Analysis of several different methods for trading Tesla stocks with a focus on natural language processing techniques applied to twitter data.

Research report

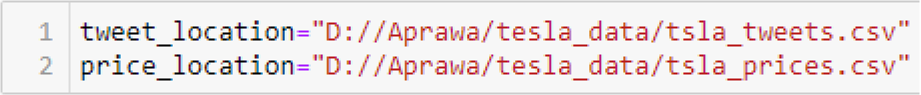
Author:

Eilder Jorge García

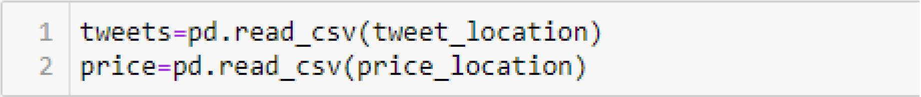
Introduction

Initial operations and data preparation

The first step is obtaining the data, since in this case it was already available in .csv format and it's relatively small I can just load it in a dataframe.

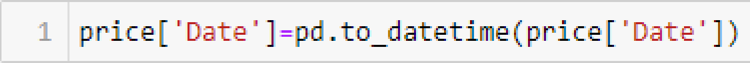


```
tweet_location="D://Aprawa/tesla_data/tsla_tweets.csv"
price_location="D://Aprawa/tesla_data/tsla_prices.csv"
```



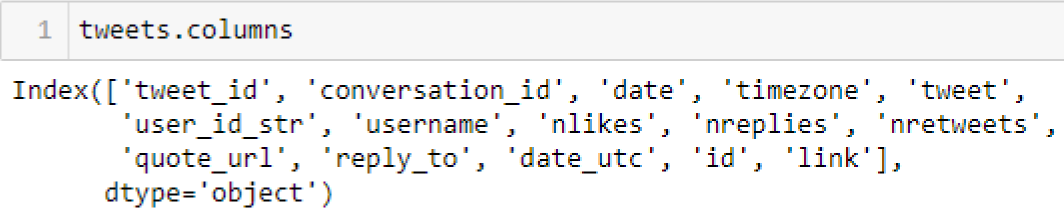
```
tweets=pd.read_csv(tweet_location)
price=pd.read_csv(price_location)
```

Next is cleaning up the data. In the case of price there isn't much to do except just check if there are missing values or any other weird problem with the data. After checking the data I conclude it's clean and I only need to convert the dates to a date time object.



```
price['Date']=pd.to_datetime(price['Date'])
```

Cleaning up the tweets is a two-fold operation. The first part is cleaning up columns that aren't of interest:



```
tweets.columns
```

```
Index(['tweet_id', 'conversation_id', 'date', 'timezone', 'tweet',
       'user_id_str', 'username', 'nlikes', 'nreplies', 'nretweets',
       'quote_url', 'reply_to', 'date_utc', 'id', 'link'],
      dtype='object')
```

Date is redundant with date_utc, so I'll keep date_utc only. Timezone is also redundant as date_utc already accounts for it.

tweet_id, link, and id are also redundant as we are interested in the content of the tweets and not in replying or operating with those tweets in any fashion.

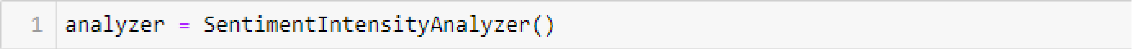
Conversation_id, reply_to and quote_url could be interesting if we are going to be doing extended social analysis network but the cost of doing such an analysis and the relation to the tweets is not practical in real time operations.

Finally user_id_str and username could be useful if we knew the ids of influential figures beforehand as we could assign a bigger weight to their tweets, but in this case we'll assume everyone is equally popular and judge by number of likes, retweets, and the sentiment of their tweet.

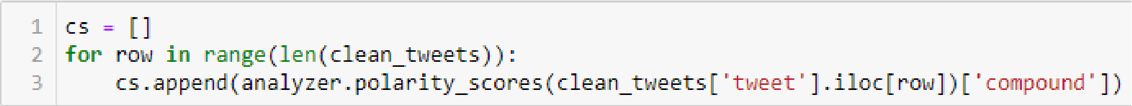
This leaves us with tweet text, the number of likes, the number of replies, the number of retweets and the date in UTC.

We then clean the tweets that are missing values and proceed to the 2nd part.

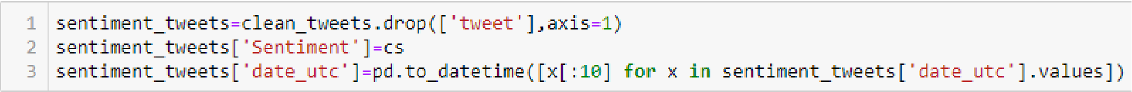
At this part we are interested in extracting a score of sentiment for each sentiment and how they affect TSLA stock. In a first iteration the idea is to use the VADER library in the NLTK package to extract the compound sentiment score for each tweet. We then select the maximum and minimum compound score per day, as well as the sum over the day and prepare a table with the scores for each day, this table is what's going to guide our trading signals.



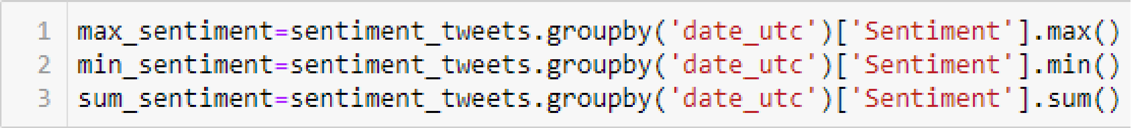
```
analyzer = SentimentIntensityAnalyzer()
```



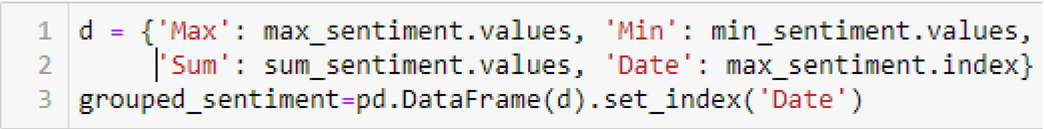
```
cs = []
for row in range(len(clean_tweets)):
    cs.append(analyzer.polarity_scores(clean_tweets['tweet'].iloc[row])['compound'])
```



```
sentiment_tweets=clean_tweets.drop(['tweet'],axis=1)
sentiment_tweets['Sentiment']=cs
sentiment_tweets['date_utc']=pd.to_datetime([x[:10] for x in sentiment_tweets['date_utc'].values])
```

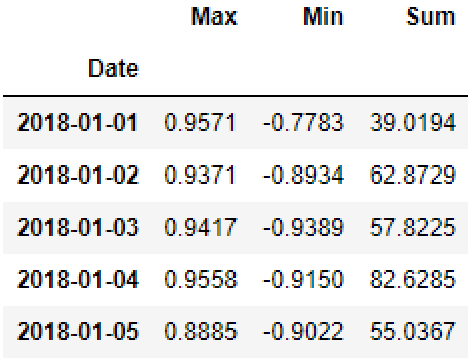


```
max_sentiment=sentiment_tweets.groupby('date_utc')['Sentiment'].max()
min_sentiment=sentiment_tweets.groupby('date_utc')['Sentiment'].min()
sum_sentiment=sentiment_tweets.groupby('date_utc')['Sentiment'].sum()
```



```
d = {'Max': max_sentiment.values, 'Min': min_sentiment.values,
     'Sum': sum_sentiment.values, 'Date': max_sentiment.index}
grouped_sentiment=pd.DataFrame(d).set_index('Date')
```

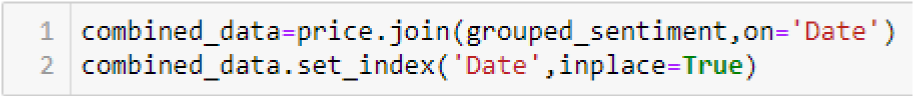
At this point we got a table with our overall sentiment each day, as well as our most positive and most negative sentiment.



	Max	Min	Sum
Date			
2018-01-01	0.9571	-0.7783	39.0194
2018-01-02	0.9371	-0.8934	62.8729
2018-01-03	0.9417	-0.9389	57.8225
2018-01-04	0.9558	-0.9150	82.6285
2018-01-05	0.8885	-0.9022	55.0367

A quick look shows that while the most positive and most negative sentiments are similar, the overall total is strongly positive.

We combine this sentiment data with our price data, removing tweets that lie in non-trading days for this first iteration.



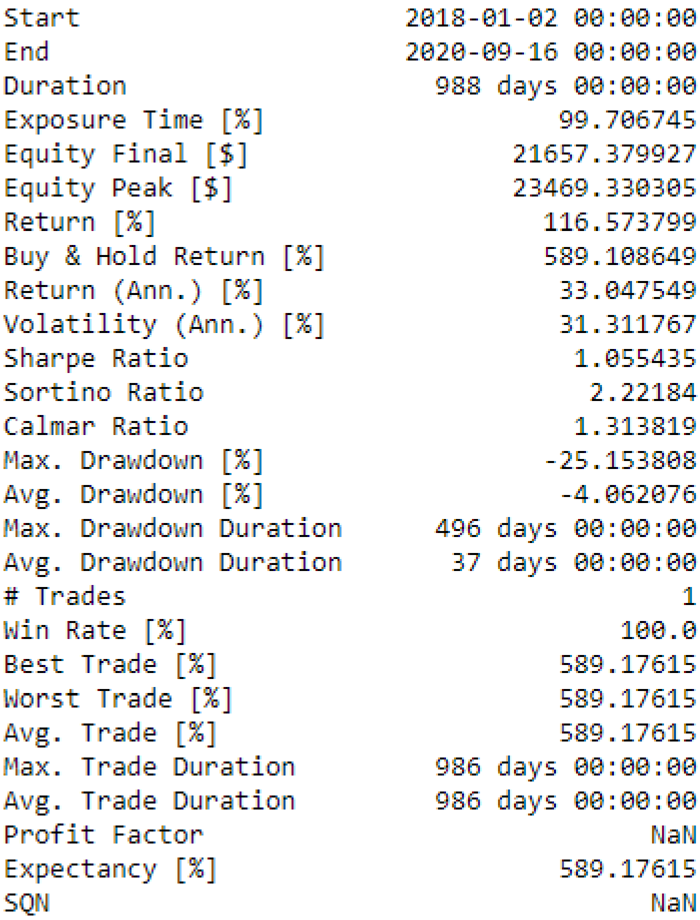
```
combined_data=price.join(grouped_sentiment,on='Date')
combined_data.set_index('Date',inplace=True)
```

We set the date as index because we need it to be set as an index for the backtesting library.

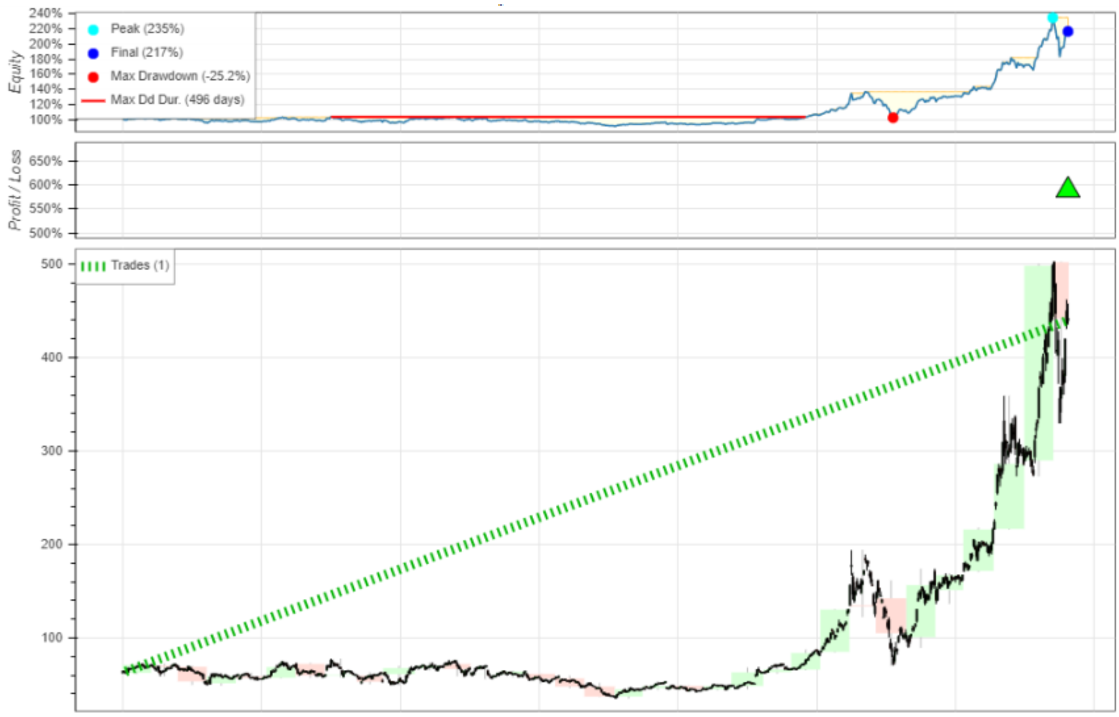
At this point all our data is together and we can test some money management and trade control operations.

