**Inequality measures-based bitcoin price prediction**

Project submitted to the

SRM University – AP, Andhra Pradesh

for the partial fulfillment of the requirements to award the degree of

**Bachelor of Technology**

In

**School of Engineering and Sciences**

Submitted by

**D.Harika (AP19110010499)**

**K.Lakshmi Priya (AP19110010032)**

**K.Tejaswi (AP19110010158)**

**V.V. Manas (AP19110010506)**

**A picture containing text

Description automatically generated**

Under the Guidance of

**Dr. Sowkuntla Pandu**

**SRM University–AP**

**Neerukonda, Mangalagiri, Guntur**

**Andhra Pradesh – 522 240**

**May- 2023**

I

# 

# 

# **Certificate**

Date: 25-May-23

This is to certify that the work presented in this Project entitled “**INEQUALITY MEASURES-BASED BITCOIN PRICE PREDICTION**” has been carried out by D.Harika, K.Lakshmi Priya, K.Tejaswi, V.Manas under my supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of Bachelor of Technology/Master of Technology in **School of Engineering and Sciences**.

**Supervisor**

(Signature)

Dr. Sowkuntla Pandu,

Assistant Professor,

Department of Computer Science and Engineering

II

**Acknowledgements**

I want to convey my sincere gratitude and appreciation to everyone who helped my final project to be completed successfully. The completion of this project would not have been possible without their assistance, direction, and encouragement.

First and foremost, I want to express my sincere gratitude to **Dr. Sowkuntla Pandu**, who served as my supervisor, for their important advice and ongoing assistance during the whole project. Their knowledge, understanding, and helpful criticism have been crucial in guiding my research and improving the standard of my work. I sincerely appreciate their effort and continuous commitment.

I want to express my sincere gratitude to the head of the department **Dr. Jatindra Kumar Dash** and the course coordinator **Dr. Sambit Kumar** Mishra to facilitate an atmosphere of academic excellence and for providing the resources and support we needed for our final project. Your vision and leadership have been instrumental in creating an atmosphere conducive to learning and research.

I heartily thank all the staff members and librarians at SRM University, who helped me, find research papers, journals, and other resources for my project. Their assistance and willingness to go higher than before were essential for the success of my project.

I want to express my gratitude for the assistance I received from friends on this journey. Their encouragement, discussions, and brainstorming sessions have been incredibly beneficial in shaping my ideas and keeping me motivated during challenging times.

III

**Table of Contents**

Certificate……………………………………………………………………………………II

Acknowledgements III

Table of Contents IV

Abstract V

Abbreviations VI

List of Tables VII

List of Figures VIII

List of Equations IX

1. Introduction…………………………………………………………………………........10

2. Existing Work……………………………………………………………………..……...11

3. Proposed work 13

3.1. Motivation………………………………………………………………………………...13

3.2. Proposed inequality measures………………………………………………………….14

3.2.1 .Inequality Measures……………………………………………………………...15

3.2.2. Algorithm………………………………………………………………….............18

3.2.3. Limitations of proposed work …………………………………………………..20

4. Experimental Analysis…………………………………………………………………...21

4.1. Experimental setup……………………………………………………………………....21

4.2. Data description…………………………………………………………………………..21

4.3. Discussion…………………………………………………………………………..…....22

4.3.1. Mean and standard deviation……………………………………………...23

4.3.2. Lorenz curve……………………………………………………………….24

4.3.3. Gini index and k value…………………………………………………..25

4.3.4. Evaluating Results………………………………………………………………26

5. Conclusion……………………………………………………………………………….29

6. Reference…………………………………………………………………………………30

IV

# **Abstract**

# Bit coin, is the leading crypto currency, has gained significant attention and popularity in recent years. As its market value continues to fluctuate, accurately predicting Bit coin prices has become a topic of great interest. This project aims to explore the feasibility of using machine learning models to forecast Bit coin prices and contribute to the field of crypto currency analysis. Due to the volatility behavior, decentralized nature, and liquidity of this digital asset, trading is attracting attention. The extreme price volatility of crypto currencies has an impact on international relations and trade. Virtual currencies include things like bit coin, ripple, Ethereum, Ethereum classic, lite coin, and others. Many prediction models were done using the time series forecasting, neural network models and some other machine learning techniques over the direct data of the bit coin prices that is available. These models were not so accurate because of the price volatility, lack of indexes, technological progress, and political factors. In our study, we especially focused on inequality measures as the features to train the ML model. Here, we explored Lorenz curve, Gini index and K-value, as inequalities measures and skewness, Kurtosis as statistical measures to train the machine learning based prediction model. We have done an extensive experimental analysis on a real time dataset. From the experimental results we have observed that our proposed model is predicting the price of bit coin accurately.

# 

# 

V

# **Abbreviations**

ML Machine learning

ARIMA Autoregressive Integrated Moving Average

SD Standard Deviation

RNN Recurrent Neural Network

LSTM Long-Short Term Memory

KNN K-Nearest Neighbors

# 

VI

# 

# **List of Tables**

Table 1: 4.1 Experimental setup → Hardware Requirements………………………..21

Table 2: 4.1 Experimental setup → Software Requirements………………………….21

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

VII

# **List of Figures**

Figure 1: Proposed model....................................................................................................14

Figure 2: Lorenz curve……………………………………………………………………..15

Figure 3: Gini index………………………………………………………………………...16

Figure 4: K-Value…………………………………………………………………………...17

Figure 5: Visualizing the price distribution over the Date …………………………….23

Figure 6: Rolling Mean & Standard Deviation among data…………………………....24

Figure 7: A graphical representation of the distribution of price & days…………….25

Figure 8: Calculating the Gini-index & K-value of the Bitcoin stable market………..26

Figure9: Random Forest Regressor Model with 93% Accuracy……………………….27

Figure 10: Linear Regression accuracy with 99%..............................................................28

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

VIII

# 

# **List of Equations**

Eq-1: K-value……………………………………………………………………17

Eq-2: Standard deviation (SD) ………………………………………….23

Eq-3: Gini coefficient..........................................................................................25

IX

**1. Introduction**

Bit coin is the most famous crypto currency, which was created in January 2009. The main concept of Bit coin is block chain, in which blocks are decentralized and a secure shell algorithm is used to link the blocks, and they serve as non-editable data that is recorded while a transaction is in progress. Block chains are organizational techniques for preserving the integrity of transactional data, and they are crucial to the development of various crypto currencies. The bit coin currency has shown a significant increase in market capitalization in recent years and is on over 40 exchanges worldwide. Bit coin facilitates quick payments by using peer-to-peer technology, in which all transactions are not regulated or controlled by any third party, unlike the other currencies. There is no government, no central bank that can shut down or arbitrarily raise or lower the value. Miners are in charge of processing block chain transactions and are motivated by repo fees for finding the combination of nonce codes.

