Effect of President’s tweets on the S&P 500 index

Sid Assawaworrarit
sca@Stanford.edu
Department of Electrical Engineering, Stanford University

Summary

Can a machine be trained to predict the likely stock market index movement from President Trump’s tweets? Here we build machine-learning models that predict the likely movement of the S&P 500 index-tracker SPY in response to a tweet released by President Trump. Models based on LSTM with word-to-vector encoder and on neural network with sentence-to-vector encoder are presented. We train the models on President Trump’s tweet record from 2017 to late 2019 and find that the trained models yield no significant difference in their prediction errors compared to a baseline model which always predicts no change. The lack of accuracy improvement from the machine learning models likely stems from the weakness or nonexistence of any direct relationship between Trump’s tweet text and SPY movement.

Data & Features

Data sources:
1. Trump tweet archive: www.trumptwitterarchive.com
2. Wharton Research Data Services (WRDS): TAQ (consolidated trades) – raw SPY traded price history data

Apply resampling to the raw SPY price data by averaging trade prices over 1 minute.

Tweet and SPY data sets are combined and reformulated into input: tweet; output: SPY percentage change 15 min after tweet.

Train/Validation/Test: 1145/200/300

Models

Apply transfer learning from learned word and sentence encoders: glove.twitter.27B.50d and universal sentence encoder (USE) USE + NN

Text-only with GloVe + LSTM

Input: list of words

Embedding: glove.twitter.27B.50d

LSTM (4 units, ret seq)

Dropout (0.25)

LSTM (4 units)

Dropout (0.25)

FC (4 units, sigmoid)

Dropout (0.1)

FC (1 unit, linear)

Output: prediction

Text and price history with GloVe + LSTM

Input: list of words with price history

Embedding: glove.twitter.27B.50d

LSTM (4 units, ret seq)

Dropout (0.25)

LSTM (4 units)

Dropout (0.25)

LSTM (4 units)

Dropout (0.1)

LSTM (4 units)

Dropout (0.1)

FC (4 units, sigmoid)

Dropout (0.2)

FC (1 unit, linear)

Output: prediction

Results

- None of the constructed models yield significantly lower prediction error than the baseline model for the test data set.
- All trained models output predictions that are typically within a ±2 basis points.

Discussion & Future Work

- Our results suggest that the direct effect of Trump’s tweets on SPY movements is limited to non-existent.
- Larger, more complicated models with additional inputs such as tweets and news from all sources could give lower prediction errors. However, it would be difficult to separate the effect belonging to Trump tweet for such a model.
- Select other stocks or indices as predictor output instead of SPY. The effect of president’s tweets may show up noticeably in the movements of other stocks or sector-specific indices, such as manufacturing, which tend to be more volatile than the S&P 500.
- Data augmentation may help with training the models. Available data such as Trump’s tweets outside market hours and tweets from other sources may lead to prediction improvement if one can realize a way to integrate these data into the training of the models.

References

- Daniel Car et al. Universal sentence encoder. arXiv, 2018

Table 1: Sample look at the data sets

<table>
<thead>
<tr>
<th>Time (ET)</th>
<th>Input (tweet)</th>
<th>Output (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-08-03</td>
<td></td>
<td>3.16%</td>
</tr>
<tr>
<td>09:30:44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019-08-01</td>
<td></td>
<td>3.267</td>
</tr>
<tr>
<td>15:00:10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019-11-27</td>
<td></td>
<td>0.1112</td>
</tr>
<tr>
<td>09:31:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019-11-27</td>
<td></td>
<td>0.0281</td>
</tr>
<tr>
<td>14:00:38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019-12-24</td>
<td></td>
<td>0.0218</td>
</tr>
<tr>
<td>09:31:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019-08-05</td>
<td></td>
<td>0.0157</td>
</tr>
<tr>
<td>12:00:14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: SPY changes relative to the time Trump tweets are released. Upper: 100 random traces Lower: all training
Video link: https://youtu.be/lD31uveVypU