First iteration: Using sentiment without modification

We start with a very simple idea, if our overall sentiment is positive we buy, otherwise we sell, and this in turn gives us the following result:

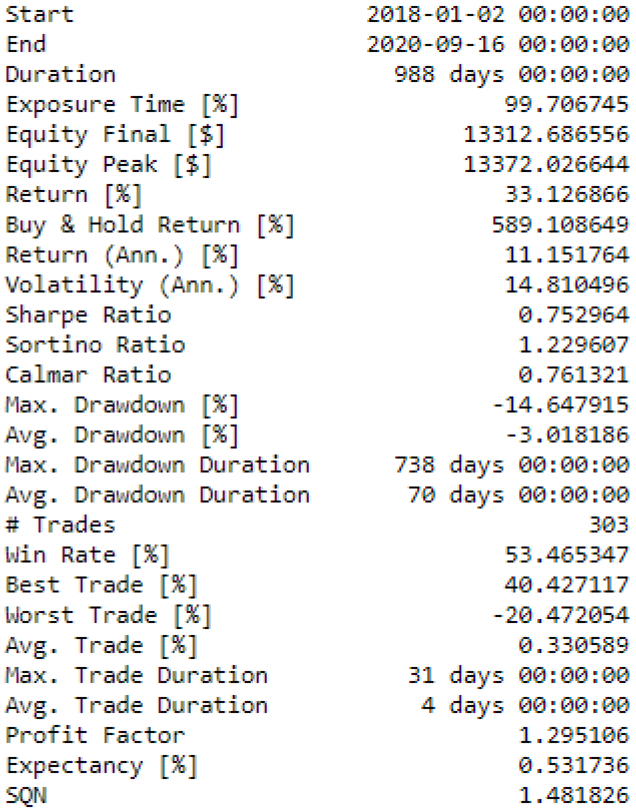


```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 21657.379927
Equity Peak [$]                  23469.330305
Return [%]                         116.573799
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   33.047549
Volatility (Ann.) [%]               31.311767
Sharpe Ratio                         1.055435
Sortino Ratio                         2.22184
Calmar Ratio                         1.313819
Max. Drawdown [%]                  -25.153808
Avg. Drawdown [%]                   -4.062076
Max. Drawdown Duration      496 days 00:00:00
Avg. Drawdown Duration       37 days 00:00:00
# Trades                                    1
Win Rate [%]                            100.0
Best Trade [%]                      589.17615
Worst Trade [%]                     589.17615
Avg. Trade [%]                      589.17615
Max. Trade Duration         986 days 00:00:00
Avg. Trade Duration         986 days 00:00:00
Profit Factor                             NaN
Expectancy [%]                      589.17615
SQN                                       NaN
```

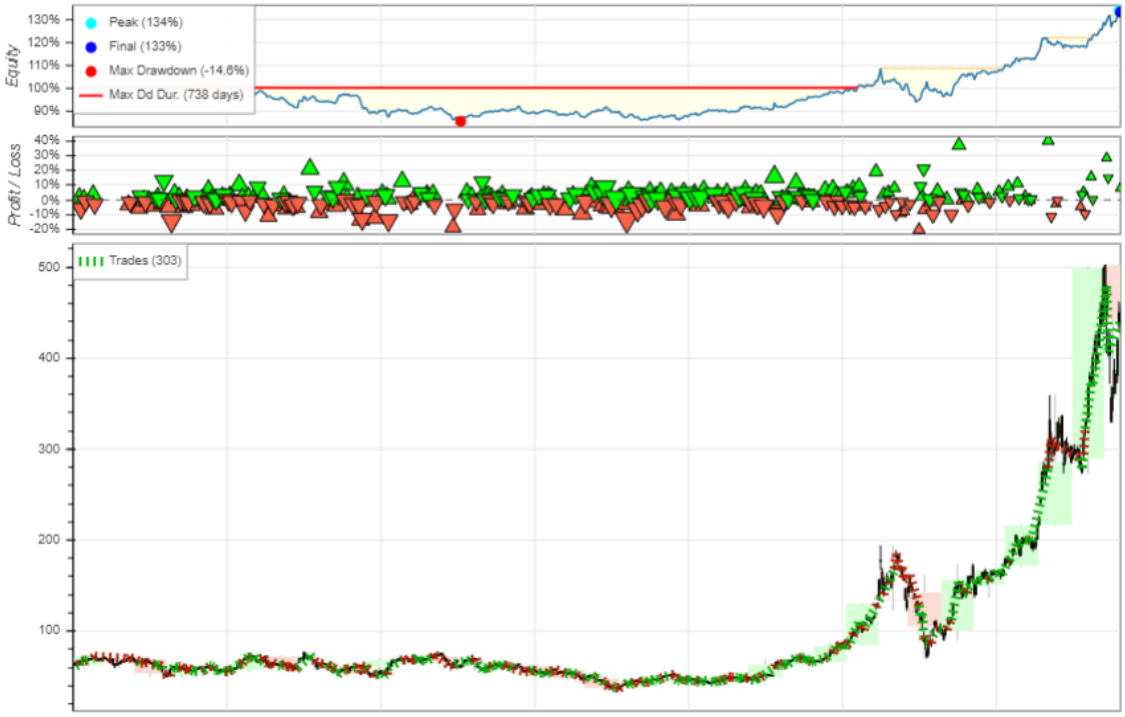


This means that the overall tweet sentiment was positive during the entire life of the stock, indicating to buy & hold.

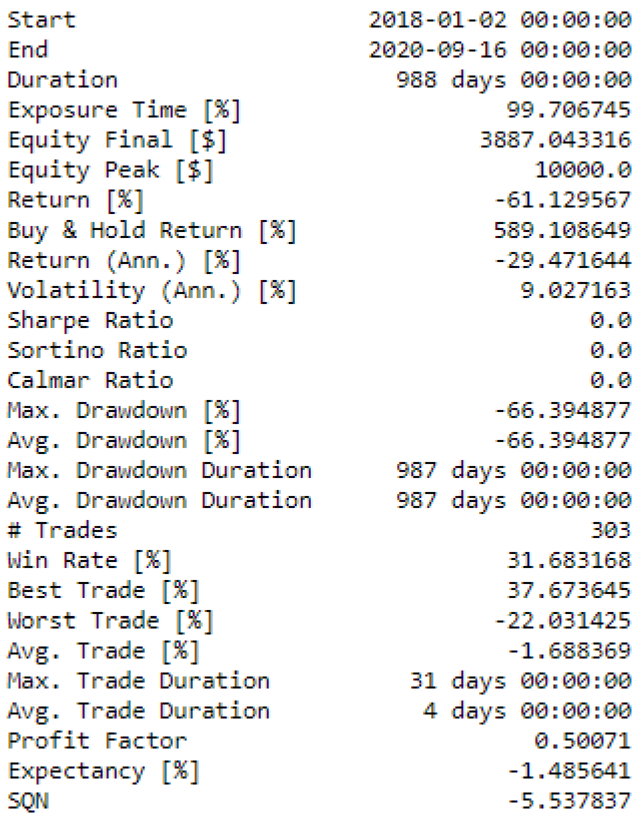
While this could be an exception of TSLA and for other stocks we could have more variety of sentiment, we can also try comparing the maximum and minimum sentiment only, and this in turn gives significantly more trades:



```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 13312.686556
Equity Peak [$]                  13372.026644
Return [%]                          33.126866
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   11.151764
Volatility (Ann.) [%]               14.810496
Sharpe Ratio                         0.752964
Sortino Ratio                        1.229607
Calmar Ratio                         0.761321
Max. Drawdown [%]                  -14.647915
Avg. Drawdown [%]                   -3.018186
Max. Drawdown Duration      738 days 00:00:00
Avg. Drawdown Duration       70 days 00:00:00
# Trades                                  303
Win Rate [%]                        53.465347
Best Trade [%]                      40.427117
Worst Trade [%]                    -20.472054
Avg. Trade [%]                       0.330589
Max. Trade Duration          31 days 00:00:00
Avg. Trade Duration           4 days 00:00:00
Profit Factor                        1.295106
Expectancy [%]                       0.531736
SQN                                  1.481826
```



The overall win rate is good too, but there is a catch: "commission". Once we add a 2% commission for trading the stock (not to mention possible slippage) our profit plummets:



```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                  3887.043316
Equity Peak [$]                       10000.0
Return [%]                         -61.129567
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                  -29.471644
Volatility (Ann.) [%]                9.027163
Sharpe Ratio                              0.0
Sortino Ratio                             0.0
Calmar Ratio                              0.0
Max. Drawdown [%]                  -66.394877
Avg. Drawdown [%]                  -66.394877
Max. Drawdown Duration      987 days 00:00:00
Avg. Drawdown Duration      987 days 00:00:00
# Trades                                  303
Win Rate [%]                        31.683168
Best Trade [%]                      37.673645
Worst Trade [%]                    -22.031425
Avg. Trade [%]                      -1.688369
Max. Trade Duration          31 days 00:00:00
Avg. Trade Duration           4 days 00:00:00
Profit Factor                         0.50071
Expectancy [%]                      -1.485641
SQN                                 -5.537837
```

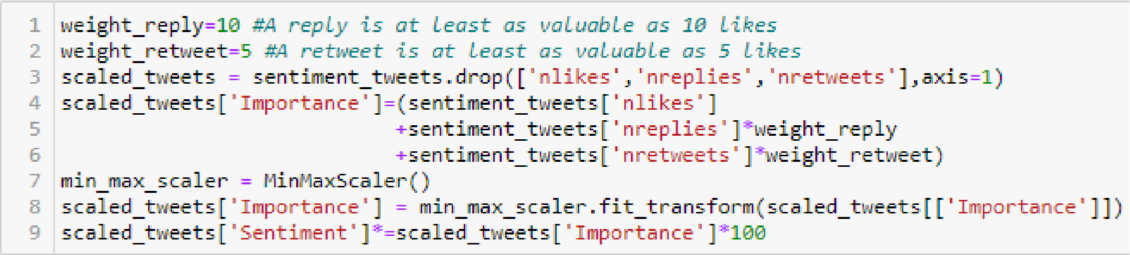
Quick points we can explore and improve upon:

1- We can experiment with weights using the number of likes, retweets and replies.
2- We can experiment with the sentiment analyzer lexicon, adding more words and tuning weights for existing words, or trying another sentiment analyzer.
3- We can test on other price data that isn’t extremely bullish to check if the overall tweet sentiment is positive, and if we can confirm that, on average, tweet data is positive, apply a bigger multiplier to negative tweets to offset the positive ones and create some equilibrium.
4- We can experiment with better money management techniques, like trailing stops or incremental positions.
5- We can experiment combining our sentiment with technical indicators.
6- We can experiment combining our sentiment with other ML models such as RNN or random forest ensembles.
7- We can experiment adding the weekend sentiment onto the Monday candle.
8- We can experiment with weights using the username, but this is not scalable to other markets or even other news sources.
9- We can add market specific data, such as the beta of the stock, their fundamental data (such as earnings, etc…), TICK data, compare it with other similar stocks in the sector or aggregate their data, etc…

Second iteration: Using weighted sentiment based on likes/retweets/replies

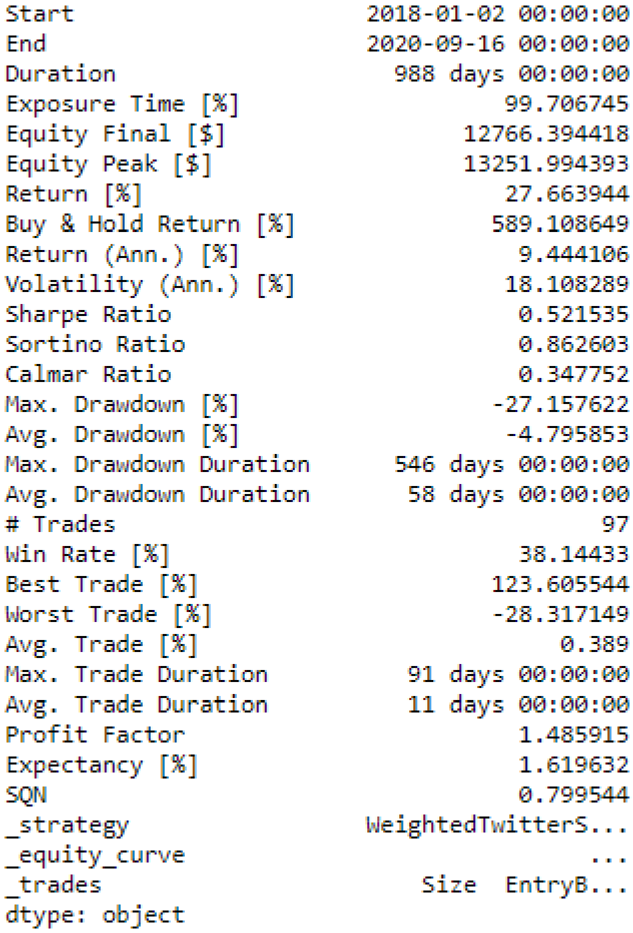
On a second iteration we can experiment with multiplying the compound score based on the number of likes, retweets and replies, as we assume more popular tweets have more impact on the underlying price.

