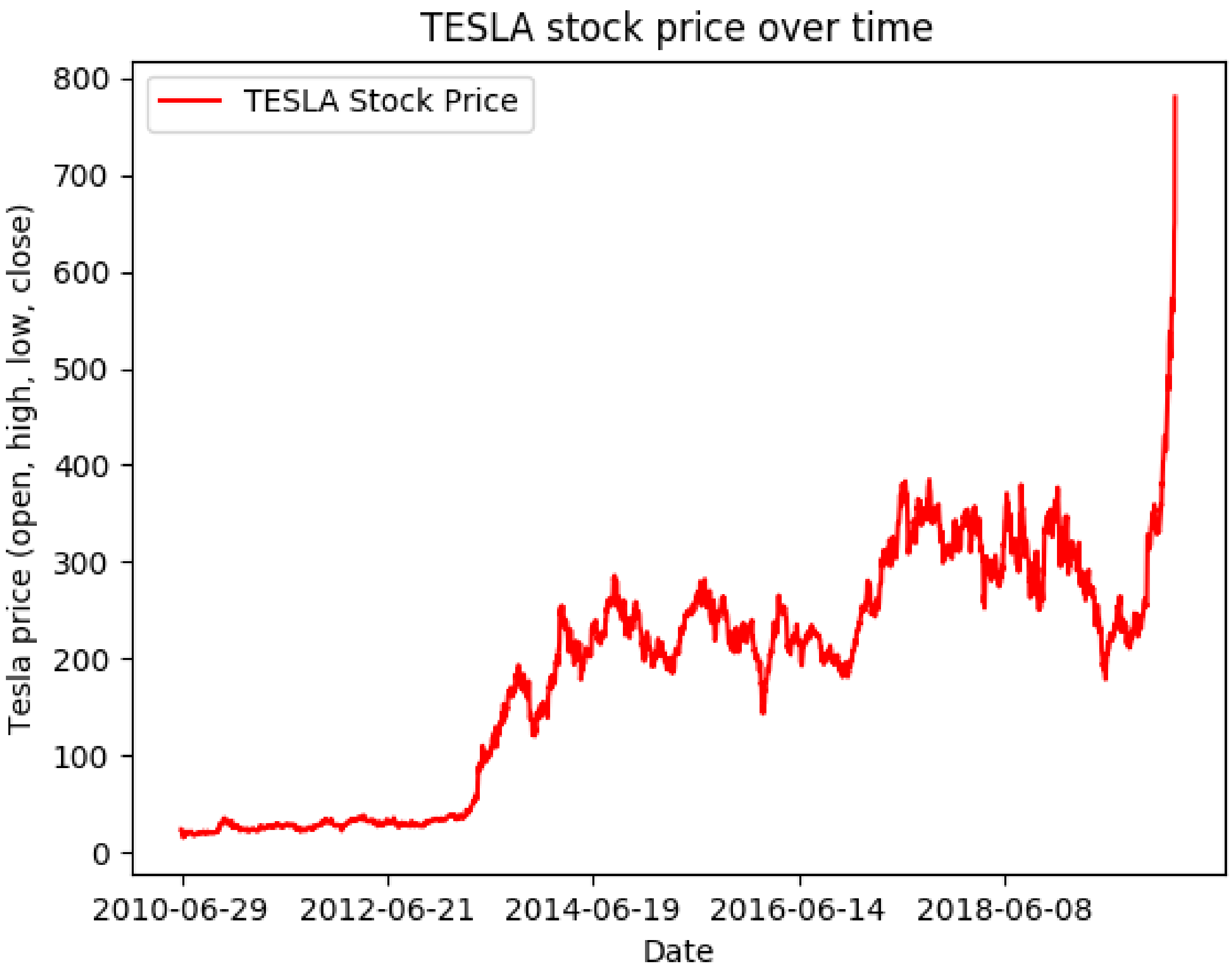
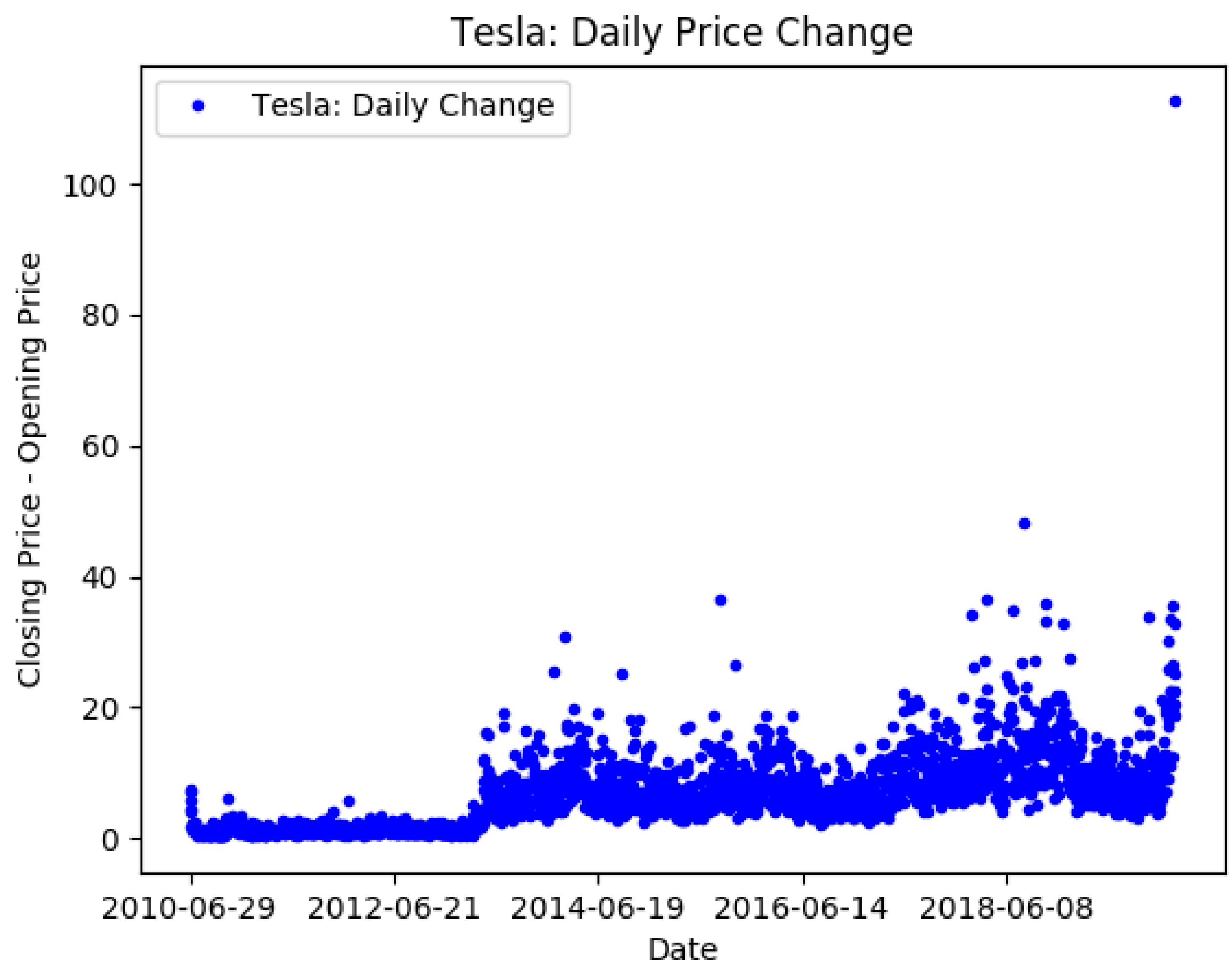
PREDICTING TESLA'S STOCK PRICE WITH MACHINE LEARNING

