A Cryptocurrency Evaluation & Trading System

Final Presentation

Ibrahim Alshibani

University of Freiburg
Department of Computer Science
Chair of Algorithms and Data Structures

Examiners: Prof. Dr. Hannah Bast
Prof. Dr. Dirk Neumann

August 6, 2019

Contents

Contents

Problem

- Forecast the future price of an asset in a financial market.
- The focus is on a new type of asset: cryptocurrencies.
- Two types of approaches were applied to solve this problem:
  1. Utilizing trading strategies based on Technical Analysis.
  2. Learning-based approach using a state of the art sequence prediction model.

Problem (2)

- Complementary to the first approach, a trading system was developed that can:

1. Run trade simulations on historic price data.
2. Launch a trading bot that trades automatically on behalf of the user.

What is a Cryptocurrency?

- Cryptocurrencies are a form of digital currencies.
- Fairness and Legitimacy of transactions is enforced via Cryptography.
- This nullifies the need for a central authority to process and verify the transactions.

What is a Moving Average

Formally:

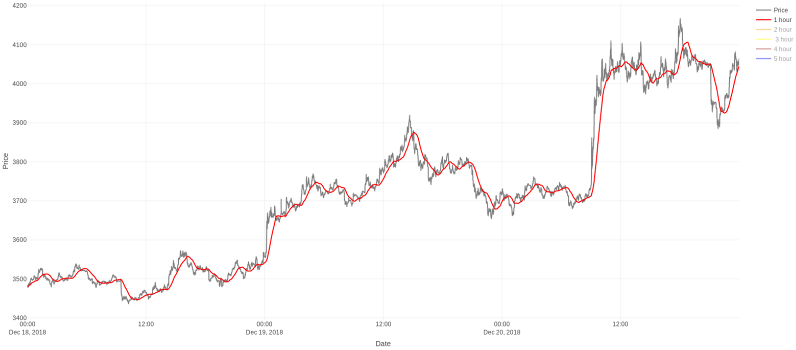
- The moving average at time step t for window size N:
- $MA_t(N) = \frac{1}{N} \sum_{i=0}^{N-1} P_{t-i}$.
- where P_t is the price of an asset at time step t.

Intuitively:

- Smoothing technique.
- Trend Indicator.

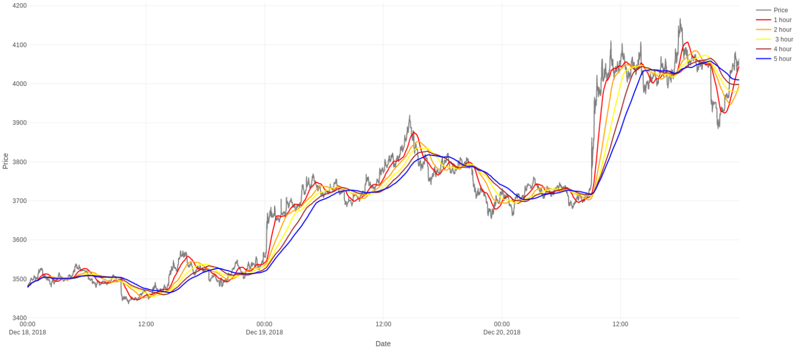
Visualizing MAs

Smoothing effect of 1 hour moving average on bitcoin price



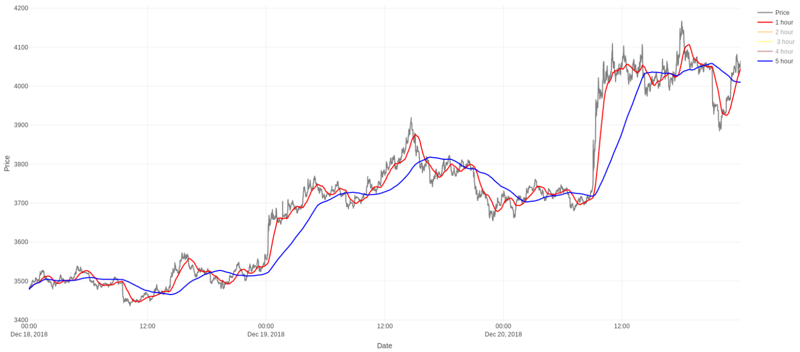
Bitcoin price data from 18th to Dec 20th 2018

Visualizing MAs (2)



1 to 5 Hour Moving Averages

Visualizing MAs (3)



