Modernized Crypto Assets

Reducing Volatility to Produce an ETF Style Commodity for Crypto

STANFORD GRADUATE SCHOOL OF BUSINESS | Center for Entrepreneurial Studies

Quick Refresher

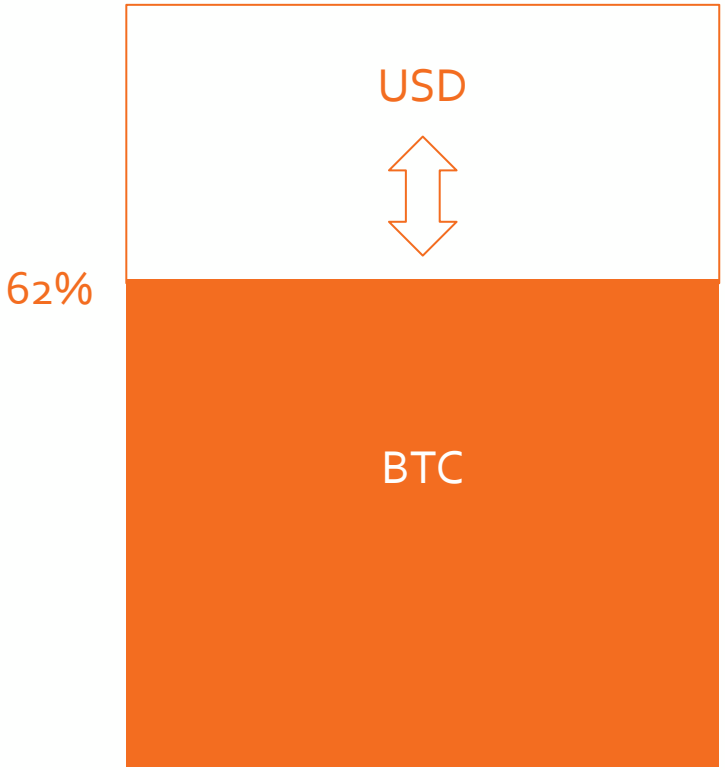
- Problem
- Project

Problem - Refresher

- Crypto is massively volatile
- No crypto asset w. smooth index effect*
  - Bundles are too highly correlated
  - Contain undesirable assets

S&P 500 Chart. Unlike stocks no crypto asset exists which smoothly tracks the movement of the market.

Project - Refresher



- Automated money management for crypto
- Intelligently moves $ between USD & crypto
- Works for all major currencies
- Money moves based upon prediction algorithms

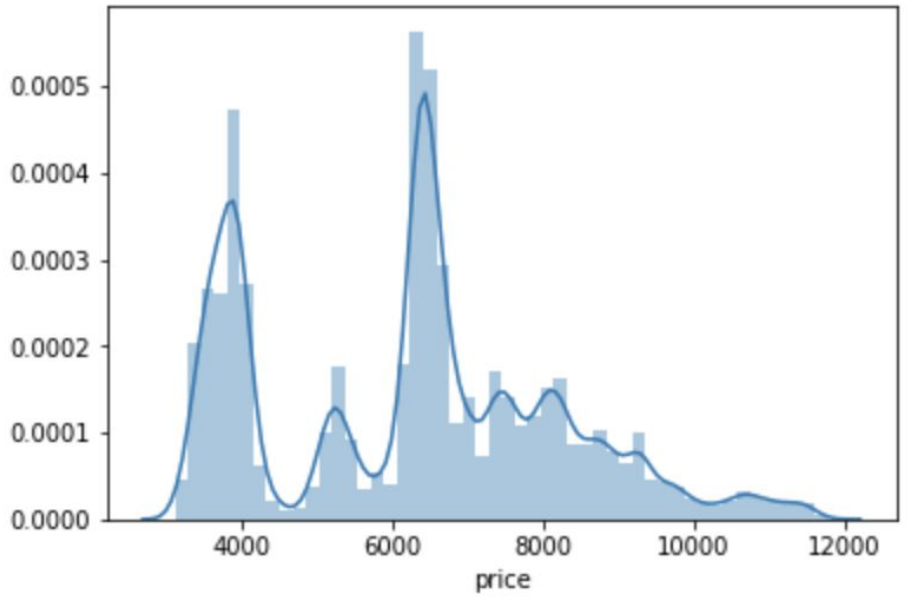
Presentation Structure

- Data
- Sentiment
- RNN Training
- RNN Analyzing
- RNN Trading

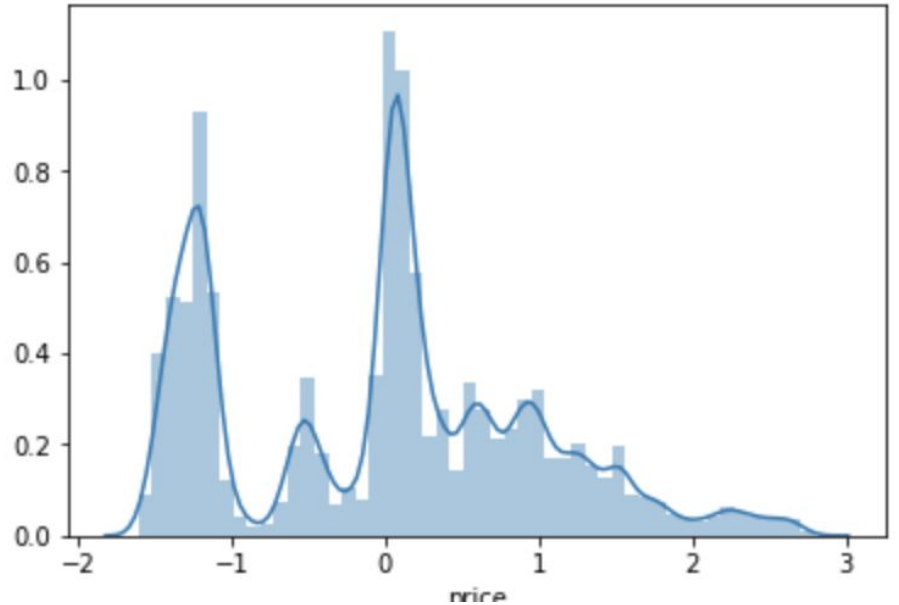
Data Exploration -> Bitcoin

Data Exploration -> Bitcoin

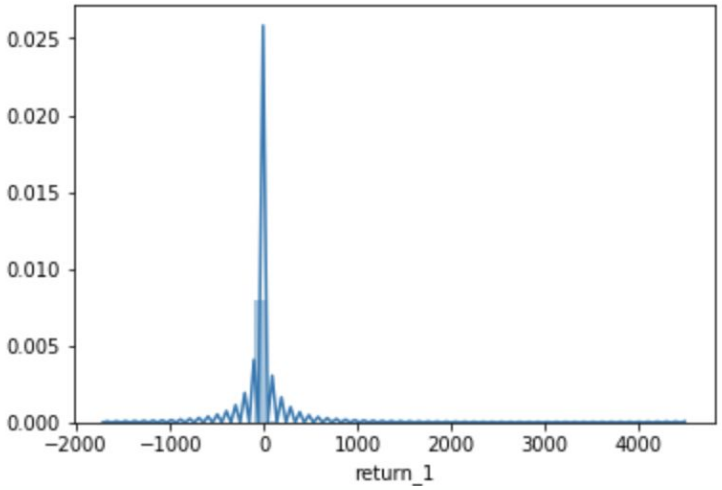
```
Plotting price distribution
```

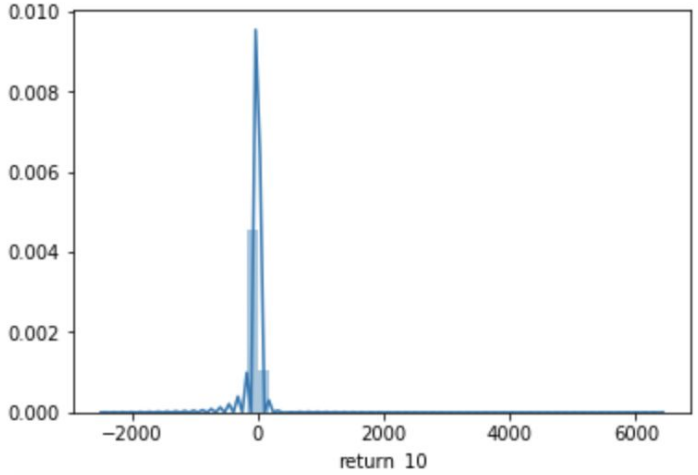


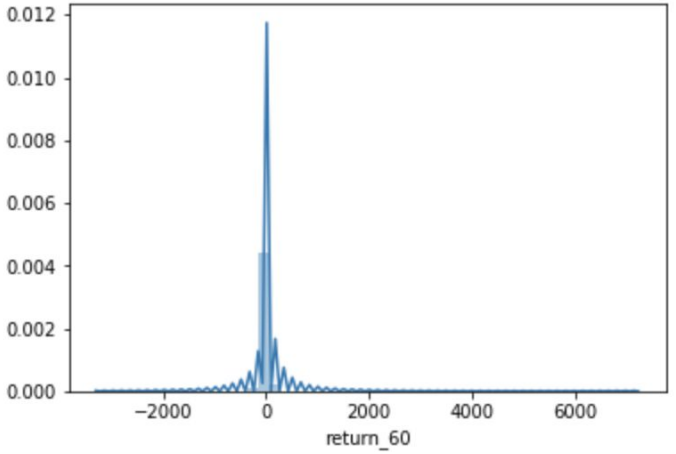

```
Plotting distribution of price z-scores
```

