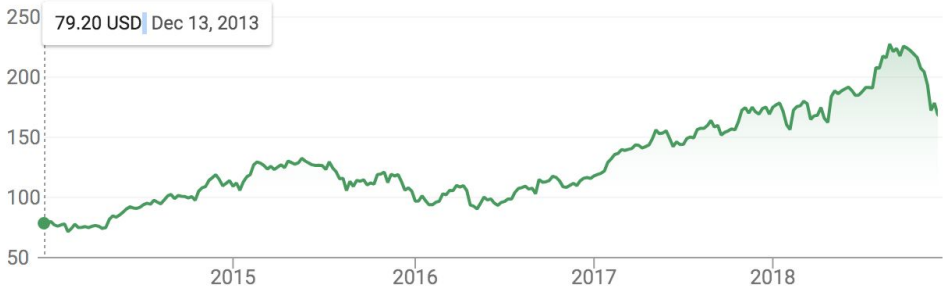
Bitcoin Stock Price Prediction

Manthan Thakkar & Bharat Vaidhyanathan

The Problem

- Predicting stock market is very very hard
- Time series data
- Stock market also affected by outside news

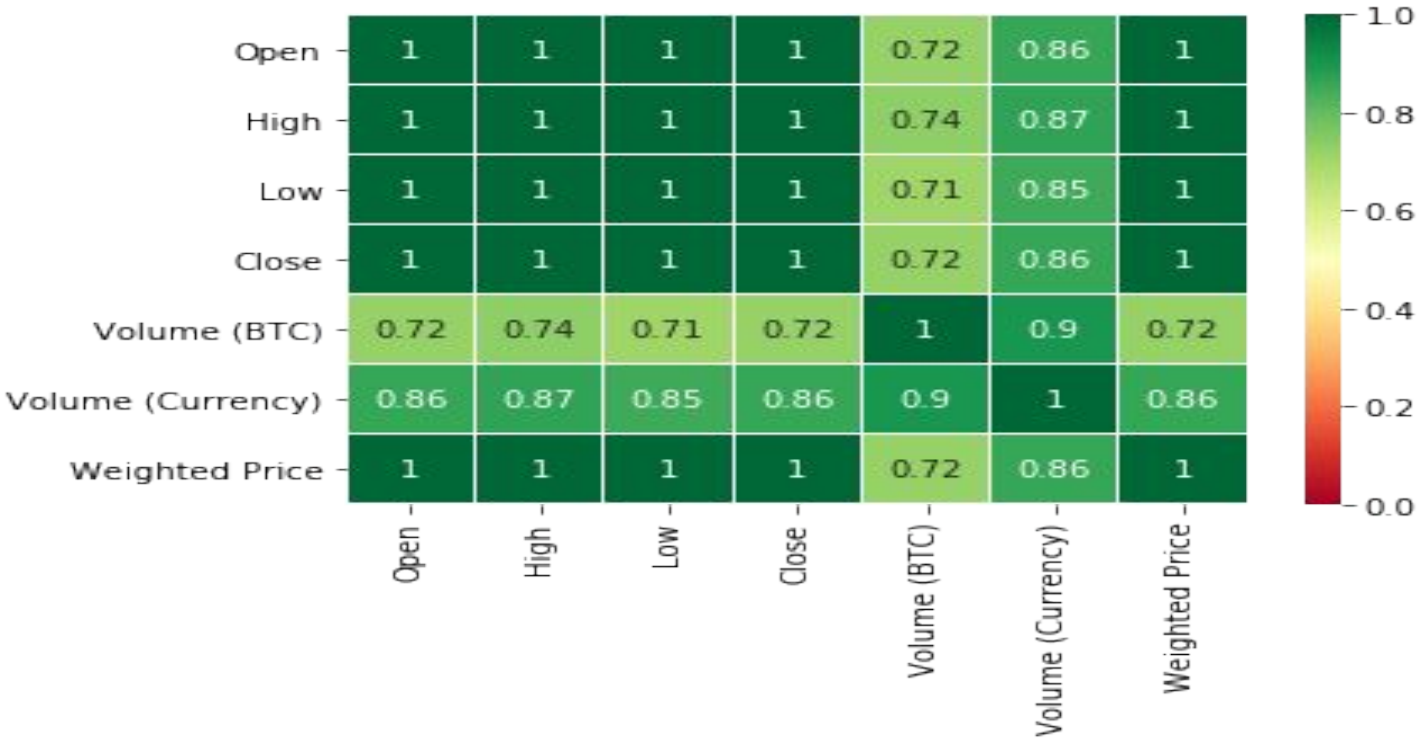
Why Bitcoin?