Although bit coins are widely used for various purchases and transactions throughout the world, there is no consensus on what constitutes bit coins or their legal standing. Investing in the Bit coin market is comparable to investing in the stocks market in that none of the risk factors that explain stock price swings apply to crypto currencies. Bit coin has lately received much attention in economics, cryptography, and computer science due to its inherent nature of combining encryption technology with financial units. So with the increase in the number of people owning bit coins and with the high price volatility of bit coins, people started using them as an investing market along with using it as a trade option.

More people started investing in the crypto currencies to make profits from them. But, as the price of bit coin is more volatile, there is a chance that people can even get losses rather than getting a profit. So, an intelligent investor will always look over the predictions based upon the past data available, before making any investments. These Predictions will allow the Investors to avoid risks and improve the returns. To forecast the price changes of Bit coin, experts and traders employ a variety of strategies and techniques. In this context, algorithms for machine learning and deep learning are being used more frequently to predict Bit coin prices. Predictions are made using these techniques by analyzing historical data, market movements, news sentiment, and other relevant elements. We will examine some of the approaches and strategies, such as fundamental analysis, technical analysis, and machine learning algorithms, used to forecast Bit coin values in this project.

10

**2. Existing works**

Since bit coin is the most expensive crypto currency in the world right now and an innovative technology, there are some methods for predicting its value. Amjad et al. [1] analyzed historical time series pricing data to predict prices using ARIMA model and develop an exchanging strategy. L. Kristoufek[2], Later analyzed the Bit coin market as being made up entirely by speculative traders without any fundamentalists, and determined the relationship between Bit coin and the volume of research on Wikipedia and Google Trends. His analysis established an apparent difference in the relationship between the price of Bit coin and search terms, proposing that trend-chasing and hypothetical behavior drive Bit coin’s price dynamics in the crypto currencymarket [2].

Using machine learning methods such as recurrent neural networks (RNNs) and long short-term memory (LSTM), McNally [3], forecasts the Bit coin price mechanism and compares the outcomes to those produced using autoregressive integrated moving average (ARIMA) models. According to H. Jang [4] the author is reportedly evaluating what has been done to forecast the U.S. Bit coin market. In his research, the prediction network's mean square error was equal to the excess return's standard deviation. However, the author is demonstrating that a number of fundamental market excess returns can be predicted by financial and economic parameters.

The author Zheshi Chen, Chunhong Li. [8], "Predicting the Price of Bit coin Using Machine Learning": This study examines how to forecast Bit coin prices using machine learning techniques like Support Vector Machines (SVM), Random Forests, and Gradient Boosting. Technical indicators, market psychological states info, and Google Trends data are just a few of the many features that are taken into consideration.

The author Z. Chen et al. [15], "An Ensemble Learning Approach for Bit coin Price Prediction": In order to forecast Bit coin values, the authors suggest an ensemble learning framework that integrates various machine learning algorithms, including ARIMA, SVM, and LSTM. In order to choose the features that are most useful for prediction, they also offer a feature selection technique.

The author D.Shah. [1], "Predicting the Price of Bit coin Using Data from Twitter and Google Trends": This study investigates the use of sentiment analysis to forecast Bitcoin prices using data from Google Trends and Twitter. It makes use of

11

machine learning methods like LSTM networks and Random Forests and evaluates how adding search and social media data affects prediction accuracy.

From the aforementioned existing works, we have observed that the primary difficulty with the bit coin exchange rate is its fast pace of price volatility. Because of the high price volatility, it is necessary to take precautions to estimate bit coin’s price effectively. To predict the future price of bit coin and to establish confidence and acceptability throughout the world, it is vital to understand forecasting activities. A range of elements, such as a country's political system, public relations, and market policy, can influence bit coin’s economic role and international interactions on various market strategies. Lastly, one more concern is that there is no clear description of the exchange platforms on which the transactions related to buying and selling Bit coins are regulated. So, the objective of our current project is to forecast the future price of bit coin with improved efficiency using some of the statistical features such as the Gini Index and the K-Value and the Machine learning models to minimize the risks for investors and policy-makers.

12

**3. Proposed Work**

This section discusses the proposed inequality measure based bit coin price prediction model. In Subsection 3.1 the motivation to propose this model is given. And in Subsection 3.2, the proposed model is discussed in detail.