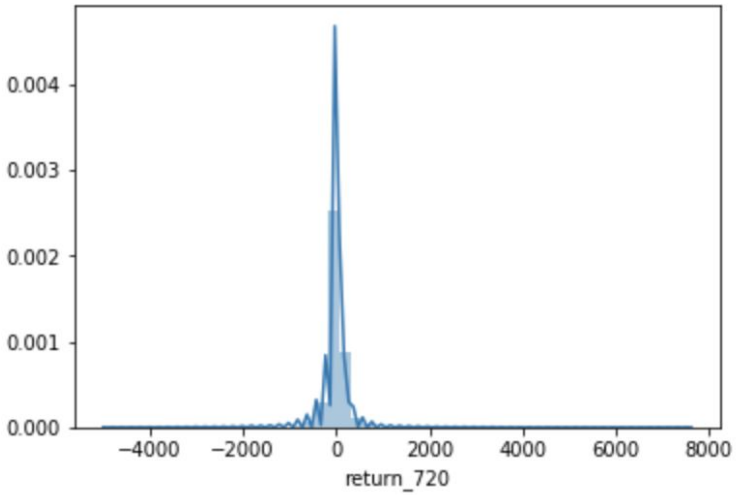


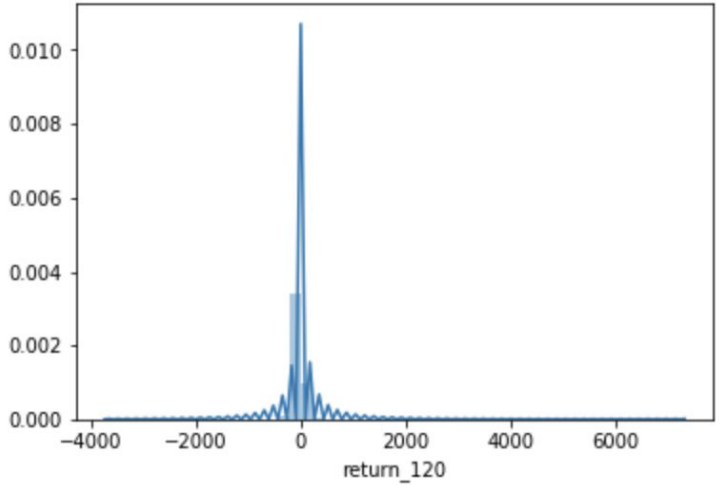
Bitcoin -> Return Distribution

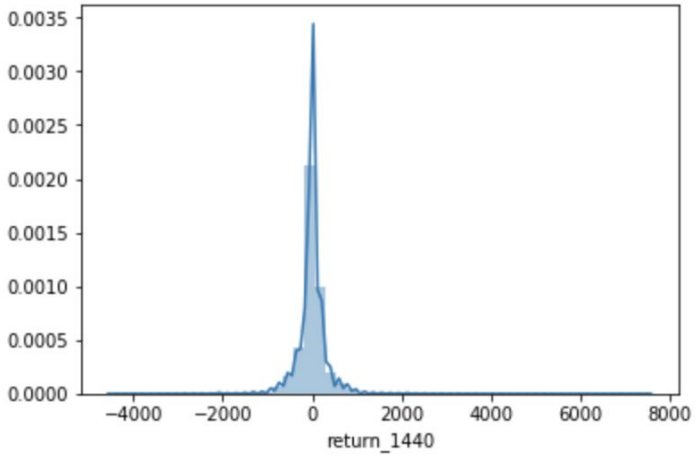






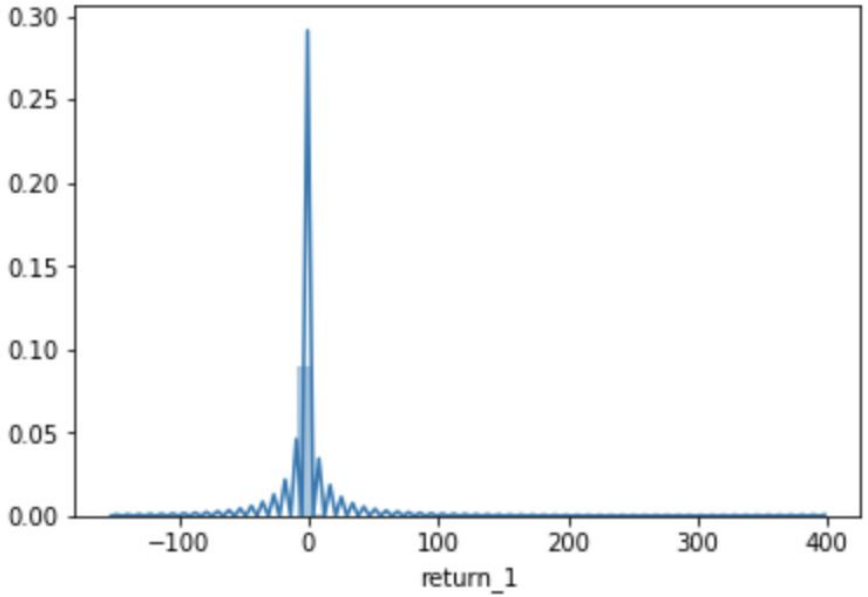






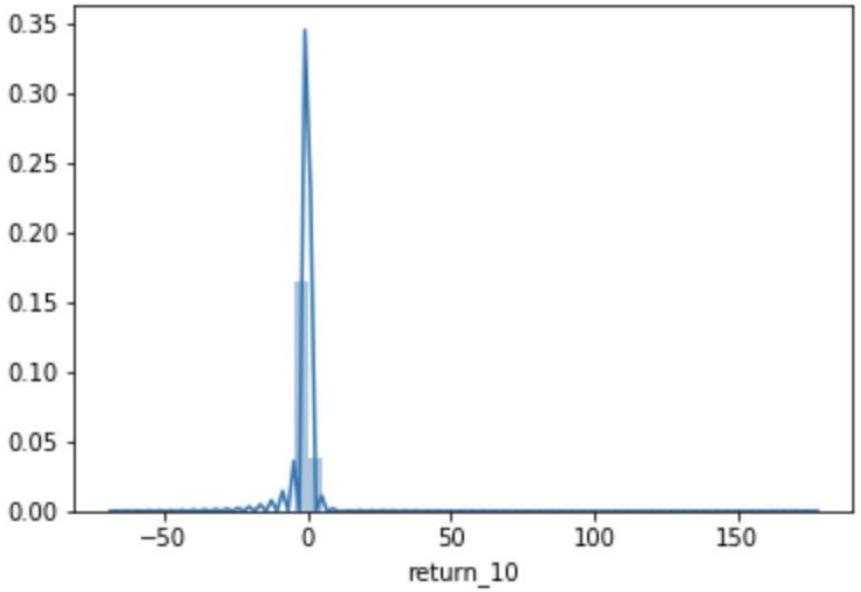
Bitcoin -> Z-score Return Distribution

```
When returns were calculated using return_1:
There were 679429 possible events
There were 29170 positive and 29627 negative 1-sigma events
There were 7185 positive and 7506 negative 2-sigma events
There were 2576 positive and 2815 negative 3-sigma events
There were 567 positive and 623 negative 5-sigma events
There were 117 positive and 125 negative 10-sigma events
There were 11 positive and 12 negative 50-sigma events
There were 3 positive and 1 negative 100-sigma events
```



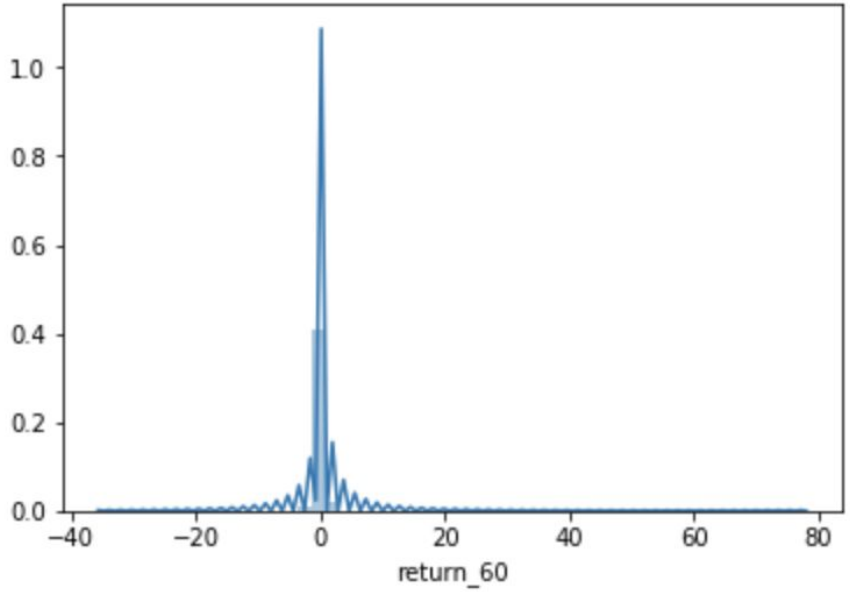
Bitcoin -> Z-score Return Distribution

```
When returns were calculated using return_10:
There were 679429 possible events
There were 23156 positive and 23822 negative 1-sigma events
There were 5923 positive and 6577 negative 2-sigma events
There were 2308 positive and 2712 negative 3-sigma events
There were 771 positive and 802 negative 5-sigma events
There were 144 positive and 242 negative 10-sigma events
There were 15 positive and 9 negative 50-sigma events
There were 10 positive and 0 negative 100-sigma events
```



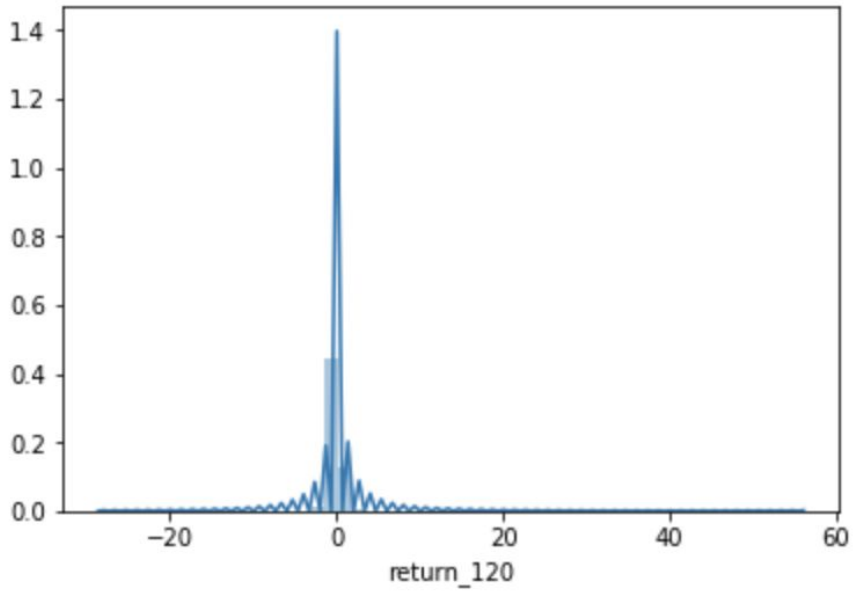
Bitcoin -> Z-score Return Distribution

```
When returns were calculated using return_60:
There were 679429 possible events
There were 24123 positive and 25595 negative 1-sigma events
There were 6226 positive and 7684 negative 2-sigma events
There were 2280 positive and 2804 negative 3-sigma events
There were 656 positive and 968 negative 5-sigma events
There were 241 positive and 378 negative 10-sigma events
There were 55 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



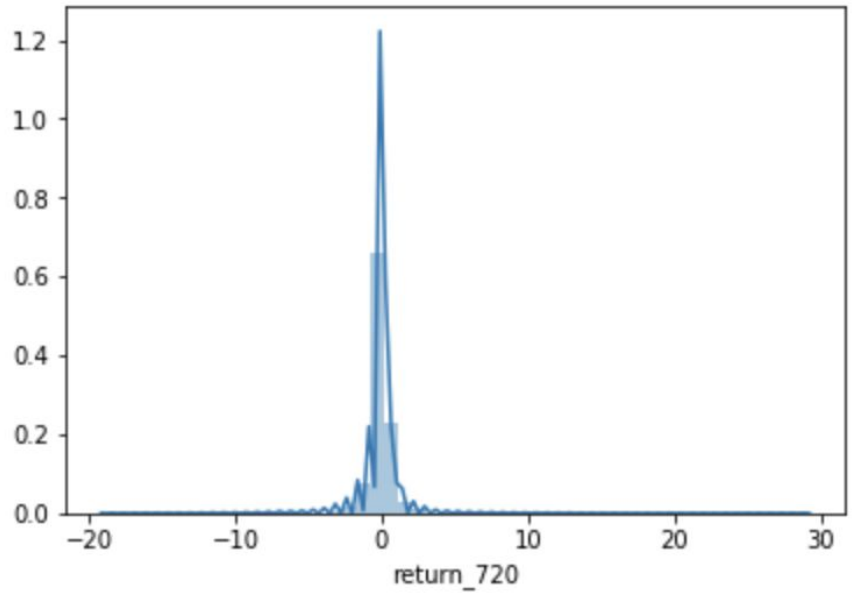
Bitcoin -> Z-score Return Distribution

```
When returns were calculated using return_120:
There were 679429 possible events
There were 25665 positive and 29084 negative 1-sigma events
There were 5984 positive and 7786 negative 2-sigma events
There were 1917 positive and 2665 negative 3-sigma events
There were 618 positive and 1080 negative 5-sigma events
There were 194 positive and 247 negative 10-sigma events
There were 68 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



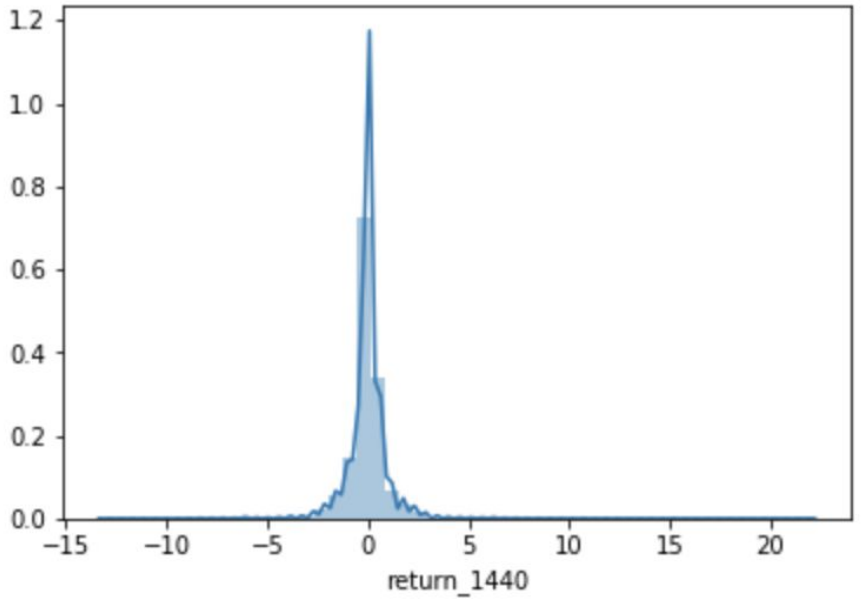
Bitcoin -> Z-score Return Distribution

```
When returns were calculated using return_720:
There were 679429 possible events
There were 40352 positive and 47558 negative 1-sigma events
There were 10057 positive and 11838 negative 2-sigma events
There were 2549 positive and 4106 negative 3-sigma events
There were 798 positive and 1078 negative 5-sigma events
There were 695 positive and 238 negative 10-sigma events
There were 0 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



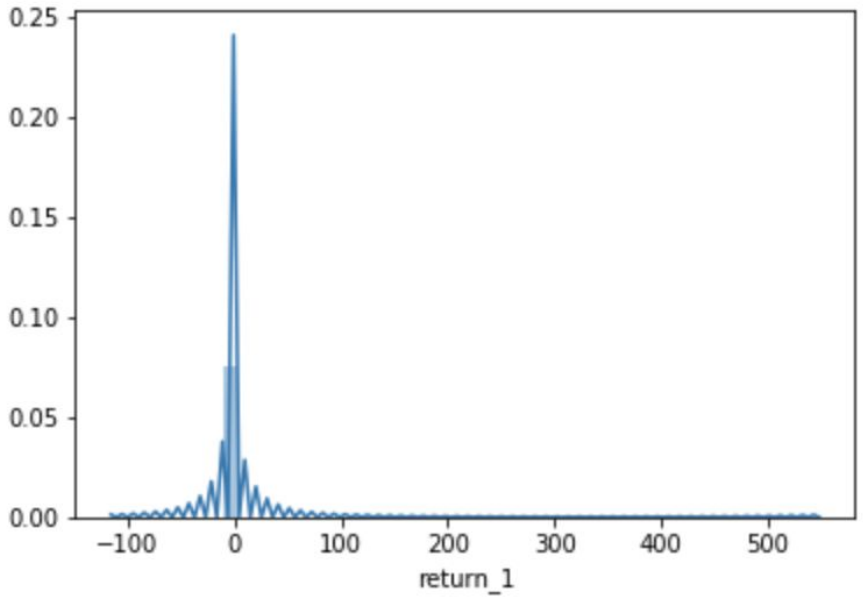
Bitcoin -> Z-score Return Distribution

```
When returns were calculated using return_1440:
There were 679429 possible events
There were 47361 positive and 59974 negative 1-sigma events
There were 14723 positive and 15790 negative 2-sigma events
There were 2539 positive and 4167 negative 3-sigma events
There were 1440 positive and 1295 negative 5-sigma events
There were 700 positive and 68 negative 10-sigma events
There were 0 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



