Transition from Observation To Knowledge To Intelligence (TOKI)

Editors
Dr. Victor ODUMUYIWA, Dr. Olufade ONIFADE,
Prof. Amos DAVID & Prof. Charles UWADIA

Victor ODUMUYIWA
Department of Computer Sciences,
University of Lagos
Nigeria

ISBN: 978-978-976-000-8

Transition from Observation to Knowledge to Intelligence

3rd Biennial International Conference on Transition from Observation to Knowledge to Intelligence (TOKI)
15-16 August 2019
University of Lagos, Nigeria

Editors

Dr. Victor ODUMUYIWA
Dr. Olufade ONIFADE
Prof. Amos DAVID
Prof. Charles UWADIA

Cryptocurrency Price Prediction Using Long Short Term Memory Modeling and Social Media Sentiment

O.P Akomolafe

Department of Computer Science
University of Ibadan, Ibadan. Nigeria.

Oluwafikayo Sanni

Department of Computer Science
University of Ibadan, Ibadan. Nigeria.

Abstract. Predicting cryptocurrency prices and market directions is not a new problem, however existing solutions in literature generally have wide margins of errors in prediction caused by conservative loss functions and auto-regressive models, and have a general inability to scale to predict prices and market directions for multiple tokens while keeping performance relatively stable. This work aims to provide an optimized solution that provides low margins of errors in predictions and scalable performance for multiple cryptocurrencies. To achieve this aim, we designed our prediction model around a long short term memory model which we trained and tested with historical price data which we enriched with social media sentiment data. Our Hyperparameters (linear activation function, mean square error loss function, adam optimizer and dropout of 0.25) ensure that our model provides a tight fit to the problem without overfitting our training data. We also penalized conservative models by excluding models with Mean Square Errors less than 0.05 when finally approximating our target function. As a result of our efforts, our model was trained to predict three cryptocurrencies, bitcoin, ripple and ethereum. . It predicted bitcoin prices to within 1%, 2%, 5%, 10%, 20%, 30% accuracy 15.9%, 33.8%, 65.9% , 87.8%, 98.1% and 99.7% of the time respectively. Ethereum prices to within 1%, 2%, 5%, 10%, 20%, 30% accuracy 9.7%, 20.6%, 54.1%, 85.6%, 97.2% and 99.7% of the time.1%, 2%, 5%, 10%, 20%, 30% accuracy 18.4%, 33.6%, 69.7%, 88.8%, 99.06% and 100% of the time respectively. These numbers showed relatively uniform performances by our model in predicting multiple cryptocurrencies. On minimizing the margins of error, we were able to achieve 3.7% and 6.0% margins of error predicting bitcoin price increases and decreases respectively, for ripple it was 5.9% and 9.1% while for ethereum the margins of error was 4.44% and 4.78% respectively. These numbers showed a comparative improvement on existing models in literature

Keywords- cryptocurrency, social media sentiment, long short term memory modeling, bitcoin, ripple, ethereum

1. Introduction

Cryptocurrencies are digital assets that are designed to work as media of exchange that use cryptography to secure transactions, control the creation of additional units, and verify the transfer of assets (Corbet et al 2018). Bitcoin was the first digital cryptocurrency, published in 2009. Since then, more than 2,000 cryptocurrencies have emerged; such as Ripple, Ethereum etc. These currencies have been used for various use cases, ranging from In-app purchases, game rewards, some even serve like traditional currencies (Corbet et al 2018).

Cryptocurrencies have become a social and economic phenomenon, and the cryptocurrency market grew to be worth as large as $500bn in may 2018. Predicting market prices and price fluctuations is not new. The connection between social media sentiment and market directions has been well established (Colianni et al., 2015, Stenqvist an Lonno 2017), also there have been multiple research work done on creating models that predict prices and price fluctuations (Lamon etal 2016, Hegazy and Mumord 2016), and comparative analyses of how various regression models fit the prediction problems.

This work fixes a gap in knowledge in existing literature by providing a framework model for predicting cryptocurrency prices and market directions accurately and uniformly across multiple cryptocurrencies. Existing limitations observed that made existing models unable to fit a generic target function for accurately making predictions are also addressed, including the use of conservative loss functions and hyperparameters and discouraging the use of conservative Auto Regression Models in price prediction. By performing both of the above, we created a model that aggressively detects spikes and troughs in prices of any cryptocurrency, with proper training.

