Cryptovisor: A Cryptocurrency Advisor Tool

Matthew Baldree, Paul Widhalm, Brandon Hill,

Matteo Ortisi

{mbaldree, pwidhalm, bdhill}@smu.edu,

matteo.ortisi@gmail.com

**Abstract.** In this paper, we present a tool that provides trading recommendations for cryptocurrency using a stochastic gradient boost classifier trained from a model labeled by technical indicators. The cryptocurrency market is volatile due to its infancy and limited size making it difficult for investors to know when to enter, exit, or stay in the market. Therefore, a tool is needed to provide investment recommendations for investors. We developed such a tool to support one cryptocurrency, Bitcoin, based on its historical price and volume data to recommend a trading decision for current or past days. This tool is 95.50% accurate with a standard deviation of 0.54%. From our analysis, we conclude that Bitcoin is a unique asset with similarities to gold. As a young asset, it lacks economic fundamentals making it very difficult to predict. By leveraging technical momentum indicators to provide buy, sell, and hold indicators a tool can be developed that performs better than a buy and hold trading strategy in a bear market.

# 1 Introduction

The first cryptocurrency, Bitcoin, based on blockchain technology, is a secure method to make financial transactions that was introduced in 2009 [14]. Bitcoin is a peer-to-peer cryptographic digital currency that was created in 2009 by an unknown person using the alias Satoshi Nakamoto [14]. Bitcoin is unregulated and hence comes withbenefits (and potentially a lot of issues) such as frictionless transactions and pseudo anonymity. Transactions can be purchased through exchanges or can be ‘mined’ by computing/solving complex cryptographic/mathematical puzzles [1].

Since 2009, the cryptocurrency market was dominated by Bitcoin with 100% market share. In 2011, Litecoin was created as a fork[[1]](#footnote-1) of Bitcoin. In 2015, Vitalik Buterin co-founded Ethereum to a blockchain technology alternative to Bitcoin. Today, there are over 1,500 cryptocurrencies traded on 5,400 plus exchanges throughout the world. Collectively, this represents a total market capitalization of $311 billion and a 46-fold ($821B/$18B) increase at its all-time high (ATH) [2]. In 2017, Bitcoin reached its ATH at $19,536 representing a 20-fold ($19,536/$973) price increase. As a comparison, the S&P 500 Index took over forty years to achieve a similar growth. As of March 12, 2018, this explosive growth market has reduced its market cap with a 62% reduction. The stock market has nothing to compare with in terms of this drop as a price correction is signified by a 10% drop and a bear market is signified by a 30% drop [25].

The volatility of this new emergent market makes it impossible to forecast. Factors such as regulatory news, investor hype, exchange volatility, and large liquidity sell offs can greatly affect the price [25]. The recent severe market downturn has revealed there is no market fundamental economic indicators to provide investors with a way to contextualize price movements. For the market to level out, economic fundamentals must be established. In order to better model cryptocurrency, we need to understand how and why it behaves such as it does. One of the main issues with cryptocurrency is that each currency is built somewhat differently. Unlike fiat currency where the units are backed by the government and it has value because the government says it has value, cryptocurrency has value because others who hold the crypto-coin say it has value. Therefore, financial markets are trying to decide whether cryptocurrencies are a hedge, safe haven or a real currency like the dollar or yen [4].

Recently, mainstream financial institutions like Fidelity are allowing its customers to add cryptocurrencies to their portfolios [3]. In addition, global exchanges like CBOW and CME are offering Bitcoin Futures to their investors allowing them to get pricing exposure to Bitcoin without holding it. More cryptocurrency financial vehicles are on the way in 2018 that will provide more liquidity and help establish market fundamentals. In the meantime, the new investor is left with little tools outside of price time series to make an informed decision.

Investors that do decide to enter this emerging market, a common question asked on social media platforms is, “When is a good time for me to buy in the market?” Or, the opposite question will be asked, “When is a good time for me to sell out of the market?” To aid the new investor in making an informed decision and answer these questions, we created a novel cryptocurrency tool, *Cryptovisor*, that recommends for the investor if he should buy, sell, or hold their investment. This tool is built with financial technical momentum indicators to label historical cryptocurrency data for optimum time to buy, sell, or hold the asset. Labeling the dataset is a trial and error procedure of adjusting the indicators so the results are as good or better than a buy and hold trading strategy. With the labeled dataset, a stochastic gradient boost classifier was used to learn the labeled data and provide current or past event recommendations. The baseline classifier was then iterated reducing down the features from 18 to a minimal set of two that resulted in an accuracy improvement of 0.6% with 0.54% standard deviation.

