



singularitydao

Sentiment and Cognitive Distortions Mining in Social Media

**Mukul Vishwas, Ali Raheman, Anton
Kolonin, Igors Fridkins, Ikram
Ansari, Juan Pinzon, Alice Ho**
(Based on work done at SingularityDAO and Antonio
Foundation)

mukul.vishwas@singularitydao.ai

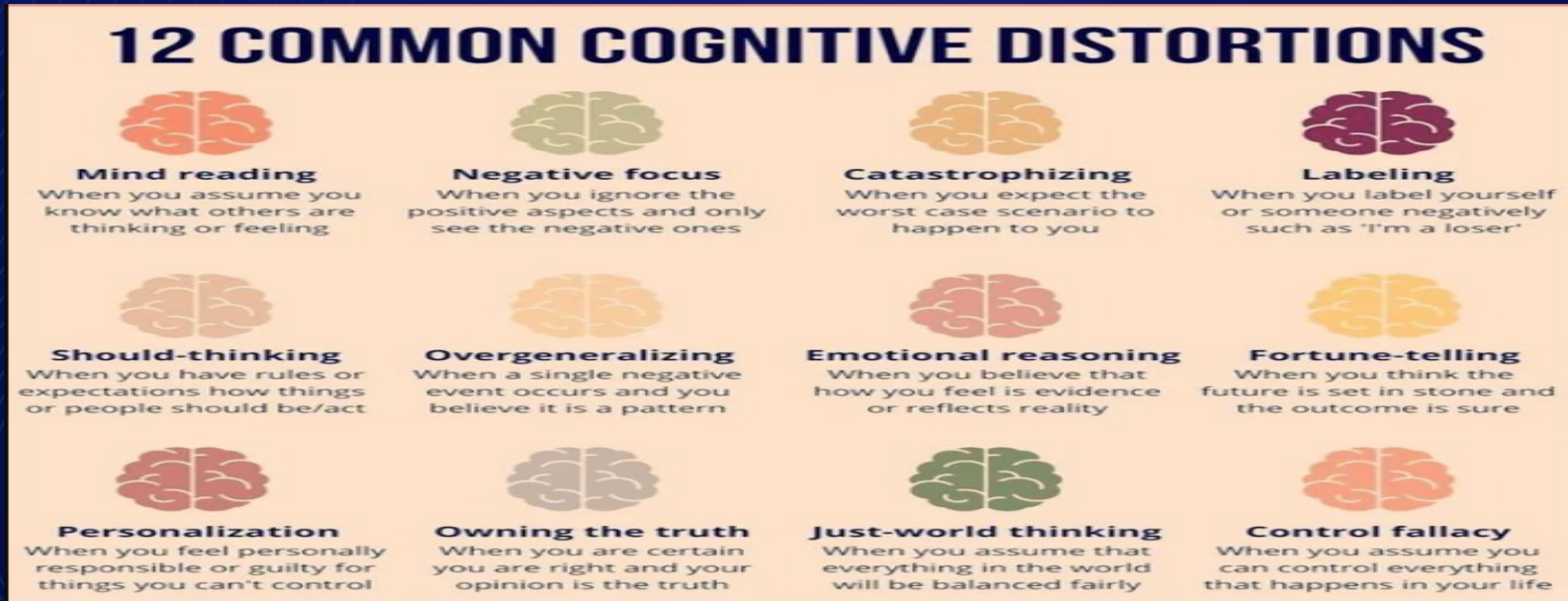
anton@singularitynet.io

Sentiment and Cognitive Distortions Mining in Social Media

- GOAL: Analysis of social media sentiment (polarity) and cognitive distortions for the purpose of crypto market prediction.
 - **Collecting social media feeds** from about 77 feeds on Twitter and Reddit, then computing about 19 metrics for them around the clock.
 - **Using prior work** “Historical language records reveal a surge of cognitive distortions in recent decades” for cognitive distortions mining model. <https://doi.org/10.1073/pnas.2102068881>
 - **Exploring different open source sentiment analysis solutions** and choosing Aigents® model because of its Interpretable nature, Its more practical to use and easy to adjust model performance with domain-specific corpus.
 - **Papers are published:** <https://arxiv.org/pdf/2204.10185> and <https://arxiv.org/pdf/2204.12928>
 - **Analyzing the potential causal connection** between different metrics and the price movement.