Christopher Pearson



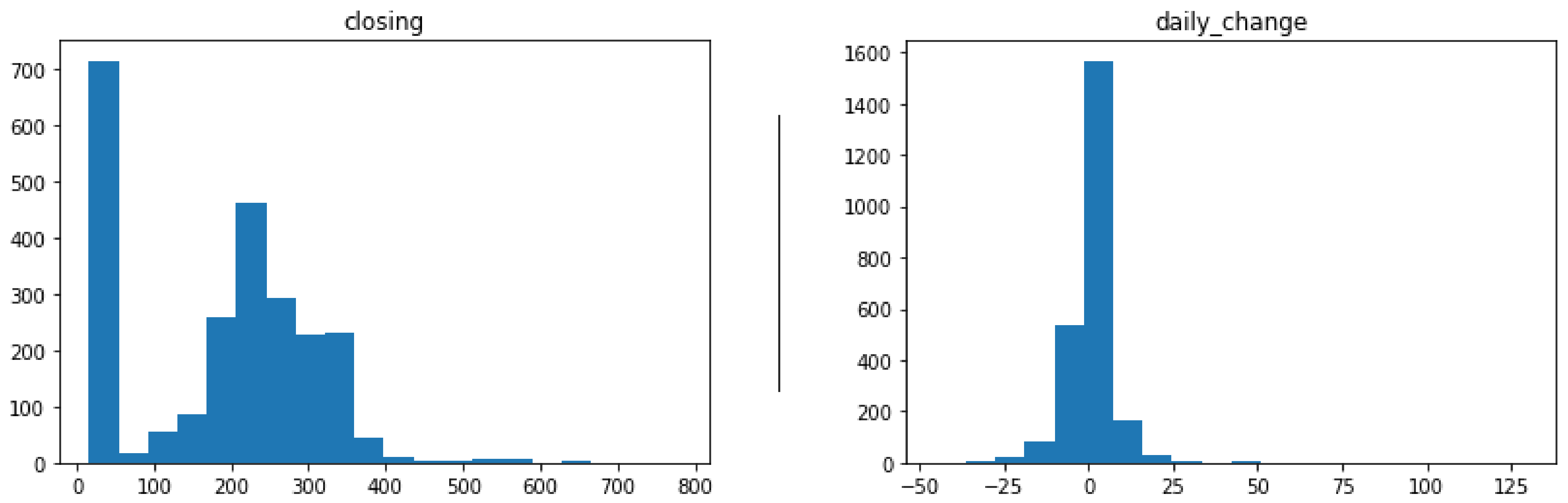
Our Data: Tesla Stock Prices

- Clean Data: No missing Data
- 2416 Rows
- Features
  - Date
  - Opening
  - Closing
  - High
  - Low
  - Volume



