# Value Relevance of Social Media Sentiment:

# An Experiment with Steem using Association Analyses

Debra VanderMeer, Sankalp Singh

## Abstract

Most popular social media platforms support a voting mechanism to capture an assessment of how much the network values a particular post, where more votes or "likes" implies greater valuation among network participants. Such voting mechanisms are subject to confounding factors, such as relative popularity of the poster, as well as outright acts of manipulation to increase vote counts. We hypothesize that post sentiment plays a role in content valuation. We expect that participants will value posts with positive sentiment more than posts with negative sentiment. Further, we postulate that the degree of positive sentiment matters, such posts with a lesser degree of positive sentiment will be more highly valued than posts demonstrating a greater degree of positivity. We base our hypothesis on theories associated with risk aversion, where users are more interested in content that may signal a need to act to avoid potential negative consequences.

## Introduction and Related Works

Social media has become an increasingly popular forum for individual investors to post and exchange their own analyses of financial securities, especially since the inception of Twitter in 2006. As evidenced in the short squeeze phenomena the financial market experienced in the first quarter of 2021, investor opinions in social media can have significant impacts on security valuations. The information contents of social media and their value relevance (i.e., the extent to which the audience finds value in the content) have gained growing recognition among both academia and practitioners (e.g., Chen et al., 2014; Jame et al., 2016; Da and Huang, 2020; Jame et al., 2021). Accordingly, institutional investors are now subscribing to social media reports to seek investment ideas and/or support their investment decisions.

While it is clear from prior research that value relevance is important, few studies have suggested appropriate mechanisms to identify and measure the extent to which an audience finds value in a particular post. More thorough investigation is necessary to determine which instances of content are valued. This is particularly challenging on typical social media platforms, where the value of a post is often judged by the number of likes or followers a user has, which can be easily manipulated.

Research in text analysis shows that sentiment is an important element in an audience's reaction to text. Sentiment refers to the emotional tone or attitude expressed in a piece of text (Pang and Lee, 2008), which can be positive, negative, or neutral.

Sentiment analysis uses natural language processing and machine learning to understand people's opinions and attitudes towards a particular topic or entity. There are many studies using sentiment analysis in a variety of applications, e.g., in finance (Bartov et al., 2018; Smailović et al., 2013; Ren et al., 2019), hospitality (Zvarevashe and Olugbara, 2018, Valdivia et al., 2019), and retail (Fang et al., 2015). Most of these studies attempt to link sentiment to a specific outcome. For example, many applications across Finance and Accounting correlate the sentiment to a measure or proxy for the firm valuation, such as share price or stock return or to the extent of impact on stock analysts. These proxies represent indirect measures of the value of the content of a post.

In contrast, few, if any, studies attempt to draw a more direct link between sentiment and the extent to which a reader values the content of a post. One example of such a direct link appears in Amazon product reviews, where the site allows users to mark individual reviews as “helpful,” and displays the number of people who found the review helpful. Here, Amazon reports on the direct utility site users found in a review’s content. In this example, the feedback is limited to positive feedback (i.e., there is no corresponding “unhelpful” button to press). Further, the granularity of feedback is coarse: there is no way for a user to rate the helpfulness of a review on a more nuanced basis.

This research looks at the sentiment of a social media post and relates it back to actual value generated using a direct measure of the extent to which the audience values the content of the post. Typically, the classical risk aversion assumption (Arrow, 1971; Pratt, 1964) will point to the tendency of investors to process negative information more efficiently. Basically, people would like to know more if there is less-positive information. Thus, (Hypothesis 1) we expect that the lesser the positivity the higher the monetary value of the information will be.

To consider the extent to which users value content, we consider posts to the Steemit social network. Steemit operates differently from other social media platforms when it comes to voting manipulation. This is because it utilizes a distinct consensus mechanism known as Delegated Proof of Stake (DPoS). DPoS ensures that votes are distributed fairly and transparently, with a network of trusted validators verifying transactions and upholding the blockchain's integrity. Furthermore, Steemit offers a token called "Steem Power" that users can earn by creating high-quality content, commenting on posts, and upvoting other users' content. This provides users with greater influence over the voting process, but accumulating Steem Power requires a significant investment of time and effort. Steemit also enables users to "flag" content that they believe to be spam or in violation of the platform's rules, which reduces the post's visibility and Steem Power rewards. Overall, Steemit's blockchain-based system and unique consensus mechanism make it more challenging for users to manipulate votes compared to traditional social media platforms. This provides a quantified measure of the audience valuation of a post.