The advisor tool trained with trading signals identified by the defined algorithm against a bear market dataset resulted in a trading performance better than a buy and hold strategy. Extending this tool with multiple models for various cryptocurrencies will allow an investor to surface new opportunities that they might otherwise ignore or miss. This extension is left as a future exercise. The rest of the paper is structured as follows. Section 2 provides background of cryptocurrency. Section 3 provides details about the cryptocurrency market and provides a visualization of the market structure for Bitcoin. Section 4 provides background of sentiment analysis. Section 5 describes our Cryptovisor tool. Section 6 details the data used in the research. Section 7 provides the results of our research. Section 8 analyzes the results. Section 9 discusses the ethics. We draw the relevant conclusions in Section 10 and discuss future work in Section 10.

# 2 Cryptocurrency

Cryptocurrency is an asset designed to be used as digital currency. The asset is a chain of digital signatures that exist in binary format secured by cryptography with the right to use [18]. There are two kinds of assets in the cryptocurrency world; i.e., coin and token. Both assets may be acquired and traded on public exchanges and used in the exchange of goods and services. Coins are more general purpose and require more effort to create and support the ecosystem. Tokens run on top of a platform such as Ethereum and are designed for a specific ecosystem such as eSports. The reference cryptocurrency is Bitcoin, the largest and oldest coin, see Fig. 1 for a timeline that includes Litecoin, Vertcoin, Ethereum, and subsequent forks. There are more than 1,500 cryptocurrencies to date.



**Fig. 1.** Cryptocurrency timeline.

Since the 1990s, numerous cryptographers and companies have repeatedly attempted to create digital currencies to compete with cash and credit. These attempts failed because their solutions were not substantially better than the current state due to several factors. 1) They relied on cryptographic certificates to tie parties to real-life identities. The hassle to acquire, install and use these certificates lowered the utility of these systems and therefore adoption. 2) They wanted merchants to adopt this new technology as opposed to going directly to consumers. Incumbent solutions are entrenched and have money to defend their marketplace. 3) They solved the “double-spend” problem through a third-party clearinghouse adding cost and time to the process. All these systems still required a network of trust. In the end, these systems added innovation but did not improve the whole user experience [13].

In 2008, after almost two years of development, an anonymous person or group under the alias “Satoshi Nakamoto” released a white paper describing the production of a peer-to-peer electronic cash system named Bitcoin [14]. This system allows irreversible, secure, digital payments directly between parties through a non-trusted network built on cryptographic proof-of-work. Transactions are assembled and hashed by bitcoin nodes or miners that use CPU power to solve cryptographic puzzles for coin and or transaction rewards by adding blocks to the ledger creating the blockchain. Incentivizing, miners to be honest, keeps the network secured and introduces coins to the marketplace. Miners may sell or trade their coins for goods and services and as long as there are 51% honest nodes on the network, the system is secure[[2]](#footnote-2).

The bitcoin network broadcasts all new transactions to all nodes on the peer-to-peer network. Nodes may come and go as they please and the active nodes collect new transactions into a block and work to solve a computational puzzle or proof-of-work for its block. Nodes will compete to work on the next block for a reward by extending the longest ledger. When a node solves the puzzle, the node will broadcast its block to all nodes on the network. The block will be accepted if all transactions are valid and are not already spent which prevents double-spending. Today there are many derivatives of Bitcoin but the main principles of a decentralized peer-to-peer digital currency network remain the same, see Fig. 2 [14].



