



Submission Number: 3

Group Number: 08

Group Members:

Full Legal Name	Location (Country)	E-Mail Address	Non-Contributing Member (X)
Monika Singh	India	ms19626@gmail.com	
Ananta Narayana Satapathy	India	ansatapathy@yahoo.com	
Harshil Sumra	India	harshilsumra1997@gmail.com	

Statement of integrity: By typing the names of all group members in the text box below, you confirm that the assignment submitted is original work produced by the group (*excluding any non-contributing members identified with an "X" above*).

Monika Singh
Ananta Narayana Satapathy
Harshil Sumra

Answer 9 and 10

Intuitively, we can see that the most ideal model for handling such data should have characteristics of both time series and multivariate characteristics. Out of above three models, only VARMA model fulfils these characteristics. Furthermore, interpretability of VARMA model is much easier and higher than neural networks.

Final model characteristics of each neural network model:

a. Classification Neural Networks

Optimal model: - Model 2 with regularised weights [34,10,10,1]

Cost function: - Cross-entropy

Output activation function: - Sigmoid

Hidden layer activation function: - ReLU

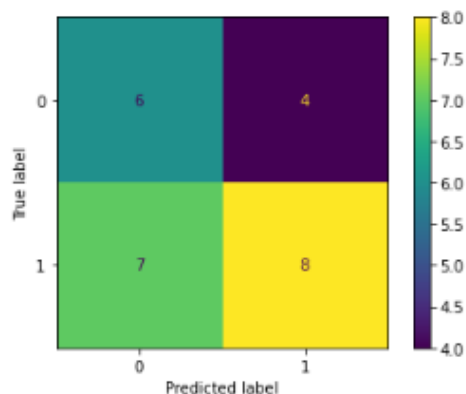
Learning rate: - 0.006

#iterations: - 150000

Performance Measure: - Test data prediction accuracy, precision, recall and f1 score.

```
In [99]: #For classification
y_pred=[1., 1., 0., 0., 0., 1., 0., 0., 0., 1., 0., 1., 1., 1., 1., 0.,
        0., 0., 0., 1., 1., 1., 0., 0., 1.]
y_actual=y_test_clsnn.to_list()
ConfusionMatrixDisplay.from_predictions(y_actual,y_pred)
print('Accuracy Score : {}'.format(accuracy_score(y_actual,y_pred)))
print('Precision Score : {}'.format(precision_score(y_actual,y_pred)))
print('Recall Score : {}'.format(recall_score(y_actual,y_pred)))
print('F1 Score : {}'.format(f1_score(y_actual,y_pred)))

Accuracy Score : 0.56
Precision Score : 0.6666666666666666
Recall Score : 0.5333333333333333
F1 Score : 0.5925925925925926
```



b. Regression Neural Networks

Optimal model: - [34,5,5,1]

Cost function: - Root Mean Square Error

Output activation function: - linear

Hidden layer activation function: - ReLU

Learning rate: - 0.04

#iterations: - 300000

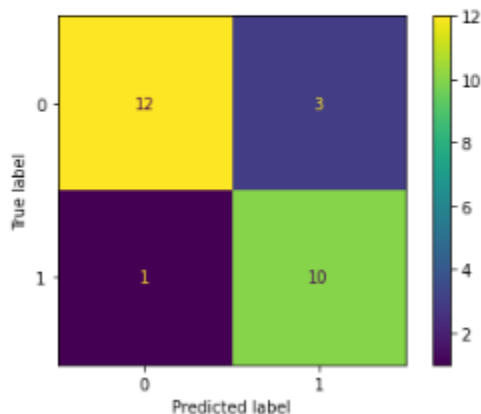
Performance Measure: - Test data prediction RMSE, in order to make comparison with classification neural network we convert the predicted return data into 1's and 0's based on previous weeks returns respectively.