Based on the small sample available I estimated that a retweet is roughly 5 times more important than a like and that a reply is roughly 10 times more important in determining the intensity of the tweet. Using these weights and multiplying this with our compound sentiment score:



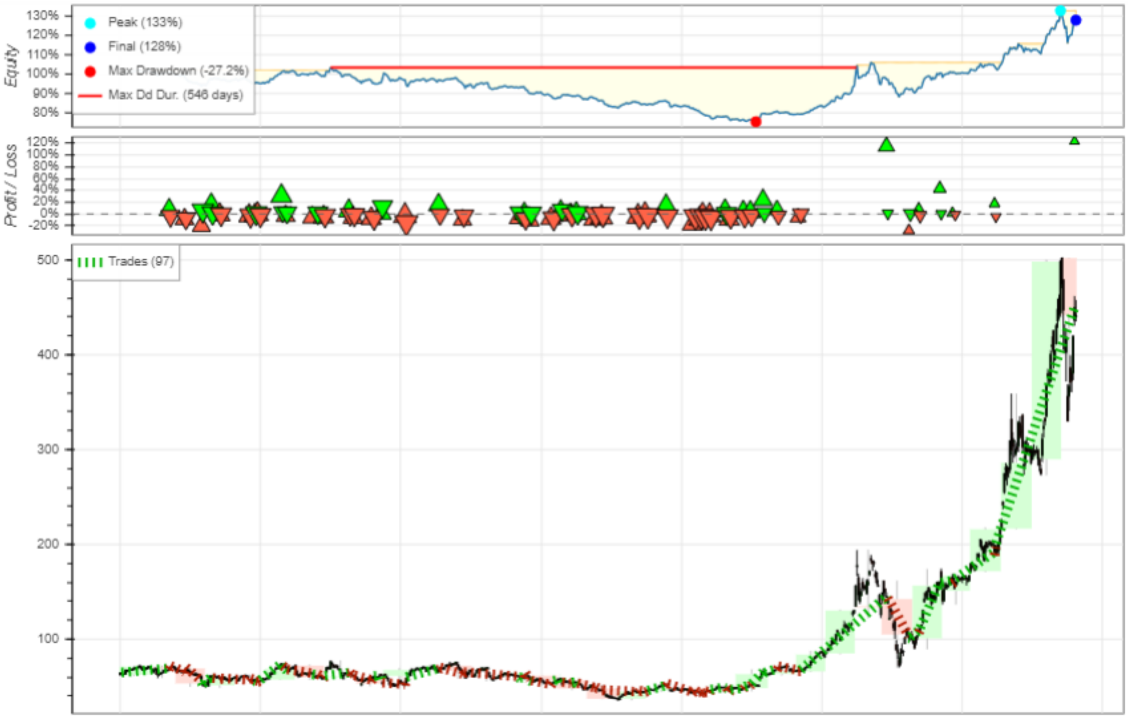
```
weight_reply=10 #A reply is at least as valuable as 10 likes
weight_retweet=5 #A retweet is at least as valuable as 5 likes
scaled_tweets = sentiment_tweets.drop(['nlikes','nreplies','nretweets'],axis=1)
scaled_tweets['Importance']=(sentiment_tweets['nlikes']
                             +sentiment_tweets['nreplies']*weight_reply
                             +sentiment_tweets['nretweets']*weight_retweet)
min_max_scaler = MinMaxScaler()
scaled_tweets['Importance'] = min_max_scaler.fit_transform(scaled_tweets[['Importance']])
scaled_tweets['Sentiment']*=scaled_tweets['Importance']*100
```

Using a simple “buy if sentiment is positive and sell if sentiment is negative” strategy:

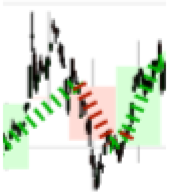


```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 12766.394418
Equity Peak [$]                  13251.994393
Return [%]                          27.663944
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                    9.444106
Volatility (Ann.) [%]               18.108289
Sharpe Ratio                         0.521535
Sortino Ratio                        0.862603
Calmar Ratio                         0.347752
Max. Drawdown [%]                  -27.157622
Avg. Drawdown [%]                   -4.795853
Max. Drawdown Duration      546 days 00:00:00
Avg. Drawdown Duration       58 days 00:00:00
# Trades                                   97
Win Rate [%]                         38.14433
Best Trade [%]                     123.605544
Worst Trade [%]                    -28.317149
Avg. Trade [%]                          0.389
Max. Trade Duration          91 days 00:00:00
Avg. Trade Duration          11 days 00:00:00
Profit Factor                        1.485915
Expectancy [%]                       1.619632
SQN                                  0.799544
_strategy                 WeightedTwitterS...
_equity_curve                             ...
_trades                       Size  EntryB...
dtype: object
```

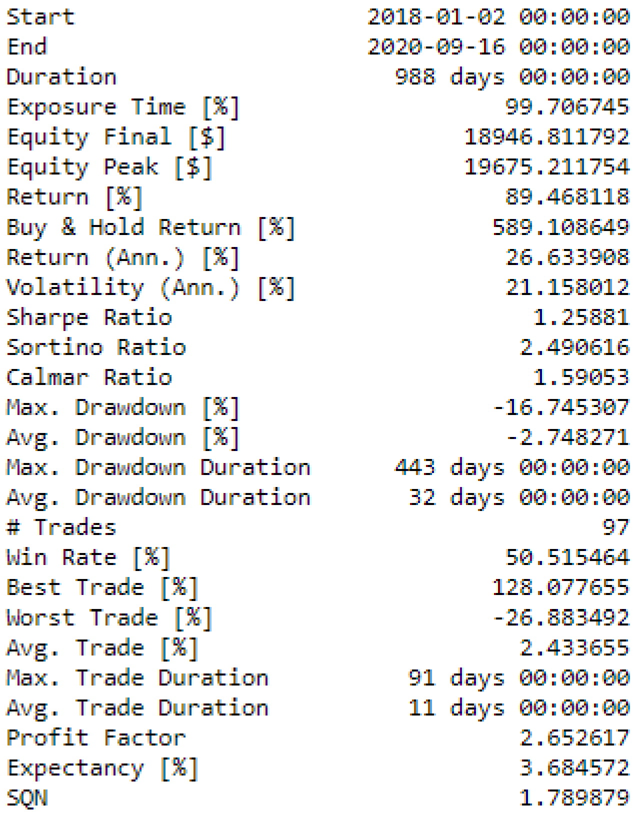
Considering that we’re paying 2% commission and that we manage to end up with positive equity I think it’s on overall pretty decent.



We notice that before the stock went parabolic there was indecision in the trades, mostly costing us money due to commissions, but afterwards the overall sentiment was very positive and the algorithm fully traded it, with a small short trade during a bearish moment:

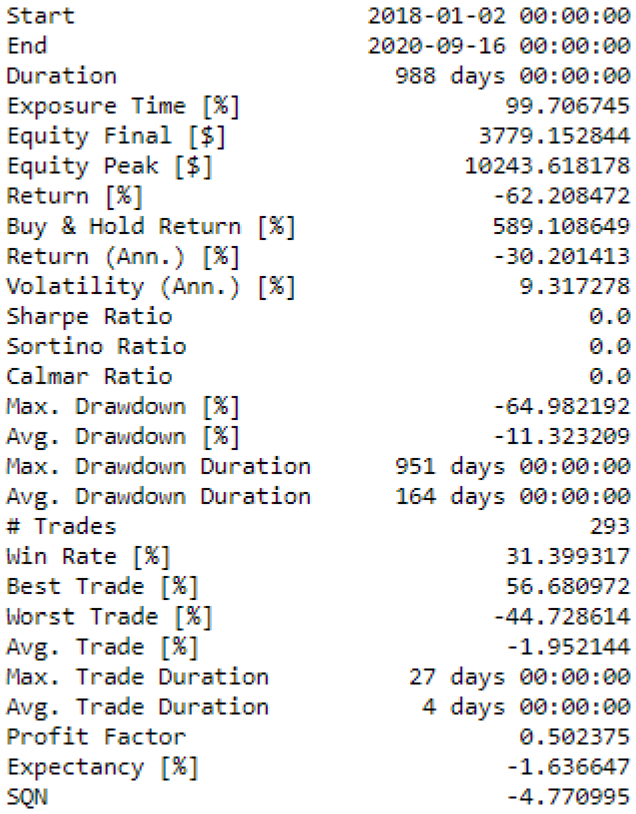


Removing commission gives us a win rate slightly above 50% but a much higher final equity:



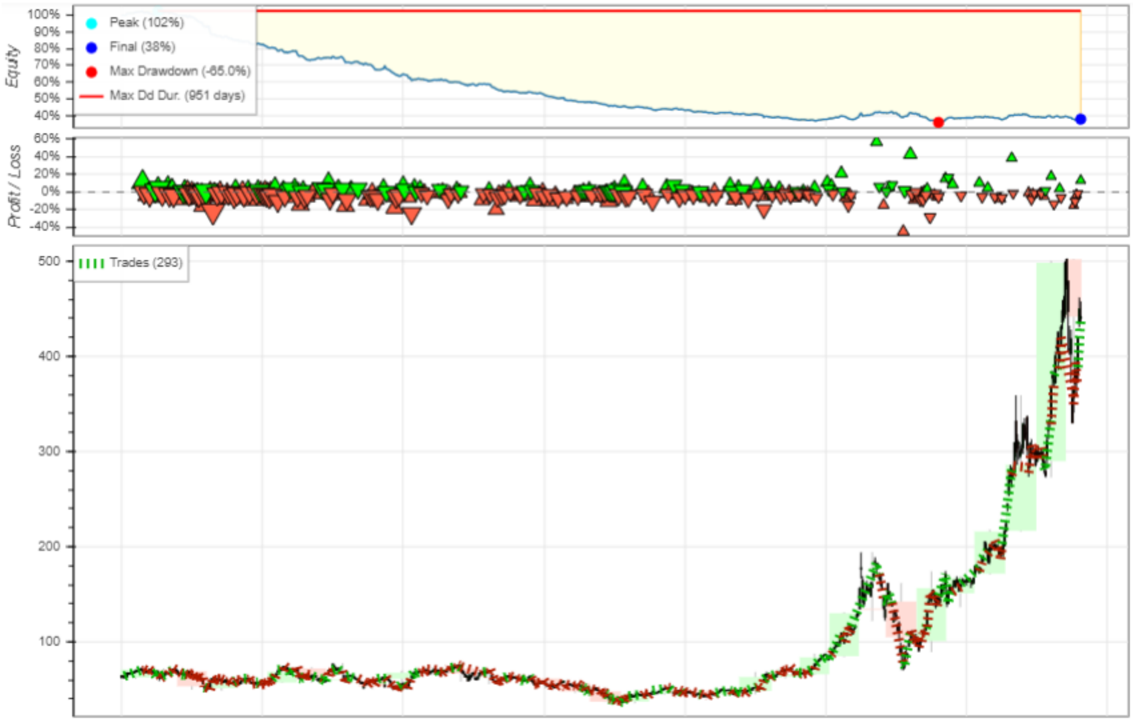
```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 18946.811792
Equity Peak [$]                  19675.211754
Return [%]                          89.468118
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   26.633908
Volatility (Ann.) [%]               21.158012
Sharpe Ratio                          1.25881
Sortino Ratio                        2.490616
Calmar Ratio                          1.59053
Max. Drawdown [%]                  -16.745307
Avg. Drawdown [%]                   -2.748271
Max. Drawdown Duration      443 days 00:00:00
Avg. Drawdown Duration       32 days 00:00:00
# Trades                                   97
Win Rate [%]                        50.515464
Best Trade [%]                     128.077655
Worst Trade [%]                    -26.883492
Avg. Trade [%]                       2.433655
Max. Trade Duration          91 days 00:00:00
Avg. Trade Duration          11 days 00:00:00
Profit Factor                        2.652617
Expectancy [%]                       3.684572
SQN                                  1.789879
```

Now, how does this change affect the maximum over minimum sentiment strategy that we had before?



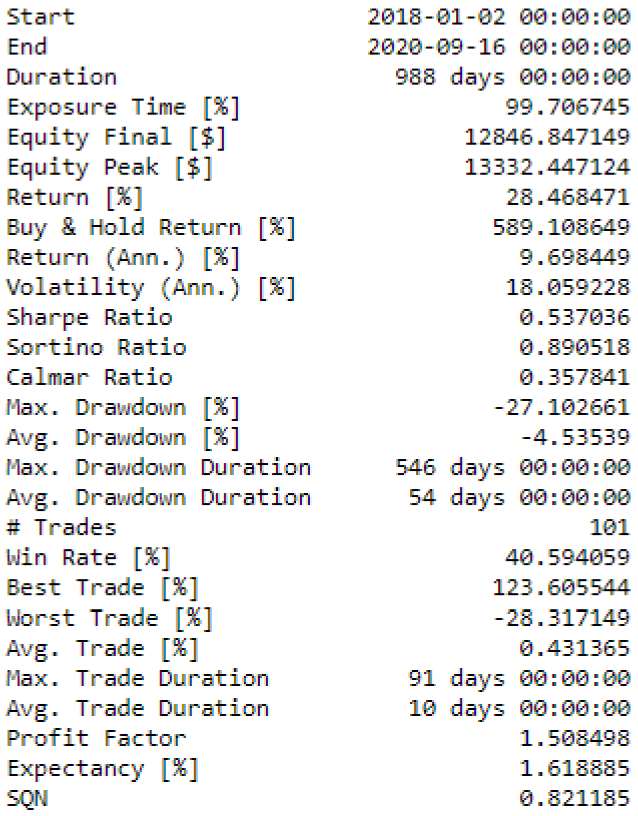
```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                  3779.152844
Equity Peak [$]                  10243.618178
Return [%]                         -62.208472
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                  -30.201413
Volatility (Ann.) [%]                9.317278
Sharpe Ratio                              0.0
Sortino Ratio                             0.0
Calmar Ratio                              0.0
Max. Drawdown [%]                  -64.982192
Avg. Drawdown [%]                  -11.323209
Max. Drawdown Duration      951 days 00:00:00
Avg. Drawdown Duration      164 days 00:00:00
# Trades                                  293
Win Rate [%]                        31.399317
Best Trade [%]                      56.680972
Worst Trade [%]                    -44.728614
Avg. Trade [%]                      -1.952144
Max. Trade Duration          27 days 00:00:00
Avg. Trade Duration           4 days 00:00:00
Profit Factor                        0.502375
Expectancy [%]                      -1.636647
SQN                                 -4.770995
```

It got worse, and we can see some very drastic short trades while the stock is going up:



At this point it’s very safe to conclude that using the maximum and minimum of the sentiment in a day isn’t nearly as helpful as obtaining the overall sentiment, so no more research will be done using those two.

A grid search for the optimal weights returned 19 for replies and 4 for retweets, as those values are probably overfitting to the data, I’m going to pick 20 and 5 for the following iterations. A quick test using these weights improves very slightly the sum of sentiment model.



