1 Introduction

Volatility has always been an important problem in financial time series and modeling. Accurate forecasting of volatility is a challenging task, given the complexity of volatility data as well as highly nonlinear patterns and the stochastic nature of the process. Typically, in the finance industry, volatility is forecasted instead of price or returns, not only because it is deemed to be easier, but because risk management is crucial for the industry.

Volatility has typically been forecasted using GARCH models, with the drawbacks being the linearity assumption and lower accuracies. Deep learning models have been shown to lead to higher accuracies (Xiong et al., 2015; Liu, 2019) but as a drawback, they are not easily interpretable and are considered “black boxes”, where it is impossible to know how the models work internally and which features affect the predictions and how.

Cryptocurrencies are still relatively new and are often considered to be an object of speculation. Also, unlike most assets tied to companies and economies, cryptocurrencies do not generate quarterly reports, but rather act similar to currencies, with most of the data being available in real-time. Instead, cryptocurrencies have been sensitive to policy changes regarding them. Due to these reasons, they will generally have high volatility compared to other assets, and it is reasonable to assume that news affects the volatility rather strongly.

By knowing how different factors affect the volatilities of cryptocurrencies, it is easier to know how they may react to different events and market conditions. In this paper, we focus especially on the effect of news from a cryptocurrency-focused news website, by including the average daily sentiment as a feature. To our knowledge, academic research regarding partial dependence of the volatility of cryptocurrencies has not been yet done.

This paper will be structured as follows: we will first motivate why the topic is important, and introduce the coins we analyze. After explaining the methods, the data and the features, we will perform exploratory analysis. Lastly, we will discuss the models, the results and then conclude the paper.

2 Motivation

An investor or a hedge fund’s performance can be measured in many ways, but the most commonly used measure of performance is the Sharpe Ratio. The reason why the Share ratio has been implemented as a gold standard of investor performance is that returns in themselves do not take into account the risk the investor is exposed to, as measured by volatility.

$$Sharpe\ Ratio = \frac{Return - Risk\text{–}free\ rate}{Volatility} \tag{1}$$

The equation shows that an investor can increase the Sharpe Ratio by either increasing returns or reducing volatility since the risk-free rate cannot be changed. The past values are easy to calculate, while the future values are unknown and only forecasts can be made.

Investors in general and hedge funds in particular, want to adopt all the financial instruments available to reduce their portfolio volatility. If the volatility of the portfolio is low, the investor may use leverage to increase their returns without taking too much risk if the Sharpe Ratio is high enough. Previously, portfolios have consisted primarily of options, equities, and commodities, but it can be seen that especially in recent years, large hedge funds have also begun to include cryptocurrencies in their portfolios. According to Business Insider (2020), the SkyBridge hedge fund is one of the funds that have invested heavily in Bitcoin over the past year.

3 Cryptocurrencies

We have chosen to analyze five coins that have different characteristics and functionality and may react differently to the sentiment of the news regarding the coin in question. It is conceivable that the price of these coins depends on different factors, and thus the volatility will probably follow different patterns. The cryptocurrencies we analyze are Bitcoin, Ether, XRM, IOTA, and the Binance Coin

3.1 Bitcoin

Bitcoin is the first implementation of a blockchain and cryptocurrencies. The main essence behind the creation of Bitcoin is that the currency is decentralized, that is to say, that there is no central bank that controls the money supply nor a third party authorizing the transactions. Instead, the transactions are authorized using blockchain technology. Because Bitcoin is the first cryptocurrency, it is also the most recognized, which is often a kind of gateway to other currencies. It is also a cryptocurrency with the largest market capitalization.

3.2 Ether

Ethereum is a technology that allows users to send money to each other for a small fee. Ethereum is an open-source that allows anyone to create applications that are powered by it. These are mostly applications that facilitate payments between end-users. Ether is also the currency paid as a fee to those who maintain the Ethereum network.

3.4 XRP

XRP is not built on the same ideals as Bitcoin. Ripple Labs is a for-profit company that aims to serve banks and payment providers. The main essence is that Ripple should function as a faster, more transparent, safer, and cheaper alternative than traditional payments such as SWIFT. An important feature of Bitcoin is that the network based on blockchain technology is maintained by enthusiasts all over the world. Ripple, on the other hand, is not based on blockchain and is controlled by a private company. Ripple is AML compliant, which means that all abnormal activity on the network is reported to the relevant authorities. XRP is the cryptocurrency used on the Ripple network.

3.5 IOTA

IOTA is a distributed ledger designed to handle transactions between the devices in the Internet Of Things (IoT) network. The payment between the units is by the currency mIOTA. The main difference between IOTA and other cryptocurrencies is that IOTA is not based on blockchain technology, but on a technology called Tangle. Tangle is designed to be faster and more efficient than the classic blockchain but has also been shown to be more vulnerable to hacks, which often leads to increased volatility.

3.6 Binance coin

Binance is the world's largest exchange platform that specializes in buying and selling cryptocurrencies. To finance its operations, Binance requires commission per trade, and the commission paid using the Binance coin (BNB). Binance has thus issued its own currency whose sole purpose is to act as a trading fee. BNB is powered by the Ethereum network.

4 The data, target, and features

As data sources, we are using Coindesk for the textual data that the sentiment analysis was conducted on and CoinMarketCap for the daily price data. As a target for our model, we use the rolling average of the future 10-day volatility, calculated from the closing price of each day. In a similar way to the target, the features we consider are mostly related to the price of each cryptocurrency. We decided to use daily data to keep the amount of data manageable while having a strong enough signal. We perform some preprocessing for each of the variables, usually to make the model more accurate and to fit better. We do not use a split into training and test sets, since they are not commonly used when similar analysis as ours is performed, but we still have a validation set that includes the last 20% of the data for each of the epochs of the model fits.

4.1 Future 10-day volatility as a target

Volatility forecasts are usually computed for a rather short period, typically only 2 weeks ahead. Therefore we chose to forecast the average 10-day future realized volatility of different cryptocurrencies. Standard deviation is referred to as “volatility” in a financial context, and measures how much returns vary in a period. Our is measure is looking ten days into the future and is therefore calculated as follows:

$$Future\ volatility_t = sdev(volatility_{t+1} +\ volatility_{t+2} + ...\ volatility_{t+10}) \quad (2)$$

We further transform the feature by taking the percentage change of the future volatility to make it stationary. The main reason to do this is however to get the scale in ALE in such a way that it shows how much the volatility changes based on the features.

$$Transformed\,future\ volatility_t = .\frac{Future\ volatility_t}{Future\ volatility_{t-1}} - 1 \quad (3)$$

4.2 Price-related variables as the features

The features derived from the price for each cryptocurrency are the closing and opening prices of each day, daily high and low, and market capitalization. We will transform the raw values for each coin into a suitable format by changing them into logarithmic changes with the following formula:

$$Logarithmic\ change_t = \ log(x_t) - log(x_{t-1}) \tag{4}$$

Since these variables are highly correlated, using neural networks can be assumed to yield more accurate results as they tend to manage multicollinearity well. This is also the reason for choosing ALE as our method instead of PDP, as they give a more accurate presentation of variables that are highly correlated with each other.

4.3 10-day rolling average of past sentiment as a feature

Sentiment analysis is a type of textual analysis where one systematically quantifies text and rates the text as positive or negative on a continuous scale. Each piece of text can be given a score without knowing the context by scoring each individual term and calculating the average of the scores in the text.

We used textual data of news articles that were scraped from the Coindesk website, which focuses on cryptocurrency-related news and data. The data ranges from the 28th of March 2016 until the 19th of March 2021, spanning a period of 5 years. To further clean the text for the sentiment analysis, we removed all punctuation, numbers, and stop words from the text, converted the text into lower case, and performed stemming on the resulting text. We further limited the required number of characters between 4 and 20, to avoid words that may contain too little or too specific information. Many of these steps do not necessarily increase the accuracy of the analysis, but rather reduces the time and data sizes that the analysis requires.

The most positive news regarding Bitcoin was as follows before performing the cleaning steps as described above:

About 93% of all bitcoin address balances are estimated to be in the black, according to Glassnode , as bitcoin continues to trade above $11,000. \"\"Addresses in profit\"\" measures balances of assets transferred into a wallet at an average price lower than the current price. In other words, their value has gone up, creating a profit, at least on paper.Balances in profit are up 29 percentage points from the 72% mark recorded on July 20.More than 90% of bitcoin addresses were last in profit through July and August 2019 when bitcoin traded around $11,500, a local top for the leading cryptocurrency.The 11-month high for in-profit addresses comes amid bitcoinâ€™s rally to $11,400 and follows steady long-term accumulation by investors at lower prices, with fewer than 40% of all bitcoins having been moved in the past year.Â A high percentage of addresses \"\"in profit\"\" increases the probability of investors selling those bitcoins to realize their profit, said Josh Olszewicz, cryptocurrency trader atÂ TechemyÂ Capital. That doesn't mean the price won't *continue to appreciate, he added.Percentage of bitcoin addresses in profit since July 2019 Update (July 29, 22:18 UTC): Â This article has been updated with a comment from Josh Olszewicz .*

After performing the cleaning steps on the text above, we obtain the following text:

bitcoin address balanc estim black accord glassnod bitcoin continu trade address profit measur balanc asset transfer wallet averag price lower current price word valu gone creat profit least paperbal profit percentag point mark record juli bitcoin address last profit juli august bitcoin trade around local lead cryptocurrencyth month high inprofit address come amid bitcoinâ€™ ralli follow steadi longterm accumul investor lower price fewer bitcoin move past yearâ a high percentag address profit increas probabl investor sell bitcoin realiz profit said josh olszewicz cryptocurr trader atâ techemyâ capit doesnt mean price wont continu appreci addedpercentag bitcoin address profit sinc juli updat juli â this articl updat comment josh olszewicz

The paragraph looks abnormal mostly due to the stemming, which usually removes the ending of a word to make it more machine-readable and consistent since otherwise, the sentiment model would not understand that two words with different endings are related. Also, a lot of the text has been removed, since due to not using a model trained to understand the context of the text, the removed text contains no useful information for the sentiment analysis.

