



ABDUL KHALIQ
AND
NS RAGHAV

Money is all we want to make

An analysis of Reddit comments and replies leading
to Bitcoin's All Time High (now previous ATH)



Bitcoin is a decentralised currency that operates transactions over a blockchain.

Bitcoin has seen mass adoptance and popularity surges in the recent years, especially the 2017 boom. In the year 2020, Bitcoin has gone from a low of \$4679 to a current trading price of around \$20,500 owing to people having lost trust in regulated and centralised currencies due to the Covid-19 Pandemic.

Our Bitcoin Story. and Motivation



Can we use social media data to gauge at Bitcoin Price ?

An attempt at using sentiment analysis
on Bitcoin Data



Our Dataset

r/BitcoinMarkets

Our Dataset is based on Daily Discussion Thread of a popular Bitcoin related subreddit. Inclusive of comments and its replies.

r/Bitcoin Markets

Daily Discussion Thread

- Our dataset comprises of daily discussions on this thread
- The dataset we will experiment with comprises of all comments and replies for these daily discussion thread from Nov,1, 2020 to Dec, 1, 2020
- We chose this period since it coincides with Bitcoin reaching its all time high on Nov. 28
- This would provide a good analysis of the sentiment trends before and after reaching all-time high.

↑ r/BitcoinMarkets - Posted by u/AutoModerator \$17991.00 Offline €14826.90 ¥117441.95 49 minutes ago

↓ [Daily Discussion] Friday, December 11, 2020

Daily Discussion

Thread topics include, but are not limited to:

- General discussion related to the day's events
- Technical analysis, trading ideas & strategies
- Quick questions that do not warrant a separate post

Thread guidelines:

- Be excellent to each other.
- Do not make posts outside of the daily thread for the topics mentioned above.

Other ways to interact:

- Get an invite to live chat on our Slack group
- Set your flair to trade pretend money in the Flair Game
- Weekly threads: Newbie Mondays and Fundamentals Fridays
- Altcoins Discussion: Altcoin Discussion

3 Comments Share Save Hide Report

100% Upvoted

↑ InDebtBigTime Bullish 1 point · 22 minutes ago

↓ I'm going to watch this one play out for the next few days before taking a position again. I can't make heads or tails out of this, but I'm expecting a trend downward to 16K

↑ diydude2 4 points · 37 minutes ago

↓ A truly massive dump would take us to 16.5K.

Six months ago we were in four figures.

Do the math. There is no other place to put your money now except gold, lead, or land. Hydrogen fuel cell technology and especially companies leading the way in decentralized microgrid tech might be wise, but I'd steer clear of biotech for karma/ethical reasons and because that is a ship that already hit the iceberg.

BTC FTW.

↑ jarederaj 2013 Veteran 2 points · 28 minutes ago

↓ Consolidation was required. The fact that we haven't corrected back to 14k after the months we've had is what's surprising.

The new year will bring new money. December is typically a consolidation month, anyway.

All I can is business as usual: move up, consolidate, repeat.

DECEMBER 2020

Presentation Flow

Content

- 01 **Polarity Analysis**
- 02 Fear vs Greed
- 03 Papers and History
- 04 Research Potential
- 05 Challenges
- 06 Drawbacks



