

SENTIMENT ANALYSIS OF CRYPTO CURRENCIES: ETHEREUM CLOSE PREDICTION MODEL

Presented By: Eric Romano





01
OBJECTIVES

02
METHODOLOGY

03
RESULTS ANALYSIS

04
CONCLUSIONS

An illustration of a person with dark hair in a bun, wearing a yellow shirt with white polka dots and dark pants with white stripes, standing next to a large yellow rectangular whiteboard. The person is gesturing towards the whiteboard with their right hand. On the whiteboard, the word 'INTRODUCTION' is written in bold black capital letters. Below it, a paragraph of text describes a volatile cryptocurrency market. In front of the whiteboard is a yellow desk with a black laptop. The entire scene is set against a light gray background. There are four yellow circular icons with a black dollar sign (\$) floating around the whiteboard: one in the top right, one on the left side, one in the bottom left, and one on the right side near the person's head.

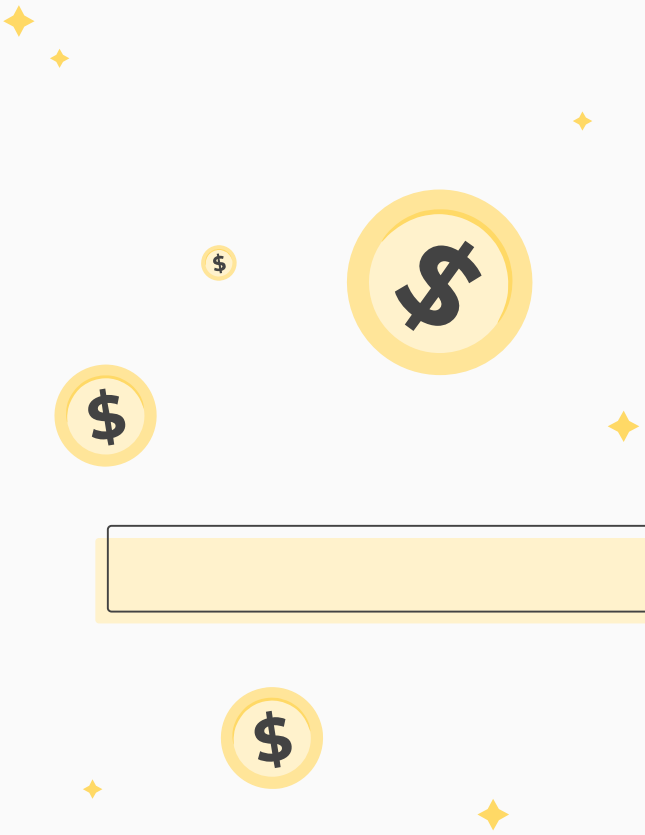
INTRODUCTION

During the week of May 15th , the market experienced a volatile environment for cryptocurrencies with Bitcoins plugging to half of it all time high of \$60,000. A total of \$830 billion lost by investors causing many investors to go bankrupt.

BUSINESS PROBLEM

Many retail investors would like to have tools to have at their disposal to make predictions that can lead to better returns or better exit strategies. The crypto market is strongly connected to retail investors and their emotions. Using Sentiment Analysis and other metrics such as Relative Strength Index we believe these will be useful in creating tools that will classify if the market closes higher or lower than the day before. The first crypto currency that will be explored in is Ethereum and will use these models to use for other top trending crypto currencies.

1. Can we successfully use sentiment analysis features to classify if the market will close higher or lower than the previous day?
2. Can we extract features using financial technical analysis tools such as Relative Strength Index that is used when evaluating momentum? To visualize and isolate conditions of overbought and oversold
3. Can previous historical data, such as the day before, be a strong feature used in modeling?
4. What are the important features of the models, and can we extract promising insight for these features that can be used in creating investment strategies?

A decorative graphic on the left side of the slide. It features several gold coins of varying sizes, each with a black dollar sign (\$) in the center. The coins are scattered around a central, larger coin. There are also several small, four-pointed gold stars scattered throughout the graphic area.

