Crypto Board

A PROJECT REPORT

Submitted by

Ashutosh Nivrutti Kesarkar

in partial fulfilment for the award of the degree

of

M.Sc. Data Science and Artificial Intelligence – Part 2



RAMNIRANJAN JHUNJHUNWALA COLLEGE OF ARTS, SCIENCE & COMMERCE (AUTONOMOUS), GHATKOPAR W

March - 2022



RAMNIRANJAN JHUNJHUNWALA COLLEGE OF ART'S, SCIENCE & COMMERCE (AUTONOMOUS), GHATKOPAR (W)



(Affiliated to University of Mumbai)

Certificate

This is to certify that the Project entitled Crypto Board is bonafide work of Ashutosh Nivrutti Kesarkar bearing Seat No 714 submitted in partial fulfilment of the requirements for the award of Degree Master of Science in Data Science & Artificial Intelligence.

Signature of Internal Guide

Signature of Co-Ordinator

College Seal and Date

Signature

Examiner

Abstract

"Prediction" refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome, such as whether or not a customer will churn in 30 days. The algorithm will generate probable values for an unknown variable for each record in the new data, allowing the model builder to identify what that value will most likely be. The word "prediction" can be misleading. In some cases, it really does mean that you are predicting a future outcome, such as when you're using machine learning to determine the next best action in a marketing campaign. Other times, though, the "prediction" has to do with, for example, whether or not a transaction that already occurred was fraudulent. In that case, the transaction already happened, but you're making an educated guess about whether or not it was legitimate, allowing you to take the appropriate action.

The objective of this article is to predict and analyse the trend of crypto coin market based on the value specified in the report which is hard to interpret if one has no knowledge from the stock market. Data used in this article is publicly available at Kaggle. This will be a regression problem and the target or dependent variable changes as per the price.

Acknowledgement

Before we get into thick of things, we would like to add a few heartfelt words for the people who were part of the project in numerous ways, people who gave unending support right from the stage the project idea was conceived.

A project report is such a comprehensive coverage; it would not have been materialized without the help of many. The four things that go on to make a successful endeavour are dedication, hard work, patience and correct guidance. Able and timely guidance not only helps in making an effort fruitful but also transforms the whole process of learning and implementing into an enjoyable experience.

In particular, I would like to thank our Mentor/Director **Dr.(Mrs.) Usha Mukundan**, R.J. College. I would like to give a very special honour and respect to our teacher, **Prof. Tejaswee Pol** who took keen interest in checking the minute details of the project work and guided us throughout the same. A sincere quote of thanks to the non-teaching staff for providing us software their time. I appreciate outstanding co-operation by them, especially for the long Lab timings that we could receive.

Declaration

I hereby declare that the Project entitled, "Crypto Board" done at R. J. COLLEGE, Ghatkopar(W), Mumbai, has not been in any case duplicated to submit to any other university for the award of any degree. To the best of my Knowledge other than me, No one has submitted to any other University. The Project is done in partial fulfilment of the requirements for the award of degree of Master of Science Data Science and Artificial Intelligence to be submitted as mini project as part of our curriculum.

Table of Contents

Index	Headings	Page Number			
1	Abstract	3			
2	Acknowledgement	4			
3	Declaration	5			
5	Introduction	7			
6	Proposed Methodology	8-9			
7	Description of Data	10			
8	Code and Output	11-19			
9	Conclusions	20			
10	Future Scope	21-23			
11	References	24-25			

Introduction

The A cryptocurrency is a digital or virtual currency that is secured by cryptography, which makes it nearly impossible to counterfeit or double-spend. Many cryptocurrencies are decentralized networks based on blockchain technology—a distributed ledger enforced by a disparate network of computers.

A defining feature of cryptocurrencies is that they are generally not issued by any central authority, rendering them theoretically immune to government interference or manipulation.

Bitcoin had a rocky first half of the year, but experts still say it will eventually hit \$100,000 — and that it's more a matter of when, not if.

Despite the volatility and recent slumping price, many experts still say bitcoin is on its way to passing the \$100,000 mark, though with varying opinions on exactly when that will happen. And a recent study by Deutsche Bank found that about a quarter of bitcoin investors believe bitcoin prices will be over \$110,000 in five years.