Cognitive Distortions, Cognitive-Behavioral Schema (CBS)

- Text mining model based on 'n-grams' (lexemes)
- Examples:
 - Catastrophizing: Exaggerating the importance of negative events (will fail, will go wrong..)
 - Fortune-telling: Making predictions, usually negative ones, about the future (I will not, we will not...)
 - Mindreading: Believing you know what others are thinking (everyone knows, everyone thinks...)



Aigents®

- Aigents® is an “interpretable” model based on “n-grams” (lexemes), available as part of <https://github.com/aigents/aigents-java> distribution, and written in Java.
- It is based on positive and negative dictionaries of words and lexemes (n-grams) and has over 8,200 negative and 3,800 positive n-grams.
- The “priority on order” feature helps the model to get a more reliable sentiment score.

<https://github.com/aigents/aigents-java/blob/master/src/main/java/net/webstructor/data/LangPack.java#L363>

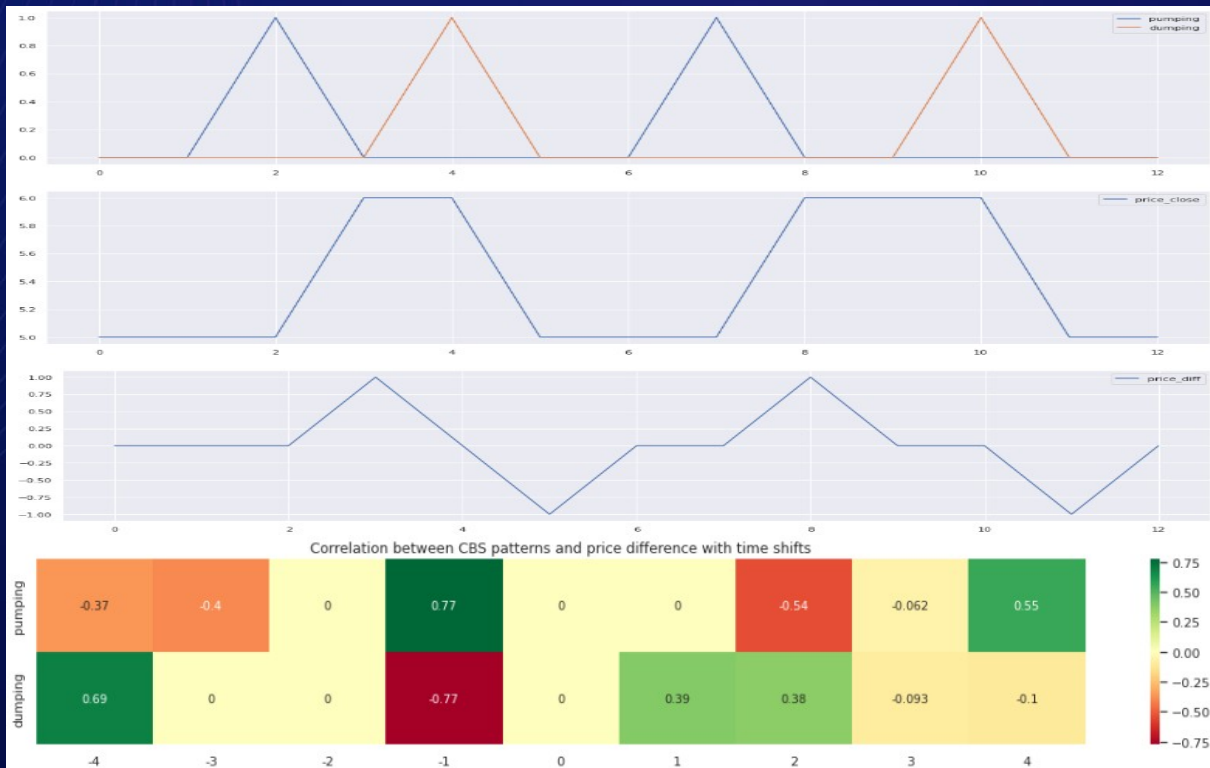
Aigents[®] Sentiment metrics

Example	Positive (P)	Negative (N)	Seniment (P+N)	Contradictive $\text{SQRT}(P*N)$
The weather is good.	1	0	1	0
The movie was bad.	0	-1	-1	0
Honestly, I don't like Crypto stuff but I am happy that it makes me richer.	1	-1	0	1

