Forecasting Cryptocurrency Volatility

Department of Physics, Washington University in St. Louis

Daniel Wasserman
Advisors: Mina Lee and Li Yang

Spring 2017

Abstract

The advent of Bitcoin has sparked massive growth in cryptocurrencies in recent years. The prices of Bitcoin and other cryptocurrencies are relatively volatile, which deters people who are risk-averse from investing in them. Many smaller cryptocurrencies are even more volatile than Bitcoin. A model that can predict cryptocurrency volatility based on price motion would benefit potential cryptocurrency investors by providing them with extra information about risk. A random forest machine learning model is used to predict cryptocurrency volatility; this model is compared to the GARCH volatility model that is widely used in modern finance as a benchmark. A discussion regarding model accuracy reveals how predictable cryptocurrency volatility actually is. Finally, comparisons of models with different feature selections give insight as to what factors can significantly influence cryptocurrency volatility.

Contents

1 Introduction

Although many people may view cryptocurrencies as trivial internet gadgets, this new digital asset class could have significant implications for the world economy. Cryptocurrencies pose several advantages over conventional fiat money. The most significant advantage of cryptocurrencies is decentralization of the ledger, which allows transactions to be verified by a network of nodes rather than one central authority.

Decentralization of the ledger means that no single authority validates transactions. This property improves the security of a decentralized network, as a compromise of the single central party cannot shut down the network. Implementations of blockchain technology allow for such decentralization. The specific implementation of the blockchain varies between cryptocurrencies, but the essence of the technology is that every node on the blockchain can validate transactions and broadcast these verifications to the rest of the network. Decentralization of the ledger also implies that network users are not exposed to the whims of a central authority. For example, if the US government decides to increase inflation considerably, all holders of the US Dollar would suffer. Such an unforeseeable action would not occur on a blockchain; the inner workings of most cryptocurrencies are transparent.

Bitcoin is one of many cryptocurrencies available for trading on several global exchanges. Most cryptocurrencies target specific purposes. Monero, for instance, focuses on transaction anonymity, while Ethereum is a platform that allows users to build decentralized apps which run on the network's blockchain. There is a plethora of ideas that are currently being implemented or are yet to be implemented with blockchain technology. Additionally, the supply of every cryptocurrency is determined by an algorithm that is coded into the currency's blockchain. Users know exactly how new coins will be introduced to the network, a property that conventional governments do not offer. Algorithmic supply control could be one fundamental difference that causes prices of cryptocurrencies to move differently as compared to other currencies.

Naturally, cryptocurrencies are difficult to price because most do not yield any revenue streams; most generate value by providing utility to holders in the form of spending power on their respective blockchains. While financial analysts can inspect Price/Earnings ratios or use structures like the Dividend Growth Model to arrive at valuations, cryptocurrency valuation is not as simple. Today, cryptocurrencies are far from reaching their envisioned full-scale operations, which implies that much of the value represented in current markets is based on expectations that they will become widely adopted. Such forward-looking prices lead many to believe that cryptocurrency price motion is driven purely by speculation.

One way that Bitcoin (and potentially other cryptocurrencies) provides value to investors is the fact that it acts as a new asset class. Bitcoin's price is not highly correlated with any other major asset, which suggests that it may present value as a hedge. [1] The fact that other cryptocurrencies may have different correlations with Bitcoin's price means that cryptocurrency could provide a set of hedging tools that investors can utilize to construct more efficient portfolios. In general, cryptocurrency prices are not directly dependent on cash flows or a country's economic performance, although these dynamics could change in the future as more institutional investors adopt these digital assets.

The drivers behind cryptocurrency price motion are not completely clear. One study postulates that Bitcoin's price is mostly driven by long-term fundamentals such as supply and demand and exchange rates. [2] A different study finds that events such as the Mt. Gox exchange hack are the major source of Bitcoin price volatility. [3] While analysis of a machine learning algorithm used to predict cryptocurrency volatility will provide some information as to what the best predictors of volatility are, this study's primary goal is not to identify volatility's drivers. Such a task is more qualitative in nature and is the subject of an entirely different study.

This research is focused on understanding cryptocurrency pricing by evaluating the predictability of cryptocurrency volatility. A machine learning model will be used to forecast cryptocurrency volatility, and the R^2 metric will be used as the primary measure of model accuracy. This study then compares results from the machine learning model with a volatility forecasting model commonly used in modern finance. Finance literature gives estimates of volatility predictability in conventional assets; these estimates are used as a benchmark when analyzing the results of this study.

2 Research Methodology

2.1 Data

Data for this research was obtained from the Poloniex cryptocurrency exchange API. The exchange offers trade data for over 100 cryptocurrencies dating back to the exchange's inception in 2014. Data is available via a REST API accessible through Python. In conventional financial research publications, raw trade data either is not available or is simply too large to permit a practical analysis. In this case, however, cryptocurrency trade volume on Poloniex is low enough relative to that of stock exchanges to allow one to calculate price parameters with high precision. For example, many research publications use daily close data for stock price analysis, which omits large amounts of intra-day information that is exposed in day-to-day trading. Possession of raw trade data allows one to calculate returns, volatility, volume, and many more price metrics of cryptocurrencies over arbitrarily small intervals.

Another important caveat is that Poloniex does not offer trading with US Dollar (USD) pairs. Instead, it supports a cryptocurrency called USD Tether (USDT), which is a token tied to the US dollar. This lets the exchange refrain from interacting with users' bank accounts. Because the exchange rate between USD and USDT is 1:1, the USDT markets on Poloniex provide valid price data with respect to USD. Note that in cryptocurrency markets Bitcoin (BTC) is used as the base pair for nearly all cryptocurrencies. Thus, whenever we present cryptocurrency prices or returns in USD, these prices are actually calculated by multiplying the price with respect to BTC by the last traded price for the BTC/USDT pair on Poloniex. In general, cryptocurrencies tend to vary less percent-wise with respect to USD than BTC.