Additional Feature: Daily Change

- Looking at Closing – Opening Price
- Some incredible spikes, no sudden drops.
- This is NOT today's closing – yesterday's closing
  - Yesterday's closing – today's opening lost


closing
700
600
500
400
300
200
100
0
0
100
200
300
400
500
600
700
800
daily_change
1600
1400
1200
1000
800
600
400
200
0
−50
−25
0
25
50
75
100
125

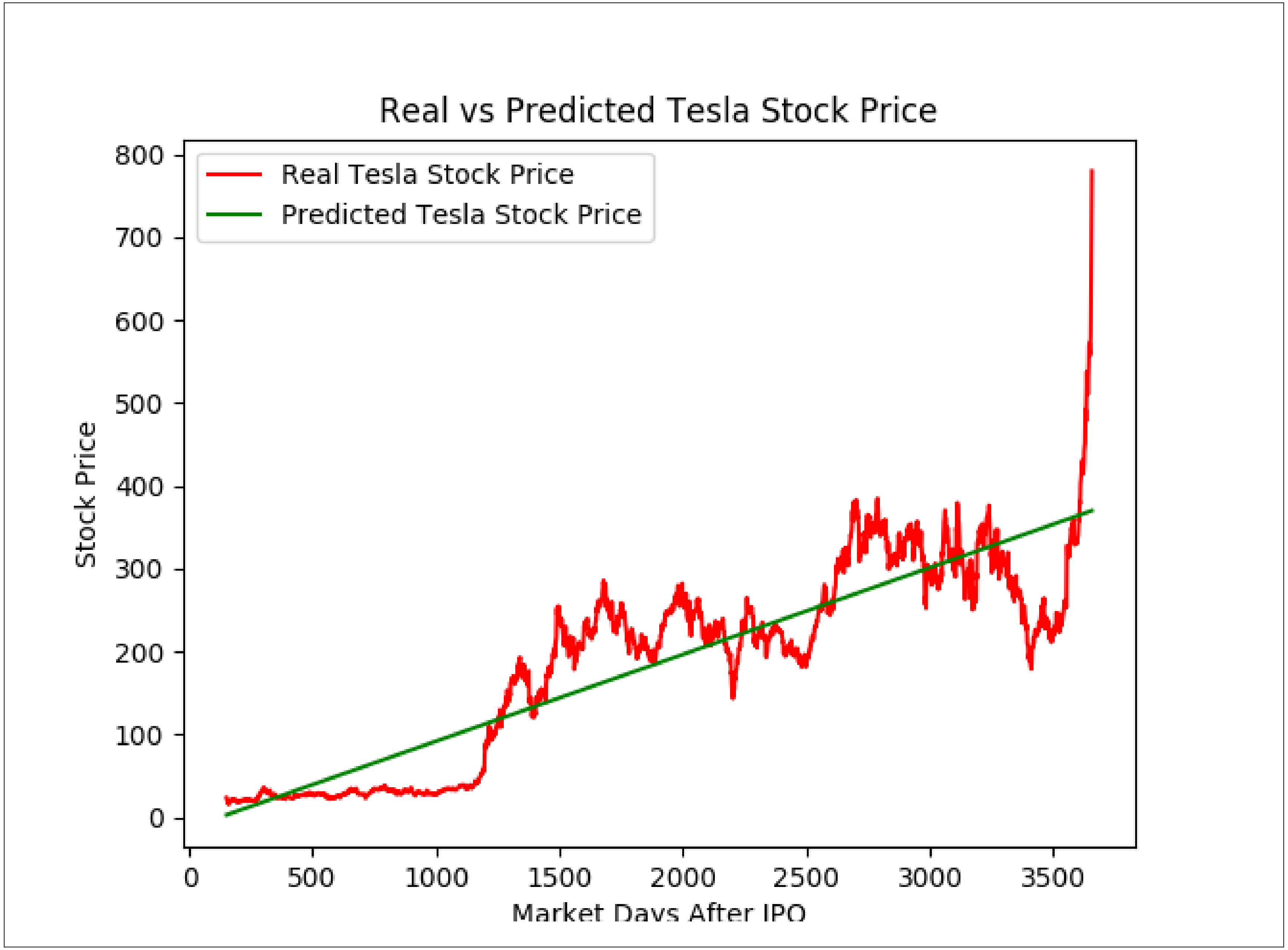
HISTOGRAMS

	open	high	low	close	adj_close	volume	days_after_ipo	daily_change
count	2416	2416	2416	2416	2416	2416	2416	2416
mean	186.2714	189.5782	182.9166	186.4037	186.4037	5572722	1902.465	0.132504
std	118.7402	120.8923	116.8576	119.136	119.136	4987809	1012.571	5.628115
min	16.14	16.63	14.98	15.8	15.8	118500	151	-28.08
25%	34.3425	34.8975	33.5875	34.4	34.4	1899275	12025.75	-1.76251
50%	213.035	216.745	208.87	212.96	212.96	4578400	1903.5	-0.015
