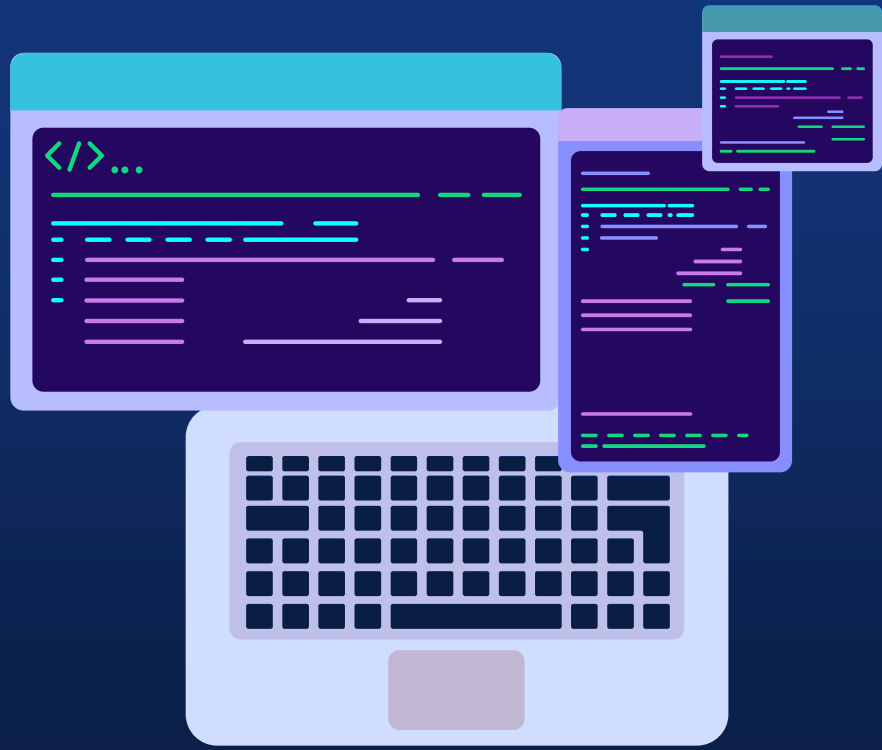
PREDICTION OF CRYPTOCURRENCY PRICES USING HISTORICAL & TWITTER DATA



Team Data Pirates
(17)

1900%

Bitcoin increased its value in 2017 before plunging in 2018

100,000

Average number of tweets shared per day
related to Bitcoin

yahoo!finance | Yahoo Finance UK

Bitcoin hits $50,000 as Elon Musk's tweet sends Shiba Inu soaring

Kotaku

Genshin Impact Cancels Elon Musk Event Following Fan Pushback

I am convinced that I have died and gone to hell. Last night, the official Genshin Impact Twitter account tweeted a contest in which the...

2 days ago

Reuters

Musk tweets he is in talks with airlines to install Starlink broadband

Oct 14 (Reuters) - Billionaire entrepreneur Elon Musk said in a tweet on Thursday he was in talks with airlines about installing Starlink,...

2 days ago

Yahoo Finance

Elon Musk dogecoin tweet lifts bitcoin

Cryptocurrencies were broadly up on Thursday morning as Tesla (TESLA) CEO Elon Musk tweeted in support of the meme-based cryptocurrency...

2 days ago

TABLE OF CONTENTS

PROBLEM STATEMENT

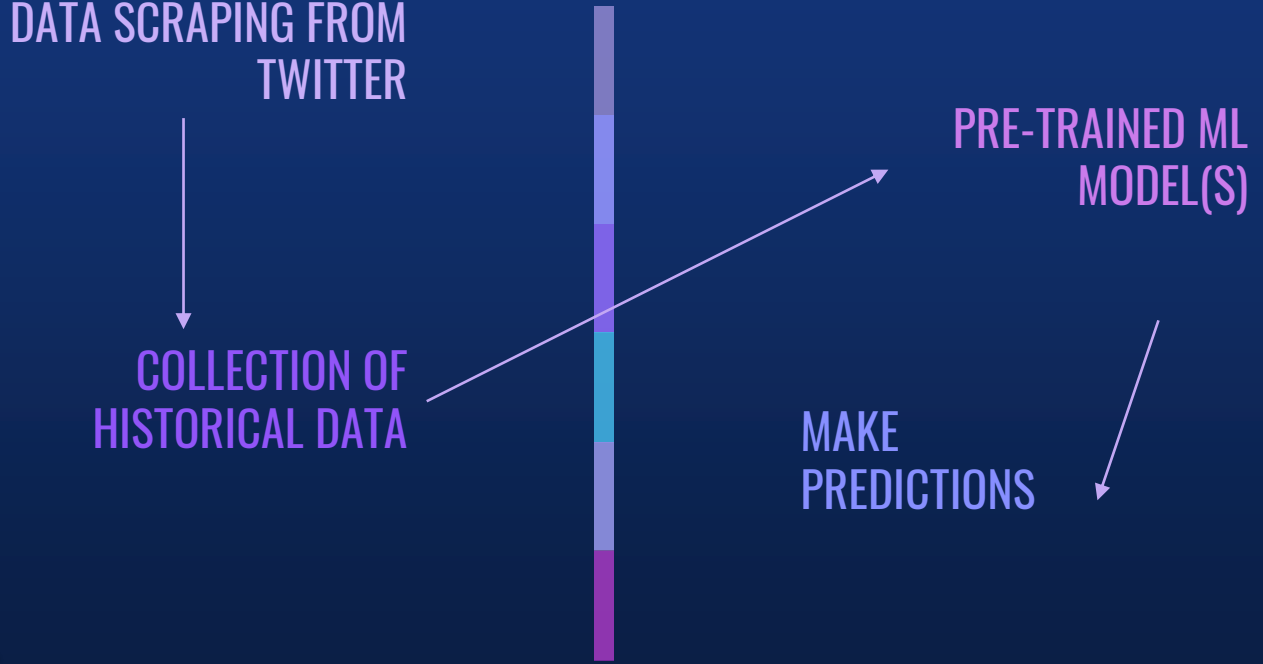
We believe that AI has endless possibilities, and it should not be limited to a certain theme. Therefore, the task is to build an AI-based solution that we are passionate about.

