



PREDICTING INFLUENCE OF NEWS ARTICLES ON CRYPTO PRICES

Sentiment Analysis, Financial Theory & Neural
Networks



PARADIGM & DATA X


Sentiment Analysis, Financial Theory & Neural
Networks

WHAT IS PARADIGM?



Paradigm provides a conversational marketplace for the crypto trading market. It features AI driven tools like automated over-the-counter trading and provides a sophisticated chat tool for institutional traders.

The chat is a combination of several trader tools including counter party discovery, text-to-trade recognition, etc.

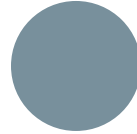


HOW DO WE CONTRIBUTE?



Analyzing Sentiment of Articles

Understand relation of new articles to cryptocurrency prices using Natural Language Processing. Using Recurrent Neural Networks, we will classify articles as having a positive, negative or no affect.



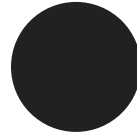
Predict Cryptocurrency Price

Articles will be ranked to identify the articles that are most relevant and use neural networks to predict prices of the cryptocurrency.



Predict Effect of Articles on Crypto Prices

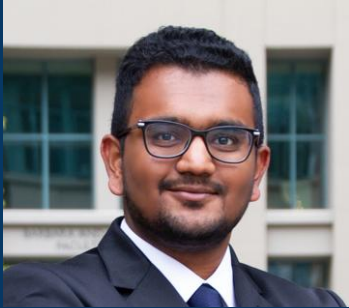
Using time series analysis and RNN the price change on the cryptocurrency price will be predicted based on certain types of articles.



Dashboard with Top Articles for Traders

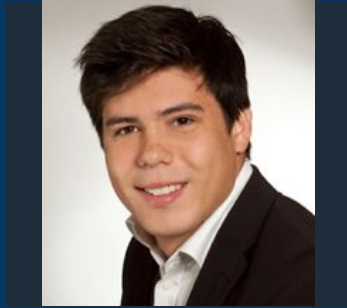
Within the chat box, articles that are most relevant will be recommended to traders on the Paradigm platform.

