FINAL REPORT

Do-Young Chung
ICS4U
Ms. Garima
03/12/2021

Statement of the problem

The stock market is where investors connect to buy and sell investments to make a large amount of money. Nowadays, more people are turning to stock market investment for profits and even entertainment. But the problem is that most likely average day trader loses money. Furthermore, as the cryptocurrency market rises and enjoys unprecedented success, now investment is very common in our life and the intrinsic values of mathematics and computing to investment are emerging. Within this situation, I used ARIMA (autoregressive integrated moving average) model to predict future points in the price of time series data.

Purpose

The purpose of the program is to explore a dataset of large company stocks, and analyze them in order to boost up the probability of winning profits by predicting prices using the ARIMA module. The program analyzes stock market data to predict the following price through five stages: Data Set Selection, Data Set Cleaning, Data Set Visualization, and Data Set Prediction. Ultimately, compare the accuracy of predicted data by using different size of data and different types of data.

Dataset

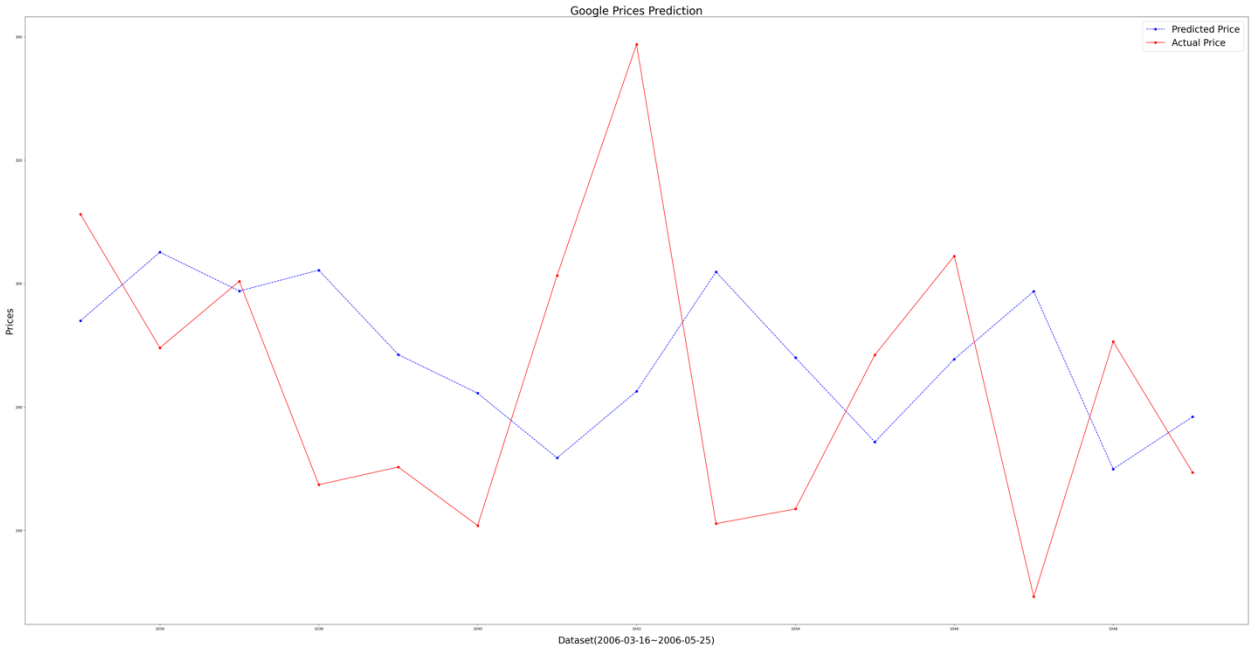
- Google (2006-01-03 to 2017-12-29)
- Amazon (2006-01-03 to 2017-12-29)

Technology/Package

- Python (Google Collab)
- NumPy and Pandas Data frame
- Data Set Visualization Tools (matplotlib, lag_plot, bootstrap_plot)
- Data Set Prediction Tools (ARIMA, mean_squared_error)