75%	266.45	270.9275	262.1025	266.775	266.775	7361150	2778.25	1.762506
max	673.69	786.14	673.52	780	780	47065000	3657	106.31

Corrrelation	open	high	low	close	adj_close	volume	days_after_ipo	daily_change
open	1	0.999425	0.999575	0.998886	0.998886	0.501762	0.98111	0.046754
high		1	0.999389	0.99964	0.99964	0.512944	0.890536	0.07485
low			1	0.999447	0.999447	0.493496	0.89096	0.067596
close				1	1	0.505169	0.890294	0.093839
adj_close					1	0.505169	0.890294	0.093839
Volume						1	0.477066	0.107403
days_after_ipo							1	0.045402
daily_change								1

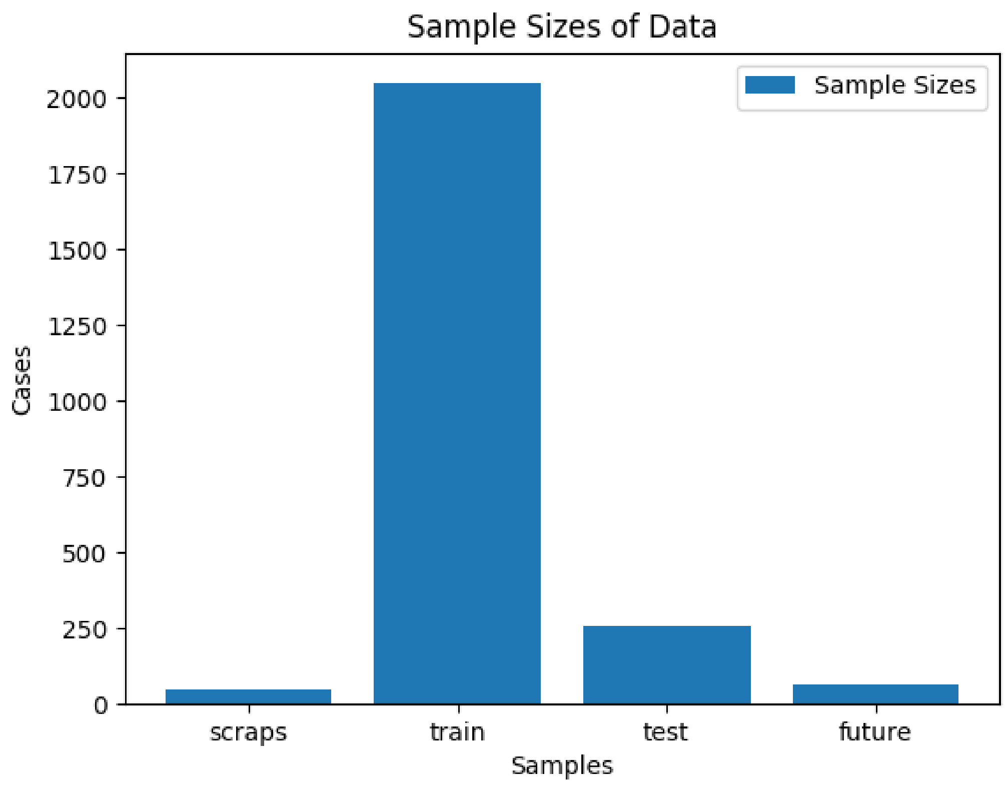
Key Figures

- N = 2416
- No missing data
- Range: $14.98 – $786.14
- Average: $186
- Open, high, low, close and adjusted close all highly correlated
- Volume and daily change not very correlated w/ other columns



