Cryptocurrency Day Trading and Framing Prediction in Microblog Discourse

Anna Paula Pawlicka Maule
Computer Science and Engineering
Michigan State University
East Lansing, MI 48824
apmaule@gmail.com

Kristen Marie Johnson
Computer Science and Engineering
Michigan State University
East Lansing, MI 48824
kristenj@msu.edu

Abstract

With 56 million people actively trading and investing in cryptocurrency online and globally in 2020, there is an increasing need for automatic social media analysis tools to help understand trading discourse and behavior. In this work, we present a dual natural language modeling pipeline which leverages language and social network behaviors for the prediction of cryptocurrency day trading actions and their associated framing patterns. This pipeline first predicts if tweets can be used to guide day trading behavior, specifically if a cryptocurrency investor should buy, sell, or hold their cryptocurrencies in order to make a profit. Next, tweets are input to an unsupervised deep clustering approach to automatically detect trading framing patterns. Our contributions include the modeling pipeline for this novel task, a new Cryptocurrency Tweets Dataset compiled from influential accounts, and a Historical Price Dataset. Our experiments show that our approach achieves an 88.78% accuracy for day trading behavior prediction and reveals framing fluctuations prior to and during the COVID-19 pandemic that could be used to guide investment actions.

1 Introduction

Beginning with the 2008 introduction of Bitcoin (BTC) (Nakamoto, 2008), a cryptocurrency for a Peer-to-Peer cash system, the use of cryptocurrencies and their corresponding blockchains have increasingly gained in popularity. In 2019, the number of Americans owning cryptocurrency doubled from 7% in 2018 to 14%, representing about 35 million people trading and investing with cryptocurrency (Partz, 2019).

This increase is largely due to the capability of cryptocurrency to improve various applications ranging from increased security of smart contracts to facilitating less expensive, faster cross-border international payments. Another contributing factor to this growth is that digital coins fulfill the property of storing value similar to other *fiat currencies*, which are government-issued currencies not backed by physical commodities, e.g., the American dollar or euro. Finally, cryptocurrency popularity can be associated with its high day trading volume. As of January 2021, the combined worth of all cryptocurrencies was $1 trillion [1], with Bitcoin accounting for $650 billion of this amount. To put this in perspective, the average trading volume of Amazon Inc. is $13 billion per day – less than one-fifth of the BTC daily volume of $70 billion. [2]

Cryptocurrencies were born on the internet, gained their visibility through online and social media coverage, and many investors follow the advice of well-known cryptocurrency experts on Twitter to guide their personal investment strategies (Mone, 2019). Because cryptocurrency prices can fluctuate quickly, resulting in real-life financial gains or losses, models that can rapidly analyze trending discourse on Twitter can be harnessed to guide and benefit investors.

Additionally, work in computational linguistics and the social sciences have shown the benefit of studying *framing*, which is how someone discusses a topic in order to influence or alter the opinion of the public, for understanding microblog discourse (Card et al., 2015; Johnson et al., 2017; Li and Goldwasser, 2019). Specifically, framing in Twitter can be used to understand social phenomena, such as political maneuvering or epidemiology coverage. However, few works exist which study the relationship between framing and cryptocurrency trading, especially during times of economic stress.

Currently, it is estimated that the COVID-19 pandemic has negatively impacted the global economy by hindering economic growth worldwide between 4.5% and 6.0%, with the potential for future global trade to fall up to 9.2% (CRS, 2020). Similar to

---

[1]https://coinmarketcap.com/charts/
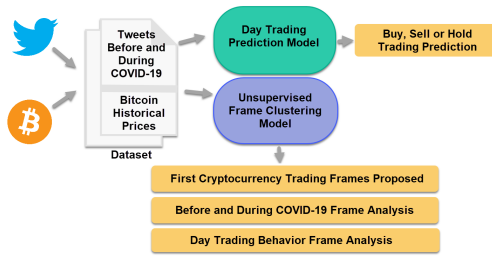[2]https://finance.yahoo.com/quote/AMZN

82

Figure 1: Cryptocurrency Day Trading Prediction and Framing Discovery Pipeline.

the pandemic's effect on Wall Street (i.e., the New York Stock Exchange and NASDAQ), the cryptocurrency market reflected a drastic 47.8% drop on March 12, 2020, one day after the World Health Organization (WHO) announced that COVID-19 could be characterized as a pandemic. Around the same time, a similar drop occurred in stocks worldwide. Therefore, we hypothesized that microblog discourse about day trading behaviors (i.e., buy, sell, hold) would be a useful predictive feature in understanding cryptocurrency trading and framing.

Our contributions include a cryptocurrency day trading behavior modeling pipeline (Figure 1) that leverages language and social network behavior extracted from tweets to implement: (1) a predictive model for investment actions, specifically, whether to buy, sell, or hold cryptocurrency based on daily discussions on Twitter, and (2) an unsupervised deep-learning clustering model to determine the underlying framing patterns used to discuss these investment actions. We have also compiled a Cryptocurrency Tweets Dataset (divided into Pre-COVID and during COVID portions) and Bitcoin Historical Price Dataset.[3] Lastly, we propose novel frames for economic and financial analysis. Our models are able to predict trading actions with 88.78% accuracy, while also revealing a distinction between how day trading behaviors are framed before and during the COVID-19 pandemic.