Short and Long Moving Averages

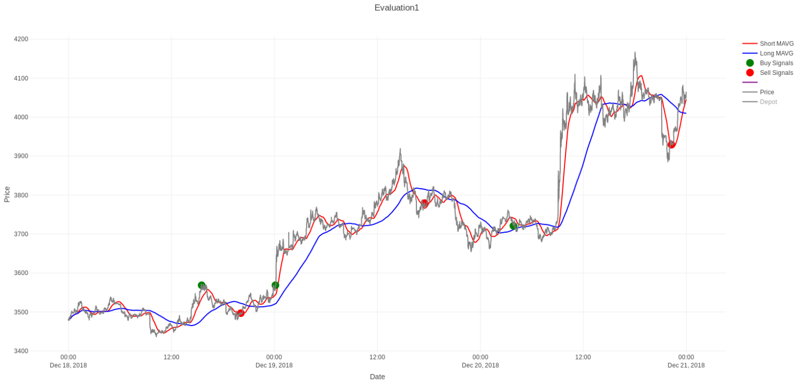
Cross Over Moving Average (Trend Following)

MA(short_window=60, long_window=300)

($100 - 21 Trades - $106.5)

Cross Over Moving Average (Trend Following)

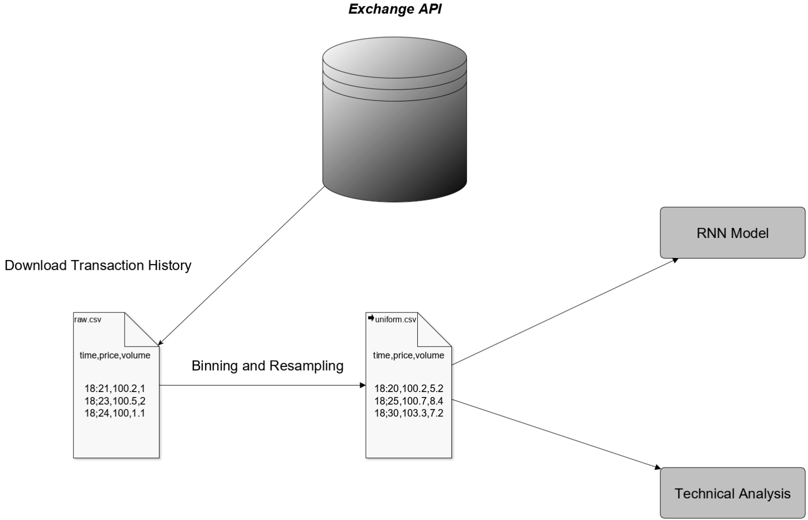
MA(short_window=60, long_window=300, filter=0.5%, holding_period = 60)



($100 - 6 Trades - $110.2)

Contents

Origin & Preprocessing



How to pick the right parameters?

- The performance of a strategy is dependent on parameters.
- Grid search was used to find optimal trading strategies.
- Return on Investment (ROI) = $\frac{\textit{Total Asset Value} - \textit{Starting Capital}}{\textit{Starting Capital}}$.
- The evaluation period is from March 2018 to March 2019.

Supported Currencies and Strategies

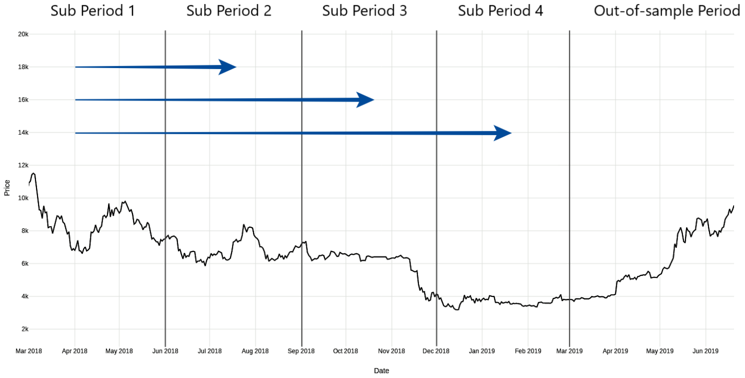
Cryptocurrencies:

- Bitcoin.
- Ether.
- Litecoin.
- ZCash.
- Dash.

TA Trading Strategies:

- Crossover Moving Average (Both forms).
- Exponentially Weighted Crossover Moving Average (Both forms).
- Bollinger Bands (Both forms).
- Relative Strength Index.