MEET OUR TEAM



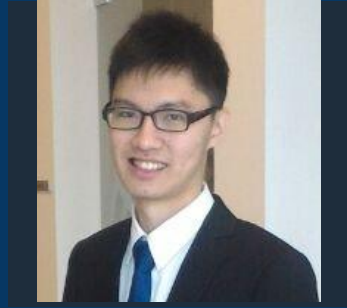
**Sudarshan
Gopalakrishnan**

Sophomore, EECS + IEOR



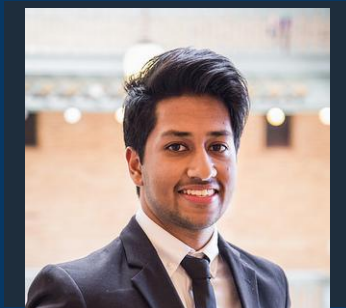
Johannes Kukula

Masters, Business



Kenneth Choong

Junior, IEOR



Srujay Korlakunta

Sophomore, EECS

MEET OUR TEAM



Kshiti Bachlaus

MEng, IEOR



Parth Sanghvi

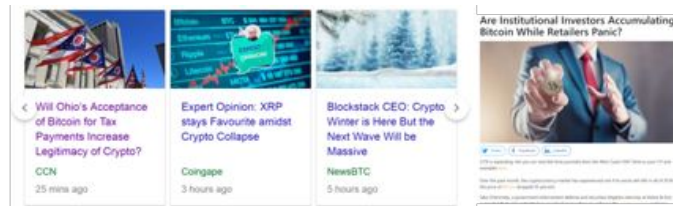
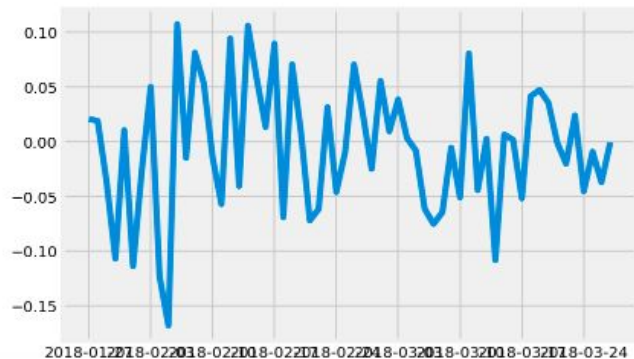
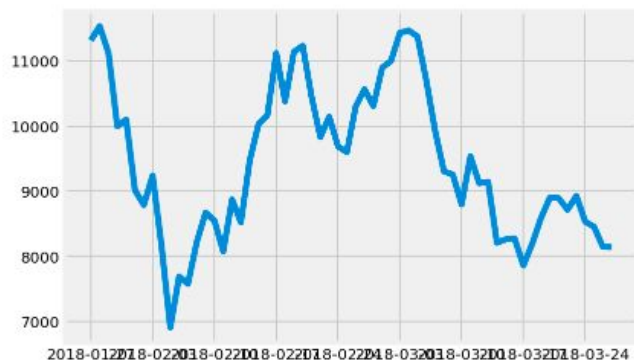
MEng, IEOR

THE PROJECT TASK

Predict which news articles have the most significance in understand the price fluctuations of the crypto-asset mentioned in the article.

- Event Detection
- Quantify Significance of Event on Crypto Prices
- Rank Importance / Influence of Article on Event
- Visually Display this Importance

DATA SAMPLE



Article Data: Full news articles related to bitcoin were scraped from the web

- **Sources:** News APIs (CNBC, Bitcoin Magazine etc.), Kaggle
- **Information:** Author, source, timestamp, title, content
- **Timeframe:** 2018-01-06 to 2018-04-22

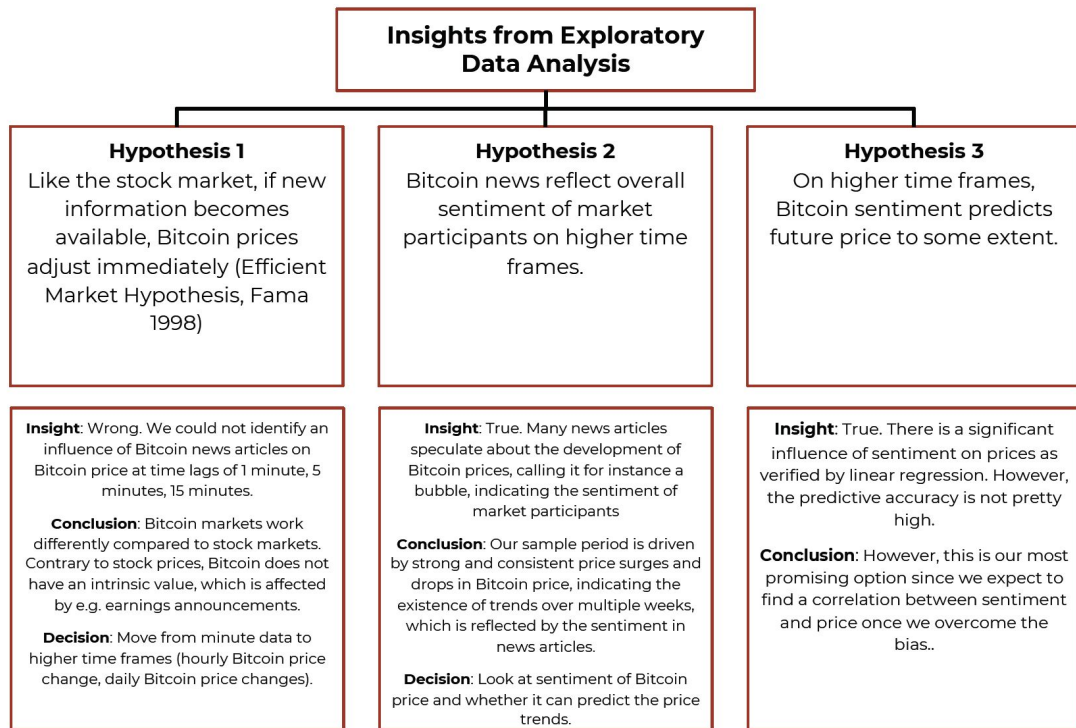
	Timestamp	Open	High	Low	Close	Volume_(BTC)	Volume_(Currency)	Weighted_Price
1686589	2018-03-26 23:56:00+00:00	8155.00	8155.00	8154.99	8154.99	0.617945	5039.342643	8154.997667
1686590	2018-03-26 23:57:00+00:00	8154.99	8154.99	8154.00	8154.01	40.655410	331543.193980	8154.958865
1686591	2018-03-26 23:58:00+00:00	8154.00	8154.01	8150.00	8150.00	9.856911	80340.432933	8150.670628
1686592	2018-03-26 23:59:00+00:00	8150.01	8150.01	8122.82	8145.00	68.274269	555026.852280	8129.370847
1686593	2018-03-27 00:00:00+00:00	8144.99	8145.01	8140.00	8140.00	37.842674	308202.442620	8144.309384