Bitcoin

OUR AIM

We aim to analyze the historical data, i.e., the open, high, low, and close prices of the Bitcoin cryptocurrency, along with real-time sentimental analysis of the tweets centered around the same for a similar time to predict the weighted price of the token by the end of the trading market.

OUR APPROACH



TECHNICAL EXPOSITION

Collection and Curation

Obtain data from Twitter, perform Sentiment Analysis and merge with Alpha Vantage API

Data Processing

Organize the dataset in such a manner that can be used for analysis further.

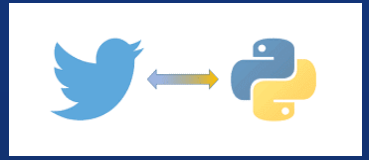
Model Fitting

Train various models such as MLP, LSTM, RF, Linear Regression, SVM

Checking Accuracy

RSME and MAE of fit of models

COLLECTION, CURATION AND PROCESSING



kaggle

	Open	High	Low	Close	Compound_Score	Total Volume of Tweets	Weighted_Price
0	2763.23	2763.24	2761.41	2762.00	0.082893	1027.0	2761.710702
1	2768.07	2772.97	2768.07	2768.07	0.053160	778.0	2772.411512
2	2779.77	2779.78	2779.77	2779.78	0.124251	836.0	2779.774992
3	2790.55	2793.25	2790.55	2790.55	-0.021037	984.0	2792.693685
4	2837.44	2837.44	2831.40	2831.40	0.055437	751.0	2832.734750

Final, cleaned dataset prepared

MACHINE LEARNING MODELS

LINEAR REGRESSION

```
linreg = LinearRegression()
linreg.fit(X_train, y_train)
```

RANDOM FOREST

```
rforest = RandomForestClassifier(n_estimators = 100,  max_depth = 10)
rforest.fit(X_train, y_train.Weighted_Price.ravel())
```

SUPPORT VECTOR MACHINE

```
clf = svm.SVR(kernel='rbf')
clf.fit(X_train, y_train)
```

MACHINE LEARNING MODELS

MULTILAYER PERCEPTRON

```
mlp = MLPClassifier(hidden_layer_sizes=(10, 10, 10), max_iter=100)

mlp.fit(X_train, y_train.Weighted_Price.ravel())
```

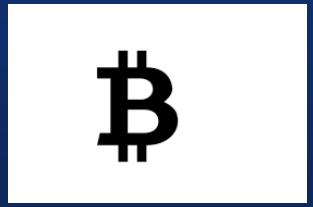
LONG-SHORT TERM MEMORY

```
regressor = Sequential()

regressor.add(LSTM(units = 50, return_sequences = True, input_shape = (X_train.shape[1], 1)))
regressor.add(Dropout(0.2))

regressor.add(LSTM(units = 50, return_sequences = True))
regressor.add(Dropout(0.2))

regressor.add(LSTM(units = 50, return_sequences = True))
regressor.add(Dropout(0.2))

regressor.add(LSTM(units = 50))
regressor.add(Dropout(0.2))

regressor.add(Dense(units = 1))

regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')

regressor.fit(X_train, y_train, epochs = 100, batch_size = 32)
```

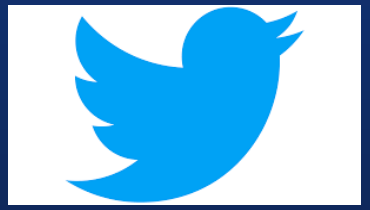
PERFORMANCE METRICS

MODEL	ROOT MEAN SQUARED ERROR	MEAN ABSOLUTE ERROR
Multi-variate Linear Regression	4.731893281133085	2.6001870657470194
Random Forest Regression	0.0	813.6578657865787
Support Vector Machine	2631.633668587859	1669.4916573941057
Multilayer Perceptron	340.324490855283	170.01980198019803
Long-Short Term Memory	0.0041880845267927985	0.003205464143801855

AREAS OF FURTHER EXPLORATION



AND SO ON..



Bloomberg

AND SO ON..

Thank you!