The volatility is nothing new, and is a big reason experts say new crypto investors should be extremely cautious when allocating part of their portfolio to cryptocurrency. bitcoin has shown as steady a rise in value over the years as any other cryptocurrency on the market.

It's only reasonable for bitcoin investors to be curious about how high it can ultimately go.

Proposed Methodology

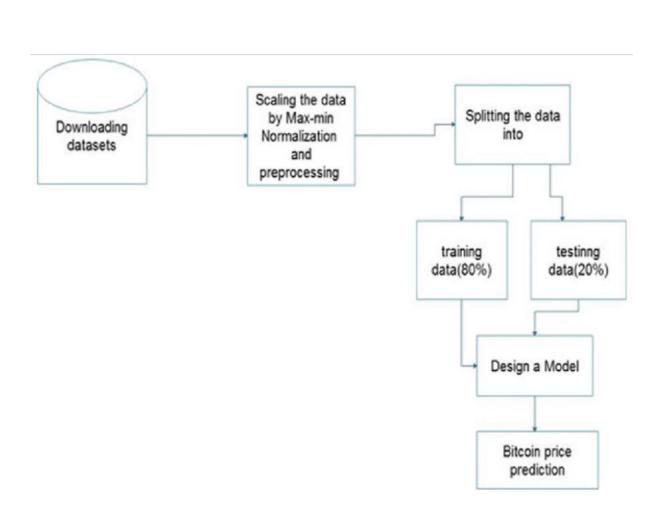
Gathering the Data: Data preparation is the primary step for any machine learning problem. We will be using a dataset from Kaggle for this problem. This dataset consists of two CSV files one for training and one for testing.

<u>Cleaning the Data</u>: Cleaning is the most important step in a machine learning project. The quality of our data determines the quality of our machine learning model. So it is always necessary to clean the data before feeding it to the model for training. In our dataset all the columns are numerical, the target column i.e. prognosis is a string type and is encoded to numerical form using a label encoder.

<u>Model Building</u>: After gathering and cleaning the data, the data is ready and can be used to train a machine learning model. We will be using this cleaned data to train the Linear Regression. We will be using a confusion matrix to determine the quality of the models.

<u>Inference</u>: After training all models we will be predicting the price of a number of crypto currencies in the near future . This makes our overall prediction more robust and accurate.

At last, we will be publishing the dashboard created using power BI on the Microsoft platform.



Description

The Dataset which we are going to use have a csv file of each type of currency. The currencies which we have picked were the 20 currencies in the market at the time of making of this project. The price history is available on a daily basis from April 28,2013. This dataset have historical price information of some of the top crypto currencies by market capitalization.

Crypto Currencies used:





SOLANA COIN



UNISWAP COIN



STELLAR COIN



USD COIN



TETHER COIN



WRAPPED BITCOIN



TRON COIN



XRP COIN

Source Code

Eda:

```
#importing libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pandas_datareader.data as web
import plotly.express as px
import plotly.graph_objects as go
from datetime import datetime, timedelta

import matplotlib.ticker import ScalarFormatter
import warnings
warnings.filterwarnings("ignore")

#for bold
start='\033[1m'
end='\033[0;0m'
```

Merging all the data csv of crytocurrencies

```
df = pd.DataFrame()
for file in os.listdir("Data"):
    if file.endswith(".csv"):
        df = pd.concat([df , pd.read_csv(os.path.join("Data", file),infer_datetime_format=True)], axis=0 )

df.reset_index(drop=True, inplace=True)
df.drop('SNo',axis=1 ,inplace=True)
df
```