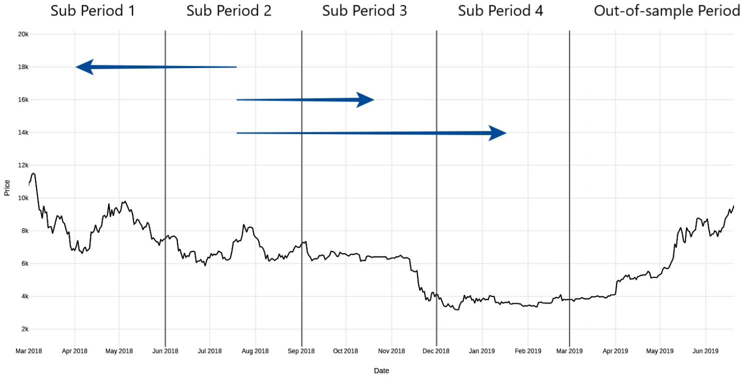
Cross-validated Grid Search



1- Run grid search on period 1.

2- Validate top 5 strategies against other periods.

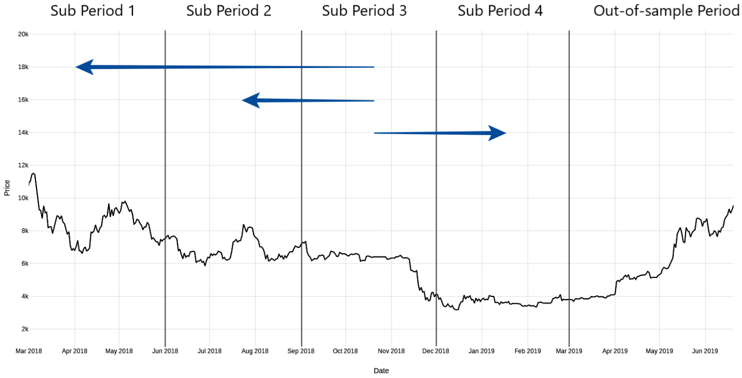
Cross-validated Grid Search



1- Run grid search on period 2.

2- Validate top 5 strategies against other periods.

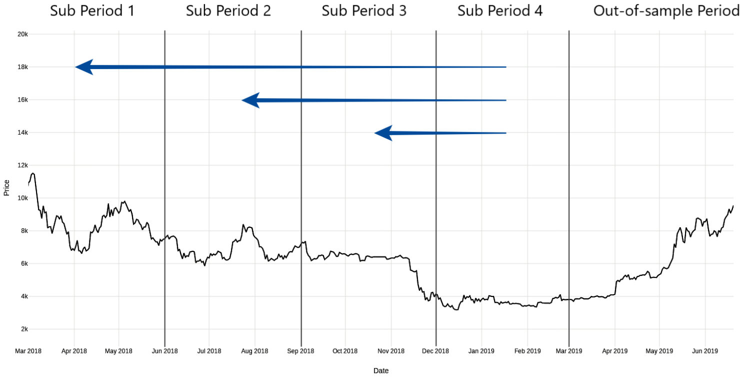
Cross-validated Grid Search



1- Run grid search on period 3.

2- Validate top 5 strategies against other periods.

Cross-validated Grid Search



1- Run grid search on period 4.

2- Validate top 5 strategies against other periods.