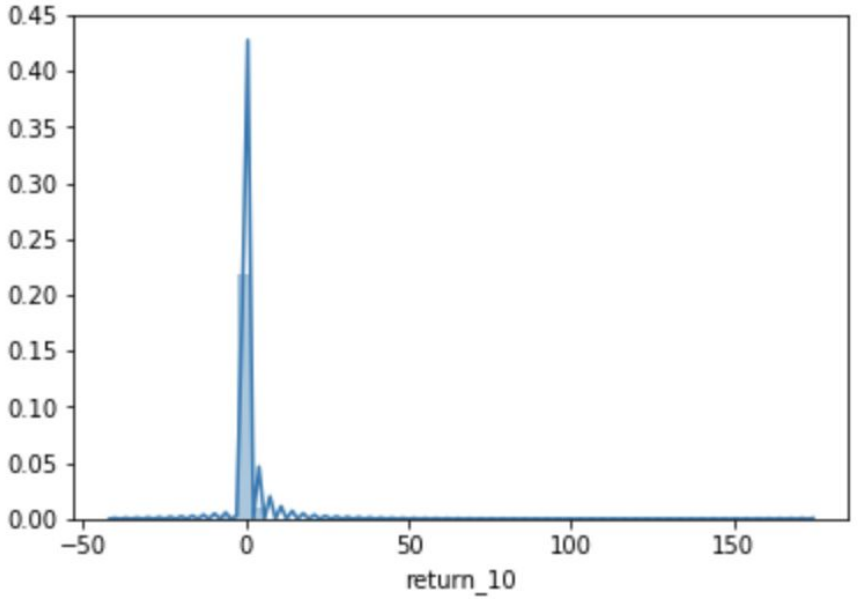
Bitcoin -> Weekly Z-score Return Distribution

```
When returns were calculated using return_1 and a weekly moving z_score:
There were 679429 possible events
There were 57932 positive and 59470 negative 1-sigma events
There were 13458 positive and 13801 negative 2-sigma events
There were 4860 positive and 4857 negative 3-sigma events
There were 1311 positive and 1384 negative 5-sigma events
There were 226 positive and 258 negative 10-sigma events
There were 4 positive and 3 negative 50-sigma events
There were 2 positive and 1 negative 100-sigma events
```



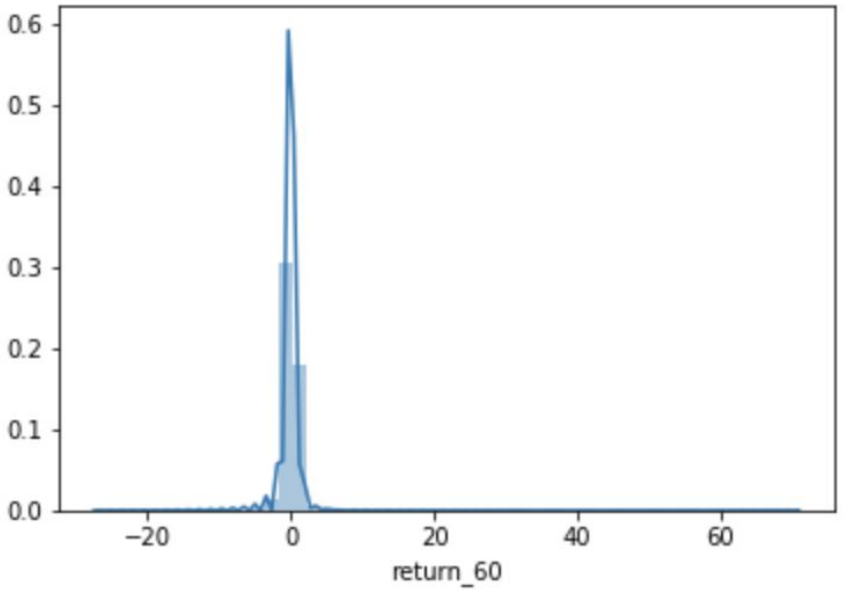
Bitcoin -> Weekly Z-score Return Distribution

```
When returns were calculated using return_10 and a weekly moving z_score:
There were 679429 possible events
There were 53001 positive and 54395 negative 1-sigma events
There were 12424 positive and 13609 negative 2-sigma events
There were 4899 positive and 5527 negative 3-sigma events
There were 1540 positive and 1734 negative 5-sigma events
There were 337 positive and 327 negative 10-sigma events
There were 6 positive and 0 negative 50-sigma events
There were 1 positive and 0 negative 100-sigma events
```



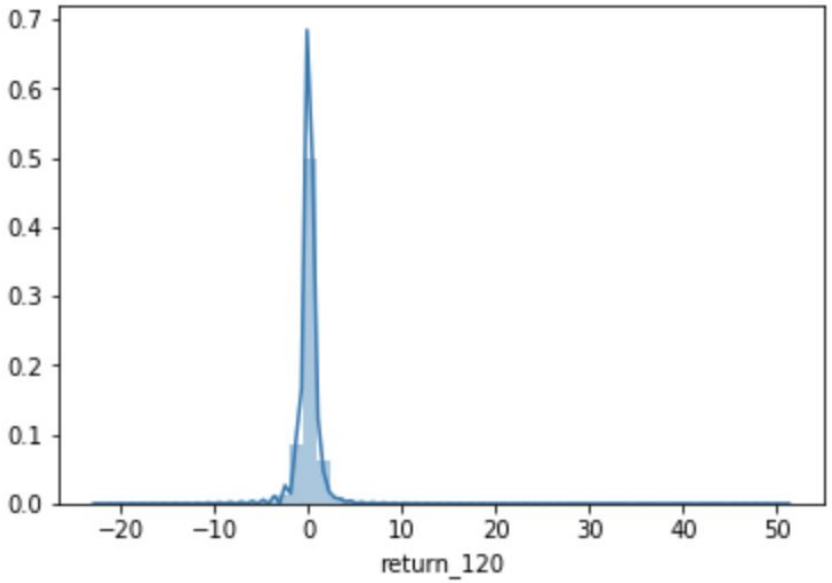
Bitcoin -> Weekly Z-score Return Distribution

```
When returns were calculated using return_60 and a weekly moving z_score:
There were 679429 possible events
There were 52940 positive and 54831 negative 1-sigma events
There were 14143 positive and 16909 negative 2-sigma events
There were 6169 positive and 7187 negative 3-sigma events
There were 2243 positive and 1913 negative 5-sigma events
There were 337 positive and 380 negative 10-sigma events
There were 2 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



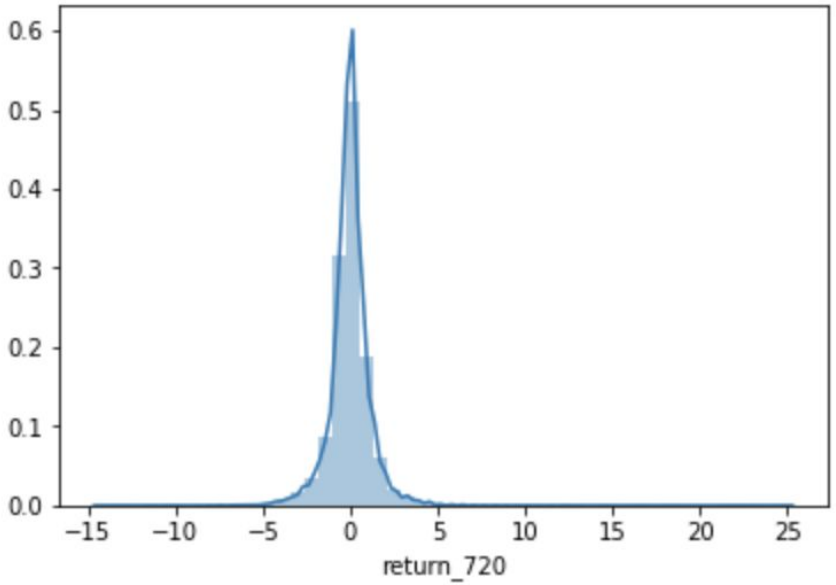
Bitcoin -> Weekly Z-score Return Distribution

```
When returns were calculated using return_120 and a weekly moving z_score:
There were 679429 possible events
There were 55603 positive and 59267 negative 1-sigma events
There were 16312 positive and 18632 negative 2-sigma events
There were 7562 positive and 7650 negative 3-sigma events
There were 2055 positive and 2069 negative 5-sigma events
There were 371 positive and 239 negative 10-sigma events
There were 1 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



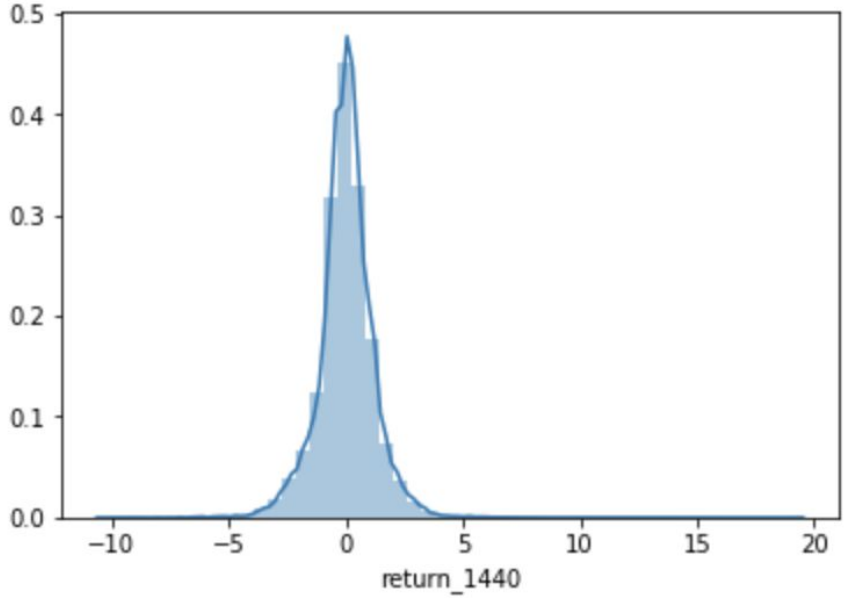
Bitcoin -> Weekly Z-score Return Distribution

```
When returns were calculated using return_720 and a weekly moving z_score:
There were 679429 possible events
There were 84376 positive and 79936 negative 1-sigma events
There were 24110 positive and 26916 negative 2-sigma events
There were 9575 positive and 9420 negative 3-sigma events
There were 1233 positive and 1129 negative 5-sigma events
There were 213 positive and 55 negative 10-sigma events
There were 0 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



Bitcoin -> Weekly Z-score Return Distribution