Aigents[®] Priority on order

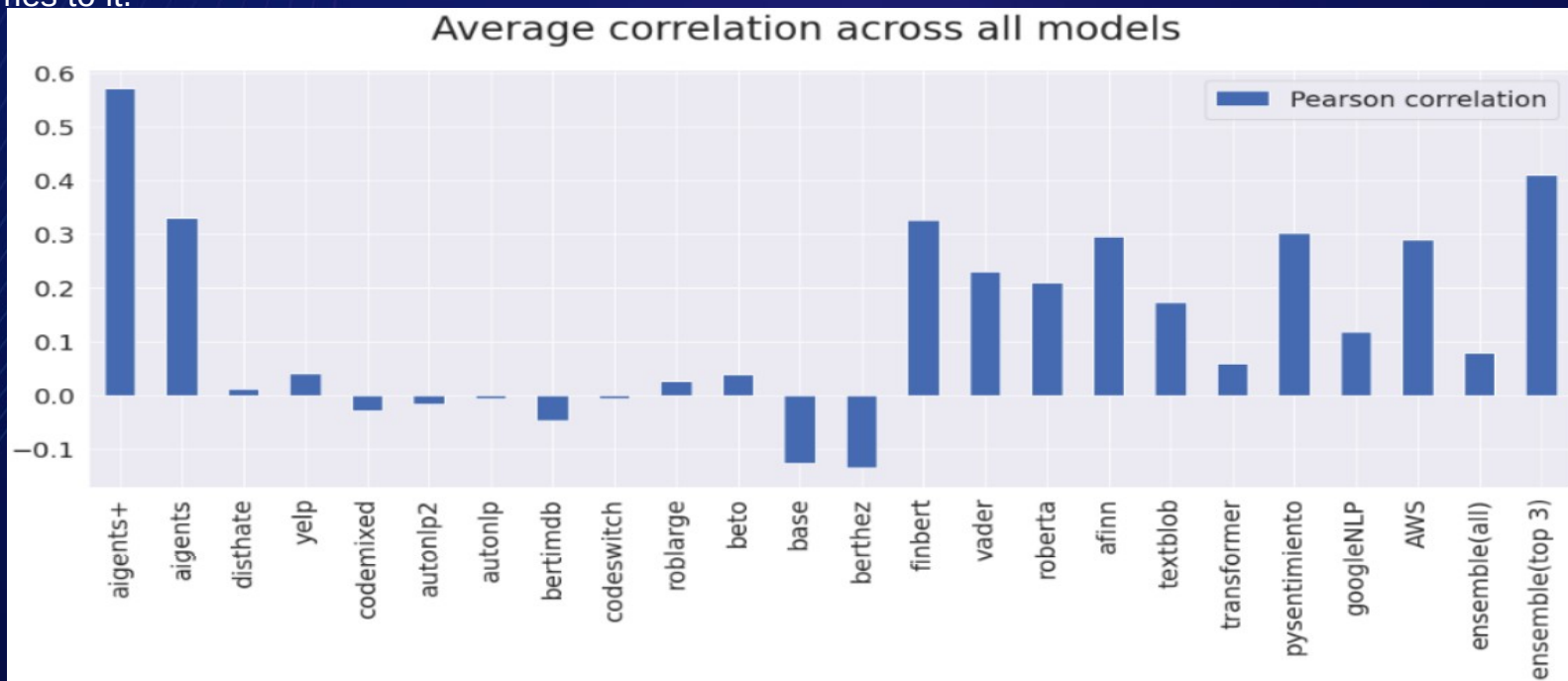
- Aigents[®] is a '**n-gram**' (lexemes) based algorithm.
- In the 'Priority on order' method precedence is given for n-grams with higher order "n". For example if tetragram ["not a bad thing"] matched then lower 'n-grams' like ['bad thing'], ['not'] and ['bad'] inside of it are discarded.
- Using this method resolves the **negation** problem (not good, not bad).