**Fig. 2.** Cryptocurrency workflow.

The process to acquire cryptocurrencies usually entails signing up with an exchange, proving who you are[[3]](#footnote-3) and submitting assets to the exchange to credit your account. The assets could be other cryptocurrencies or fiat currency. After submitting assets, the exchange will credit your account and you may then submit buy and sell orders into the marketplace through market or limit orders. A market order authorizes the exchange to make the trade on your behalf. A limit order places the order in an order book waiting for the order’s criteria to be met. The liquidity of an asset in the exchange can affect the volatility in pricing. Large purchases on exchanges can run up the price from 1% to 10% because the exchange does not have enough liquidity or amount of assets to satisfy the order without bumping up the price. In addition, cryptocurrencies are traded across many exchanges further reducing the liquidity of the asset. Large trades could exhaust an exchange causing a “flash crash” similar to the one experience by GDAX in June of 2017. On that day, a multi-million dollar sell order caused the price of Ethereum to drop 99.9% within a second. The steep drop triggered cascading stop loss orders until the price reached $0.10 within a second. This price volatility is one of the main barriers to cryptocurrency adoption [15].

Exchanges make money through transaction fees that include converting between assets and supporting the inflow and outflow of assets such as the US dollar fiat from trader’s exchange wallets. The trader may move their assets off the exchange to other exchange wallets or even his own local wallet. The exchange of digital assets is performed by digital wallets which is an application that supports storing and sending a digital asset such as bitcoin. Wallets can be for a single asset or support multiple assets like Jaxx. Public and private key pairs for your digital asset are stored in the wallet. The public key is the address you use to send and receive assets. The private key is used to claim your assets on the blockchain. Transferring money or digital assets is a transfer of ownership between parties on the public blockchain ledger. After the transaction is sufficiently confirmed, it is essentially permanent. The major security risk of this system is losing your private keys and thereby losing ownership of your assets.

# 3 Cryptocurrency Market

To determine if Bitcoin can be considered an asset class similar to the world’s government-backed currencies, Bitcoin would need to satisfy three questions. 1) Can it be used as a medium of exchange? 2) Can it be used as a unit of comparability between two goods and 3) it must store value over time [4]. Price fluctuations in Bitcoin and other cryptocurrencies are dependent on both internal and external factors[4]. The internal factors are supply and demand but since the supply is deterministic this means that the only internal driver is the demand for Bitcoin. The demand for Bitcoin is determined by the hash rate. External factors affecting the price of Bitcoin is the adoption rate and how it is being used as an investment vehicle. In the short-term, Bitcoin acts as a safe-haven investment and in the long run acts as a hedge [4].

In an effort to better understand correlation of Bitcoin with other assets, we modeled various foreign exchange rates to determine their influence on the price of crypto-coin. Interestingly, Bitcoin and Litecoin are positively correlated with various fiat currencies (as compared to the US dollar) while Ethereum is slightly negatively correlated with the same currencies. Ethereum is a ‘smart contract’ cryptocurrency while Bitcoin and Litecoin are conventional cryptocurrencies.



*(correlation matrix for change in prices between Jan 2017 and March 2018)*

Also, we review the correlation of the cryptocurrencies with gold and found that the cryptocurrencies are negatively correlated to the price movements in gold.



*(correlation matrix for change in prices between Jan 2017 and March 2018)*

# 4 Sentiment Analysis

There has not been a significant research towards sentiment analysis focused on cryptocurrency market. Instead, we look to a wealth of knowledge gained from the numerous papers focused on sentiment analysis focused on another financial market, the stock market index. “As more and more personal Opinions are made available online, recent research indicates that analysis of online text such as blogs, web pages, and social networks can be useful for predicting different economic trends [17]. As such, sentiment analysis is performed using various data sources/tools such as Twitter, google and yahoo search trends, or message boards/blogs such as Reddit. These studies look for trends from public tools to understand the public sentiment in order to directionally predict the stock market. Similarly, we look to use this public sentiment to predict the cryptocurrency market.

Each of these analyses follows a very similar pattern. First, they seek to gather a consistent data source from the public tool that meets the frequency in which the prediction method requires. These feeds typically come in the form of an API provided by the tool of choice. This makes it somewhat trivial to plug into the tool to capture the transactional data that is provided through the API subscription.

Second, the data retrieved through the APIs is filtered using terms that identify the different exchanges such as names, IPO ticker symbols, or associated businesses. This allows for a much more efficient algorithm process as it only performs further processing on transactions that are relevant to data requirements of the research analysis.