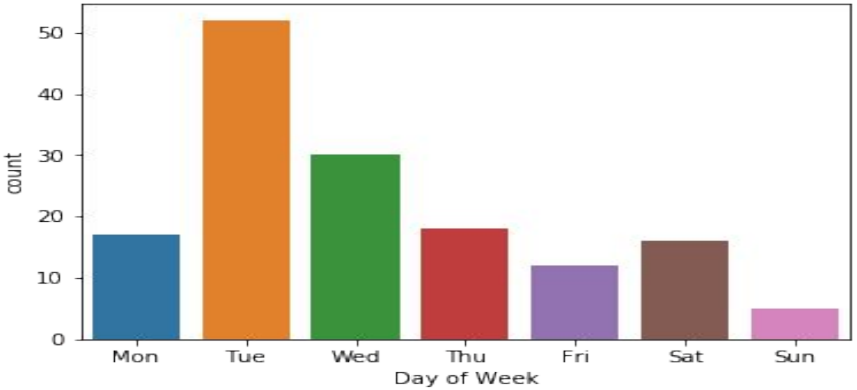
Apple:
$75 to $230

Bitcoin:
$300 to $20K!

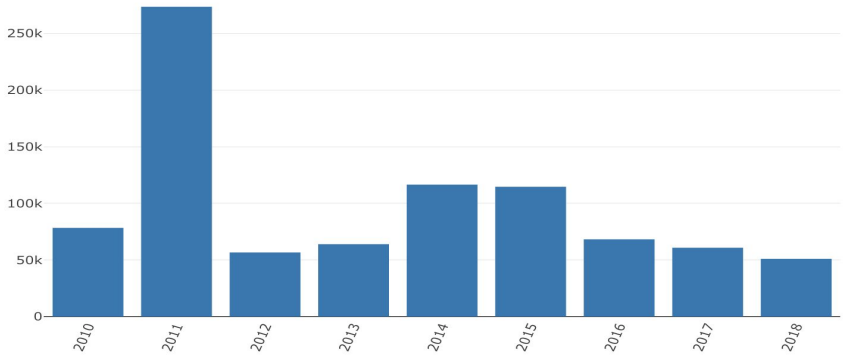
Challenge: Why is it difficult?



Challenge: Why is it difficult?



Single Highest Valued Transaction Average Per Year



- Seasonality
- Volatile market
- Time Series
- Random Walk

Literature Overview

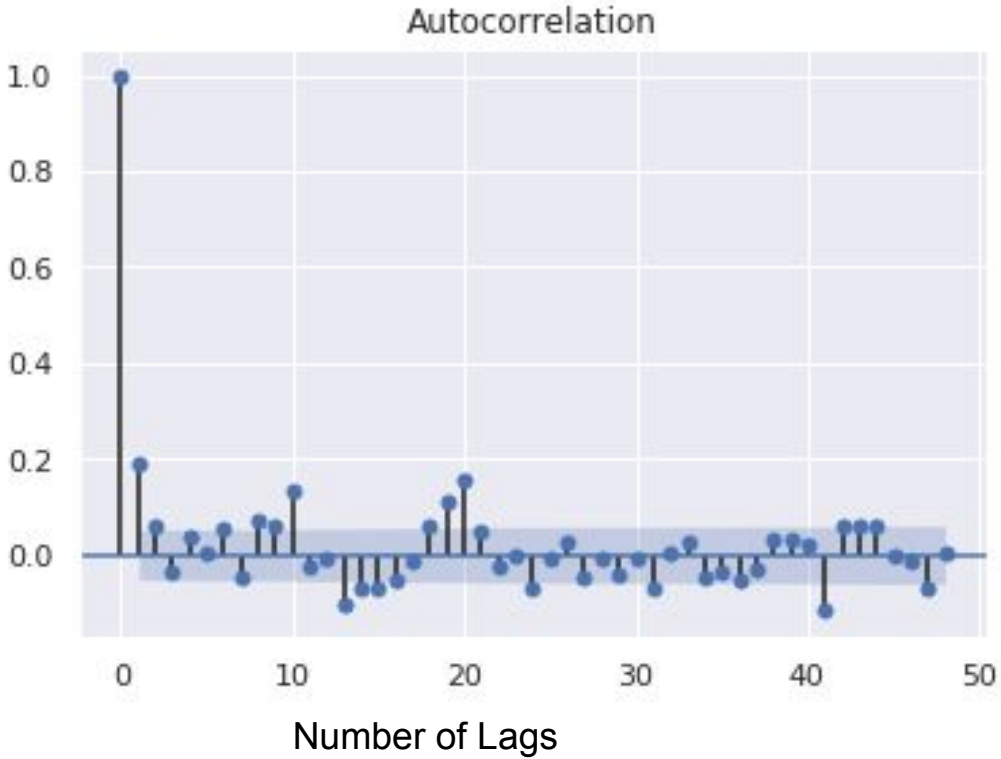
Financial models ARIMA, AR,MA, speech recognition