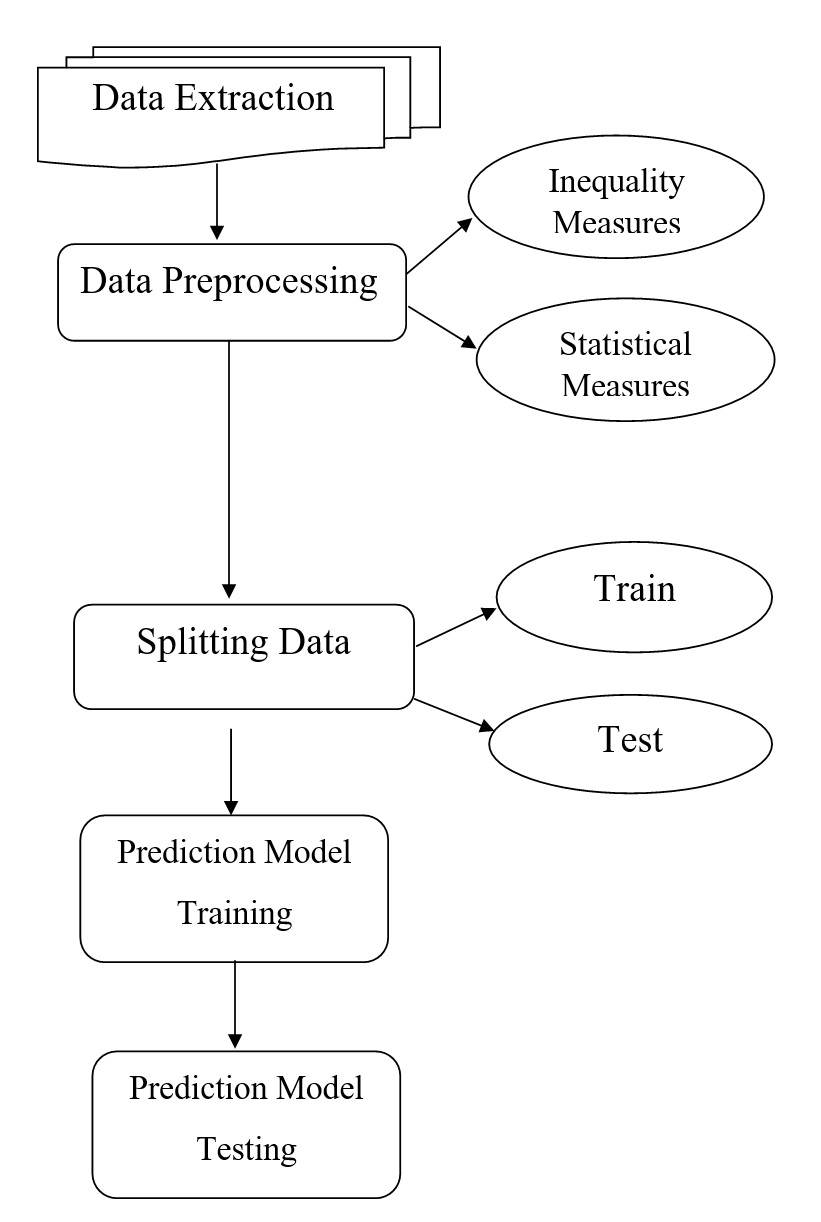
**3.1. Motivation**

Bit coin, being the first and most well-known crypto currency, has attracted significant attention from individuals, investors, and researchers. Bit coin has become a desirable asset for traders and investors because of its volatile nature and potential for large fluctuations in prices. Machine learning advancements: Predicting the price of bit coin is an excellent usage for these techniques. Scholars can use the project to investigate and use different machine learning techniques, such as regression, time series analysis, or deep learning, to forecast and examine changes in the price of bit coin.

The Bit coin market is an unrestricted market; hence the currencies which are not regulated will follow a certain pattern in their inequality measures. These patterns, when used as features, help in the machine learning model to improve the accuracy of the prediction.

To train the machine learning model, we will collect the bit coin dataset from Investing.com [19] in a period from 1st January 2012 to 31th January 2023. Then we will prepare our dataset by using data processing techniques which is one of the most crucial parts of the machine learning techniques. In the dataset, we have chosen six attributes: the open price, high price, low price, close prices,change,volume and the market cap of publicly traded outstanding shares. So, first, we will import the data and then data has been processed by removing null values and adding the additional required features that are needed for predictions to be more accurate. Economic inequality measures like the Gini Index and the K-Value are included here as additional features that are important to consider when analyzing the pricing of any economic entity. By integrating these extra features as variables when training the ML Model, we will be able to accurately anticipate the stability of the bit coin market. Here, we have used the random forest machine learning model to forecast time series using the features mentioned above, with the result being the closed price. 13

**3.2. Proposed inequality measures based bitcoin prediction model**



**Figure 1**: Architecture of proposed model

As shown in Figure-1, the imported data is processed by removing all the null values from the dataset and adding the additional required features like inequality measures and statistical model to obtain more accurate predictions. After that, we have created a new array with the values consisting of the closed price difference between the consecutive days over which the Inequality measures will be calculated.

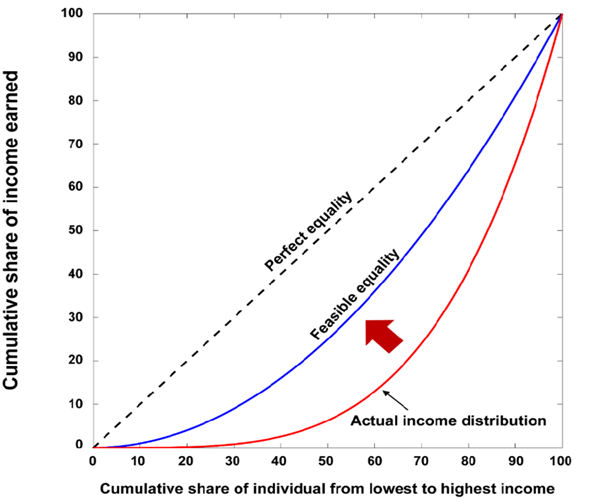
Here we have economic inequality measures, such as the Gini Index and the K- Value as the additional features which are considered as important factors while studying the price of any economic entity and adding the statistical models to overcome the biased models. 14

These additional features will determine the stability of the bit coin market, thereby helping us to make accurate predictions by using them as features in training the ML model. In the following subsection 3.2.1, a brief description of the inequality measures that are used in this project is discussed.

**3.2.1. Inequality measures**

**Lorenz Curve**

Max Lorenz, a famous mathematician, introduced the Lorenz curve in 1905 to measure income inequality. It is one of the important statistical features which are a widely used measurement in statistics of economic prices. It is known as a graphical presentation of the distribution's inequalities. It is plotted between the cumulative price (closed prices differences) on the vertical axis and the cumulative population (number of days) on the horizontal axis. The Lorenz curve is accompanied by a straight line with a slope one that describes that the observations are perfectly equally distributed. From the curve we can conclude that the closer the curve is to the straight, which is perfect equality, the less income inequality. The Gini coefficient is a scalar value which is the measurement of Inequality. This curve is often used in calculating the inequalities of a country.