After cleaning the text, we transformed the data from vectors of text into document-term matrices, containing 15,582 different news and 117,801 different terms. We chose to use the *SentimentAnalysis* package, developed by Proellochs and Feuerriege (2021) for the R language to perform the sentiment scoring, and we based the scoring on the Loughran-McDonald dictionary with stemmed words, which is suitable for financial data. We assume that the dictionary may be missing some cryptocurrency-specific terms, but would most likely yield results that are good enough to achieve the goal of this paper.

As a result, we obtain a sentiment score for each news item. For this data, the sentiment score ranged from a negative sentiment of -0.11 to a positive sentiment of 0.08, with a value of 0 being a neutral sentiment. We further classified whether the specific coin was mentioned in a news item, and filtered only those to be included for that coin. We also filtered out sponsored posts, since they may be biased when compared to other posts on Coinbase.

The best way to illustrate how sentiment analysis works, we will provide an example. The following article from October 2020 was classified as having a sentiment score of -0.10, which is one of the most negative of all articles from the data:

Over $26 million worth of bitcoin (BTC) associated with the massive 2016 ***hack*** *of crypto exchange Bitfinex was moved around across seven transactions on Wednesday, according to Twitter-based blockchain tracker bot Whale Alert .Â The* ***security breach*** *at Bitfinex in August 2016 led to the* ***theft*** *of over 120,000 BTC (worth approximately $1.2 billion today). Similar movements of* ***stolen funds*** *were also reported in July .Â Six of the transactions* ***flagged by Whale Alert*** *ranged between $4.1 million and $4.8 million, with one for a* *relatively small amount of $12,000.Â   Read more: Whale Alert: $27M From 2016 Bitfinex Hack Is on the Move*

The news article states that a large amount of bitcoin is in motion, and the bitcoin can be traced back to a hack from 2016. The reason why this news article is rated as very negative is probably due to words like hack, security breach, stolen funds, and theft. This article can also illustrate weaknesses in the method. It is unlikely that this article is in fact the article that has the most negative sentiment, given that it is about a hack that happened 4 years before. These are small details that the sentiment analysis probably fails to capture, but we still believe that averaging the sentiment within and between dates will be closer to the actual sentiment of that time.

To create a feature based on the sentiment for each coin, we filled days with no sentiment with zero and took the average of the sentiment from the past 30 days for each day. To avoid possible leakage related to different time zones between the news data and the price data, we excluded one day in between, and therefore use $_{t-1}$ in the calculation instead of $_t$. We limited the data to a period where there was both price data and sentiment data available for the specific coin.

The formula for this feature variable is calculated for each coin separately as follows:

$$Average\ sentiment_t = \frac{sentiment_{t-1} + sentiment_{t-2} + \dots\ sentiment_{t-30}}{30} \tag{5}$$

After this, the variable is transformed by taking the percentage change the same way as in Equation 3 for the future volatility. This was done for similar reasons, but we ended up analyzing the effect of the sentiment score with ALE instead of the change of the sentiment score.

It is important to keep in mind that since Coinbase is focused on cryptocurrencies and is, therefore, most likely to not publish only articles with objective views of different cryptocurrencies due to a conflict of interest, we must keep in mind that the sentiment obtained from the website may not present the overall sentiment regarding cryptocurrencies.

4.4 Common feature engineering

In addition, we will perform some transformations for all of the features and the target. We replaced missing or infinite values with zeros, since those were caused by some of the previous steps, and neural networks can only handle data that is ultimately in numeric form. Before shaping the data to be suitable for the LSTM network, we perform min-max normalization to rescale the data to have a range between 0 and 1. The transformation is shown in the equation below:

$$Scaled\ value = \frac{value - min(value)}{min(value) - max(value)} \tag{6}$$

Doing this typically leads to faster convergence, which speeds up the training process of neural networks. (Ioffe et al., 2015) The downside is that we have to back-transform the predictions to get them in the original scale, which we need to do in our case with ALE.

5 Accumulated Local Effects (ALE)

In order to understand how the particular feature affects predictions, one could want to estimate the average effect of the feature on the prediction. For that purpose, Partial Dependence Plots (PDP) are commonly used. However, the PDPs could be biased if the predictors are correlated, given that it could utilize the predictions of unlikely artificial events as averaging living area of the house (the feature of interest) to 30 m^2 and use it for all observations even for the houses with 10 rooms (highly correlated feature), and include these unlikely instances in the feature effect estimation and not actually consider the effect of a single feature at once. As an alternative to PDPs, which averages the predictions over the marginal distribution, Accumulated Local Effects (ALE), which averages changes in the predictions and accumulates them over the grid, could be used. ALE plots describe how the prediction would change in response to the change of the feature, by isolating the effect of the feature and excluding the effect of correlated features. (Apley, 2016) The ALE plots the effect of our features on our target, the future 10-day volatility.

6 Exploratory analysis

One of the most important steps in all modeling is the exploratory analysis, which we perform on all of the data that goes into the models. It may help us to understand things that affect the modeling process, such as if some coins behave differently from others. All of the data is limited to the period from the 28th of March 2016 until the 19th of March 2021, since we will only use that period for training the models since we only have the scraped news data from that period.

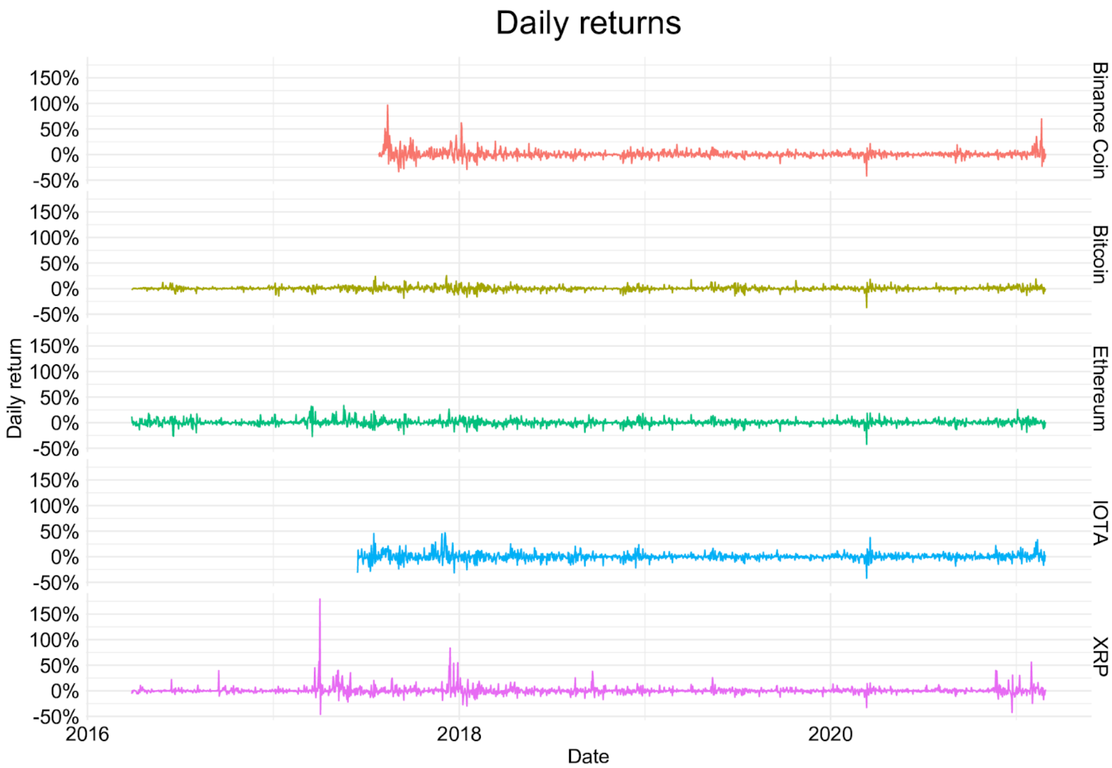


Figure 1. Daily returns of the different cryptocurrencies

Figure 1 illustrates the historic time series data for the daily returns. The fluctuations in the plot indicate similar patterns and dispersion of data for all coins. The plot also shows that XRP seems to have larger sudden spikes. The data availability varies by coin, due to some of the coins being rather new.