- **AR**: *Autoregression* dependent relationship between some number of lagged observations.
- **I**: *Integrated*. To make the data stationary
- **MA**: *Moving Average*. A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Feature Engineering

- Time series data:
  - Windowing
  - Shifting
  - Forward & Backward Fill
- News Parser:
  - NLTK tokenizer to cleanse the data
  - Doc2Vec and Word2Vec for word embeddings

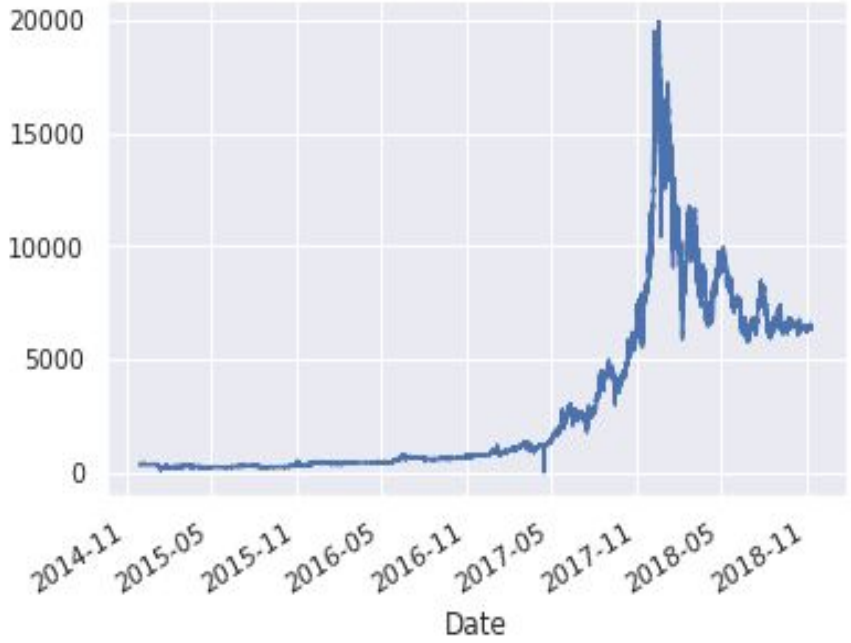
How many days is today's price dependent on?

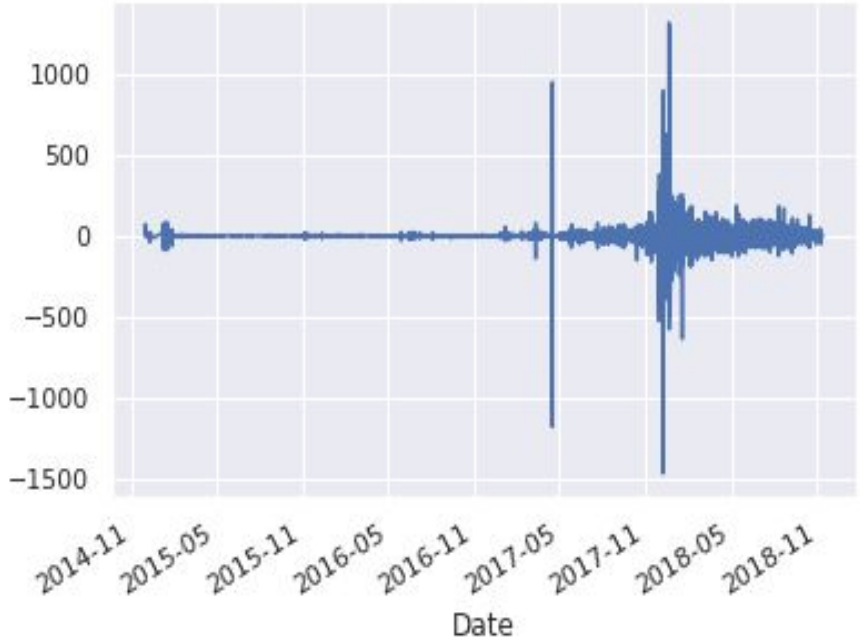


Autocorrelation : similarity between observations as a function of the time lag between them.