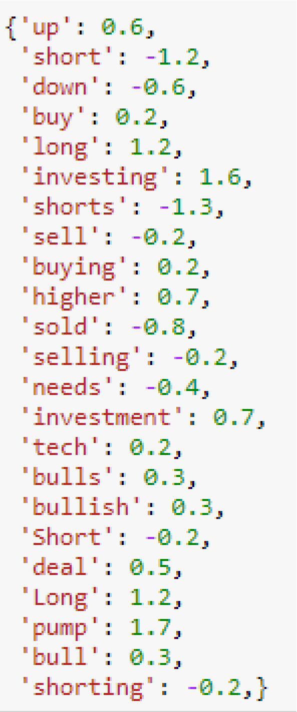
```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 12846.847149
Equity Peak [$]                  13332.447124
Return [%]                          28.468471
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                    9.698449
Volatility (Ann.) [%]               18.059228
Sharpe Ratio                         0.537036
Sortino Ratio                        0.890518
Calmar Ratio                         0.357841
Max. Drawdown [%]                  -27.102661
Avg. Drawdown [%]                    -4.53539
Max. Drawdown Duration      546 days 00:00:00
Avg. Drawdown Duration       54 days 00:00:00
# Trades                                  101
Win Rate [%]                        40.594059
Best Trade [%]                     123.605544
Worst Trade [%]                    -28.317149
Avg. Trade [%]                       0.431365
Max. Trade Duration          91 days 00:00:00
Avg. Trade Duration          10 days 00:00:00
Profit Factor                        1.508498
Expectancy [%]                       1.618885
SQN                                  0.821185
```

Third iteration: Lexicon expansion

Research on publicly available sentiment dictionaries to replace VADER’s original one doesn’t provide any upgrade. VADER has 7502 words in its dictionary, while most sentiment lexicons have at most 500.

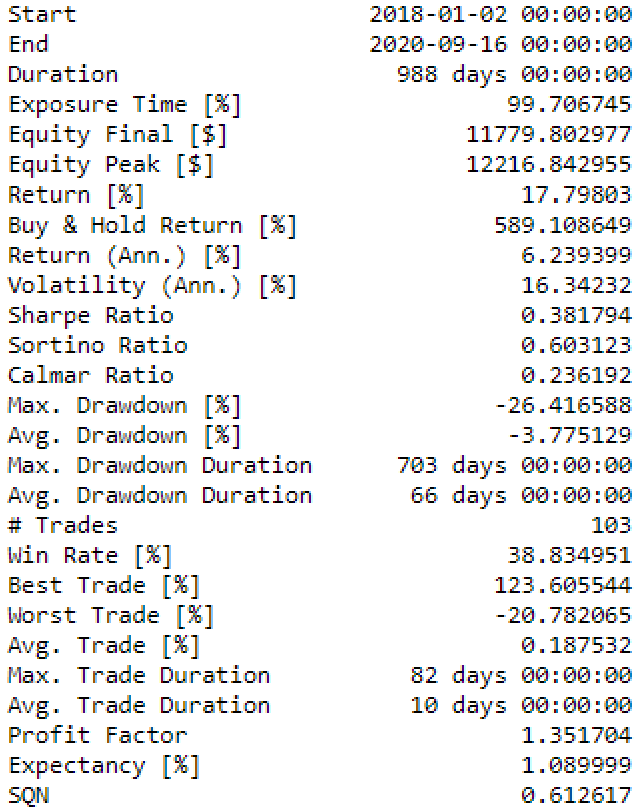
A first step would be to search for common words in our sentences that aren’t in VADER’s lexicon and add them in manually.

There are 825 common words that aren’t part of VADER, many of them are objective neutral words and thus won’t influence sentiment in any way. A small extension to the lexicon will be done with the following words:



```
{'up': 0.6,
 'short': -1.2,
 'down': -0.6,
 'buy': 0.2,
 'long': 1.2,
 'investing': 1.6,
 'shorts': -1.3,
 'sell': -0.2,
 'buying': 0.2,
 'higher': 0.7,
 'sold': -0.8,
 'selling': -0.2,
 'needs': -0.4,
 'investment': 0.7,
 'tech': 0.2,
 'bulls': 0.3,
 'bullish': 0.3,
 'Short': -0.2,
 'deal': 0.5,
 'Long': 1.2,
 'pump': 1.7,
 'bull': 0.3,
 'shorting': -0.2,}
```

The extension of the lexicon wasn't exactly great as we can observe:



```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 11779.802977
Equity Peak [$]                  12216.842955
Return [%]                           17.79803
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                    6.239399
Volatility (Ann.) [%]                16.34232
Sharpe Ratio                         0.381794
Sortino Ratio                        0.603123
Calmar Ratio                         0.236192
Max. Drawdown [%]                  -26.416588
Avg. Drawdown [%]                   -3.775129
Max. Drawdown Duration      703 days 00:00:00
Avg. Drawdown Duration       66 days 00:00:00
# Trades                                  103
Win Rate [%]                        38.834951
Best Trade [%]                     123.605544
Worst Trade [%]                    -20.782065
Avg. Trade [%]                       0.187532
Max. Trade Duration          82 days 00:00:00
Avg. Trade Duration          10 days 00:00:00
Profit Factor                        1.351704
Expectancy [%]                       1.089999
SQN                                  0.612617
```

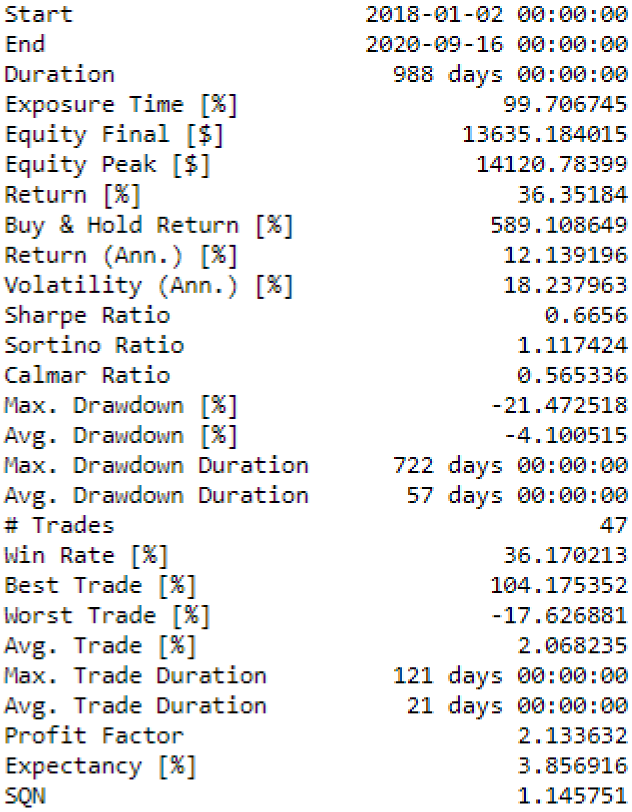
Manually fine tuning the weights or modifying other words could probably improve the results but this proves that just adding missing words won't influence the final results much when you're reading hundreds of thousands of words. Possible alternatives would include generating our own sentiment lexicon (a non-trivial matter) or do a massive grid search of optimal weights (an extremely expensive computational task that is likely to overfit).

Fourth Iteration: Using other sentiment analyzers.

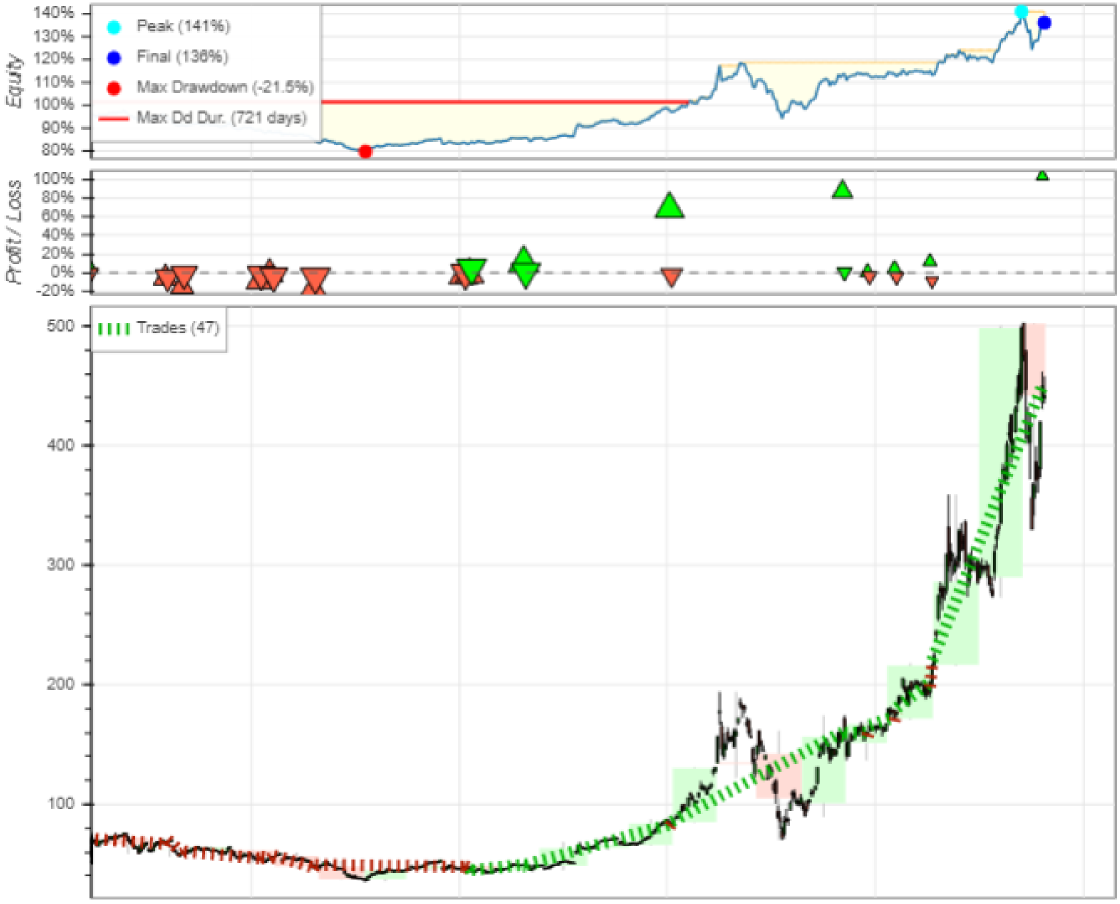
There are two approaches: the first one is using sentiment lexicon that work on a variety of text like VADER, the other approach is to build a supervised learning model, but this would need to first label the tweets as positive or negative. As we don’t have the label for each tweet the only way is to use a generic sentiment analyzer.

SentiWordnet

SentiWordnet is extremely slow and in messy twitter data it gives a lot of neutrals, as a result we end up with significantly less trades, but slightly better performance:



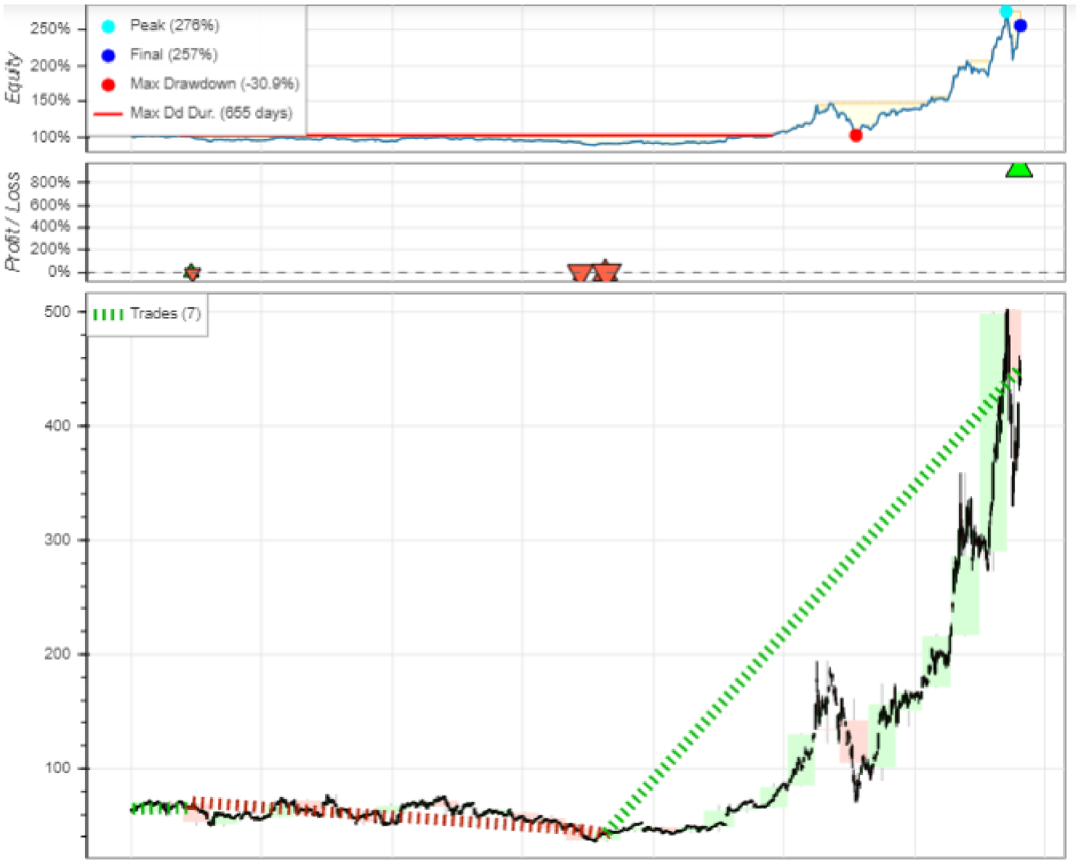
```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 13635.184015
Equity Peak [$]                   14120.78399
Return [%]                           36.35184
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   12.139196
Volatility (Ann.) [%]               18.237963
Sharpe Ratio                           0.6656
Sortino Ratio                        1.117424
Calmar Ratio                         0.565336
Max. Drawdown [%]                  -21.472518
Avg. Drawdown [%]                   -4.100515
Max. Drawdown Duration      722 days 00:00:00
Avg. Drawdown Duration       57 days 00:00:00
# Trades                                   47
Win Rate [%]                        36.170213
Best Trade [%]                     104.175352
Worst Trade [%]                    -17.626881
Avg. Trade [%]                       2.068235
Max. Trade Duration         121 days 00:00:00
Avg. Trade Duration          21 days 00:00:00
Profit Factor                        2.133632
Expectancy [%]                       3.856916
SQN                                  1.145751
```



Textblob

Textblob is also widely used for NLP and sentiment analysis. It’s simple to use and gives decent results. We can see here that it makes 7 trades, with the big ones making a huge impact on the profit.