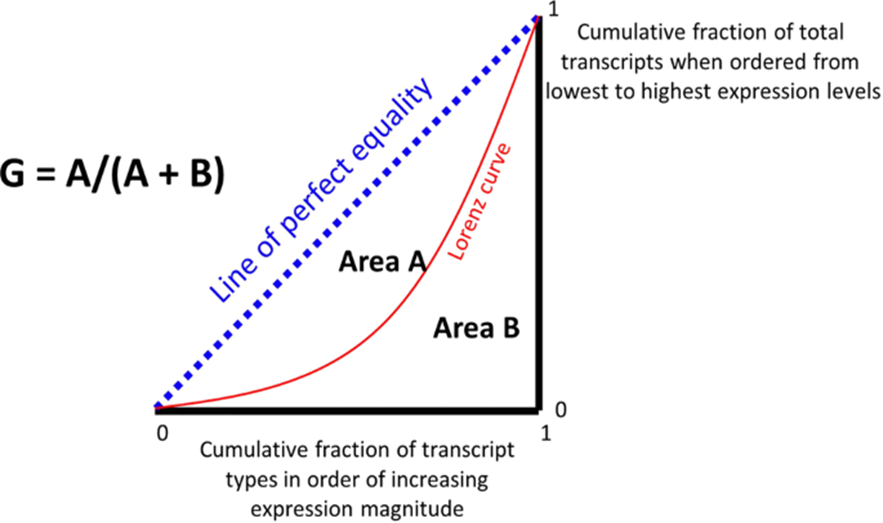


**Figure 2:** Lorenz Curve

15

The above Fig 2 is the Lorenz curve graphical representation of the distribution of price over the date, but more of the scalar measurements are needed to analyze and predict the bitcoin prices. In this situation, Gini Coefficient and K-Value come in as reasonable helping measures, which are scalar that make use of the Lorenz curve to give a scalar measure of the Inequality.

#### **Gini Coefficient**

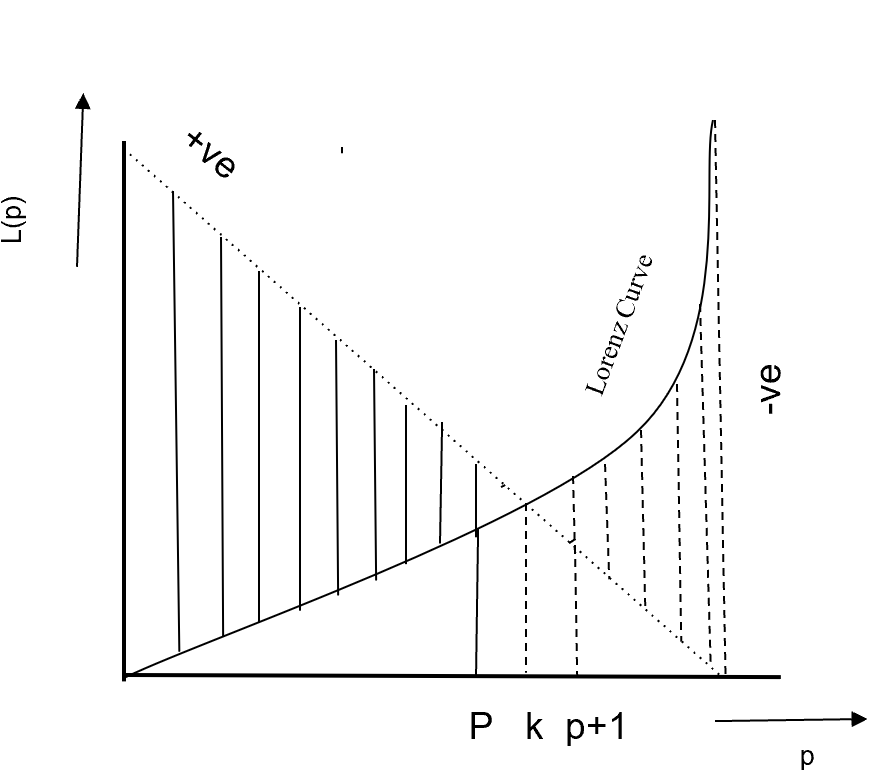
An Italian statistician Collado Gini developed the Gini coefficient (G) in 1912. It is often used as income distribution, or less commonly, a measure of economic Inequality for measuring the distribution of population wealth. The Gini coefficient values range from 0 to 1, where 0 is perfect equality, and 1 is perfect Inequality. Gini Coefficient uses the Lorenz curve to calculate from with the formula mentioned below in Fig 3. If every person in a country earns the same income, as there is no price distribution, we will have a Gini Coefficient of 0, but if a person earns the most income and others earn nothing, they will have a Gini Coefficient of 1. Gini Coefficient is known as the area between the Line of perfect equality divided by the triangle area.

**Figure 3:** Gini index

**K-Value**

K-Value is also calculated from the Lorenz Curve. A 1 - K fraction of an event has a K fraction of value. When a Gini Coefficient and K-Value meet each other, then that situation is known as a stable market condition. That means the market is stable when the Gini coefficient and K-values meet each other. 16

Formulae for k-value is π+1)/2

****

**Figure 4:** K-Value

**Plotting Lorenz curve, Gini Index and K-Value**

As we have the data of the Bit coin Price extracted between the years 2012 to 2023. We used the above mentioned formulas and techniques to calculate the Gini Index and K-Value from the Lorenz Curve. The dataset that we extracted contains the features such as date, opened price, highest value, and lowest value. Now as mentioned above we use the absolute difference between the two consecutive Closed prices and store those values in an array from which we will plot the Lorenz curve. In order to plot the Lorenz curve, we calculate the cumulative sum of closed prices difference in the time series divided by the total sum of closed prices difference along the Y-axis and the time series of days divided by the total number of days in the X-axis. After calculating these values we will store them in a separate array and use a plot function between those two arrays, which gives us the needed Lorenz curve.