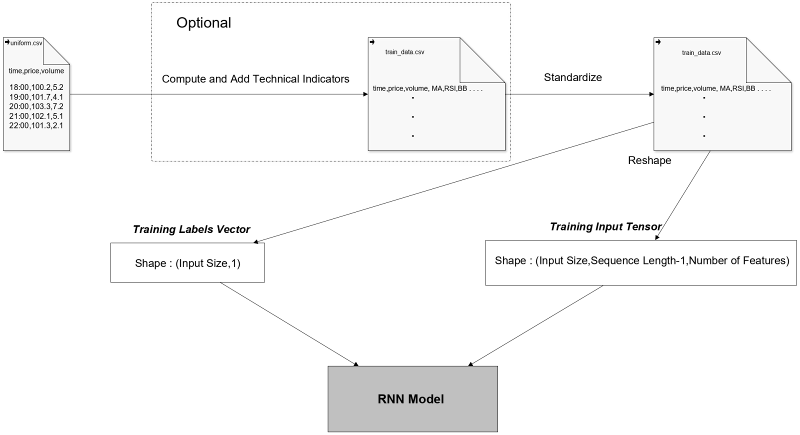
Candidate Strategies

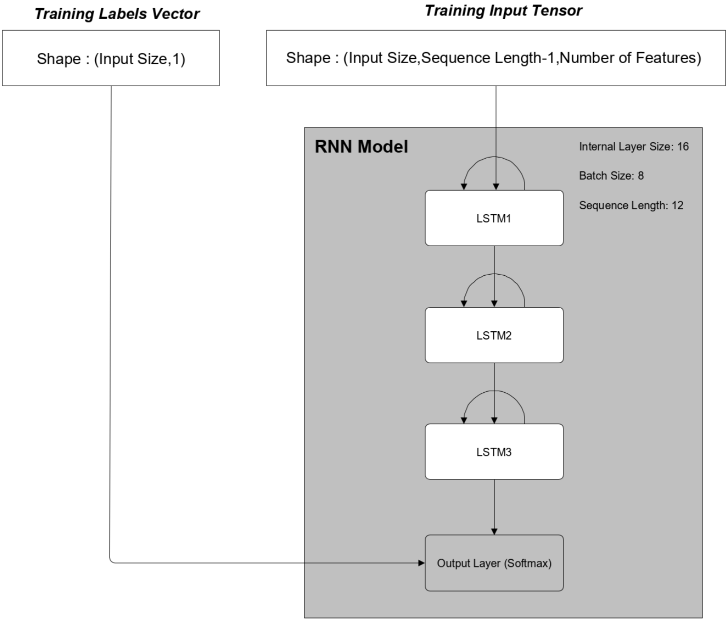
- After performing the cross validation, strategies are chosen based on a pre-defined criteria:
  - Profitability across all folds.
  - Highest return on investment.

Overview

RNN Model

- Recurrent Neural Networks, a variation of Neural Networks.
- Structure modified to process temporal data.
- Problem formulated as a binary classification problem:
  - Class 0: Price stays the same or goes down.
  - Class 1: Price goes up.
- Long Short-term Memory architecture (1) used in the network.
- Entire transactional history of currencies fed in RNN as input.
- Data was resampled into 1 hour intervals.


uniform.csv
time,price,volume
18:00,100.2,5.2
19:00,101.7,4.1
20:00,103.3,7.2
21:00,102.1,5.1
22:00,101.3,2.1
Optional
Compute and Add Technical Indicators
train_data.csv
time,price,volume, MA,RSI,BB . . . .
Standardize
train_data.csv
time,price,volume, MA,RSI,BB . . . .
Reshape
Training Labels Vector
Shape : (Input Size,1)
Training Input Tensor
Shape : (Input Size,Sequence Length-1,Number of Features)
RNN Model


Training Labels Vector
Shape : (Input Size,1)
Training Input Tensor
Shape : (Input Size,Sequence Length-1,Number of Features)
RNN Model
Internal Layer Size: 16
Batch Size: 8
Sequence Length: 12
LSTM1
LSTM2
LSTM3
Output Layer (Softmax)

Contents

Results (1)

- Grid search approach filtered out 8 trading strategies.
- 2 were found for Bitcoin, 2 for Ether, 3 for Zcash and 1 for Litecoin.
- Trading bot was launched in the beginning of out-of-sample period with candidate strategy to trade bitcoin.

Results (2)

Bitcoin	Parameters	Average ROI
MA(Trend Following)	**(24h, 120h, 5%, 4.5h)**	**7.8%**
BB(Trend Following)	(2h, 0.5, 5%, 0h)	7.1%
Ether	Parameters	Average ROI
MA(Trend Following)	(12h, 72h, 1%, 1.5h)	16.3%
MA(Trend Following)	(12h, 72h, 1%, 1h)	16.5%
Litecoin	Parameters	Average ROI
BB(Trend Following)	(120h, 1, 1%, 3h)	14.7%
Zcash	Parameters	Average ROI
BB(Trend Reversing)	(672h, 1, 0.05%, 24h)	29.3%
BB(Trend Reversing)	(336h, 2, 1%, 12h)	35.9%
BB(Trend Reversing)	(336h, 1, 5%, 12h)	24.1%

Candidate Strategies Filtered Out by Cross Validation.