(Snapshot of Bitcoin data from Coinbase)

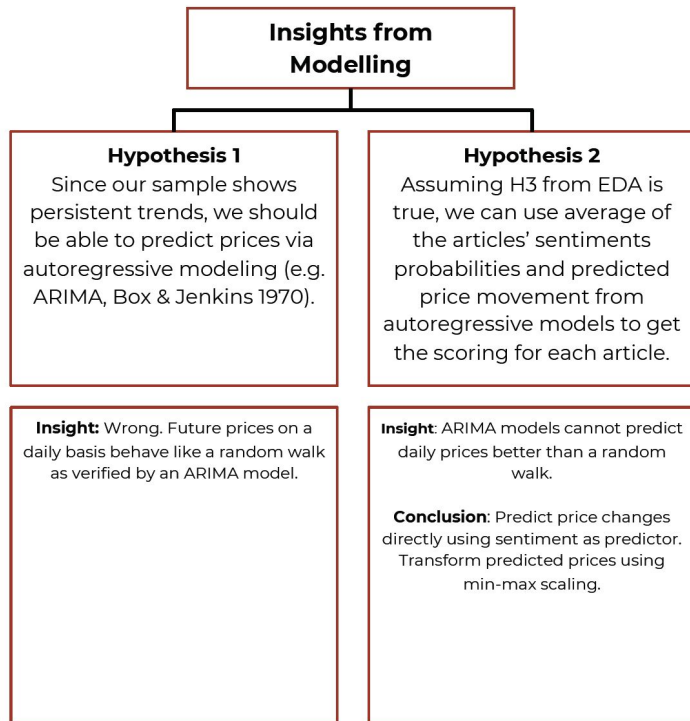
Bitcoin Price Data: Minute-by-minute Bitcoin (BTS) price data in USD

- **Information:** Timestamp, open, high, low, close, vol., BTC price
- **Timeframe:** 2014-12-01 to 2018-03-27

HYPOTHESES FROM EXPLORATORY DATA ANALYSIS



HYPOTHESES FROM MODELING



HYPOTHESES

CONCLUSION

- Move to daily price data.
- Use sentiment analysis of news articles to identify influence of news article on Bitcoin price.
- Predict Bitcoin prices using sentiment.

A collage of various posters and advertisements, including political slogans like "ANKARA'DAYIZ NURIYE VE SEMİH İŞE GERİ ALINSIN" and "23 TEMMUZ PAZAR", and a large silhouette of a person's head in the foreground. The background is a dense, overlapping collection of posters, some with text like "HAYIR", "TANER", "YÜRÜYÜŞ", "BİRLEŞ", "KABARE", and "Neset". The overall theme is social and political activism.

SENTIMENT ANALYSIS USING RNN



Primary Analysis Method 1 – Training RNN using Human-Assigned Data

- Manually Labelled 300 article Dataset, Normalized.
- RNN trained on above dataset.
- RNN along with Softmax regression assigns a probability to each article.



Primary Analysis Method 2 – Training RNN using Word Dictionaries

- The RNN trained word sets for positive and negative sentiments in financial news articles
- Each article is assigned a sentiment between 0 and 1 using softmax regression.
- RNN was then tested on 300 manually labeled articles to create a training set for supervised learning.