2.2 Machine Learning

2.2.1 Features

Features are independent variables that are used to predict a dependent variable. Future volatility is the dependent variable in this study. Ideally, features present a comprehensive view of the data to the machine learning model; they should contain information that is useful for predicting the dependent variable. Feature selection is critical to the performance of machine learning models. A supervised machine learning model is trained on the data set consisting of features for every data point. The model then makes predictions based on a separate test data set, in which the model's predicted values are regressed with the actual dependent variables. The features used in this random forest regressor model include various trading metrics and parameters consisting of physics equations applied to price motion:

- *Position*: the Volume-Weighted Average Price ($VWAP$) of the cryptocurrency
- *Volatility*: the standard deviation of past $VWAP$
- *Velocity*: change in $VWAP$ over time
- *Inertia*: $1/Volatility$
- *Momentum*: $Velocity * Inertia$
- *Acceleration*: change in Velocity over time
- *Force*: $Acceleration * Inertia$
- *Kinetic Energy*: $*Inertia * Velocity^2$
- *Power*: change in *Kinetic Energy* over time
- *Volume*: volume (in BTC) traded
- *Volume Change*: change in volume traded over time
- *Number of Trades*: the number of trades in the set period
- *Market Cap*: included to help differentiate between cryptocurrencies. Market cap data is pulled from a REST API provided by `www.coinmarketcap.com`.

2.2.2 Machine Learning Model

The machine learning model chosen for this study is a Random Forest Regressor. Random forests can be used for classification or regression. Regression was chosen for this study because we require continuous (rather than discrete) prediction values. The random forest utilizes a series of decision trees, or "forest," to arrive at a predictive model. The generation and analysis of the individual trees is subject to an algorithm with inputted parameters. Each decision tree is computed using only a subset of the provided features. The number of features to be considered in each tree is determined by an input parameter. Features for each tree are selected at random, without replacement. This process is called "bagging," which reduces overfitting in the model. [4] Each tree calculates the best split using the selected features for that tree. Predictions generated by each tree are then averaged. Random forests are popular because they are powerful predictors, are resistant to overfitting even with large numbers of features, and are relatively easy to tune.

2.2.3 Model Parameters

Tuning random forest parameters is relatively straightforward and can be done using simple grid search methods. In general, the parameters that most affect random forest performance are the number of estimators and the maximum number of features to select for each tree.

In this study, model performance plateaued with 300 estimators. When running the random forest regressor model with more than 300 estimators, R^2 and mean squared error did not improve and training runtime was greater. The optimal maximum number of features was found to be about a third of the features for this study. This metric is relatively standard for regression based models; for classifiers the standard maximum number of features is the square root of the number of features in the model.

3 Return Characteristics

3.0.1 Bitcoin Return Distribution

A study that investigates the return characteristics of Bitcoin yields surprising results. It finds that, up to May of 2014, Bitcoin had expressed annualized log returns of 517% and volatility of 86%. [5] These metrics are extremely high and likely unsustainable moving forward. In addition, Bitcoin returns exhibited -1.35 skew and 16.51 kurtosis, implying that large downward moves are more likely than large upward moves, and that extreme returns are much more likely than the standard normal distribution describes. These findings are not too surprising given Bitcoin's massive crashes due to hacks of the Mt. Gox exchange. The fact that Bitcoin exhibits negative skew is rather interesting because this property is the reverse that one would expect of a commodity. Negative skew could imply that investors are subject to "panic selling," which is a case that could be

Cryptocurrency	Mean Return	Volatility	Skew	Kurtosis
Bitcoin	0.003351	0.025002	-0.3741748603538985	5.414141
Monero	0.007105	0.087155	4.055324909821741	31.966679
Ethereum	0.000480	0.058248	0.24324664223346534	4.712194
Factom	-0.000061	0.074331	0.8868238690632719	5.471923
Ripple	-0.002720	0.047665	3.7902486435655676	32.924243
Dash	0.004747	0.053010	2.0601163940736487	16.288484
MAID	0.000164	0.063889	0.8342783098904394	6.347442
Litecoin	-0.002197	0.021105	1.1383151879943874	7.325364
Bitshares	-0.002275	0.053899	2.2061968807149324	12.335859
Stellar	-0.001742	0.057826	3.8579385402108577	31.353760
Dogecoin	-0.002573	0.033436	1.4044512328087584	8.051820
Clams	-0.000505	0.060176	1.822810314736624	18.097290

Figure 1: Descriptive statistics for daily cryptocurrency returns

made for indexes like the SP 500. All in all, these return characteristics suggest that Bitcoin (and potentially other cryptocurrencies) lies somewhere between an index and a commodity in terms of price behavior.

3.0.2 Historical Cryptocurrency Return Statistics

The following statistics are for cryptocurrencies available for margin trading on Poloniex computed using daily returns over the past year. These cryptocurrencies generally have the most daily volume, although cryptocurrency volume can fluctuate significantly from one day to the next. It is important to note that some cryptocurrencies have experienced considerable growth over the past year; this growth is likely unsustainable in the long term, meaning that these return characteristics may change as their underlying assets mature.

One important note is that the kurtosis for every cryptocurrency listed above is greater than 3. This implies that each of these cryptocurrencies presents fatter tails than that of the normal distribution; extreme returns are more likely than the normal distribution predicts.

3.0.3 Conventional Asset Return Statistics

Here, we compute the same return statistics over the past year's daily returns for conventional assets to give a perspective for the return characteristics of cryptocurrencies.

Two important metrics to note are lower volatilities and lower kurtosis for these assets. Conventional assets do not present as many extreme returns as cryptocurrencies.