Polarity

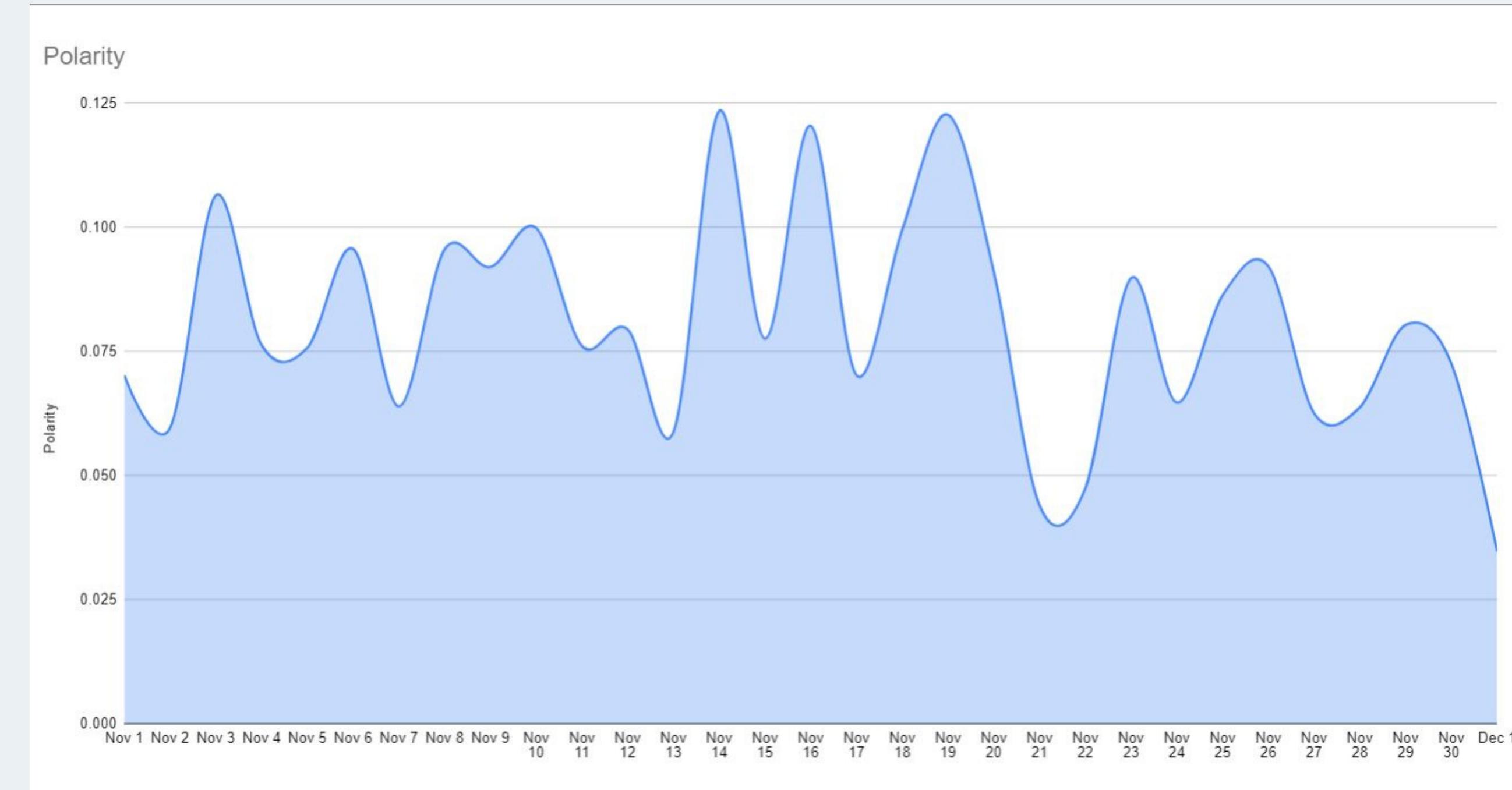
What is Polarity

Polarity is the measure of positiveness and negativeness in a statement and a sentence. Polarity is measured on a scale of -1 to +1 with +1 being positive and -1 and negative.

Bitcoin Polarity (TextBlob)

Upon Reddit Data from
Nov 1 - Dec 1, 2020

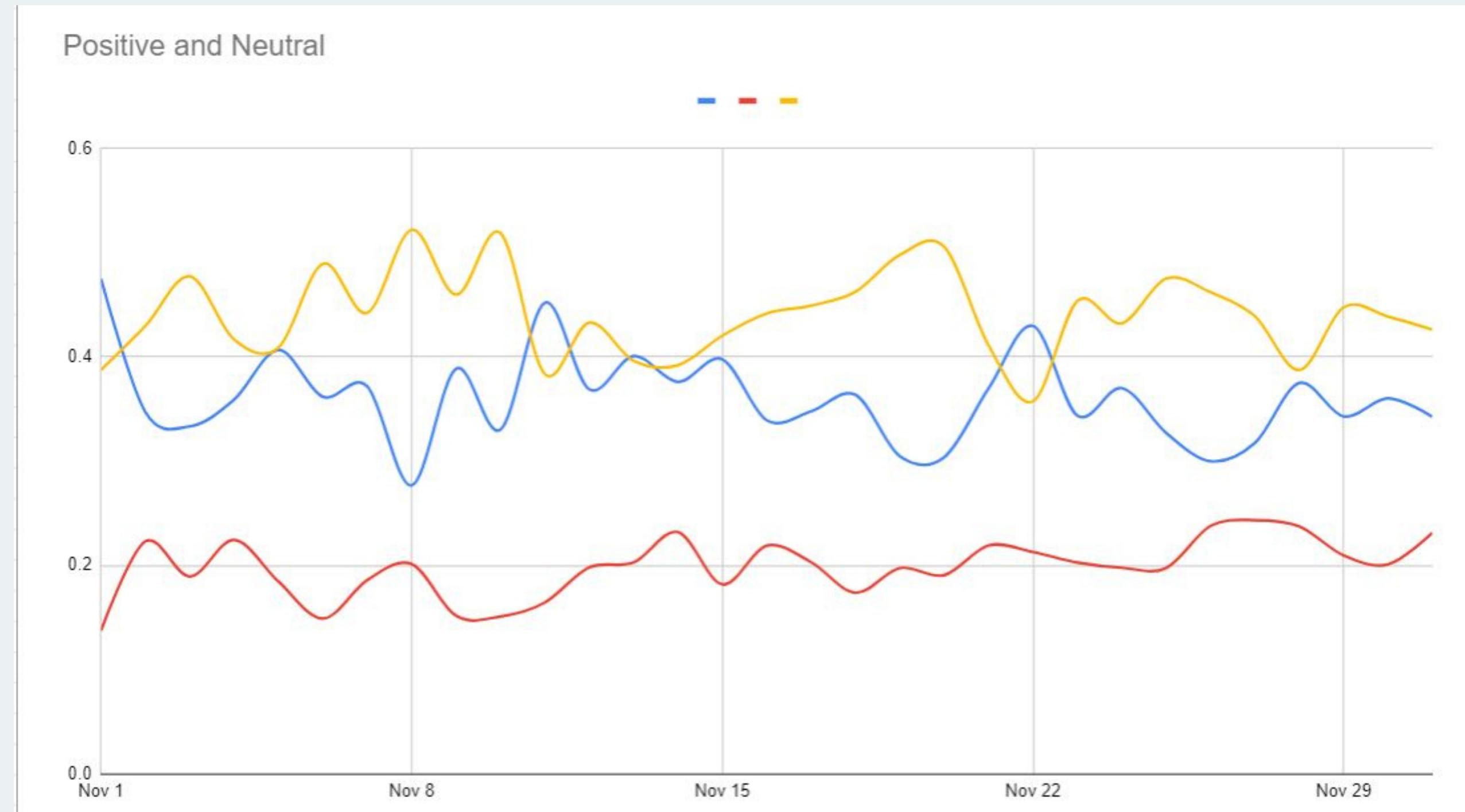
- For November, mean polarity lies in the range of 0.032 (min) and 0.122 (max).
- As can be realised from this data, the polarity of bitcoin related comments and replies remains relatively neutral and slightly skewed to the positive side.
- The 3 large spikes between Nov 14 to Nov 21 coincide with Bitcoin breaking the \$18,000 mark.
- This is followed by a steep decline in polarity which arises from the fear mindset of Bitcoin retaining its \$18,000 mark for longer than usual.