2 Related Work

Most relevant to this work are previous works that have studied public opinion, sentiment, stock market predictions, and framing using online discourse. Modeling social media microblogs, specifically Twitter, to show connections between online discourse and its effects on public opinion has been widely studied in NLP (Ritter et al., 2010; Walker et al., 2012; Abu-Jbara et al., 2013; Hasan and

Ng, 2014; West et al., 2014; Sridhar et al., 2015) and the social sciences (Bollen et al., 2011; Harlow and Johnson, 2011; Meraz and Papacharissi, 2013; Burch et al., 2015; Jang and Hart, 2015). There are many works on Twitter sentiment analysis, but closest to our work are those concerning the use of Twitter sentiment for stock market predictions (Kouloumpis et al., 2011; Rao and Srivastava, 2012; Si et al., 2013; Derakhshan and Beigy, 2019).

There are relatively few works concerning cryptocurrency analysis and prediction. Of these, a majority use social media sentiment (Jain et al., 2018; Li et al., 2019), volume of tweets (Vidal, 2020), or both (Abraham et al., 2018) as the main feature for prediction. Furthermore, the prediction tasks are typically to predict prices or whether those prices will rise or fall. However, sentiment is known to be difficult to predict on Twitter. Furthermore, the volume of tweets can be falsely inflated by bots reporting currency prices, but not contributing to the discourse. Therefore, instead of sentiment or tweet volume, we aim to use the language directly extracted from tweets, their context, and features representing the social network behavior for a buy, sell, or hold investment action prediction.

Research has also shown that given an adequate amount of historical data, such as stock values and indices, it is possible to forecast future currency exchanges (Walczak, 2001). Different from this work, we focus on predicting cryptocurrency investment *actions*, instead of fiat currency prices, by extracting patterns from historical *tweets* rather than stock values and indices.

Previous works have shown the effectiveness of using frames to predict various social sciences phenomena, such as political framing of Twitter discourse, congressional speeches, and news coverage of current events (Boydstun et al., 2014; Baumer et al., 2015; Card et al., 2015; Tsur et al., 2015; Jang and Hart, 2015; Fulgoni et al., 2016; Johnson et al., 2017; Field et al., 2018). Framing has also been used to understand the role of Twitter discussions in influencing public opinion of events such as riots and protests (Harlow and Johnson, 2011; Meraz and Papacharissi, 2013; Burch et al., 2015). Despite this coverage, to the best of our knowledge we are the first to study the role of framing in economics, specifically concerning stocks or cryptocurrency day trading, or associated correlations with the current pandemic. This work presents a first step in understanding both cryptocurrency day

---

[3]https://github.com/MSU-NLP-CSS/crypto-framing

trading and how framing can reveal insights about cryptocurrency trading.

3 Data Annotation

This section describes the collection and preprocessing steps of the tweets and historical Bitcoin (BTC) transaction prices. Section 3.4 describes how tweets were annotated for use in the day trading behavior prediction model. The non-annotated version of these tweets were used in the framing clustering models.

3.1 Twitter Data Collection

For this work we collected tweets related to cryptocurrency, including BTC and other popular coin types such as Etheureum (ETH) and XRP, because prices of different cryptocurrencies are highly correlated (Magas, 2020). Rather than collect based on hashtags or keywords alone, we narrowed our search to specific time frames and user accounts.

Tweets were scraped from January 2017, when Bitcoin surpassed $1,000 per coin, until February 2020. This time range covers times of frequent changes in cryptocurrency trading and adheres to the finding that an optimal dataset for financial time series prediction consists of information from the past two years (Walczak, 2001). These tweets form our Pre-COVID Dataset.

Within these time frames, three types of user accounts were identified for tweet collection (details in A.1) to maximize the presence of discourse for analysis and minimize tweet noise. These include influential cryptocurrency Twitter accounts, or *influencers*, which are well known as sources for investment information and thus should provide features for message propagation. This category also includes users who frequently tweet about cryptocurrency and have at least ten thousand followers. Similarly, *media* accounts from traditional or online news sources, such as *@CNNBusiness* and *@BitcoinMagazine*, are used. Lastly, we include *company* accounts, e.g., *@IBMBlockchain* and *@BitPay*.

The majority of tweet activity comes from influencer accounts that have between 10,000 and 499,999 followers. There are fewer media accounts, however, these have a much broader reach, e.g., *@nytimes* potentially reaches 48.2 million people. More details of the number of tweets collected per account type are presented in Appendix A.1.

Using the same accounts, we then collected additional cryptocurrency tweets which occurred during the COVID-19 pandemic time frame: from March 2020 until June 2020. These tweets comprise our COVID Dataset. Table 1 summarizes the amount of unique tweets per account type in the two portions comprising the overall dataset.