Asset	Mean Return	Volatility	Skew	Kurtosis
S&P 500	0.000702	0.006398	-0.6591100372675858	5.047792
NASDAQ	0.000895	0.007758	-0.6014286499758614	3.569517
Gold	-0.000106	0.009161	0.41321283205133413	3.512763
Silver	0.000597	0.014931	-0.2905229221430751	2.214206
Crude Oil	0.000485	0.009622	-0.7918932322867047	6.417185
Apple Stock	0.001410	0.012518	0.3665756163661318	6.575413
Google Stock	0.000749	0.010283	-0.7329942548172184	3.546139

Figure 2: Descriptive statistics for conventional asset daily returns (from Yahoo Finance)

3.1 Volatility Prediction Methods in Modern Finance

3.2 Volatility Predictability

A study of the volatility in major markets asserts that about 50 percent of volatility is predictable up to a four-week horizon. [6] This theory holds for smaller equity markets as well, but these markets are still considerably larger than that of cryptocurrencies (to give a sense of scale, the Toronto Stock Exchange has a market capitalization of $2.8 trillion; today's cryptocurrency market cap is roughly $24 billion). It should also be noted that equities are fundamentally easier to price than cryptocurrencies, which may affect the predictability of cryptocurrency volatility relative to that of equities.

3.3 Volatility Models Used in Modern Finance

There are several volatility forecasting methods utilized in modern finance. The most effective, unsurprisingly, is utilizing implied volatility. Implied volatility is financial markets' expectation of volatility and can be backed out from options prices using the Black-Scholes pricing model. No market for cryptocurrency options exists at the moment- therefore implied volatility is not an option here. The next best methods are classified as historical volatility models and Autoregressive Conditional Heteroskedasticity (ARCH) models. [6] These models are roughly equal in effectiveness. Historical volatility models include a wide variety of models, including random walk historical averages of returns, moving averages, exponential weights, and autoregressive models. ARCH models also host a wide family of models, including Generalized ARCH (GARCH), NGARCH, TGARCH, IGARCH, LGARCH, EGARCH, and GJR-GARCH. It is generally accepted that GARCH outperforms ARCH in terms of forecasting volatility of conventional financial assets.

Another study compares the forecasting capabilities of the ARCH and GARCH models along with others. The research uses these models to predict volatility of Nordic equities, using roughly 30% of the price data as a test set. Although determining relative performance in predicting volatility in the test set was inconclusive, it was clear that the GARCH model outperformed the ARCH model. [7] In some cases, EGARCH outperformed the GARCH model as well. A major conclusion asserts that that good in-sample fits do not necessarily produce accurate out-of-sample forecasts. These results could be due to overfitting in some of these models or changes in the underlying dynamics of Nordic equity volatility over time.

3.3.1 Volatility Model Performance

A comprehensive analysis of 66 volatility studies provides benchmarks for volatility forecasting accuracy. [6] Many of these studies incorporate the GARCH model (or its variants), but do not have standardized metrics. For instance, some studies attempt to forecast volatility using daily returns, while others use returns gathered in 5-minute intervals. The R^2 metric for these models increases monotonically with decreasing sample intervals. Actual R^2 values for the GARCH models vary significantly due to sample size, return intervals, underlying financial assets, and forecasting period. R^2 for day-ahead forecasts using daily generated returns is generally on the order of .1; one study predicting day-ahead volatility achieved an R^2 of .5 using Deutsche Mark/USD and Yuan/USD pairs and the GARCH(1,1) model. Results provided from these studies yield reference points for any volatility predictions to be made for cryptocurrencies.

3.3.2 The GARCH Model

The Generalized ARCH model is an extension of the ARCH model. The "AR" in ARCH stands for Autoregressive, meaning that the ARCH model is autoregressive in squared returns. The model is conditional in that the next forecasted volatility is conditional on data present in the current period. Heteroskedasticity is part of the model in that it assumes that volatility is not constant. [8] If the return on a cryptocurrency is denoted as:

$$r_t = \alpha + \sigma_t \epsilon_t$$

where a is the mean return on the cryptocurrency, t is the standard deviation of returns in the same period, and t is the error for that period, then the residual return can be represented as:

$$R_t = \sigma_t \epsilon_t$$

The ARCH model then determines forecasted volatility by:

$$\sigma_t^2 = \omega + \alpha R_{t-1}^2$$

where and are model parameters, $\omega > 0$ and $0 \leq \alpha < 1$. Squared returns are used to ensure that predicted volatility is positive. We can see that the last period's residual return Rt-1 is proportional to the next period's forecasted volatility, with the constant of proportionality being . A high indicates a heavy weighting on the last period's residual return (high conditionality).

As we can see, the ARCH model only assumes future volatility conditionality on past residual returns. We could try incorporating the last period's volatility in addition to residual returns in order to feed more information to the model. This is the aim of the Generalized ARCH (GARCH) model. The GARCH(1,1) model adds an extra term to the forecasted volatility equation of the ARCH model:

$$\sigma_t^2 = \omega + \alpha R_{t-1}^2 + \beta \sigma_{t-1}^2$$

where ω, α, and β are model parameters, $\omega > 0$, $\alpha > 0$, $\beta > 0$, and $\alpha + \beta < 1$. Now, the model incorporates information from the previous period's residual returns and volatility. Note that this GARCH(1,1) model only forecasts volatility for the next period; to achieve further forecasts, one can use the model recursively to generate forecasts for the next n periods. The GARCH(1,1) model optimizes these parameters for maximum likelihood and forecasts using these optimized parameters. [8]

4 Results

4.1 Model

4.1.1 Model Setup

The final volatility prediction model is designed to predict cryptocurrency volatility for the following periods: 4 hours, 12 hours, 1 day, and 1 week. We would expect prediction accuracy to decrease as we predict further out, as postulated in the literature. In addition, the predictive power of our model is more useful if it can predict volatility further out into the future. Feature calculation periods were chosen to be equal to that of forecasted volatility periods.

4.2 Model Accuracy

To arrive at the final random forest regressor model 120 days of raw cryptocurrency trade data was pulled from the Poloniex REST API (ideally, more data would be used for training and testing the model, but this was the most that our systems were able to handle given RAM limitations). 70% of the data was allocated towards training, and the remainder towards testing. Separate models were developed for predicting future volatility with periods of 4 hours, 12 hours, 1 day, and 1 week. We expect the usefulness of predictions to increase as prediction period increases, while accuracy may decrease as prediction period increases. The R^2 values for each of the predicted volatility periods are displayed in Figure 3.