Daily Polarity Distribution

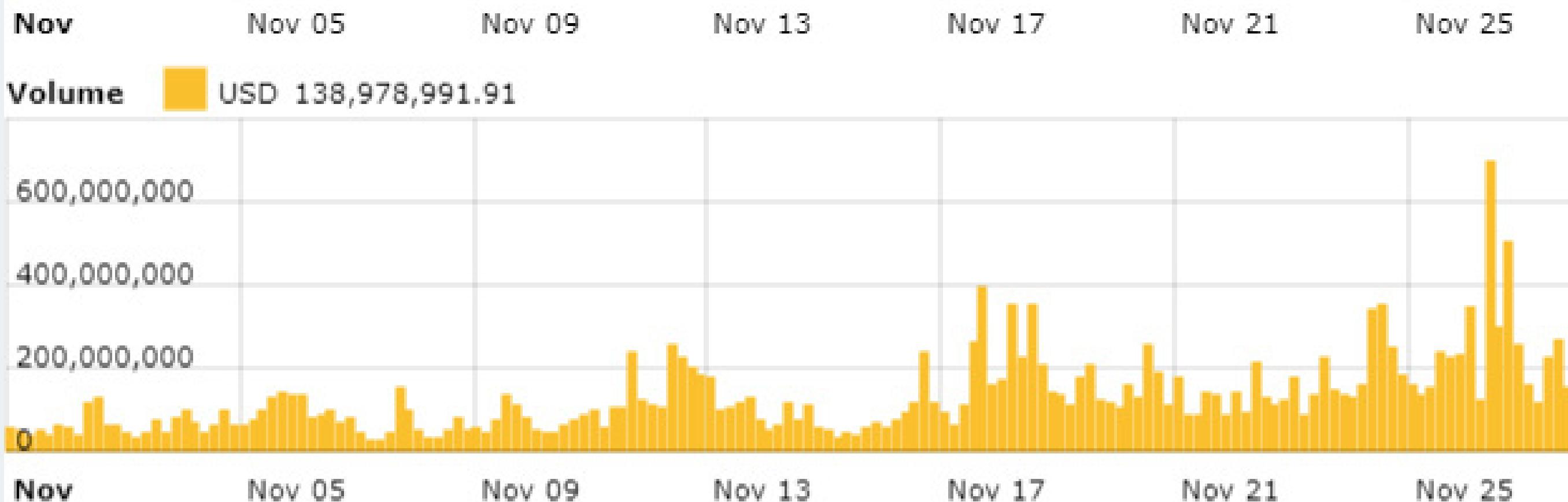
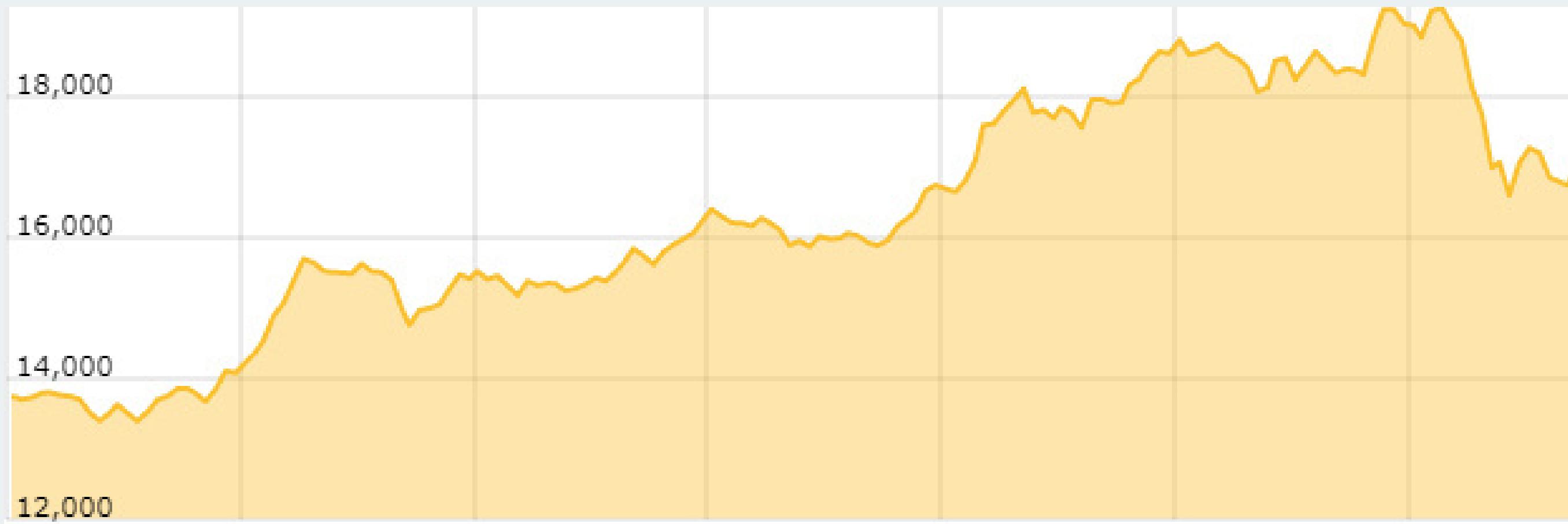
Upon Reddit Data from Nov 1 - Dec 1, 2020