Mining Causal Connections Between Social Media Metrics and Price Movement



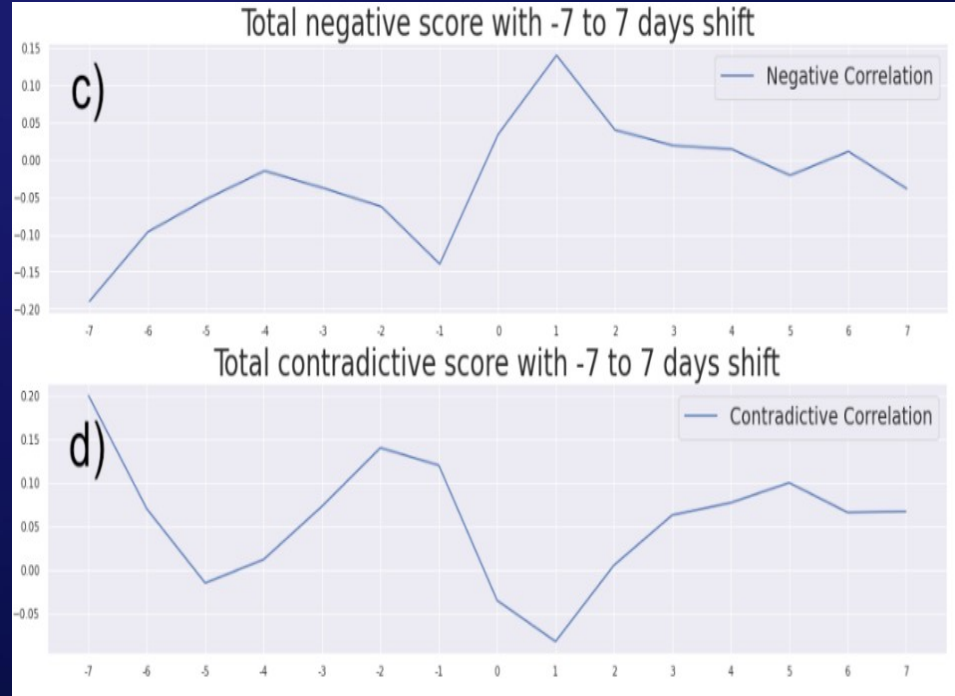
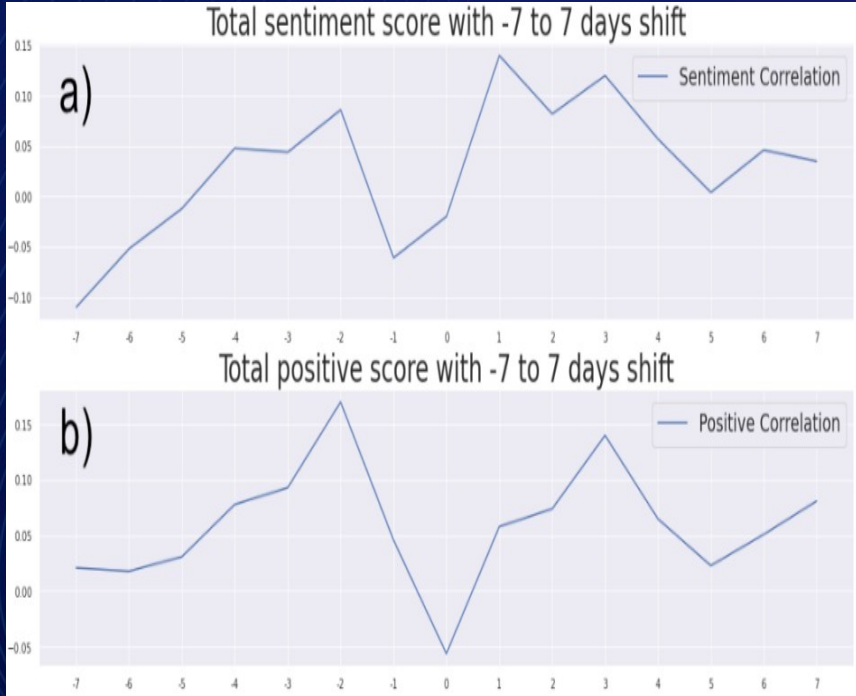
Results: Finding the best sentiment analysis model

- Below is the Pearson correlation between outputs provided by different models and manual sentiment markup in the test corpus, the Aigents® model outperformed all others.
- Aigents®+ model is the extended version of the Aigents® baseline one, we have added crypto domain-specific lexemes to it.