The random forest model appears to predict future cryptocurrency volatility very well. These R^2 values are considerably higher than the highest of those

Volatility Period (s)	R^2
14400	0.634
43200	0.681
86400	0.679
604800	0.604

Figure 3: R^2 for various volatility periods

provided in the literature, which lie around 0.5. It is possible, however, that this high accuracy is simply due to an increased data collection frequency relative to those used in conventional finance studies. We must compare these metrics to that of a benchmark, which in this case is the GARCH(1,1) model. Analyzing accuracy of a random forest with a limited subset of the original features will also give insight as to how additional information affects volatility predictability. As expected, the lowest R^2 value corresponds to the longest volatility period. This result makes sense because volatility predictability should theoretically decrease as the prediction horizon increases.

4.3 Plots

4.3.1 Scatter Plots

Figures 4-7 illustrate the accuracy of the model by plotting predicted volatility versus actual volatility for the selected volatility periods.

4.4 Feature Analysis

4.4.1 Gini Importance

Gini importance is the primary metric we will use to evaluate feature importance in this model. To compute Gini importance, a criterion called the Gini index is tracked throughout iterations of model training. When a feature is used to split a node, the descendent nodes' Gini index values decrease. The Gini importance is determined by the sum of Gini decreases invoked by a given feature. [4] In theory, a higher Gini importance for a given feature implies that the feature contributes more towards model accuracy. This measure gives insight as to what variables are most useful for predicting the volatility of cryptocurrencies.

4.4.2 Random Forest Feature Importances

The top ten most important features in the random forest regressor model for 1-week future volatility predictions are listed below in order of Gini importance:

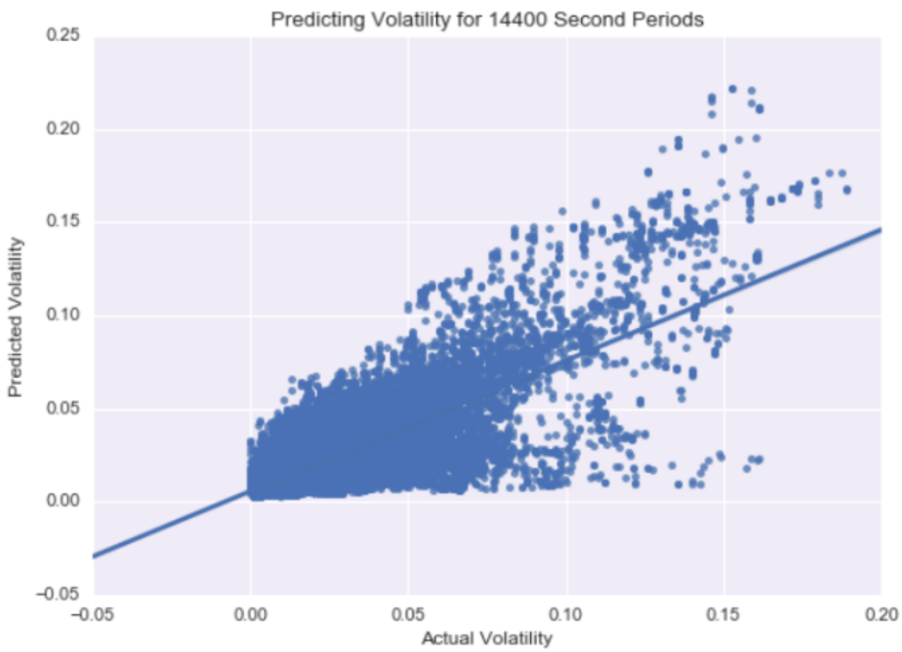

Figure 4: Predicting cryptocurrency volatility over 4 hour periods

1. Volume Change [1-week] (Gini importance = 0.177211)
2. Kinetic Energy [1-week] (Gini importance = 0.133530)
3. Volatility [1-week] (Gini importance = 0.071513)
4. Volume Change [4-day] (Gini importance = 0.060003)
5. Inertia [1-week] (Gini importance = 0.057734)
6. Market Capitalization (Gini importance = 0.051718)
7. Number of Trades [1-week] (Gini importance = 0.049895)
8. Momentum [1-week] (Gini importance = 0.049229)
9. Volatility [1-day] (Gini importance = 0.043205)
10. Power [1-week] (Gini importance = 0.037394)

Most of these feature importances make sense from an intuitive standpoint. In the spirit of the GARCH model, we can expect future volatility to be conditional on past volatility to some extent. Thus, it is no surprise that many of the features that contribute most to model accuracy include past volatility information (namely Kinetic Energy, Volatility, Inertia, Momentum, and Power). One