Used extensively on Wall Street to make money

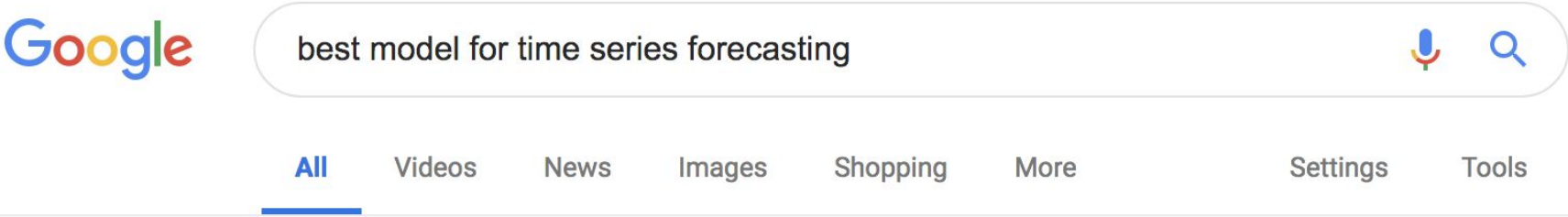
Feature Engineering





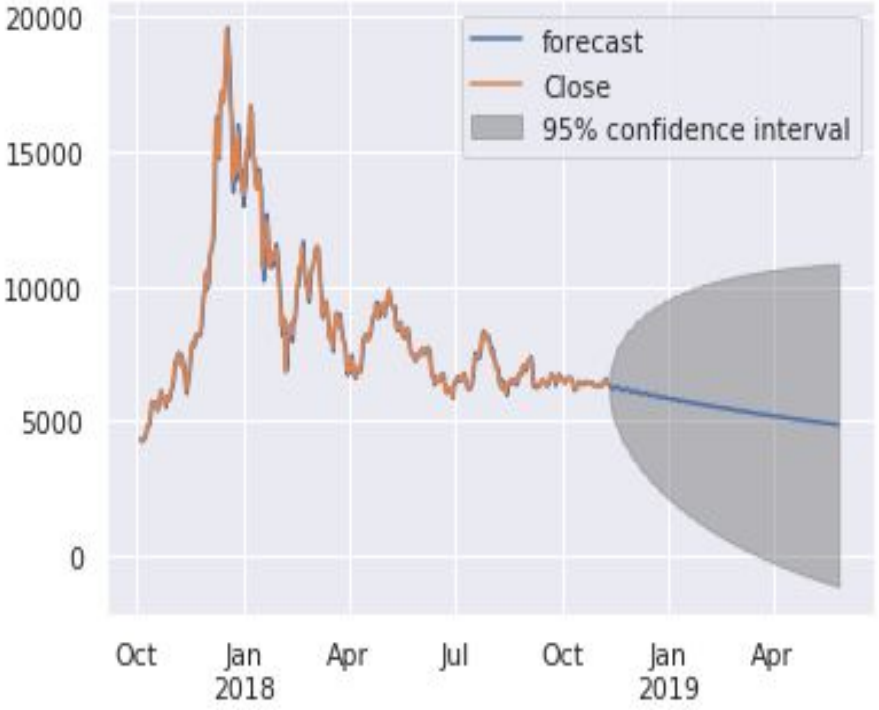
Stationary

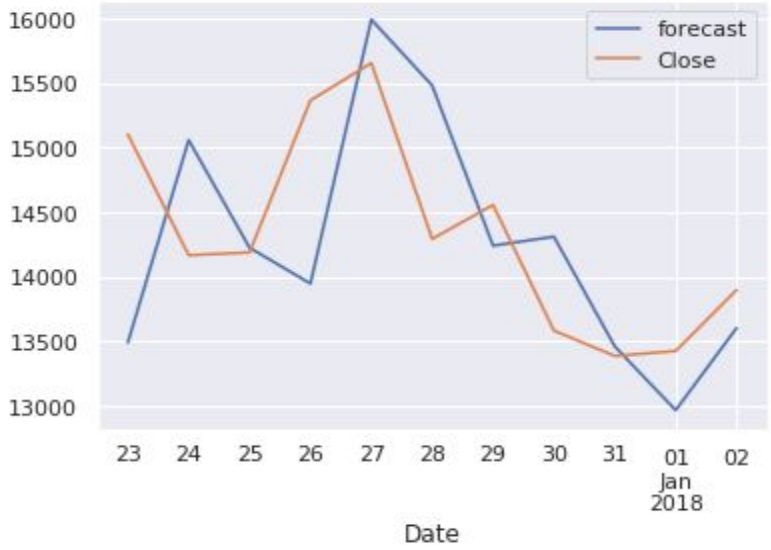
Hmm, so what next?



Baseline: ARIMA

Data: last 7 days **Predict:** Today's Price **RMSE: $395.24 (Nov 2017 to 2018)**