17

**3.2.2. Algorithm: Inequality measure based prediction model**

1. Start
2. Read the **Bitcoin.csv** file into a **Pandas DataFrame**.
3. Calculate the difference between the **closed prices** of two days.
4. Calculate the percentage change in the closed price from one day to the next.
5. Create a list of the percentage changes in the **closed price**.

**For each price** in **Closed\_diff**

**Sum ← 0**

**Sum ← Sum + price**

**y\_axis.append(sum/total\_sum);**

**x\_axis.append(count/n);**

1. Create a list of the cumulative percentage changes in the **closed price**.
2. Plot the cumulative percentage changes in the closed price as a **Lorenz curve**.
3. Create a list of days to plot the **gini coefficients** and **k values** over the time.
4. Create a list of **Gini coefficient**s by using the gini function over the closed prices

**for each** iteration over the number of days

**sum1 ← sum2+((( y\_axis1[i+1]+ y\_axis1[i] )/2)\*(x\_axis1[i+1]-x\_axis1[i]))**

**ginival ←1- (2 \* sum1)**

1. Create a list of k values by using the kv function over the closed prices.

**For each** iteration over the number of days

**D ← (1-x\_axis1 [i]) - y\_axis1 [i]**

**If d <0 then**

**Value ← (x\_axis1 [i-1] + x\_axis1 [i])/2**

**Break;**

**diff.append(d)**

**End if**

1. Plot the **Gini coefficients, k values,** and **0.87** values over time.
2. Append the newly calculated values as features to the existing data frame.
3. Calculating the statistical measures as a rolling window over the price distribution.
4. Split the dataset into training and testing and then define the **X\_train** and **y\_train** to store the training data and labels.
5. Create a **RandomForestRegressor** object with **n\_estimators** set to **100** and random\_sate set to **0**.
6. Fit the model to the training data.

**regressor.fit(x\_train, y\_train.values.ravel())**

18

1. Predict the target values for the test data.

**regressor.predict(x\_test)**

1. Calculate the **r2**score

**r2\_score (y\_test,e\_dataframe)**

1. End

Calculate total\_sum

As given in the above algorithm, statistical measurements and machine learning techniques play a major role in analyzing Bit coin price changes. Read the historical price data from the Bitcoin.csv file into a Pandas DataFrame. Determine the daily price change by calculating the difference between the closed prices of two consecutive days. Calculate the percentage change in the closed price from one day to the next, which represents the daily price return. Make a list of the percentage changes in the closing price that will be used to determine the cumulative percentage changes. Iterate through the list, adding each price change to the previous amount to compute the cumulative sum. By dividing the cumulative amount by the sum of all price changes, you can calculate the cumulative percentage change.

To see the distribution of price changes, plot the total percentage changes in the closed price as a Lorenz curve. To plot the Gini coefficients and k values across time, make a list of the days. Apply the gini function on the closed prices to determine the Gini coefficients. The Gini coefficient is a measure of price distribution inequality. Apply the kv function to the closed prices to determine the k values. The k value shows where the distribution crosses the 0.87 level on the Lorenz curve. To examine the price distribution and inequality, plot the k values, the 0.87 values, and the Gini coefficients with time. Add the newly generated k values and Gini coefficients as features to the current DataFrame. Calculate statistical measurements to create a rolling window over the price distribution that reveals trends and volatility.

Calculate statistical measurements to create a rolling window over the price distribution that reveals trends and volatility. Divide the dataset into training and testing sets, and then define the variables X\_train and y\_train to store the labeled training data. Fit a Random Forest Regressor model to the training set using 100 estimators. Apply the trained model to the test data to predict the target values (price changes). Measure the quality of fit between the expected and actual values by computing the r2 score. The algorithm concludes with the results and an examination of the fluctuations in the price of bit coin.

19

**3.2.3. Limitations of proposed work**

Inequality is often measured using techniques that provide a simple view of inequality, such as the Gini coefficient or income quintiles. Important distinctions and variances within multiple categories or geographical areas may be missed by these measurements.

The availability and quality of data have a significant impact on inequality measurement. Many times data is missing, incorrect or out of date, making it difficult to correctly represent the true amount of inequality.

20

**4. Experimental Analysis**

In this section, extensive experimental analysis is provided.

**4 4.1 Experimental setup**

* **Table 1: Hardware Requirements**

| **Category** | **Description** |
| --- | --- |
| Processor | Minimum i3 Dual core |
| Internet | Ethernet connection(LAN)/wireless adapter(Wi-fi) |
| Hard Drive | Minimum 100GB; Recommended 200 GB or more |
| Memory(RAM) | Minimum 8GB; Recommended 32 GB or above |

* **Table 2: Software Requirement**

| **Category** | **Description** |
| --- | --- |
| Domain | Machine Learning |
| Programming Language | Python |
| Tools & Libraries | Numpy, pandas, math,matplotlib |
| IDE | Anaconda / Jupiter Notebook |
| prerequisites | Python Programming language, Plotting & visualization, Machine Learning Evaluation. |

**4.2. Dataset description**

The data that we are using here to train a machine learning model is extracted from the website known as investing.com. From that website, we have collected the data between the years 2012 and 2023. The data consists of features such as the date, price, open price, highest value, lowest value, and change in percentage, volume. In addition to them, we have calculated the inequality measures & statistical measures as additional features such as the Gini Index, K-Value, Mean, Standard Deviation, and Kurtosis. 21

These features are very important in the study of the price predictions of economic entities that are not regulated by any government organizations such as crypto currencies.