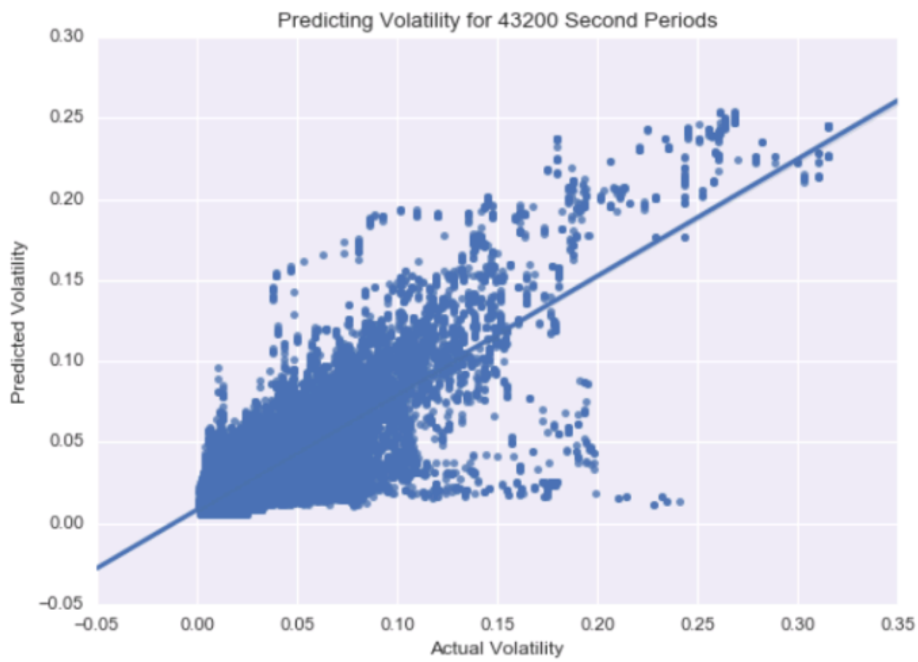

Figure 5: Predicting cryptocurrency volatility over 12 hour periods

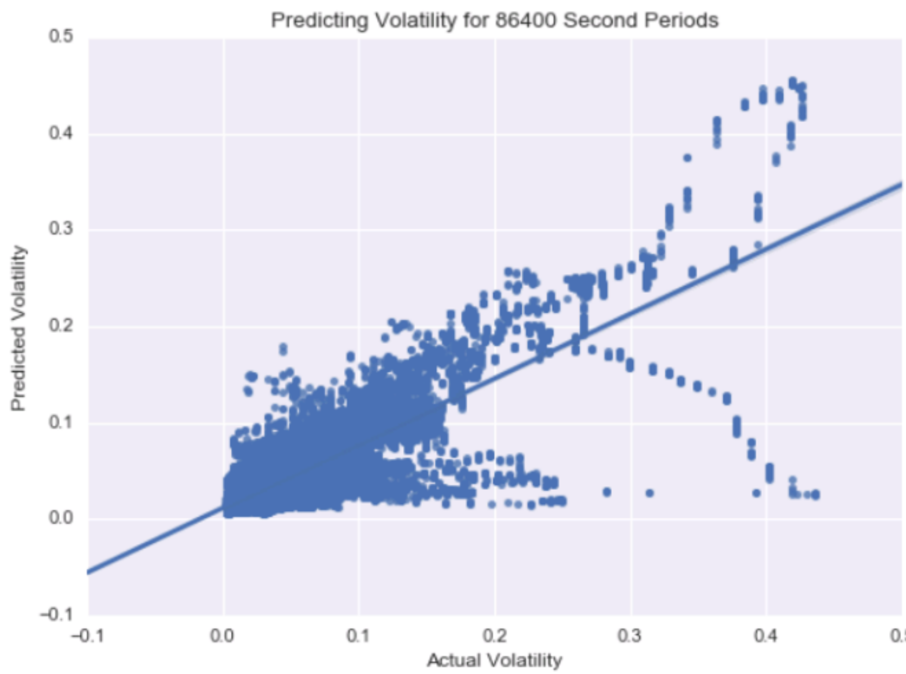

Figure 6: Predicting cryptocurrency volatility over 1 day periods

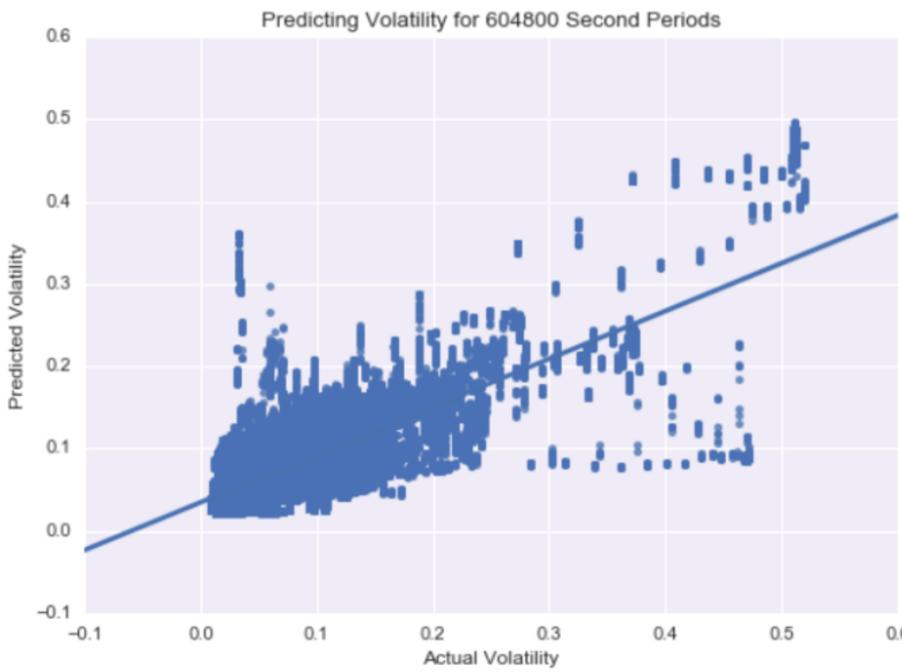

Figure 7: Predicting cryptocurrency volatility over 1 week periods

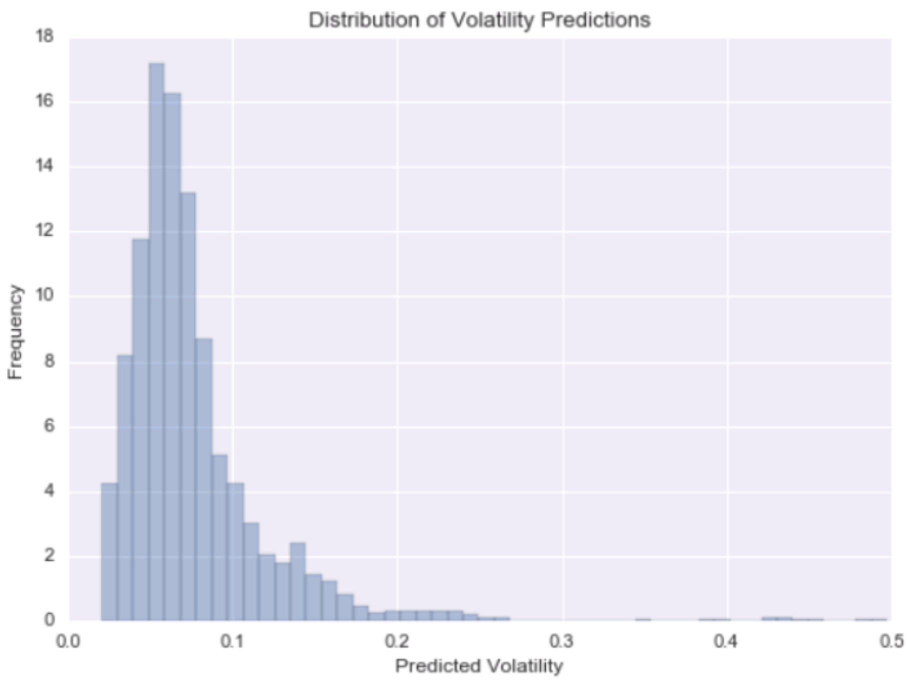

Figure 8: Histogram of 1 week volatility predictions

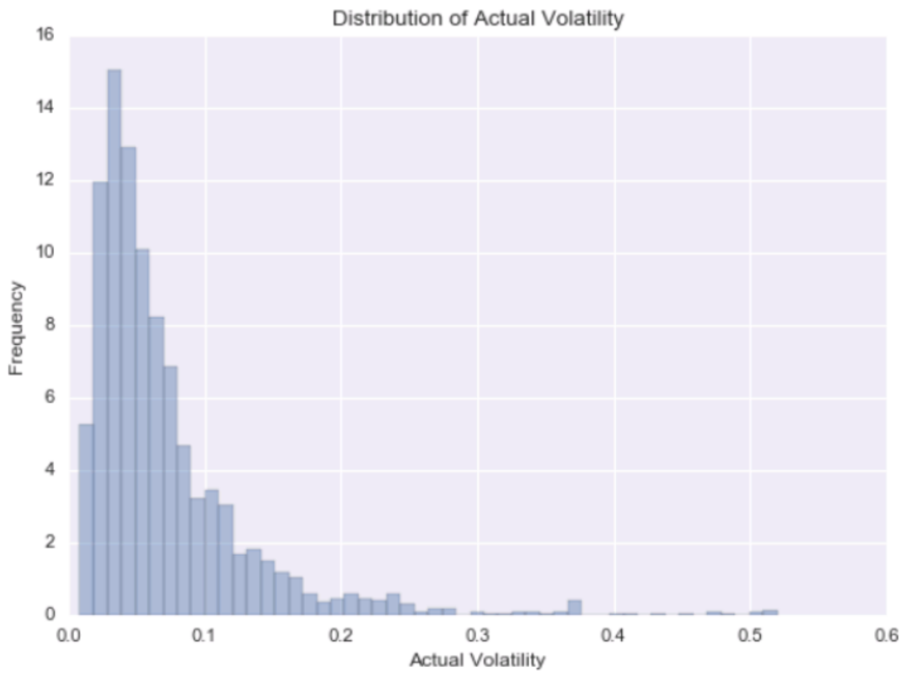

Figure 9: Histogram of 1 week actual volatility

Volatility Period (s)	R^2
14400	0.560
43200	0.480
86400	0.306
604800	0.122

Figure 10: R^2 values for random forest with only returns and volatility features

notable result is the extremely large importance of volume change in this model, which includes information that the GARCH model does not utilize. Finally, it is important to note that these results were produced using a sample size of only 120 days. Utilizing greater amounts of data could remove noise in feature importances and reveal more accurate information about which variables contribute most to future volatility in cryptocurrencies.

In order to directly gauge features' contribution to accuracy, the same random forest regressor was run on two additional sets of data. The first included only returns and volatility as features. This model has access to the same information as the GARCH model- running this model allows us to compare methodology effectiveness. Accuracy metrics for this random forest model are listed in Figure 10:

It is not surprising that all of the R^2 measures are lower for this volatility forecasting model, as the model has access to less information. Nevertheless, the random forest is still able to attain strong R^2 values, especially for shorter time frames. Model accuracy drops off quickly as the prediction horizon increases; this result suggests that returns and volatility are effective predictors in the short term, but other metrics have greater prediction power for longer time horizons.