Results (3)

Bitcoin	Parameters	Out-of-sample ROI
MA(Trend Following)	**(24h, 120h, 5%, 4.5h)**	**51.34%**
BB(Trend Following)	(2h, 0.5, 5%, 0h)	8.81%

Ether	Parameters	Out-of-sample ROI
MA(Trend Following)	(12h, 72h, 1%, 1.5h)	30.36%
MA(Trend Following)	(12h, 72h, 1%, 1h)	29.69%

Litecoin	Parameters	Out-of-sample ROI
BB(Trend Following)	(120h, 1, 1%, 3h)	78%

Zcash	Parameters	Out-of-sample ROI
BB(Trend Reversing)	(672h, 1, 0.05%, 24h)	27.2%
BB(Trend Reversing)	(336h, 2, 1%, 12h)	6.58%
BB(Trend Reversing)	(336h, 1, 5%, 12h)	16.68%

Out-of-Sample Performance of Candidate Strategies.

Trading Bot Performance

Performance of Trading Bot. ROI: 47.7%.

MA(Trend Following)(24h,120h, 5%, 4.5h)

Overview

Models

- A separate model was trained for each currency.
- Split into Train:Test (80:20).
- Train further split into Train:Validation(90:10).

Data Set Sizes

Currency	Starting Date	Training Set	Validation Set	Test Set
Bitcoin	2014-01-07	34034	3782	9454
Ethereum	2015-08-07	24023	2670	6674
Litecoin	2013-10-24	31573	3509	8771
Dash	2017-04-12	13455	1495	3738
Zcash	2016-10-29	16317	1814	4533

Number of Rows of Training, Validation and Test sets.

Evaluation Metrics

$$Recall = \frac{TP}{TP + FN} = \frac{Number\ of\ items\ of\ class\ identified}{Total\ number\ of\ class\ members\ in\ test\ set}$$

$$Precision = \frac{TP}{TP + FP} = \frac{Number\ of\ items\ of\ class\ identified}{Total\ items\ assigned\ to\ class}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Results

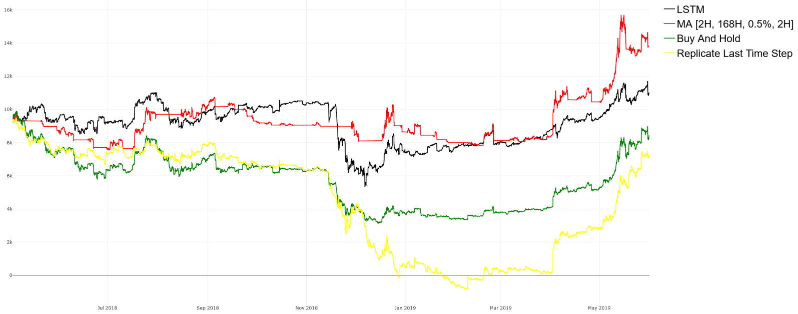
Test Set	Balanced Accuracy	Avg. F1
Bitcoin	54.9%	54.8%
Ether	**55.2%**	**55.1%**
Litecoin	53.6%	53.6%
Dash	50.8%	43.7%
Zcash	50.4%	46.0%

Performance of LSTM model trained with only historical prices.

Trading Simulations

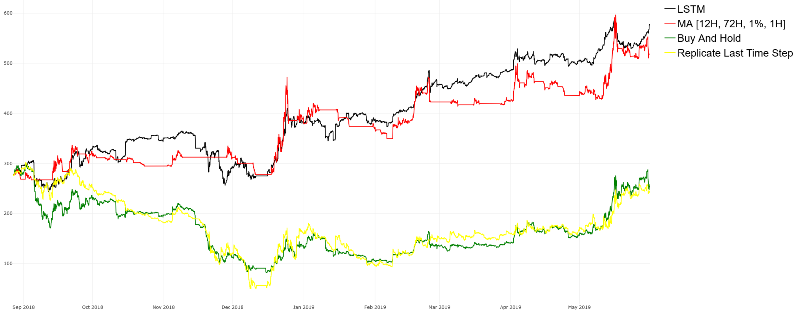
- Simulation Rules: when the model predicts up, buy one unit of asset and vice versa.
- Compared to multiple baseline strategies:
  1. Buy and Hold.
  2. Replicate Last.
  3. MA strategy optimized by cross validation.