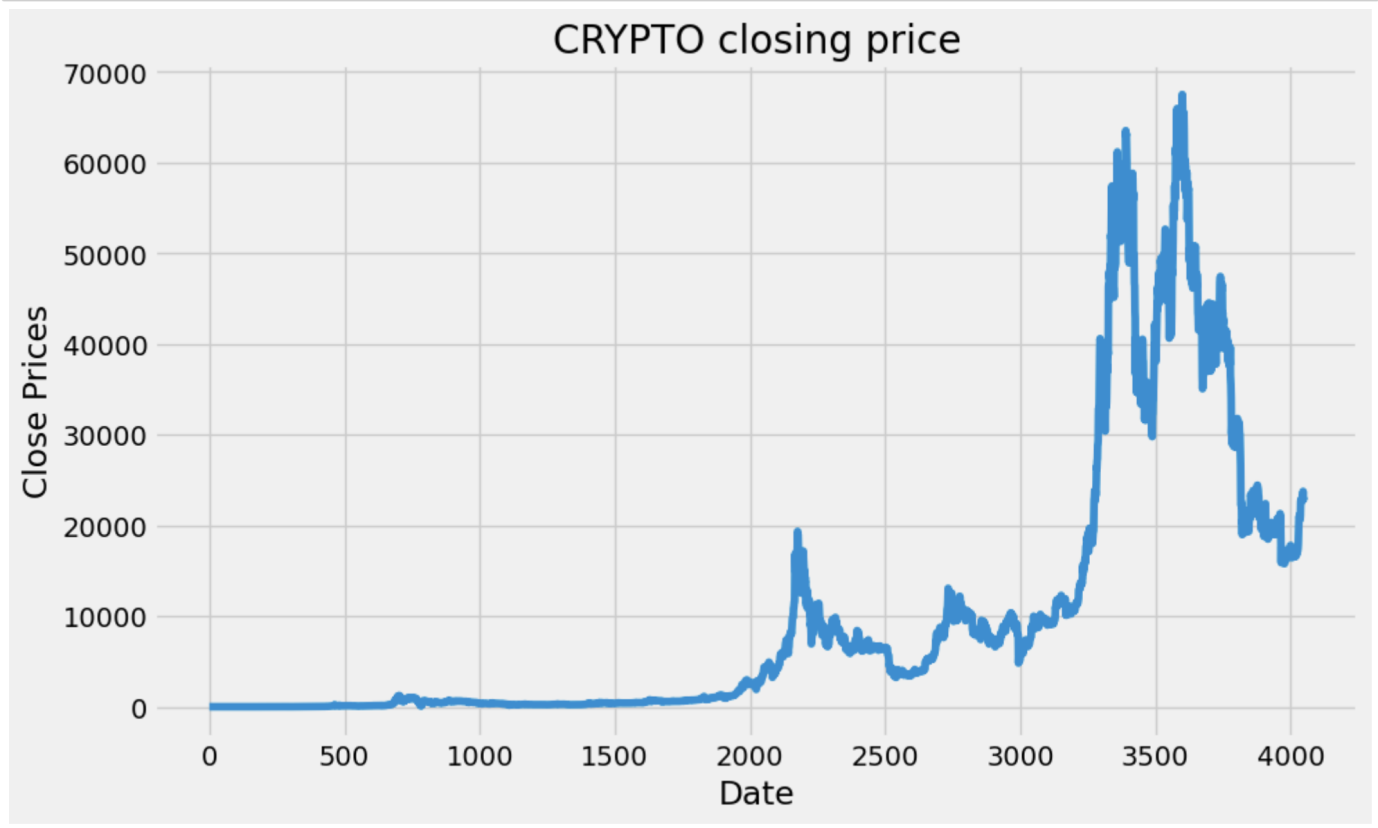
Once the data is extracted, then required features are added to the dataset to preprocess and clean the data. All the null values in the data are dropped and the data type of the features is transformed such that they can be used to train the regressor models.

**4.3. Discussion**

As we can see in Fig-5, the Bit coins market is an unrestricted market, which means that it is not regulated by any external organizations or government bodies. Many other crypto currencies, like Bit coins, are not regulated. So, all these currencies which are not regulated will follow a certain pattern in their inequality measures. These patterns, when used as features, will help the machine learning model improve the accuracy of the prediction.

The curve that is plotted between the Gini value and the K-Value will tend to coincide near the value of 0.87, which is the characteristic that these types of unrestricted currencies will follow. The graph will also show how crypto currency prices behave; if both the Gini and K-Value patterns tend to converge, the market is becoming more stable; if they diverge, the market's price will suddenly rise or fall. So, as these two values describe the nature of the crypto currency market for all unregulated currencies, they are prominent to use along with the other features while training the machine learning algorithm.