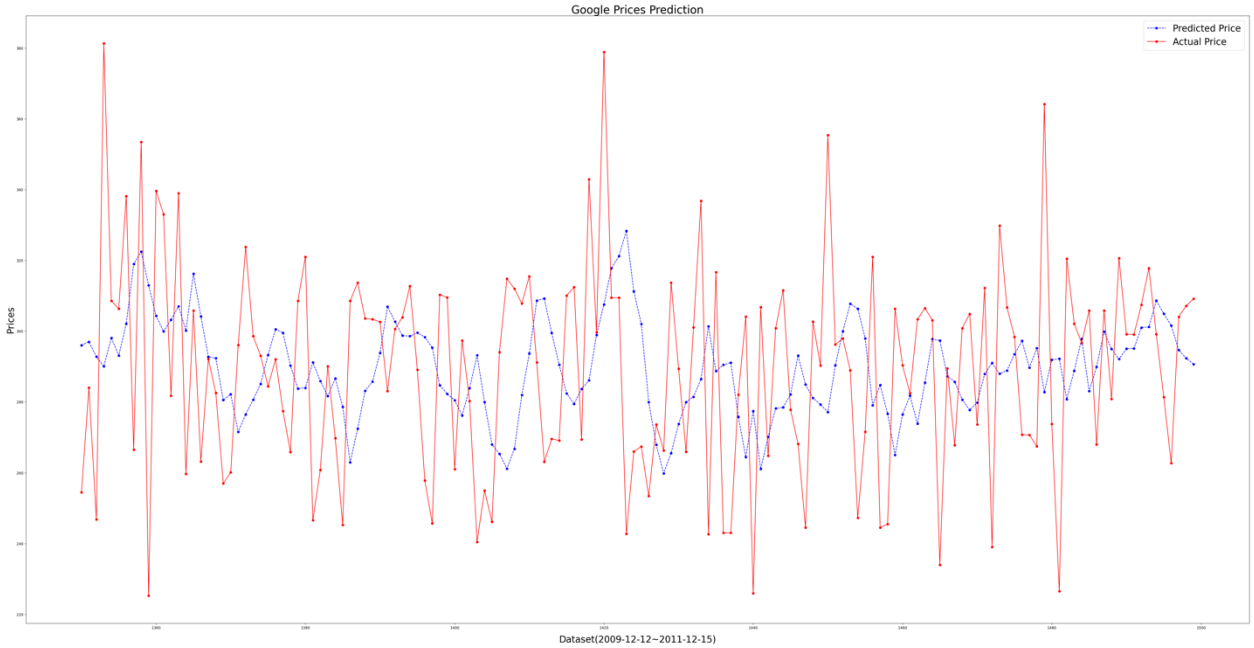
Result

50 Size Google Stock Data (2006-03-16 to 2006-05-25)



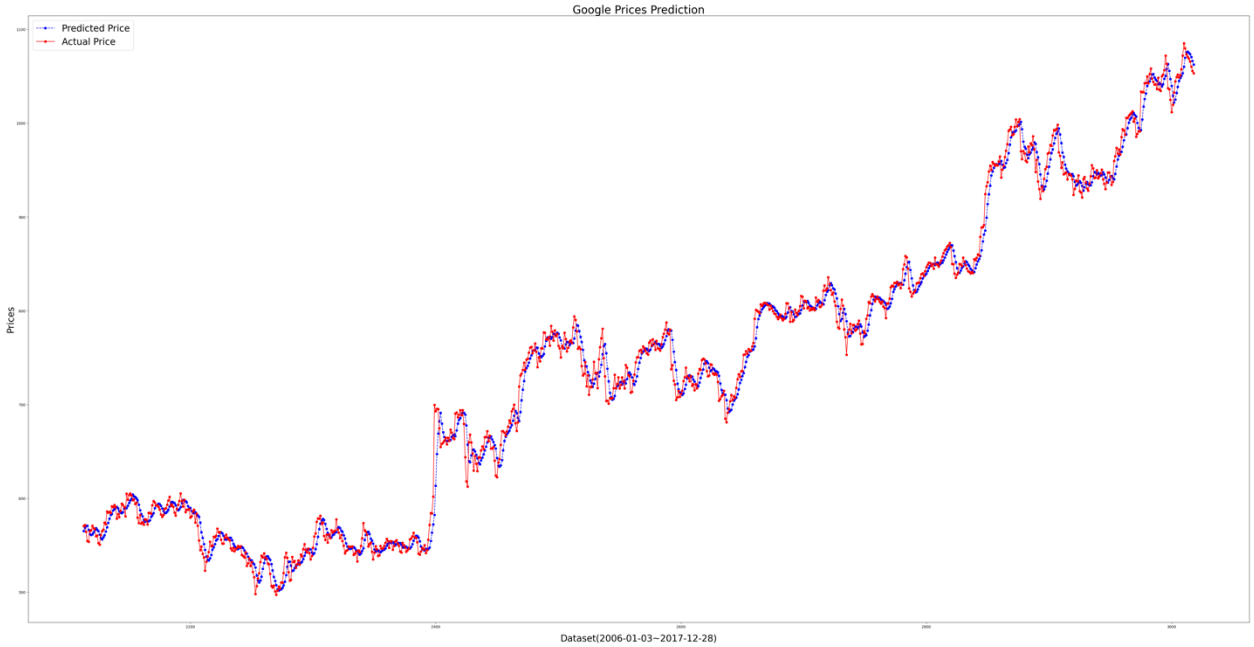
The accuracy of the prediction data is 54.743104437698776%

