	**Trin Chavalittumrong**

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed.
Note: You may be required to provide proof of your outreach to non-contributing members upon request.

REPORT

1. STEP_1(d)

1.1 Original Time Series:

Stationarity Test (ADF Test):

The ADF statistic is 0.08561421816059317, and the p-value is 0.9650650736640431. Since the p-value is greater than the significance level of 0.05, we fail to reject the null hypothesis of non-stationarity for the original time series.

1.2 First Difference of Time Series:

Stationarity Test (ADF Test):

ADF Statistic (Differenced Series): -14.671503268207804

p-value (Differenced Series): 3.2612279131691476e-27

Analysis: The first difference is taken to remove the trend component in the time series. The differenced series shows stationary behavior as evidenced by the significant ADF statistic and a very low p-value.

1.3 Fractionally Differenced Time Series:

Stationarity Test (ADF Test):

ADF Statistic (Fractionally differenced Series): -20.732432678046003

p-value (Fractionally Differenced Series): 0.0

Analysis: Fractional differencing is applied to the time series to achieve stationarity. The resulting series exhibits stationarity, as indicated by the highly significant ADF statistic and a p-value of 0.0.

Summary

In summary, the first difference and the fractionally differenced series both show stationarity, while the original time series does not. Fractional differencing appears to provide a more effective transformation in achieving stationarity compared to the simple first difference.

2. STEP_2(d)

Results

2.1.1 MLP for predicting the level of the time series (Model 1):

MSE - Levels Prediction: 106.83220330817257

2.1.2 MLP for predicting the stationary time series (Model 2):

MSE - Stationary Prediction: 30.268425850075268

2.1.3 MLP for predicting the fractionally differenced time series (Model 3):

MSE - Fractionally differenced Prediction: 31.248156390896945

2.2 Analysis

Based on the MSE values, we can see that Model 2, which predicts the stationary time series, has the lowest MSE, indicating better performance compared to the other models. Model 1, which predicts the level of the time series, has the highest MSE, indicating relatively poorer performance.

2.3 Possible reasons for the performance differences:

2.3.1 Stationarity

Model 2 works with the differenced (stationary) time series, which can simplify the underlying patterns and make the prediction task more manageable. By removing the trend and seasonality, the model can focus on capturing the remaining patterns, leading to better predictions.

2.3.2 Fractional differencing

Model 3 also operates on a differenced time series but uses fractional differencing, which allows capturing long-term dependencies and non-linear patterns more effectively than simple differencing. However, in this particular scenario, Model 3 performs slightly worse than Model 2, possibly because the fractional differencing order (1.2) may not be the optimal choice for this specific dataset.

2.3.3 Model architecture and input representation

The MLP models in all three cases have similar architectures with two hidden layers, but the input representations differ. Model 1 takes the lagged values of the original time series,

while Model 2 and Model 3 use the differenced versions. The differences in input representation can affect how well the models capture and learn the underlying patterns in the data, leading to variations in prediction performance.

2.3.4 Training data availability

The training data used for each model might vary due to the nature of the transformations applied. Model 2 and Model 3 have fewer data points compared to Model 1 due to the differencing operations. The availability of more training data can positively impact the model's ability to learn and generalize, which could contribute to better predictions.

3. STEP_3(d)

Results

The MSE values obtained for each model are as follows:

3.1.1 Levels Prediction

Levels Prediction: 23127.907630225563

3.1.2 Stationery Prediction

Stationary Prediction: 34.54555303646466

3.1.3 Fractionally Differenced Prediction

Fractionally differenced Prediction: 32.80798607421771

3.2 Analysis

Based on the MSE values, the Stationary Prediction and Fractionally differenced Prediction models perform better than the Levels Prediction model. Here are some arguments explaining why:

3.3 Possible reasons for the performance differences:

3.3.1 Levels Prediction:

- The Levels Prediction model uses a CNN architecture with GAF representation, which transforms the input time series data into Gramian Angular Field (GAF) images.
- However, the high MSE value of 23127.91 suggests that this model may not capture the underlying patterns and dynamics of the time series accurately.
- It is possible that the CNN architecture and GAF representation might not be well-suited for capturing the level predictions in this particular scenario.

3.3.2 Stationary Prediction:

- The Stationary Prediction model uses an MLP architecture with GAF representation for predicting the stationary time series.
- The MSE value of 34.55 indicates that this model performs better than the Levels Prediction model.
- The MLP architecture might be more suitable for capturing the stationary patterns in the data, leading to improved predictions.

3.3.4 Fractionally differenced Prediction:

- The Fractionally differenced Prediction model also uses a CNN architecture with GAF representation, similar to the Levels Prediction model.
- However, it focuses on predicting the fractionally differenced time series, which captures the long-range dependencies in the data.
- The MSE value of 32.81 suggests that this model performs slightly better than the Stationary Prediction model.
- The long-range dependencies captured by the fractionally differenced series might benefit from the CNN architecture and GAF representation, leading to improved predictions.

In Summary, the Stationary Prediction and Fractionally differenced Prediction models outperform the Levels Prediction model in terms of MSE. The choice of model architecture and representation technique plays a crucial role in capturing the patterns and dynamics of the time series accurately. Different types of data transformations, such as GAF representation, can have varying effects on the performance of different models.

STEP_4

The results obtained between the CNN and MLP architectures in the given scenario show some differences in their prediction performance. The CNN architecture is used for both Levels Prediction and Fractionally differenced Prediction, while the MLP architecture is used for Stationary Prediction**.**

4.1.1 CNN Architecture:

- The CNN architecture is well-suited for processing grid-like data, such as images or GAF representations.
- In the Levels Prediction task, where the goal is to predict the level of the time series, the CNN architecture may not perform as well as expected. The high MSE value of 23127.91 suggests that the CNN might not effectively capture the underlying patterns and dynamics of the time series.
- In the Fractionally differenced Prediction task, the CNN architecture performs slightly better, with an MSE value of 32.81. This could be attributed to the CNN's ability to capture spatial patterns in the GAF representation, which may be helpful in predicting the fractionally differenced time series with long-range dependencies.

4.1.2 MLP Architecture:

- The MLP architecture is a traditional feedforward neural network that is well-suited for learning complex relationships in data.
- In the Stationary Prediction task, where the goal is to predict the stationary time series, the MLP architecture performs relatively better, with an MSE value of 34.55.
- MLPs are known for their ability to capture non-linear relationships, which might be more suitable for capturing the stationary patterns in the data compared to the CNN architecture.

4.2 Educated guesses regarding the differences in results:

4.2.1 Levels Prediction:

- The high MSE value obtained for the Levels Prediction using the CNN architecture suggests that the level of the time series may not be well-captured by the GAF representation and the CNN's spatial pattern detection capabilities.
- The time series might have complex non-linear relationships that cannot be effectively captured by the CNN architecture alone.

4.2.2 Stationary Prediction:

- The MLP architecture, with its ability to learn complex non-linear relationships, may be better suited for capturing the stationary patterns in the data.
- The stationary time series may exhibit more localized patterns that can be effectively learned by the MLP architecture.

4.2.3 Fractionally-Differenced Prediction:

- The CNN architecture, with its ability to capture spatial patterns in the GAF representation, performs slightly better in predicting the fractionally differenced time series.
- The fractionally differenced series captures long-range dependencies, and the spatial pattern detection capability of CNNs might help in capturing these dependencies effectively.