	Name	Symbol	Date	High	Low	Open	Close	Volume	Marketcap
0	Aave	AAVE	2020-10-05 23:59:59	55.112358	49.787900	52.675035	53.219243	0.000000e+00	8.912813e+07
1	Aave	AAVE	2020-10-06 23:59:59	53.402270	40.734578	53.291969	42.401599	5.830915e+05	7.101144e+07
2	Aave	AAVE	2020-10-07 23:59:59	42.408314	35.970690	42.399947	40.083976	6.828342e+05	6.713004e+07
3	Aave	AAVE	2020-10-08 23:59:59	44.902511	36.696057	39.885262	43.764463	1.658817e+06	2.202651e+08
4	Aave	AAVE	2020-10-09 23:59:59	47.569533	43.291776	43.764463	46.817744	8.155377e+05	2.356322e+08

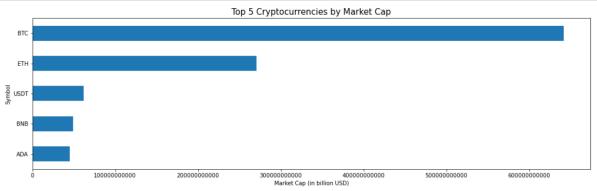
37077	XRP	XRP	2021-07-02 23:59:59	0.667287	0.634726	0.659890	0.656763	2.061607e+09	3.030759e+10
37078	XRP	XRP	2021-07-03 23:59:59	0.683677	0.644653	0.655639	0.672888	1.872820e+09	3.105172e+10
37079	XRP	XRP	2021-07-04 23:59:59	0.707783	0.665802	0.673218	0.694945	1.885242e+09	3.206960e+10
37080	XRP	XRP	2021-07-05 23:59:59	0.695653	0.648492	0.695653	0.654300	2.076373e+09	3.019395e+10
37081	XRP	XRP	2021-07-06 23:59:59	0.679923	0.652676	0.653055	0.665402	1.938959e+09	3.072284e+10

```
df['Date']=pd.to_datetime(df['Date']).dt.date

df.to_csv('final.csv')
```

Finding the top 5 crytpocurrencies by market cap

```
plt.figure(figsize=(18,5))
ax = df.groupby(['Symbol'])['Marketcap'].last().sort_values(ascending=False).head(5).sort_values().plot(kind='barh')
ax.set_xlabel("Market cap (in billion USD)")
ax.ticklabel_format( style='plain', axis='x')
plt.title("Top 5 Cryptocurrencies by Market Cap", fontsize=15)
plt.show()
```



Observation:

Here we can see that Bitcoin is the highest value according to market cap and etherum is 50% of it. and all other are very less in comparison to it and USDT is 3 times less then the Etherum