3.2 BTC Historical Price Data Collection

In addition to cryptocurrency related tweets, we also collected the historical transaction prices of Bitcoin from CoinMarketCap. [4] This BTC Historical Price Dataset contains the following information: the opening price of Bitcoin (Open), the highest price (High), the lowest price (Low), and the closing price (Close) of Bitcoin on that particular day. This dataset also includes the date and the dollar volume of BTC traded that day.

3.3 Preprocessing

Before processing, a total of 484,025 tweets were scraped to collect text with meta-information, including number of replies, number of retweets, and date. Preprocessing consisted of three main steps. First, all tweets were standardized, i.e., we controlled for capitalization, applied stemming, and removed URLs, white space noise, and stop words. Second, we removed irrelevant tweets by filtering for the presence of cryptocurrency-based keywords or hashtags (e.g., bitcoin, btc, ethereum, crypto, cryptocurrency, blockchain, etc.) reducing the dataset to 64,685 tweets. Lastly, in order to create a balanced dataset for training and testing the models, we retained equal proportions of tweets that corresponded to each momentum-labeled group, as described in Section 3.4. After processing, a total of 18,900 Pre-COVID tweets were used in the day trading movement prediction experiments.

For the framing clustering experiments, an additional 123,515 tweets were collected during the beginning of the pandemic. Preprocessing for these tweets consisted of removing: duplicate tweets, English stop words, and references to other users, emails, or website links.

3.4 Annotation

Recall that we are first interested in predicting whether an investor should buy, sell, or hold their cryptocurrency based on tweets discussing cryptocurrency that day. However, such a trading prediction is a challenging task, requiring extensive

---

[4]https://coinmarketcap.com/

CRYPTOCURRENCY TWEETS DATASET	INFLUENCERS	MEDIA	COMPANY	TOTAL
PRE-COVID DATASET	136,637	128,041	110,846	375,524
COVID DATASET	48,254	24,014	36,233	108,501
TOTAL	184,891	152,055	147,079	484,025

Table 1: Quantity of Unique Tweets Per User Account Type.

domain knowledge, yet also susceptible to different investing strategies and conflicting knowledge from Twitter discussions.

Therefore, we used the price information in the BTC Historical Price Dataset (Section 3.2) to define a *momentum* metric that represents the fluctuation of cryptocurrency costs on a given day:

$$momentum = \frac{Price_{close} - Price_{open}}{Price_{open}} \quad (1)$$

If the momentum on a given day increases or decreases by 5% on the following day, then we label tweets of that given day as *buy* or *sell*, respectively. If there is less than a 5% change, these tweets are neutral in terms of buying or selling, and are therefore labeled as *hold*, to represent that an investor should take no action with their cryptocurrency. This 5% cutoff was chosen because BTC volatility during 2017 and 2018 was around 8% (Reiff, 2020), while between 2019 and 2020 it was 4.66% (Tuwiner, 2020). For comparison, the average day trading volatility of stocks is 3.3%, which is considered to be high (Kyröläinen, 2008). This annotation was automated with a Python script that cross referenced the date of the tweet with the BTC Historical Price Dataset and is dependent only on price data.

We then attempted manual annotation. One inexperienced investor and one long-term experienced investor were asked to label a randomly generated subset of the Pre-COVID dataset. They were instructed to label tweets as *buy*, *sell*, or *hold* based on the tweet content and BTC price percentage fluctuation from the previous day (details in A.2).

Table 2 reports the results of two different annotation approaches, where the true labels are those generated by the momentum equation (Eqn. 1). First, annotators were asked to label the tweets based on their content and give an overall label for that particular day based on all their individual tweet annotations (shown in the OVERALL DAY columns of Table 2). Both annotators performed significantly below random guessing, e.g., where the expected label was *sell*.

Second, they were asked to give another overall annotation for a particular day with the addi-

	OVERALL DAY		TWEET + PRICE	
LABEL	INEXP.	EXP.	INEXP.	EXP.
SELL	17%	20%	50%	0%
BUY	28%	36%	30%	0%
HOLD	36%	31%	34%	53%

Table 2: Annotation Precision Experiments.

tional information about the BTC price percentage change from the previous day (shown in the TWEET + PRICE columns). This required annotators to take into consideration the price movement from the previous day to decide on what trading action, *buy*, *sell*, or *hold*, to take. The annotators have very contrasting results, with the inexperienced annotator outperforming random guessing by over 15%. This is likely because their strategy was to sell when prompted with a strongly worded tweet combined with big BTC price drops. By contrast, the experienced annotator did not perform well because a long term investing strategy, i.e., to invest with the goal of profiting in the next 20 years, was applied.

The results of these annotation experiments illustrate that day trading is a non-trivial task for people with and without prior trading and investing experience. Given the high quantity of tweets and highly dynamic language of Twitter, the subjectivity of choosing to buy or sell under different investing strategies, and the large variance in labeling via human annotators, we used the momentum metric as a weak form of supervision to generate labels for investment actions in the day trading prediction.