RMSE= 0.024937372696291305

```

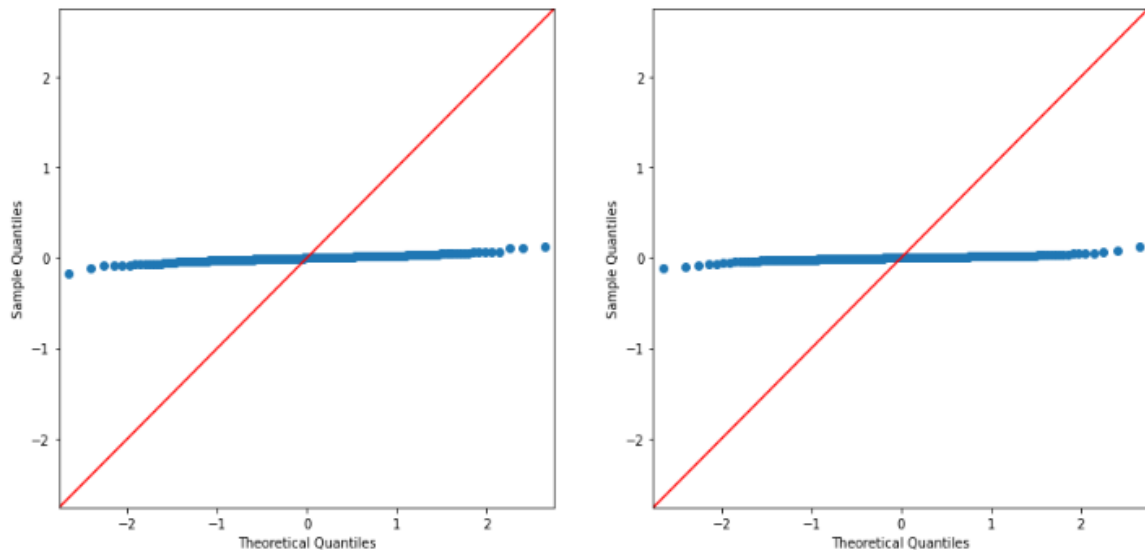
Accuracy Score : 0.8461538461538461
Precision Score : 0.7692307692307693
Recall Score : 0.9090909090909091
F1 Score : 0.8333333333333333

```



Therefore, Regression Neural Network gives us much better prediction based on all classification performance parameters as compared to classification neural networks. But still the most ideal of these models is VARMA due to its model definition, ease of implementation and its interpretability.

Fundamentally, we observe that variances and co-variances in our VARMA model have a p-value of 0, indicating a certain degree of stationarity characteristic. This fact is also shown through the q-q plots.

**Answer.11.**

Monika Singh= Parts 1 to 5

Ananta Narayana Satapathy=Parts 6 and 7

Harshil Sumra= Part 8 + compilation of code

Rest of the parts (9-13) were worked on together by all three through rigorous discussion about interpretation of the code and its output.

Answer.12.**Technical Report****Problem Statement**

We try to model 'LUX' (Luxembourg stock market index) index weekly increment/decrement based on itself and other MSCI indexes.

Understanding the data

a. Skewness (1) test

We observed that almost all the variables are negatively skewed with negative mean-median difference except for Brazil, Mexico and South Africa where skewness and mean-median difference have opposite signs. It is due to impact of outliers in both positive and negative directions and the values in the opposite direction of skewness are more extreme than those in the same direction of skewness.

b. Kurtosis (2) test

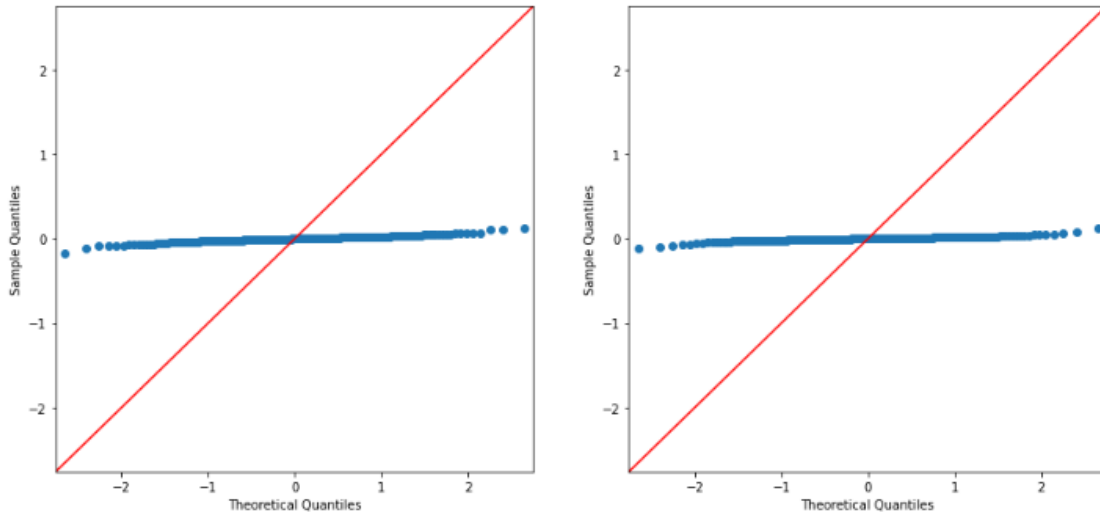
Our response variable 'LUX's return series is slightly leptokurtic, that is, it has slightly fatter tails than normal distribution

c. Structural breaks (3)

We ran Markov Regression with two regimes to identify 1 structural break in the time series to discover that p-values for both regimes is 0 indicating stationarity of the time series.

d. Distribution

We utilized distribution plots (with kde = True) and (quantile-quantile) Q.Q plots to observe 'LUXXX' and 'MSCI INDIA' series to discover that our data followed a normal distribution, indicating that both the series have common distribution characteristics.