- Fluctuations in Neutral opinions has interestingly been the most correlated data with Bitcoin Prices for each day.
 - Positive and Negative sentiments regarding Bitcoin are exclusive to each other during this period. The rise of one kind of sentiment leads to a fall in the other. There have been far less speculative opinions during this month (bullrun ?)
 - Negative emotions regarding bitcoin show steady rise when bitcoin approaches all time high. This is due to its sudden bullish nature with no corrections that speculated analysts to see a massive Bitcoin price drop.
 - The most erratic data has been the positive comments, which ran frenzy amongst holders during this bull run.



Bitcoin Price Chart

Nov 1 - Dec 1





Subjectivity.

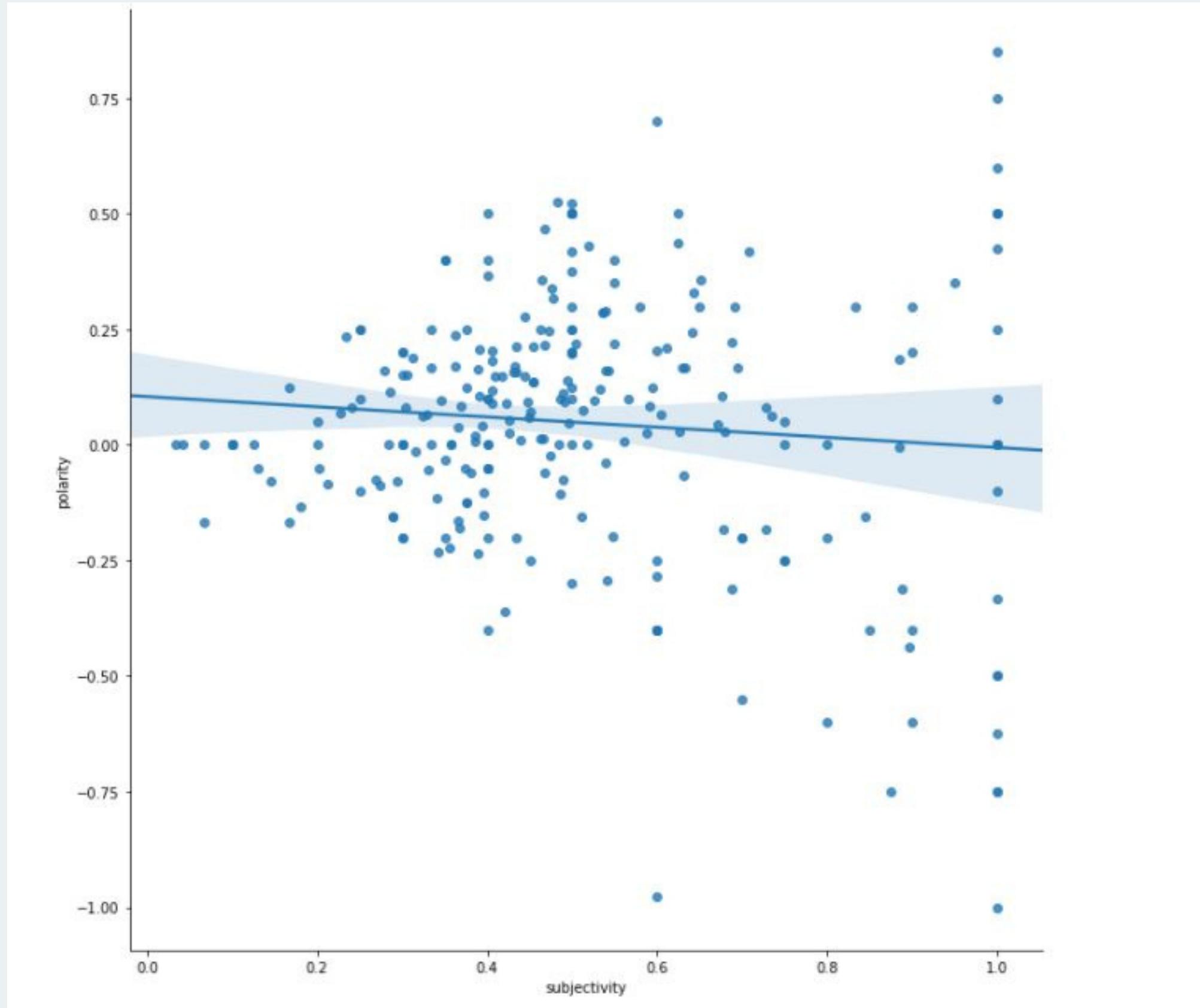
What is Subjectivity ?