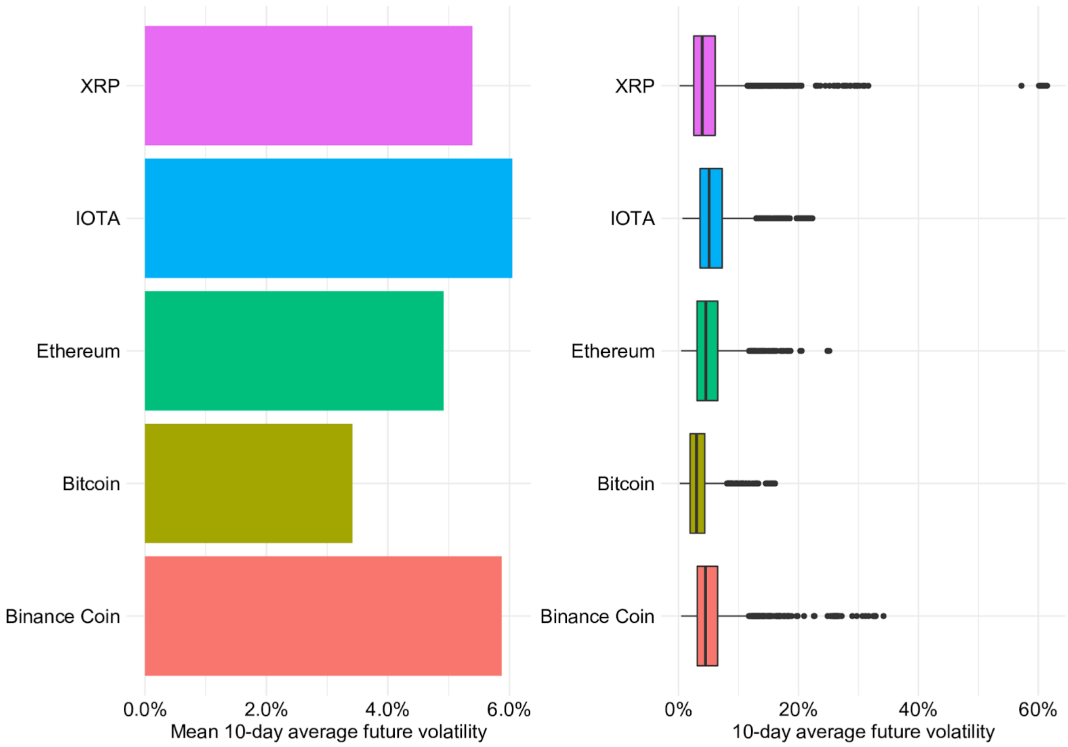


Figure 2. 10-day average future volatilities of the different cryptocurrencies

From the box plot of *Figure 2*, one could observe some strong outliers for Ripple, with the 10-day future volatility reaching 60% in March 2017, possibly due to growing bank membership. Bitcoin has a considerably smaller average volatility as compared to the other coins and seems to have the smallest outliers, most likely due to being the oldest, largest, and most well-known cryptocurrency. IOTA and Binance Coin had the highest volatilities, but the latter one had larger and more outliers.

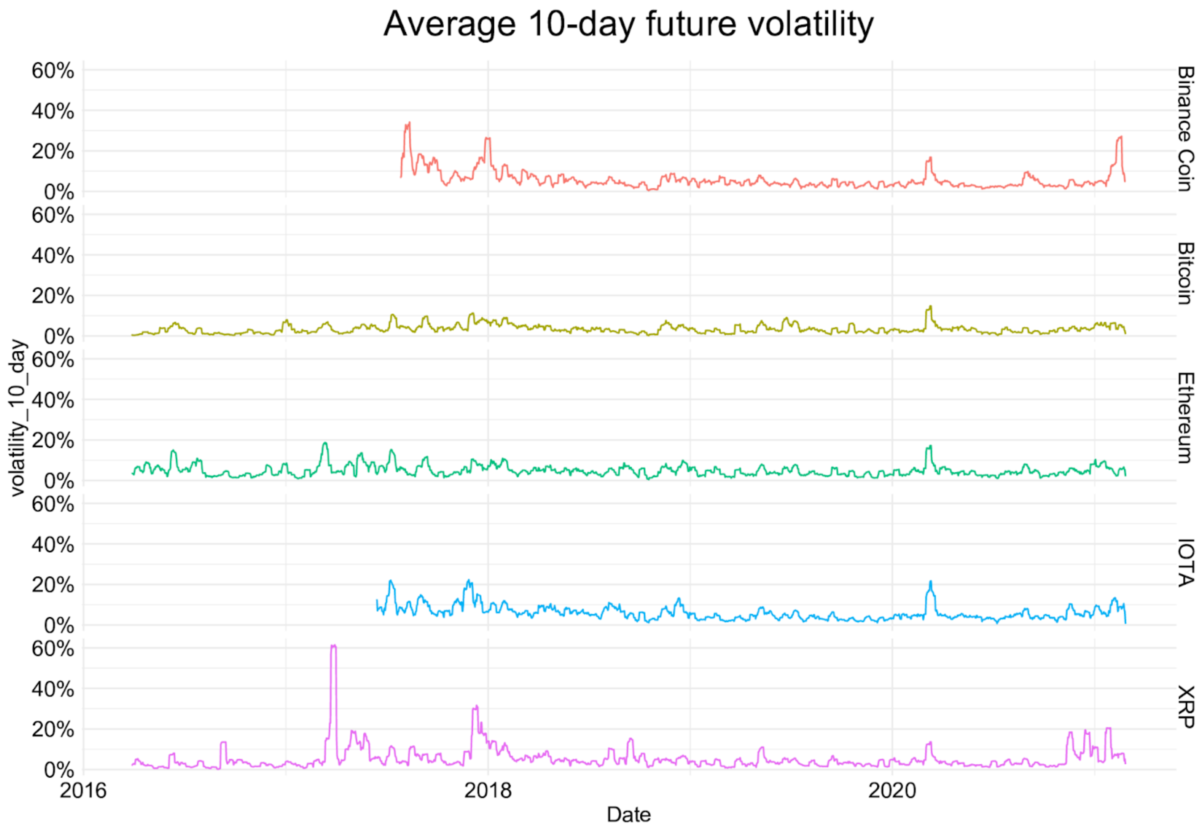


Figure 3. 10-day average future volatilities of the different cryptocurrencies

Visual examination of *Figure 3* suggests that the volatility of coins tends to appear in sudden spikes which could be positively correlated with each other, as coins follow significantly similar patterns. In order to achieve formal results, correlation coefficients are obtained.

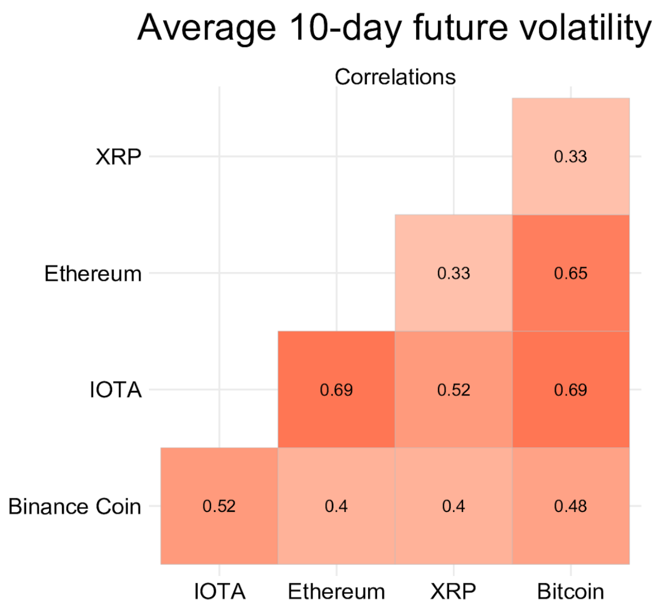


Figure 4. Correlations of the 10-day average future volatilities of the different cryptocurrencies

There is a quite strong linear relationship between all the coin volatilities as shown in *Figure 4*, however, following the p-values show that correlations are not statistically significant at a 5% significance level.

