



Stock Price Prediction

APMA 4990

Ho Hin Lee, Kevin Jin, Tiancheng Li, Yuchen Li



Problem Statement

- 1 Stock prices can be volatile and fluctuate daily
- 2 Quantitative trading is used by many financial institutions and hedge funds
- 3 Goal is to build a machine learning model to predict future stock prices based on past data to form a buy/sell decision



>>> Dataset - NASDAQ: GOOG

Date	Open	High	Low	Close	Adj Close	Volume	50 Day MA	200 Day MA
1/3/2017	778.81	789.63	775.80	786.14	786.14	1657300	778.11	751.87
1/4/2017	788.36	791.34	783.16	786.90	786.90	1073000	777.86	752.10
1/5/2017	786.08	794.48	785.02	794.02	794.02	1335200	777.48	752.36
1/6/2017	795.26	807.90	792.20	806.15	806.15	1640200	777.45	752.70
1/9/2017	806.40	809.97	802.83	806.65	806.65	1272400	777.60	753.06
1/10/2017	807.86	809.13	803.51	804.79	804.79	1176800	777.79	753.42
1/11/2017	805.00	808.15	801.37	807.91	807.91	1065900	778.04	753.73
1/12/2017	807.14	807.39	799.17	806.36	806.36	1353100	778.48	754.01
1/13/2017	807.48	811.22	806.69	807.88	807.88	1099200	778.96	754.33
1/17/2017	807.08	807.14	800.37	804.61	804.61	1355800	779.68	754.60
1/18/2017	805.81	806.21	800.99	806.07	806.07	1294400	780.56	754.90
1/19/2017	805.12	809.48	801.80	802.17	802.17	919300	781.37	755.22
1/20/2017	806.91	806.91	801.69	805.02	805.02	1670000	781.82	755.52
1/23/2017	807.25	820.87	803.74	819.31	819.31	1963600	782.39	755.92
1/24/2017	822.30	825.90	817.82	823.87	823.87	1474000	783.16	756.34
1/25/2017	829.62	835.77	825.06	835.67	835.67	1494500	784.62	756.84
1/26/2017	837.81	838.00	827.01	832.15	832.15	2973900	786.19	757.28
1/27/2017	834.71	841.95	820.44	823.31	823.31	2965800	787.93	757.64
1/30/2017	814.66	815.84	799.80	802.32	802.32	3246600	788.81	757.89



»» Dataset

1. CSV file downloaded from Yahoo Finance from January 2017 to March 2018



2. Basic Features from downloaded data include

- Open
- Close
- High
- Volume
- Low



3. Additional Features are added:

- 50 day moving average
- 200 day moving average
- Daily Percent Change
- Daily High-Low fluctuation



- Several S&P500 stock datasets will be used for final web app

- 1 Several Machine Learning models will be tested with the dataset
 - Random Forest Regression
 - Gradient Boosting
 - ARIMA Model
 - Recurrent Neural Network
- 2 Model will predict price (y) for a number of days beyond the current dataset
- 3 With the predicted price, a buy/sell decision will be output



Thanks for listening!
