

Trading Strategy Using Linear Models (2020–2024)

1. Objective

To use historical weekly return and volatility data of NVIDIA stock to train linear classification models and implement trading strategies based on predicted labels (Green or Red weeks).

2. Dataset and Preprocessing

The dataset includes weekly features: mean_return, volatility, and manually assigned labels for the years 2020–2024. Labels were encoded as binary values (Green = 1, Red = 0). The data was split into training (2020–2022) and testing (2023–2024) sets.

3. Models Trained

Two logistic regression models were trained:

- Model A: Using only mean_return as input
- Model B: Using both mean_return and volatility as input

4. Accuracy on Testing Set (2023–2024)

- Model A Accuracy: 66.67%
- Model B Accuracy: 86.67%

5. Trading Strategy Results

Using predicted labels, a trading strategy was simulated over the test period (2023–2024):

- Invest \$100 on the first Green week of 2023
- Stay invested on consecutive Green weeks
- Sell and hold cash on Red weeks

Final Portfolio Values

- Model A: \$1302.60
- Model B: \$1390.12

6. Conclusion

Model B outperformed Model A in both classification accuracy and final portfolio value. The inclusion of volatility as a feature significantly improved the predictive power and trading performance. A rule-based strategy driven by predicted labels using linear models can outperform naïve

```
print("\nFinal Portfolio Value (Model A): $", round(f
print("Final Portfolio Value (Model B): $", round(f
```



```
Model A Accuracy: 0.6666666666666666
Model B Accuracy: 0.8666666666666667
```

```
Final Portfolio Value (Model A): $ 1302.6
Final Portfolio Value (Model B): $ 1390.12
```

e investment strategies.

Attachments

1. Processed_Stock_Data.csv
2. NVDA.csv
3. Python code file for training and evaluation