Subjectivity is the measure of opinionated vs analysed comments. It measures whether a comment is subjective to the person commenting or is an analysed comment which is not as opinionated. It is measured on a scale of 0 to 1, with 0 being the least subjective and 1 being the most subjective.

Subjectivity vs Polarity

Upon Reddit Data from
Nov 1 - Dec 1, 2020

- As can be seen from the trend across Reddit, the graph has a mean subjectivity of around 0.5 which is a good sign.
- Comments that are less subjective also tend to be less polar often sticking close to 0. This is due to the balanced stances of such comments.
- As subjectivity increases beyond 0.5, the polarity values spread towards the extremes either being highly positive or highly negative, showcasing the personal viewpoint of a trader with respect to the market.



Presentation Flow

Content

- 01 Polarity Analysis
- 02 **Fear vs Greed**
- 03 Papers and History
- 04 Research Potential
- 05 Challenges
- 06 Drawbacks



Fear vs Greed

What is Fear and Greed in Stock Market ?

Fear and greed are the sentiment metrics used in Stock Markets to determine the sentiment of the traders. A fearful market would mean that traders assume the prices are going to drop and are in a rush to sell their assets, whereas a greedy market means traders predict the market to rise and are in a rush to buy.

How do we look at Fear v Greed ?

Upon Reddit Data from
Nov 1 - Dec 1, 2020

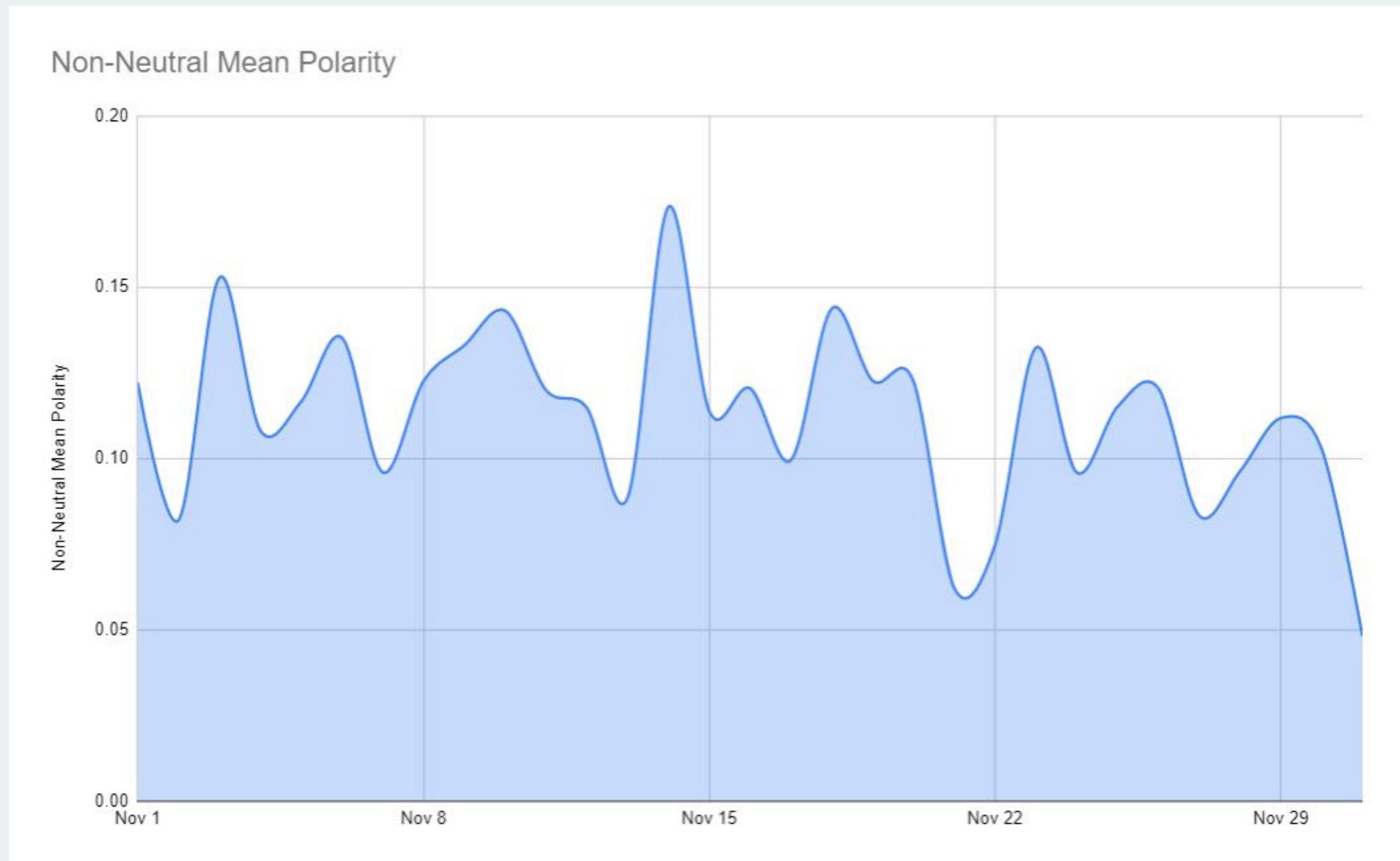
- We use our polarity values and remove the neutral polarities.
- This leaves us only with those opinions that are either positive or negative.
- We use these values to gauge at Fear v Greed on a day to day basis for the month of November.