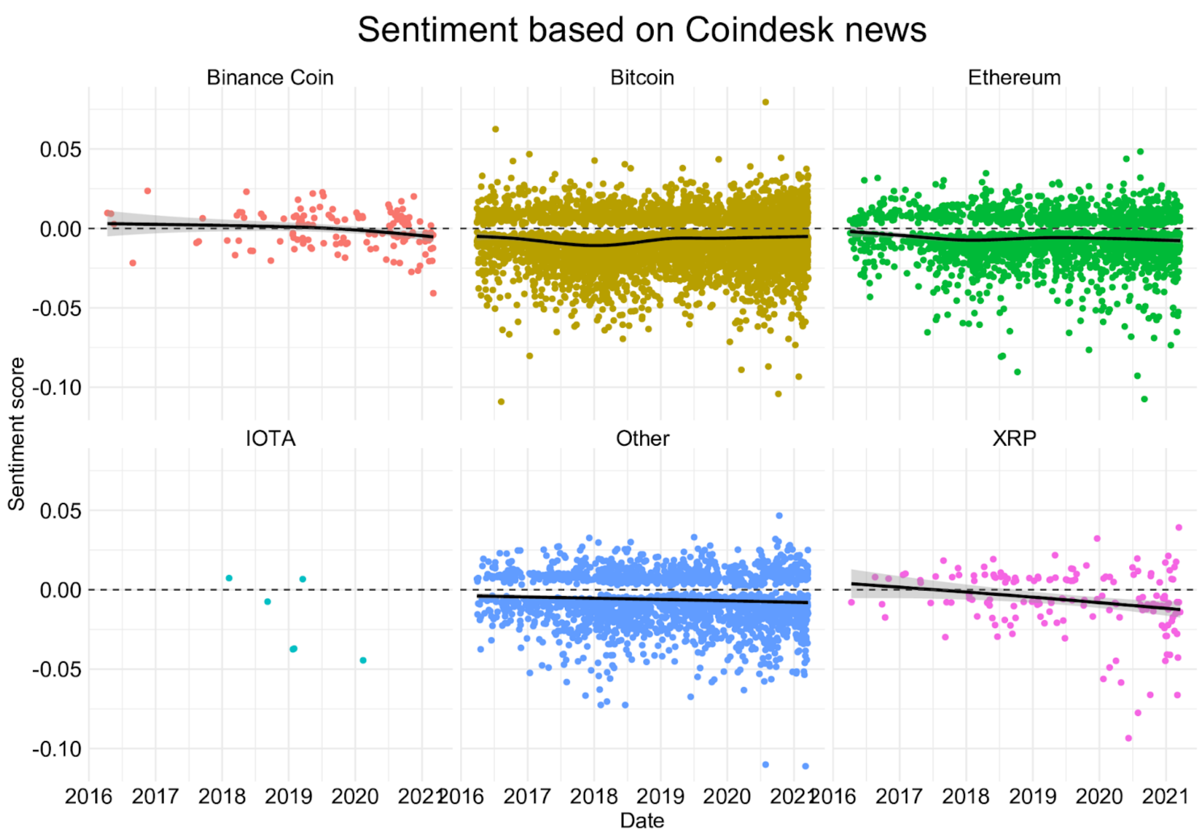


Figure 5. Sentiments of each news item by mentioned cryptocurrency

The sentiments for most coins have been negative on average, as illustrated in *Figure 5*. This is surprising considering that the news source could be assumed to depend on the positive performance of cryptocurrencies, and therefore post positive articles about them. Most of the cryptocurrencies were not mentioned too often as compared to Bitcoin and Ethereum. It is important to note that if a news article mentioned all of these coins, it would show as a dot in all except the “Other” category since we did not filter out articles that mentioned multiple coins. Some of the coins have rather few observations, which will affect the abilities of the corresponding models to fit and the trustworthiness of the results.

7 The models

We experimented with both LSTM and GRU, by fitting multiple models with different hyperparameters for each coin separately. Our goal was to minimize the validation set accuracy as measured by MSE, since we trained the entire model without using a test set due to the nature of our analysis, where we are not maximizing the accuracy of future forecasts but rather finding the relationships between the features and the target as given by the model.

7.1 Iterations

We mainly tested models for the different coins separately with both LSTM and GRU layers. It was therefore important to be very careful and systematic with the hyperparameter tuning so we do not waste time and computing power. To get an overview, we manually tested some hyperparameters with 3000 epochs and saw that the models almost always started overfitting before 1000 epochs. Furthermore, we saw that 1 layer was generally better than 2 and that a batch size of 64 reduced computational loads without reducing accuracy too much. Common to all the models that were systematically tested were using the ADAM optimizer, batch size of 64, one LSTM layer with the TANH activation function, MSE as loss, a validation split of 20%, a sigmoid recurrent activation and a hard sigmoid for GRU.

Otherwise the following hyperparameters were tuned:

Models: {LSTM, GRU}
Units: {3, 8, 16, 32}
Dropout: {0, 0.25, 0.5}

This means we end up with a total of 24 iterations for each coin. We used the *tfruns* package to automate the hyperparameter tuning and logging. To make sure that the iterations stopped before the models were overfitted, we used early stopping so that if *val_loss* was reduced by less than 0 during the last 250 iterations, then the training would stop and the next run would be started. This led to some overfitting still because the validation loss could increase by a substantial amount in the last 250 epochs. This happened with the Ethereum GRU model as can be seen in the appendix.

All models reached peak accuracy before 1000 epochs, which means that the models stop improving relatively quickly. The GRU models generally performed better with more units than the LSTM. GRU also generally performed better when using *dropout* for regularization, while LSTM performed best without the regularization. Bitcoin performed the best with 16 units for both the models, XRP performed best without dropout, and the Binance coin performed best on both models with a 0.25 dropout rate. Otherwise, the models did not differ in the optimal parameters.

LSTM performed best for Binance Coin and IOTA, while GRU performed best for Ethereum, Bitcoin and XRP. However the performance is just marginally better between them, so there could be insignificant differences in improvement. It is also important to keep in mind that by using the validation set to make decisions regarding the hyperparameters, the validation set cannot be anymore trusted as much as an accuracy measure due to leakage.

8 Results

Table 1. Results from the iterations

Coin	Model	Loss	Validation Loss	Units	Dropout	Optimal epochs
IOTA	LSTM	0.0284	0.0273	8	0	787
XRP	GRU	0.3384	0.0767	3	0	315
Binance Coin	LSTM	0.0277	0.022	3	0.25	307
Bitcoin	GRU	0.0656	0.0777	16	0	990
Ethereum	GRU	0.0379	0.0431	32	0.5	994
IOTA	GRU	0.0277	0.0285	16	0.25	354
XRP	LSTM	0.3384	0.0769	3	0	355
Binance Coin	GRU	0.0278	0.0365	8	0.25	409
Bitcoin	LSTM	0.0661	0.0786	16	0	691
Ethereum	LSTM	0.0368	0.0456	16	0.25	266

The results of the best iterations of the models are presented above, and the plots of the fitting processes are shown in the appendix. Rest of our results are based on these models.

We ran ALE using the code from the library *ALEplot*, which we modified the source code of to our needs to get just the data of ALE instead of plots. Since any data that can be transformed for the model can be given to the ALE function, we decided to use the data where log-returns have been taken out of the variables related to price, but no transformation was not yet done to the sentiment score or the future 10-day volatility. This way we obtained results that show the effect of the sentiment score itself and the changes in the price-related variables on the actual 10-day volatility. To do this, we had to write a custom function that preprocesses the data given to ALE the same way as to how the data is preprocessed before modeling, including all the preprocessing steps as described in *Chapter 4*. We ran the ALE for 100 different values of the features, with an even distance.

We will first inspect the effects of the price-related features on the future volatility. We view these results as secondary as our primary focus is in the effect of the sentiment score.