Third, comes the identification of opinion and weighting of each sentiment instance. Now there are various ways to identify and weight each instance, but a common theme is to utilize a lexicon of financial terms to determine whether the opinion/text should be reviewed as positive or negative. This is necessary because as noted, “the Harvard dictionary is not structured for the vocabulary of traders. [16].” Then one can use other attributes provided by the API for the instance such as times searched for search engines such as Google and Yahoo, or the number of followers for the poster for instances gathered from Twitter, Facebook, or bloggers.

Our research focuses on Twitter as our initial source of public opinion based on the amount of sentiment research found for this tool. Many of these analyses found that the opinions provided by the tweeters provided good indicators that could be used by investors trying to determine economic trends. As noted “Twitter has become a major source of information and an effective communication tool for investors and public companies [16].”

# 5 Cryptovisor – A Cryptocurrency Advisory Tool

Machine learning cryptocurrency markets is hard. The markets themselves are inherently unpredictable and lack fundamentals [26]. Search and news hype, follow the heard, liquidity, and large investors play a major role in price fluctuation. This wild fluctuation causes inconsistent data distribution making it near impossible for the train and test datasets to have similar distributions. Other factors come into play when the data is at various scales; e.g., daily, hourly, or minute. Data sampling less than a day is influenced by high frequency trading or algorithmic trading. Opening and closing prices have their own patterns. News and rumors happen at any time and are a driving force to multi-day horizons. In the end, you do not know what will happen tomorrow as no one has a complete picture at any point in time. We have minimal information with a distribution of data that is constantly changing. Therefore, we took a recommender or advisor approach based historical observable buy and sell signals to predict today’s decision into an advisory tool.

Cryptovisor is an advisory or recommendation tool for a cryptocurrency investor to query current or past periods for a buy, sell, or hold position regarding a cryptocurrency. The tool is trained on past historical pricing and volume information and labeled for ideal buy, sell, and hold positions based on an algorithm utilizing both leading and lagging financial technical indicators.

A technical indicator is “any class of metrics whose value is derived from generic price activity in a stock or asset [20].” There are two kinds of technical indicators, leading and lagging, that try to predict the future or general price direction of a security by looking at past patterns. Leading indicators signal future events. Lagging indicators follows an event. The importance of a lagging indicator is its ability to confirm that a pattern is occurring. There are many, many indicators. For this paper, two popular indicators, relative strength indicator (RSI) and Bollinger bands (BB), are used to determine a trading strategy [21]. Through trial and error, the indicators were adjusted to fit the pattern of Bitcoin close price for one-hour resolution. An algorithm was developed to incorporate both indicators to determine a trading strategy of buy, sell, or hold. This strategy was then applied to buy or sell the asset and the result compared to a buy and hold strategy.

The resulting labeled data for one hour closing price trading strategy was then used to train a stochastic gradient boosting machine learning algorithm [22] to predict buy, sell, or hold strategy based on time series closing price and volume plus derived data. In addition, a feature ranking and example decision tree plots are provided for deeper understanding [23].



**Fig. 3.** Cryptovisor system diagram.

The tool, *Cryptovisor*, is comprised of six main components, see Fig. 3. This paper focuses on primarily on three components: *labeler*, *sentimentor*, and *modeler*. The *labeler* component reads in the cryptocurrency data and up samples or down samples it appropriately to a one-hour time period. Then, data is used to generate over 1,970 data points to train a machine learning classifier. Next, Bollinger Bands (BB) are calculated for a 4-hour moving average with a standard deviation of 1.3. BB are volatility lines created from a close price moving average (1) and its standard deviation (2). The bands are defined by (3), (4), and (5). N is the number of days to compute the moving average. Next, Relative Strength Indicator (RSI) is calculated for a 8-hour period. RSI is a momentum oscillator that measures speed and change of price movements between zero and 100, see (6). N is the average days up closes or down closes. Next, lags are calculated for each attribute and the trading algorithms (7) and (8) are applied. Once the trading signals are generated, the returns are generated and compared with a buy and hold strategy to ensure trading signals yield comparable results. A comprehensive chart is generated to allow for results investigation. Finally, the labeled data plus calculated data is generated.

|  |  |
| --- | --- |
| . | (**1**) |
| . | (**2**) |
| . | (**3**) |
| . | (**4**) |
| . | (**5**) |
| . | (**6**) |
| . | (**7**) |
| . | (**8**) |

The modeler component reads in labeled data, cleans it, and separates out the features from the labels. A stochastic gradient boost classifier, XGBoost, is then trained with a stratified 5-fold cross validation to determine accuracy. Next, the importance of the features is plotted for examination. Features are then pruned to yield the simplest model without sacrificing accuracy. The simplified model is then used to provide a trading strategy for past and current time period for the cryptocurrency.