Plotting Graphs of Closing Prices of Top 4 Cryptocurrencies as per Market Cap

```
dx=df.copy()

top_4_currency_names = dx.groupby(['Symbol'])['Marketcap'].last().sort_values(ascending=False).head(4).index

top_4_currency_names_except_first=dx[dx['Symbol']!='BTC'].groupby(['Symbol'])['Marketcap'].last().sort_values(ascending=False).he
 top_4_currency_names_except_first_two=dx[(dx['Symbol']!='BTC') & (dx['Symbol']!='ETH')].groupby(['Symbol'])['Marketcap'].last().s
 top_4_currency_names_except_first_two_three=dx[(dx['Symbol']!='BTC') & (dx['Symbol']!='ETH') & (dx['Symbol']!='USDT')].groupby(['Sumbol'].groupby(['Sumbol'].groupby(['Sumbol'].groupby(]'])

data_top_4_currencies = dx[dx['Symbol'].isin(top_4_currency_names)]
 top_4_currencies_after_BTC_ETH = dx[dx['Symbol'].isin(top_4_currency_names_except_first_two)]
 top_4_currencies_after_BTC_ETH_USDT = dx[dx['Symbol'].isin(top_4_currency_names_except_first_two_three)]
```

```
plt.figure(figsize=(20,25))

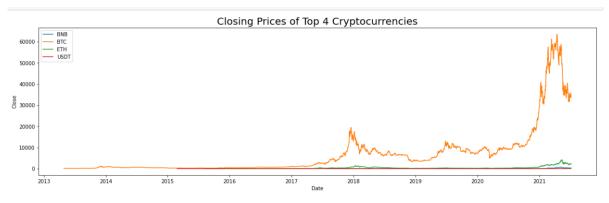
plt.subplot(4,1,1)
sns.lineplot(data=data_top_4_currencies, x="Date", y="Close", hue='Symbol')
plt.title("Closing Prices of Top 4 Cryptocurrencies", fontsize=20)
plt.legend(loc='upper left')

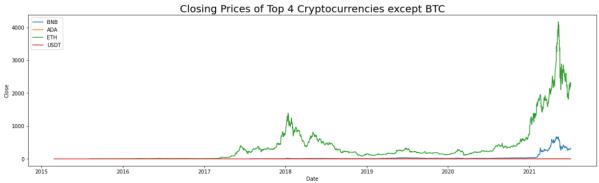
plt.subplot(4,1,2)
sns.lineplot(data=top_4_currencies_after_BTC, x="Date", y="Close", hue='Symbol')
plt.title("Closing Prices of Top 4 Cryptocurrencies except BTC", fontsize=20)
plt.legend(loc='upper left')

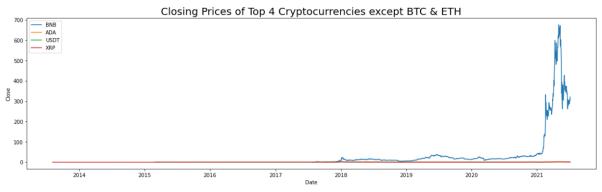
plt.subplot(4,1,3)
sns.lineplot(data=top_4_currencies_after_BTC_ETH,x="Date", y="Close", hue='Symbol')
plt.title("Closing Prices of Top 4 Cryptocurrencies except BTC & ETH", fontsize=20)
plt.legend(loc='upper left')

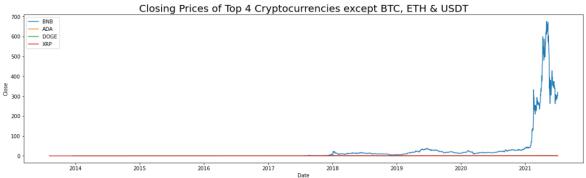
plt.subplot(4,1,4)
sns.lineplot(data=top_4_currencies_after_BTC_ETH_USDT,x="Date", y="Close", hue='Symbol')
plt.subplot(4,1,4)
sns.lineplot(data=top_4_currencies_after_BTC_ETH_USDT,x="Date", y="Close", hue='Symbol')
plt.title("Closing Prices of Top 4 Cryptocurrencies except BTC, ETH & USDT", fontsize=20)
plt.legend(loc='upper left')

plt.show()
```









Observation:

In 2021 the value of BTC was at all-time high of nearly 60,000\$ which is almost 15 times more than the second-highest cryptocurrency ETH.

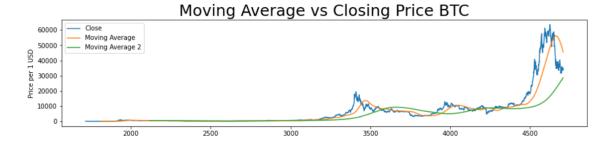
Among the top cryptocurrencies, the growth of BTC, ETH, and USTD over the last five years was beneficial for the investors.

Observation:

Investment in penny cryptocurrencies should be avoided, as depicted by the candlestick chart of USDT.

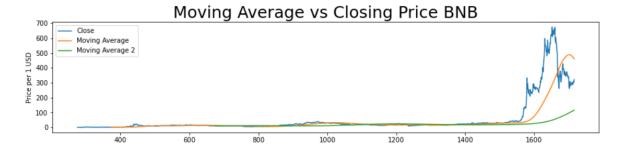
It is visible in the candlestick chart of USDT that this cryptocurrency is volatile for a longer period.