500 Size Google Stock Data (2009-12-12 to 2011-12-15)



The accuracy of the prediction data is 75.29703101252647%

3018 Size Google Stock Data (2006-01-03 to 2017-12-29)

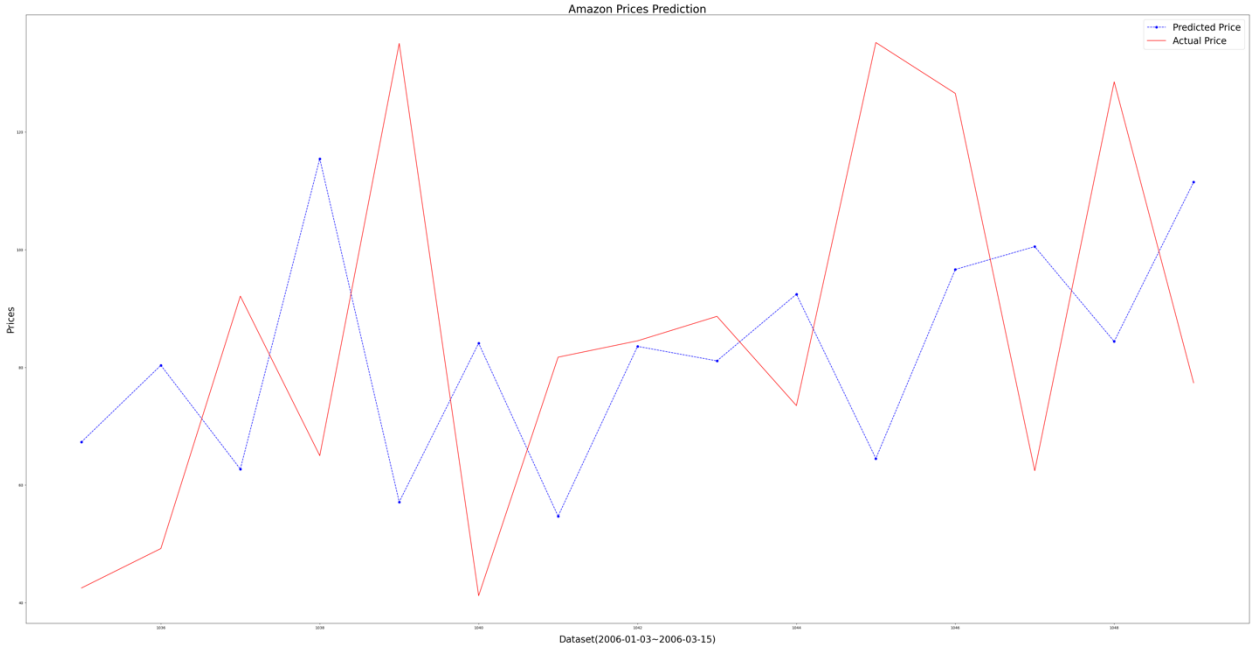