"There is nothing that Bitcoin can do which
Ethereum can't. While Ethereum is less
battle-tested, it is moving faster, has better
leadership, and has more developer
mindshare."

—GIL PENCHINA

METHODOLOGY

STEP 01

Gathered data from a reputable source. Most of the data webscraped from Coinmarketcap.com and yahoofinance

STEP 04

Evaluate models and make recommendations. Recall and accuracy were prioritized



STEP 02

Explore, clean, and analyze. Find insights

STEP 03

Build predictive model: Linear Discriminant Analysis and Random Forest





POSITIVE SENTIMENT

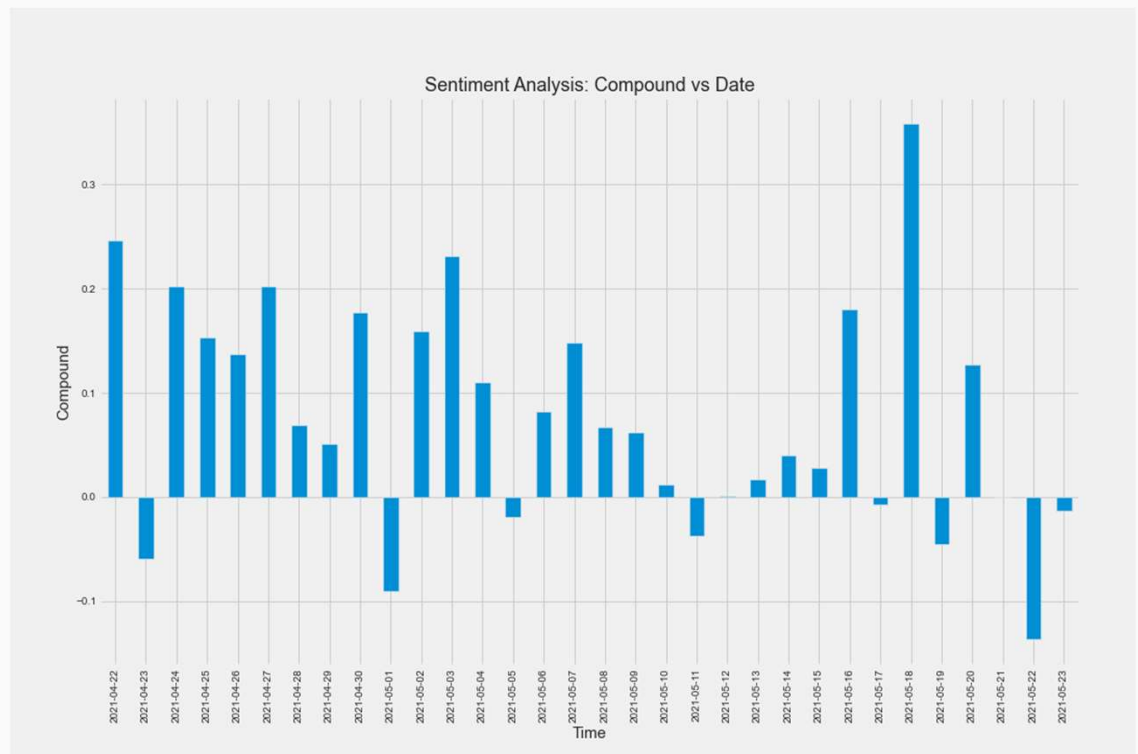
Leading up to the crash you see a overall positive outlook of the market based on the sentiment of News headlines



