JPX Stock Exchange

A Time-Series Project

Predicting the Sharpe Ratio of Stocks

Jim Petoskey

Flatiron School

Capstone Project

Outline

- Business Opportunity
- Summary
- Data
- Methods
- Results
- Conclusions
- Further Study

Business Opportunity

- JPX Stock Exchange – Increase Trading
  - Increase trading to increase revenue
  - Provide institutional investors with machine learning models that predict daily stock direction.

Summary

- It is possible to predict the daily direction of stocks from the Sharpe Ratio.
- Machine Learning Models make it possible to accurately predict stock direction and accrue returns when stocks are bought and shorted at scale.
- Returns are generated by buying 200 and shorting 200 stocks on a daily basis.
- Sharpe Ratio is predicted on a given day, then:
    - Stocks are bought at the close of the following business day and sold at close the day following that.

Data

- 2000 stocks with 4.5 years of data
  - Data from additional stocks provided to improve modeling
- Financial Data Utilized
  - Trading Information, Income Statements, Market Sector, Issued Shares,
- New Information Produced – 12 Features
  - Improvised Sharpe Ratio, Daily Return, Weekly Moving Average, etc.

Methods

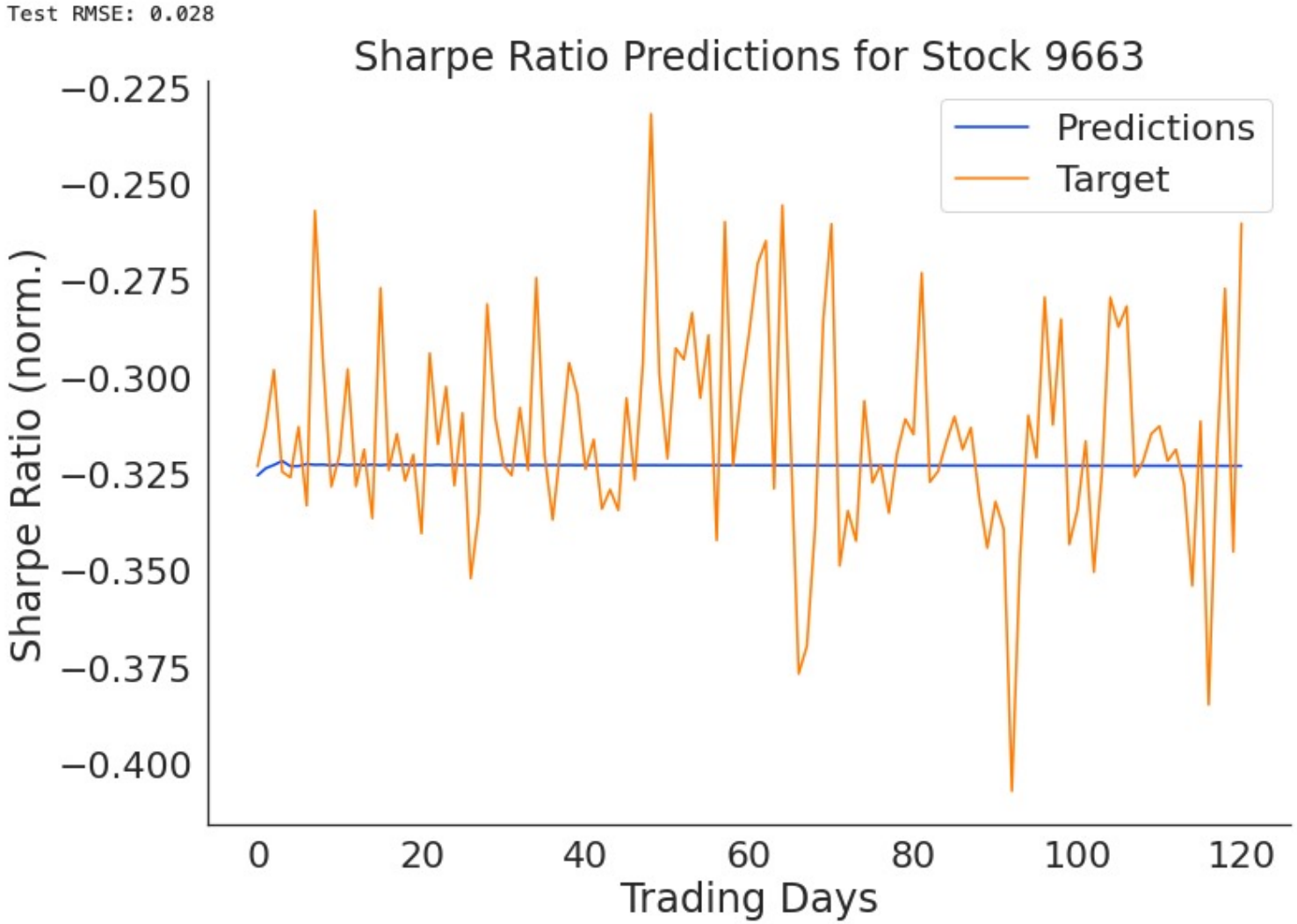
- Find best method of modelling data to predict Sharpe Ratio
  - Use ARIMA, LSTM Neural Network, and Random Forest Regressor
- RMSE used as method for calculating error and comparing models.
- Model Forecast predicts Sharpe Ratio on daily basis
  - Sharpe Ratio prediction greatly enhanced by calculating an expected Sharpe Ratio using Sharpe's equation for Ex-Post Sharpe Ratio.