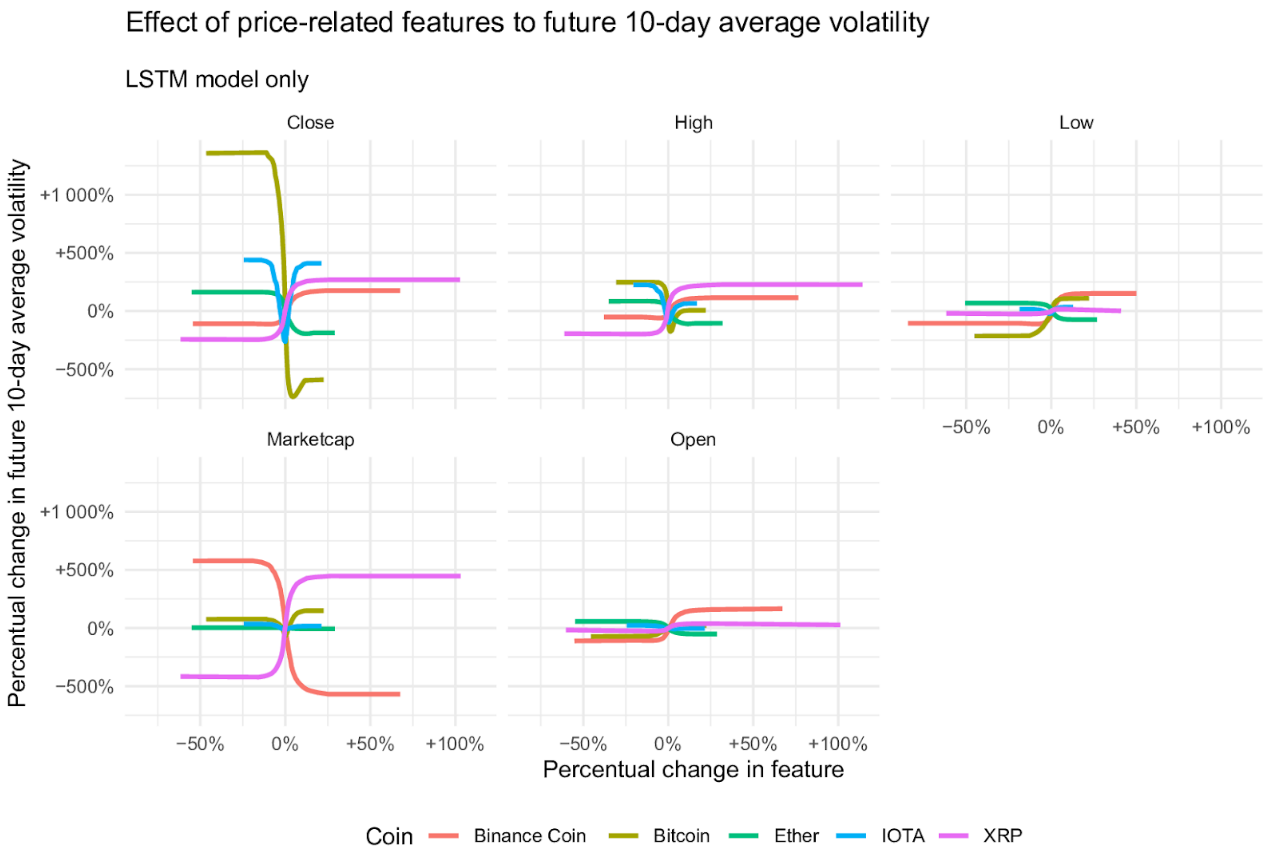


Figure 6. Effect of the price-related features on the future volatility of each coin using LSTM

The resulting figures show interesting patterns that a strictly linear model would not be able to capture. We do not consider these results to contain a lot of information, since the features are highly correlated and every feature should be looked at the same time since they are all related to the prices. However, it gives us a look inside how the model takes each feature into account.

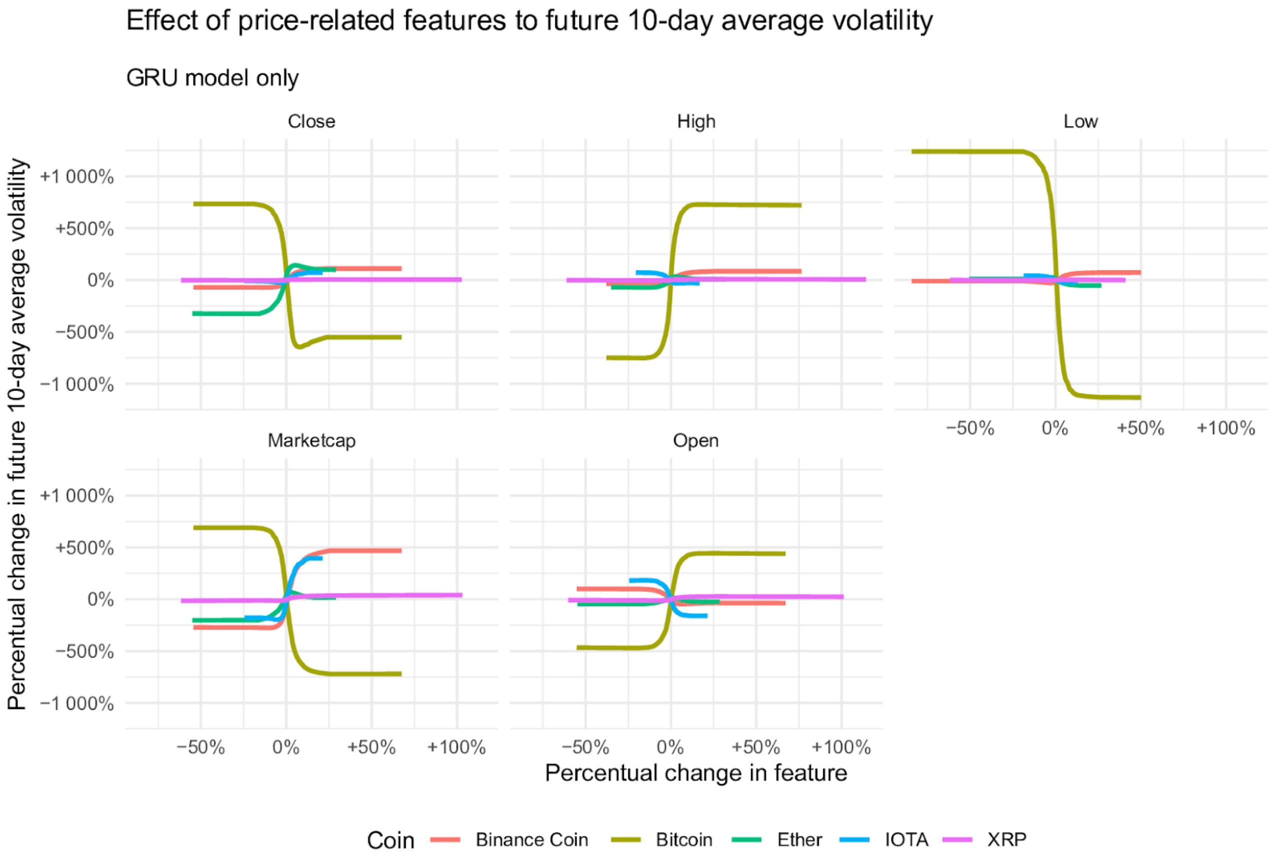


Figure 7. Effect of the price-related features on the future volatility of each coin using GRU

We can see that the features affected the GRU model of Bitcoin the most.

In general, the closing price has a positive effect on the predictions for all coins except for bitcoin, which means that the higher volatility is more likely to be predicted when the price goes up.

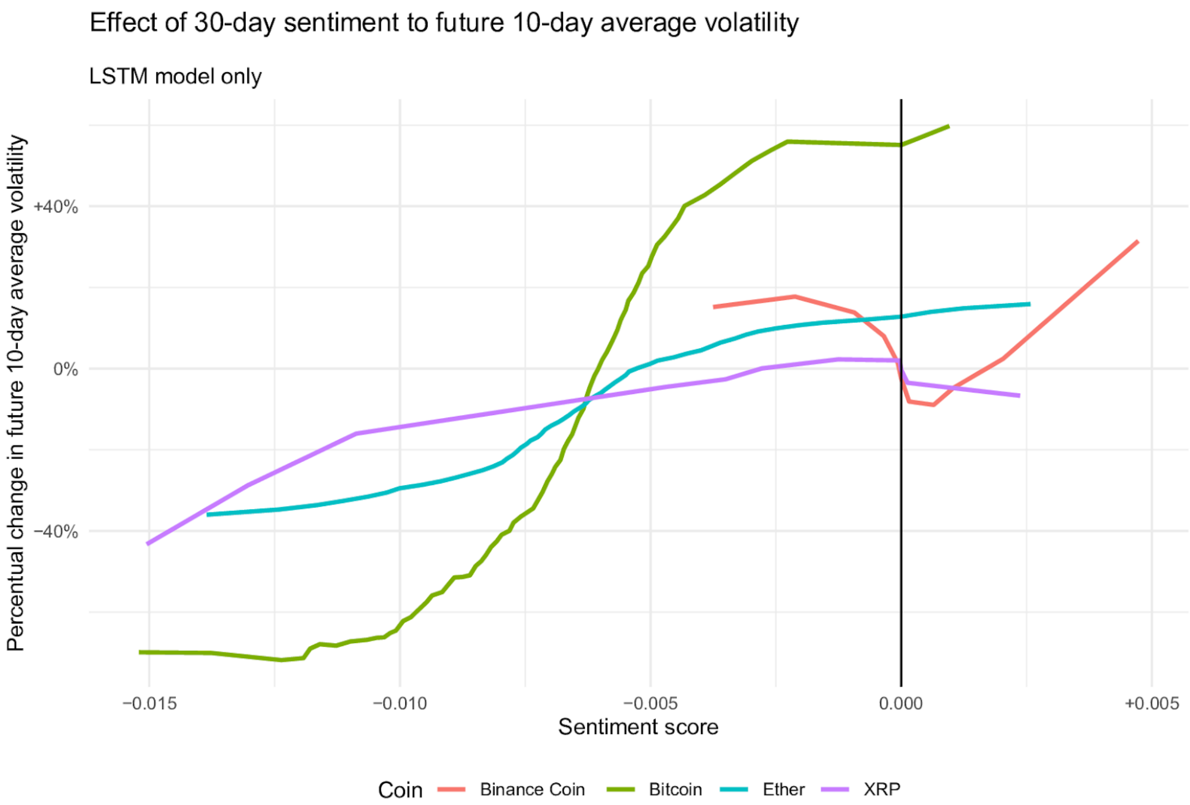


Figure 8. Effect of the sentiment score on the future volatility of each coin using LSTM

The LSTM-based ALE shows that generally as the average 30-day sentiment score of the Coindesk news regarding the specific coin increases, usually more volatility will follow over the next 10 days. We assume that this is usually upwards price movements caused by positive news. It is however interesting that the negative sentiment score seemed to decrease the future volatility. It might be that these news are misclassified as negative in case Coindesk does not typically give negative views of the coins, but this would require further analysis.