2. Existing Work

In previous research works, (Hegazy and Mumord 2016) were able to predict bitcoin prices to within 1% margin of error 57% of the time using historical prices and supervised learning techniques. (Jiang and Liang 2016) took a different approach and used deep reinforcement learning algorithms to make trading decisions on a bitcoin portfolio that resulted in 10 times increase in the portfolio's value.

Existing work are mostly bitcoin centric, with the handful non-bitcoin centric literature performing comparative analysis of how well different approaches fit the predictions of different cryptocurrencies. (Lamon et al., 2017) performed one of these comparative analysis and tried to predict the prices of Bitcoin, Litecoin and Ethereum using news data and social media data labelled with actual price fluctuations they caused, by size, rather than actual sentiment expressed in the tweets. They had some interesting results.

They found out that while Logistic regression was the best for predicting bitcoin prices, it was only able to predict 43.9% of price increases correctly, and 61.9% of price decreases correctly. Logistic Regression was the best at predicting the prices of Litecoin, but only when predicting decreases, not increases. Bernoulli Naive Bayes's Classifier was very effective for predicting price increases with 75.8% prices on increases predicted correctly whilst it could only predict 16.1% of price decreases correctly.

What should be noteworthy about all these, is that the variance in prediction performance varies from token to token and from model to model, which raises the need for a generic model with uniform performance that can scale to predict multiple tokens with proper training. And market directions were used to label sentiment, not the actual sentiment expressed in the tweets.

This means since we have an inability to know how large tomorrow's price increase would be beforehand, these models are useless for real life applications. Another major fault was the lack of evidence of an attempt to curb model conservatism by improving the models ability to aggressively find increments and decrements in prices, these led to the margins of error while predicting cryptocurrency prices during increments and decrements to be quite large, which meant that the existing models would perform very poorly during periods of very unstable prices

In other research work have used tweets and other social media data to predict prices. Laskowski, Kim & Lee (2016) were able to find a direct correlation between number of mentions on twitter and the prices of bitcoin and dogecoin. (Coliammi et al., 2015) were able to predict the market directions of a handful of cryptocurrencies including bitcoin using supervised learning algorithms for sentiment analysis on tweets.

3. Datasets

Data input into the model was sourced in two forms, the first was sentiment data collected from tweets about a cryptocurrency during the course of the day, the second was historical price data collected from coinmarketcap.com.

To get the first data set, we built a web scraping script with a python library called GetOldTweets, which allowed us to find tweets about specific tokens from specific days. This allowed us to generate a time series dataset of tweets for every 24 hours. Each of these tweets was cleaned to remove some commonly occurring regular expressions such as hashtags, mentions, links, pictures, etc.

Using this method, we were able to fetch 311,810 tweets over a period of 1280 days between June 15, 2015 and December 15, 2018. This time period provided 128,000 tweets for bitcoin, 77,084 tweets for ethereum and 106,726 tweets for ripple.

A total of 8 features are extracted from these data sets, these features include Date, Open, High, Low, Close, Volume, MarketCap, Daily Sentiment and Tweet Num. LSTMs don’t like data with huge variances, so each data point had to be normalized to within a variance of 0 and 1. We also had to reduce the dimensions of the problem. To achieve this, we replaced the high & low fields with volatility which can be expressed as

$$volatility = (high-low)/open$$

This was computed on a scale of -1 and 1. CloseOffHigh is used to represent the gap between the closing price and the high for the day on a scale of -1 to 1. Where -1 means the closing price was the lowest price of the day and 1 means the closing price was the highest of the day. This can be expressed mathematically as

$$closeOffHigh = 2(high-close)/(high-low)-1$$

4. Methodology

4.1. Sentiment Analysis

Sentiment Polarity was favoured over sentiment subjectivity. Sentiment polarity is a float value between -1 and 1 where -1 represents an absolutely negative sentiment and 1 represents an absolutely positive sentiment. Subjectivity on the other hand, refers to personal opinion. Its a measure on a scale of [0,1] on whether the opinion expressed in a text are subjective (0) or objective (1).

In computing sentiment polarity for each tweet, we would be using a prebuilt generic sentiment classifier built with the Textblob library of python. Each sentiment was collected and summed for every day in our 1,280 day time series. This sentiment summation was then averaged by the total number of tweets to give the average daily sentiment of the token for each of the 1,280 days.

4.2. Train-Test Split

After performing our sentiment analysis and consolidating our historical price data with daily average sentiment features; we performed a 60-40 train-test split on our time series. Since our time series was over a period of 1280 days, this translated into a 768-512 split. The training set covered the time period between 15 June, 2015 and 21 July, 2017. The test set covered the time period between 22 July, 2017 and 15 December, 2018.