The next iteration of the model was trained on features that contained extra information about the cryptocurrencies, such as volume, volume change, number of trades, and market capitalization, in addition to past returns and volatility. Omitted features are the physics-based features that are essentially transformations of returns and volatility information. The aim of this model was to quantify the predictive power of including extra information as features, and the results are listed in Figure 11.

The R^2 values for this model show that extra cryptocurrency trading information improves the model significantly, as R^2 increased by about 73% on average. Conversely, R^2 values for this model were slightly worse than the model that had access to transformations of the returns and volatility features as the R^2 was about 5% lower than the original model on average. This model shows that including extra information about cryptocurrencies can increase forecasting ability dramatically, while including transformations of existing features can improve performance marginally. Additionally, we can observe that the R^2 values

Volatility Period (s)	R^2
14400	0.624
43200	0.646
86400	0.663
604800	0.552

Figure 11: R^2 values for random forest without physics-based features

	14400s	43200s	86400s	604800s
ω	2.62E-05	2.60E-05	2.42E-05	3.10E-04
α	0.573	0.57	0.552	0.821
β	0.42	0.43	0.448	0.108

Figure 12: GARCH model parameters for different volatility periods

for this model drop off much more slowly relative to that of the model with only returns and volatility as features. This result suggests that the extra information included in this model exhibits considerable predictive power for long-term volatility.

4.5 Comparison to the GARCH Model

4.5.1 GARCH Implementation

The GARCH(1,1) model was implemented using the GARCH package for MATLAB, using the same amount of train and test data (70% train and 120 days of overall trade data). The parameters for the GARCH(1,1) model are shown below for the corresponding volatility periods. These parameters are listed in Figure 12.