Results Summary

- Most volatile stock's predictions shown for each model.
    1. ARIMA Model
    2. Prophet Model
        - Not elaborated upon as very similar to ARIMA
    3. LSTM Neural Network Model
    4. Random Forest Regressor – 1 Model for all Stocks
    5. Random Forest Regressor – Individually Modelled
- RMSE Plot and Kaggle Scores Presented

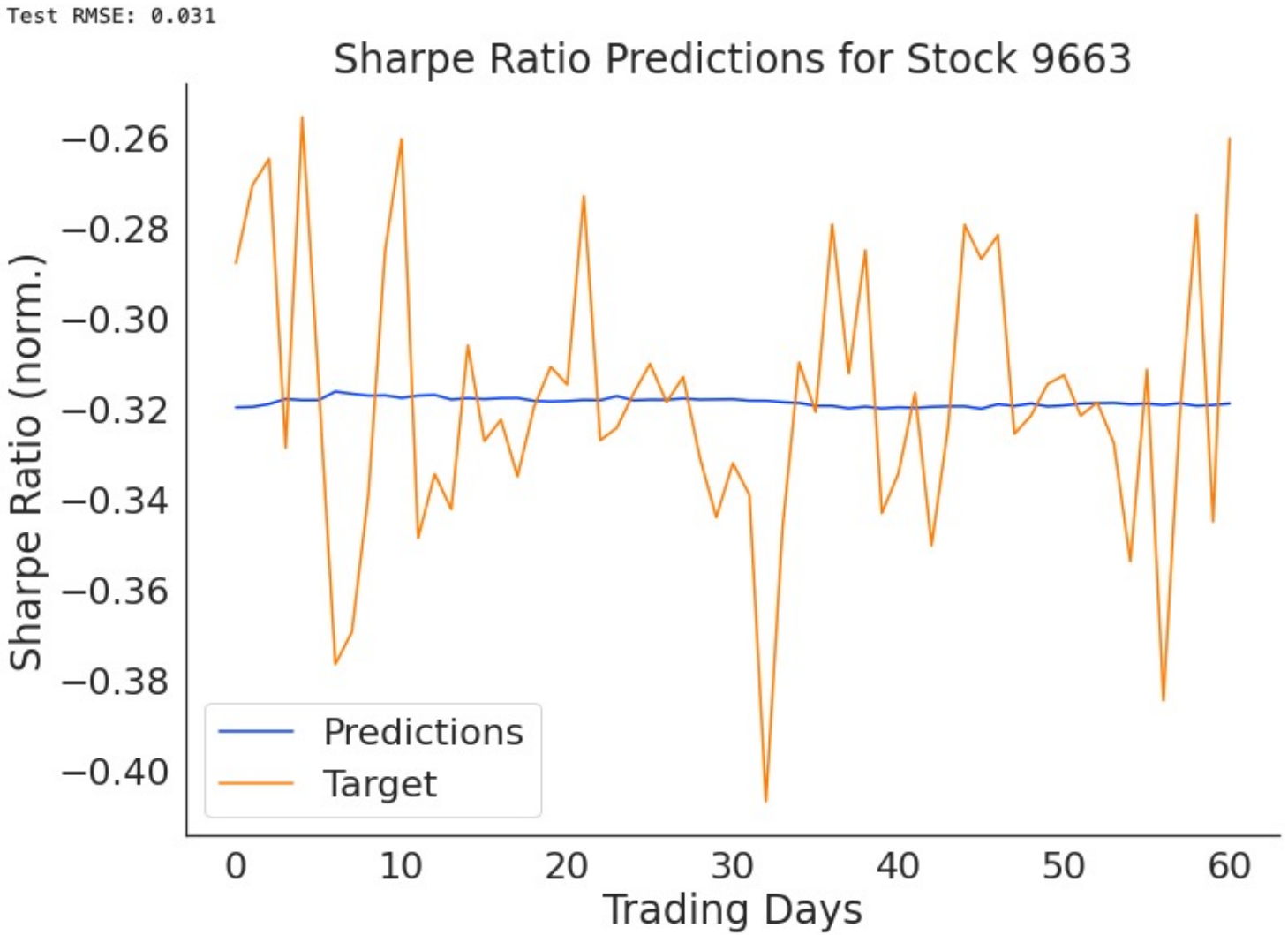
Results
ARIMA Model

- ARIMA had a low Kaggle score of -1.43.
- Model predicted near the mean because it could not determine any pattern, trend, or seasonality to follow.