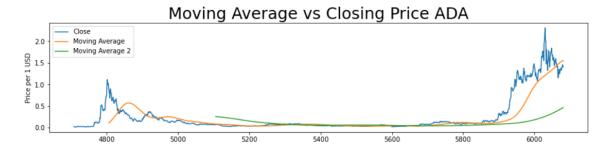
```
for i in ['BTC', 'ETH', 'USDT', 'BNB', 'ADA']:
    j=1
    dft=df[df['Symbol']==i]
    top_currency_name = dft.groupby(['Symbol'])['Marketcap'].last().sort_values(ascending=False).index
    top_currency = dft[dft['Symbol'].isin(top_currency_name)]
    top_currency['Moving Average']=top_currency['Close'].rolling(window=50).mean()
    top_currency['Moving Average 2']=top_currency['Close'].rolling(window=200).mean()
    plt.subplot(5,1,j)
    top_currency['Close'].plot(figsize=(15,18))
    ax=top_currency['Moving Average'].rolling(window=50).mean().plot()
    ax=top_currency['Moving Average 2'].rolling(window=200).mean().plot()
    ax.set_ylabel("Price per 1 USD");
    plt.title(f"Moving Average vs Closing Price {i}", fontsize=25);
    plt.legend()
    j=j+1
    plt.show()
```











Observation:

In BTC and ETH

They have shown a pattern of 2 golden cross and 2 death cross over the years. The last crossing point of moving averages was golden cross hence it was bullish.

In USDT

There are too many golden cross and death cross. Hence it is a volatile stock and should be avoided for the time being.

Conclusion:

Bitcoin is the highest value according to market cap and etherum is 50% of it and all other are very less in comperision to it and USDT is 3 times less then the Etherum

Investment in penny cryptocurrencies should be avoided, as depicted by the candlestick chart of USDT.

It is visible in the candlestick chart of USDT that this cryptocurrency is volatile for a longer period.

In 2021 the value of BTC was at all-time high of nearly 60,000\$ which is almost 15 times more than the second-highest cryptocurrency ETH .

Among the top cryptocurrencies, the growth of BTC, ETH, and USTD over the last five years was beneficial for the investors.

In BTC and ETH -- They have shown a pattern of 2 golden cross and 2 death cross over the years. The last crossing point of moving averages was golden cross hence it was bullish.

In USDT -- There are too many golden cross and death cross. Hence it is a volatile stock and should be avoided for the time being.

Web Scrapping:

```
datetime_start = datetime_end - relativedelta(hours = 12)
          #Api for the scrapping
         url = 'https://production.api.coindesk.com/v2/price/values/'+ coin +'?start_date='+datetime_start.strftime("%Y-%m-%dT%H:%
          #we are using the request to fetch the data from the api in the json format and then storing it into the dataframe.
          temp_data = requests.get(url).json()
         df = pd.DataFrame(temp_data['data']['entries'])
df.columns = ['Timestamp', 'Open', 'High', 'Low', 'Close']
          # To handle the Missing Data
         insert_ids_list = [np.nan]
         '''In numerous occasions, I observed that CoinDesk did not collect every minute of data. According to observation, in a normal circumstance, the timestamp would have a discrepancy
          value of 60000 for 1 minute. As a result, I can quickly tell that the missing period is the
         time gap if the row difference for the timestamp is larger than 60000. To deal with it, a hot-deck imputation procedure is used. To put it another way, the nearest minute data will
         be substituted for the missing one.
         while len(insert ids list) > 0:
               timestamp_checking = np.array(df['Timestamp'][1:]) - np.array(df['Timestamp'][:-1])
               insert_ids_list = np.where(timestamp_checking!= 60000)[0]
if len(insert_ids_list) > 0:
    print(str(len(insert_ids_list)) + ' mismatched.')
                    insert_ids = insert_ids_list[0]
                    temp_df = df.iloc[insert_ids.repeat(int(timestamp_checking[insert_ids]/60000)-1)].reset_index(drop=True)
temp_df['Timestamp'] = [temp_df['Timestamp'][0] + i*60000 for i in range(1, len(temp_df)+1)]
df = df.loc[:insert_ids].append(temp_df).append(df.loc[insert_ids+1:]).reset_index(drop=True)
                    insert_ids_list = insert_ids_list[1:]
          #adding datetime and symbol to dataframe
         df = df.drop(['Timestamp'], axis=1)
         coin_df['Symbol'] = coin
main_df = main_df.append(coin_df)
1 mismatched.
1 mismatched.
1 mismatched.
1 mismatched.
1 mismatched.
1 mismatched.
```