We compiled a dataset of Steemit posts and associated SteemCoin earnings. This provides an opportunity to analyze how the value of a post depends on its content, as opposed to its popularity or authorship. Specifically, the data allows us to determine the correlation between sentiment and money gained. We used a set of advanced tools for sentiment analysis, such as VADER, Google sentiment analysis, and IBM Watson sentiment analyzer to determine the degree of positive or negative sentiment in a post, providing a quantified measure of sentiment. We then went on to explore which type of sentiment is associated with greater monetary value.

In our analysis, we found that a particular sentiment tends to accrue more money. Our results demonstrate that, among posts identified as having positive sentiment, the lesser the positivity of sentiment, the higher the dollar amount. This trend has the potential to help social media posters predict how consumers will react to the sentiment in their posts.

**Data Description**

We generated our Steem dataset using the SteemOps database of operations on the Steem blockchain described by Chao et al., (2021). This dataset contains three sub-datasets that capture three types of blockchain operations: (1) the social-network operation dataset, which captures user actions for posting content; (2) the witness-election operation dataset, which holds information about voting operations for the blockchain; and (3) the value-transfer operation dataset, which captures information about value transfers on the blockchain.
We considered operations from the social-network operation dataset because it contains post-related information. We considered post details for three randomly-selected months from the available dataset (August 2016, October 2016, and November of 2017). Here, a post could be either a standalone post (i.e., not replying to another post), or a reply to a prior post. Both types of contributed content can earn money in the Steem platform, so we treat them the same way in our analyses.

This SteemOps dataset does not contain the text of posts, but does provide sufficient information to create a URL that returns the full description of the post details, including data attributes of interest in our study: timestamp, title, author, content, monetary value earned, and Steem blockchain block number. We coded a scraping algorithm to gather post details, based on the Steem operation data.

The initial analysis of operations in the SteemOps dataset for our selected months yielded pointers to 641172 Steem posts. Based on this, we attempted to gather post details using our scraping code for our selected months. Within this dataset, some operations in the SteemOps dataset were missing necessary descriptors to construct a URL (e.g., missing a post author or title). Since we could not associate comments or monetary value data with these posts, these are excluded from our sentiment and correlation analyses. This resulted in a final count of n=1,578 posts used for analysis in this work.

## 3. Analysis and Results

### 3.1 Analysis Method

Our analysis method consists of two steps: (1) associating sentiment with each post through sentiment analysis, and (2) correlating sentiment with monetary value earned.

In the sentiment analysis step, we identified a number of tools in common use for associating sentiment with instances of text. These sentiment engines use Natural Language Processing and Machine Learning Algorithms to determine a sentiment value for a set of text

Generally, these tools take a set of text instances as input, and produce a numeric sentiment score for each text instance in the range [-1, 1], where scores in the range [-1, 0) indicate negative sentiment, a score of 0 indicates neutral sentiment, and scores in the range (0, 1] indicate positive content. A score's distance from 0 indicates the degree of sentiment detected. For example, a score of 0.82 indicates a more strongly positive sentiment than a score of 0.43, and a score of -0.64 indicates a more strongly negative sentiment than a score of -0.33.

We performed sentiment analysis using five sentiment engines, namely VADER, Google, IBM Watson, TextBlob, and Flair to analyze the Steem posts. We used the default parameters in each tool, and selected the post content as the text to be analyzed. This resulted in a sentiment score for each post.

In the correlation analysis step, we identified the sentiment score as the independent variable and monetary value earned as the dependent variable for the correlation study. We ran separate correlation analyses for each sentiment analysis engine, and within these, we separated the analysis for positive-scored posts from negative-scored posts.

### 3.2 Descriptive Statistics

The descriptive statistics of sentiment value received by sentiment analyzer engines are given in Table 1.

Table 1: Basic Statistic Tables for Each Sentiment Engine

| Engine | Sentiment | Mean Sentiment Value | Standard Deviation for Sentiment Variable | 50th Percentile of Sentiment Data | Mean Monetary Value | Standard Deviation for Monetary Value | 50th Percentile of Monetary Data |
|---|---|---|---|---|---|---|---|
| VADER | Negative | 0.547769 | 0.302488 | 0.520400 | 5.531189 | 33.906036 | 0 |
| VADER | Positive | 0.739857 | 0.236677 | 0.809250 | 3.654264 | 20.25803 | 0 |
| Google | Negative | 0.312329 | 0.205221 | 0.300000 | 5.447400 | 30.815341 | 0 |
| Google | Positive | 0.519575 | 0.291021 | 0.500000 | 2.256757 | 0.291021 | 0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| IBM | Negative | 0.591530 | 0.205713 | 0.555805 | 5.341929 | 30.149836 | 0 |
| IBM | Positive | 0.780429 | 0.209469 | 0.841410 | 3.495036 | 20.321763 | 0 |
| TextBlob | Negative | 0.198972 | 0.186496 | 0.130000 | 1.344741 | 6.187580 | 0 |
| TextBlob | Positive | 0.337388 | 0.242827 | 0.271383 | 4.442621 | 24.696343 | 0 |
| Flair | Negative | 0.924123 | 0.128778 | 0.994073 | 5.259891 | 28.653539 | 0 |
| Flair | Positive | 0.953272 | 0.098668 | 0.995366 | 2.706277 | 0.098668 | 0 |