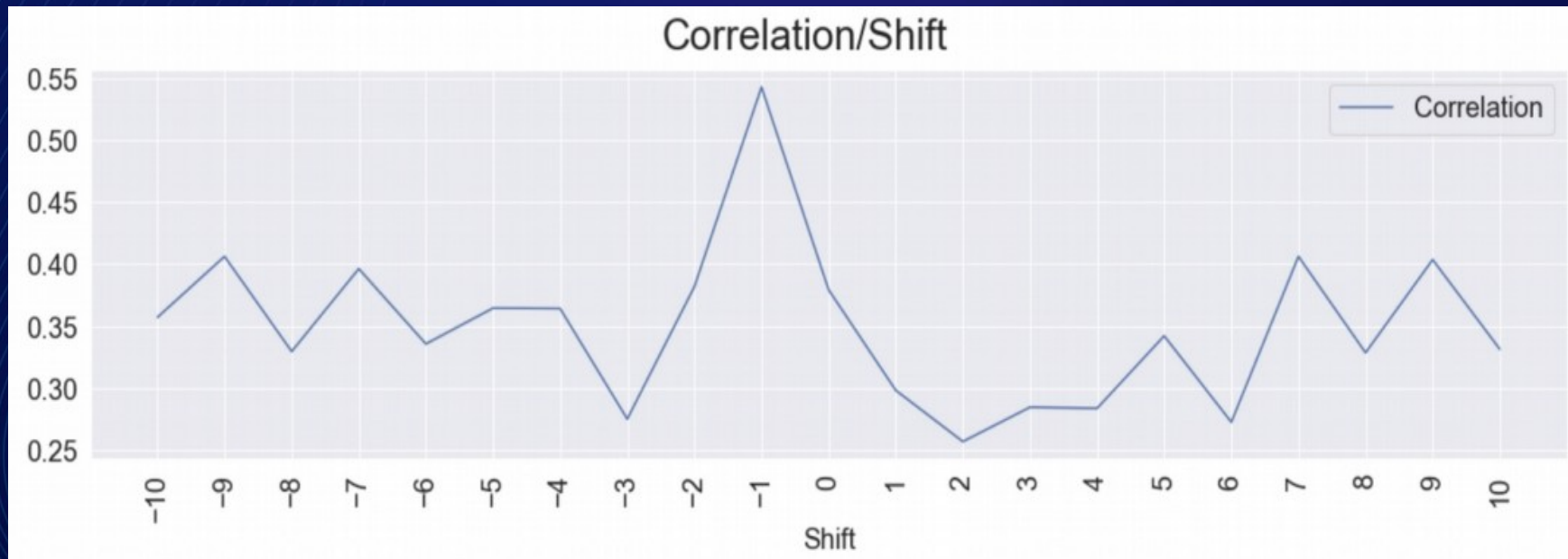
Results: Causal connections between Sentiment and BTC/USD price difference

- Temporal correlation analysis for different sentiment metrics with mutual Pearson correlation computed between
- the daily Bitcoin price difference (derivative) and respective metrics over six months from July to December 2021.



Results: Causal connections between Compound Indicator and BTC/USDT Price difference

- Temporal correlation analysis for different sentiment metrics (Channel weight included) with mutual Pearson correlation computed between the daily Bitcoin price difference (derivative) and compound sentiment indicator (built on combination of Channel and metric weights, maximising the overall correlation value)
- $Y(l,d) = \sum X(c,m,d) * P(l,c,m) * W(c)$