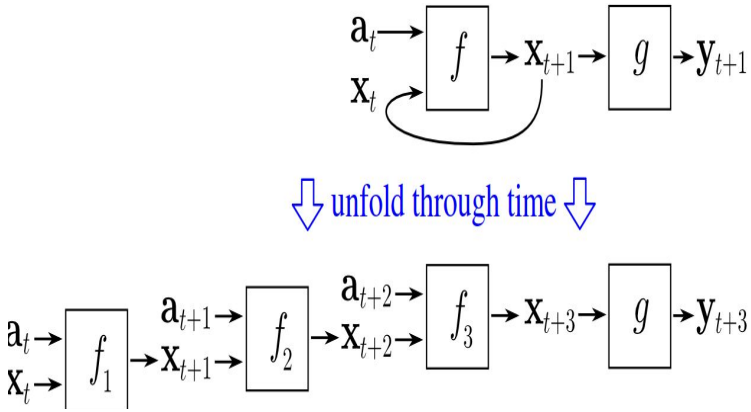




BitCoin Crash in Dec 2017 (Lag)

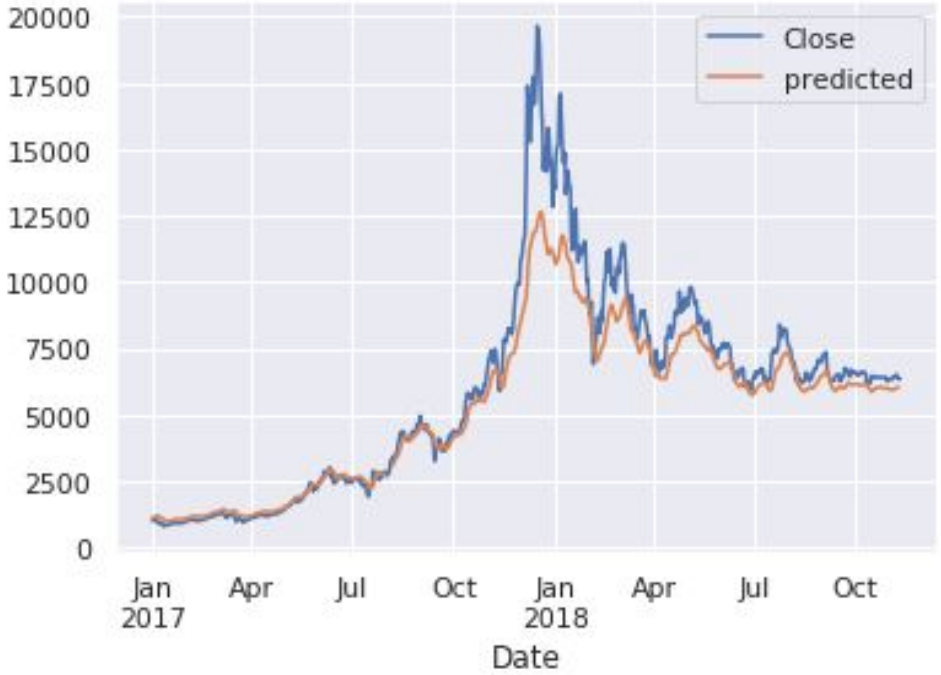
Deep RNN (Recurrent Neural Networks) with LSTM

Data: Price for last 7 days Predict: Price: today



2 Layers with 230 LSTM

RMSE: $358.11



Bitcoin Gold rush in Dec 2017

Do cryptocurrency news affect bitcoin prices?