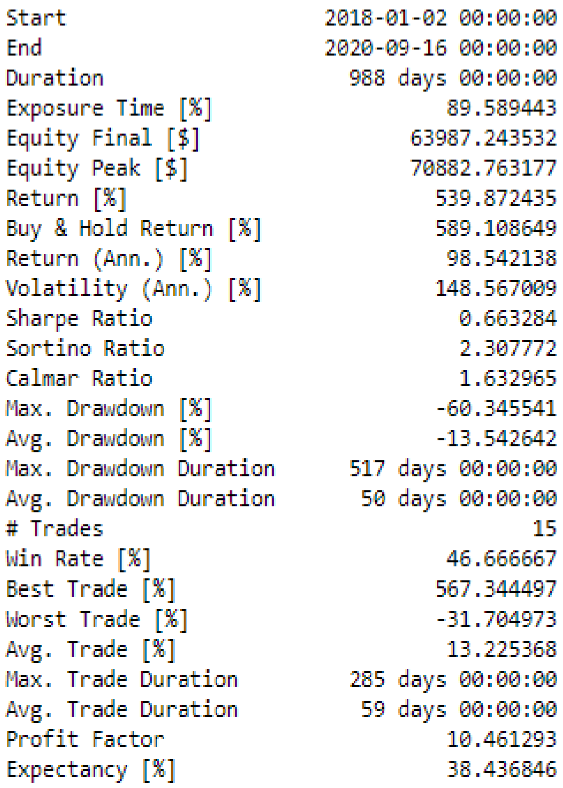
```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 25657.477331
Equity Peak [$]                  27648.437229
Return [%]                         156.574773
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   41.646274
Volatility (Ann.) [%]               39.532599
Sharpe Ratio                         1.053467
Sortino Ratio                        2.392818
Calmar Ratio                         1.345964
Max. Drawdown [%]                  -30.941587
Avg. Drawdown [%]                   -4.635817
Max. Drawdown Duration      658 days 00:00:00
Avg. Drawdown Duration       40 days 00:00:00
# Trades                                    7
Win Rate [%]                        28.571429
Best Trade [%]                     925.826294
Worst Trade [%]                     -34.18368
Avg. Trade [%]                      27.942965
Max. Trade Duration         459 days 00:00:00
Avg. Trade Duration         141 days 00:00:00
Profit Factor                       17.466916
Expectancy [%]                     124.837607
SQN                                  0.931023
```



The biggest demerit of textblob is the huge amount of neutrals. As such we end up with very little trades. To try another approach I changed the conditions for opening and closing trades based on the intensity of the sentiment, and not only on if being positive or not. I also decided to close trades when there wasn't a prevalent sentiment and we were mostly flat.

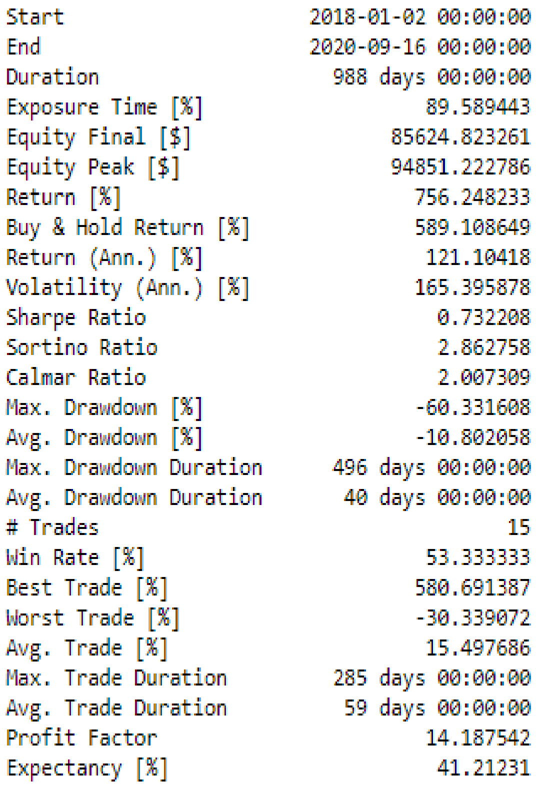
This doubles the number of trades although we also end up paying more in commission. In a commission free world where we use all our equity we even beat the buy & hold approach.

2% commission per trade

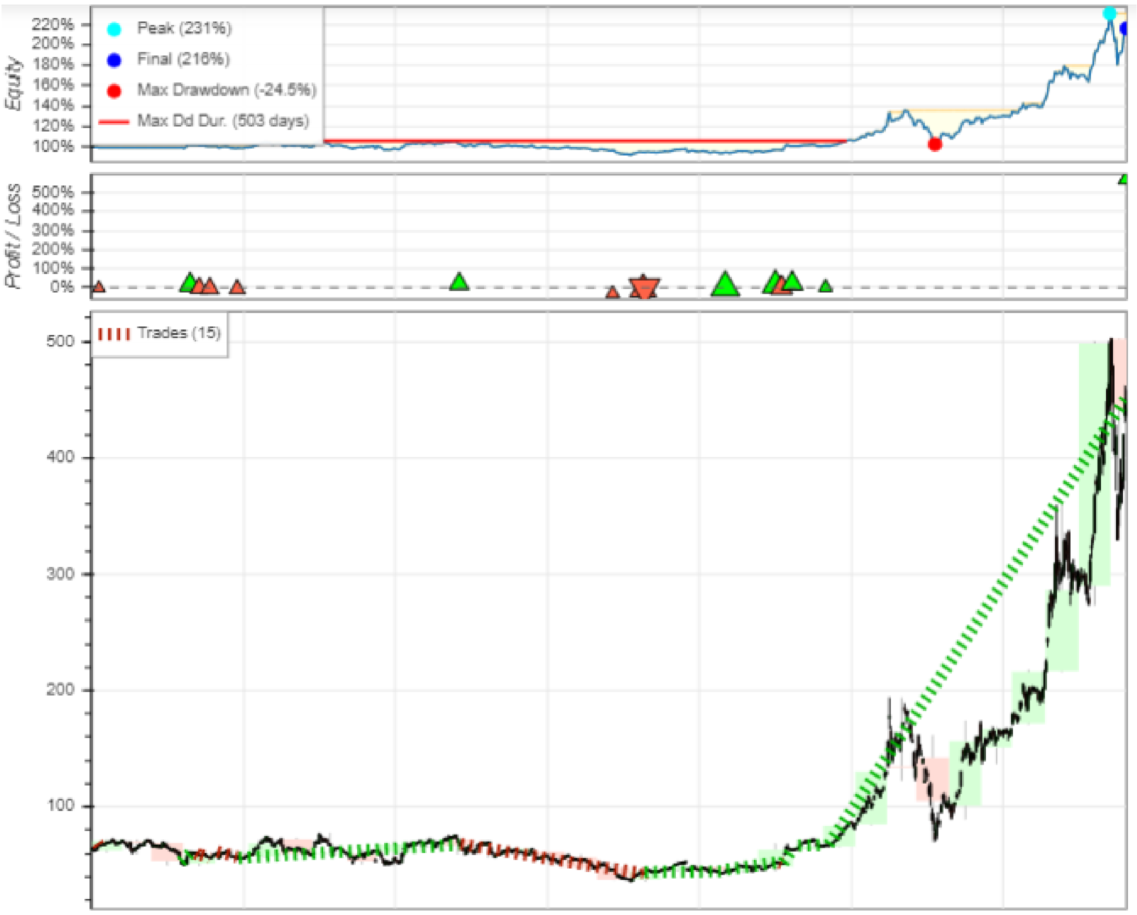


```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   89.589443
Equity Final [$]                 63987.243532
Equity Peak [$]                  70882.763177
Return [%]                         539.872435
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   98.542138
Volatility (Ann.) [%]              148.567009
Sharpe Ratio                         0.663284
Sortino Ratio                        2.307772
Calmar Ratio                         1.632965
Max. Drawdown [%]                  -60.345541
Avg. Drawdown [%]                  -13.542642
Max. Drawdown Duration      517 days 00:00:00
Avg. Drawdown Duration       50 days 00:00:00
# Trades                                   15
Win Rate [%]                        46.666667
Best Trade [%]                     567.344497
Worst Trade [%]                    -31.704973
Avg. Trade [%]                      13.225368
Max. Trade Duration         285 days 00:00:00
Avg. Trade Duration          59 days 00:00:00
Profit Factor                       10.461293
Expectancy [%]                      38.436846
```

No commission per trade.



```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   89.589443
Equity Final [$]                 85624.823261
Equity Peak [$]                  94851.222786
Return [%]                         756.248233
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   121.10418
Volatility (Ann.) [%]              165.395878
Sharpe Ratio                         0.732208
Sortino Ratio                        2.862758
Calmar Ratio                         2.007309
Max. Drawdown [%]                  -60.331608
Avg. Drawdown [%]                  -10.802058
Max. Drawdown Duration      496 days 00:00:00
Avg. Drawdown Duration       40 days 00:00:00
# Trades                                   15
Win Rate [%]                        53.333333
Best Trade [%]                     580.691387
Worst Trade [%]                    -30.339072
Avg. Trade [%]                      15.497686
Max. Trade Duration         285 days 00:00:00
Avg. Trade Duration          59 days 00:00:00
Profit Factor                       14.187542
Expectancy [%]                       41.21231
```

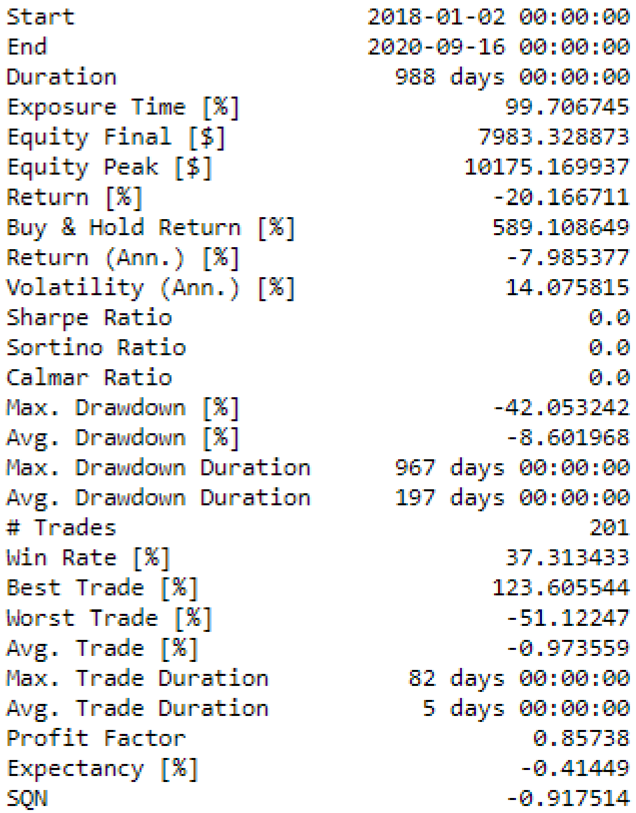


We can observe that while the general idea is good, it could use some fine tuning, specially on closing and entering, this is the same issue that VADER and Sentinet had.

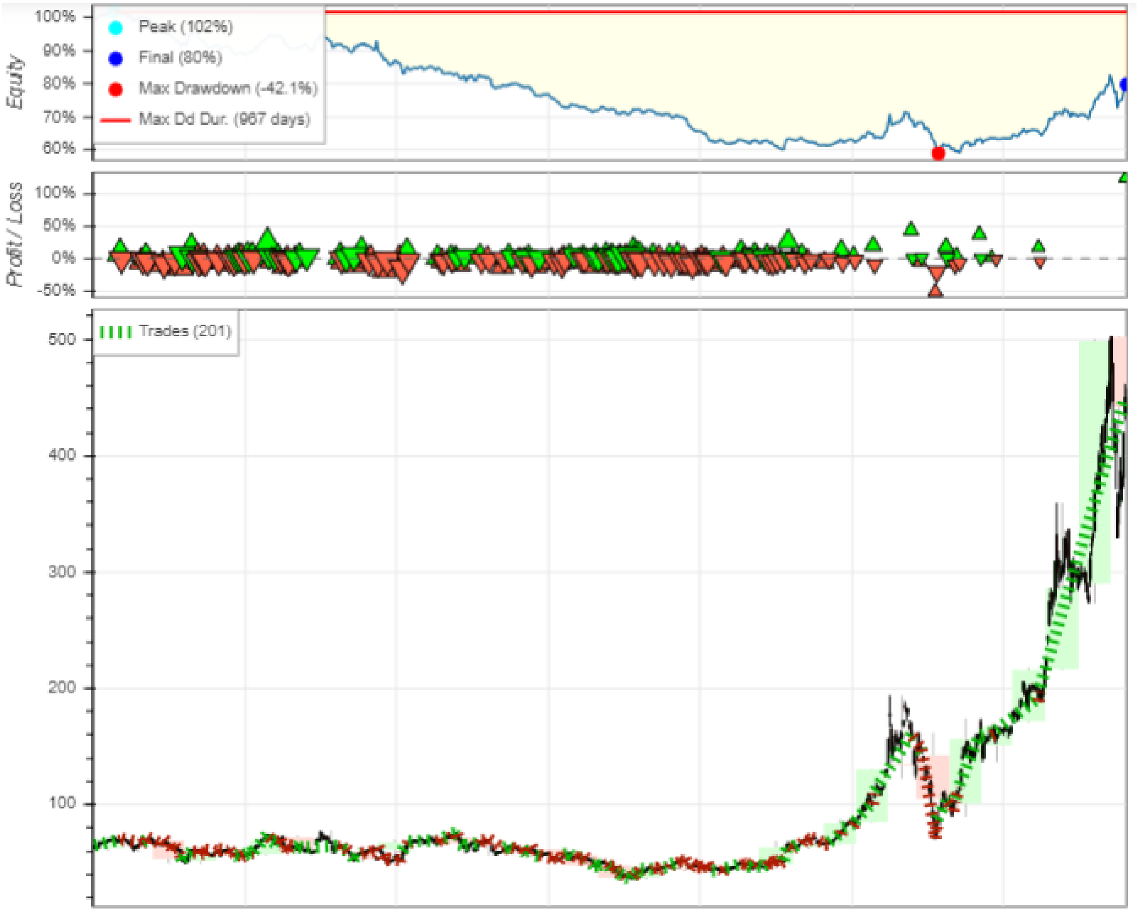
As such the core improvement will surely be based on polishing those entry and exit conditions.