1 mismatched.

main_df = main_df[['Datetime', 'Symbol', 'Open', 'High', 'Low', 'Close']].reset_index(drop=True)
main_df

	Datetime	Symbol	Open	High	Low	Close
0	2021-07-01 00:00:00	втс	35049.045484	35056.817222	34991.326658	34993.994267
1	2021-07-01 00:01:00	BTC	34995.349202	35030.583041	34989.167695	34991.266464
2	2021-07-01 00:02:00	BTC	34991.950939	34994.431719	34929.671801	34952.450112
3	2021-07-01 00:03:00	BTC	34954.642324	34980.657519	34954.642324	34961.649826
4	2021-07-01 00:04:00	BTC	34960.350436	34996.245435	34960.350436	34976.586395
27355	2021-07-01 23:55:00	EOS	3.945295	3.945295	3.936671	3.939916
27356	2021-07-01 23:56:00	EOS	3.939917	3.940801	3.933735	3.933735
27357	2021-07-01 23:57:00	EOS	3.933920	3.934184	3.928797	3.933092
27358	2021-07-01 23:58:00	EOS	3.933122	3.939846	3.933122	3.939811
27359	2021-07-01 23:59:00	EOS	3.939591	3.939752	3.927281	3.927281

27360 rows × 6 columns

main_df.to_csv('main_df.csv', index=False)

Conclusion and Future Scope

In this project, I have proposed several methods in which comparative analysis was done and promising results were achieved for different crypto currency prediction and analysis. Many researchers have previously suggested that we should use ML where the dataset is not that large. The traditional methods which are used for comparison are confusion matrix, precision, specificity, sensitivity, and F1 score all of them have been implemented in algorithms itself for ease and faster calculations. There are several analysis mentioned in this project where different currencies performed best for different ones.

Obstacles

Cryptocurrencies such as BitCoin still have numerous significant obstacles to overcome before they could totally replace current currency systems. The most immediate is the simple opposition from existing financial institutions, which wield great power and have incentives to discourage the proliferation of cryptocurrencies. Other large corporations, even when amenable to the idea of cryptocurrencies, do not currently consider them stable enough to keep as assets for long periods of time. In addition to battling the current economic system, cryptocurrencies have some internal challenges to overcome. Attempting to convert the entire world financial system to the BitCoin model, for example, could cause such a massive growth in blockchain size that the distributed ledger model would become impractical. It is also still unclear whether blockchain technology could be successfully adapted to use cases which require very high speeds with high volumes (on the order of seconds instead of hours), and would be poorly suited for any application which required some degree of reversibility . Finally, because of the substantial energy costs and diminished rewards over time associated with the "mining" process, users may eventually be forced to bear increasingly high and unreasonable transaction costs.

Short-term (3-5 years)

Increasing Efficiency in the Financial Industry

Since the 2008 financial crisis, large banks are increasingly feeling pressure to increase efficiency and cut costs wherever possible. To that end, a May 2016 report from Goldman Sachs estimates that the financial industry alone could realize up to \$6 billion/year in savings through use of blockchain technology. However, this would not necessarily include decentralized cryptocurrencies such as Bitcoin, but may involve the creation of new proprietary centralized cryptocurrencies (such as the Bank of England's newly introduced RSCoin). The introduction of cryptocurrencies may also lead to increased levels of transparency and few incidents of fraud. Under current systems, the correct identification of fraud is very manual-labor intensive and prone to error. However, cryptocurrencies are designed to be explicitly transparent and automatically detect fraud, greatly alleviating the costs associated with managing associated systems.

Emerging Markets

Because cryptocurrencies require only an Internet connection, and are not dependent on established institutions such as banks, they are ideally suited for societies without a well-developed financial infrastructure . As with how many individuals emerging markets skipped over landlines and went straight for mobile phones, the same individuals may skip the overhead of the traditional banking system and engage directly in mobile banking . For these reasons, we expect cryptocurrencies to become a major influence in emerging markets over the next 3-5 years .