CHANGE IN SENTIMENT

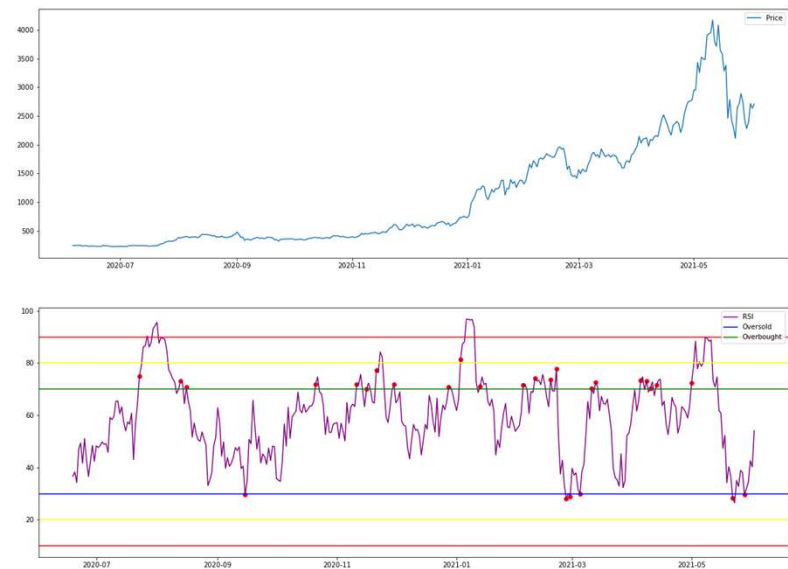
Before the crash you begin to see a drop in the magnitude of the compounded sentiment. This suggest that there is more negative sentiment spreading through the media.

SENTIMENT ANALYSIS: COMPUND VS DATE



RELATIVE STRENGTH INDEX

One of the indicators used to create engineered features was the Relative Strength Index which is a metric that can be used to understand if the current market is being overbought or oversold. From the chart above you can see when Ethereum crossed the overbought bar and touched the 90% level right before a large sell off occurred.



01 ID NEWS CENTER

This feature represents the unique id given to different news centers.

02 BOUGHT_SOLD

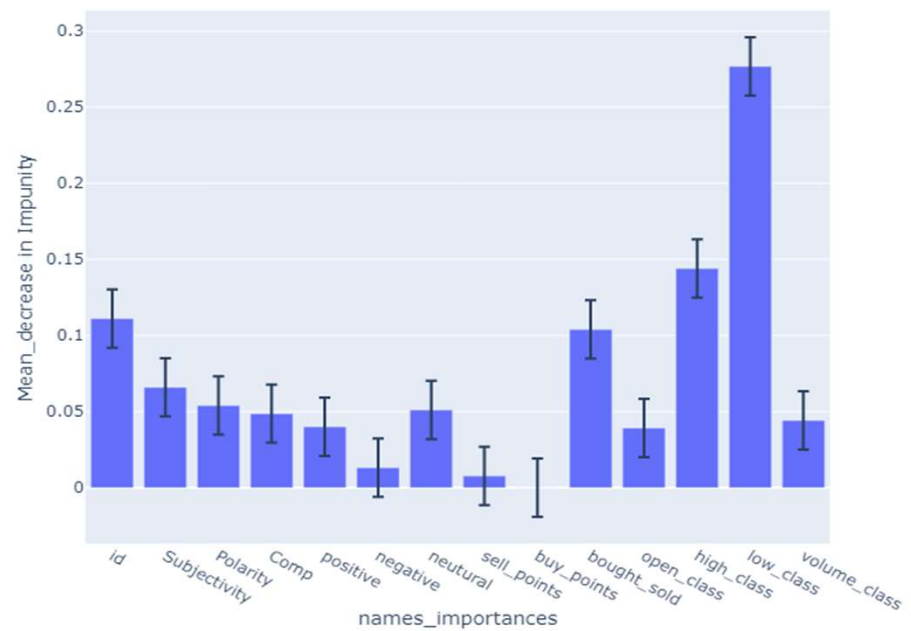
Feature engineered to indicate when there are conditions of overbought and over sold