Bitcoin seemed to be affected the most by the sentiment, even though it was the least volatile coin. We suspect it might be since there are more things driving the volatilities of other coins than just news. Interestingly, the Binance coin had a V-shape, meaning that volatility has mostly decreased based on both negative and positive news. It is the coin most affected by cryptocurrency adoption, as it is directly tied to the exchange community. IOTA had so few news items at Coindesk that its effect could not be calculated.

The volatility to sentiment relation plot can follow a straight line where sentiment does not affect volatility at all, otherwise, the form can take a v-shape where both positive and negative sentiment values lead to increased volatility, or the relationship can have a non-linear positive correlation where increased sentiment results in increased volatility. The relationship may also have a non-linear negative correlation shape. Good news positively affects the volatility of XRP GRU, Bitcoin LSTM, and Ethereum LSTM in particular. XRP GRU's volatility is positively affected by both high and low sentiment values. The volatility of the Ethereum GRU model is affected by sentiment in the opposite way to Ethereum LSTM. If

Sentiment increases, ETH GRU will predict reduced volatility, while ETH LSTM will predict increased volatility. Bitcoin GRU and Binance Coin GRU do not seem to include a sentiment to any great extent in the model.

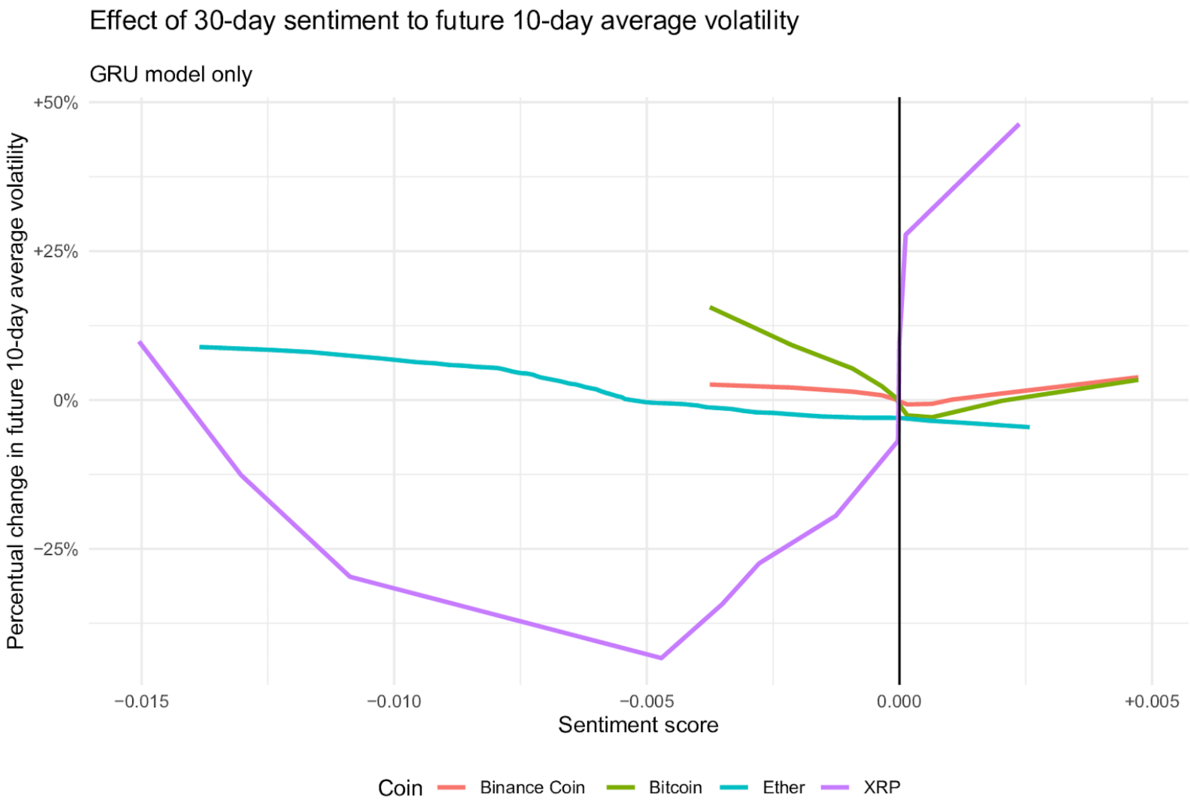


Figure 9. Effect of the sentiment score on the future volatility of each coin using GRU

The GRU-based ALE shows a different shape to the LSTM, which cannot be explained as easily. We consider this model to be less trustworthy, as it also had lower validation accuracies as measured by mean squared error.

For the sentiment feature, the ALE plot demonstrates that volatility is affected slightly differently by the different models for the five coins.

We also tried to run the previously introduced PDPs for the different models to be able to compare them to ALE. After trying the libraries *iml, pdp* and *ALEplot,* none of them worked for the PDP. We suspect that this was due to the high multicollinearity in the non-transformed variables.

9 Conclusion

The effect on future realized volatility has been observed in several aspects. In particular, how changes in different features affect the average future 10-day volatility was illustrated. In this paper, we considered price-related variables and sentiment indices as predictors. After performing normalization of inputs and textual analysis of sentiment data, neural network models LSTM and GRU were trained for each coin separately. Lastly, ALE plots were discussed. This report shows the potential of studying the factors which affect the volatility prediction of cryptocurrencies. Our main finding was that increased sentiment typically increased the future volatility of different cryptocurrencies based on an LSTM model, while the results from GRU were untrustworthy.

References

Apley, D. W. (2016) *Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models*. arXiv:1612.08468 [stat.ME]

Apley, D. W. (2018) *Package 'ALEPlot'.* Available: https://cran.r-project.org/web/packages/ALEPlot/ALEPlot.pdf

Business Insider (2020) *Anthony Scaramucci's SkyBridge hedge fund just invested $25 million in a bitcoin fund as it sees an 'avalanche of institutional investors' buying crypto in 2021. Available:* https://markets.businessinsider.com/currencies/news/anthony-scaramucci-skybridge-capital-bitcoin-fund-launch-institutional-cryptocurrency-btc-2020-12-1029915328

Ioffe, S., Christian, S. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. arXiv:1502.03167 [cs.LG]

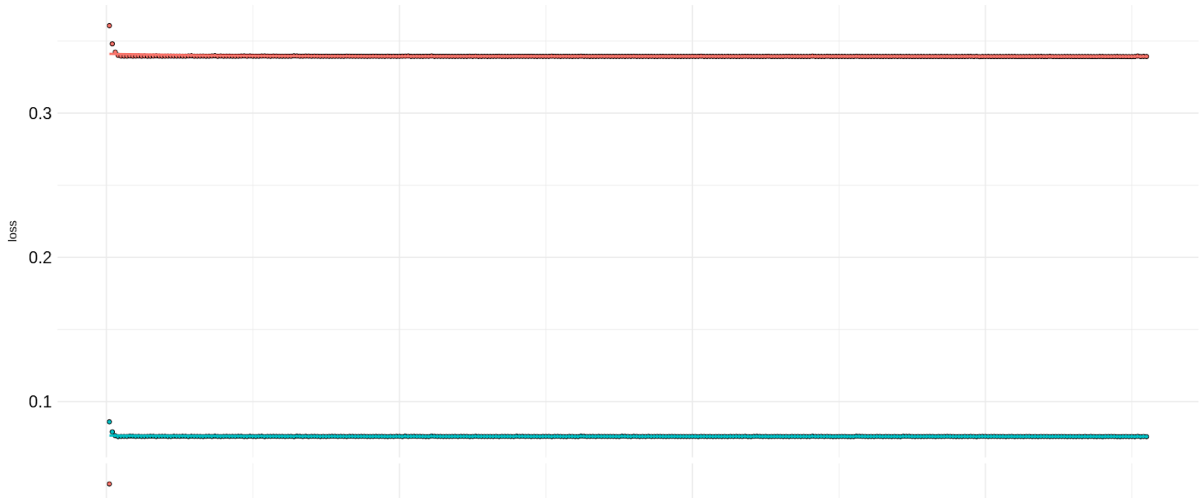
Proellochs, N., Feuerriege, S. (2021) *Package 'SentimentAnalysis'*. Available: https://cran.r-project.org/web/packages/SentimentAnalysis/SentimentAnalysis.pdf

Xiong, R., Nichols, E. P., & Shen, Y. (2015). *Deep learning stock volatility with Google domestic trends*. arXiv e-prints arXiv:1512.04916.