```
When returns were calculated using return_1440 and a weekly moving z_score:
There were 679429 possible events
There were 106740 positive and 96047 negative 1-sigma events
There were 27967 positive and 32340 negative 2-sigma events
There were 7305 positive and 8564 negative 3-sigma events
There were 1215 positive and 1065 negative 5-sigma events
There were 163 positive and 3 negative 10-sigma events
There were 0 positive and 0 negative 50-sigma events
There were 0 positive and 0 negative 100-sigma events
```



NLU + Sentiment

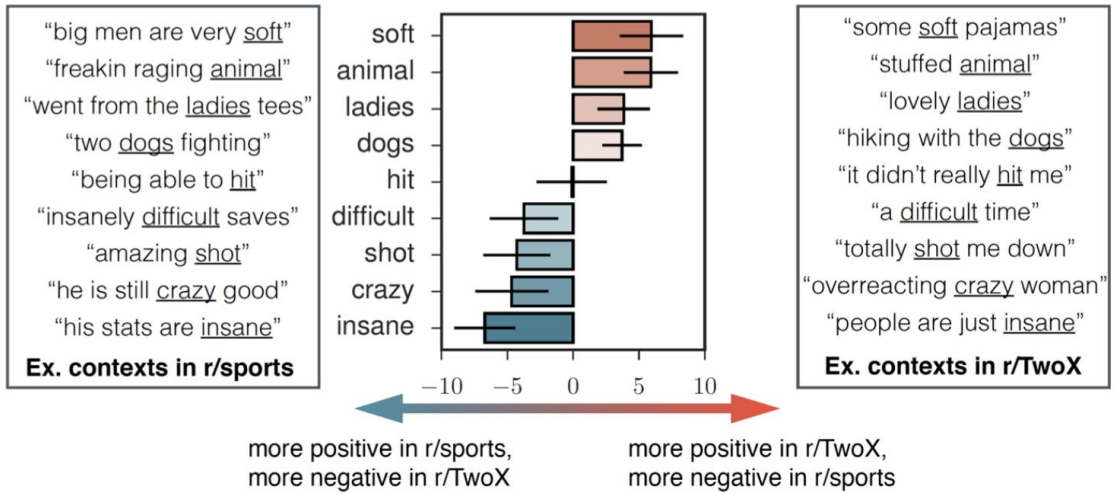
- Hypothesis
  - Public sentiment drives market prices for crypto assets
- Data
  - Reddit
- Methods
  - Off the shelf sentiment classifiers (NLTK+TextBlob)
  - SocialSent

NLU Data

- Too much data
  - All comments on cryptocurrency related subreddits **ever**
  - Narrowed to **r/btc** and **r/Bitcoin**
- Comment Text
- Timestamp
- Score (Upvotes/Downvotes)
- Author
- Metadata

NLU Methods

- Off the shelf sentiment classifiers
  - NLTK
  - TextBlob
  - **Poor performance**
- **SocialSent:** Inducing Domain-Specific Lexicons from Unlabeled Corpora
  - William L. Hamilton
  - Kevin Clark
  - Jure Leskovec
  - Dan Jurafsky



SocialSent

1. Co-Occurrence VSM
   a. Vocab 5000 most frequent words
   b. Automatically includes seed words
2. PPMI
3. SVD (Singular Value Decomposition)
   a. dimensions = 300
4. Turn the VSM into a Graph