# 6 Data

The historical pricing data for Bitcoin was obtained from CryptoCompare API from October 3, 2016 to February 10, 2018 [19]. This low-level hourly resolution allowed us to resample it for any desired resolution. For this project, the data we used the default one-hour sampling of 82 days yielding 1,986 data points. The raw Bitcoin data in Table 1 is then read by the Labeler component which calculates data in Table 2. Data from table 2 is the read into the Modeler component which fits a model to classify a record as buy, sell, of hold.

**Table 1.** Exchange dataset read into *Labeler* component.

|  |  |
| --- | --- |
| Attribute | Description |
| Timestamp | Data and time of the transaction |
| Open | Open price for the period |
| High | High price for the period |
| Low | Low price for the period |
| Close | Close price for the period |
| Volume BTC | Trade volume of asset for the period |

**Table 2.** Calculated dataset by *Labeler* component.

|  |  |
| --- | --- |
| Attribute | Description |
| Bollinger High | BB mid plus standard deviation |
| Bollinger Mid | BB n-periods moving average |
| Bollinger Low | BB mid minus standard deviation |
| RSI | Relative strength indicator for n-periods |
| Close Lag 1 | Close price for previous period |
| Bollinger High Lag 1 | BB high for previous period |
| Bollinger Mid Lag 1 | BB mid for previous period |
| Bollinger Low Lag 1 | BB low for previous period |
| RSI Lag 1 | RSI for previous period |
| Close Lag 2 | Close price two periods back |
| Bollinger High Lag 2 | BB high two periods back |
| Bollinger Mid Lag 2 | BB mid two periods back |
| Bollinger Low Lag 2 | BB low two periods back |
| RSI Lag2 | RSI two periods back |
| Signal | Buy, sell, or hold indicator |
| Portfolio | Bought or sold asset |
| Trading Return | Trading period return on asset |
| Buy & Hold Return | Buy & sell period return on asset |
| Trading Cum Return | Trading cumulative return on asset |
| Buy & Hold Cum Return | Buy & sell cumulative return on asset |

# 7 Results

With around 2,000 signal points, the Labeler identified half percent of them as sell periods, see Table 3. The Labeler than used the trading signals to build a portfolio and compared its returns with a buy and hold strategy, see Table 4. The difference in Sharpe Ratio return between the two strategies is less than 0.43 indicating the trading strategy is accurately identifying good buy and sell conditions. The algorithm was adjusted by visualizing the trading signals with the historical close price in Fig. 4. The resulting data from the Labeler was then read into the Modeler to learn the signals through a 5-fold cross validation with a resulting accuracy of 94.89% with a standard deviation of 1.95%. The contribution of the features to classifying the label is shown in Fig. 5. Then the features were pruned to a minimal set yielding an accuracy of 95.50% with a standard deviation of 0.54%. Reducing the features down to volume lag1 and RSI lag1 yielded 0.61% accuracy improvement, see Fig. 6 for the feature importance and Table 5 for interpretation of the feature importance ID’s.

**Table 3.** Labeler signal results.

|  |  |  |
| --- | --- | --- |
| Attribute | Total | Percentage |
| Hold | 1,908 | 96.07% |
| Buy | 52 | 2.62% |
| Sell | 26 | 1.31% |

**Table 4.** Labeler portfolio results.

|  |  |  |
| --- | --- | --- |
| Summary | Trade | Buy & Hold |
| Return | 1.18 | 0.007 |
| Standard Deviation | 1.3 | 1.3 |
| Sharpe Ratio[[4]](#footnote-4) (Rf=0%) | 1.26 | 0.83 |