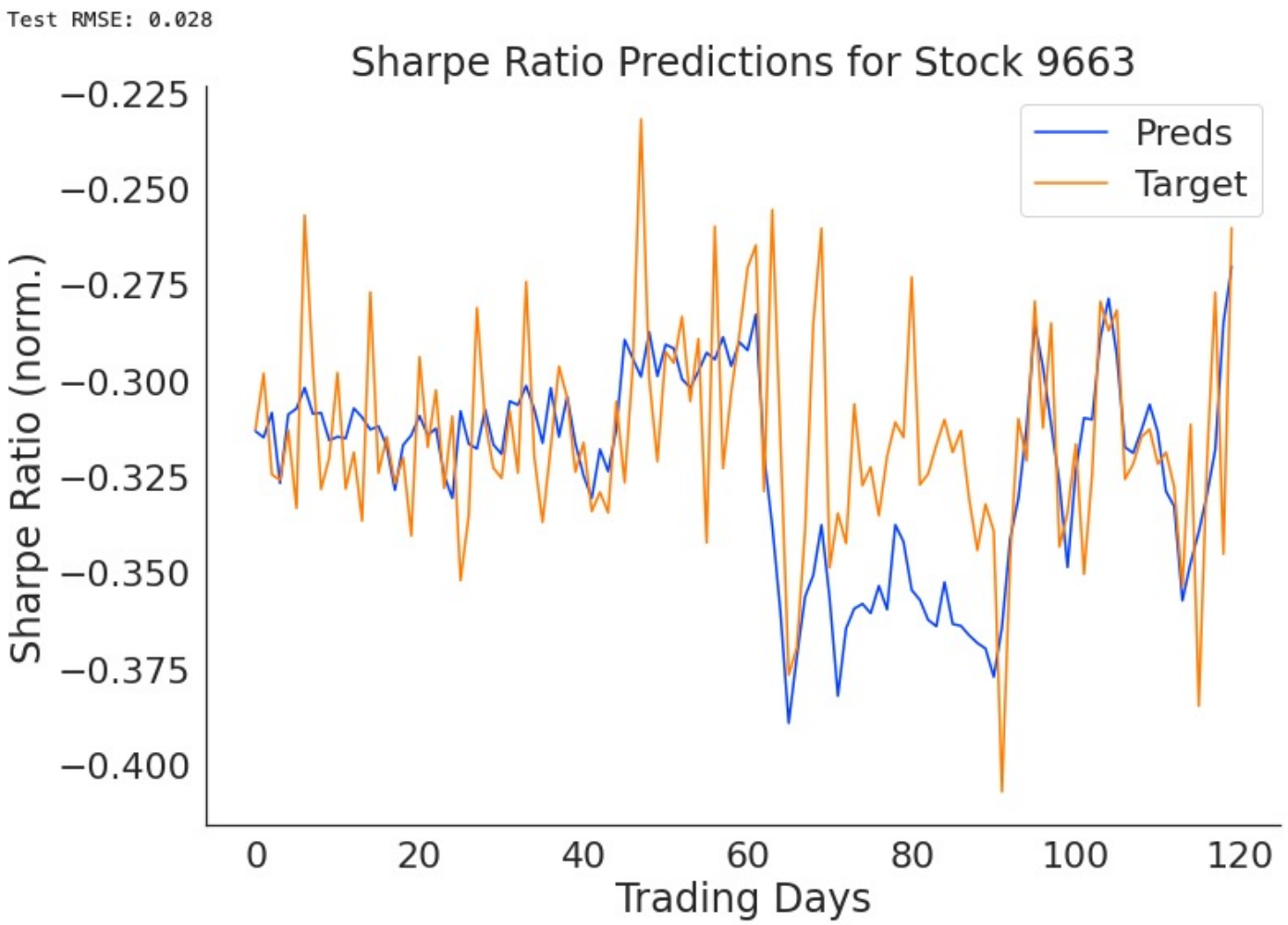
Results
LSTM Model

- LSTM was not scored due to high computational expense.
- Likely would have scored similarly to the ARIMA model due to similar predictions for each stock.
- Model predicted near the mean – likely because this was the best way to lower error.