$$\mathbf{M}_{i,j}^{PPMI} = \max\left\{\log\left(\frac{\hat{p}(w_i, w_j)}{\hat{p}(w)\hat{p}(w_j)}\right), 0\right\}$$

$$\mathbf{M}^{PPMI} \doteq \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top}$$

$$\mathbf{w}_i^{\mathrm{SVD}} = (\mathbf{U})_i \,.$$

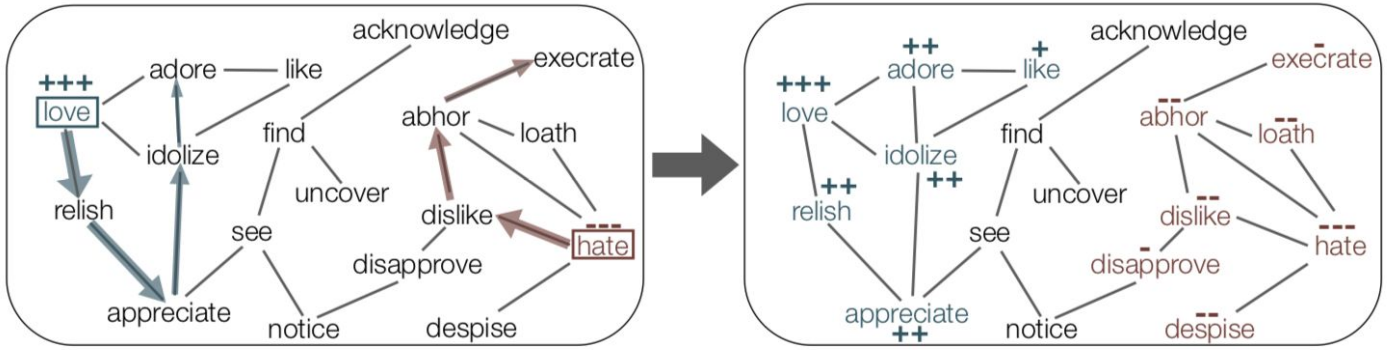
$$\mathbf{E}_{i,j} = \arccos\left(-\frac{\mathbf{w_i}^{\top}\mathbf{w_j}}{\|\mathbf{w_i}\|\|\mathbf{w_j}\|}\right)$$

SocialSent

1. Propagate sentiment from seed words in the graph
2. Use random walk to obtain positive and negative polarity scores
3. Weight those scores to obtain final polarity

Domain	Positive seed words	Negative seed words
Standard English	good, lovely, excellent, fortunate, pleasant, delightful, perfect, loved, love, happy	bad, horrible, poor, unfortunate, unpleasant, disgusting, evil, hated, hate, unhappy
Finance	successful, excellent, profit, beneficial, improving, improved, success, gains, positive	negligent, loss, volatile, wrong, losses, damages, bad, litigation, failure, down, negative
Twitter	love, loved, loves, awesome, nice, amazing, best, fantastic, correct, happy	hate, hated, hates, terrible, nasty, awful, worst, horrible, wrong, sad

Table 1: Seed words. The seed words were manually selected to be context insensitive (without knowledge of the test lexicons).



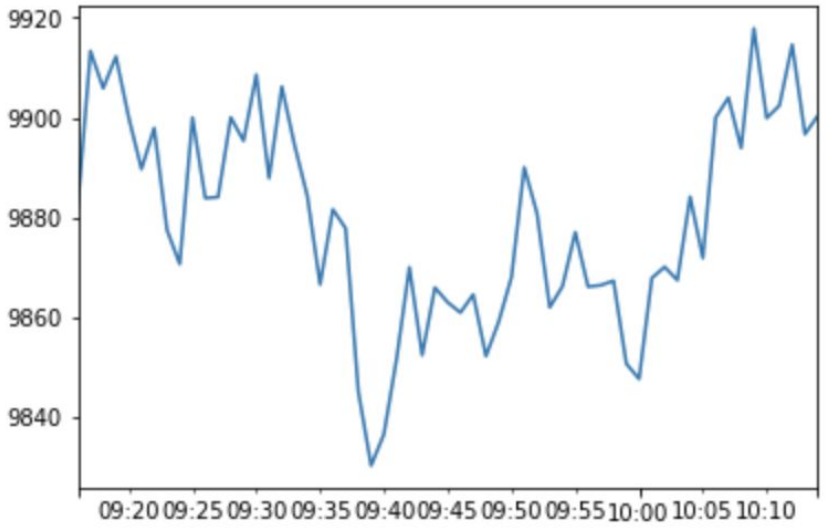
a. Run random walks from seed words.

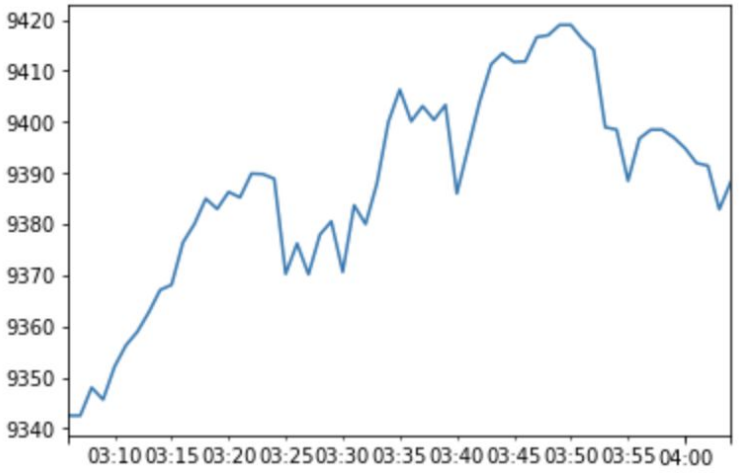