**Fig. 4.** Labeler charts.

**Table 5.** Modeler features.

|  |  |  |
| --- | --- | --- |
| ID1 | ID2 | Feature |
| f0 |  | Close price |
| f1 |  | Volume BTC |
| f2 |  | BB high |
| f3 |  | BB mid |
| f4 |  | BB low |
| f5 |  | RSI |
| f6 |  | Close price prev period |
| f7 | f0 | Volume BTC prev period |
| f8 |  | BB low prev period |
| f9 |  | BB mid prev period |
| f10 |  | BB high prev period |
| f11 |  | Close price two periods back |
| f12 |  | Volume BTC two periods back |
| f13 |  | BB low two periods back |
| f14 |  | BB mid two periods back |
| f15 |  | BB high two periods back |
| f16 | f1 | RSI prev period |
| f17 |  | RSI two periods back |



**Fig. 5.** Modeler feature importance considering all features.



**Fig. 6.** Modeler feature importance with minimal features.



**Fig. 7.** Modeler sample decision tree for minimal features.

# 8 Analysis

Blah blah

## 8.1 Market Analysis

Cryptocurrencies currently behave like a separate asset class, although closer to a hedge such as gold but in their own market space. Market movements of the S&P 500 and NASDSAQ are far removed from being similar with cryptocurrencies. This is also true with fiat currencies, investors do not appear to utilize them in the same manner as cryptocurrency. In comparison, gold moves closely, although negatively correlated, with cryptocurrencies while not directly related. The Nikkei market shows a slight connection to Ethereum which is a smart contract-type cryptocurrency. The Nikkei has a slight relation to the price of Ethereum (ETH-P) and gold’s price movement is somewhat related to cryptocurrencies. Fiat currencies are completely unrelated which would be expected. The correlation between these assets are shown in Fig. 8.

#### Remark 1. BTC-P = Bitcoin price, LTC-P = Litecoin price, ETH-P = Ethereum price, BTC-V = Bitcoin volume, LTC-V = Litecoin volume, and ETH-V = Ethereum volume.



**Fig. 8.** Cryptocurrencies correlation with fiat currencies and market indices.

Fig. 9 shows a zoin view of tIn this view, w



**Fig. 9.** Zoomed in view of the relationship between cryptocurrencies, treasury, gold and crude oil prices.

Fig. 10 focuses on the relationship between fiat currencies and cryptocurrencies. This chart emphasizes the point that investors view cryptocurrency more like gold and not fiat currencies at this point in time. Because gold is used to hedge volatile markets, it would make sense that gold too is not correlated with fiat currencies.



**Fig. 10.** Relationship that fiat currencies have with cryptocurrencies.

Fig. 11 shows a



**Fig. 11.** Modeler sample decision tree for minimal features.

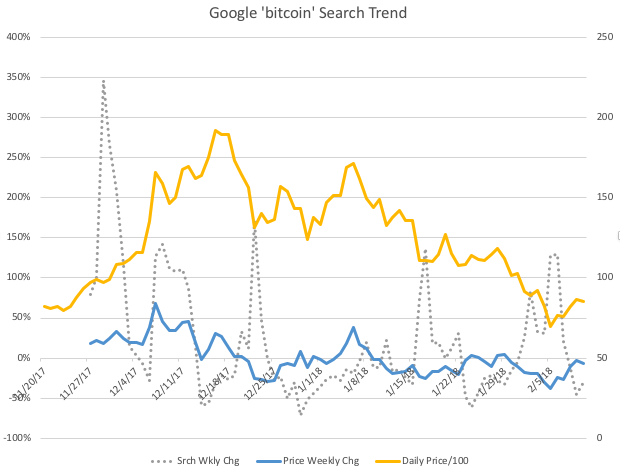
Fig. 12, shows the relationship between various sentiment indicators and cryptocurrencies. While the prices of the coins do not appear to be related to the various sentiment movements, their trade volumes are close but not directly related.



**Fig. 12.** Modeler sample decision tree for minimal features.

## 8.2 Sentiment Analysis

Provided a search trend chart located in google-search-trend.xls. Might be helpful if we don’t have anything better to discuss.