$3K Ahead? Bitcoin Price Bounce Is Again Losing Steam

Dec 10, 2018 at 11:00 | Omkar Godbole

Bitcoin still has the potential to drop towards $3,000, despite a minor bounce from 15-month lows seen on Friday.

Bitcoin's Share of the Crypto Market Is Nearing a 3-Month High

Dec 10, 2018 at 16:25 | Sam Ouimet

Bitcoin's share in the cryptocurrency market reached an 11-week high late last week.

Bitcoin Price Consolidates Sub-$3.5K With Bulls and Bears in Stalemate

Dec 13, 2018 at 11:30 | Omkar Godbole

Bitcoin is consolidating below $3,500 for the third day straight, but is a bull move building?

Cheap Power Is Luring Battered Bitcoin Miners to Iran

Dec 12, 2018 at 10:00 | Wolfie Zhao

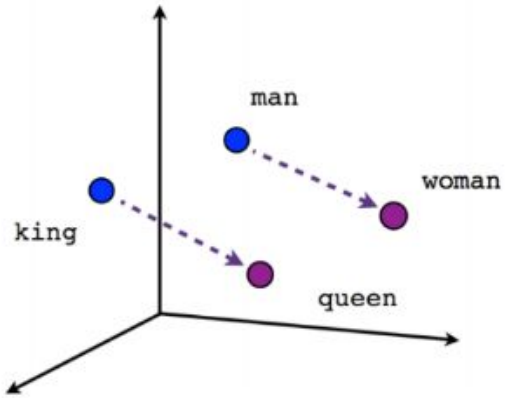
Some crypto miners are looking to Iran for its low-cost power – but the path to setting up shop in the country is anything but simple.

Aggregating and Parsing News

- Manually parsed news for the past 6 years
- Filter specific to “Bitcoin”, “Crypto”, and more related terms.
- Aggregated for each day