b. Assign polarity scores based on frequency of random walk visits.

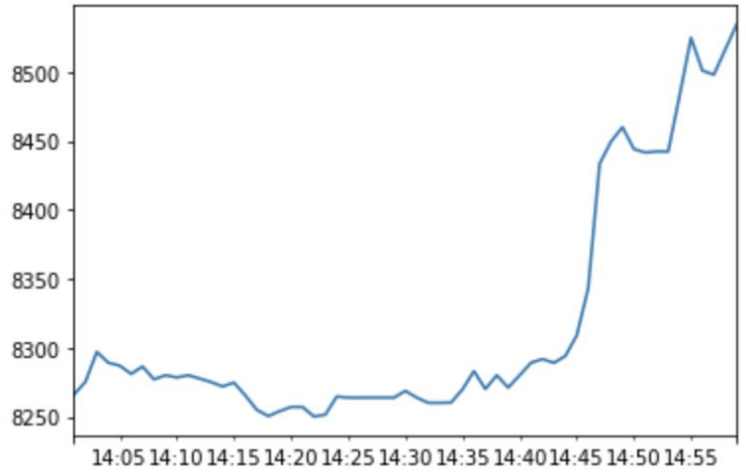
Figure 3: Visual summary of the SENTPROP algorithm.

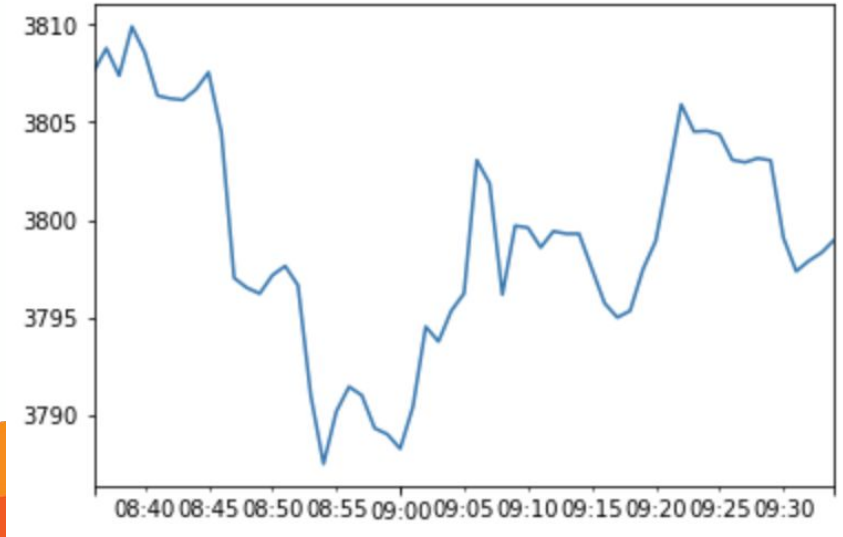
$$\frac{\mathbf{p}^P(w_i)}{\mathbf{p}^P(w_i)+\mathbf{p}^N(w_i)}$$

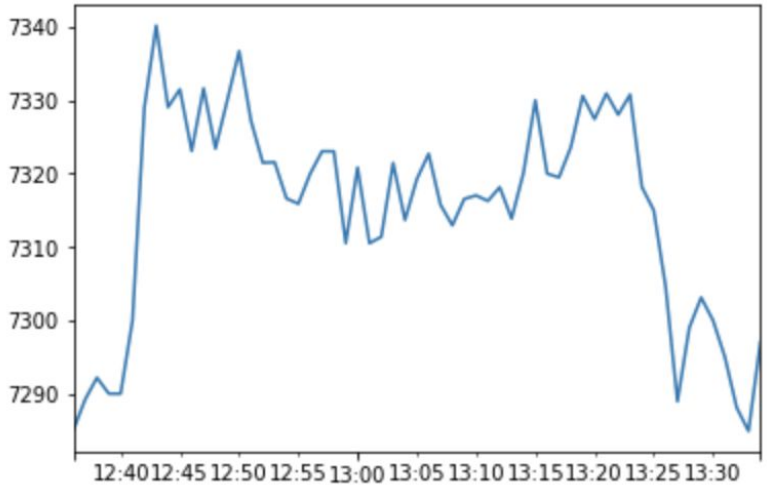
Price 1 hour after 25 sigma event using weekly moving z-score

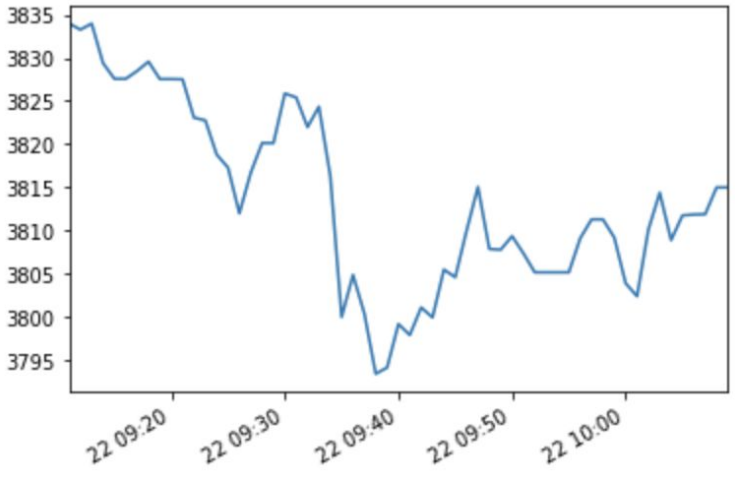








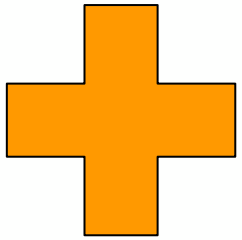




NLU Further Work

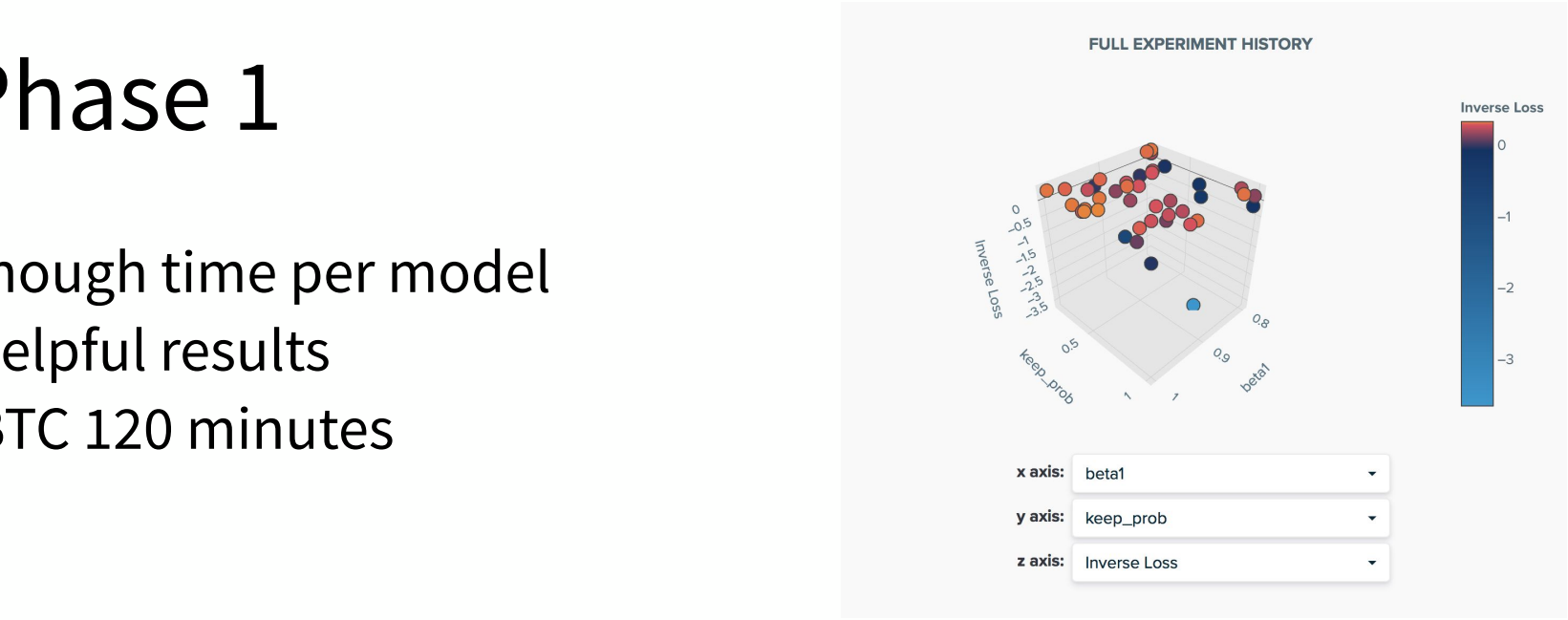
- Use of sentiment time series as feature in broader RNN
- Lower hanging fruit -> current methods are complex and rely on assumptions
- Time
  - Tuning hyperparameters, exploring data further
- Using a densifier training method for subreddits with 100B+ tokens of data
- Entity recognition and resolution
  - “How often is Bitcoin mentioned?”
  - “Are BCH, BTC, BCHABC, and BSV different things?”

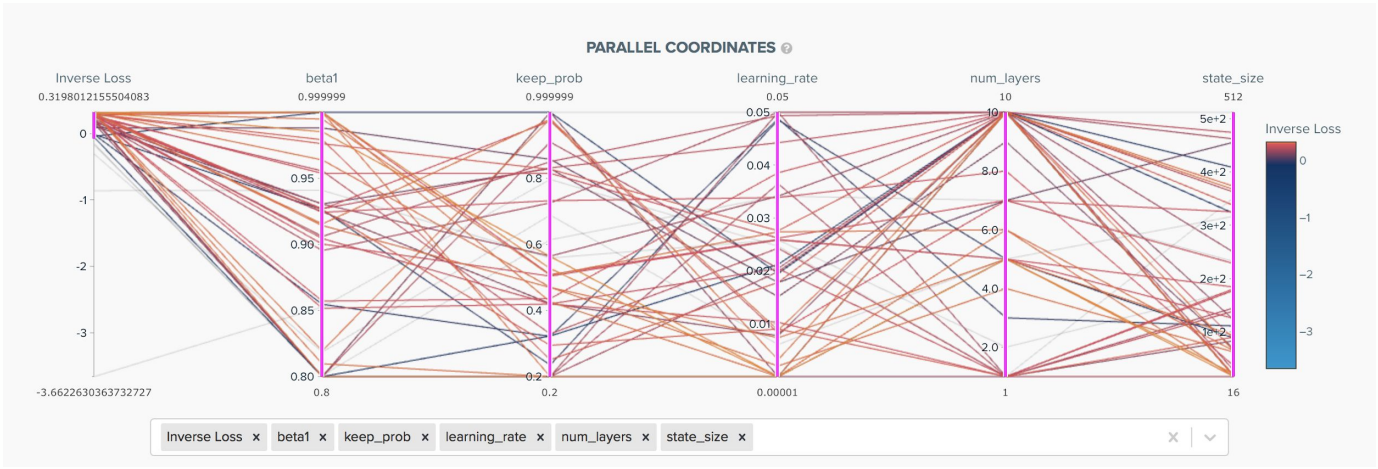
Training

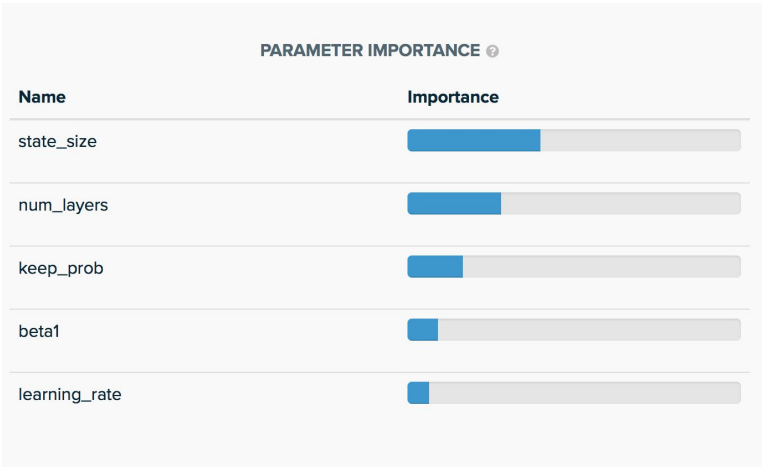


Training Phase 1

- Didn’t give enough time per model
- Some early helpful results
- Training on BTC 120 minutes

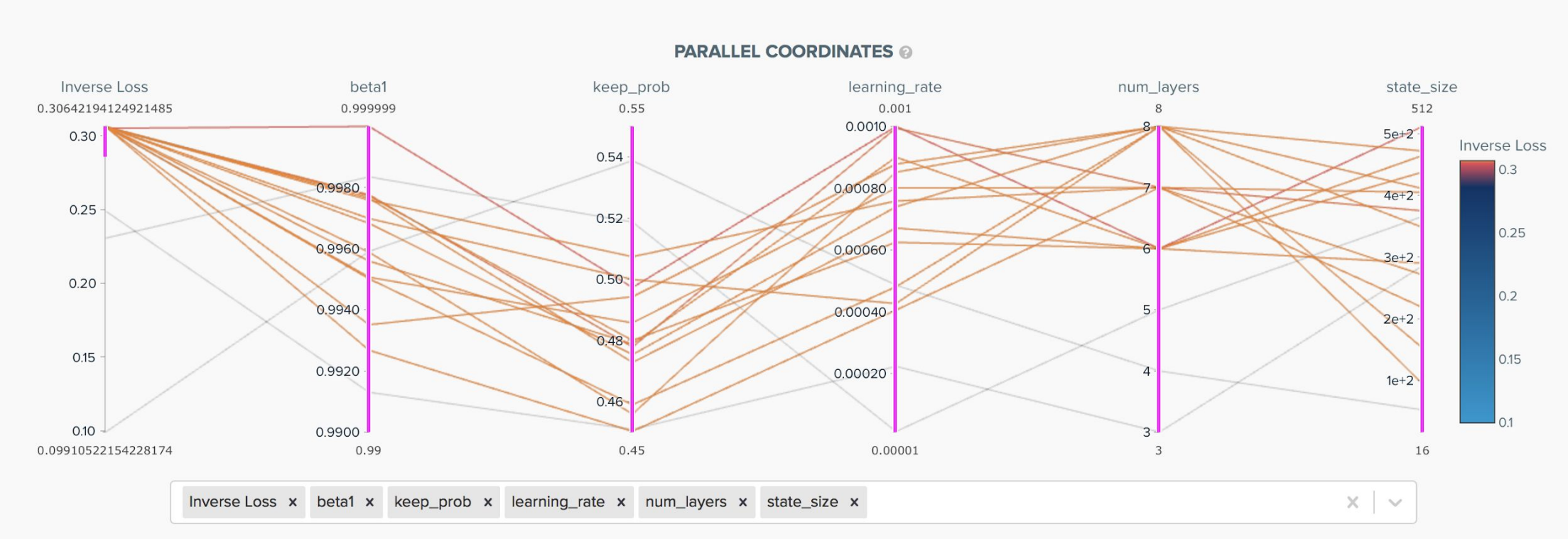




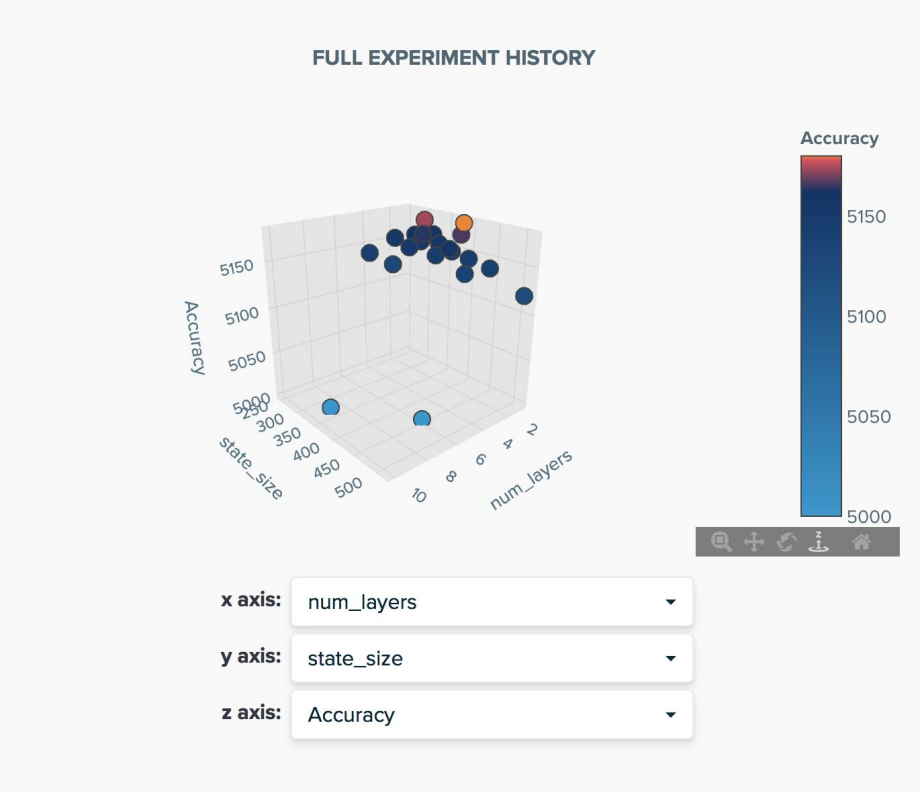


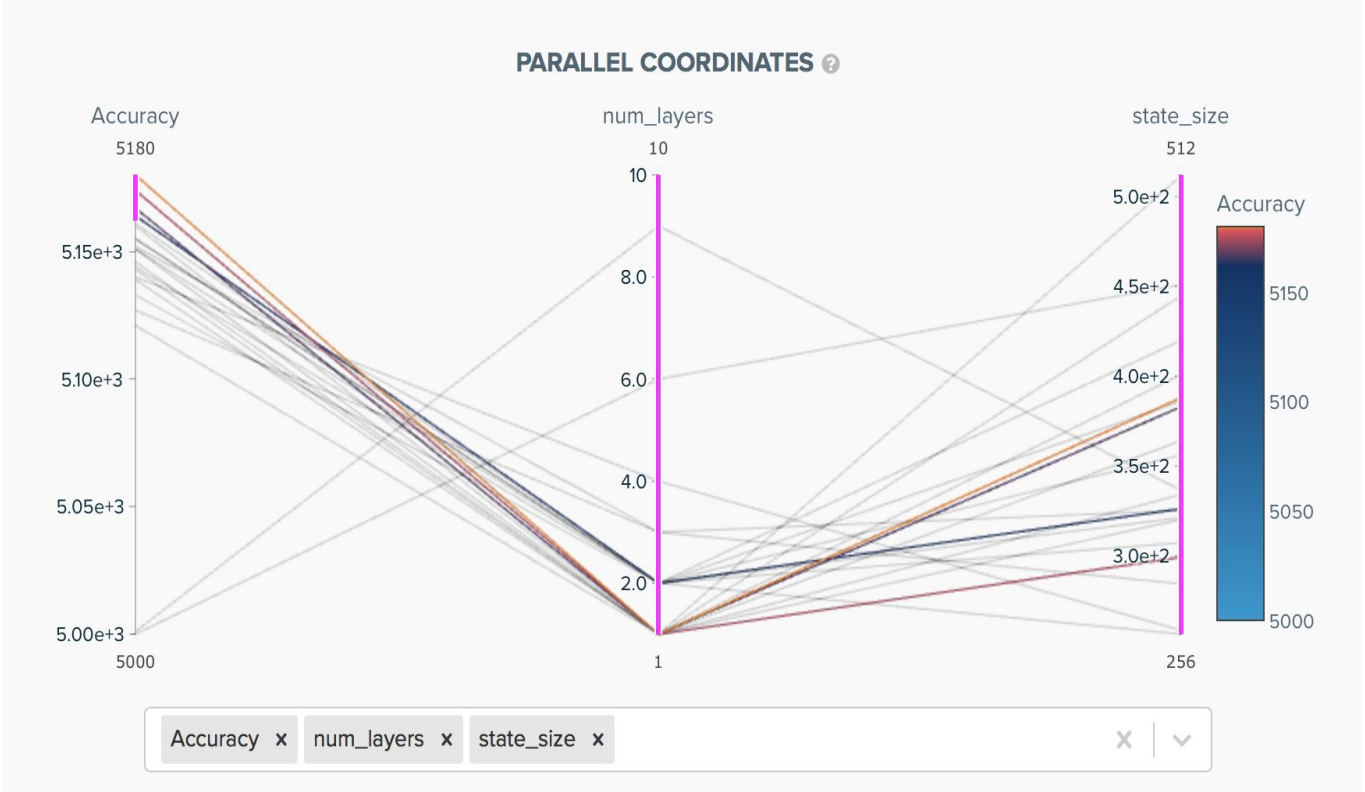
Training Phase 2

- Identified some key hyperparameters
- More training necessary



Training Phase 3





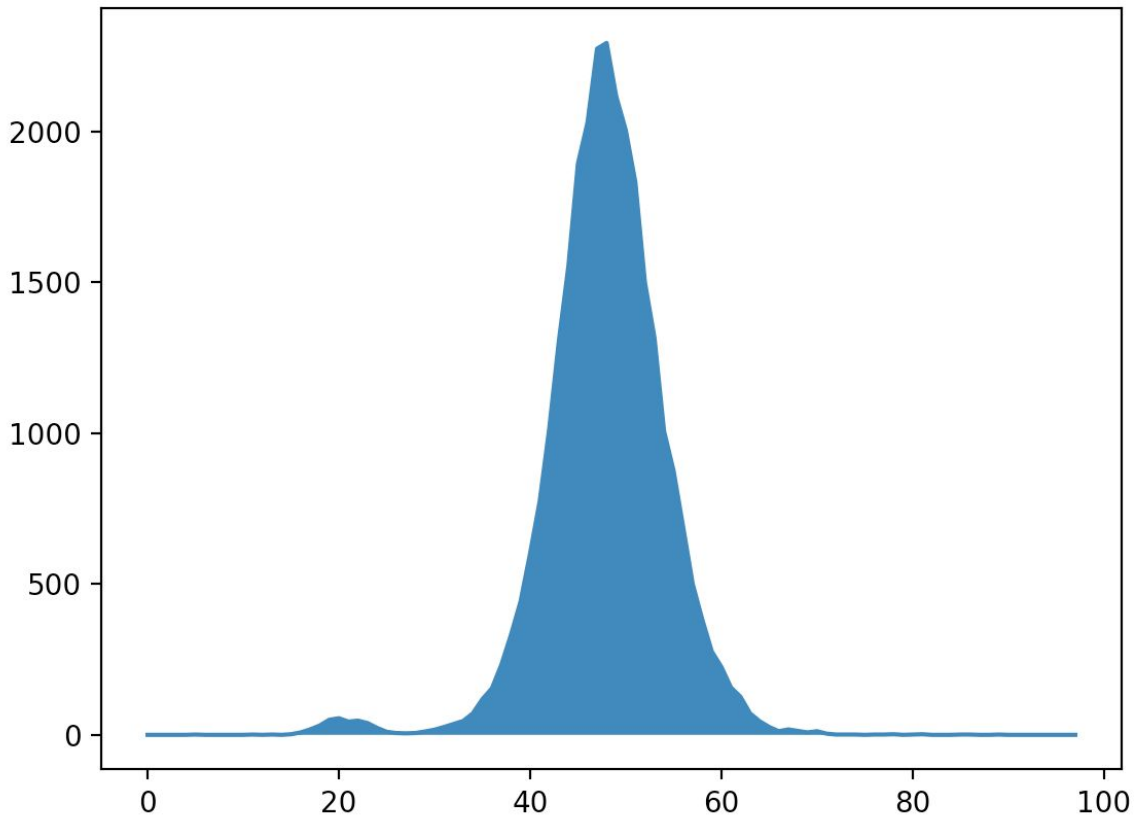
Model Results - multiple time periods