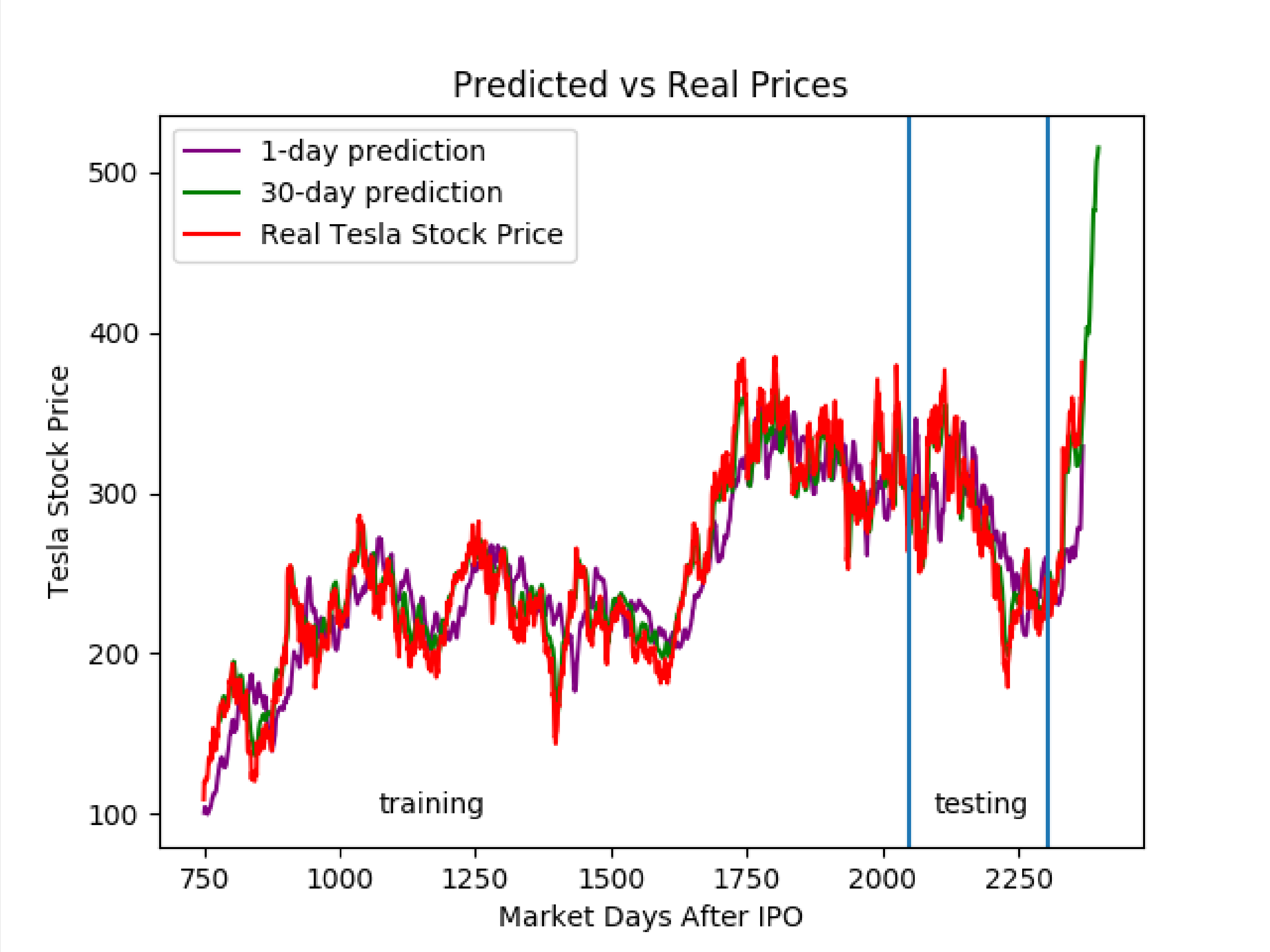
Basic Regression Model

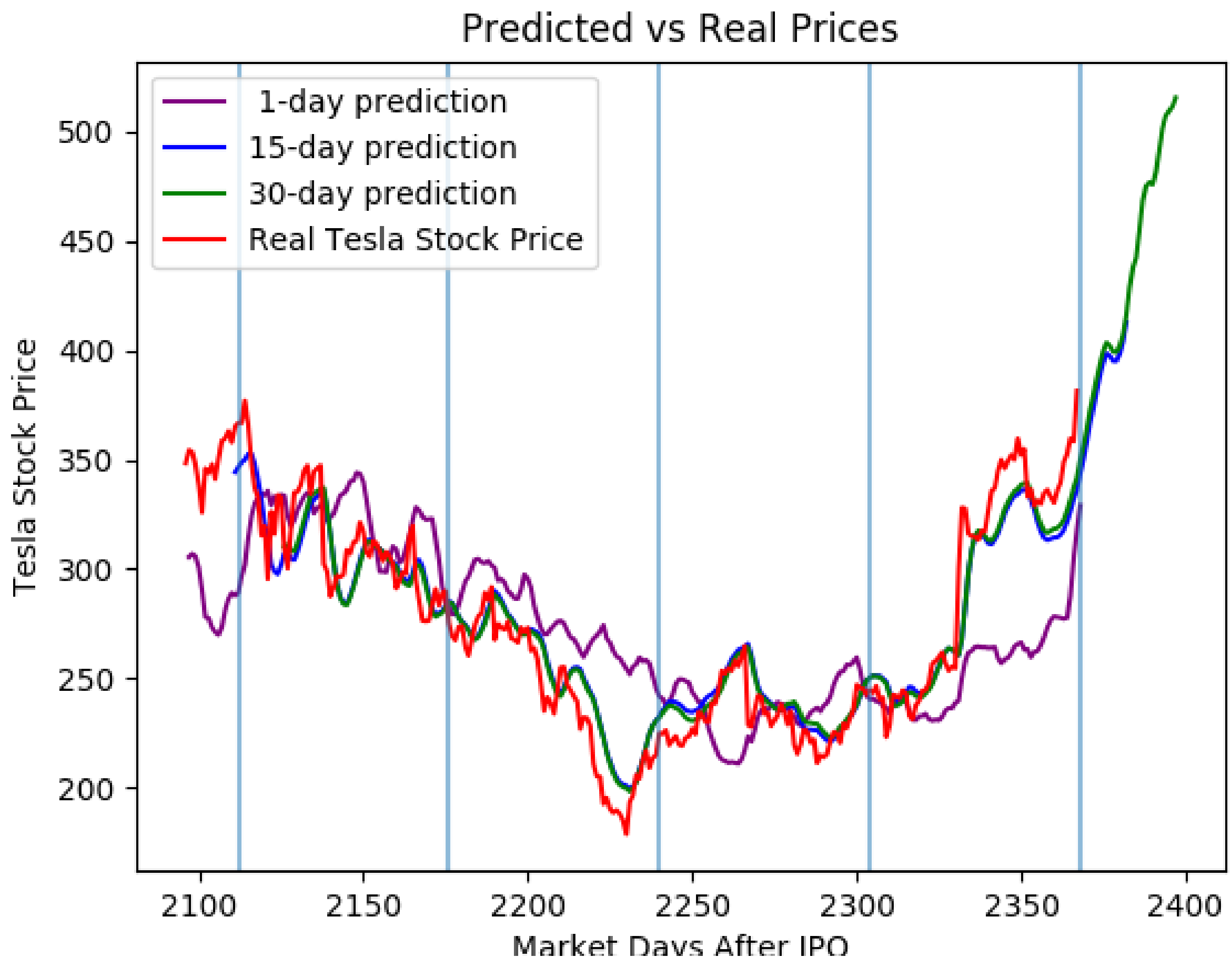
- R^2 = 0.88
- OK for summary
- Not for Prediction
- Regression Model w/ more independent variables would be better



Machine Learning Model: Can we improve?

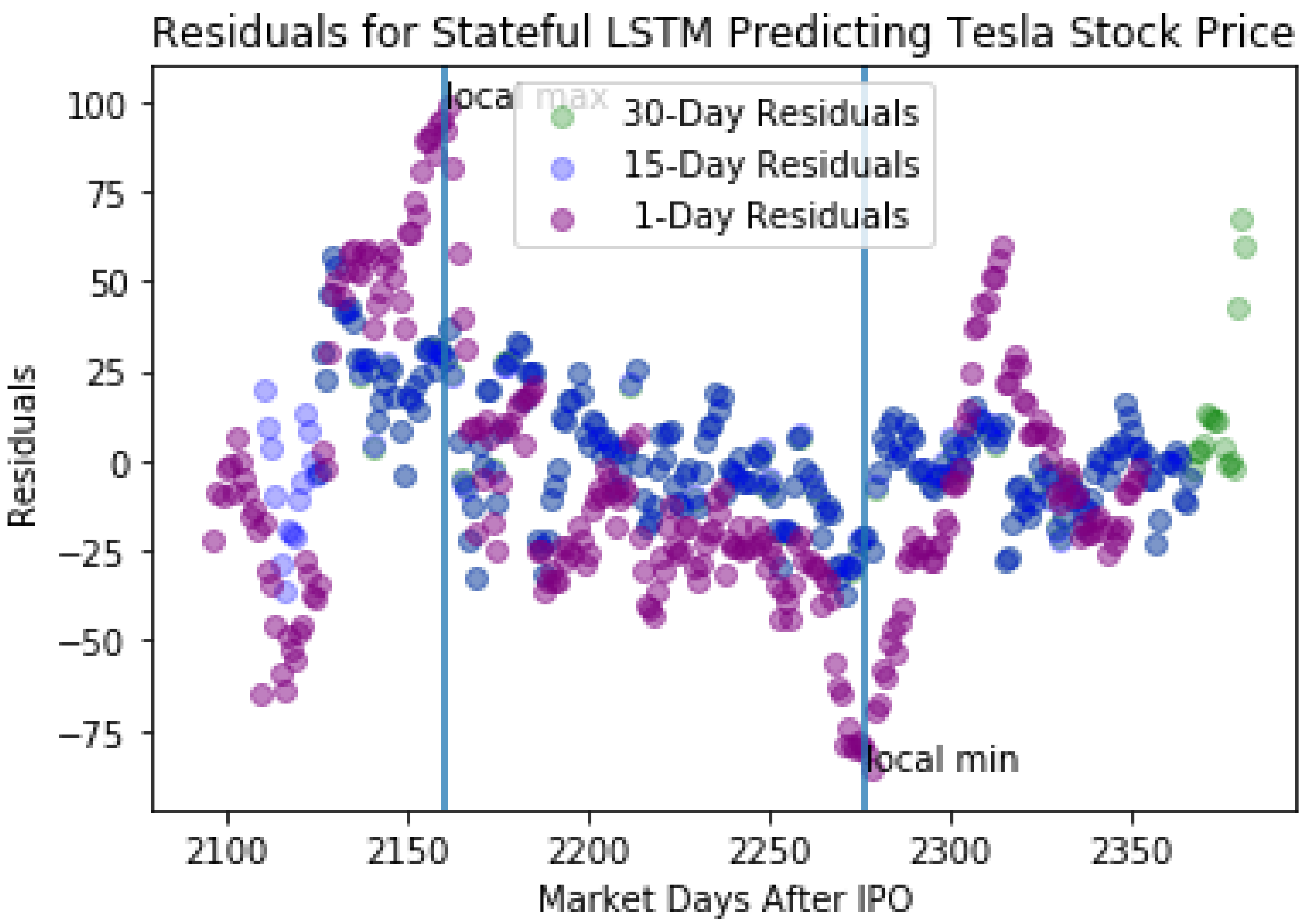
- Long Short-Term Memory (LSTM)
- Stateful
- Batch Size = 64
- Timesteps = 32
- Epochs = 120
- Data must be in full batches of 64, and each batch is used to predict 32 market days into the future. So we have four types of cases:
  - Scraps: data at beginning of dataset we cannot put in a batch
  - Training: Data to build our model
  - Test: Data we can test our model on
  - Future: Data to make predictions that cannot be verified until we get more data