Polarity	Non-Neutral Mean Polarity
0.070235	0.122148
0.059658	0.082029
0.106494	0.152795
0.076321	0.108236
0.075777	0.116558
0.095717	0.135299
0.063973	0.095959
0.095625	0.122616
0.092055	0.133038
0.099923	0.143133
0.076123	0.119935
0.079489	0.114857
0.058808	0.088621
0.123489	0.17344
0.077589	0.113798
0.120495	0.120495
0.070243	0.099553
0.099574	0.143786
0.122714	0.122714
0.091139	0.122648
0.04404	0.062224
0.047435	0.074813
0.089795	0.13236
0.064766	0.095929
0.086359	0.115146
0.092187	0.120528
0.06259	0.083307
0.063778	0.096269
0.080409	0.111827
0.072589	0.102834
0.034739	0.048152

Fear v Greed and its inferences.

Upon Reddit Data from
Nov 1 - Dec 1, 2020

- We see polarity waver across the month of November quite drastically.
- We can find reason for Nov 14- Nov15 high and the Nov21 - Nov 22 low in polarity.
- Nov14 high can be attributed to rising bitcoin prices breaking \$18,000 and forming a upward triangle on graph.
- The dip on Nov 21 can be attributed to skeptical concerns of traders for Bitcoin making a sharp rise without any corrections or stabilisations in prices.





But is all this conclusive ?

Has the dataset for the month provided
useful results?

As you may be thinking, was all this data concrete. Were we able to concretely figure any correlations or were they extrapolated inferences. Lets analyse the prior work done in the field to understand the field of Economics and NLP to figure what could be done better

Presentation Flow

Content

- 01 Polarity Analysis
- 02 Fear vs Greed
- 03 Papers and History**
- 04 Research Potential
- 05 Challenges
- 06 Drawbacks



Papers and History

Has this been done before?

Yes. Numerous times. Bitcoin is related to money. Money is something people are drawn to for some inexplicable reason. ^_(ツ)_/^\n

What did we do different?

We looked at a different source of data.

Algorithmic Trading of Cryptocurrency Based on Twitter Sentiment Analysis

Stuart Colianni, Stephanie Rosales, and Michael Signorotti (2015)

- Tweets
- text classification and sentiment analysis algorithms
- After creating the feature vector, this is tested against Naive Bayes, Logistic Regression and SVMs
- hour-to-hour data & day-to-day data
- First (words in tweets as elements of vector):
 - Bernoulli Naive Bayes - day-to-day - 95%
 - Bernoulli Naive Bayes - hour-to-hour - 76.23%
- Second (calculating negativity, neutrality, positivity):
 - Logistic Regression - day-to-day - 86%
 - Logistic Regression - hour-to-hour - 98.58%

$$x = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ bright \\ \vdots \\ economy \\ \vdots \\ future \\ \vdots \\ 0 \\ \vdots \\ worlds \\ \vdots \\ 0 \end{bmatrix}$$

Using sentiment analysis to predict interday Bitcoin price movements

Vytautas Karalevicius, Niels Degrande and Jochen De Weerdt (2017)

- Tweets
- Built their own sentiment analysis program
 - four word window approach
 - list of negators, list of intensifiers
- Finance focused lexicon created by Loughran and McDonald (2011)
- identified that interaction between media sentiment and the Bitcoin price exists
 - there is a tendency for investors to overreact on news in a short period of time

JRF
19,1

64

Table I.
Performance of the
interday sentiment
strategy using
Harvard
Psychosocial
Dictionary