Afinn

Afinn is the final library for sentiment analysis I'm going to try in this project. It's considerably slow (on par with sentiwordnet) and doesn't really perform too well on messy social media data like tweets. We can confirm that using our standard test bed it doesn't really perform too well, although it can improve a bit tuning the entry conditions to only do trades when there is a substantial amount of sentiment, but still below Textblob.

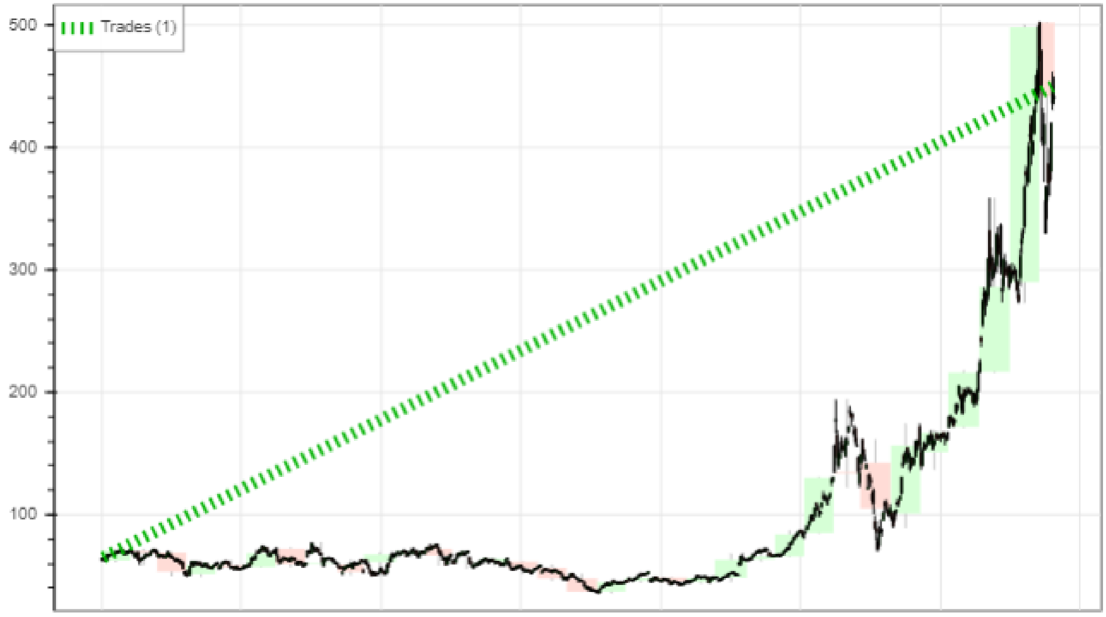


```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                  7983.328873
Equity Peak [$]                  10175.169937
Return [%]                         -20.166711
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   -7.985377
Volatility (Ann.) [%]               14.075815
Sharpe Ratio                              0.0
Sortino Ratio                             0.0
Calmar Ratio                              0.0
Max. Drawdown [%]                  -42.053242
Avg. Drawdown [%]                   -8.601968
Max. Drawdown Duration      967 days 00:00:00
Avg. Drawdown Duration      197 days 00:00:00
# Trades                                  201
Win Rate [%]                        37.313433
Best Trade [%]                     123.605544
Worst Trade [%]                     -51.12247
Avg. Trade [%]                      -0.973559
Max. Trade Duration          82 days 00:00:00
Avg. Trade Duration           5 days 00:00:00
Profit Factor                         0.85738
Expectancy [%]                       -0.41449
SQN                                 -0.917514
```



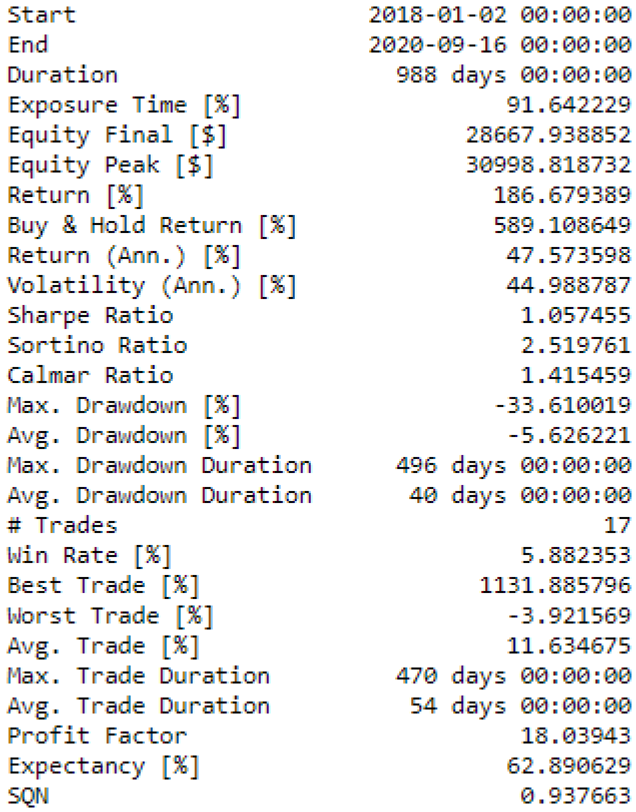
Ensemble system, combining multiple sentiment analyzers

A final idea would be to combine the analysis of the different analyzers and scale them, so none of them is more important than the others. Doing this process gives us a very simple buy & hold strategy again:

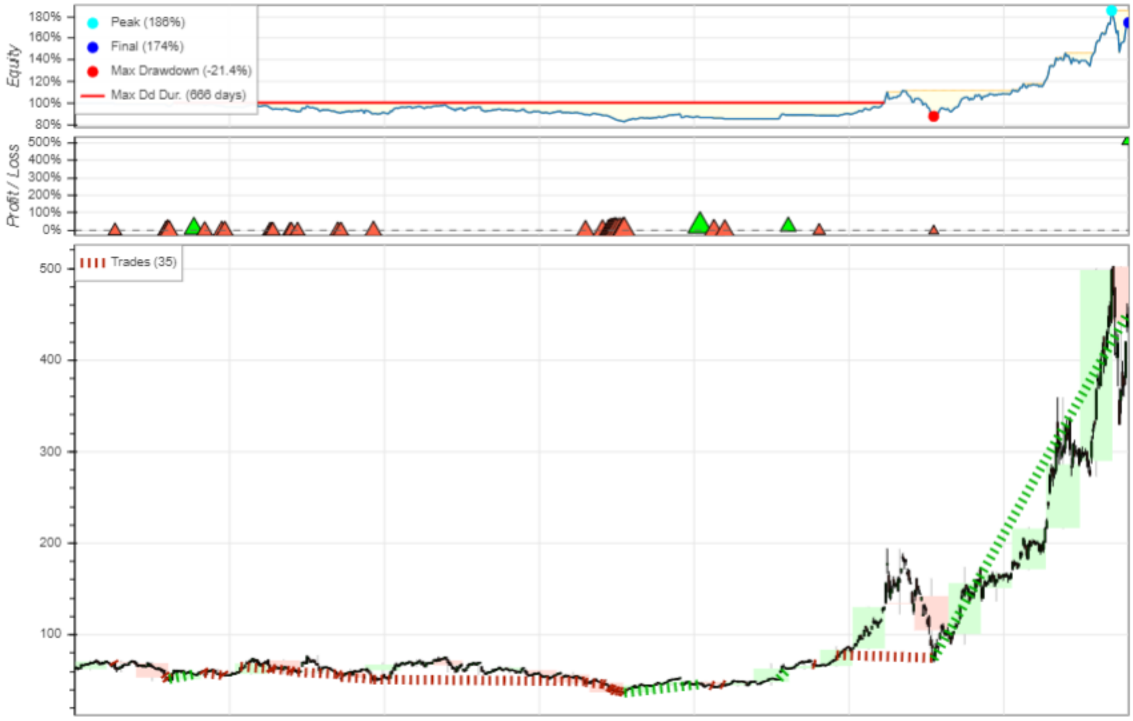


This is a very interesting conclusion, and it would certainly be interesting to apply the same system to a company that has gone bankrupt instead.

However, while this could just be the case of TSLA tweets being extremely positive, we could adjust a bit our entries as we have done before, and add some stop loss for good measure as well. By adding a stop loss at 2% price variation and buying or shorting only on a significant sentiment change we can achieve a strategy with very low win rate (thanks to the stop loss) but very high P/L.



```
Start                       2018-01-02 00:00:00
End                         2020-09-16 00:00:00
Duration                      988 days 00:00:00
Exposure Time [%]                     91.642229
Equity Final [$]                   28667.938852
Equity Peak [$]                    30998.818732
Return [%]                           186.679389
Buy & Hold Return [%]                589.108649
Return (Ann.) [%]                     47.573598
Volatility (Ann.) [%]                 44.988787
Sharpe Ratio                           1.057455
Sortino Ratio                          2.519761
Calmar Ratio                           1.415459
Max. Drawdown [%]                    -33.610019
Avg. Drawdown [%]                     -5.626221
Max. Drawdown Duration        496 days 00:00:00
Avg. Drawdown Duration         40 days 00:00:00
# Trades                                     17
Win Rate [%]                           5.882353
Best Trade [%]                      1131.885796
Worst Trade [%]                       -3.921569
Avg. Trade [%]                        11.634675
Max. Trade Duration           470 days 00:00:00
Avg. Trade Duration            54 days 00:00:00
Profit Factor                          18.03943
Expectancy [%]                        62.890629
SQN                                    0.937663
```



Conclusion

All four sentiment analysis tools give a pretty good idea of general direction, even Afinn. It's safe to conclude that the information obtained from sentiment analysis can be a powerful tool for deciding our trades, but it needs to be fine-tuned in order to maximize profits with better entries and exits.

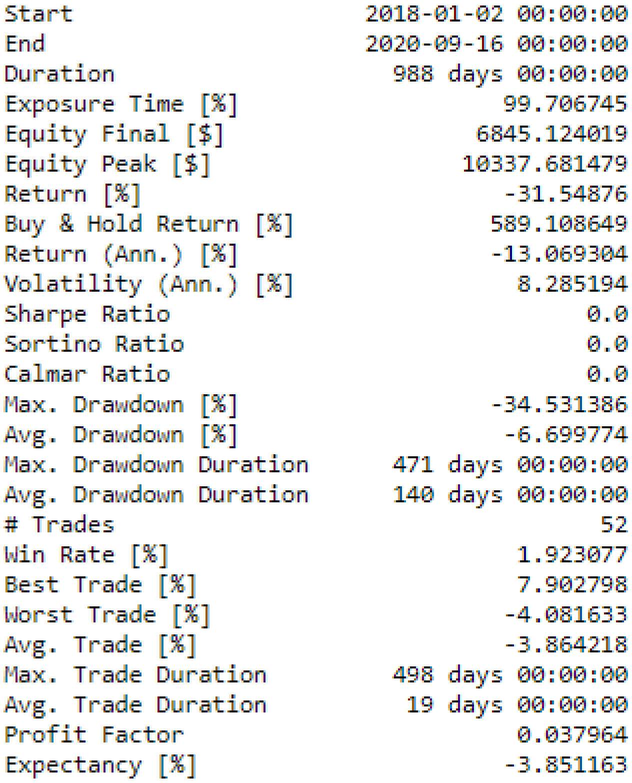
Vader is solid overall and works fast.
SentiwordNet is very good and gives a balanced amount of sentiment information but it's very slow.
Textblob gave us the best results, but the very small amount of trades could lead to overfitting or even coincidence.
Afinn was disappointing but modifying the conditions for opening the trades themselves it does give decent results.

Combining them all into an ensemble is a very powerful idea, which combined with some basic money management concepts provided impressive returns, even with a 2% commission per trade.

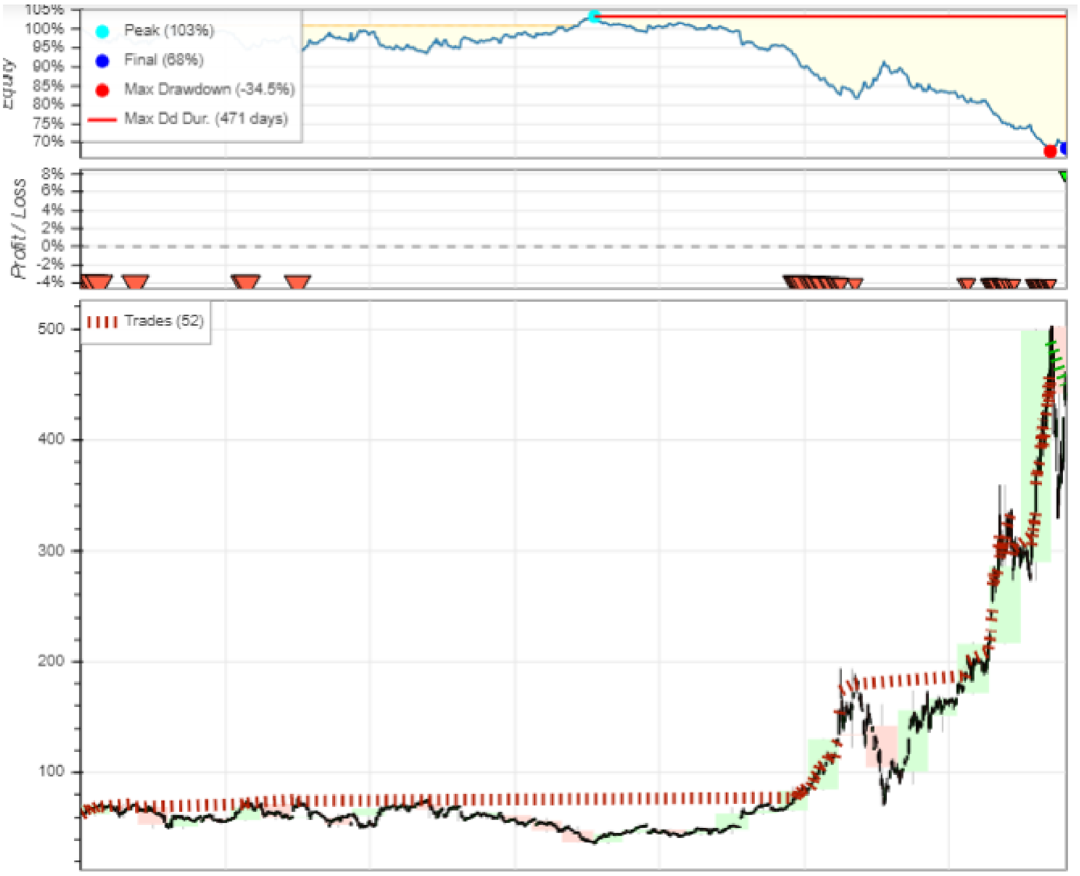
There is more that can be done in the field of sentiment analysis, like using Deep neural network systems or supervised learning systems, but lexicon approaches gave good results overall.