4 Models & Features

This section describes the two models and their associated features, which represent social network aspects of Twitter and the actual language and context of tweets. Both models are incorporated into a pipeline (Figure 1) that enables us to make trading predictions and discover and analyze new frames.

4.1 Day Trading Behavior Prediction

Model. For the day trading prediction model, we experimented with a combination of features, mod-

els, and a balanced dataset where there is an equal number of *sell*, *buy*, and *hold* labels. Naive Bayes with Bag-of-Words (BOW) features was used for the baseline model. We then tested Random Forest, RNN, and LSTM models that resulted in final accuracies above 85%. We conclude in Section 5 that the best performing model for this task [5] is the RNN with three layers.

Features. Social network related features are extracted directly from the meta-information of the cryptocurrency tweets. This includes the number of retweets and the number of replies. In our experiments, we found that the number of retweets provided some information gain when weighting the tweet feature representation. The type of user account, either influencer, media, or company, that posted the tweet is also used as a feature.

In addition to social features, we also used features directly related to the language of the tweet. First, we implemented an LDA topic model (Jelodar et al., 2019) and used the presence of a top 10 topic in a given tweet as a feature. Next, the tweets were transformed into 768 language features using DistilBERT. [6] All of the tweets were concatenated according to their momentum label and for each group (buy, sell, or hold), DistilBERT was used to extract high-quality language features to represent each of the three tweet groups. Finally, we calculated the cosine similarity of each tweet to these three group representations. We selected the match between a tweet and group with the highest cosine similarity to be used as a feature for that tweet. More concretely, each tweet is compared to the DistilBERT representation of the buy, sell, and hold concatenated tweet groups and the highest similarity group is chosen to be used as a feature.

4.2 Discourse Framing Clustering

Model. From an NLP perspective, frames represent latent abstractions of a discussion and are *not* equivalent to topics. We hypothesized that *how* a topic is discussed, or framed, could be identified in an unsupervised manner by analyzing how the tweet content clusters together. In order to extract the clusters which represent such frames, we tried two modeling approaches. First, we used a basic k-means clustering. Second, we implemented the

MODEL	BOW	ALL FEATURES
NAIVE BAYES	49.72%	61.58%
RANDOM FOREST	63.81%	86.61%
RNN	33.67%	**88.78%**
LSTM	31.57%	88.18%

Table 3: Experimental Results. The columns represent the accuracy of each model when using either a bag-of-words (BOW) or all features of Section 4.1 combined with a DistilBERT (Sanh et al., 2019) representation of the tweets as features.

unsupervised Deep Embedded Clustering (DEC) approach of Xie et al.; Hadifar et al., which combines both an autoencoder and k-means clustering to achieve a more precise separation. DEC simultaneously learns feature representations and cluster assignments.

Features. The features used for the basic k-means and DEC models were sparse representations of the word counts for each tweet. Both BOW and TF-IDF features were used as input to the k-means model and autoencoder of the DEC pipeline.

5 Experimental Results

In this section, we present our experimental setup and an analysis of our modeling results.

5.1 Day Trading Behavior Prediction

We conducted supervised experiments using five-fold cross-validation with random shuffling and an 80% training and 20% testing split. For the Neural Networks, we experimented with 50 epochs because the dropout after each layer was 0.001.

Table 3 shows the results of using the following models: Naive Bayes, Random Forest, Recurrent Neural Network, and an LSTM. Both the RNN and LSTM use three dense layers. The columns of Table 3 correspond to the tweet feature representations used with each model: a baseline where tweets are represented as BOW and the combination of all features described in Section 4.1.

From Table 3, we can see that using an RNN with all combined features has the highest accuracy of 88.78% across all three classes. Details of the label prediction distribution are shown in Table 8 in Appendix A.3. Predicting day trading behavior, i.e., whether to buy or sell stock, is a complicated task, especially in a volatile asset such as cryptocurrency. Ablation studies revealed that the most informative features for prediction were the

---

[5]Our task is classification for future use in downstream applications. Thus we do not perform regression or time-series analysis in this paper, but will use time-series in future work.

[6]DistilBERT had a 0.6% better performance than BERT.

language features, specifically the combination of DistilBERT representations with cosine similarity.

(a) Using k-means.
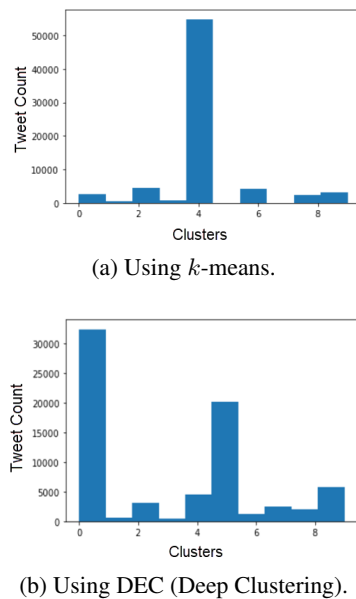

(b) Using DEC (Deep Clustering).