Long-term (5-10 years)

Financial Market Disruption

Within the cryptocurrency community, one of the most popularized goals is the total replacement of banks and other centralized financial intermediaries [4]. Although such institutions may never be fully replaced by a democratized network, their role (and associated profitability) may steadily diminish with rise of cryptocurrencies, hopefully leading to the prevention of future financial catastrophes on the scale of the 2008 crisis [7].

IoT Integration

Although cryptocurrencies have the possibility to replace functions of the existing financial infrastructure, their greatest potential may be in incorporating with other technologies to facilitate a true revolution . The blockchain model is ideally suited for Internet of Things (IoT) transactions, which require both efficient simplicity and robust security . For example, imagine if every time you needed to fill up a car with gas, your car could pay the gas station automatically.

Expanding Industries

In addition to revolutionizing the financial system, the blockchain technology of underlying cryptocurrencies has the potential to expand across nearly any industry that involves large-scale record-keeping. Blockchain could be a massive boon to proponents of effective protection of intellectual property rights, such as with music and film. New companies such as Ascribe are pioneering methods for creating secure limited copies of digital media, in order to ensure that artists are properly compensated for their work, instead of being financial damaged by pirates. Other examples include the growing "Sharing Economy" (including AirBnB) which can use blockchain to ease identity and reputation management, and "Smart Grid" utility companies which could use blockchain to introduce efficient microtransactions for energy consumption.

Far Future (10+ years)

In the very far future, global and democratized cryptocurrencies have the potential to replace government-backed fiat currencies as the primary means of conducting financial transactions. With that end in mind, Microsoft has also begun facilitating large-scale simulation tests on behalf of banks and other large corporations interested in understanding the potential ramifications for such a large-scale shift in the global economy .

References

Deployment Link:

 $\frac{https://app.powerbi.com/groups/me/reports/d8648abb-7f39-43b7-afb9-70ecbc68f418/ReportSection 6290258e7775596934b0}{afb9-70ecbc68f418/ReportSection 6290258e7775596934b0}$

Github Link:

https://github.com/AshuKesarkar25/CryptoBoard

How to deploy power bi project on web:

https://www.youtube.com/watch?v=_3YAB6dNQ18

Special reference for the project:

https://www.youtube.com/c/SatyajitPattnaik

Paper referenced in the project:

- 1. Ascribe home page. Ascribe. Accessed 29 September 2016 from https://www.ascribe.io/
- 2. Dutch Central Bank Presents Results of Cryptocurrency Experiments, by Michael del Castillo. CoinDesk. Accessed 29 September 2016 from http://www.coindesk.com/dutch-central-bank-cryptocurrency-experiments/
- 3. Citi: Bitcoin is an Opportunity for Banks, Not a Threat, by Pete Rizzo. CoinDesk. Accessed 29 September 2016 from http://www.coindesk.com/citi-bitcoin-banks-remittances-wont-disrupt/
- Goldman Sachs on blockchain, by Pete Rizzo. CoinDesk. Accessed 29 September 2016 from http://www.coindesk.com/goldman-sachs-blockchain-tech-save-capital-markets-12-billion/
- 5. https://hbr.org/2016/05/the-impact-of-the-blockchain-goes-beyond-financial-services
- Banks find Blockchain hard to put into practice, by Kadhim Shubber. Financial Times. Accessed 29 September 2016 from https://www.ft.com/content/0288caea-7382-11e6-bf48-b372cdb1043a

- 7. Microsoft Bets That Bitcoin-Style Blockchains Will Be Big Business, by Tom Simonite. MIT Technology Review. Accessed 20 September from https://www.technologyreview.com/s/545806/microsoft-bets-that-bitcoin-style-blockchains-will-be-big-business/
- 8. No, Big Companies Aren't Really Accepting BitCoin, by Jacob Davidson. Time. Accessed 29 September 2016 from http://time.com/money/3658361/dell-microsoft-expedia-bitcoin/
- 9. Central banks beat Bitcoin at own game with rival supercurrency, by Ambrose Evans-Pritchard. The Telegraph. Accessed 20 September 2016 from http://www.telegraph.co.uk/business/2016/03/13/central-banks-beat-bitcoin-at-own-game-with-rival-supercurrency/