Long short-term memory (LSTM) in RNNs provide a more effective sentiment analysis, given that they solve the vanishing gradient problem and hence are able to remember and process long term contexts. Given the subtle tone of financial articles, LSTMs would yield better results on our dataset.

DATA



CLASSIFIED DATA
FROM PREVIOUS
TEAMS

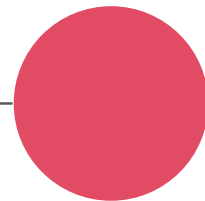


DICTIONARIES

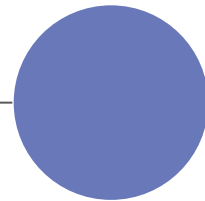


MANUALLY
LABELLED DATA

MODELS



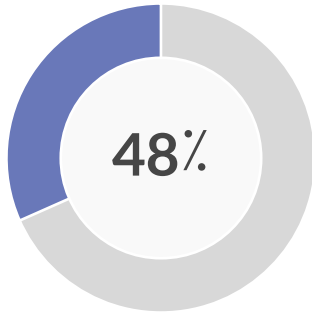
RECURRENT
NEURAL
NETWORK



RANDOM
FOREST
CLASSIFIER

ACCURACY MATRIX

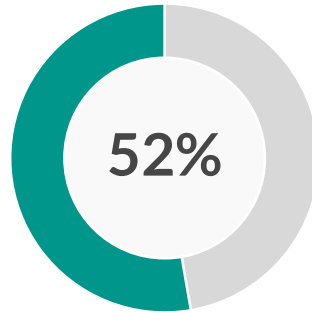
40000 Data Points
1000 Features



RNN

Trained using various
sentiment dictionaries

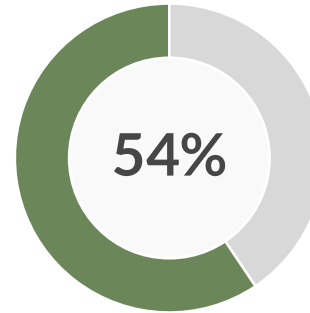
300 Data Points



Random Forest

Manually Labelled DataSet

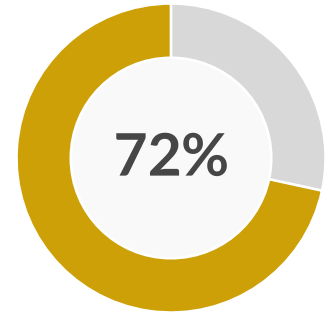
40000 Data Points



Random Forest

Major Dataset

300 Data Points



RNN

Manually Labelled DataSet



PRICE MOVEMENT



- Positive correlation for extreme price changes
- Transform predicted prices into a score between 0 (not significant) and 10 (highly significant) using min-max scaling.

Price prediction using Random Forest

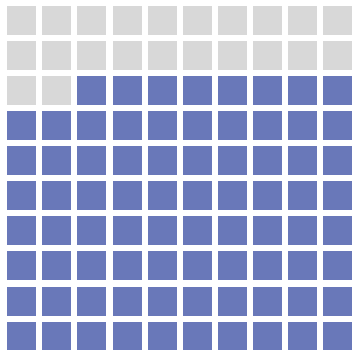
`predicted_price = f(sentiment, vectorized news text) + noise`



SCORING

Scoring of news article reflects the impact of
the news article on Bitcoin price

SCORING METRIC



Article 1

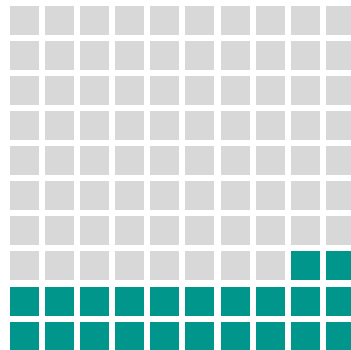
9.0

Headline:

“Experts Believe Bitcoin Crash will continue”

Author

Source



Article 2

3.0

Headline:

“Arun Jaitley has just killed India’s Crypto Party”

Author

Source

```
def create_score(prediction):
```

```
    min_ = np.percentile(abs(prediction), 2)
    max_ = np.percentile(abs(prediction), 98)
```

```
    new_list = []
```

```
    for i in abs(prediction):
        if i < min_ or i > max_:
            new_list.append(10)
        else:
            new_list.append(round((i-min_)/(max_-min_)*10))
    return new_list
```