Models

a. VARMA

Here, we use our Q-Q plot conclusion along with feature importance from LASSO model application in GWP 1 to find exogenous variable 'MSCI ITALY' for VARMA model. The order of our VARMA model is (2,0,2). We found the variables to have distinctive linear relationship along with the fact that p-values for variances and covariance are zero which indicates of stationarity of series of market data.

b. Classification neural network

We followed the step by step design process to reach our most optimal model of classification. Here, our dependent variable is increment/decrement of returns when compared with previous weeks returns.

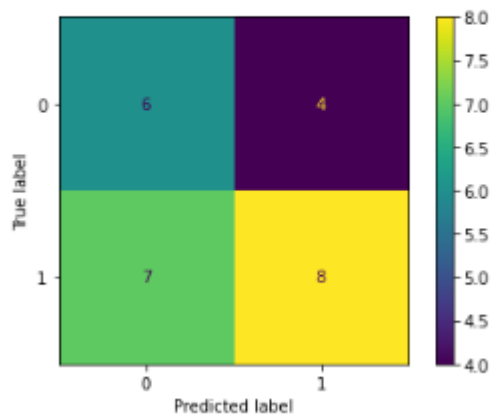
Model development process: -

1. Hyper-parameter tuning for learning rate and number of iterations to get intuition for the preferred value of these parameters in case of given data. Learning rate = 0.006 and number of iterations=150000
2. Model-1 has the structure [34,5,5,1], general model to set benchmark for all other models

3. Model 2 has the structure [34,10,10,1], Cross-Validation scores are:
Accuracy: 0.7352941176470589
Precision: 0.8421052631578947
Recall: 0.7272727272727273
4. Model 3 has the structure [34,10,10,10,1], Cross-Validation scores are:
Accuracy: 0.7142857142857143
Precision: 0.7894736842105263
Recall: 0.7142857142857143
5. Model 4 is model 2 with regularized weights, Cross-Validation scores are:
Accuracy: 0.7575757575757576
Precision: 0.8947368421052632
Recall: 0.7391304347826086
6. Model 5 is model 3 with regularized weights, Cross-Validation scores are:
Accuracy: 0.6578947368421053
Precision: 0.7058823529411765
Recall: 0.6

Based on above scores we find that Model-4 as our most optimal classification neural network with following confusion matrix on test data.

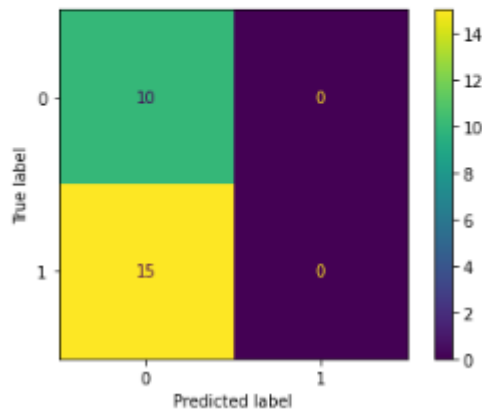
Accuracy Score : 0.56
Precision Score : 0.6666666666666666
Recall Score : 0.5333333333333333
F1 Score : 0.5925925925925926



Which is better than naïve approach output shown below:

```
Accuracy Score : 0.4
Precision Score : 0.0
Recall Score : 0.0
F1 Score : 0.0
```

```
c:\users\harshil sumra\appdata\local\programs
nedMetricWarning: Precision is ill-defined an
ontrol this behavior.
_warn_prf(average, modifier, msg_start, len
```



c. Regression neural network

Here, our dependent variable is weekly returns of 'LUX' and independent variables are weekly returns of other MSCI indexes.