22



**Figure 5:** Visualizing the price distribution over the Date

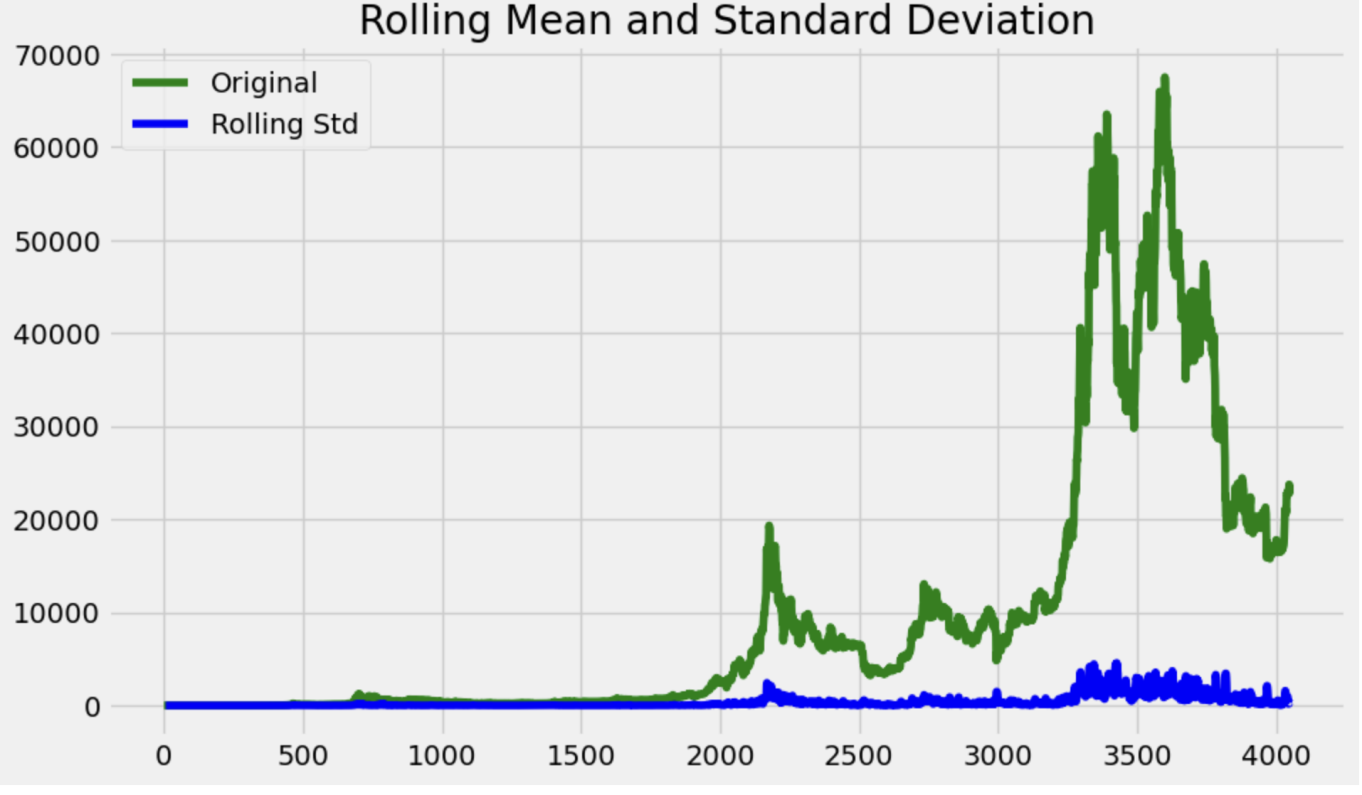
**4.3.1 Mean and Standard deviation**

The rolling mean, commonly referred to as the moving average, is a statistical technique used to analyze data over a given time period. When the rolling mean is examined over time, it becomes easier to identify long-term trends and changes in the data. An upward or downward slope in the rolling mean indicates a sustained increase or decline in the data, respectively. Future values can be predicted using the rolling mean as a base. By extending the rolling mean into the future, it provides a simplified prediction of the data's future behavior, assuming the underlying trend remains consistent.

A statistical measure known as the standard deviation can be used to determine how much variation or dispersion there is in a set of data. It reveals how far apart the data points are from the mean value. . A lower standard deviation denotes less variability, whereas a higher standard deviation predicts more variability.

Standard deviation (SD)= √/N

23



**Figure 6:** Rolling Mean & Standard Deviation among data

In the above figure, X-axis indicates the Date and Y -axis indicates the Price

**4.3.2. Lorenz curve**

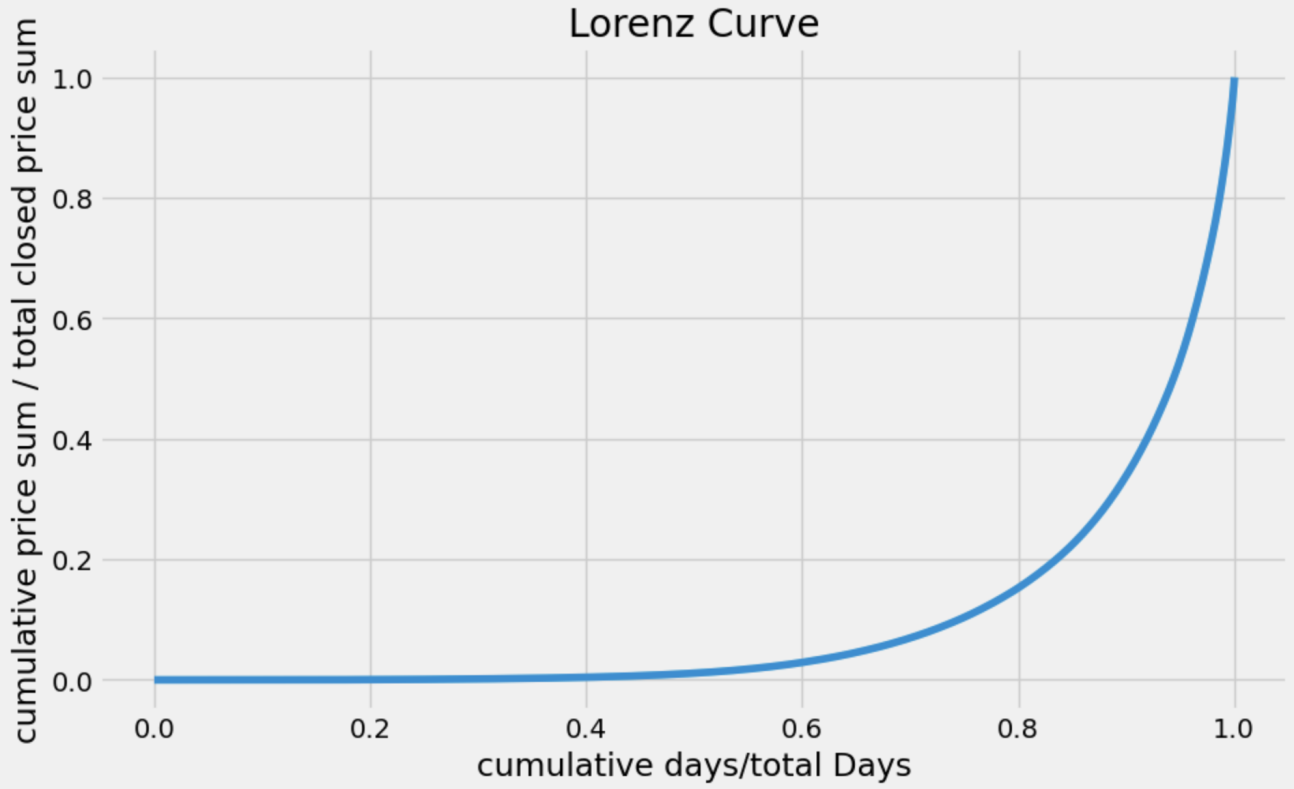
In the fields of society and economics, the Lorenz curve is a typical graphical representation of income distribution within a population. It assists in visualizing the level of income inequality in a society. Here, we have implemented this Lorenz curve as the price change of bit coin over the date and considered as data on the X-axis and price on the Y-axis.