Figure 2: Number of Tweets Per Cluster. Both figures show the number of tweets per cluster using 10 initial clusters and BOW features for the Pre-COVID dataset.

5.2 Discourse Framing Prediction

We conducted unsupervised clustering experiments using (1) a basic k-means clustering and (2) deep clustering with autoenconders (DEC) as described in Section 4.2. The encoder outputs were used as inputs to the deep clustering layer, and similar to Hadifar et al., the k-means center clusters were used as initial weights for the deep clustering model. The tweets were randomly shuffled for training. The autoencoder ran for 100 epochs, achieving an accuracy of 99.99% with both training and validation loss on the order of 5.5453e-04 and without overfitting.

We initially experimented with 32 clusters because 32 is the default number of features that get compressed by the autoencoder. However, we observed that several clusters had similar, overlapping themes and keywords. Therefore, we conducted the rest of our experiments with 10 clusters. Figure 2 shows the number of tweets that fall into each of the 10 initial clusters for each modeling approach.

Figure 2a shows six clusters identified in our Pre-COVID Dataset by k-means clustering. Using Singular Value Decomposition (SVD) (Figure 3a) and an analysis of the most frequent words appearing in each cluster, we were able to extract three main clusters: the first cluster included tweets discussing

(a) Using k-means.
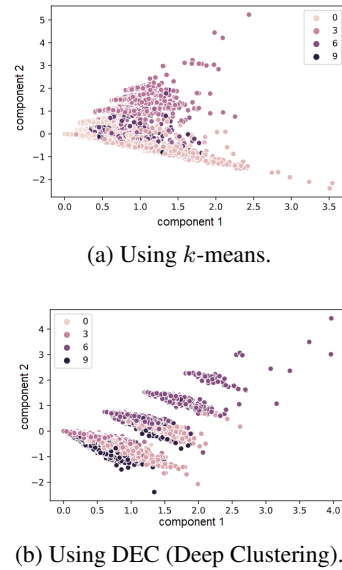

(b) Using DEC (Deep Clustering).

Figure 3: Pre-COVID Dataset Cluster Visualization on Reduced Dimensions Using SVD. SVD is used to reduce the clusters (0 to 9) to two dimensions to better visualize the frame groupings.

Bitcoin halving [7], the second concerns trading and investing cryptocurrency, and the third discussed how trading is affected by politics.

The DEC clustering of the Pre-COVID dataset (Figure 3b) also identified four main clusters: one discussing halving but with more emphasis on long term store value, one discussing political effects, and two discussing cryptocurrency trading and applications. This latter cluster splits the *cryptocurrency trading and investing* cluster identified as one large cluster by k-means into two clusters. Section 6 provides more analysis of the frames these clusters represent, how they are associated with trading behaviors, and how these associations change during the pandemic.

Cluster Verification. Given the novel aspect of this task, the difficulty of determining frames within tweets, and that both clustering approaches operate in an unsupervised setting, we conducted three cluster verification experiments. First, we asked an evaluator to determine how well the clusters represent how cryptocurrency discussions are framed. Given a subset of tweets, the evaluator was asked to "label" if cryptocurrency was discussed in the tweet with one of the DEC-identified frames using the following guidelines:

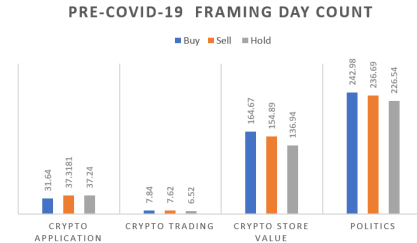• *Trading Frame*: Does the tweet discuss how

---

[7]About every 4 years or so, the amount of BTC that can be mined (mining capacity) decreases by half.

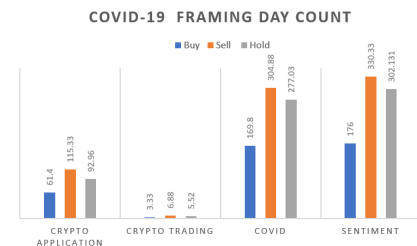| Topic	Top Words
Knowledge	know, bitcoin, time, blockchain, market, world, buy, change, people, point, today
Business	year, thank, start, problem, business, write, stop, plan, risk, reason, check
Support	make, think, work, want, day, people, need, use, year, week, support, happen, read
Hold	look, price, money, try, build, econ, think, end, tell, idea, people, term, win, hold

Table 4: Pre-COVID Dataset Top 4 LDA Topics and Most Frequent Keywords.

or why to buy or sell cryptocurrency?

- *Application Frame*: Does the tweet emphasize the uses of cryptocurrency?
- *Store Value Frame*: Does the tweet discuss cryptocurrency in terms of long term value?
- *Political Frame*: Does the tweet put a political spin on cryptocurrency trading actions?

The evaluator's manual annotation was compared to the actual cluster (or frame) the tweet was assigned to by the DEC model. With this evaluation approach, we found the clustering to be 69.23% accurate. Given the lack of previous work on cryptocurrency framing, we compared this result to a previous work which found an annotator agreement of 73.4% on a tweet dataset labeled for political frames (Johnson et al., 2017).