	Start	End	Days	Articles	Start price	End price	Annual return	Sharpe	Meaning
01/01/14	31/12/15	730	11,488	754.22	430.89	0.45	1.14	Good	
01/01/14	31/12/14	365	5,689	754.22	321.00	0.70	1.49	Good	
01/01/15	31/12/15	365	5,799	313.81	430.89	0.10	0.45	Poor	
01/01/14	30/06/14	181	3,192	754.22	641.11	0.60	1.14	Good	
01/07/14	31/12/14	184	2,497	638.25	321.00	0.99	2.85	Very good	
01/01/15	30/06/15	181	3,010	313.81	262.50	-0.26	-0.58	Poor	
01/07/15	31/12/15	184	2,789	257.62	430.89	0.87	1.95	Good	
01/01/14	31/03/14	90	1,449	754.22	454.83	-0.60	-1.27	Poor	
01/04/14	30/06/14	91	1,743	478.98	641.11	3.38	3.08	Excellent	
01/07/14	30/09/14	92	1,363	638.25	391.00	0.51	1.91	Good	
01/10/14	31/12/14	92	1,134	384.86	321.00	0.88	2.61	Very good	
01/01/15	31/03/15	90	1,458	313.81	244.24	-0.67	-2.00	Poor	
01/04/15	30/06/15	91	1,552	246.69	262.50	0.23	1.03	Good	
01/07/15	30/09/15	92	1,573	257.62	236.20	0.03	0.22	Poor	
01/10/15	31/12/15	92	1,216	237.15	430.89	1.55	2.46	Very good	
01/01/16	31/03/16	91	1,248	433.82	415.35	-0.36	-1.37	Poor	
06/05/14	24/06/14	50	943	428.01	572.62	0.20	0.66	Poor	

Notes: All returns are compounded. Annual return shows the annualized return. For example, for the first quarter of 2016, the strategy would have yielded 36% annual loss. The Sharpe ratio of -1.37 indicates that we are taking excessive risks

Table II.
Performance of the
interday sentiment
strategy using
Loughran–McDonald
finance-specific
dictionary

	Start	End	Days	Articles	Start price	End price	Annual return	Sharpe	Meaning
01/01/14	31/12/15	730	11,488	754.22	430.89	0.82	1.84	Good	
01/01/14	31/12/14	365	5,689	754.22	321.00	1.28	2.23	Very good	
01/01/15	31/12/15	365	5,799	313.81	430.89	0.32	1.03	Good	
01/01/14	30/06/14	181	3,192	754.22	641.11	0.77	1.47	Good	
01/07/14	31/12/14	184	2,497	638.25	321.00	2.06	3.31	Excellent	
01/01/15	30/06/15	181	3,010	313.81	262.50	-0.15	-0.28	Poor	
01/07/15	31/12/15	184	2,789	257.62	430.89	1.10	2.62	Very good	
01/01/14	31/03/14	90	1,449	754.22	454.83	1.20	1.99	Good	
01/04/14	30/06/14	91	1,743	478.98	641.11	1.94	2.39	Very good	
01/07/14	30/09/14	92	1,363	638.25	391.00	-0.40	-3.14	Poor	
01/10/14	31/12/14	92	1,134	384.86	321.00	9.45	5.46	Excellent	
01/01/15	31/03/15	90	1,458	313.81	244.24	-0.07	0.07	Poor	
01/04/15	30/06/15	91	1,552	246.69	262.50	-0.17	-0.61	Poor	
01/07/15	30/09/15	92	1,573	257.62	236.20	1.85	4.16	Excellent	
01/10/15	31/12/15	92	1,216	237.15	430.89	0.59	1.55	Good	
01/01/16	31/03/16	91	1,248	433.82	415.35	1.71	2.51	Very good	
06/05/14	24/06/14	50	943	428.01	572.62	0.01	0.16	Poor	

Notes: All returns are compounded. Annual return shows the annualized return. For example, for the first trimester of 2016, the strategy would have yielded more than 171% annual return. The Sharpe ratio of 2.51 indicates that we do not take excessive risks

Extracting Cryptocurrency Price Movements From the Reddit Network Sentiment

Stephen Wooley, Andrew Edmonds, Arunkumar Bagavathi, Siddharth Krishnan (2019)

- A List of 24 subreddits
- A rank of top 100 users
- sentiment time series, average node degree, and average clustering coefficient features are shown to be correlated with observed price changes.

Presentation Flow

Content

- 01 Polarity Analysis
- 02 Fear vs Greed
- 03 Papers and History
- 04 **Research Potential**
- 05 Challenges
- 06 Drawbacks



Research Potential

Has this been done before?

Yes. Numerous times. Bitcoin is related to money. Money is something people are drawn to for some inexplicable reason. ^_(ツ)_/^-

Using a finance specific dictionary.