For example, these parameters determine that MATLAB's mathematical GARCH(1,1) model for a volatility period of 4 hours is $\sigma_t = .000026 + .57R_{t-1}^2 + .42\sigma_{t-1}^2$. We see that predicted volatility is heavily weighted towards both the previous period's residual return and volatility, which roughly matches findings from the random forest regressor model implemented earlier. It is important to note that large values imply dependence on the previous period's squared mean return, which is the basis for the ARCH model.

4.5.2 GARCH Accuracy

Below are accuracy metrics comparing the GARCH models' predicted volatility with actual volatility. All of the GARCH(1,1) models seem to exhibit similar predictive powers for different volatility periods, and are shown in Figure 13.

Volatility Period (s)	R^2
14400	0.524176
43200	0.532900
86400	0.519841
604800	0.490000

Figure 13: R^2 values for the GARCH model

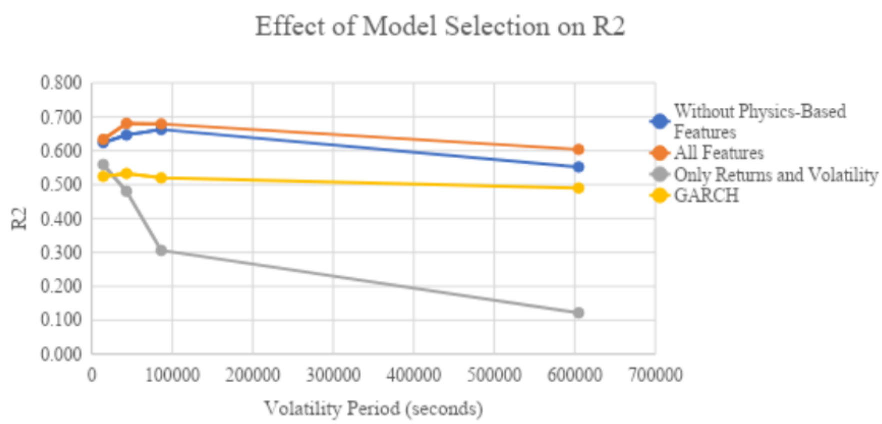

Figure 14: Comparing R^2 values for different models

We can observe that the GARCH model also performs well when predicting future cryptocurrency volatility. These R^2 values are above average relative to those achieved in conventional finance. One possible explanation for this result is the fact that these GARCH(1,1) models are using raw trade data; literature suggests that using more frequent intra-day returns improves the predictability of models significantly. The other explanation would be that cryptocurrency volatility could simply be more dependent on past returns and volatility than the volatilities of other financial assets. One interesting result is that the predictive power of the GARCH model appears to decrease very slowly as the volatility period grows from one day to one week. This behavior is more pronounced in the random forest model.

4.5.3 Comparing Model Accuracy

Using these three methods of feature selection, we can observe the effect of feature selection on random forest performance.

Figure 14 shows the remarkable increase in accuracy that volume, trade

Model	*Relative Runtime*	*R^2 (period=1 week)*	*R^2 / Relative Runtime*
RF (only returns, volatility)	1.00	0.12	0.12
GARCH	1.14	0.49	0.43
RF (without physics features)	4.04	0.55	0.14
RF (all features)	5.07	0.60	0.12

Figure 15: Runtime analysis for different models

frequency, and market cap provide if these factors are included as features. Including the physics-based features provides a much smaller performance boost, likely because these features do not incorporate additional outside information (they are transforms of the returns and volatility features). The random forest models with the extra features clearly outperform the GARCH model. When the random forest has access to the same information as the GARCH model, however, the GARCH model provides superior performance for 3 out of 4 volatility periods. This random forest's R^2 drops off rapidly as volatility period increases. This characteristic is less pronounced when adding in the other features, which suggests that this extra information exhibits greater predictive power for longer volatility periods. Although the GARCH model does not perform as well as the random forest models, it generates the least performance decay with increasing volatility period out of all of the random forest models, which could explain why it is so prominent in financial research.

4.5.4 Comparing Model Runtime

It is clear that the random forest model trained on all of the discussed features presents the highest R^2 value when compared to models trained on less features. This performance boost comes with a feature calculation runtime cost, however. In general, the feature calculation runtime increases as the size of the training set increases, but the percent runtime for each feature stays relatively constant. Comparing the runtimes shows how the marginal performance boost per second of runtime changes as features are added to the model.

Figure 15 shows that the GARCH model is by far the most efficient in terms of runtime. Including the GARCH model only increases runtime by 14% and dramatically increases R^2. Interestingly, R^2 per unit runtime actually stays relatively constant within the random forest models. All in all, if one desires a short runtime, the GARCH model is the best volatility forecasting option. If one disregards runtime, the random forest with all features is the best predictor.

5 Discussion

The most obvious conclusion from this research is that cryptocurrency volatility is quite predictable, at least in the short term. We are able to predict cryptocurrency volatility computed over a 1-week period with an R^2 of .60, which means that 60% of the variation in future volatility is explained by the features of the model. Both the random forest model and the GARCH model heavily weight past volatility in predicting future cryptocurrency volatility. Past volatility is a relatively strong predictor of future volatility in other financial assets as well. The random forest model requires significantly more run time and computation than MATLAB's GARCH(1,1) model implementation; this model, however, outperforms the GARCH model by about 26% on average in terms of R^2. Calculations of predictions for the GARCH model have a very short runtime due to the simplicity of the GARCH model, which entails multiplying return and volatility values by parameters and summing them. Feature calculation for the random forest has a longer runtime due to the increased numbers of features and numbers of computations on those features.

Some conclusions from this study reveal information about the nature of the random forest machine learning model. The fact that the random forest R^2 per runtime ratios stay relatively constant as features are added suggests that feeding more features to the random forest can improve model performance in an efficient way. There is a limit to model accuracy, however; eventually the accuracy per runtime will decrease as new features essentially become noise to the model. In addition, the fact that random forest R^2 values improved when adding in the physics-based features shows that transforms on inputs can present new information to the model. The random forest algorithm is not sophisticated enough to permute through transformations of the given features. In this case, the physics-based features provided sufficient information and added on such little runtime that R^2 per runtime did not change significantly.