Bitcoin Simulation



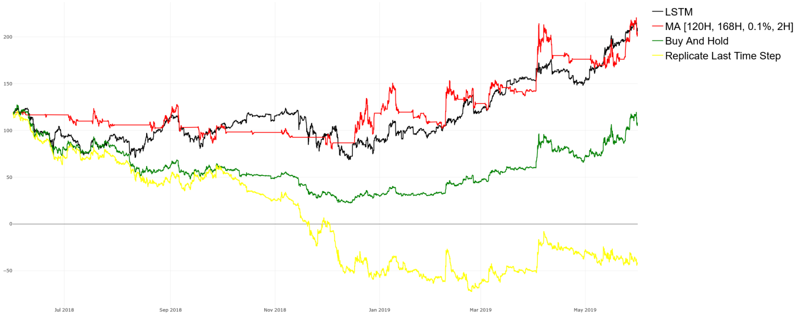
Equity Lines of Trading Simulations on Bitcoin Data. **Avg. F1 = 54.8%**

Ether Simulation



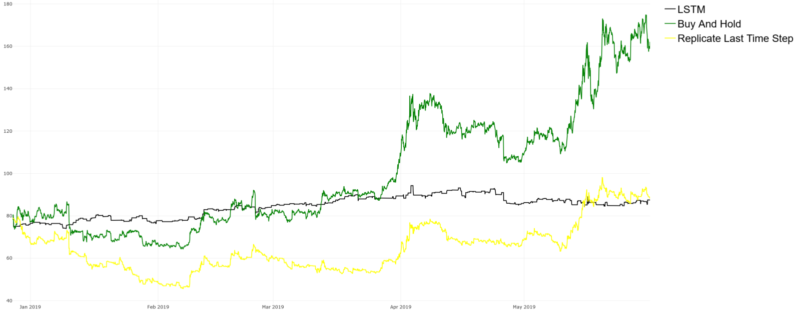
Equity Lines of Trading Simulations on Ether Data. **Avg. F1 = 55.1%**

Litecoin Simulation



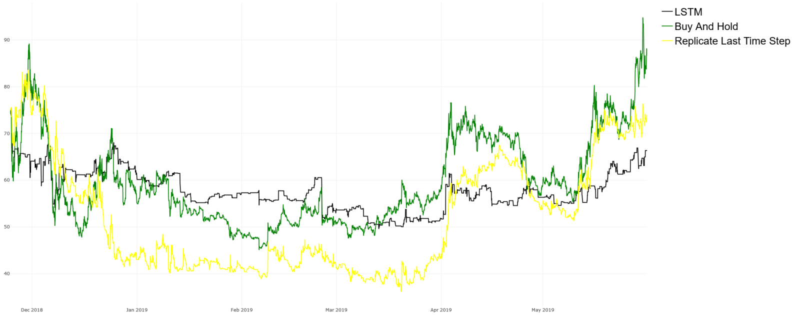
Equity Lines of Trading Simulations on Litecoin Data. **Avg. F1 = 53.6%**

Dash Simulation



Equity Lines of Trading Simulations on Dash Data. **Avg. F1 = 43.7%**

Zcash Simulation



Equity Lines of Trading Simulations on Zcash Data. **Avg. F1 = 46.0%**

Concluding Remarks

- Accurately predicting price in a financial market is a difficult task.
- Limited accuracy improvement in comparison to a random classifier.
- This limited improvement over a long term has resulted in profitable trade simulations.
- Strategies based on technical analysis are simpler than the RNN model.
- TA Strategies perform comparatively to more sophisticated RNN model with fewer signals generated.

Future Work

1. Investigating the effect of financial news on cryptocurrency prices. There has been already some work in this direction for traditional stock markets by Shumaker and Chen (2). They achieved 57.1% directional accuracy using this approach.
2. Experiment further with different RNN architectures and to further tune the parameters of the model.
3. Enabling the trading bot to be able to connect to multiple cryptocurrency exchanges.
4. Enhance the bot by implementing live trading based on the signals generated by the LSTM model.

Thank you for your time and attention!

References

- Hochreiter, Sepp and Schmidhuber, Jürgen
  Long short-term memory.
  Neural computation, 1997.

- Schumaker, Robert P and Chen, Hsinchun
  Textual analysis of stock market prediction using breaking financial news: The AZFin text system.
  ACM Transactions on Information Systems (TOIS), 2009.

Bollinger Bands

What is it?

- Calculate and upper and lower bound for the price of an asset at a given time step.
- Based on volatility of the asset.
- Assumes that the price should always be within these bounds.

Input

BB(window_size, width)

Formally

Standard Deviation

- Standard deviation at a given time step is :
- $S_t(N) = \sqrt{\frac{1}{N-1} \sum_{i=0}^{N-1} (P_{t-i} - MA_t)^2}$
- where P_t is the price at time step t.