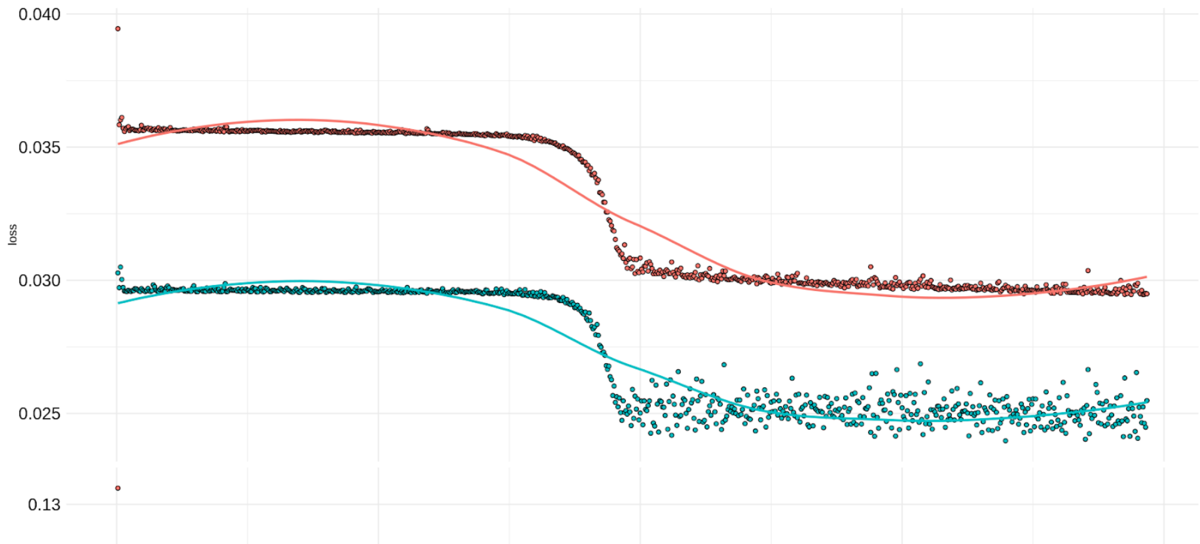
Y. Liu (2019) *Novel volatility forecasting using deep learning–Long short term memory recurrent neural networks*. Expert Systems with Applications, 132 , pp. 99-109

Appendix

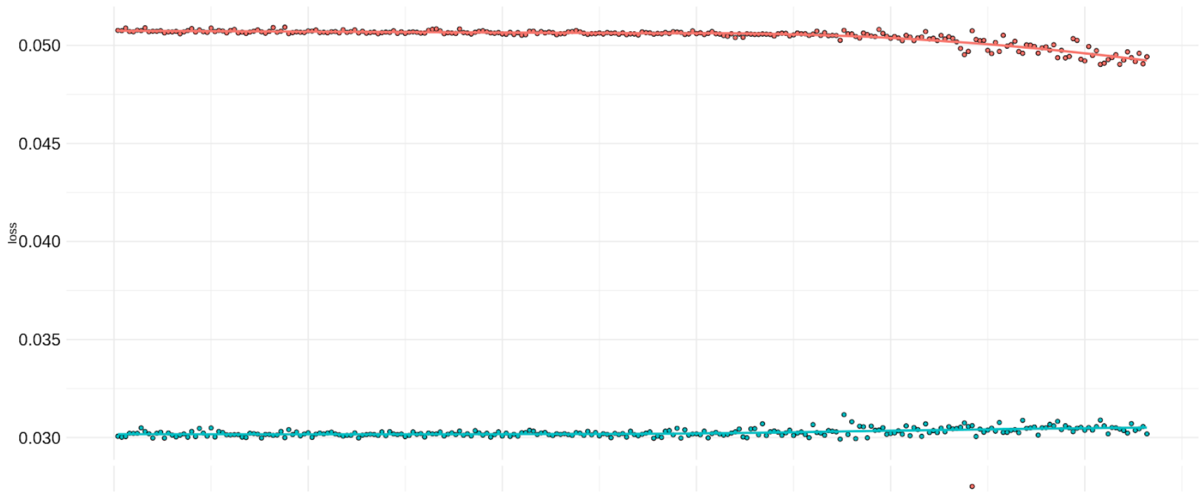
All the Iteration plots follow this structure: The plots show MSE validation loss (blue line) and MSE loss (red line).



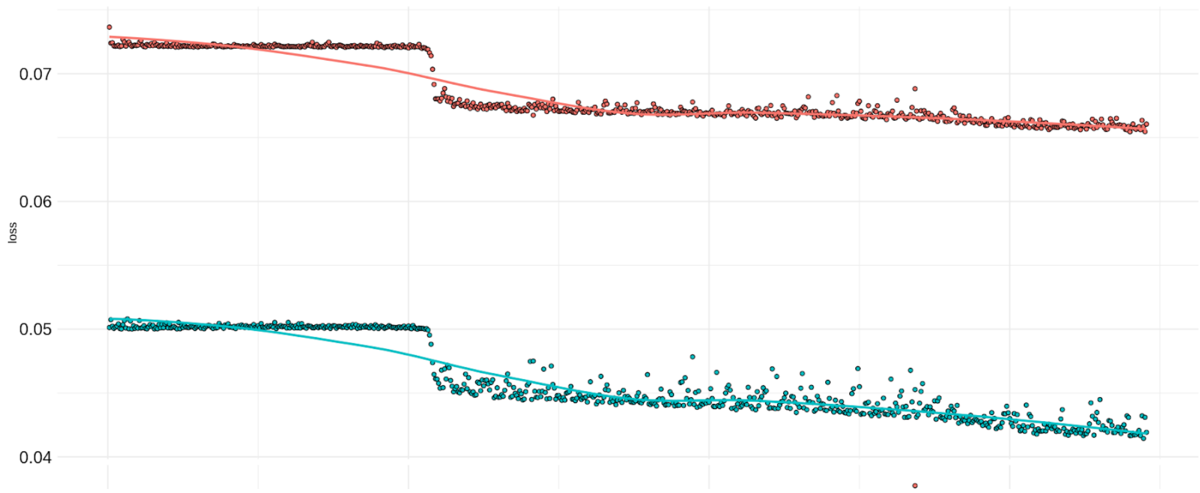
Appendix Figure 1. Fit of the XRP LSTM model



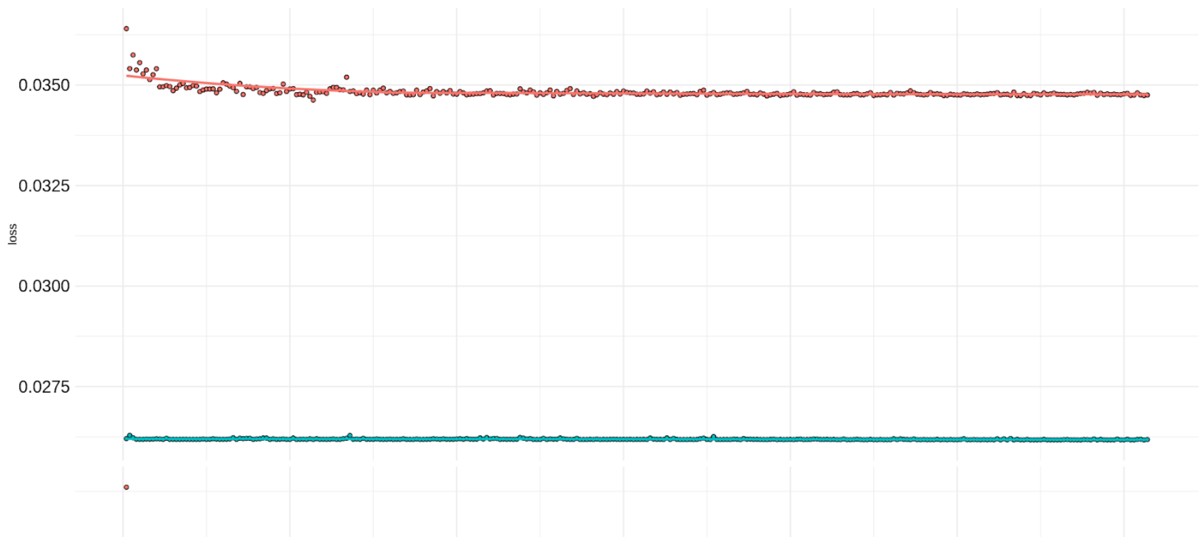
Appendix Figure 2. Fit of the IOTA LSTM model



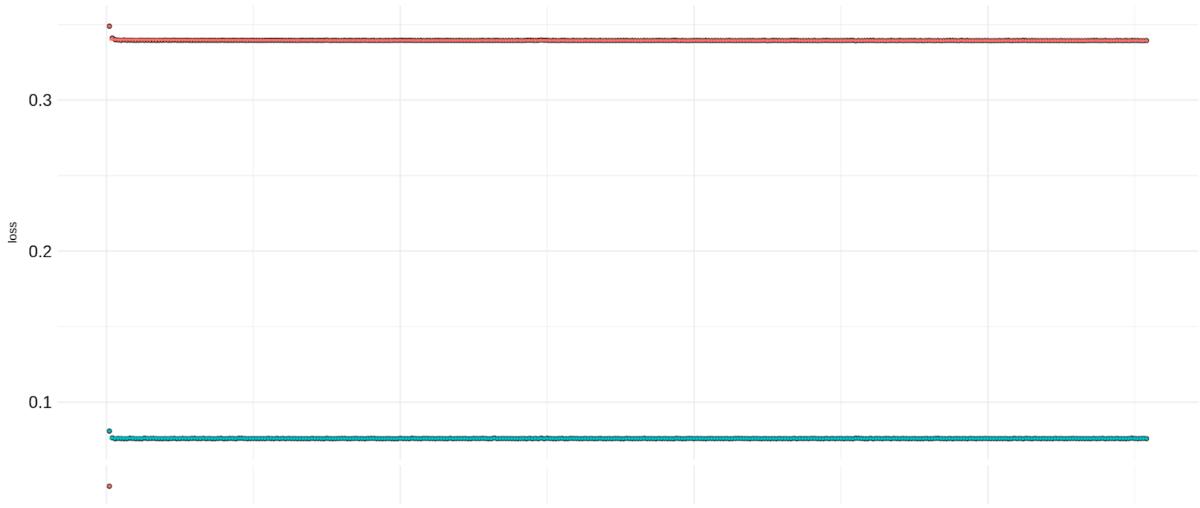
Appendix Figure 3. Fit of the Ethereum LSTM model



