

Capstone Project 2

Stock Market Trend Predictions using Random Forests and Gradient Boosting

Springboard - Vincent Tseng

Goal for the project

- Predict the price movement from 1 to 20 days ahead and decide how many days ahead has a better prediction
 - ✓ Technology Stocks: Apple, Microsoft
 - ✓ Retail Stocks: Costco, JNJ
- Compare the performance between Random Forest and Gradient Boost
 - ✓ Evaluate the model by Accuracy and F1 Score
- Compare technology stocks and retail stocks

Contents

- **Data Preparation**

- Get “closing prices”, “High”, “Low”, and “Volume” from Yahoo finance for all stocks
- Use the TA-Lib package to get stock indicators

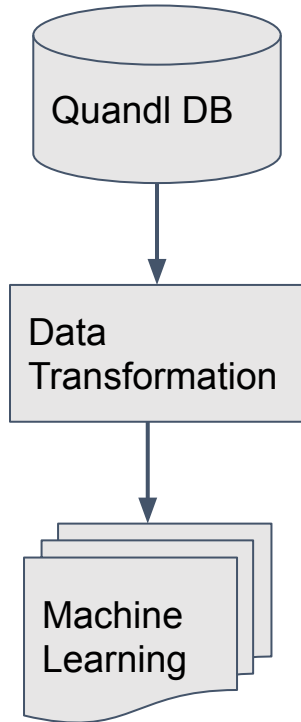
- **Machine Learning Algorithms**

- Random Forest
- Extreme Gradient Boosting (XG Boost)
- Evaluation: F1 Score Comparison

- **Results**

- Conclusion

Data Preparation

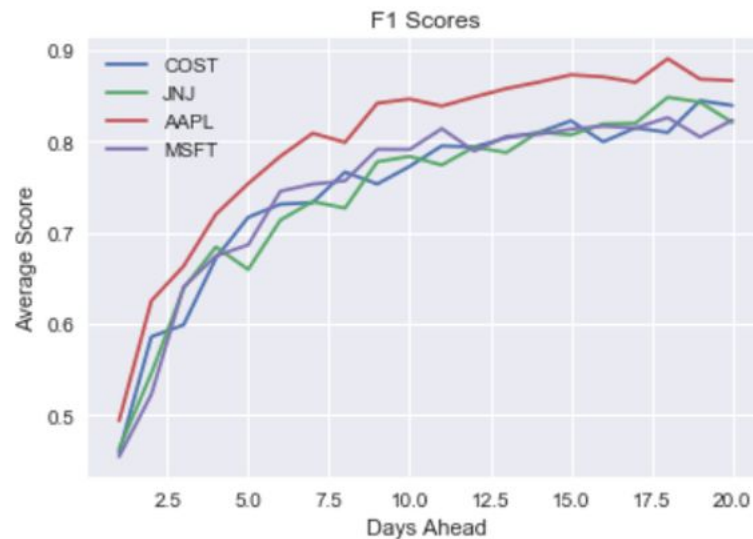
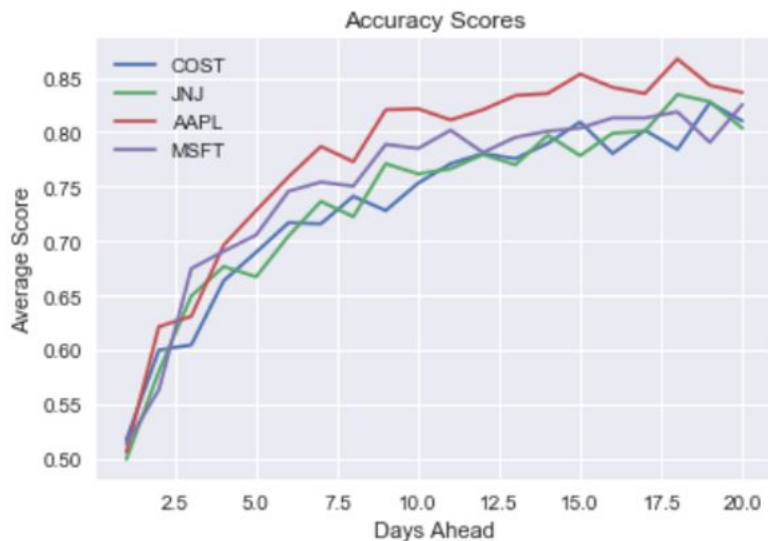


- Data Acquisition
 - ✓ Technology Stocks: Apple, Microsoft
 - ✓ Retail Stocks: Costco, JNJ
 - ✓ Data Range: 1/1/2006 to 12/31/2019
- Data Transformation
 - ✓ Get "stock indicators" by TA-lib package
- Machine Learning Algorithms
 - ✓ Random Forest & XG Boost
 - ✓ Model Evaluation

Stock Indicators

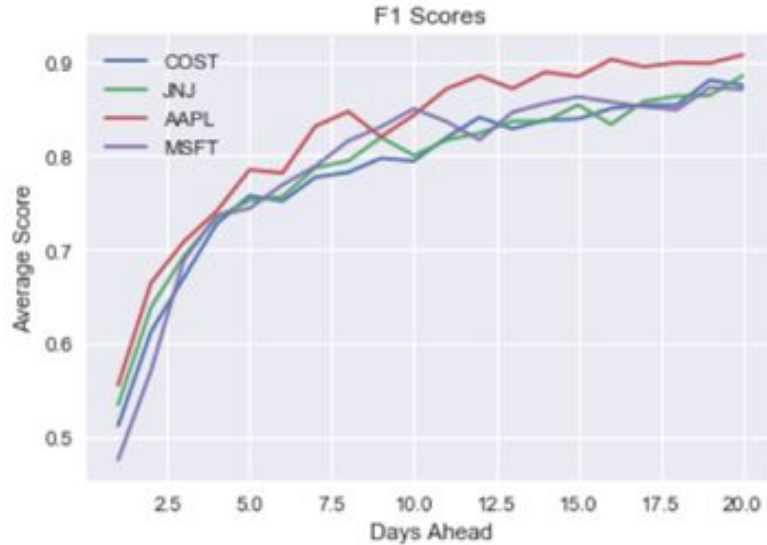
Type	Abbreviation	Name
Overlap Studies	BBANDS	Bollinger Bands
	SMA	Simple Moving Average
	WMA	Weighted Moving Average
Momentum Indicators	MOM	Momentum
	MACD	Moving Average Convergence Divergence
	RSI	Relative Strength Index
	WILLR	Williams' Percentage R
	CCI	Commodity Channel Index
	ADX	Average Directional Movement Index
Volume Indicators	ADOSC	Chaikin A/D Oscillator
	OBV	On Balance Volume

Evaluate Random Forest Algorithm



- The highest scores are between 18-20 days ahead.

Evaluate Extreme Gradient Boost



- The highest scores are between 18-20 days ahead.

Conclusion

Highest F1 Score Comparison	AAPL	MSFT	COST	JNJ
Random Forest	0.8902	0.8258	0.8444	0.8480
Extreme Gradient Boosting	0.9074	0.8727	0.8805	0.8851

- Predictions within 18 to 20 days ahead with F1 scores ranging from 0.80 to 0.91.
- Extreme gradient boosting performs better than random forest since the scores are a little higher.
- No huge difference between technology stocks and retail stocks.