Next, a chi-square test was performed to verify the hypothesis that the frames, represented by clusters, were independent of each other. In order to perform the test, the top word count was collected for each cluster, as well as their count in every other cluster. The resulting *p-value* was less than 0.05, indicating that the clusters are independent.

Lastly, to justify that these clusters represent how tweets are framed, we also performed an LDA topic analysis to ensure that clusters were not finding topics. Table 4 shows the top 4 LDA topics which are different from those extracted for frames (more details in Section 6). Topics represent the content of the tweet, e.g., the topic *Hold* represents holding cryptocurrency. Frames, however, are fundamentally different and represent *how* someone discusses a topic, e.g., how or why to hold.

6 Qualitative Results

In this section, we explore how cryptocurrency frames change over time and their correlation with cryptocurrency day trading behavior. Section 6.1 shows the effects of the pandemic on day trading discussions and behaviors. Section 6.2 discusses how day trading behaviors are framed.

(a) Pre-COVID Frames with Actions.

(b) COVID-19 Frames with Actions.

Figure 4: Predicted Frames and Investment Actions. Each figure shows the quantity of tweets using a certain frame (separated by a grey line) associated with each investment movement action: buy, sell, or hold.

6.1 Frames Before and During the Pandemic

Tables 5 and 6 show the most frequent words appearing in each of the four clusters extracted from the Pre-COVID or COVID Dataset, respectively. Prior to the pandemic, Table 5 shows that the cryptocurrency tweets were framed in terms of aspects important to cryptocurrency itself, i.e., trading actions, applications or uses, and long term store value. Table 6 shows that once the pandemic was occurring, the discussion shifted. People still discussed cryptocurrency in terms of trading and applications, however, there was a shift from focusing on long term value and political effects on cryptocurrency to sentiment concerning cryptocurrency and the pandemic.

One interesting event captured by the *Trading* frame in the COVID-19 Dataset was the BTC halving event on May 11, 2020. This halving marks the first quarter of the year as a historical event in the cryptocurrency world because this is the third halv-

88

Frame	Most Frequent Words
CRYPTO TRADING	price, bitcoin, usd, market, trading, value, action
CRYPTO APPLICATION	blockchain, btc, business, use, tech, crypto
CRYPTO STORE VALUE	bitcoin, people, need, want, use, market, value, years
POLITICS	world, man, president, america, china, work, government, time

Table 5: Most Frequent Words Per Cluster Prior to COVID-19 (Pre-COVID Dataset).

Frame	Most Frequent Words
CRYPTO TRADING	money, crypto, btc, trading, finance, investment, halving
CRYPTO APPLICATION	btc, crypto, time, right, know
SENTIMENT	like, look, things, dont, good, time, feel
COVID	people, coronavirus, covid, pandemic, bitcoin, world, dont

Table 6: Most Frequent Words Per Cluster During COVID-19 (COVID Dataset).

ing to take place. The past two times that halving occurred, Bitcoin later experienced an all-time high price jump. In future work, we aim to track this frame and use it to predict potential price jumps.

6.2 Frames and Momentum Patterns

From observing the frames and momentum patterns prior to the pandemic shown in Figure 4a, we can see that *Store Value* frames have a higher frequency when the momentum pattern suggests a *Buy* movement. This correlation makes sense because if there is a belief that some asset will store value it creates more confidence in buying and holding the cryptocurrency. It is also not surprising that there is an increase in *Political* frames associated with the *Buy* movement. Countries and economies often cited as being politically unstable, such as Botswana, Ghana, Venezuela, and India, have seen an increase in BTC interest because it is more stable than fiat currencies from those countries [8]. Another potential association with the slight increase in *Political* frames during a *Buy* movement is the increase of government adoption and additional regulation of cryptocurrencies. These patterns suggest that prior to the pandemic, if Twitter cryptocurrency discussions were framed in terms of store value or politics, an investor might consider buying more cryptocurrency.

During the COVID-19 time span (Figure 4b), all frames decrease during an indicated *Buy* movement. However, the opposite occurs, i.e., all frames increase, when the indicated movement is to *Sell*.

Regarding both *Trading* and *Application* frames, it makes sense to purchase cryptocurrency when nobody is talking about it, and sell it when the interest in those topics rises. The *COVID* frame having a lower frequency during a *Buy* movement could indicate that investors feel less threatened by the market instability introduced by the pandemic, which is the opposite of the general sentiment of investors dealing with physical stock exchange markets.

7 Conclusion

We have presented a dual modeling pipeline to understand how the way influential people and news sources frame cryptocurrency discussions on Twitter affects cryptocurrency day trading. Using classic NLP techniques and cosine similarity between the DistilBERT representations of tweet features and cryptocurrency tweets, we provide a day trading prediction model that is capable of distinguishing between day trading actions such as buy, sell, or hold. Using our features and modeling approach we are able to achieve an accuracy of 88.78% over a 49.72% traditional baseline. Furthermore, we are first to present an unsupervised deep clustering approach to reveal the latent frames used to discuss these day trading behaviors. Our work shows interesting relationships between investment actions and how cryptocurrency discussions are framed on Twitter, as well as how these framing patterns change in response to a pandemic.

