GROUP WORK PROJECT # 1
Group Number: 6888

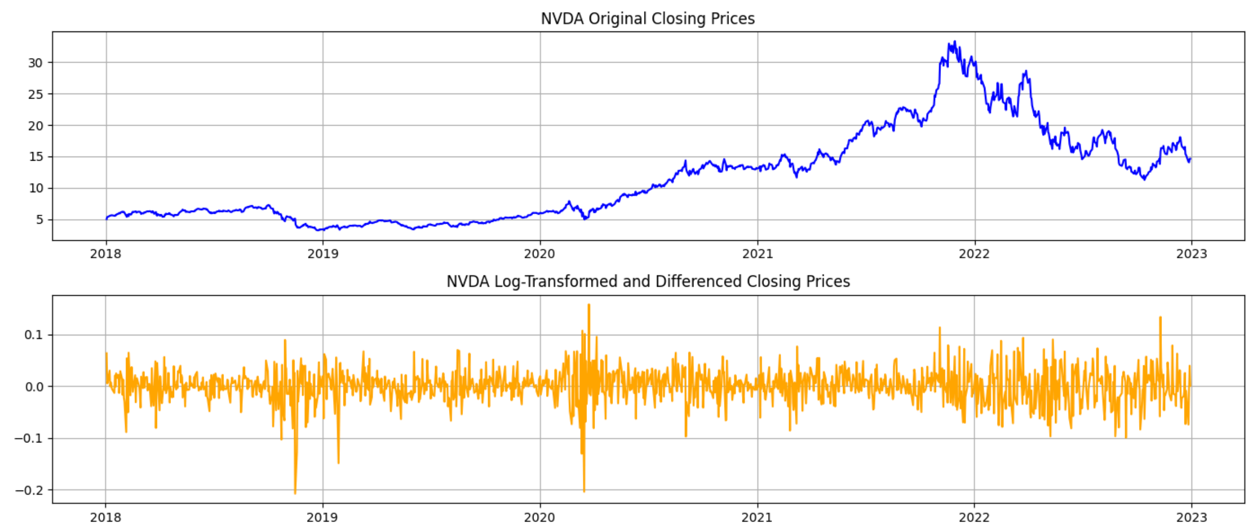
MScFE 642: Deep Learning for Finance

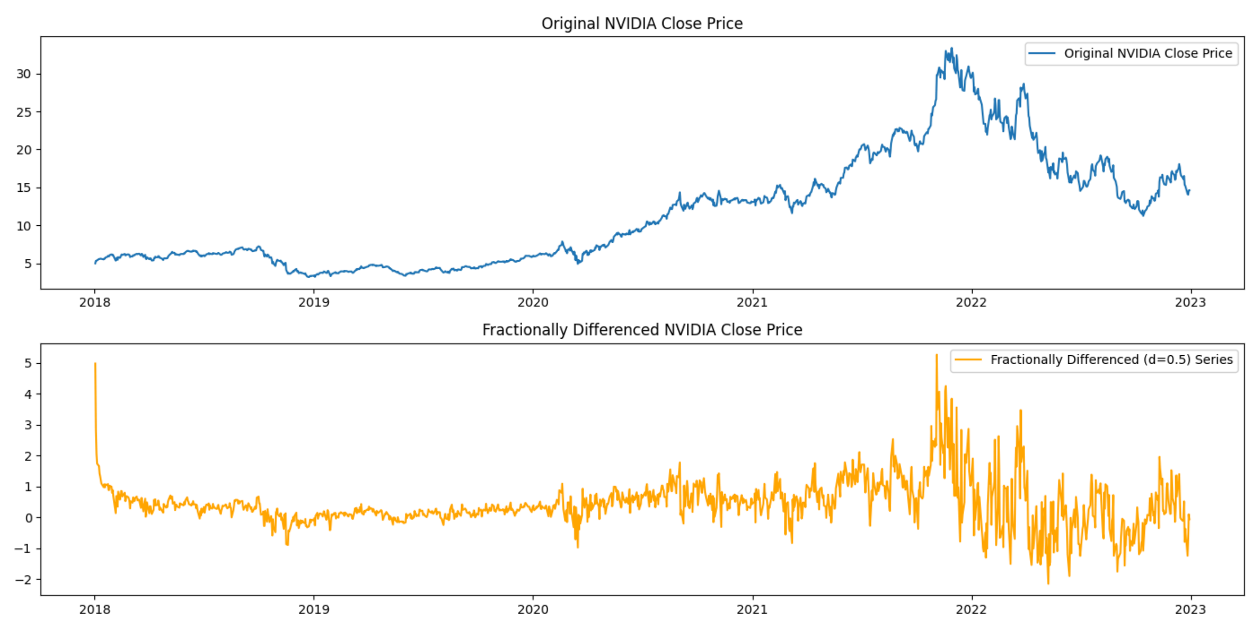
FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
Ebenezer Yeboah	Ghana	ebenezeryeboah46@gmail.com	
Jatin Rai	India	jatinrai198@gmail.com	

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	
Team member 1	**Ebenezer Yeboah**
Team member 2	**Jatin Rai**
Team member 3	

Use the box below to explain any attempts to reach out to a non-contributing member.
(N/A)

STEP 1





In the first representation which has to do with the original form of the times series, the price of NVIDIA showed an upward trend which explains that the stock price has been increasing over time. With the second representation which has to do deal with the differencing of the stationary times series, the statistical properties remain the same over time which explains the stationarity here but vital information is lost which may affect the analysis of long term trends. The third representation, which is the fractional differencing, aims to achieve stationarity, that is the statistical properties such as the mean and variance being constant overtime and retain much of the original information which will be necessary to make relevant decisions over time. This explains the variations in the graphical representation of the three.