4.3. Training Hyper-Parameters

The choice of hyperparameters in training our LSTM were informed by our requirement of aggressive detection of price increments and decrements. We used a linear activation function, a dropout of 0.25 and an adam optimizer, 40 neurons are used and our choice of loss function was the mean square error.

4.4. Training LSTM

Training the LSTM was carried over 25 randomizations, each randomization reduced the loss function over 50 epochs. We used random seeds between 775 & 800. Weights were initialized with the simple weight guessing algorithm for each randomization

4.5. Penalizing Conservative Models

To discourage conservatism, we automatically excluded all randomizations of the models whose loss function (MSE) could not be reduced below 0.05. This was an arbitrary number chosen by us.

4.6. Averaging the Target Function

After training the 25 random initializations of the LSTMS and weeding out the models with loss functions that fell above our arbitrary threshold. We averaged the loss functions obtained from all the initializations to get our target function, we used on our test set.

5. Experiments

5.1. Evaluation Parameters

In evaluating our model, we focused on accuracy metrics, such as the accuracy in predicting price increases, accuracy in predicting price decreases and various percentage margin of errors of prediction of prices. The Margins of Error evaluated are 1%, 2%, 5%, 10%, 20% and 30%. of actual prices.

In evaluating scalability, the primary indicator whether our model would be scalable to predict more cryptocurrencies; if trained, with similar performance would be a uniform performance by our model in predicting all three tokens we have trained it with. Large variations in performance would indicate non-scalability, while small variations would indicate scalability.

6. Results

5.2 Result Summary

Table 1: Result Summary

Token	1%	2%	5%	10%	20%	30%	INC MOE	DEC MOE	INC ACC	DEC ACC
BTC	15.9%	33.8%	65.9%	87.8%	98.1%	99.7%	3.7%	6.0%	26.6%	26.3%
ETH	9.7%	20.6%	54.1%	85.6%	97.1%	99.7%	4.4%	4.8%	18.4%	27.2%
XRP	18.4%	33.6%	69.7%	88.8%	99.06%	100%	5.9%	9.1%	18.1%	30.9%

5.3 Bitcoin Results

For Bitcoin, we Observed that our model achieved a 26.6% accuracy predicting price increases and 26.3% accuracy predicting price decreases. It predicted bitcoin prices to within 1%, 2%, 5%, 10%, 20%, 30% accuracy 15.9%, 33.8%, 65.9% , 87.8%, 98.1% and 99.7% of the time respectively. Overall, accuracy at predicting prices during price reductions at a 6.0% margin of error, while price rises was at a 3.7% margin of error.

5.4 Ethereum Results

For Ethereum, we observed that our model achieved an 18.4% accuracy predicting price increases and 27.2% accuracy predicting price decreases. It predicted ethereum prices to within 1%, 2%, 5%, 10%, 20%, 30% accuracy 9.7%, 20.6%, 54.1%, 85.6%, 97.2% and 99.7% of the time respectively. Overall, accuracy at predicting prices during price reductions at a 4.8% margin of error, while price rises was at a 4.44% margin of error.

5.5 Ripple Results

For Ripple, we observed that our model achieved an 18.1% accuracy predicting price increases and 30.9% accuracy predicting price decreases. It predicted ripple prices to within 1%, 2%, 5%, 10%, 20%, 30% accuracy 18.4%, 33.6%, 69.7%, 88.8%, 99.06% and 100% of the time respectively. Overall, accuracy at predicting prices during price reductions at a 9.1% margin of error, while price rises was at a 5.9% margin of error

6. Existing Models

The table below shows optimal results in previous works as collated in the work of Lamon et al (2016) which are shown in the table below.

TABLE 2: Best Results in Literature

Token	Increment Accuracy	Decrement Accuracy	Increment Margin of Error	Decrement Margin of Error	Model
Bitcoin (BTC)	43.19%	61.9%	42.24%	65.24%	Logistic Regression
Ethereum (ETH)	75.8%	16.1%	48.23%	9.74%	Bernoulli Naive Baye's

Simple Comparative Analysis of the above statistics with our model's results from the previous chapter show the following graphical results.