Descriptive statistics for posts with positive and negative sentiment analyzed through five different sentiment analysis engines

In general, it is common to see variations in sentiment values across different engines, as each engine may use different algorithms, models, or data sources to determine the sentiment of the text. Some engines may be more accurate in determining sentiment in certain contexts or domains, while others may perform better on different types of text. Therefore, it is important to use multiple sentiment analysis engines and compare their results to get a more comprehensive understanding of the sentiment of a given text.

The IBM Watson, TextBlob, and Flair engines have the tightest data distribution, as seen in their respective standard deviations in both the positive and negative sentiment categories. On the other hand, VADER has the highest standard deviation in the negative, meaning that VADER found that there was a wide range of negativity. Overall, having a larger standard deviation leads to less accuracy.

### 3.3 Correlation Analyses:

We used Spearman and Pearson Correlation analysis to examine the relationship between dollar amount and sentiment of the comment. Pearson evaluates the correlation between two continuous variables linearly, while Spearmen performs a similar evaluation, but with monotonic functions. Spearman looks for the direction of monotonicity.

Both correlations seemed accurate regarding direction, as both predicted a general negative correlation. However, on VADER, as stated before, both correlation engines (Pearson and Spearman) indicated a positive correlation between monetary value per comment and sentiment analysis. This means that the amount of money gained on a given post increases based on the positivity of the post.

The advantage of using both correlation coefficients is that we can see if the data is accurate since both correlation values should look similar. Overall, our findings find that the lesser the positivity of sentiment, the higher the dollar amount. This is true only for the positive sentiment samples, not the negative sample data, and valid for our four major sentiment analyzers (the VADER sentiment analysis engines produced contrary findings). Table 2 presents the detailed findings of our correlation analysis.

Table 2: Correlation Table

| Sentiment Engine | Positive Posts | | | Negative Posts | | |
|---|---|---|---|---|---|---|
| | Spearman Correlation | Pearson Correlation | P-value for Pearson correlation | Spearman Correlation | Pearson Correlation | P-value for Pearson correlation |
| IBM | -0.361453 | -0.138171 | 0.000655 | -0.196078 | -0.197823 | 0.198026 |
| Google | -0.349209 | -0.135522 | 0.000831 | -0.374889 | -0.273393 | 0.072548 |
| TB | -0.323644 | -0.103500 | 0.010854 | -0.209676 | -0.204294 | 0.183451 |
| FLAIR | 0.075358 | -0.090180 | 0.026550 | -0.089656 | -0.402749 | 0.006718 |
| VADER | 0.319074 | 0.109934 | 0.006796 | 0.227269 | 0.267204 | 0.079525 |

Correlation values for each of the five engines for positive and negative posts

We interpret the correlations for each sentiment analyzer below. For IBM Watson, both Spearman (-0.361453) and Pearson (-0.13817) correlation numbers are negative and have significant P values for Pearson correlation (0.0006546) for the positive sentiment group. For the negative sentiment category, although both correlations are negative, the P value (0.19802) is insignificant.

Similar findings can be seen in the Google results. For Google, the Spearman and Pearson correlation coefficients for the positive sentiment group are -0.349209 and -0.1355227 and the p value for Pearson analysis 0.0008319 is significant.

The TextBlob findings repeat the same outcome, as the Spearman and Person correlations are -0.323644 and 0.103500 and the p value for Pearson is 0.0108542.

For Flair, in the positive sentiment category, the Spearman correlation has mixed findings as Pearson correlation coefficient repeats the same outcome as above as it is -0.0901801 and has significant p value of 0.0265499. However, in the negative sentiment category the Pearson Correlation coefficient is negative and has significant p value meaning lesser the negativity higher is the monetary value attached. The finding actually supports the overall finding about monetary value being associated with more positive (less risky) posts.

The results for VADER seem to represent the only anomalous findings in our study. For VADER, in both the positive and negative sentiment category all the correlations between dollar value and sentiment are positive regardless of the positive or negative classification and the p values are significant. Overall, VADER shows monetary value being associated with heightened sentiments in either direction positive or negative.

Overall, what we find essentially, the lesser the positivity of sentiment, the higher the dollar amount. This is true only for the positive sentiment samples, not the negative sample data, and valid for four of the five major sentiment analyzers (except VADER). To summarize, we see less positivity in the positive sentiment category correlated with higher dollar values.