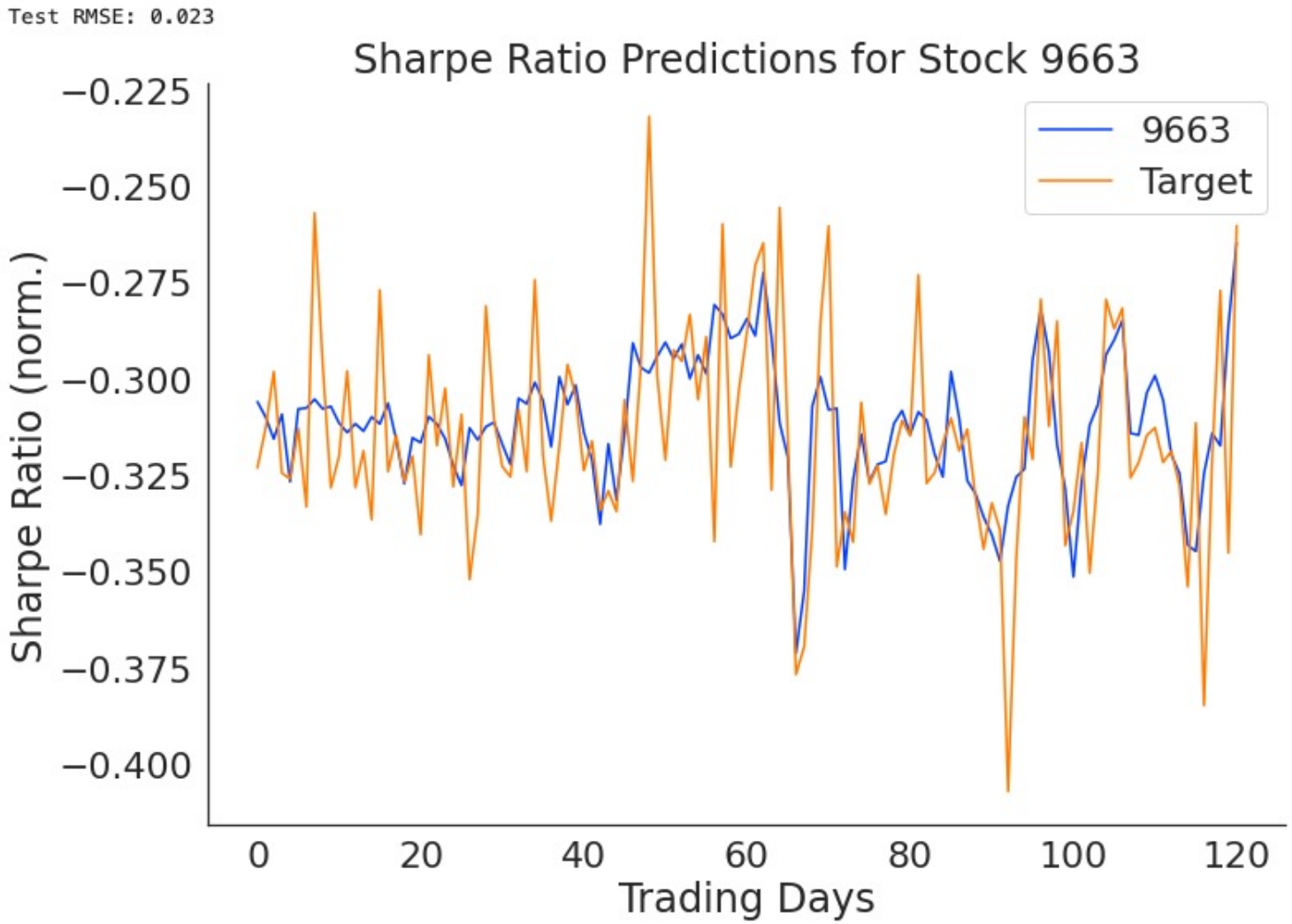
Results
Single Random Forest Regressor Model

- Model predicts much further away from the mean with similar or less error than ARIMA or LSTM
- Improvised Sharpe Ratio calculation most important feature.
    - Nearly 4 times as important as next most utilized feature.