STEP 2

The MSE which provides accuracy on the prediction gives us relevant analysis that the stationary times series has a higher accuracy than the normal times series and yes it is usually expected as the stationary times series contains trained data.

If in the attempt to remove noise using the stationary method coincidentally removes relevant data that it sees as noise, it may affect its prediction error and therefore making the normal times series better.

STEP 3

CNN Model - MSE: 26.012, MAE: 4.048, RMSE: 5.100, R^2: 0.549

MLP Model - MSE: 0.460, MAE: 0.560, RMSE: 0.678, R^2: -0.059

Performance of the CNN Model: MSE/MAE/RMSE: The higher values of these error metrics suggest that the CNN model has some difficulty in achieving precise predictions. However, compared to the MLP, the CNN still shows a relatively better fit. R^2 Value: An R^2 value of 0.538 indicates that the CNN model explains approximately 53.8% of the variance in the target variable, which is a moderately good fit for the data.

Performance of the MLP Model: MSE/MAE/RMSE: The MLP model achieves significantly lower error metrics compared to the CNN, which might initially suggest it performs better. However, this is misleading because... R^2 Value: The negative R^2 value (-0.04) indicates that the MLP model performs worse than a simple mean of the target variable. This means that the model is not capturing the underlying patterns in the data at all and is likely underfitting.

CNN Model: The CNN model performs better because it is designed to handle image-like data, which is what the GAF transformation produces. It can learn complex patterns and relationships within this transformed data, leading to better predictions despite higher error metrics compared to MLP.

MLP Model: The MLP model underperforms because it struggles with the complexity of the GAF-transformed data. It likely fails to capture the intricate patterns necessary for accurate predictions, leading to underfitting as reflected by the negative R^2 value.

STEP 4

Comparing CNN and MLP Architectures:

When we look at the results of the CNN and MLP models, the differences in their performance can be traced back to how each architecture processes data and the characteristics of the time series itself. Here's a breakdown of why the CNN outperformed the MLP in this case.

1. How Each Architecture Handles Data

CNN (Convolutional Neural Network):

Understanding Spatial Patterns: CNNs are specifically designed to identify patterns in spatial data, which makes them particularly effective for images. In this case, we transformed the time series into Gramian Angular Field (GAF) images, which capture temporal dependencies as spatial patterns. CNNs are naturally good at picking up on these kinds of patterns because they use layers that focus on different aspects of the data, such as edges or textures, before combining them into a more complete understanding.

Performance: This explains why the CNN was able to achieve a positive R^2 value (0.538), indicating it could explain a decent portion of the variability in the time series.

Complex Relationships: The GAF transformation likely introduced complex relationships that the CNN could capture well due to its ability to learn from multiple layers of data. Each layer in the CNN builds upon the previous one, allowing the model to understand both simple and complex patterns in the GAF images.

MLP (Multi-Layer Perceptron):

Flat Data Processing: Unlike CNNs, MLPs work with data that has been flattened into a single dimension. This means that when the GAF images are turned into vectors, any spatial relationships are lost, making it harder for the MLP to detect important patterns. This is a key reason why the MLP struggled to perform well, as indicated by its negative R^2 value (-0.04), which suggests that it wasn't able to capture any meaningful relationship in the data.

Underfitting: The MLP's poor performance suggests it might have underfit the data, possibly because it couldn't leverage the spatial information that was crucial in the GAF images.

Sensitivity to Preprocessing: The MLP might also be at a disadvantage because of the preprocessing applied to the time series. By applying a log transformation to make the series stationary, we might have simplified the data too much, stripping away some of the complexity that the model could learn from. MLPs typically do well when the data is linear and straightforward, but in this case, the transformation might have removed too much information.

2. The Nature of the Time Series

Trends and Seasonality:

If the original time series had strong seasonal or trending patterns, these would be encoded into the GAF images. The CNN, designed to recognize patterns in spatial data, would be able to pick up on these trends effectively. On the other hand, the MLP's approach of flattening these images might have caused it to miss out on these crucial patterns.

Stationarity and Noise:

The log transformation aimed to stabilize the variance and make the series stationary. However, while this helps in many cases, it might have oversimplified the series for the MLP, preventing it from learning anything meaningful. The CNN, with its ability to process complex patterns in the GAF images, would be less affected by such a transformation.

3. Why the CNN Performed Better

Effective Use of GAF: The GAF transformation turns the time series into an image-like format, which is exactly what CNNs are built to handle. This gave the CNN a significant advantage, allowing it to capture the intricate patterns within the data and produce better predictions.

MLP's Limitations: The MLP, which works well with simpler, more linear data, struggled with the complexity introduced by the GAF transformation. Flattening the GAF images into vectors removed spatial context, making it difficult for the MLP to find meaningful patterns. The result was underfitting, as the model couldn't capture the underlying structure of the time series.

4. Conclusion

In conclusion, the CNN outperformed the MLP because it is better suited for handling the type of data produced by the GAF transformation. While the MLP could have performed well with more straightforward data, the complexity and spatial patterns inherent in the GAF images made it difficult for the MLP to learn effectively. The CNN's ability to process and interpret these patterns allowed it to produce more accurate predictions, highlighting its strengths in dealing with complex, image-like data.