The accuracy of the prediction data is 98.38578746878025%

Google Size Accuracy Table

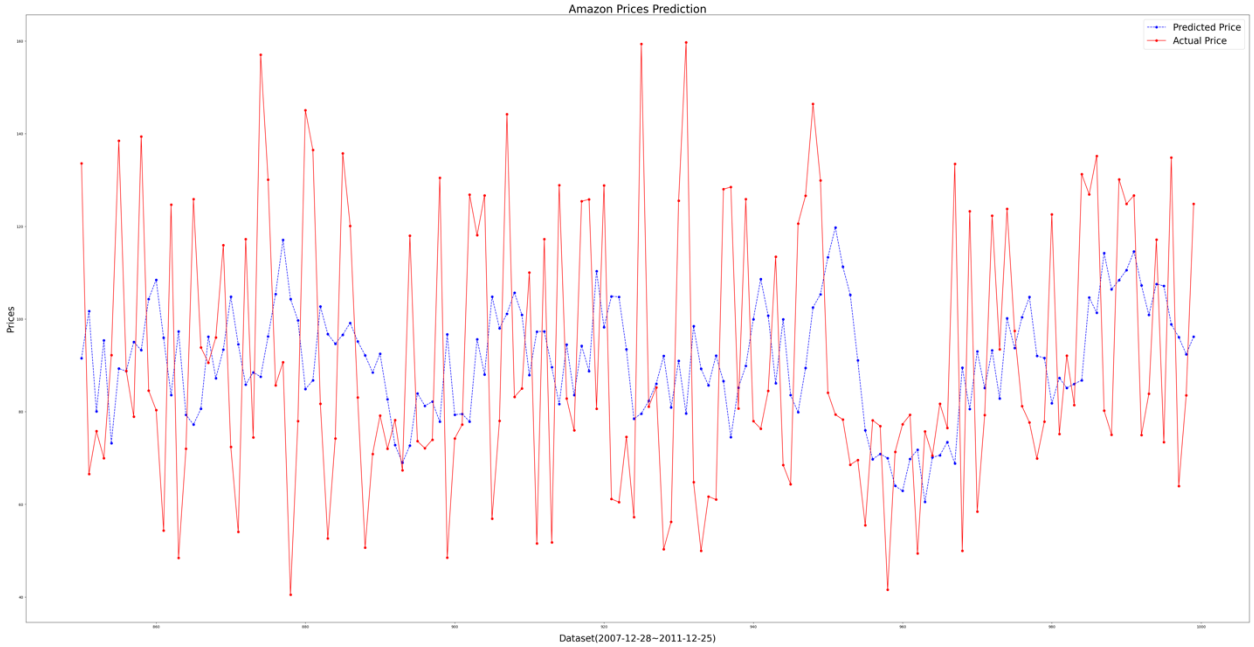
Size	50	500	3018
Accuracy (%)	63.2%	70.5%	98.4