Upper and Lower Bounds

- Upper Bound is : $MA_t(N) + (width \times S_t)$
- Lower Bound is : $MA_t(N) - (width \times S_t)$

Visually

Bollinger Bands strategy on Bitcoin data from 3 Apr to 12 Apr 2019

Parameter Space MA Strategies

Parameter	Range (In hours for time parameters)
Short Window Size	{1, 2, 3, 6, 12, 24, 36, 48, 72, 96, 120, 144, 168}
Long Window Size	{12, 24, 48, 72, 120, 168, 240, 336, 672}
Percentage Filter	{0, 0.05, 0.1, 0.5, 1, 5}
Holding Period	{0, 0.25, 0.5, 1, 1, 2, 2, 3, 4, 6, 12, 24}

Total Number of Combinations : 8424

Parameter Space for BB Strategy (Both Variations)

Parameter	Range (In hours for time parameters)
Window Size	{1, 2, 3, 6, 12, 24, 36, 48, 72, 96, 120, 144, 168, 240, 336, 672}
Band Width	{0.5,1,1.5,2,3}
Percentage Filter	{0, 0.05, 0.1, 0.5, 1, 5}
Holding Period	{0, 0.25, 0.5, 1, 1, 2, 2, 3, 4, 6, 12, 24}

Total Number of Combinations : 5760

Parameter Space for RSI Strategy

Parameter	Range (In hours for time parameters)
Window Size	{1, 2, 3, 6, 12, 24, 36, 48, 72, 96, 120, 144, 168, 240, 336, 672}
Width Offset	{10, 15, 20, 25, 30, 35, 40, 45}
Percentage Filter	{0, 0.05, 0.1, 0.5, 1, 5}
Holding Period	{0, 0.25, 0.5, 1, 1, 2, 2, 3, 4, 6, 12, 24}

Total Number of Combinations : 9216

Total Number of Combinations for All Trading Strategies : 23400

Trade Simulation of Candidate Strategy Bitcoin (1)

MA(Trend Following)(24h, 120h, 5%, 4.5h)

Trade Simulation of Candidate Strategy Ether

MA(Trend Following)(12h, 72h, 1%, 1.5h)

Trade Simulation of Candidate Strategy Litecoin

BB(Trend Following)(120h, 1, 1%, 3h)

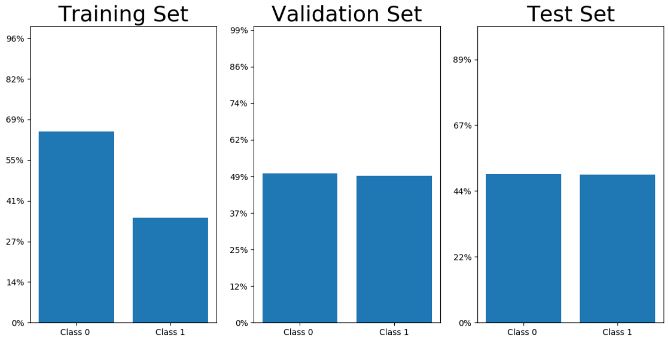
Trade Simulation of Candidate Strategy Zcash

BB(Trend Reversing)(672h, 1, 0.05%, 24h)

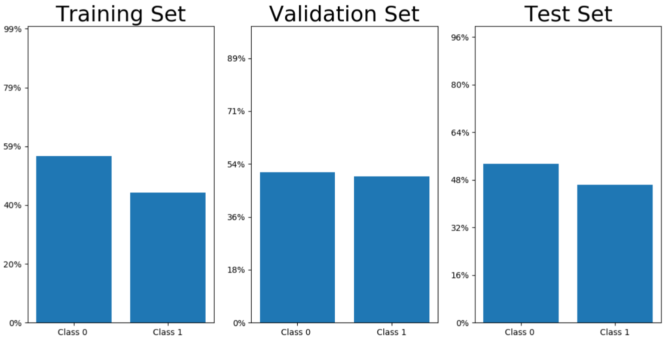
Order Initiation Times

Type	Trading Bot (Live)	Offline Simulation
Buy	Apr $2^{nd}18 : 20 CEST$	Apr $2^{nd}18 : 20 CEST$
Sell	May $18^{th}04 : 30 CEST$	May $18^{th}04 : 30 CEST$
Buy	May $27^{th}15 : 55 CEST$	May $27^{th}15 : 55 CEST$

