Investing in SPY ETF: Deep Learning on SPY Constituents’ Momentum Data
Dale Angus
dangus@stanford.edu

Introduction
Could an ETF constituents’ price and related information derived from a series of prices, e.g., momentum, be used to reliably predict the ETF’s price outcome? Two deep learning models were used to answer this question.

- Model A predicts outcome as either up or down
- Model B predicts outcome as falling within a defined percentage range

Dataset
- The SPY ETF has 504 constituents.
- The dataset is composed of the 2-year historical data, around 527 trading days.
- The dataset has around 265,600 rows.
- The dataset is split 90-10 between the Training set (239,040 rows) and Test set (26,560 rows)
- The price data was gathered from Yahoo Finance. The other features are computed.

Features
The 16 features are derived from the following:
- Weighting and Sector
- Open*, High*, Low*, Close
- Simple Moving Averages, 10 and 20-day*
- Linear Regression Slopes, 10 and 20-day*
- Awesome Oscillator* and 34-day Momentum*
- A derivative of the 21-day 2nd degree Polynomial Regression*
- Price ratio information of related ETFs, i.e., DIA, QQQ, IWM
- Value of the VIX index

*normalized by dividing by the close price

Models
<table>
<thead>
<tr>
<th>Type</th>
<th>Model A</th>
<th>Model B</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-Layer Binary Classifier</td>
<td>3-Layer Softmax Classifier</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Model A</th>
<th>Model B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer Dimensions</td>
<td>16, 15, 10, 9, 5, 4, 1/13, 12, 8, 6, 4, 2, 1</td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>0 or 1</td>
<td></td>
</tr>
<tr>
<td>Framework Used</td>
<td>n/a (NumPy)</td>
<td></td>
</tr>
<tr>
<td>Learning Rate</td>
<td>.0075</td>
<td>.0001</td>
</tr>
<tr>
<td>Iterations</td>
<td>100,000 or until difference between costs is &lt;= 7x10^-5</td>
<td>1500</td>
</tr>
</tbody>
</table>

Cost function for both models

$$ J = \frac{1}{m} \sum_{i=1}^{m} (y^{(i)} \log (\sigma(z^{(i)})) + (1 - y^{(i)}) \log (1 - \sigma(z^{(i)}))) $$

Both use the Xavier Initialization.

1 with/without related ETF information

Results

<table>
<thead>
<tr>
<th>Training Accuracy</th>
<th>91.07%</th>
<th>98.3%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Accuracy</td>
<td>90.89%</td>
<td>98.1%</td>
</tr>
</tbody>
</table>

Adding other related price ratios to the set of momentum features dramatically improved the accuracy of the two models. This could be due to the high correlations between ETFs.

Use Case and Future Work
In production, using real-time data as substitute for closing data, these can be used to make trading decisions before the market closes.

Future improvements could be done in finding better features; and to predict the next day’s outcome, next 2 days and so on. This can be replicated for other ETFs.