Coin	Pred Minutes	Real Train Acc.	Alpha Train Acc.	Real Test Acc.	Alpha Test Acc.
BTC	120	65.26%	56.24%	57.87%	50.32%
BTC	120	58.22%	54.39%	54.15%	51.45%
LTC	120	59.81%	55.24%	53.27%	50.88%
LTC	5	58.32%	53.25%	54.02%	51.80%
LTC	5	68.42%	60.25%	62.77%	54.59%

Focus On LTC-5min

Coin	Pred Minutes	Real Train Acc.	Alpha Train Acc.	Real Test Acc.	Alpha Test Acc.
BTC	120	65.26%	56.24%	57.87%	50.32%
BTC	120	58.22%	54.39%	54.15%	51.45%
LTC	120	59.81%	55.24%	53.27%	50.88%
LTC	5	58.32%	53.25%	54.02%	51.80%
LTC	5	68.42%	60.25%	62.77%	54.59%

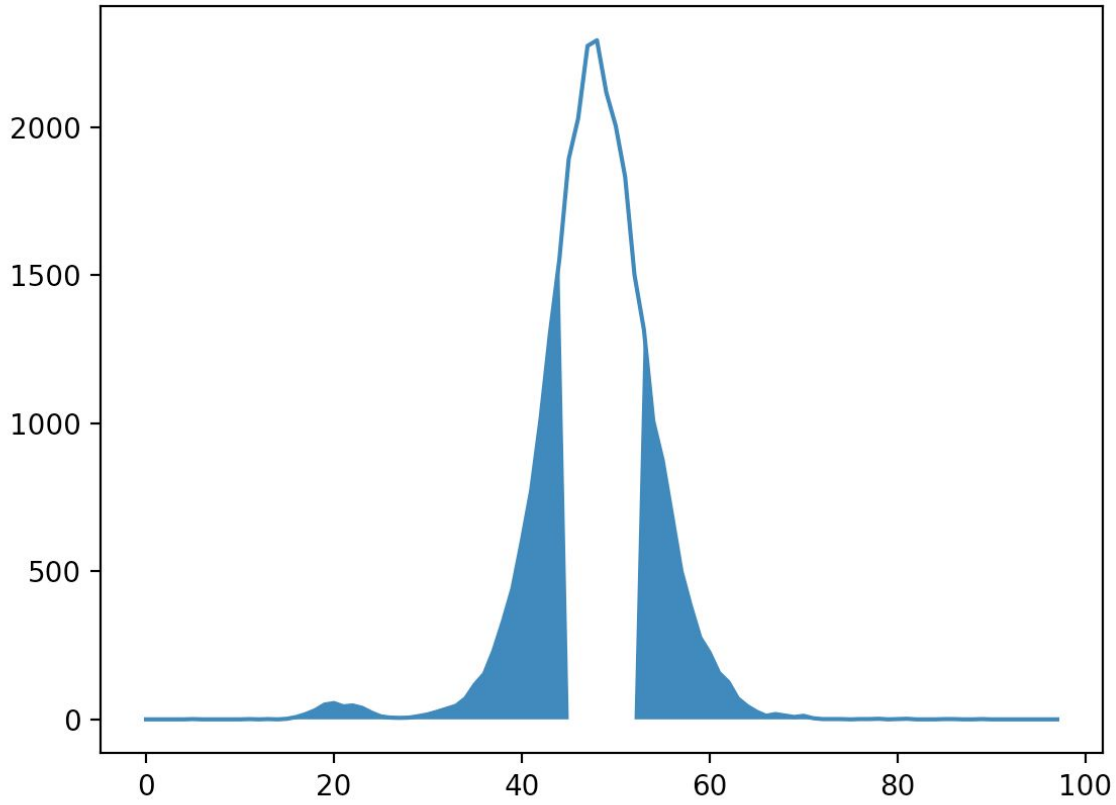
Model Analysis - Sample Space Ω



Recall @ 0.50 Investment Thresh

Positive Recall: 62.12%
Negative Recall: 63.32%

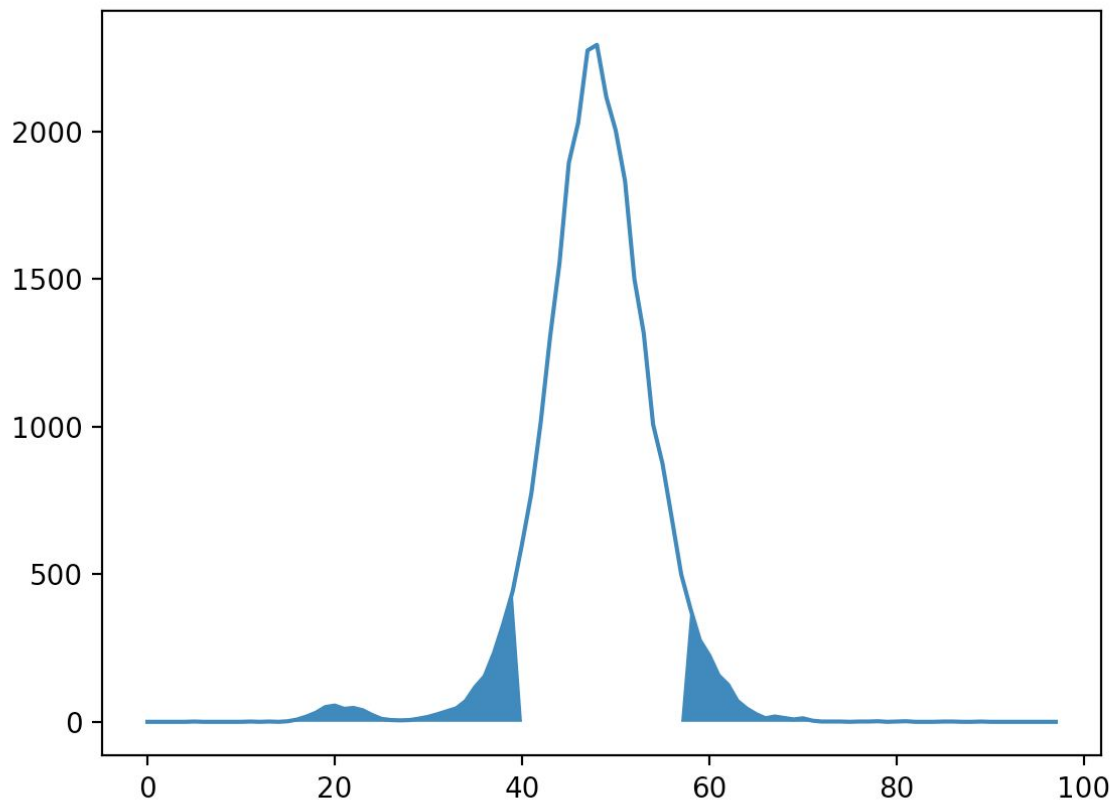
Model Analysis - Event Space $E \subseteq \Omega$



Recall @ 0.55 Investment Thresh

Positive Recall: 74.81%
Negative Recall: 70.01%

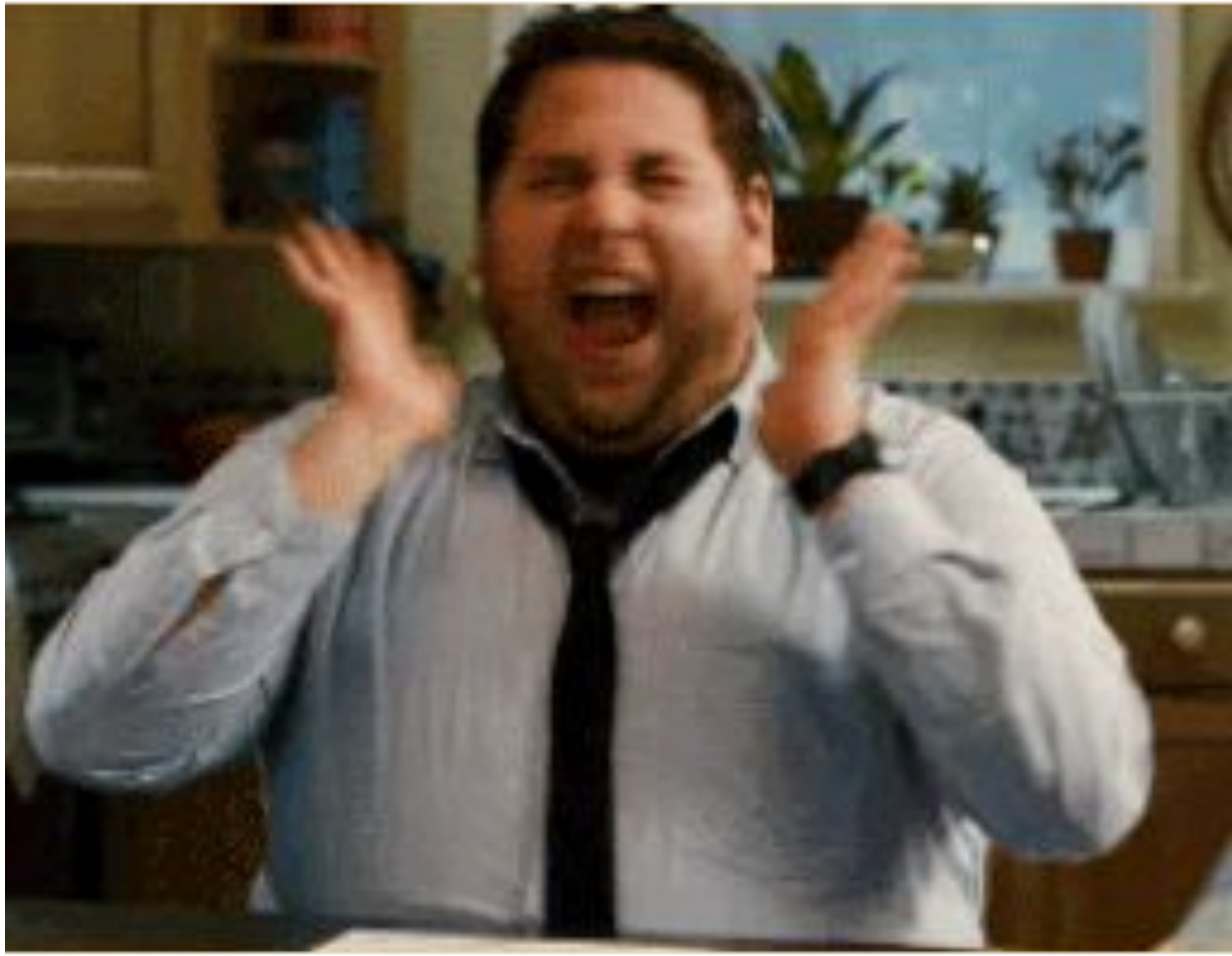
Model Analysis - Event Space $E \subseteq \Omega$



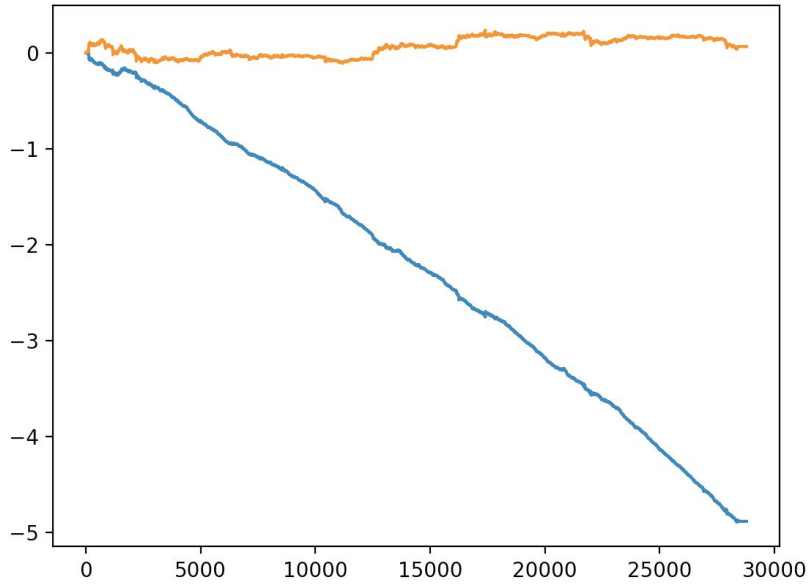
Recall @ 0.60 Investment Thresh

Positive Recall: 80.81%
Negative Recall: 78.23%

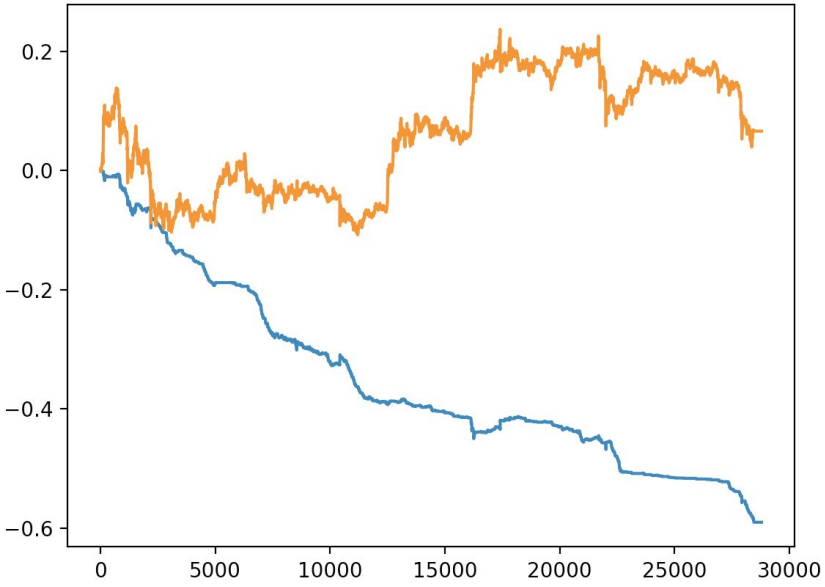
Mar ket



Market Results @ 25 basis point cost per trade

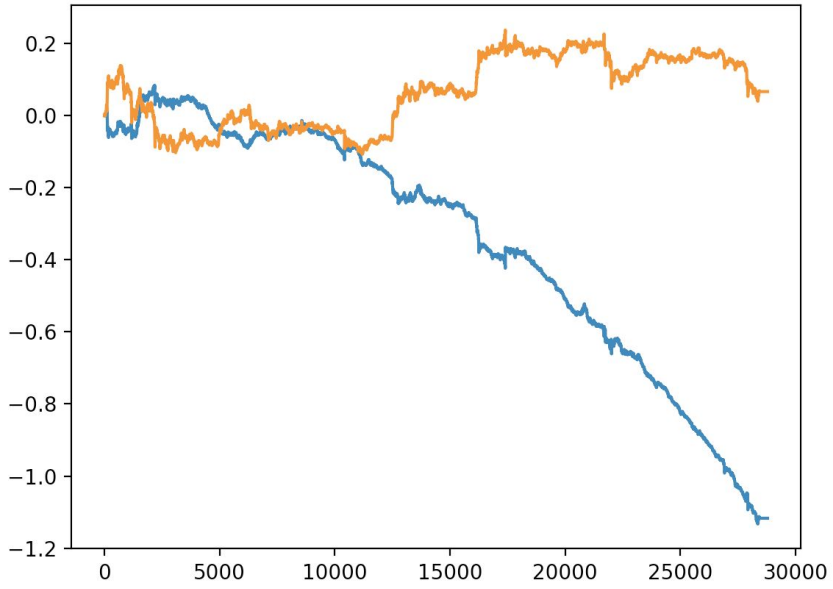


Investment Threshold: 0.50
Return %: -488.42%
Sharpe: -19.90

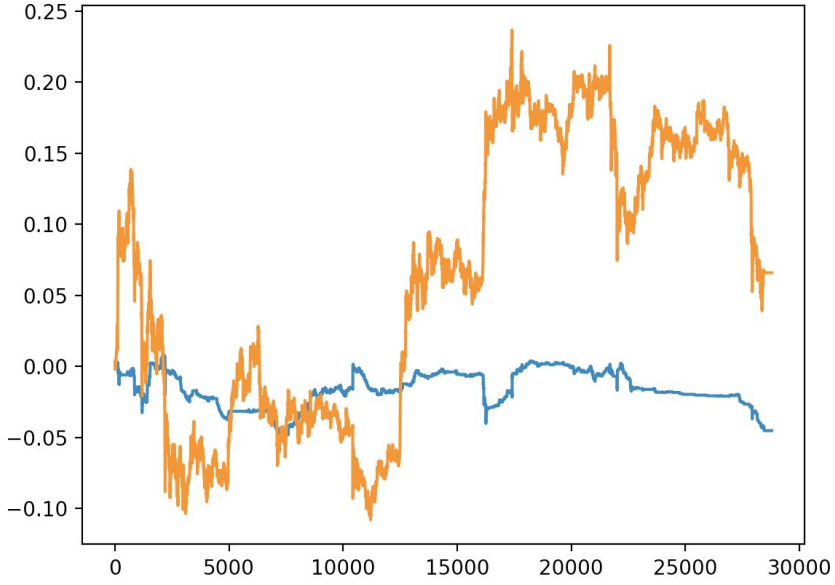


Investment Threshold: 0.60
Return %: -58.99%
Sharpe: -24.28

Market Results @ 10 basis point cost per trade