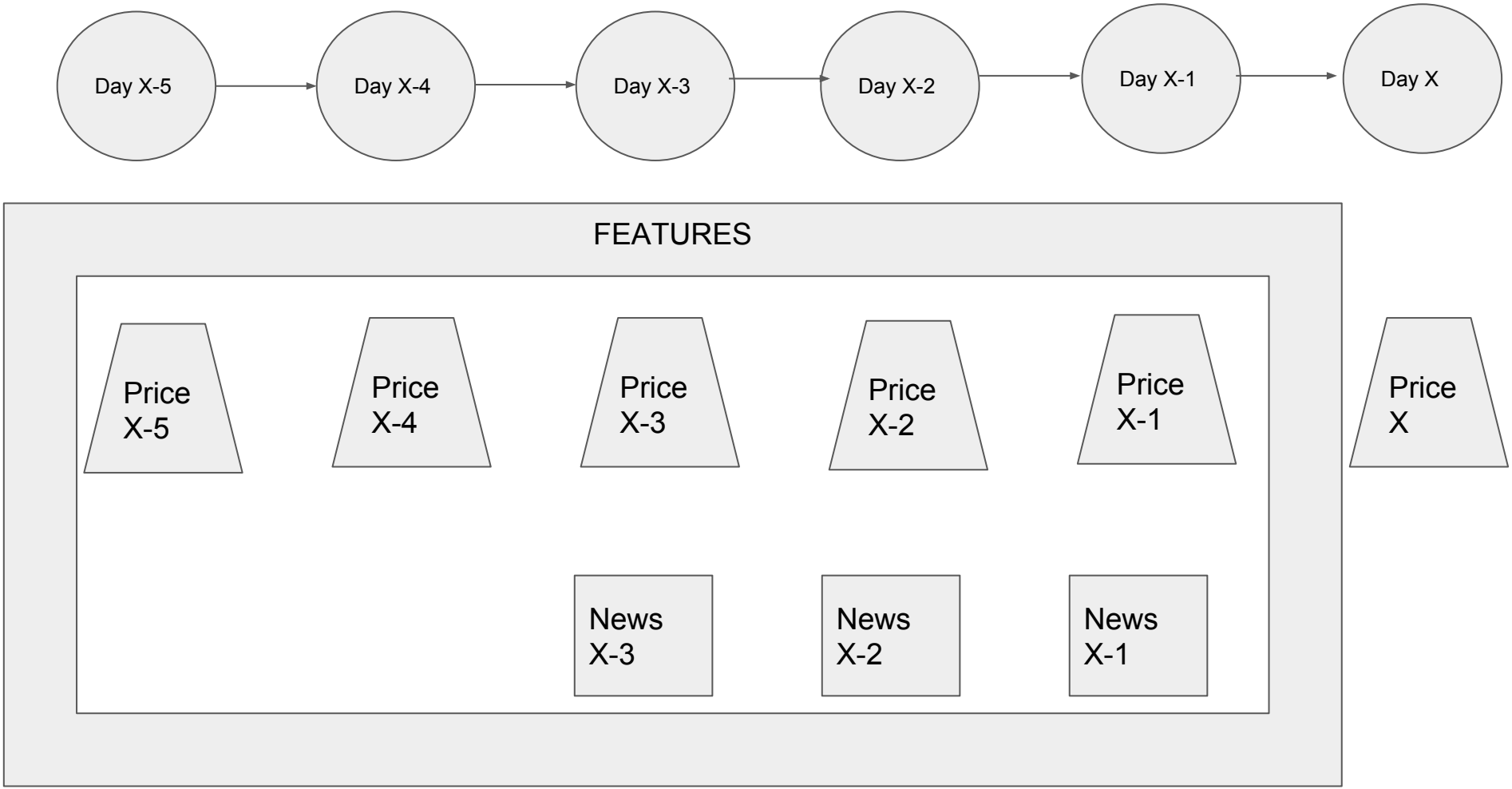
Converting Text to Vectors

- Tokenize sentences to words
- Used word embeddings to represent words in a vectorial space.
- Doc2Vec for vectorial representation of a paragraph.
- 300 dimensional vector



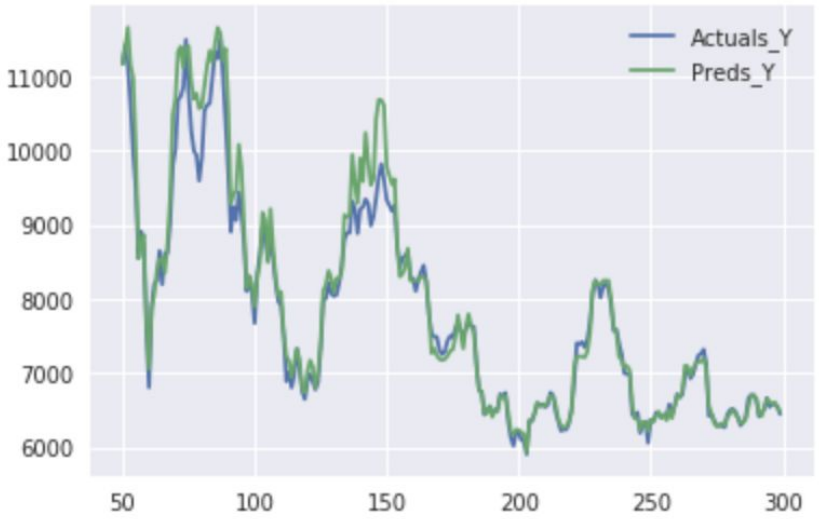
Male-Female

```
Date                                                2013-04-16
Headline    Bitcoin Investors Hang On for the Ride Bits an...
Vec         [-3.578344e-05, 0.0016333716, -0.0001520646, 0...
```