**Fig. 13.** Google ‘bitcoin’ weekly search trend

## 8.3 Portfolio Analysis

Blah blah

# 9 Ethics

Determining how to govern/steward cryptocurrencies is an evolving topic involving multiple ethical concerns. As noted in previous research, the adaption of this technology by nature brings the global society into a "new era of openness, decentralization, and global inclusion [9]." What does governance/stewardship look like for this open, global, decentralized resource with various levels of transaction anonymity? This collective challenge is our central concern of ethics as it pertains to our research. We intend to give a better understanding of the ethical concerns pertaining to each listed component of the challenge equation (open, global, decentralized, and anonymous transactions), as well as inform on current efforts to address these concerns.

## 9.1 Financial Action Task Force (FAFTA)

One organization that has taken notice of this collective challenge is the Financial Action Task Force (FATF). This organization “is an independent inter-governmental body that develops and promotes policies to protect the global financial system against money laundering, terrorist financing and the financing of proliferation of weapons of mass destruction. The FATF Recommendations are recognized as the global anti-money laundering (AML) and counter-terrorist financing (CFT) standard [10].” FATF published its first report in 2013 providing an initial guidance to online alternative currencies. In the report, the organization notes that “given the developing nature of alternate online currencies, the FATF may consider further work in this area in the future [10].”

## 9.2 Open

Cryptocurrencies are “open” because of their public, decentralized ledgers. “Because the blockchain is massively replicated by mutually-distrustful peers, the information it contains is public [12].” The source code for these currencies are open source meaning they are readily available to the public and the ledgers themselves are maintained by public miners. “No one can hide a transaction, and that makes bitcoin more traceable than cash [15].” This openness is where the ethical risk comes into play. With the “secret sauce” of each cryptocurrency being open it is far easier to reverse engineer to deduce information such as account balances and spending habits of the identities within the blockchain. With this public information, entities could learn about these transactions and try to exploit the market for their gain.

## 9.3 Global / Decentralized

Bitcoin’s network is not hindered by international borders. FAFT who are focused on money laundering and terrorist financing abuse prioritized this characteristic when noting potential risks because cryptocurrencies allow for cross-border payments and funds transfers. In their report, they state, “customer transaction records may be held by different entities, often in different jurisdictions, making it more difficult for law enforcement and regulators to access them”. Today, criminals will locate their business in “jurisdictions with weak AMF/CFT regimes” [16] to money launder. With cryptocurrencies, criminals just need to worry about converting fiat to crypto-coin and onto the network. The distributed ledger allows users to access their funds anywhere as long as they can access the internet.

A decentralized network means there is no central oversight body for cryptocurrencies. The current maturity level of governance/stewardship of this technology is similar to the early days of the internet as noted by the Tapscotts [15]. FATF is concerned that lack of a central “trusted” authority to regulate and generate currencies will lead to exploits.

## 9.4 Anonymity

Anonymous transactions mean “the participants in transactions are not explicitly identified: both the sender(s) receiver(s) are identified solely by a pseudonym, and participants in the system can use many different pseudonyms without incurring any meaningful cost [11].” This capability of the system makes it very challenging for an organization to govern the financial industry from money launderers and terrorists not to mention capital gains tax evaders.

# 10 Conclusions

Restate problem and elaborate on abstract.

What did we learn about the market? Lack of correlation of bitcoin to any other asset make them ideal to add them to a portfolio. (add to this thought and cite). Bitcoin acting as a hedge.

What did we learn about sentiment analysis? The effects are real, see Google search trend, but the correlation to price is allusive.

What did we learn about advising?

# 11 Future Work

Future work in this project would include adding additional technical indicators to determine which ones provide the most value in determining trading strategy, incorporating other cryptocurrency price history to determine if feature importance is the same, automate data acquisition, labeling, and training of algorithm, and develop a web service to provide trading strategy for today or past days.

# References

1. Shah, D., Zhang, K.: Bayesian regression and Bitcoin. Department of EECS, MIT. (2014)

2. Coin Market Cap. https://coinmarketcap.com.

3. Stern, H: Fidelity Labs Tests Digital Asset Wallet On Fidelity.com. August, 09, 2017. <https://www.fidelity.com/about-fidelity/corporate/fidelity-labs-tests-digital-asset-wallet-on-fidelity.com>

4. Poyser, O.: Exploring the Determinants of Bitcoin’s price, an application of Bayesian Structural Time Series. June 2017.