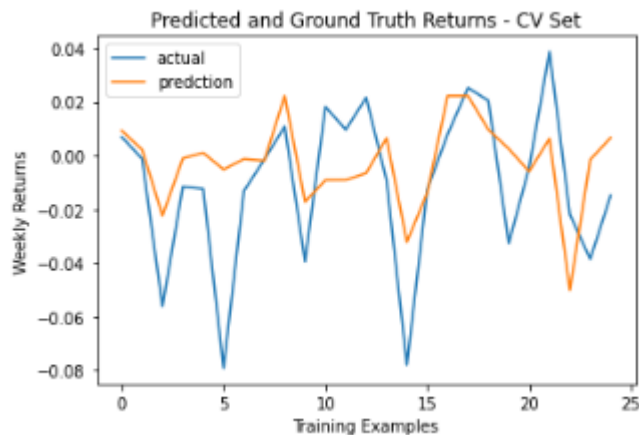
Model development process: -

1. Hyper-parameter tuning for learning rate and number of iterations to get intuition for the preferred value of these parameters in case of given data. Learning rate = 0.04 and number of iterations=300000
2. Model-1 has the structure [34,5,5,1], Cross-validation RMSE= 0.02678

Plot for cross-validation data [predictions vs actual]:

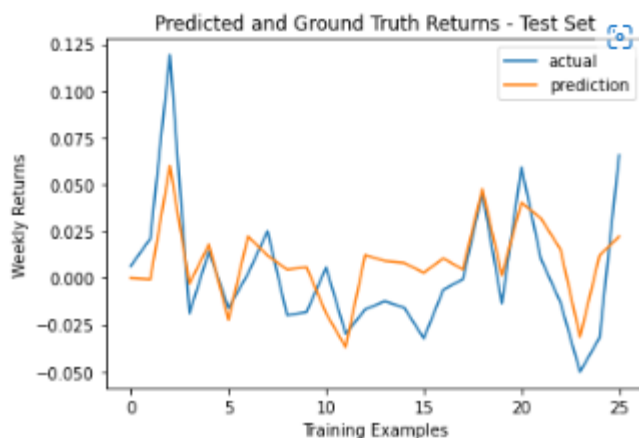


3. Model 2 has structure [34,5,5,5,1]. Cross-validation RMSE= 0.02619
Plot for cross-validation data [predictions vs actual]:



We see that model 2 has lower RMSE, so ideally, we should take it as our optimal model but we still prefer model 1 due to comparatively larger swings in predicted values as this is the only shortcoming of these regression models (plus both models RMSE is comparably close). Therefore, we choose model 1 as our optimal model. We don't try further model as that would only increase complexity and computation time while our graphs of model 1 and 2 already capture the directionality quite well which is the target of our problem statement.

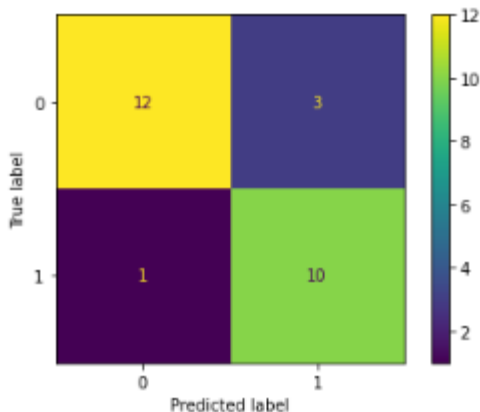
So, on test set for model 1, we get:
RMSE=0.02494



In order to compare it with the classification neural network, we transform our regression predictions and actual values to same classification data in order to assess directionality.

Output is as follows:

```
Accuracy Score : 0.8461538461538461
Precision Score : 0.7692307692307693
Recall Score : 0.9090909090909091
F1 Score : 0.8333333333333333
```



We find that Regression neural network is much better than classification neural network based on all performance parameters.

Conclusion

We proved that regression neural network is much preferred as compared to classification neural networks. The regression model is directionally accurate but is not able to catch up with the magnitude of extreme values. But if we aim to actually understand the relationship between different variables and their distributions then we prefer VARMA model, fit well with the data, due to its ease of implementation and interpretability for the stationary data (as inferred from Q-Q plots).

Answer.13.

Non-Technical Report

Sir/Ma'am,

We tried to predict increase or decrease in weekly returns of LUX index dependent on its own lagged values and various other MSCI indexes based on model requirement. We tried three different approaches which are VARMA, classification neural network and regression neural network. We performed regime analysis using Markov regression to find that p-values are 0 for both regimes, implying that the series are stationary. Among the various models we conclude the regression neural network is better than classification neural network as it has higher accuracy score. Regression neural networks is very promising when considering small swings in return values in terms of both magnitude and direction whereas if swings are large then good directional prediction is still happening but not magnitude. For precise interpretation of data and the involved relationship between variables, we found VARMA model much more effective and efficient.

References

1. <https://www.investopedia.com/terms/s/skewness.asp>
2. <https://www.investopedia.com/terms/k/kurtosis.asp>
3. <https://www.twosigma.com/articles/a-machine-learning-approach-to-regime-modeling/>