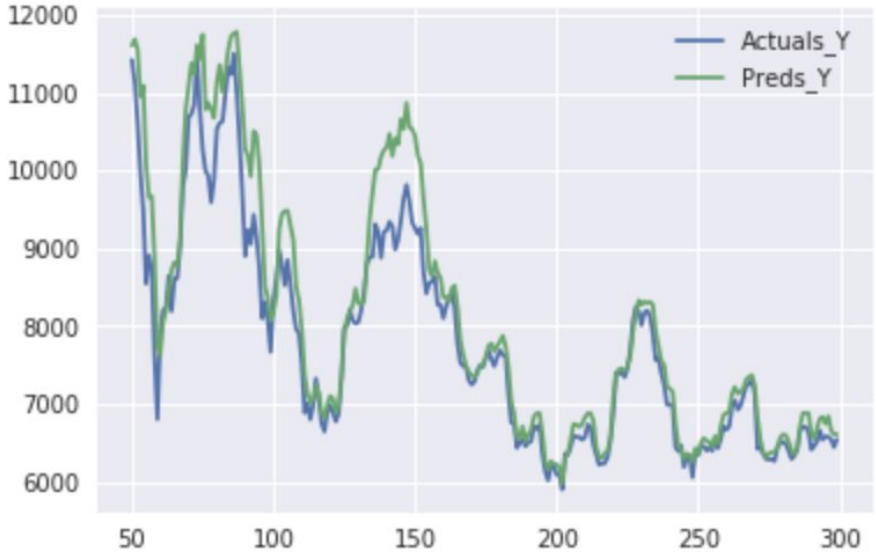

Day X-5
Day X-4
Day X-3
Day X-2
Day X-1
Day X
FEATURES
Price X-5
Price X-4
Price X-3
Price X-2
Price X-1
Price X
News X-3
News X-2
News X-1

The Model

Random Forest Regressor (RMSE)



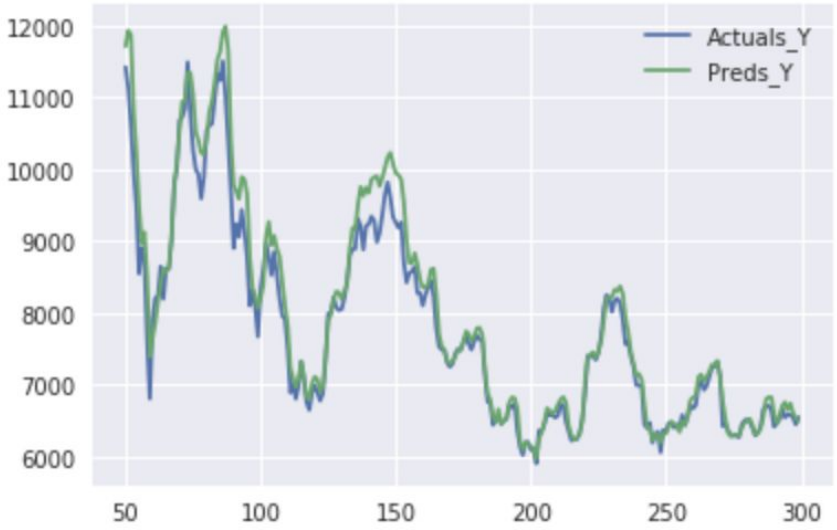
Without News: $406.23



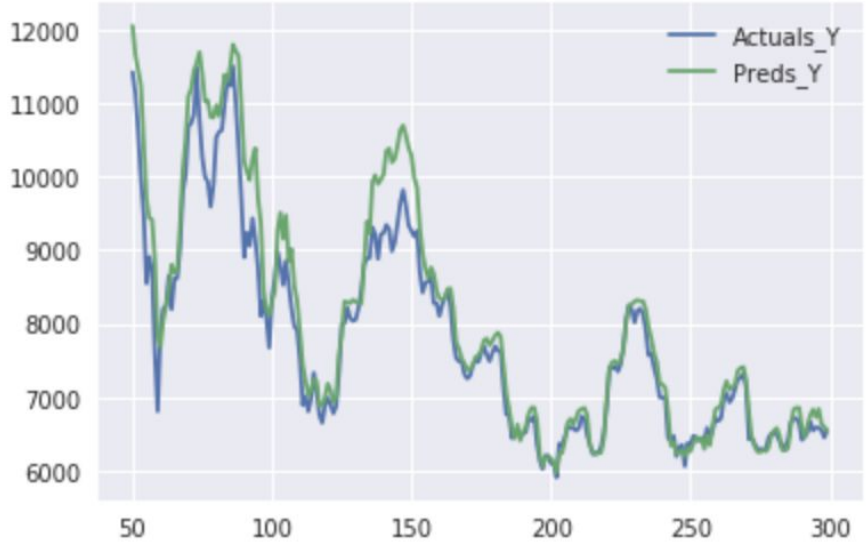
With News: $849.61

The Model

Recurrent Neural Net (RMSE)



Without News: $271.62



With News: $539.85

Results

Model	RMSE
ARIMA	$395.24
Random Forest	$406.23
Random Forest with News	$849.61
RNN	**$271.62**
RNN with News	$539.85

Conclusions and Future Work

- Price prediction is tough, especially predicting spikes
- Use a hybrid model [RNN + RF(withNews)]:
  - Moving Price prediction from RNN
  - Spike predictions from the News
  - RL for making a bot using financial trading strategies
- External factors other than news

"The four most dangerous words in investing are: "This time it's different"."

Sir John Templeton