03 LOW_CLASS

This label identifies in the low value is higher or lower than the day before

FEATURE IMPORTANCE

Feature Importance Using MDI:SA and RSI



FEATURE IMPORTANCE

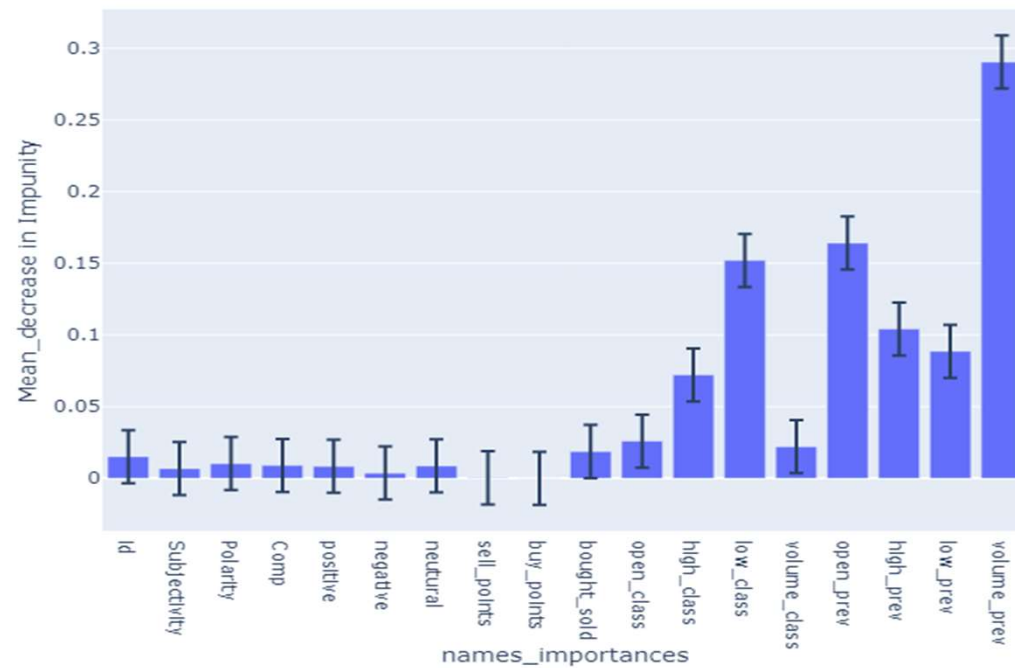
01 OPEN PREV

The open value of the day before

02 VOLUME PREV

The volume value of the day before

Feature Importance Using MDI:SA and RSI and HD





MODEL EVALUATION

BASE MODEL: SENTIMENT ANALYSIS FEATURES



Linear Discriminant Analysis

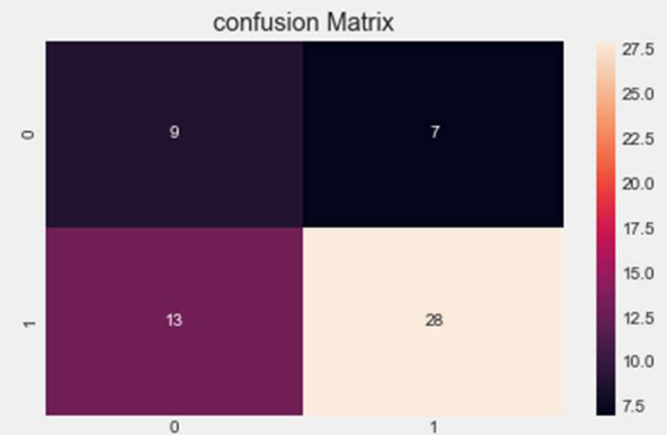
The accuracy is :0.74

AUC is :0.61

```
Classification Report:
      precision    recall  f1-score   support

     0       0.56      0.31      0.40        16
     1       0.77      0.90      0.83        41

   accuracy          0.74        57
  macro avg       0.66      0.61      0.62        57
 weighted avg       0.71      0.74      0.71        57
```