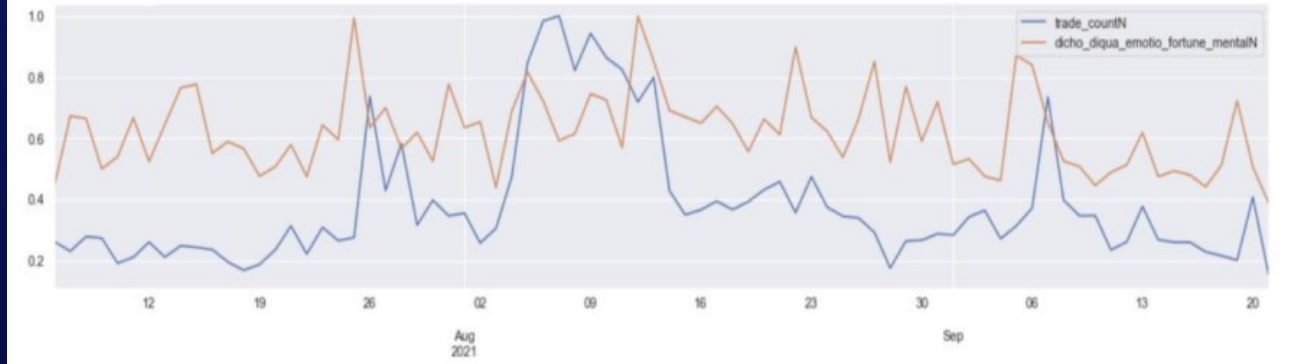


Results: Causal connections between Sentiment and BTC/USD Price Moves and Volume

- Cognitive Distortion vs Price Move

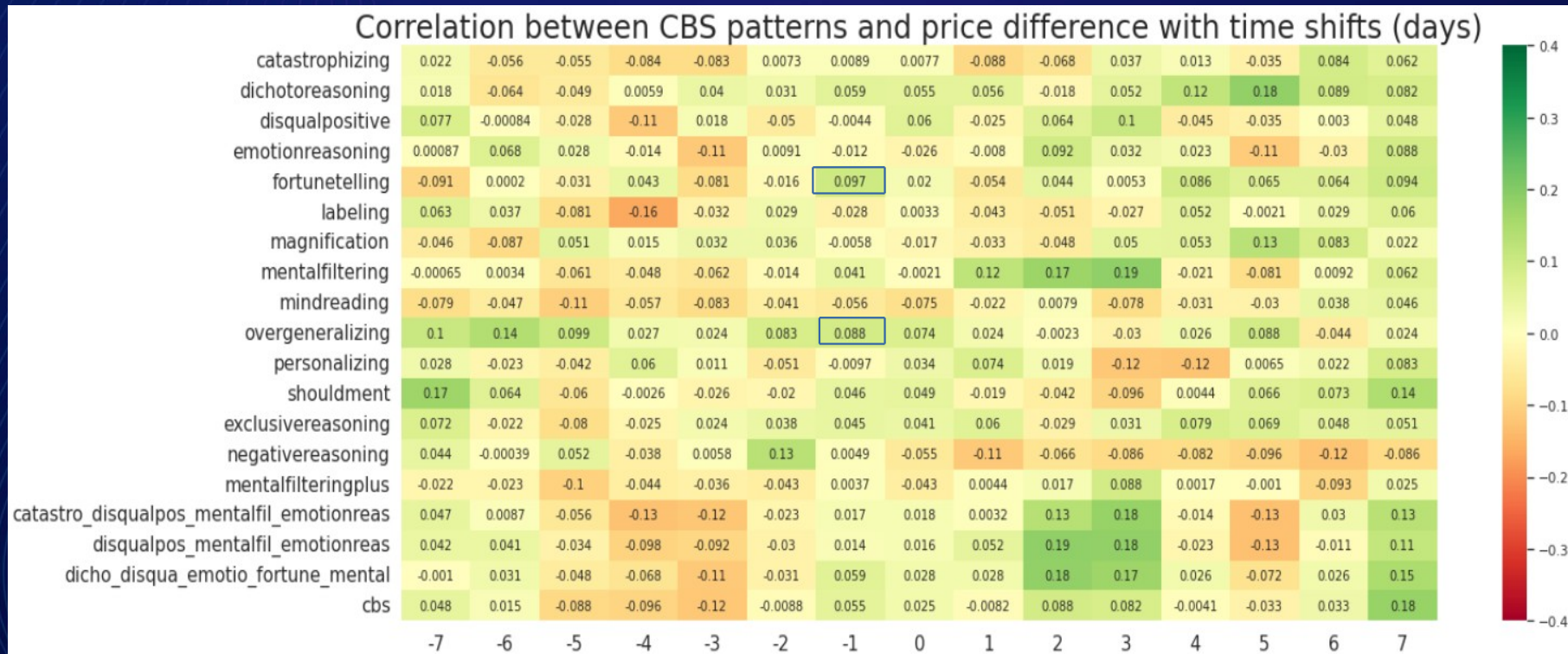


- Cognitive Distortion vs Trade Count



CBS patterns vs. Market Dynamics (Price difference)