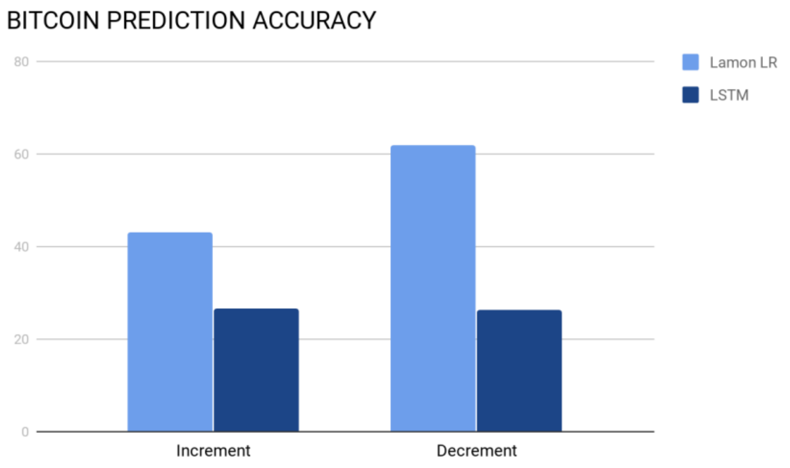

FIGURE 1: Bitcoin Prediction Accuracy Comparison

In predicting absolute price decrements and increments our model performs extremely poorly when compared to Logistic regression performed by Lamon et al (2016). Our model has a 26.56% accuracy predicting the prices of bitcoin while Lamon et al (2016) achieved a 43.19% accuracy predicting increases. Our model could predict 26.25% decreases accurately, while logistic regression achieved a 61.9% accuracy predicting decreases.

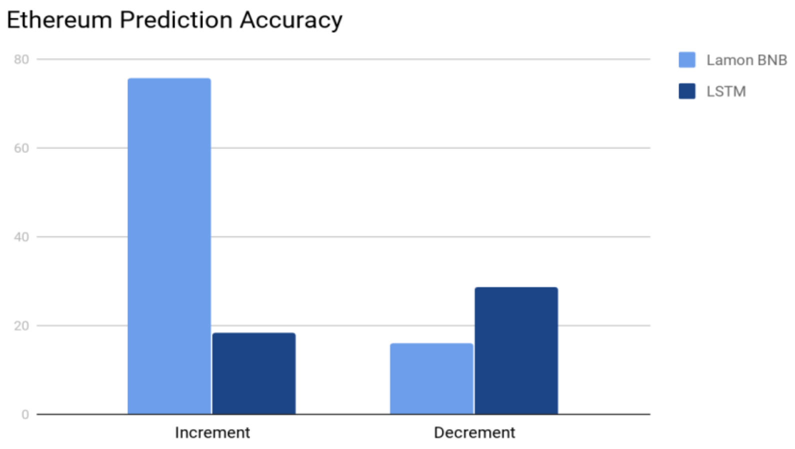

Figure 2: Ethereum Prediction Accuracy Comparison

In predicting Ethereum price fluctuations. Lamon et al (2016) achieved a 75.8% and 16,1% accuracy predicting increments and decrements respectively, while our model had a 18.44% and 28.75% accuracy predicting increments and decrements respectively.

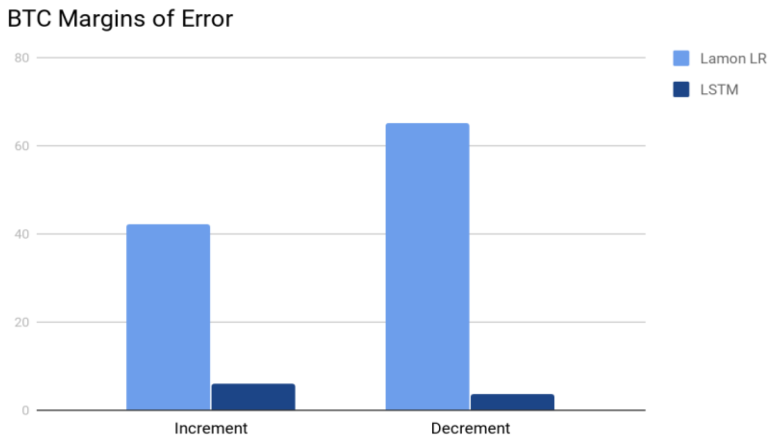

Figure 3: Bitcoin Prediction Margin of Error Comparison

When it comes to Margins of Error, Our Model showed comparable improvement over existing model. In predicting bitcoin price fluctuations; when predicting the size of an accurately predicted price increase or decrease, our model has margins of errors of 3.72% and 6.04% respectively, while Lamon et al (2016) had 42.24% and 65.24% respectively.