8 Acknowledgments

---

[8]https://news.coinsquare.com/government/government-instability-bitcoin/;
https://www.un.org/africarenewal/magazine/april-2018-july-2018/africa-could-be-next-frontier-cryptocurrency

89

References

Jethin Abraham, Daniel Higdon, John Nelson, and Juan Ibarra. 2018. Cryptocurrency price prediction using tweet volumes and sentiment analysis. In *SMU Data Science Review: Vol. 1: No. 3, Article 1*.

Amjad Abu-Jbara, Ben King, Mona Diab, and Dragomir Radev. 2013. Identifying opinion subgroups in arabic online discussions. In *Proc. of ACL*.

Abderrahim Ait Azzi, Houda Bouamor, and Sira Ferradans. 2019. The FinSBD-2019 shared task: Sentence boundary detection in PDF noisy text in the financial domain. In *Proceedings of the First Workshop on Financial Technology and Natural Language Processing*, pages 74–80, Macao, China.

Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta, and Geri Gay. 2015. Testing and comparing computational approaches for identifying the language of framing in political news. In *In Proc. of NAACL*.

Johan Bollen, Huina Mao, and Alberto Pepe. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proc. of AAAI Conference on Weblogs and Social Media*.

Amber Boydstun, Dallas Card, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2014. Tracking the development of media frames within and across policy issues.

Lauren M. Burch, Evan L. Frederick, and Ann Pegoraro. 2015. Kissing in the carnage: An examination of framing on twitter during the vancouver riots. *Journal of Broadcasting & Electronic Media*, 59(3):399–415.

Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proc. of ACL*.

Congressional Research Service CRS. 2020. Global economic effects of covid-19.

Ali Derakhshan and Hamid Beigy. 2019. Sentiment analysis on stock social media for stock price movement prediction. In *Engineering Applications of Artificial Intelligence*.

Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. Framing and agenda-setting in Russian news: a computational analysis of intricate political strategies. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3570–3580, Brussels, Belgium. Association for Computational Linguistics.

Dean Fulgoni, Jordan Carpenter, Lyle Ungar, and Daniel Preotiuc-Pietro. 2016. An empirical exploration of moral foundations theory in partisan news sources. In *Proc. of LREC*.

Amir Hadifar, Lucas Sterckx, Thomas Demeester, and Chris Develder. 2019. A self-training approach for short text clustering. In *Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019)*, pages 194–199, Florence, Italy. Association for Computational Linguistics.

Summer Harlow and Thomas Johnson. 2011. The arab spring| overthrowing the protest paradigm? how the new york times, global voices and twitter covered the egyptian revolution. *International Journal of Communication*, 5(0).

Kazi Saidul Hasan and Vincent Ng. 2014. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *Proc. of EMNLP*.

Arti Jain, Shashank Tripathi, Harsh Dhardwivedi, and Pranav Saxena. 2018. Forecasting price of cryptocurrencies using tweets sentiment analysis.

S. Mo Jang and P. Sol Hart. 2015. Polarized frames on "climate change" and "global warming" across countries and states: Evidence from twitter big data. *Global Environmental Change*, 32:11–17.

Hamed Jelodar, Yongli Wang, Chi Yuan, Xia Feng, Xiahui Jiang, Yanchao Li, and Liang Zhao. 2019. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11):15169–15211.

Kristen Johnson, Di Jin, and Dan Goldwasser. 2017. Leveraging behavioral and social information for weakly supervised collective classification of political discourse on twitter. In *Proc. of ACL*.

Katherine Keith and Amanda Stent. 2019. Modeling financial analysts' decision making via the pragmatics and semantics of earnings calls. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 493–503, Florence, Italy. Association for Computational Linguistics.

Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore. 2011. Twitter sentiment analysis: The good the bad and the omg! In *Proc. of AAAI Conference on Weblogs and Social Media*.

Petri Kyröläinen. 2008. Day trading and stock price volatility. *Journal of Economics and Finance*, 32:75–89.

Chang Li and Dan Goldwasser. 2019. Social information with graph convolutional networks for political perspective detection in news media. In *Association for Computational Linguistics (ACL)*.

Tianyu Ray Li, Anup S. Chamrajnagar, Xander R. Fong, Nicholas R. Rizik, and Feng Fu. 2019. Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model. *Frontiers in Physics*, 7:98.

Julia Magas. 2020. Truth about crypto price correlation: How closely does eth follow btc?

Sharon Meraz and Zizi Papacharissi. 2013. Networked gatekeeping and networked framing on #egypt. *The International Journal of Press/Politics*, 18(2):138–166.

Lesa Mone. 2019. I read crypto twitter for hours daily — here are the 40 accounts that really matter. In *ConsenSys Blog*.

Satoshi Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system.