MIN MAX SCALING OF PREDICTED PRICE CHANGES

HIGHLY SIGNIFICANT

Final Score
Assigned

10.0

true value: 0.11 predicted value: 0.0 predicted scoring: 0.0

Craig Wright, who once claimed to be Bitcoin founder Satoshi Nakamoto, imman's estate for stealing \$5B in bitcoin (Russell Brandom/The Verge)

true value: -0.21 predicted value: -0.02 predicted scoring: 2.0

Arun Jaitley has just killed India's cryptocurrency party

true value: -0.04 predicted value: -0.02 predicted scoring: 2.0

Why did Ethereum Drop so hard? Bitcoin is ~~Correcting~~, and Cryptocurren

true value: 0.3 predicted value: 0.16 predicted scoring: 10

Bitcoin slides below \$6000; half its value lost in 2018 - Fox Business

true value: -0.04 predicted value: 0.05 predicted scoring: 4.0

Dutch Court Finds Bitcoin A Legitimate "Transferable Value"

true value: 0.14 predicted value: 0.12 predicted scoring: 10

Bitcoin plunges below \$10,000 as major crypto exchange to share user c

true value: 0.19 predicted value: 0.11 predicted scoring: 10

Expert Believes Bitcoin Price Crash Will Continue

NOT SIGNIFICANT

Final Score
Assigned

2.0

true value: 0.11 predicted value: 0.0 predicted scoring: 0.0

Craig Wright, who once claimed to be Bitcoin founder Satoshi Nakamoto, imman's estate for stealing \$5B in bitcoin (Russell Brandom/The Verge)

true value: -0.21 predicted value: -0.02 predicted scoring: 2.0

Arun Jaitley has just killed India's cryptocurrency party

true value: -0.04 predicted value: -0.02 predicted scoring: 2.0

Why did Ethereum Drop so hard? Bitcoin is Correcting, and Cryptocurren

true value: 0.3 predicted value: 0.16 predicted scoring: 10

Bitcoin slides below \$6000; half its value lost in 2018 - Fox Business

true value: -0.04 predicted value: 0.05 predicted scoring: 4.0

Dutch Court Finds Bitcoin A Legitimate "Transferable Value"