Fifth iteration: Adding negative bias to tweets to try and compensate the fact that tweets are often positive.

Adding a small bias results in almost no change, while adding a significant bias just gives horrible results overall. It's not worth trying to account for social media bias directly by adding weights manually to the sentiment.

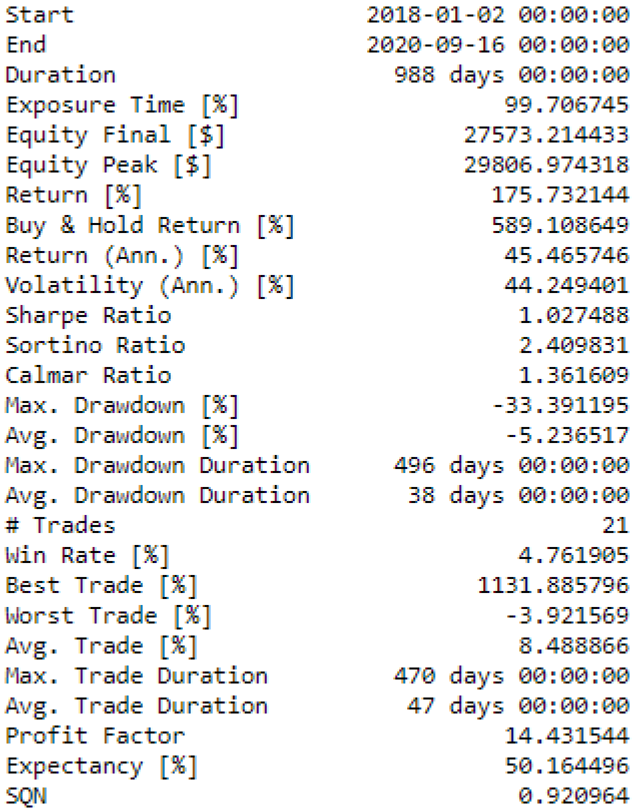


```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                  6845.124019
Equity Peak [$]                  10337.681479
Return [%]                          -31.54876
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                  -13.069304
Volatility (Ann.) [%]                8.285194
Sharpe Ratio                              0.0
Sortino Ratio                             0.0
Calmar Ratio                              0.0
Max. Drawdown [%]                  -34.531386
Avg. Drawdown [%]                   -6.699774
Max. Drawdown Duration      471 days 00:00:00
Avg. Drawdown Duration      140 days 00:00:00
# Trades                                   52
Win Rate [%]                         1.923077
Best Trade [%]                       7.902798
Worst Trade [%]                     -4.081633
Avg. Trade [%]                      -3.864218
Max. Trade Duration         498 days 00:00:00
Avg. Trade Duration          19 days 00:00:00
Profit Factor                        0.037964
Expectancy [%]                      -3.851163
```

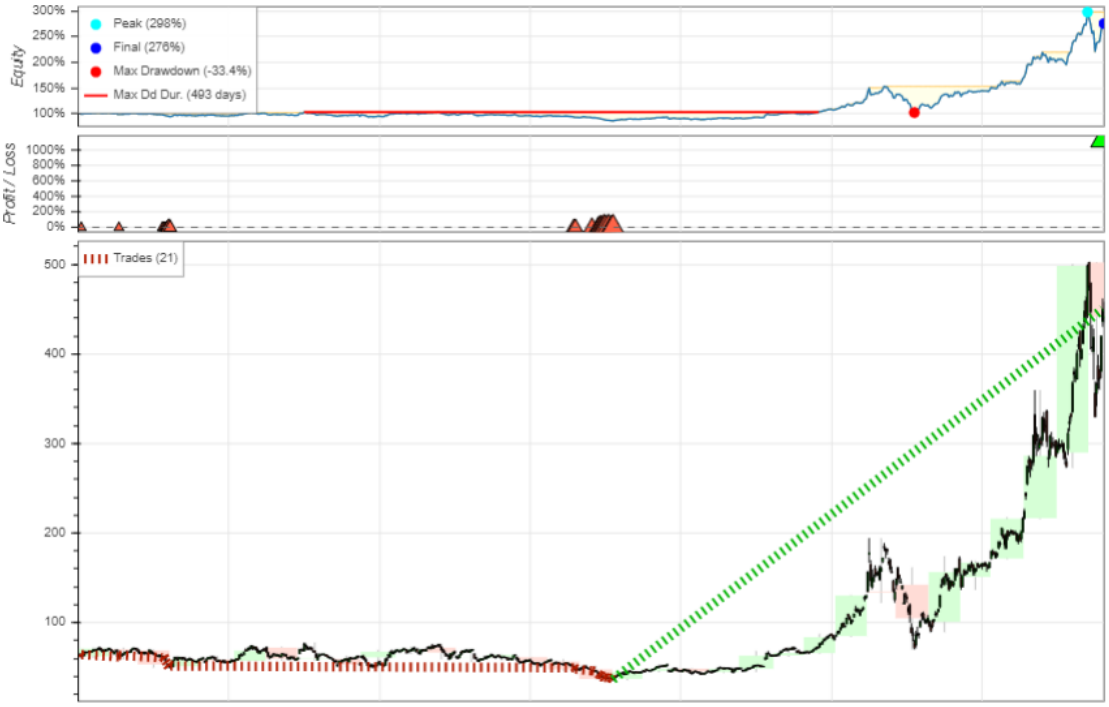


Sixth iteration. Amplifying overall sentiment based on activity

Just like on the 2nd iteration which gave good results overall, it’s of interest to amplify the overall sentiment at a given candle based on tweet activity: more tweets = more activity which should imply a bigger impact on price in either direction. Amplifying the sentiment based on the amount of tweets didn’t really affect our trades. This is due to the fact that if the overall sentiment is positive, having more tweets just means more positive sentiment, which in this case it just means that we can tune our entries based on smaller or bigger thresholds.



```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   99.706745
Equity Final [$]                 27573.214433
Equity Peak [$]                  29806.974318
Return [%]                         175.732144
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   45.465746
Volatility (Ann.) [%]               44.249401
Sharpe Ratio                         1.027488
Sortino Ratio                        2.409831
Calmar Ratio                         1.361609
Max. Drawdown [%]                  -33.391195
Avg. Drawdown [%]                   -5.236517
Max. Drawdown Duration      496 days 00:00:00
Avg. Drawdown Duration       38 days 00:00:00
# Trades                                   21
Win Rate [%]                         4.761905
Best Trade [%]                    1131.885796
Worst Trade [%]                     -3.921569
Avg. Trade [%]                       8.488866
Max. Trade Duration         470 days 00:00:00
Avg. Trade Duration          47 days 00:00:00
Profit Factor                       14.431544
Expectancy [%]                      50.164496
SQN                                  0.920964
```



Seventh iteration: Establishing more complex money management techniques.

Due to the small sample of data this is very hard to really test, but I'm going to adjust the ensemble system and try with varying stop losses and trailing stops, as well as establishing take profit values.

Multiple tests were made with stop losses values, trailing stops, take profits, incremental orders and varying the price delta or establishing windows of aggressive SL, amongst others.

The results are as following:

Too tight of a stop loss will create many losses, it's better to have a wide margin of stop loss and use a smaller amount of money instead.

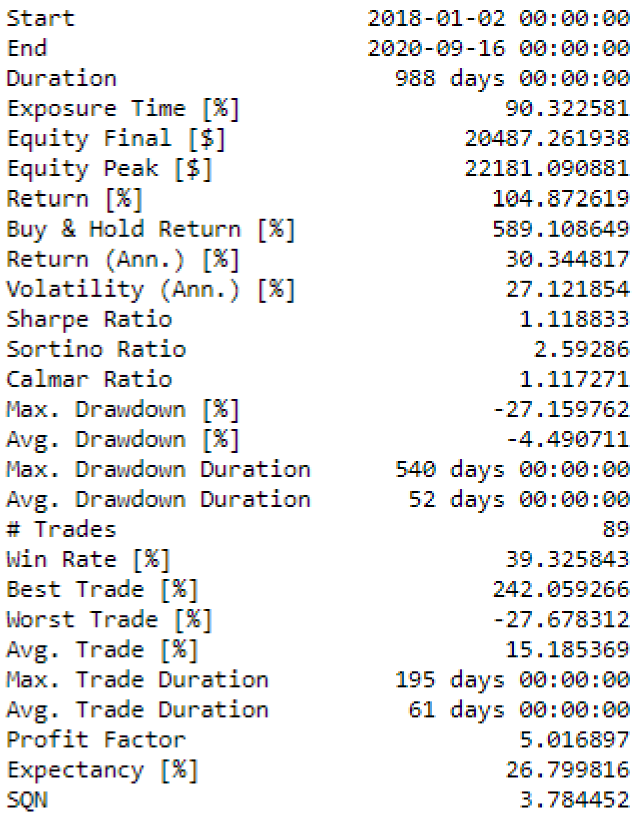
Margin can, and often is lethal and should only be considered for diversification.

In general stop losses can reduce your maximum P/L in backtesting, but that's mostly due to survivorship bias of data. Trailing stop losses can also help protect your profit once you're above break-even.

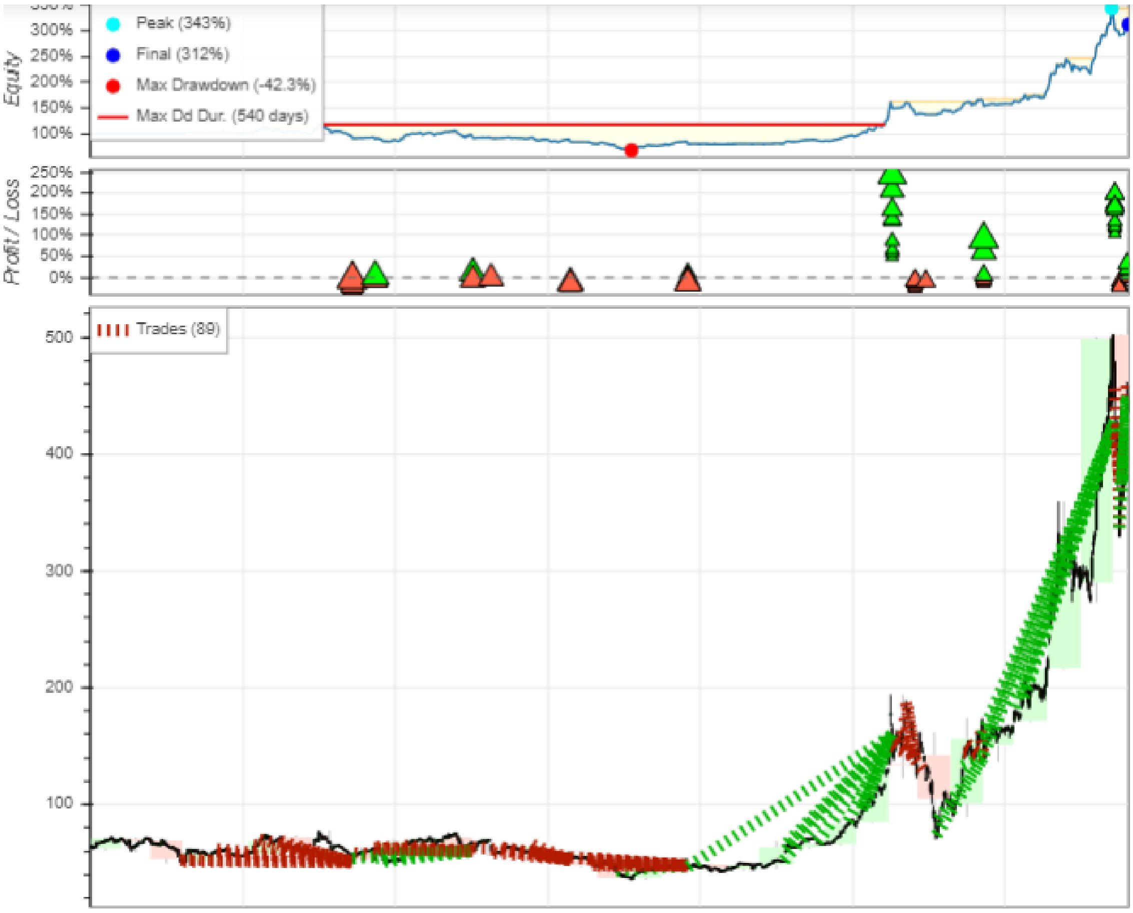
Take profits are in general detrimental, you can backtest and select optimal value for TP that magically hit some daily highs, but in general it just limits the profit of your trades and forces you to pay more commission. It can see more value during intraday trading, however.