In addition to finding that cryptocurrency volatility is both predictable and dependent on past volatility, one interesting observation is that the Volume Change feature exhibited the highest Gini importance value in the case of 1-week volatility forecasting. This result implies that past changes in volume can help explain future volatility in cryptocurrencies. Such a relationship could have significant implications for risk-averse investors when evaluating the metrics of cryptocurrencies. While these findings are potentially useful, they should be investigated further. One factor to consider is that a larger train and larger test set would make the random forest regressor reflect more of the truth behind future cryptocurrency volatility; including more data could produce a different Gini importance distribution among the known features. Secondly, the cryptocurrency space is still growing. It is possible that market conditions will change significantly in the coming years, which could potentially alter any relationships discovered in this study.

Finally, the accuracy of a machine learning model that predicts future cryptocurrency volatility is limited by the extent to which the prices of cryptocurrencies are driven by external events. Previous price motion cannot predict major price moves such as the Bitcoin crash after the Mt. Gox hack. [2] These sources of volatility are external events that cannot be predicted accurately from a technical standpoint. Nevertheless, a machine learning model with more features and a robust feature selection procedure could produce better accuracy than the models discussed in this study. The random forest model with all of the presented features provides a lower bound for the predictability of cryptocurrency volatility. These results suggest that cryptocurrency volatility is at least as forecastable as assets in conventional finance. All in all, this study has shown that investors can forecast cryptocurrency volatility with considerable accuracy given that they have access to the same intra-day trade data.

6 Future Research

The cryptocurrency sphere is continuing to grow as more people learn about blockchain technology and more developers commit to either inventing new cryptocurrencies or working on present ones. At this moment in time, not much literature exists regarding cryptocurrency analysis. Cryptocurrencies are still extremely young, and there are many avenues for research. One example of future research would be forecasting cryptocurrency volatility by using a greater span of raw trade data with a sophisticated sampling technique. Training sets for models in this study were limited due to the sheer amount of raw trade data that they contained.

As the literature suggests, the most powerful predictor of future volatility for cryptocurrencies would likely be implied volatility. Currently, no options market for cryptocurrencies exists. An options market for cryptocurrencies may develop in the future, however. Options can be coded into smart contracts, a term many coin as "programmable money." Developers within the cryptocurrency community have discussed the possibility of an options market entrenched in Ethereum's smart contracts; if this idea comes to fruition, there would be much more reliable information to be used in predicting cryptocurrency volatility. In addition, an options market for cryptocurrencies could further increase cryptocurrencies' value as hedging assets.

Another potential avenue of research is predicting kurtosis for cryptocurrencies. One large difference between cryptocurrency and conventional asset return characteristics is that cryptocurrencies generally exhibit very high kurtosis. A model that could predict kurtosis well could provide value to investors by indicating when their investments could be subject to extreme moves in either direction. Predicting skew would be a much more difficult task because this metric is directional and is more likely to be priced into the cryptocurrency markets. Forecasting these characteristics of return distributions could provide value to investors that desire risk-optimized cryptocurrency portfolios.

One final idea regarding this research relates to portfolio optimization. Much of the research presented in this study could be used to construct a portfolio of cryptocurrencies for the risk-averse investor. Modern portfolio theory suggests that not only expected volatility and returns are important; correlations between assets provide additional information that can be used to reduce idiosyncratic risk. Forecasted volatility could be used in conjunction with historical (or forecasted) correlations with other cryptocurrencies. This covariance matrix could be used in a Markowitz optimization to provide portfolios with maximized Sharpe ratios for investors in cryptocurrencies.

In conclusion, cryptocurrencies are the beginning of a new digital asset class; as this asset class grows, understanding the price motions of cryptocurrencies becomes both useful and important. The considerable volatility of cryptocurrencies has been one of the biggest criticisms of this asset class, and this study provided a foundation for understanding this phenomenon.

6.1 Acknowledgements

I would like to thank Mina Lee and Li Yang for their guidance throughout the production of this research. I would also like to thank Washington University in St. Louis for providing the resources necessary to perform this study.

References

[1] D. Baur, K. Hong, A. Lee, Kuhne Logistics University (2015).
Bitcoin: Currency or Asset?
`https://mbs.edu/getattachment/fircg/FIRCG-2016/Papers/8-Adrian -2c-KiHoon Bitcoin-Baur-et-al-2015-P.pdf`

[2] J. Bouoiyour, R. Selmi, A. Tiwari, O. Olayeni. Economics Bulletin (2016).
What drives Bitcoin price?
`http://www.accessecon.com/Pubs/EB/2016/Volume36/EB-16-V36-I2-P82.pdf`

[3] N. Stråle Johansson, M. Tjernström. Umeå School of Business and Economics (2014).
The Price Volatility of Bitcoin.
`http://www.diva-portal.org/smash/get/diva2:782588/FULLTEXT01.pdf`

[4] G. Louppe, L. Wehenkel, A. Sutera, P. Geurts. University of Liege.
Understanding variable importances in forests of randomized trees.
`https://papers.nips.cc/paper/4928-understanding-variable-impor tances -in-forests-of-randomized-trees.pdf`

[5] A. Goyal. Anderson Graduate School of Management, UCLA (2000).
Predictability of Stock Return Volatility from GARCH Models.
`http://www.hec.unil.ch/agoyal/docs/Garch.pdf`

[6] S. Poon, C. Granger. Journal of Economic Literature (2003).
Forecasting Volatility in Financial Markets: A Review.
`https://faculty.washington.edu/ezivot/econ589/PoonGrangerJELsurvey.pdf`

[7] A. Wennström. Royal Institute of Technology (KTH) (2014).

Volatility Forecasting Performance: Evaluation of GARCH type volatility models on Nordic equity indices.

https://www.math.kth.se/matstat/seminarier/reports/M-exjobb14/140616b.pdf

[8] R. Reider. New York University (2009).

Volatility Forecasting I: GARCH Models.

http://cims.nyu.edu/ almgren/timeseries/Vol$_{F}orecast1.pdf$