Random Forest

The accuracy is :0.65

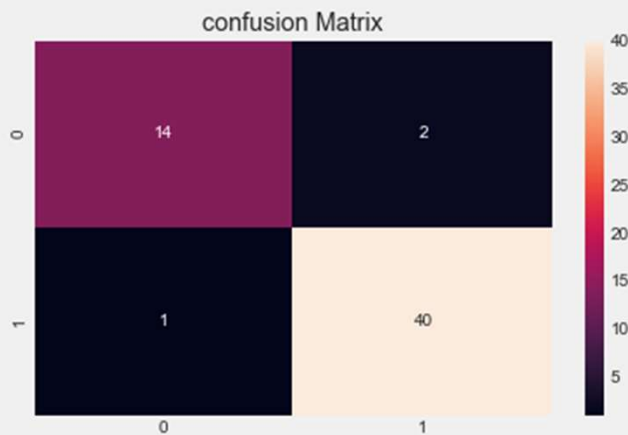
AUC is :0.62

```
Classification Report:
      precision    recall  f1-score   support

     0       0.41      0.56      0.47        16
     1       0.80      0.68      0.74        41

   accuracy          0.65        57
  macro avg       0.60      0.62      0.61        57
 weighted avg       0.69      0.65      0.66        57
```

FINAL MODEL: SENTIMENT ANALYSIS, RSI, HISTORICAL DATA FEATURES



Linear Discriminant Analysis

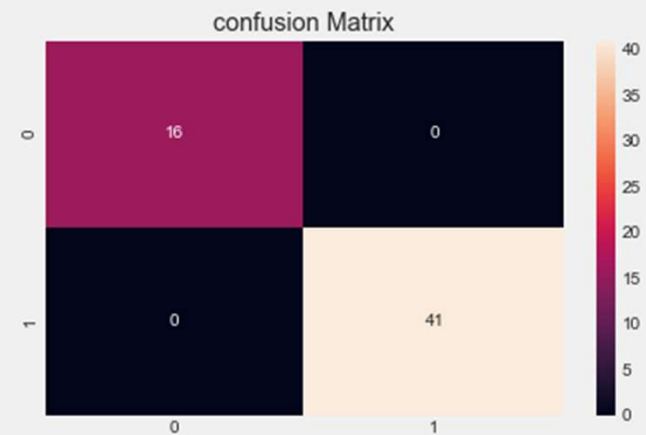
The accuracy is :0.95

AUC is :0.93

```
Classification Report:
      precision    recall  f1-score   support

     0       0.93      0.88      0.90        16
     1       0.95      0.98      0.96        41

   accuracy          0.95
  macro avg       0.94      0.93      0.93        57
 weighted avg     0.95      0.95      0.95        57
```



Random Forest

The accuracy is :1.0

AUC is :1.0

```
Classification Report:
      precision    recall  f1-score   support

     0       1.00      1.00      1.00        16
     1       1.00      1.00      1.00        41

   accuracy          1.00
  macro avg       1.00      1.00      1.00        57
 weighted avg     1.00      1.00      1.00        57
```


A collection of various coins and a Bitcoin on a dark background. The coins include Euro coins (1, 2, 5, 10, 20, 50) and a Bitcoin. The background is dark and textured.

CONCLUSION

In conclusion, the combination of engineered features from the historical data set and the features from a sentiment analysis provide a great foundation to models to predict whether the Ethereum market will close higher or lower than the previous day. Looking into just the months leading up to these all-time highs and before the crash our models provide actionable recommendation to help investors possibly make more informed investments.

RECOMMENDATION



NEWS CENTER ID

Find news centers that have a strong correlation to market. Use these to inform decisions



RSI: BOUGHT AND SOLD

By indicating when the market is either in a condition of overbought or oversold will help prediction how market will close



PREVIOUS VOLUME VALUE

Presents the most important feature. Trading volume is a technical indicator that represents the overall activity of a security or market



MONITOR LOWS

The low from the previous day is helpful when evaluate it based on the current low.

FUTURE WORK

MORE DATA

By adding more data points, we will be able to see if the model remains great or if there is a significant decrease in the metrics that were used to evaluate the performance

BITCOIN

When performing the sentiment analysis, it was found that Bitcoin was the most common word that was linked to the article. Bitcoin is the leading crypto currency. We believe the same engineered feature will help in creating a similar model

CARDANO

Cardano is trying to answer the problems that ethereum has. Being its direct challenger might mean that most of the same engineered features will help in creating a similar model

THANKS

Does anyone have any questions?

Egustavo94@gmail.com



[Linkedin.com/in/eric-romano](https://www.linkedin.com/in/eric-romano)

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik** and illustrations by **Storyset**