Modifying the sentiment entry trigger doesn't have much of an impact, that's because the overall sentiment is incredibly positive during the entire time series. Trying to close trades when there is little sentiment fails, and trying to nitpick shorts based on low (but still positive) sentiment fails as well and it shouldn't really be done in the first place.

Incremental orders are interesting, using a fixed lot amount instead of variable equity and up to a fixed amount of risk it gives decent results at mitigating the drawdown.



```
Start                     2018-01-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    988 days 00:00:00
Exposure Time [%]                   90.322581
Equity Final [$]                 20487.261938
Equity Peak [$]                  22181.090881
Return [%]                         104.872619
Buy & Hold Return [%]              589.108649
Return (Ann.) [%]                   30.344817
Volatility (Ann.) [%]               27.121854
Sharpe Ratio                         1.118833
Sortino Ratio                         2.59286
Calmar Ratio                         1.117271
Max. Drawdown [%]                  -27.159762
Avg. Drawdown [%]                   -4.490711
Max. Drawdown Duration      540 days 00:00:00
Avg. Drawdown Duration       52 days 00:00:00
# Trades                                   89
Win Rate [%]                        39.325843
Best Trade [%]                     242.059266
Worst Trade [%]                    -27.678312
Avg. Trade [%]                      15.185369
Max. Trade Duration         195 days 00:00:00
Avg. Trade Duration          61 days 00:00:00
Profit Factor                        5.016897
Expectancy [%]                      26.799816
SQN                                  3.784452
```



Conclusion

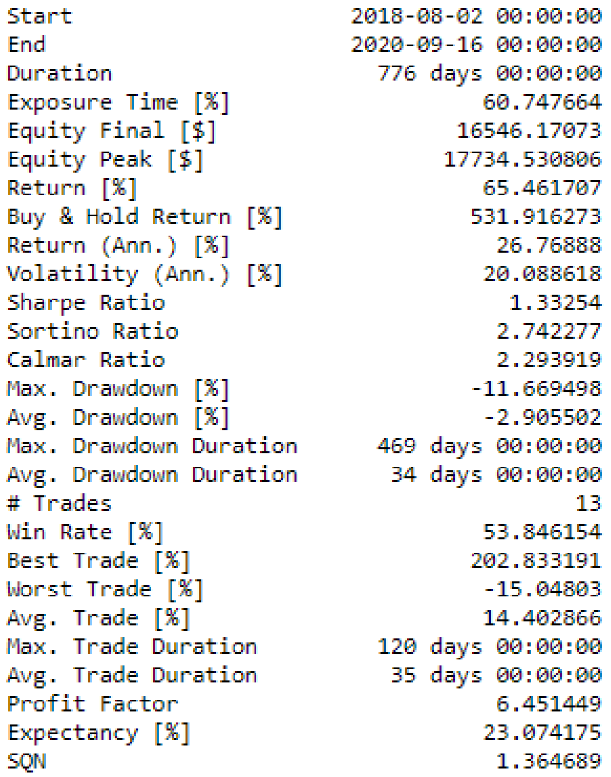
The traditional money management methods had a big impact on the models at first, during the previous iterations, but focusing completely on take profit, stop losses, trailing stops, etc… barely managed to diminish the drawdown and in general didn't help much in the trades, as it was only longs, even when it would have been better to not trade at all.

Eight Iteration: Sentiment + Technical indicator trading system

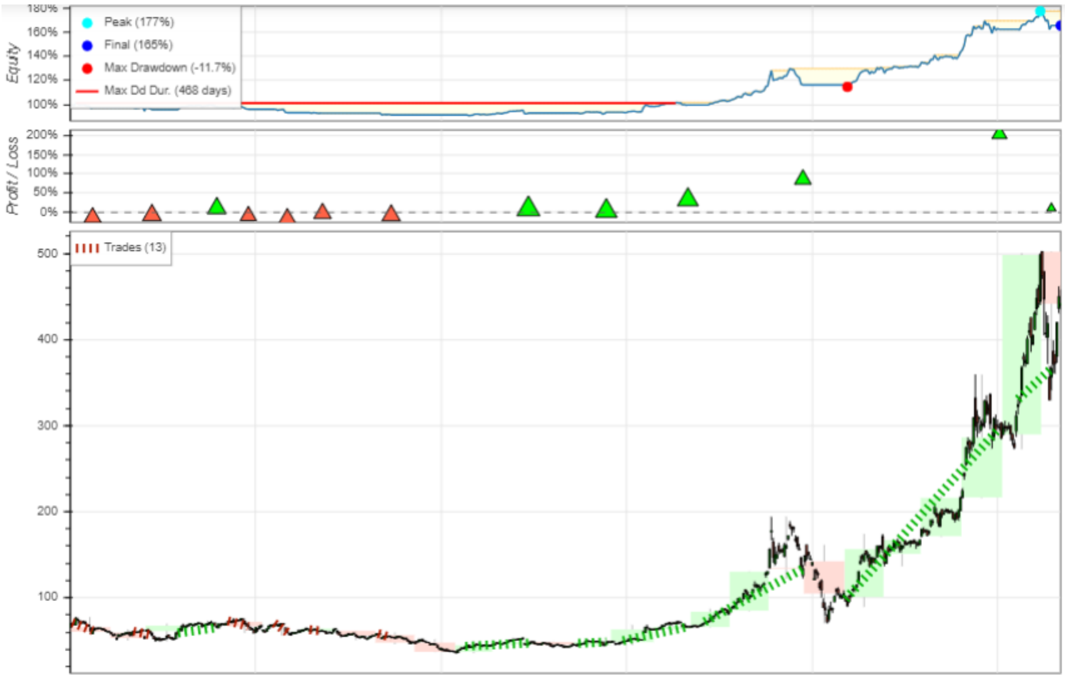
Technical indicators and trading methods based on them have been around for decades, they're simple to test and simple to execute, and can give an unbiased representation of information contained on the price.

As such, it's a logical idea to extend our information of overall direction and sentiment with the underlying price structure and try to search for better entry points, and perhaps, acceptable exit points.

Using the traditional moving averages method of a fast (10 candle) over a slow (20 candle) weighted moving average for firing the trades and using the Bollinger bands for stop loss instead of a fixed stop gives us less profit due to less big trades but much less drawdown and risk:



```
Start                     2018-08-02 00:00:00
End                       2020-09-16 00:00:00
Duration                    776 days 00:00:00
Exposure Time [%]                   60.747664
Equity Final [$]                  16546.17073
Equity Peak [$]                  17734.530806
Return [%]                          65.461707
Buy & Hold Return [%]              531.916273
Return (Ann.) [%]                    26.76888
Volatility (Ann.) [%]               20.088618
Sharpe Ratio                          1.33254
Sortino Ratio                        2.742277
Calmar Ratio                         2.293919
Max. Drawdown [%]                  -11.669498
Avg. Drawdown [%]                   -2.905502
Max. Drawdown Duration      469 days 00:00:00
Avg. Drawdown Duration       34 days 00:00:00
# Trades                                   13
Win Rate [%]                        53.846154
Best Trade [%]                     202.833191
Worst Trade [%]                     -15.04803
Avg. Trade [%]                      14.402866
Max. Trade Duration         120 days 00:00:00
Avg. Trade Duration          35 days 00:00:00
Profit Factor                        6.451449
Expectancy [%]                      23.074175
SQN                                  1.364689
```

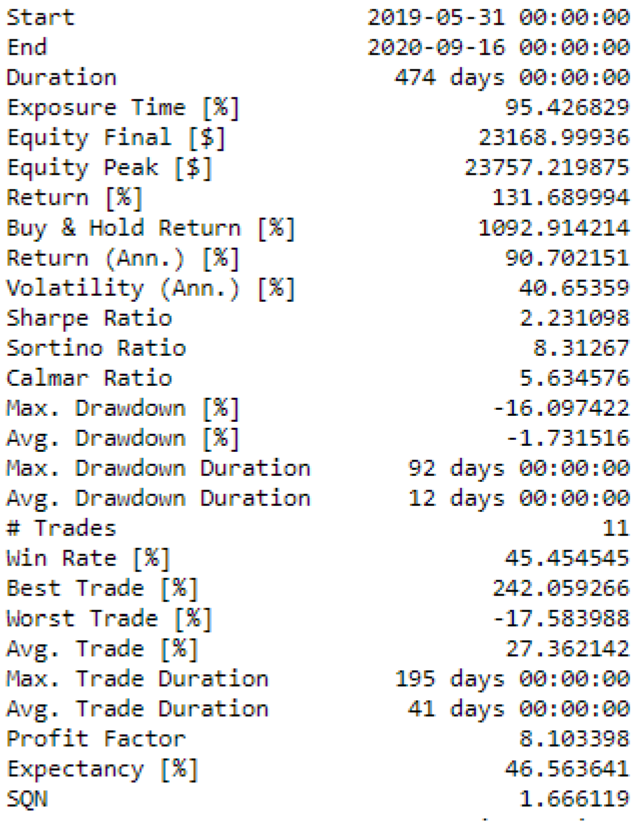


While there are many technical analysis methods we could combine with our sentiment to try and open trades only at optimal moments this process is very slow and prone to mistakes or overfitting. It's an iterative process that can take a considerable time to reach a decent result.

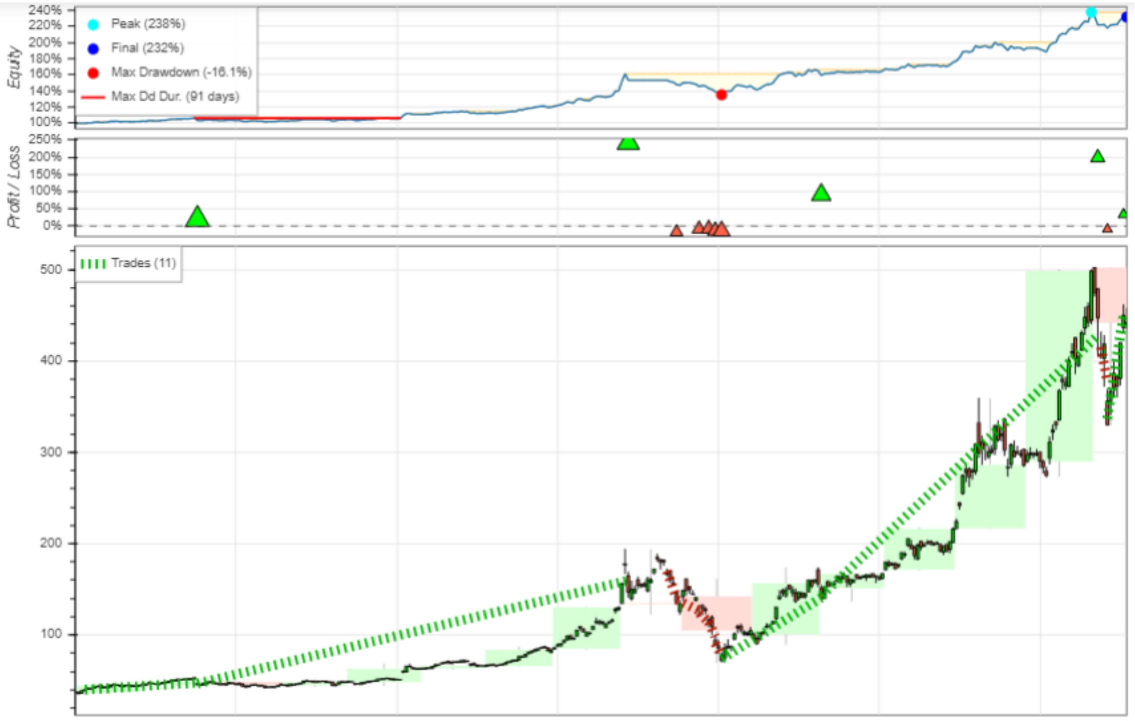
Ninth iteration: Using a classifier for automatic TA analysis

Using classifiers for these kind of operations is best done with plenty of data (up in the hundreds of thousands), but we can do a quick test with a very solid overall algorithm called random forest and set a big number of trees for automatic feature selection and classification. The approach is done using the next candle data and the next week's candle data, The first step is separating the data into train and test, the random forest has no memory or time lag so we just feed it the whole first 200 rows of data with the TA indicators and the price data. We then use it later with our sentiment and define the following rule:

“If our sentiment is positive and we predict a next day positive candle and a next week positive candle” then we open a long position, and the opposite for a short one. We add stop loss and a trailing stop loss and the resulting test is shown below:



```
Start                     2019-05-31 00:00:00
End                       2020-09-16 00:00:00
Duration                    474 days 00:00:00
Exposure Time [%]                   95.426829
Equity Final [$]                  23168.99936
Equity Peak [$]                  23757.219875
Return [%]                         131.689994
Buy & Hold Return [%]             1092.914214
Return (Ann.) [%]                   90.702151
Volatility (Ann.) [%]                40.65359
Sharpe Ratio                         2.231098
Sortino Ratio                         8.31267
Calmar Ratio                         5.634576
Max. Drawdown [%]                  -16.097422
Avg. Drawdown [%]                   -1.731516
Max. Drawdown Duration       92 days 00:00:00
Avg. Drawdown Duration       12 days 00:00:00
# Trades                                   11
Win Rate [%]                        45.454545
Best Trade [%]                     242.059266
Worst Trade [%]                    -17.583988
Avg. Trade [%]                      27.362142
Max. Trade Duration         195 days 00:00:00
Avg. Trade Duration          41 days 00:00:00
Profit Factor                        8.103398
Expectancy [%]                      46.563641
SQN                                  1.666119
```



The results are very amazing with a sharpe ratio of 2.23, although we could surely do better if we could train the classifier on much more data first and extend our test. We barely have a year of half of test data. Other classifiers could also help mitigate the errors we have on the price falls, so we could build an ensemble of those.

Tenth iteration: Using a RNN with LSTM for price prediction.