50 Size Amazon Stock Data (2006-01-03 to 2006-03-15)



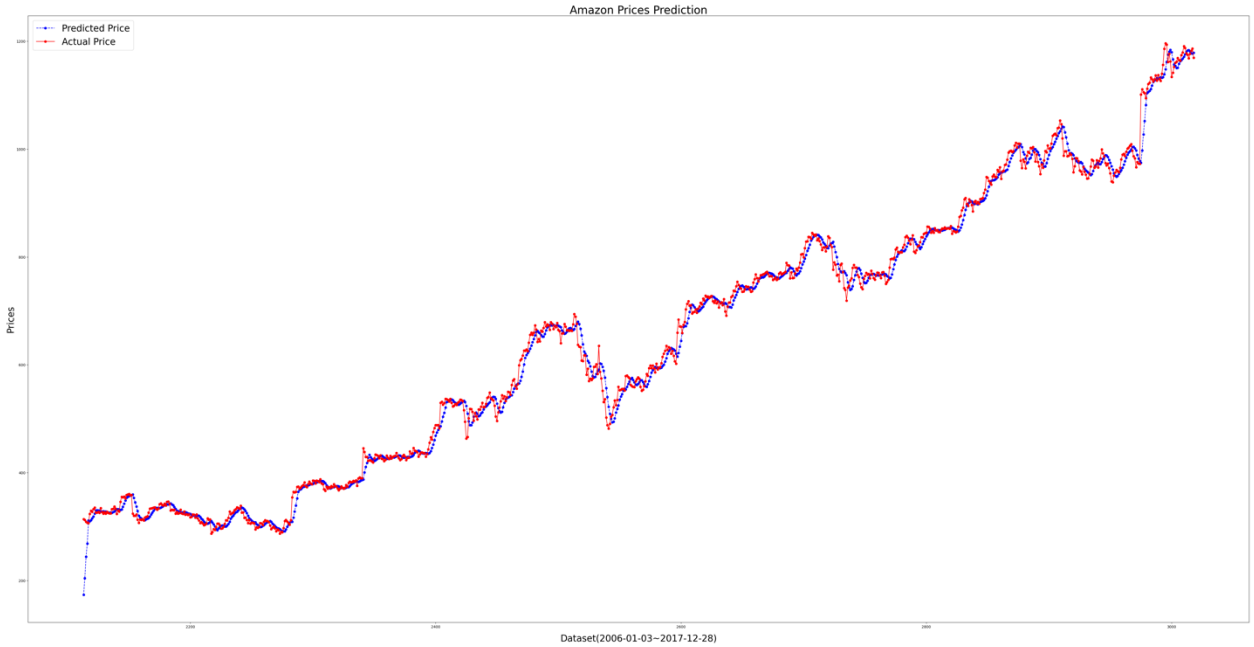
The accuracy of the prediction data is 54.86585128609253%

500 Size Amazon Stock Data (2007-12-28 to 2009-12-21)



The accuracy of the prediction data is 67.9201803636624%

3018 Size Amazon Stock Data (2006-01-03 to 2017-12-29)



The accuracy of the prediction data is 98.00405754201196%

Amazon Size Accuracy Table

Size	50	500	3018
Accuracy (%)	54.9%	67.9%	98.0%

Compare

Google Size Accuracy Table

Size	50	500	3018
Google Accuracy (%)	63.2%	70.5%	98.4
Amazon Accuracy (%)	54.9%	67.9%	98.0%
Difference (absolute %)	8.3%	2.6%	0.4%

As the size of data that is put to train increases, the accuracy of the algorithm also increases.

Final Project Report Rubrics

Achievement category	Marks Given	50–59% (Level 1)	60–69% (Level 2)	70–79% (Level 3)	80–100% (Level 4)
Knowledge & Understanding - The student demonstrates a thorough understanding of the concepts related to python programming and software development in the report.	/20	Demonstrates limited understanding of python programming and software development concepts in the report.	Demonstrates some understanding of python programming and software development concepts in the report.	Demonstrates a considerable understanding of the python programming and software development concepts in the report.	Demonstrates a high degree of understanding of the python programming and software development concepts in the report.
Thinking – The student is able to come up with a creative project idea and present the idea with the devised solution in project report	/20	Demonstrates limited ability in presenting the project idea and its solution in a creative way in the project report.	Demonstrates some critical understanding in presenting the project idea and its solution in a creative way in the project report.	Demonstrates a considerable understanding in presenting the project idea and its solution in a creative way in the project report.	Demonstrates a high degree of critical understanding in presenting the project idea and its solution in a creative way in the project report.
Communication - follows the project report template and all the chapters should be well structured and clearly written. Design diagrams should be in proper format	/40	Demonstrates limited ability to follow the project report template and structure it. Design diagrams also created with limited effectiveness.	Demonstrates some ability to follow the project report template and structure it. Design diagrams were also created with some effectiveness.	Demonstrates considerable ability to follow the project report template and structure it. Design diagrams were also created with considerable effectiveness.	Demonstrates a high degree of understanding, follows the project report template and structures it. Design diagrams also created with effectiveness.
Application – The student applies their understanding of python programming with software development concepts in their project report	/20	Apply knowledge of python programming and software development concepts with limited effectiveness.	Apply knowledge of python programming and software development concepts with some effectiveness.	Apply knowledge of python programming and software development concepts with considerable effectiveness.	Apply knowledge of python programming and software development concepts with a high degree of effectiveness.

Comments :