Helen Partz. 2019. Number of americans owning crypto doubled in 2019: Finder. In *Coin Telegraph*.

Tushar Rao and Saket Srivastava. 2012. Analyzing stock market movements using twitter sentiment analysis. In *Proc. of International Conference on Advances in Social Networks Analysis and Mining*.

Nathan Reiff. 2020. Why bitcoin has a volatile value.

Alan Ritter, Colin Cherry, and Bill Dolan. 2010. Unsupervised modeling of twitter conversations. In *Proc. of NAACL*.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter.

Ramit Sawhney, Puneet Mathur, Ayush Mangal, Piyush Khanna, Rajiv Ratn Shah, and Roger Zimmermann. 2020. *Multimodal Multi-Task Financial Risk Forecasting*, page 456–465. Association for Computing Machinery, New York, NY, USA.

Jianfeng Si, Arjun Mukherjee, Bing Liu, Qing Li, Huayi Li, and Xiaotie Deng. 2013. Exploiting topic based twitter sentiment for stock prediction. In *Proc. of 51st Annual Meeting of the Association for Computational Linguistics*.

Dhanya Sridhar, James Foulds, Bert Huang, Lise Getoor, and Marilyn Walker. 2015. Joint models of disagreement and stance in online debate. In *Proc. of ACL*.

Oren Tsur, Dan Calacci, and David Lazer. 2015. A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In *Proc. of ACL*.

Austin Tuwiner. 2020. The bitcoin volatility index price and more.

Tiago Vidal. 2020. How traders can use twitter to anticipate bitcoin price moves, volume.

Steven Walczak. 2001. An empirical analysis of data requirements for financial forecasting with neural networks. *Journal of management information systems*, 17(4):203–222.

Marilyn A. Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proc. of NAACL*.

Robert West, Hristo S Paskov, Jure Leskovec, and Christopher Potts. 2014. Exploiting social network structure for person-to-person sentiment analysis. *TACL*.

Junyuan Xie, Ross Girshick, and Ali Farhadi. 2015. Unsupervised deep embedding for clustering analysis.

91

QUANTITY	INFL.	MEDIA	CO.
$10,000 - 99,999$	45	-	13
$100,000 - 499,999$	24	2	5
$500,000 - 999,999$	2	2	1
$\geq 1,000,000$	-	5	-

Table 7: Quantity of Followers Per User Account Type. Each row represents the number of user account types (columns: influencers, media, company) that have that quantity of followers who are actively tweeting about cryptocurrency.

MODEL	BUY	SELL	HOLD
NAIVE BAYES	57 %	62 %	65 %
RANDOM FOREST	82%	87 %	90%
RNN	85%	88 %	89%

Table 8: Label Distribution Results. The columns represent the accuracy of each label based on models using all the features of Section 4.1 combined.

A Appendix

A.1 Twitter Data Collection

In order to determine our Twitter accounts subset we narrowed it down to accounts that were associated with *crypto*, *cryptocurrency*, *bitcoin* and *blockchain* keywords. We only considered accounts that had more than *10,000* followers. Further, we segregated the accounts into three distinct groups *influencers*, *media*, and *companies*.

Table 7 presents the distribution of followers for accounts collected from the different types of accounts: influencers, media, or company. Column one lists the quantity of followers, divided into four groups. The remaining columns indicate how many of the influencer, media, and company accounts have the different number of followers. From this table, we can see that the majority of tweet activity comes from influencer accounts that have between 10,000 and 499,999 followers. There are fewer media accounts, however, these accounts have much broader reach. For example, @*nytimes* reaches up to 48.2 million people when tweeting about cryptocurrencies.

A.2 Annotations

For the annotation experiments, both annotators had different levels of experience in both investing and trading stocks and cryptocurrencies. One of the annotators was an inexperienced investor, who has never bought or sold cryptocurrencies or stocks. The second annotator is an experienced investor that has been investing and following the stock market for the past 5 years, and in the past 2 years has been investing in cryptocurrencies. Annotators were given a randomly selected subset of the Pre-COVID Dataset to label for supervised experiments for the day trading actions prediction. The reduced dataset for manual annotation has 798 unique tweets, covering approximately 1% of the total dataset. There are 114 different days represented, with 7 distinct tweets per day.

A.3 Day Trading Prediction Results

Table 8 shows the prediction accuracy for each label (buy, sell, or hold) for the three different models when using all features. We have also performed ablation studies for the features, which can be included with the final draft of the paper.

A.4 Somewhat Related Work

The combination of NLP and financial applications has been gaining interest in recent years. There have been three recent somewhat related publications working in an economics or financial domain. However, these papers are not *directly* related to this work and due to page constraints, we have moved them to the Appendix for now. Azzi et al. report shared task findings for sentence boundary detection of noisy financial PDFs in the First Workshop on Financial Technology and Natural Language Processing (FinNLP). Keith and Stent compare financial analysts' decision making with fiscal quarter earning calls. Finally, Sawhney et al. uses a multimodal text and audio attention model to predict stock market prices.