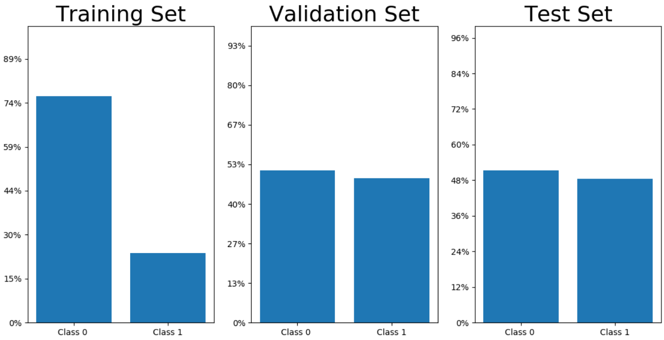
Bitcoin Class Distributions



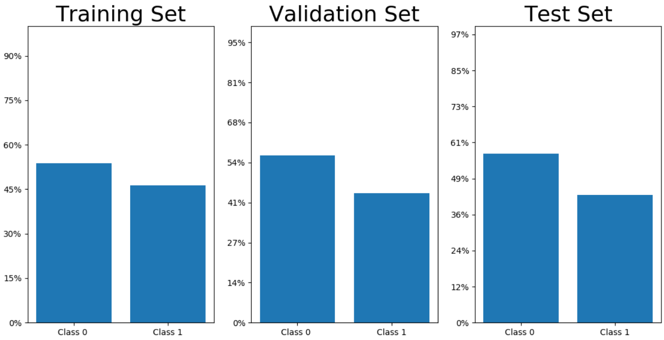
Ether Class Distributions



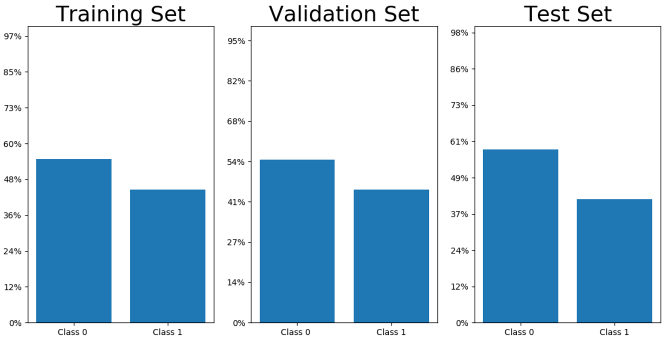
Litecoin Class Distributions



Dash Class Distributions



Zcash Class Distributions



RNN Model with TA

Test Set	Balanced Accuracy	Avg. F1
Bitcoin	54.3%	54.3%
Ether	53.4%	53.2%
Litecoin	53.0%	52.1%
Dash	50.8%	50.6%
Zcash	50.7%	48.0%

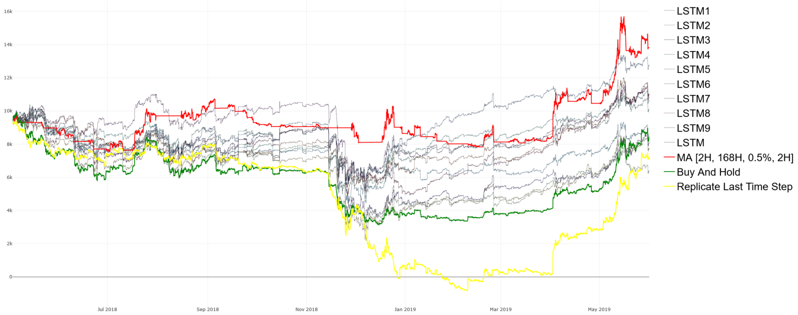
Performance of LSTM model trained with historical prices + technical indicators

RNN Model Positive Class

Test Set	Recall	Precision	F1
Bitcoin	57.9%	54.9%	56.4%
Ether	48.9%	54.3%	51.5%
Litecoin	55.9%	52.3%	54.1%
Dash	12.4%	45.5%	19%
Zcash	17.1%	45.6%	24.9%

Performance of LSTM model trained with only historical prices.

Multiple Bitcoin Models



Number of Transactions

Currency	LSTM	Moving Average	Buy and Hold	Replicate Last
Bitcoin	3778	70	1	5157
Ether	2766	66	1	3559
Litecoin	2502	55	1	4740
Dash	528	-	1	1905
Zcash	860	-	1	2313

Number of signals generated for each simulation