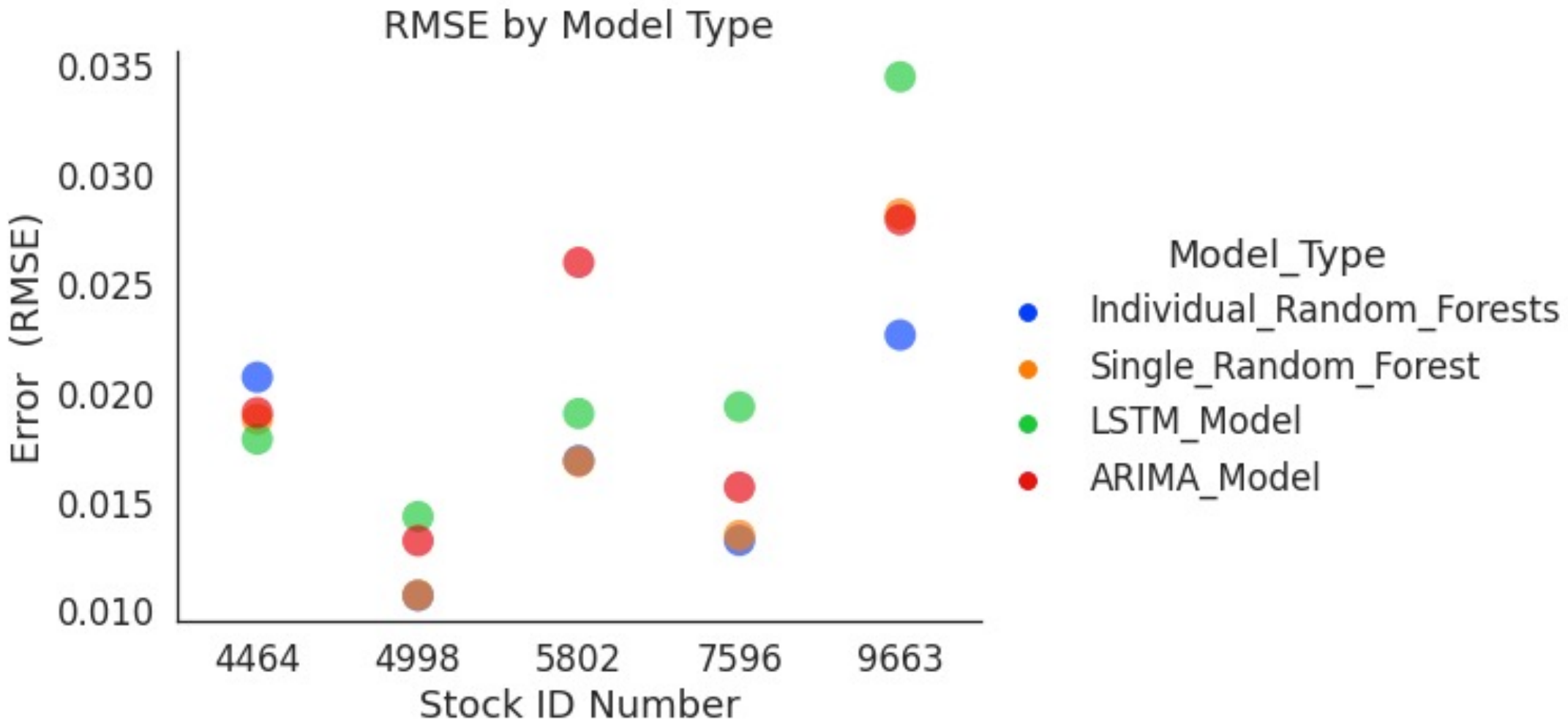


Results
Individual Random Forest Regressor Models

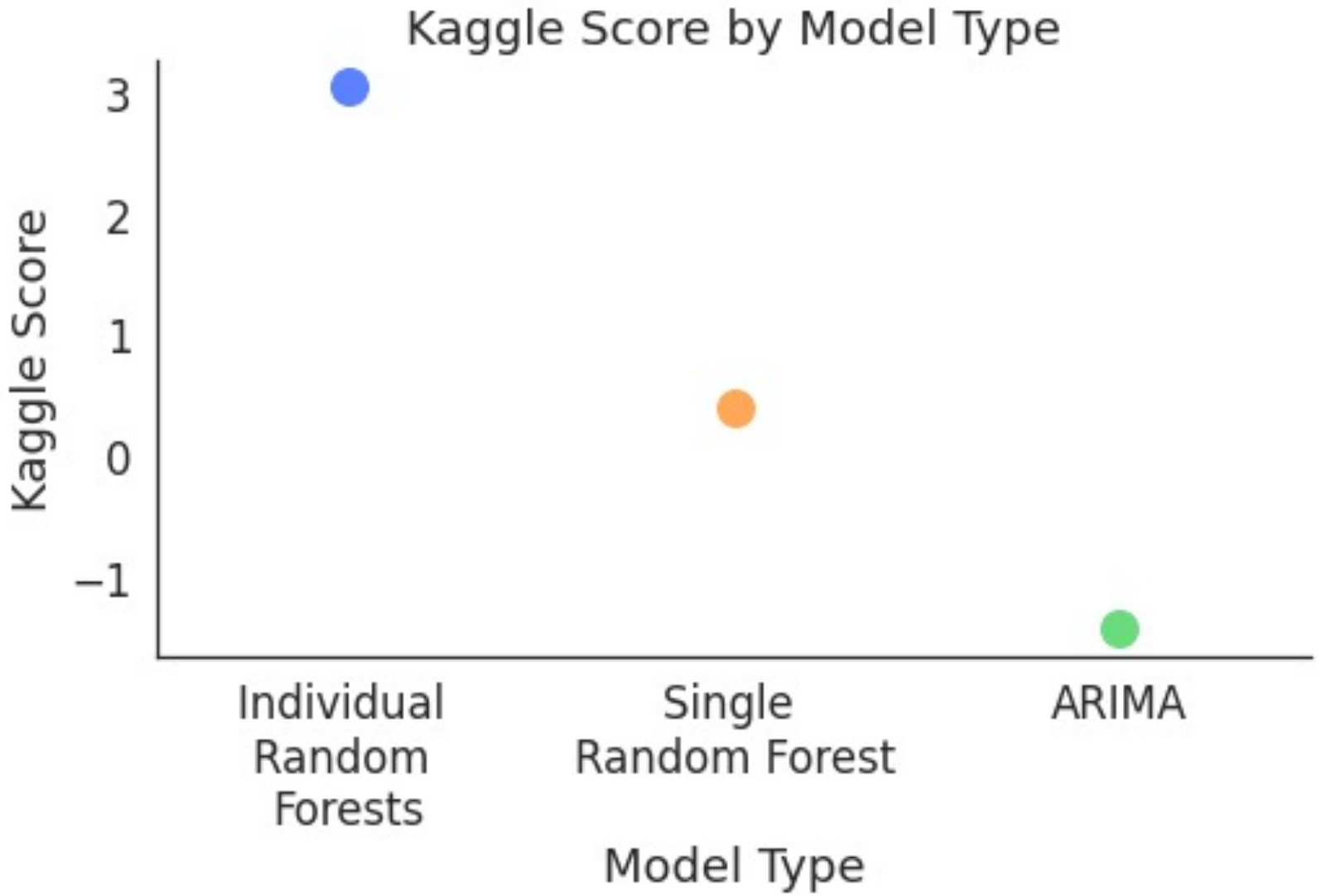
- Model with least error for Stock 9663.
- Illustrates how predicting volatile stocks important for differentiating between stocks' Sharpe Ratio.
- Predictions clearly much better than previous Random Forest Regressor between days 65 – 90.



Results


RMSE by Model Type
Error (RMSE)
0.035
0.030
0.025
0.020
0.015
0.010
4464
4998
5802
7596
9663
Stock ID Number
Model_Type
Individual_Random_Forests
Single_Random_Forest
LSTM_Model
ARIMA_Model

Results


Kaggle Score by Model Type
Kaggle Score
3
2
1
0
−1
Individual Random Forests
Single Random Forest
ARIMA
Model Type

Conclusions

- Institutional Investors and Retail Investors, in some cases, will be able to profit from daily investing at a large scale.
  - Important to understand that a large portfolio minimizes risk.
- Increased trading on JPX Stock Exchange likely
  - Random Forest Regressors are producing positive returns.
- Increasing revenue likely for JPX Stock Exchange due to competition.
- Recommend creating relationships between data scientists in competition and institutional investors.

Future Work

- **Add Options Data**
  - Information leading to model interpretation of Sharpe Ratio important
  - Need to test whether options information can improve predictions.

- **Test Additional Regressor Model Types**
  - Bayesian Regressors thought to be strong stock market model types.

- **Further Investigate Neural Networks**
  - .Many Neural Networks with options for layers are possible
  - Want Neural Network that risks predictions further from mean

Questions & Answers

Contact Information:

- Jim Petoskey – Data Scientist
  - Jim.Petoskey.146@gmail.com
  - Phone: (734) 649-3182
  - linkedin.com/in/jimpetoskey/