true value: 0.14 predicted value: 0.12 predicted scoring: 10

Bitcoin plunges below \$10,000 as major crypto exchange to share user c

true value: 0.19 predicted value: 0.11 predicted scoring: 10

Expert Believes Bitcoin Price Crash Will Continue

DASHBOARD



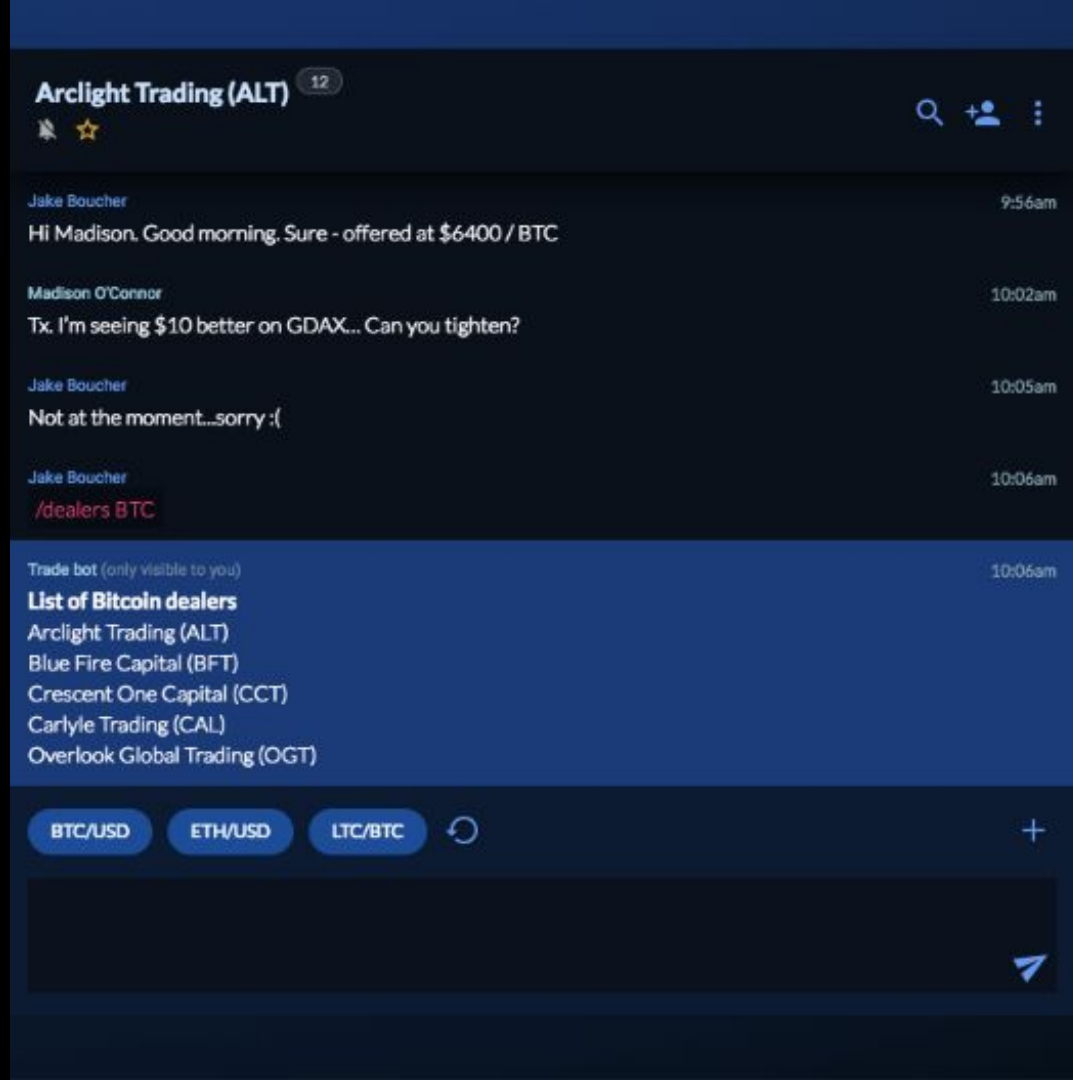
UPDATE

+ 40.25 \$

- 05.75 \$

CHAT INTERFACE

The final product is to be pipelined directly into Paradigm's chat interface. This would enable all entities using the chat interface to directly discern the relevance of news articles while trading over the chat interface.



DASHBOARD EXAMPLE

Governments Might Start Issuing Their Own Cryptocurrencies



Canada looking into Bitcoin's risk factor, Bill Morneau says



SEC, CFTC Warn of Fraud in Cryptocurrency Markets





LIMITATIONS & NEXT STEPS

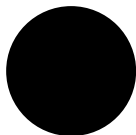
LIMITATIONS AND NEXT STEPS



Increase Quality of Data. Includes removing non relevant articles, refining articles based on authors, etc.



Create a customized dictionary that can be used to increase the accuracy of the RNN sentiment analysis.



Create a labelled dataset of articles professionally classified by cryptocurrency experts in order to explore supervised learning opportunities.



FUTURE PROSPECTS

FUTURE PROSPECTS



The final version of the model will be able to look at all articles available on the internet and sort the articles as well as rate them



The model can also be trained to evaluate the sentiment of a series of articles and we believe there is significant research potential in this area

FINANCE REPORT

ACCOUNT

REPORT

DASHBOARD

INCOME

OUT GOING

DASHBOARD > INCOME

DAILY WEEKLY MONTHLY



TOTAL INCOME

37.91%

LINE ITEMS 16.15 M\$
SHIPPING 0.15 M\$
TAXES 0%
TOTAL 16.3 M\$

31.86%

LINE ITEMS 13.5 M\$
SHIPPING 0.2 M\$
TAXES 0%
TOTAL 13.7 M\$

30.23%

LINE ITEMS 13.00 M\$
SHIPPING 0 \$
TAXES 0%
TOTAL 13.00 M\$

THANK YOU

Learn from the best to
ensure success
Reasons we will be
successful

BUS