Predicted vs Real Prices
1-day prediction
30-day prediction
Real Tesla Stock Price
Tesla Stock Price
100
200
300
400
500
750
1000
1250
1500
1750
2000
2250
Market Days After IPO
training
testing



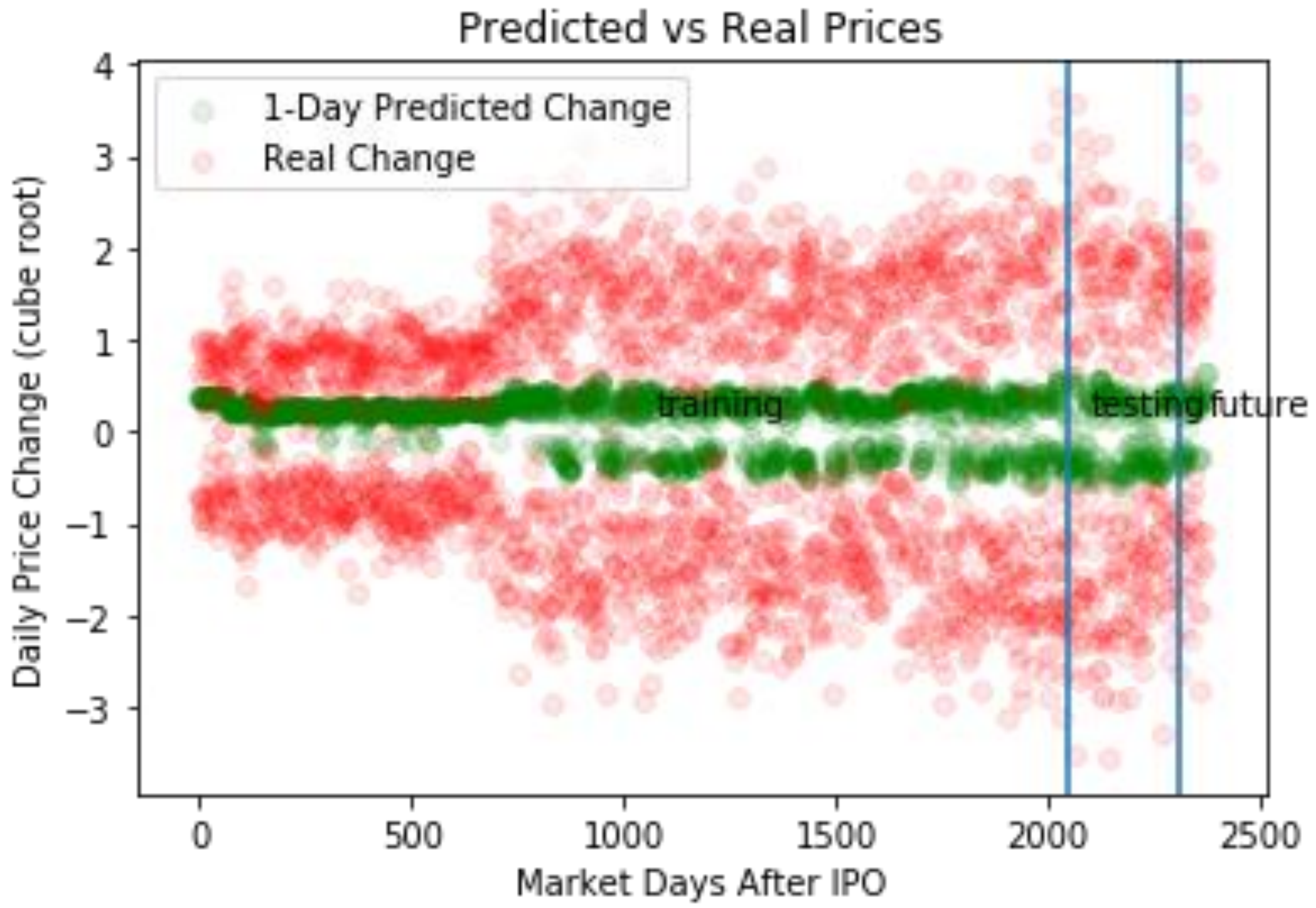
Closing Price Prediction:

- 30-Day R^2 = 0.8540
- 15-Day R^2 = 0.8599
- 1-Day R^2 = 0.2994
- Good 30 and 15 predictions
- Unreliable for next-day predictions
  - Likely need hourly data to predict next-day price



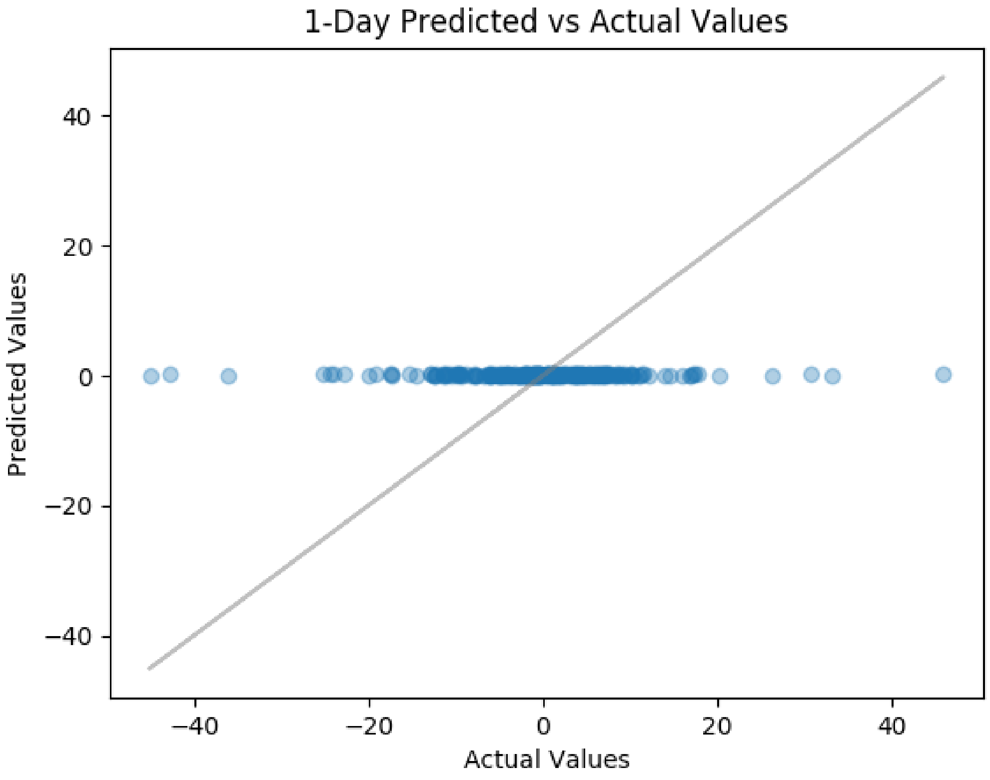
What do residuals tell us?

- Residuals for 1-Day prediction not random
  - Dropout layer may help
- Added two lines for local maximum and minimum of actual data
- The 1-day prediction model the worst—lemming-like predictions
- Better as tool for longer-term investing.
  - Model of lemming behavior can still be useful
    - Try to out-lemming the lemmings
    - Or take advantage of lemming behavior



Daily Price Prediction

- Was curious how this would perform
- Does not look like we have the data to predict one-day price changes



Checking Residuals for Daily Change Prediction

- Model seems to know it cannot predict next-day price changes
- Additional data may lead to different results
- Perhaps best not to try predicting one-day changes
- Could try to simply predict whether stock would go up or down on a given day
- 30-Day R2 = -0.00079
- 15-Day R2 = -0.00135
- 1-Day R2 = -0.00231

Consideration on Models and Data Used

- Regression learning algorithm using only prices is good for summaries, but not predictions
- Machine learning algorithm works fairly well predicting closing prices 3-6 weeks in the future
  - Could improve performance for near-future predictions with finer-grained data
- For daily change, the model is not useful
- Would be use model with additional data added.