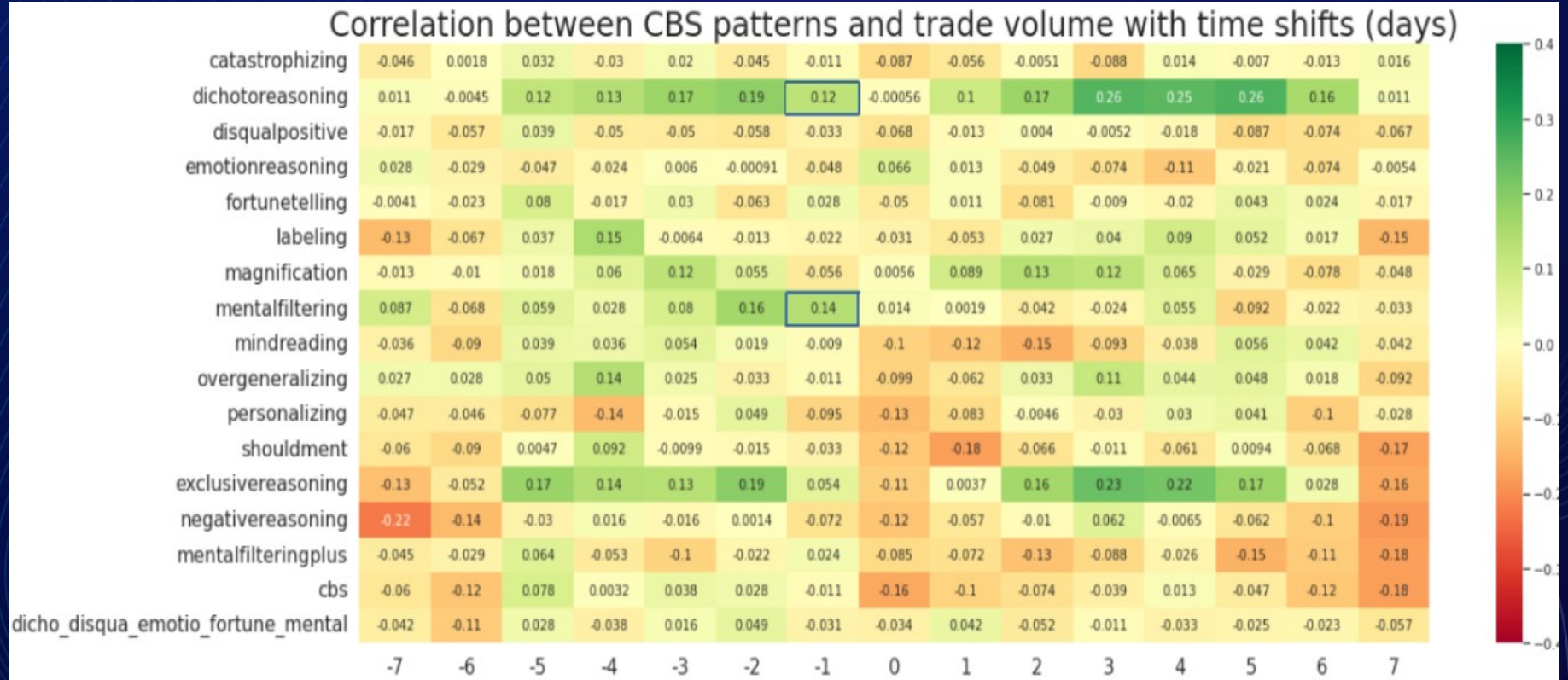
- Tracking overall Twitter and Reddit cognitive distortions as a predictor for BTC/USD price difference



CBS patterns vs. Market Dynamics (Volume)

- Tracking overall Twitter and Reddit cognitive distortions as a predictor for BTC/USD trade volume

Correlation between CBS patterns and trade volume with time shifts (days)



Challenges and Future work

- **N-grams weight:** Excellent > Good
- **Context-aware ngrams:** The wether is cold but BTC is shining (Context is about the BTC not about the wether).
- **Current news agenda aware ngrams:** This law will block the crypto industry, This law was passed (should be negative).
- **Sarcasm:** I like how these friendly senators have agreed on that law increasing the government control over crypto. (seems positive-neutral but rather negative in fact)
- Idioms.
- Non-text data.

Thank you for your attention

mukul.vishwas@singularitydao.ai
anton@singularitynet.io