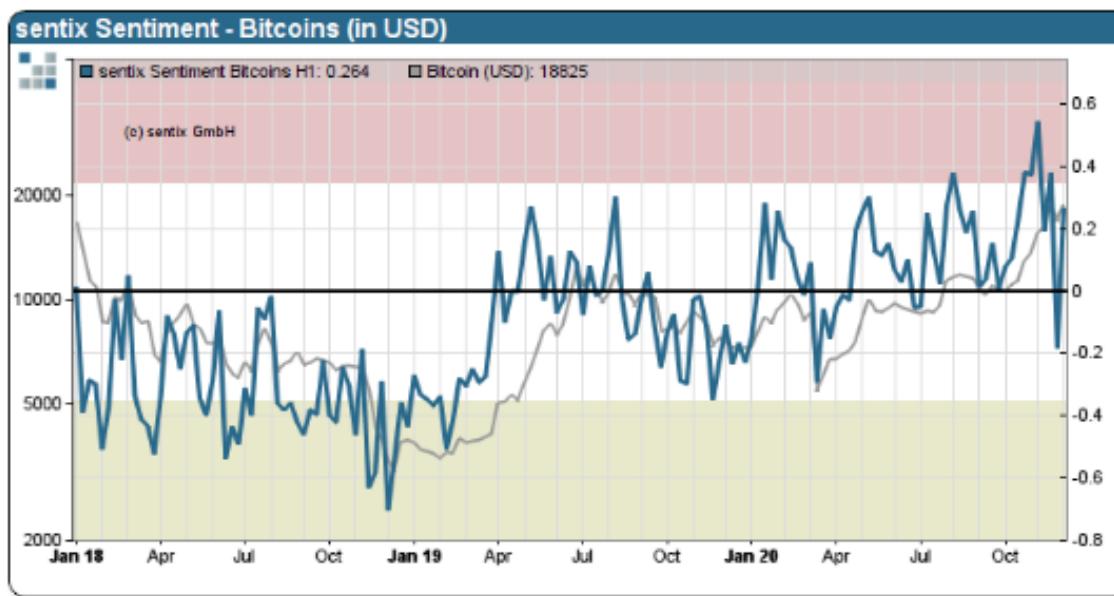
- Finance focused lexicon created by Loughran and McDonald (2011)
 - Can we expand it to Bitcoin and Altcoins?
- Use this for Sentiment Analysis

Sources of Data

- There are multiple forums on the net
 - IRC
 - Reddit
 - Twitter
 - Facebook Groups

The Actual Predictor

- Can we use Deep Learning instead of traditionally used SVMs, Decision Trees & Other Classifiers?



Bitcoin sentiment

The sentix Sentiment, investors one-month expectations for a given market, reflects the emotions – between greed and fear – of market participants. Negative sentiment extremes are usually a straight indication for rising prices. High optimism may be a warning signal for an upcoming market consolidation. A sentiment divergence mostly indicates more important turning points.

[sentix compendium](#)

whereas very high optimism is often a warning sign of impending price weakness. The sentiment indices should therefore be interpreted in a contrary way.
 Longer-term reversal points are indicated by the Sentiment in combination with Strategic Bias. The Strategic Bias (as a 6-month expectation) should point in the opposite direction of the Sentiment. The sentix Sentiment is obtained from the weekly sentix Survey (sentix Global Investor Survey) of currently more than 5,000 private and institutional investors, whose estimations of 14 financial markets are collected ([How the survey works](#)).

- [How to use the indicator](#)
- [Figures - Examples of forecast successes](#)
- [Indicator construction](#)
- [Descriptive statistics](#)

Im Paket

Complete, sentix Market sentiment (SNMS)

Anwendung

- Euro area Equities
- German Equities
- US Equities
- Japan Equities

[Other countries](#)

How to use the indicator

search...

Thursday, 10 December 2020

 Log-In / Log-Out

5.0/5 rating (4 votes)



Username

Password

Remember me

[Forgot login?](#)

[Register](#)

Indicator worldmap



Database brochure



Presentation Flow

Content

- 01 Polarity Analysis
- 02 Fear vs Greed
- 03 Papers and History
- 04 Research Potential
- 05 **Challenges**
- 06 Drawbacks

Cleaning the Data & General NLP Problems

- SO MUCH SPAM AND ADVERTISING!

You Must Know To Make #Money #Trading #Cryptocurrency #TMSmart <https://t.co/aQq1RBB3vq> #bitcoin'

FOLLOW, TO SEE DEPTH \n\n#crypto #influencermarketing #womenshistorymonth #BlackHistoryMonth #cryptocurrency #bitcoin\xe2\x80\xa6
<https://t.co/GgoBHT6lrQ>"

- Sarcasm, Trolling, Metaphors

Sentiment Analysis on Financial Terminology and slang

- Financial slang terms like "pump and dump", "bullish and bearish", "moonshot" as well as many rocket emojis involved in social media discourse needs to be accounted for during sentiment analysis.

Bitcoin is hypervolatile

- Bitcoin is termed as hypervolatile, which is very different from the normal stock market.
- A small news heading could impact the rate of BTC on exchanges very severely.

Needs lots of data to show correlation

- As we realised, lots of data across the years and lots of parameters are needed to accurately predict and show correlation between traders and market price.

Presentation Flow

Content

- 01 Polarity Analysis
- 02 Fear vs Greed
- 03 Papers and History
- 04 Research Potential
- 05 Challenges
- 06 **Drawbacks**

TextBlob doesn't perform well with negative sentiments

- Compared to NLTK Vader, TextBlob performs slightly poorly in identifying negative sentiments in statements. This might lead to results favoring more positive polarity and can skew the results.

Predictions doesn't equate to causality_(especially_for Bitcoin).

- Being hypervolatile, its a very tough task to accurately predict the growth of Bitcoin. Lots of unattended variables leads to Bitcoin making uncalled for dumps and pumps which is hard to model.



Again, Bitcoin is hypervolatile

- No. Seriously. Check out the graph on coindesk to see how much the exchange moves in less than 24 hours.



Thank You!

Ask us questions