The Genesis Model

Hatem Alghuti, Azaldin Assi, Anjana Chowdiah

Executive Summary

The Genesis Model:

Aims at projecting future prices of stocks based on the performance of staple stocks using machine learning models

Neural Network -> Linear Regression

The objective is to construct a model that effectively predicts selected share prices 24 hours in advance within a 3% error margin

Commodities

- Copper
- Nicker
- Zinc

Stocks

- Costco (COST)
- Proctor & Gamble (PG)

ETFs

- Consumer Staples Select Sector SPDR Fund (XLP)
- Vanguard Consumer Staples Index Fund ETF (VDC)

Data Collection, Cleanup & Exploration

Data was collected using an import function for the Commodity data from FRED. Alpaca API was used to get closing prices

The cleanup was done by addressing and removing NaN, homogenizing timeframe to include monthly closing prices, joining data along common columns.

Exploration of the data was done by running a correlation between individual commodity variables and stocks to get an idea of what type of correlation exists between variables

Approach

- Brainstorming the ideas of the team and narrow it to the best qualified one based on the requirements.
- Reorganizing the data to use in Neural Networks
- Develop model(s) to predict future stock prices using historical data

Model Used

Perceptron model - neural network unit

- Receives input data, weighs the information, and produces a clear output.
- Not viable for the data used because of size

Linear Regression Model -

- Establishes the linear relationship in between yesterday's commodity data and predicting tomorrow's stock data

Results

```
# Evaluate the model loss and accuracy metrics using the evaluate method and the test data
model_loss, model_accuracy = nn.evaluate(X_test_scaled,y_test,verbose=2)

# Display the model loss and accuracy results
print(f"Loss: {model_loss}, Accuracy: {model_accuracy}")
```

✓ 0.2s

```
1/1 - 0s - loss: 3046.9971 - accuracy: 0.0000e+00 - 150ms/epoch - 150ms/step
Loss: 3046.9970703125, Accuracy: 0.0
```

+ Code + Markdown

```
# The mean squared error
print("Mean squared error: %.2f" % mean_squared_error(y_test,y_pred))
# The coefficient of determination: 1 is perfect prediction
print("Coefficient of determination: %.2f" % r2_score(y_test,y_pred))
```

[41] ✓ 0.4s

```
Mean squared error: 769.20
Coefficient of determination: -1.19
```

Conclusion + Next Steps

- Split data against specific groups for a stronger correlation
- Can expand commodity list from metals to energy/agricultural sector to see its impact on different sectors of the stock market
-

Questions?