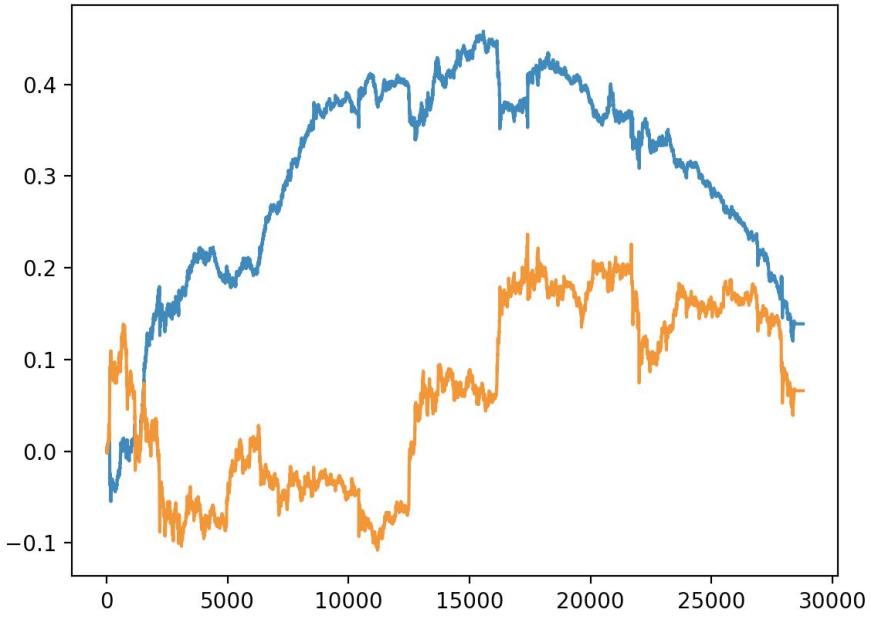


Investment Threshold: 0.50
Return %: -111.97%
Sharpe: -05.05

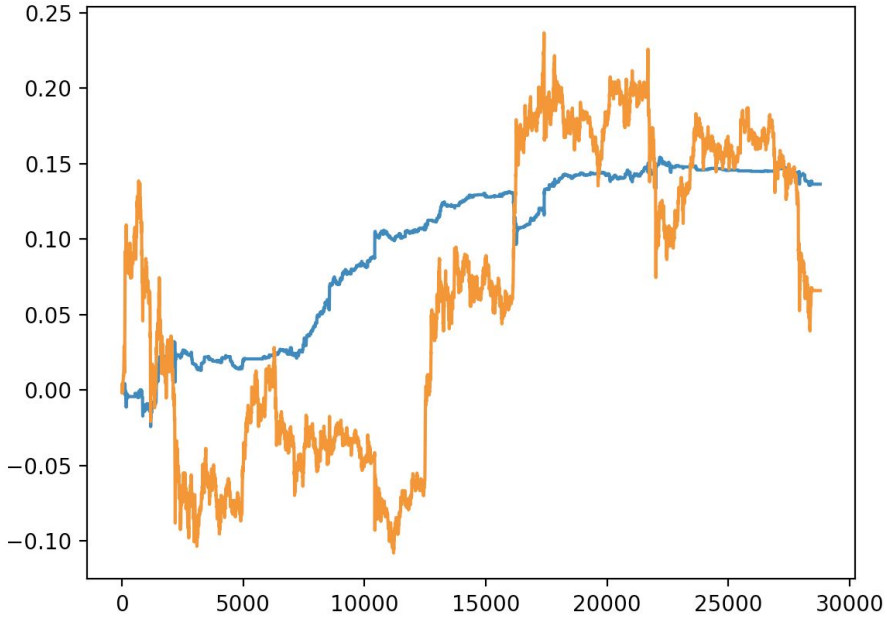


Investment Threshold: 0.60
Return %: -4.50%
Sharpe: -2.04

Market Results @ 05 basis point cost per trade



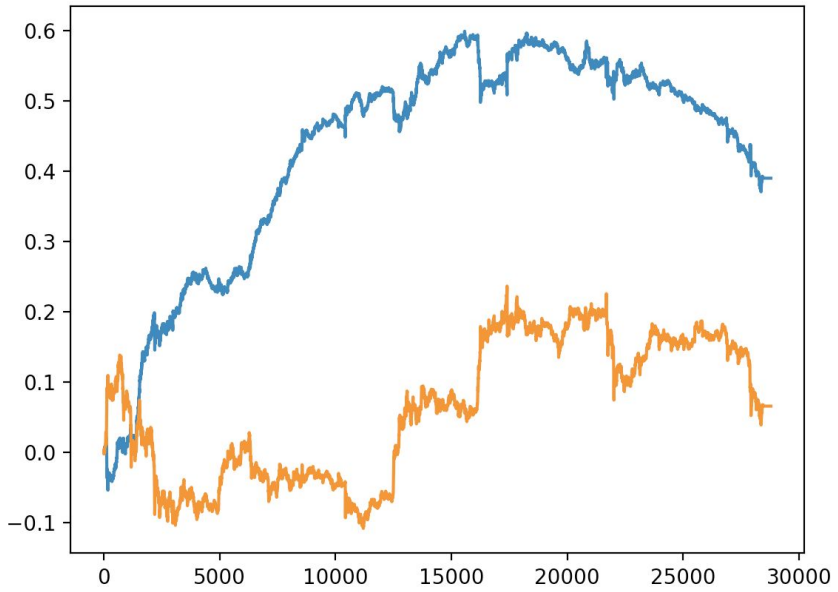
Investment Threshold: 0.50
Return %: 13.89%
Sharpe: 0.639



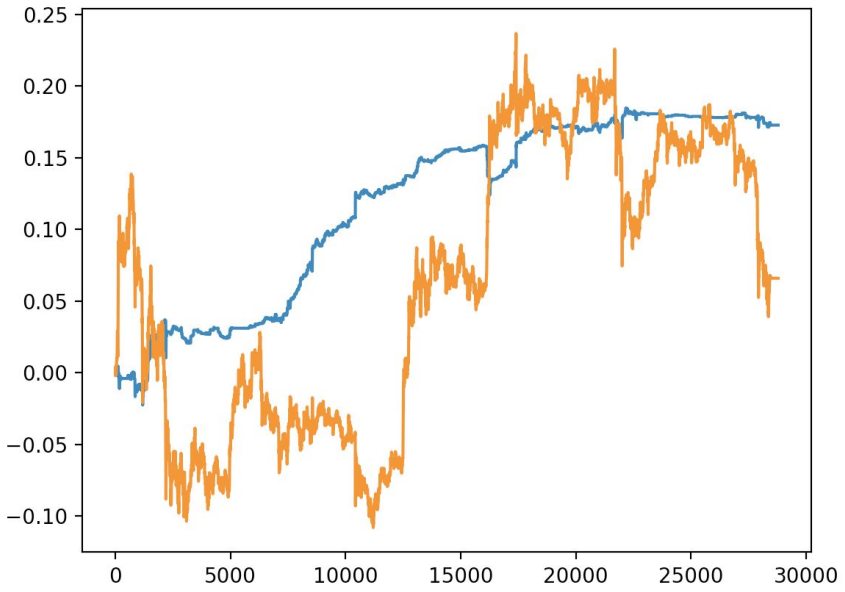
Investment Threshold: 0.60
Return %: 13.65%
Sharpe: 6.27

Best Entry

Market Results @ 04 basis point cost per trade



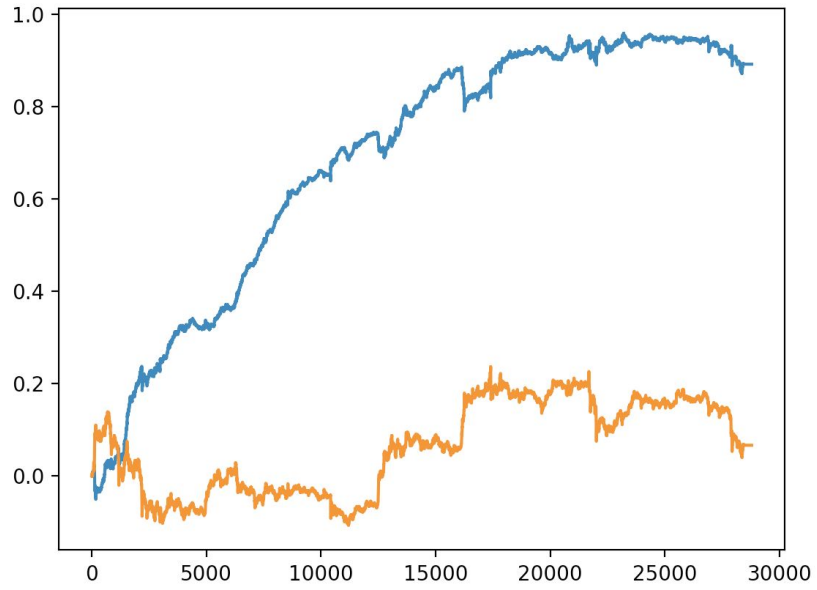
Investment Threshold: 0.50
Return %: 39.01%
Sharpe: 1.79



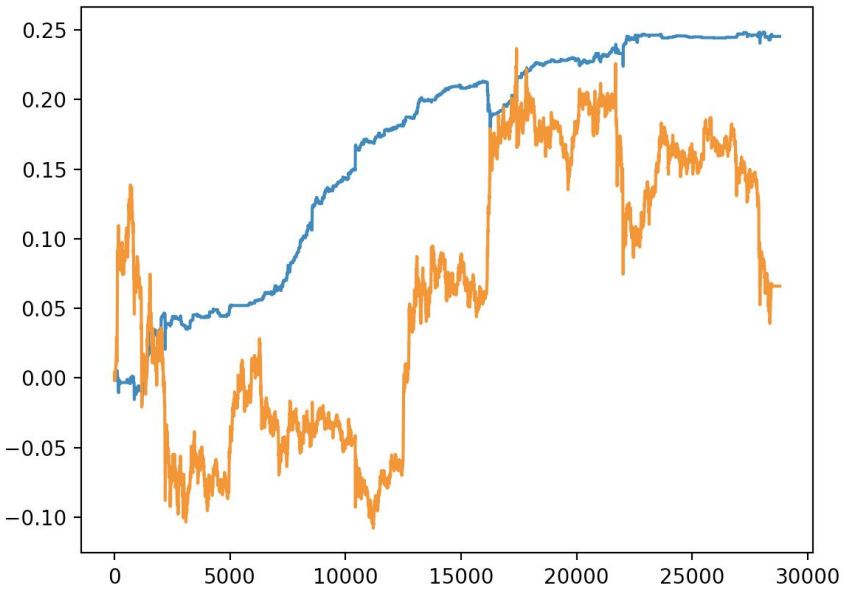
Investment Threshold: 0.60
Return %: 17.27%
Sharpe: 7.95

Theoreticals

Market Results @ 02 basis point cost per trade



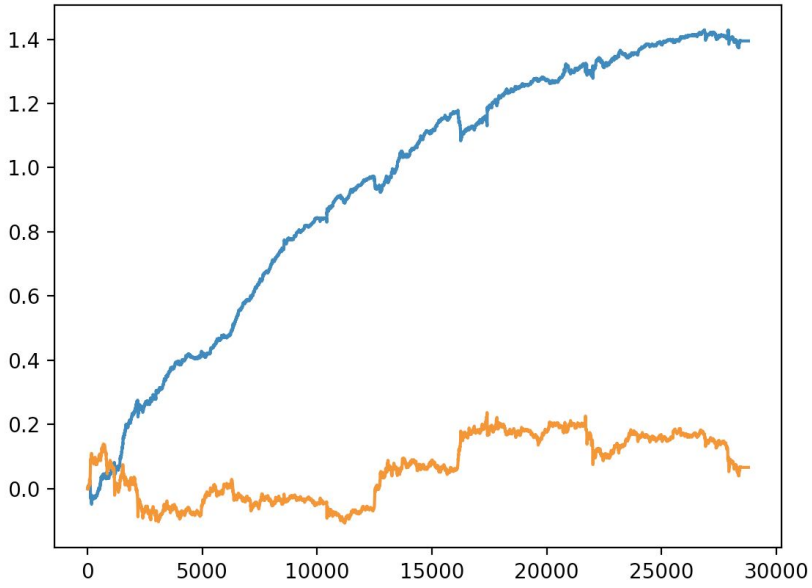
Investment Threshold: 0.50
Return %: 89.24%
Sharpe: 4.12



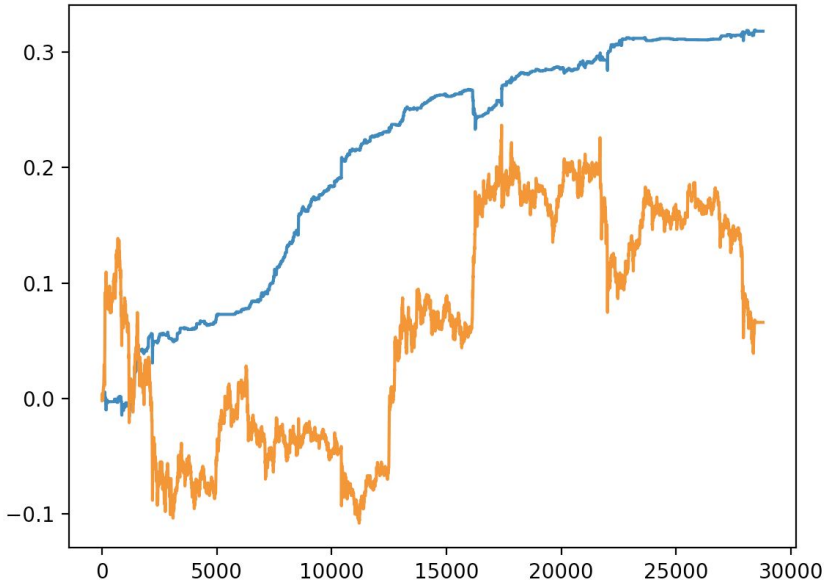
Investment Threshold: 0.60
Return %: 24.54%
Sharpe: 11.31

Theoretical Limit

Market Results @ 00 basis point cost per trade



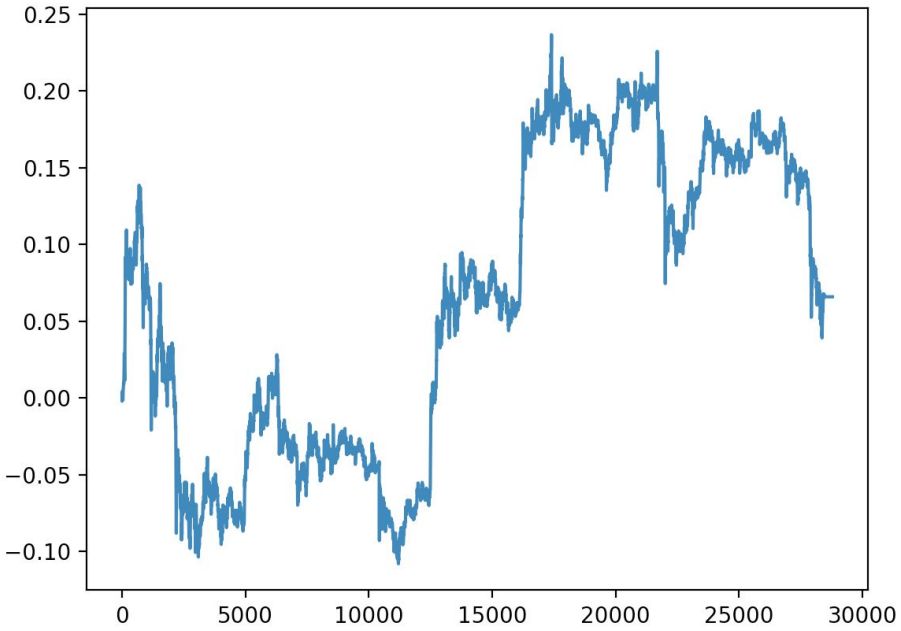
Investment Threshold: 0.50
Return %: 139.48%
Sharpe: 6.44



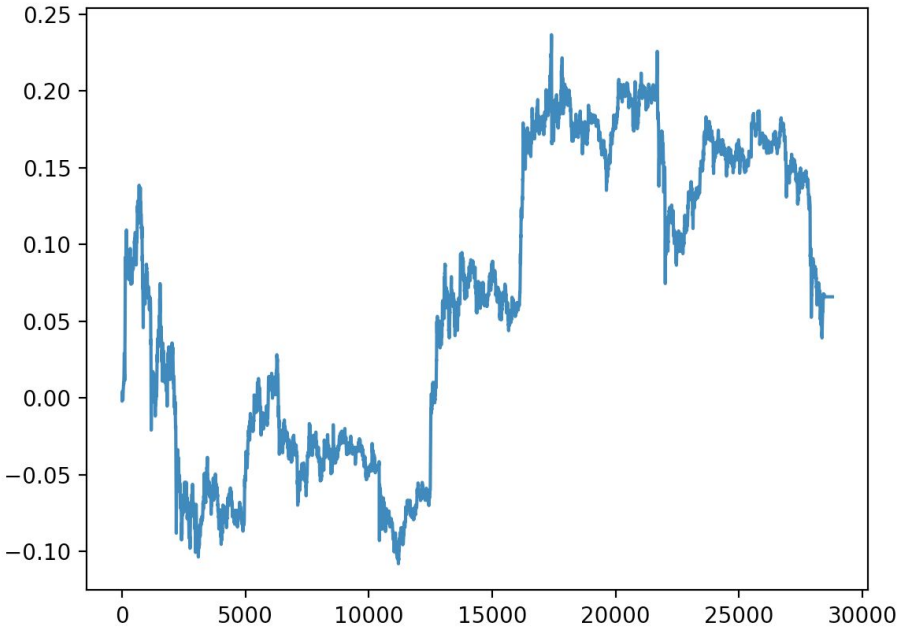
Investment Threshold: 0.60
Return %: 31.80%
Sharpe: 14.64

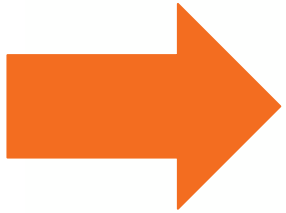
So What?

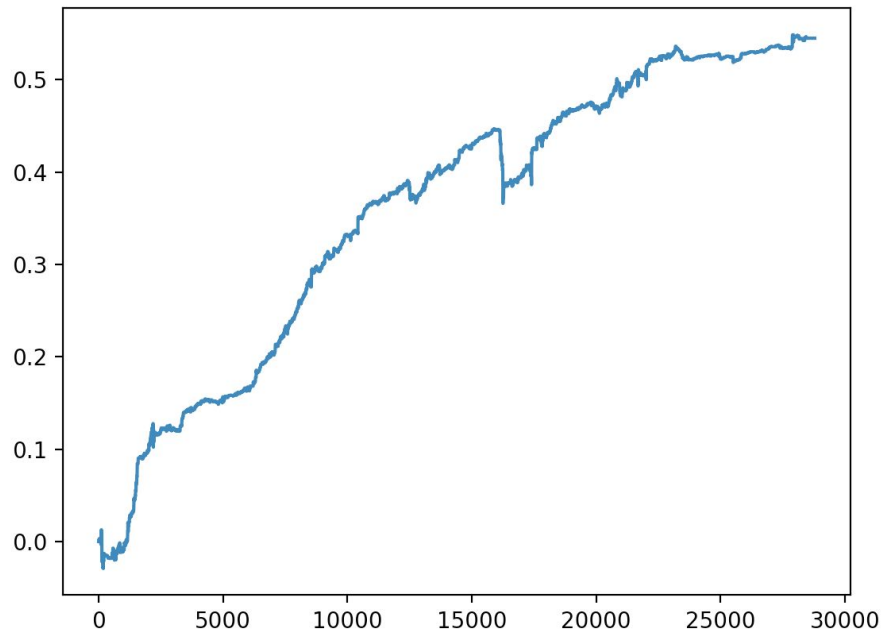
Problem Solved - Kinda



Problem Solved - Kinda







The End.