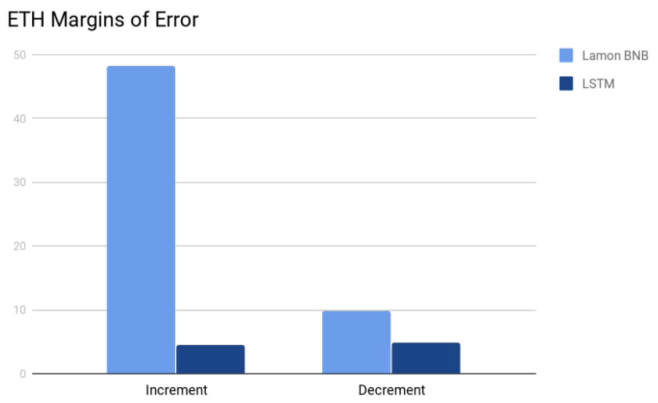

Figure 4: Ethereum Prediction Margin of Error Comparison

In predicting Ethereum price fluctuations; when predicting the size of an accurately predicted price increase or decrease, our model has margins of errors of 4.44% and 4.78% respectively, while (Lamon et al., 2016) had 48.23% and 9.74% respectively.

7. Limitations & Discussions

The efficacy of the TextBlob sentiment classifier in this study is contentious. A quick statistical analysis of the average daily sentiments returned by the classifier showed a variance of 0.003064886449 and a range of 0.5717818182 for ethereum, variance of 0.001043348426 and range of 0.561462284 for bitcoin, variance of 0.001749528775 and range of 0.367453823 for ripple which is very minute when considering TextBlob’s sentiment range of [-1,1].

This means that, despite the improvements recorded over existing models in literature, this model could be better served by having its own custom built classifier for predicting cryptocurrency related tweets.

A control experiment with a similar model, without the social media sentiment added showed a sharp increase in the Margins of Error of the Model. With margins of predicting bitcoin increasing to 5.6% for price increases and 6.8% price decreases respectively. Ripple recorded an increase to 6.4% for price increases and 9.8% for price decreases and Ethereum to 4.7% for increases and 5.2% for price decreases.

8. Conclusions

In predicting absolute number of price reductions and increments, our model performed very poorly when compared with existing models. Our Model however possess a comparative advantage of having a higher accuracy when it comes to predicting the margins of these fluctuations. Our model also has a relatively uniform performance for all three tokens. Meaning its behaviour could be scalable for predicting other tokens with proper training

9. Recommendations

We suggest that to further improve on the accuracy of predicting price increases and decreases, the following upgrades be made to this model

1) Training a sentiment classifier particularly built specifically for cryptocurrencies
2) Classifying tweets with the above classifier; instead of the textblob library
3) Reducing the MSE threshold for eliminating conservative models
4) Removing the hard cap of 100 tweets per day

List of References

Corbet, S., Lucey, B. M., Urquhart, A., & Yarovaya, L. 2018. Cryptocurrencies as a Financial Asset: A Systematic Analysis. SSRN. https://doi.org/10.2139/ssrn.3143122.

Evita Stenqvist, Jacob Lnn, 2017 Predicting Bitcoin price fluctuation with Twitter sentiment analysis.

Jiang, Z., & Liang, J. 2016. Cryptocurrency Portfolio Management with Deep Reinforcement Learning. ArXiv. Retrieved from http://arxiv.org/abs/1612.01277

Lamon, C., Nielsen, E., Redondo, E. 2015: Cryptocurrency Price Prediction using news and social media sentiment. Master's thesis, stanford.Master's thesis, Stanford http://cs229.stanford.edu/proj2017/final-reports/5237280.pdf

Kareem Hegazy and Samuel Mumford. 2016. *Comparitive Automated Bitcoin Trading Strategies.*CS229 Project, 2016 http://www.divaportal.org/smash/get/diva2:1110776/FULLTEXT01.pdf

Laskowski, M., Kim, H. M., & Ieee. 2016. Rapid Prototyping of a Text Mining Application for Cryptocurrency Market Intelligence. *Proceedings of 2016 Ieee 17th International Conference on Information Reuse and Integration (Ieee Iri)*, 448–453. **https://doi.org/10.1109/iri.2016.66**

Stuart Colianni, Stephanie Rosales, and Michael Signorotti. 2015. *Algorithmic Trading of Cryptocurrency Based on Twitter Sentiment Analysis*. CS229 Project, 2015.