5. Coin Desk. https://www.coindesk.com/solving-liquidity-challenge-decentralized-exchanges

6. Cazalet, Z., Roncalli, T.: Facts and Fantasies About Factor Investing. October 2014.

7. Fletcher, T., Hussain, Z., Shawe-Taylor, J.: Currency Forecasting using Multiple Kernel Learning with Financially Motivated Features. NIPS 2010 workshop: New Directions in Multiple Kernel Learning (2010).

8. Indexing and Performance in Crypto Assets: Bletchley Indexes. September 8, 2017. <https://www.bletchleyindexes.com/blog/idx_perf_post>

9. Tapscott D., Tapscott A. (2017). Realizing the Potential of Blockchain. A Multistakeholder Approach to the Stewardship of Blockchain and Cryptocurrencies: Whitepaper. Retrieved from World Economic Forum website:

http://www3.weforum.org/docs/WEF\_Realizing\_Potential\_Blockchain.pdf

10. Financial Action Task Force on Money Laundering. FATF report (2014). http://www.fatf-gafi.org/media/fatf/documents/reports/Virtual-currency-key-definitions-and-potential-aml-cft-risks.pdf

11. Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., & Savage, S.: A fistful of bitcoins. Proceedings of the 2013 conference on Internet measurement conference - IMC '13. doi:10.1145/2504730.2504747

12. Sasson, E. B., Chiesa, A., Garman, C., Green, M., Miers, I., Tromer, E., & Virza, M.: Zerocash: Decentralized Anonymous Payments from Bitcoin. 2014 IEEE Symposium on Security and Privacy. doi:10.1109/sp.2014.36

13. Narayanan, A., et al.: Bitcoin and Cryptocurrency Technologies. Princeton University Press (2016).

14. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. Bitcoin.org (2009).

15. Houser, K.: In the Age of Blockchain, Crytpo Has a Major Problem. Futurism. https://futurism.com/the-age-of-blockchain-crypto-has-a-major-problem (2017).

16. Gholampour, V.: What Can We Learn from Stock Market Tweets? SSRN Electronic Journal (2017). doi:10.2139/ssrn.3041880

17. Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M.: Predictive Sentiment Analysis of Tweets: A Stock Market Application. Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data, 77-88 (2013). doi:10.1007/978-3-642-39146-0\_8

18. Buntinx, JP.: Digital Assets vs Cryptocurrencies. May 17, 2017.

19. Zielak. Coinbase Bitcoin Historical Data. Kaggle, <https://www.kaggle.com/mczielinski/bitcoin-historical-data/data>.

20. Technical Indicator. Investopedia, <https://www.investopedia.com/terms/t/technicalindicator.asp>.

21. Stock Technical Analysis with Python, <https://www.udemy.com/stock-technical-analysis-with-python/>.

22. XGBoost, <https://xgboost.readthedocs.io/en/latest/>.

23. XGBoost with Python, <https://machinelearningmastery.com/xgboost-with-python/>.

24. Monaghan, Angela: Bitcoin is a fraud that will blow up, says JP Morgan boss. <https://www.theguardian.com/technology/2017/sep/13/bitcoin-fraud-jp-morgan-cryptocurrency-drug-dealers>.

25. Chainalysis Team. The Great Bitcoin Price Dip: Its Causes and a Way Forward. <https://medium.com/chainalysis/the-great-bitcoin-price-dip-its-causes-and-a-way-forward-1199e9360adf>.

26. Patel, Hardik. Why is machine learning in finance so hard? February 11, 2018. https://www.linkedin.com/pulse/why-machine-learning-finance-so-hard-hardik-patel/.

1. A fork is when a copy of the source code is used to start an independent software project. [↑](#footnote-ref-1)
2. If an entity collects 51% of the hashing power, they are able to mine faster than everybody else. This allows an entity to double spend on the network by mining their own blocks without announcing them to the network. The double spend case here is the rogue entity spends currency on the public chain but does not include it on the private blockchain. The longer blockchain is accepted by the network with nodes that do not include the rogue public spend. [↑](#footnote-ref-2)
3. Know Your Customer (KYC) [↑](#footnote-ref-3)
4. Sharpe ratio is a way to normalize returns for the risk that was taken to achieve them, with higher values being better. It is calculated here as the annualized return divided by the annualized volatility, so we are using a zero-risk free rate. [↑](#footnote-ref-4)