Intuitively, people pay relatively more attention to overall positive information, however, once the post has their attention the magnitude of positivity seems to direct the dollar amount associated. Basically, less positive posts are able to garner more dollars. This finding resonates well with risk aversion behavior of a rational investor as the lesser the positivity of the post the higher the dollar amount is according to our analyses.

These findings may help posters in the network in understanding how network participants are responding to sentiment in post content. A naive poster may intuitively believe that greater positive sentiment in content will result in higher post valuations and a larger network following, leading to overexuberance in post sentiment. A clearer understanding of how network participants react to the extent of positivity in a post could help posters to create more informed posts.

## 4. Conclusion:

In this work, we examined the relationship between post sentiment and content valuation in social media posts. We examined a dataset of posts from the Steem social media network, where participants are encouraged to provide high-quality content with SteemCoin payments that increase in value the greater the positive feedback from the network. We analyzed the posts for sentiment using multiple sentiment analysis tools to demonstrate consistency of results. Based on the sentiment analysis and corresponding post valuation on the Steem network, we considered the relationship between sentiment and valuation. Our results show that posts containing marginally positive content sentiment are more strongly correlated with stronger valuation among network participants, compared to the valuation of posts with greater positive sentiment.

## References

Arrow, K. J. (1965). "Aspects of the Theory of Risk Bearing". *The Theory of Risk Aversion*. Helsinki: Yrjo Jahnssonin Saatio. Reprinted in: *Essays in the Theory of Risk Bearing*, Markham Publ. Co., Chicago, 1971, 90–109.

Bartov, E., Faurel, L. and Mohanram, P.S., (2018). Can Twitter help predict firm-level earnings and stock returns?. *Accounting Review* 93(3), 25-57.

Chen, H., De, P., Hu, Y.J. and Hwang, B.H., (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies 27*(5), 1367-1403.

Chao Li, Balaji Palanisamy, Runhua Xu, Runhua Xu and Jingzhe Wang (2021). "**SteemOps: Extracting and Analyzing Key Operations in Steemit Blockchain-based Social Media Platform**." Proc. of 11th ACM Conference on Data and Application Security and Privacy (**ACM CODASPY'21**), Virtual Event, USA,.

Da, Z. and Huang, X., (2020). Harnessing the wisdom of crowds. *Management Science* 66(5), 1847-1867.

Fang, X., Zhan, J. Sentiment analysis using product review data. (2015) Journal of Big Data 2, 5. https://doi.org/10.1186/s40537-015-0015-2

Jame, R., Johnston, R., Markov, S. and Wolfe, M.,(2016). The value of crowdsourced earnings forecasts. *Journal of Accounting Research* 54(4), 1077-1110.

Jame, R., Markov, S. and Wolfe, M., (2022). Can fintech competition improve sell-side research quality? *Accounting Review* 97(4), 287-316.

Pratt, John W. (January 1964). "Risk Aversion in the Small and in the Large". *Econometrica*. **32** (1/2): 122–136. doi:10.2307/1913738. JSTOR 1913738.

Ren, R., Wu, D. D., and Liu, T. (2019) "Forecasting Stock Market Movement Direction Using Sentiment Analysis and Support Vector Machine," IEEE Systems Journal, 13(1), 760-770, doi: 10.1109/JSYST.2018.2794462.

Smailović, J., Grčar, M., Lavrač, N., Žnidaršič, M. (2013). Predictive Sentiment Analysis of Tweets: A Stock Market Application. In: Holzinger, A., Pasi, G. (eds) Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data. HCI-KDD 2013. Lecture Notes in Computer Science, vol 7947. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39146-0_8

Valdivia, A., Hrabova, E., Chaturvedi, I., Luzón, M. V., Troiano, L., Cambria, E., Herrera, F., (2019). Inconsistencies on TripAdvisor reviews: A unified index between users and Sentiment Analysis Methods. Neurocomputing 353, 3-16. https://doi.org/10.1016/j.neucom.2018.09.096.

Zvarevashe, K. and Olugbara O. O., (2018) A framework for sentiment analysis with opinion mining of hotel reviews, 2018 Conference on Information Communications Technology and Society (ICTAS), Durban, South Africa, 2018, pp. 1-4, doi: 10.1109/ICTAS.2018.8368746.

Sentiment analysis references:

VADER: https://github.com/cjhutto/vaderSentiment

Google: https://cloud.google.com/natural-language

IBM Watson: https://www.ibm.com/cloud/watson-natural-language-understanding

TextBlob: https://textblob.readthedocs.io/en/dev/index.html

Flair: https://github.com/flairNLP/flair