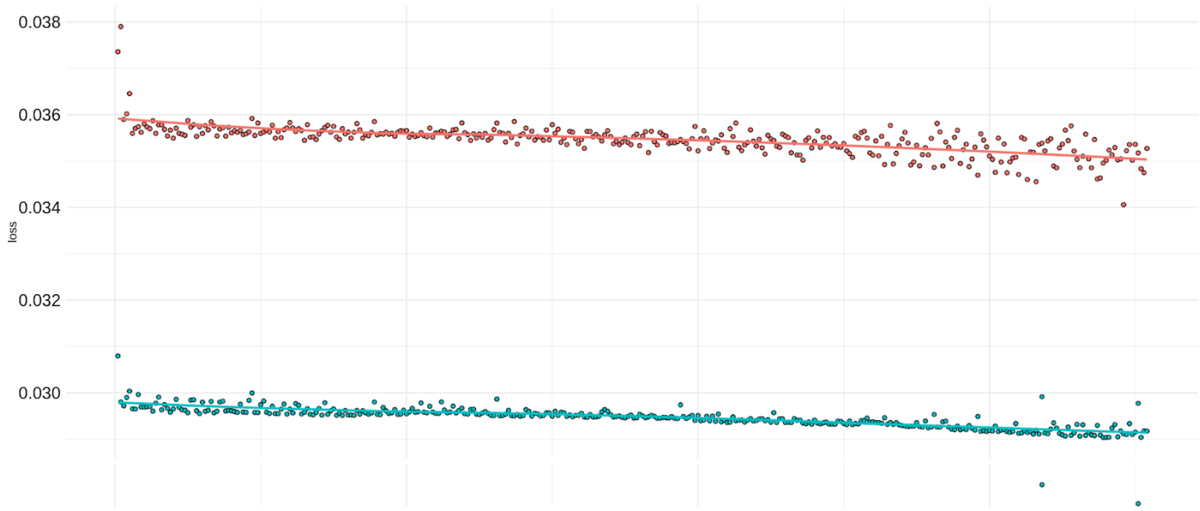
Appendix Figure 4. Fit of the BTC LSTM model



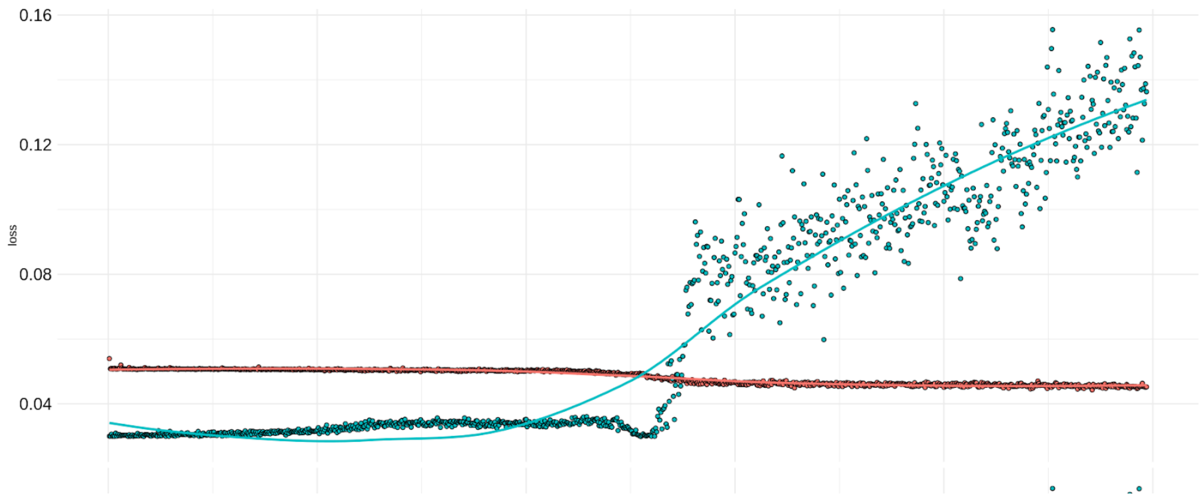
Appendix Figure 5. Fit of the Binance Coin LSTM model



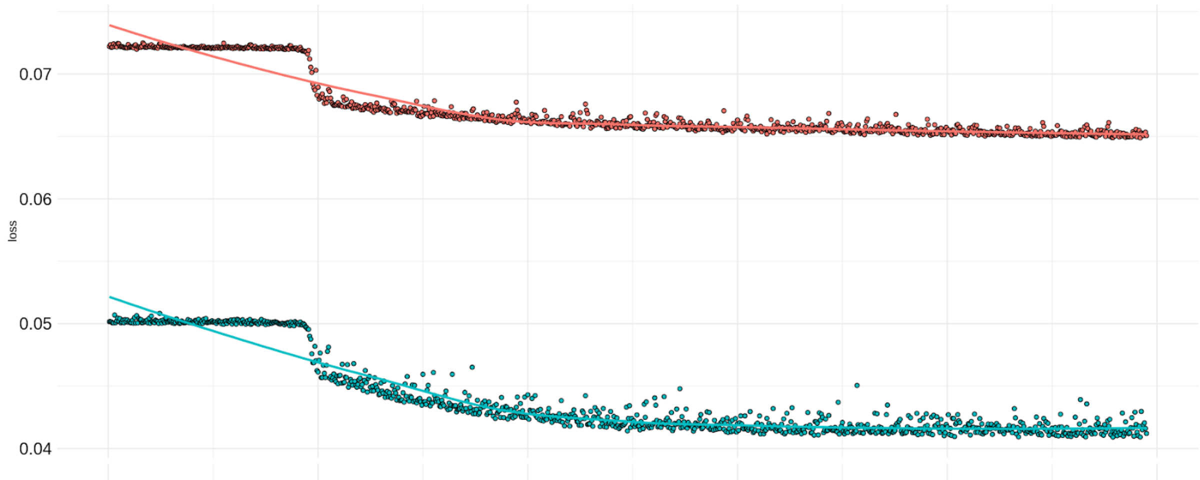
Appendix Figure 6. Fit of the XRP GRU model



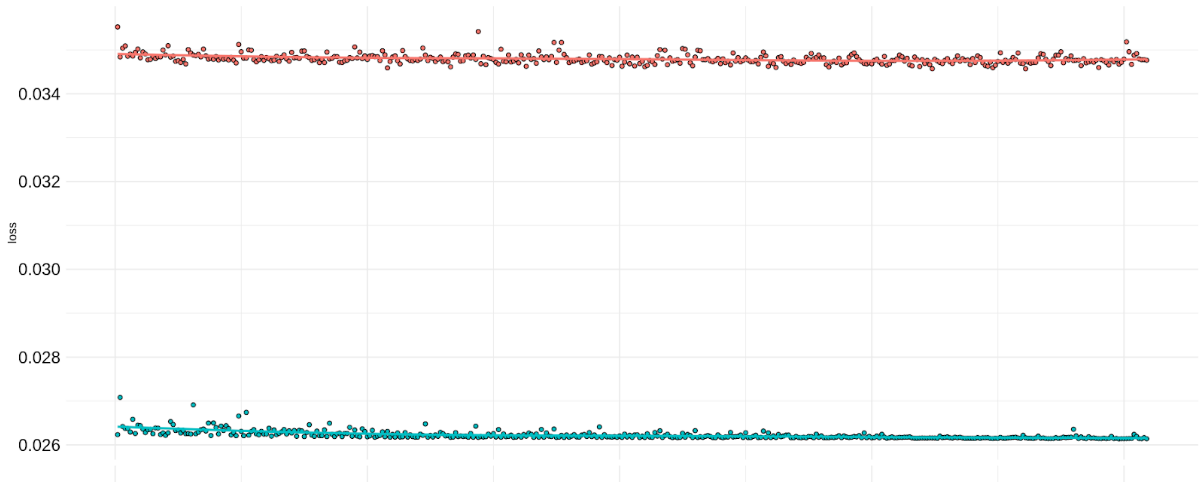
Appendix Figure 7. Fit of the IOTA GRU model



Appendix Figure 8. Fit of the Ethereum GRU model



Appendix Figure 9. Fit of the Bitcoin GRU model



Appendix Figure 6. Fit of the Binance GRU model