On the X-axis we had represented day 1 as day1/ total number of days and day 2 as day 2/ total number of days and so on ... .up to 4049 as 4049/4049=1.

On the Y-axis we had represented the price as p2-p1/total price for p1 and p3-p2/total price for p2 and so on…..

The Lorenz curve is frequently represented by an upward-sloping, but typically exponentially rising curve. It is shown as an unbroken line in Fig 7.

24



**Figure 7:** A graphical representation of the distribution of price & days

**4.3.3. Gini and K-value**

The Gini index is determined as the ratio of the area between the perfect equality line and the Lorenz curve (A) divided by the total area under the perfect equality line (A + B).

From Fig - 3 which is derived from Lorenz curve

Area of triangle = (A+B)

Area of curve = A

Gini coefficient = A/ (A+ B)

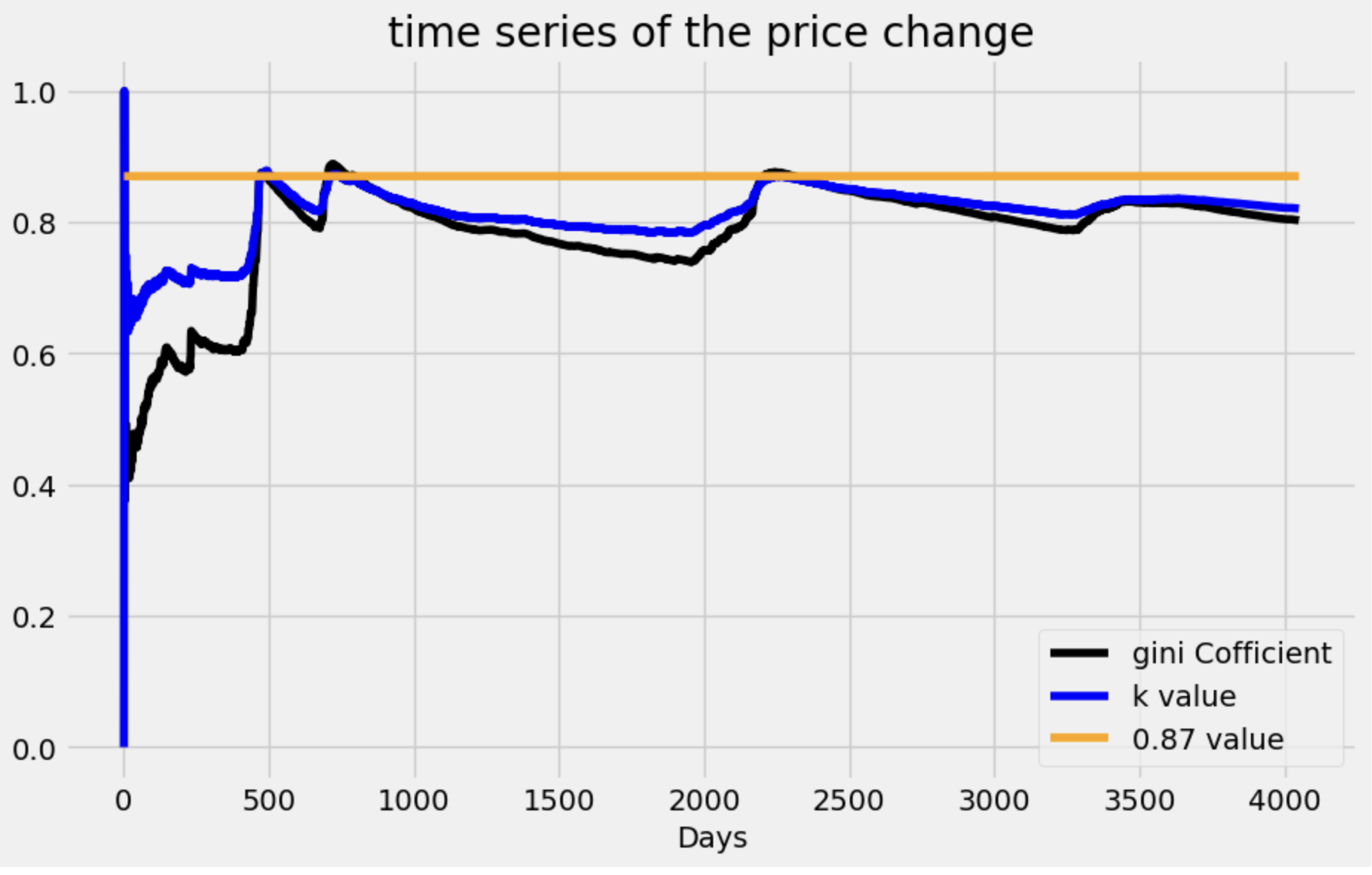
The k-value identifies a particular point on the Lorenz curve that is used to calculate the Gini index.

K-value Determination: The Lorenz curve's intersection with the vertical line leading from the point of absolute equality is known as the k-value. Mathematically, it indicates the cumulative share of income at a certain percentile of the population.

We wanted to determine when the market became stable after calculating the Lorenz curve, and for that both K-value and Gini computations are crucial. But in order to determine when they intersect, we must choose a statistical measure as a rolling window and compute the Gini and k-values for that.

25

We made a list of Gini and K-values, where index one contains the first point's Gini index and index two contains the first two points' K-values and Gini coefficient values. The K-value and Gini Coefficients of the first three points on the Lorenz curve are contained in the third index, and so on.



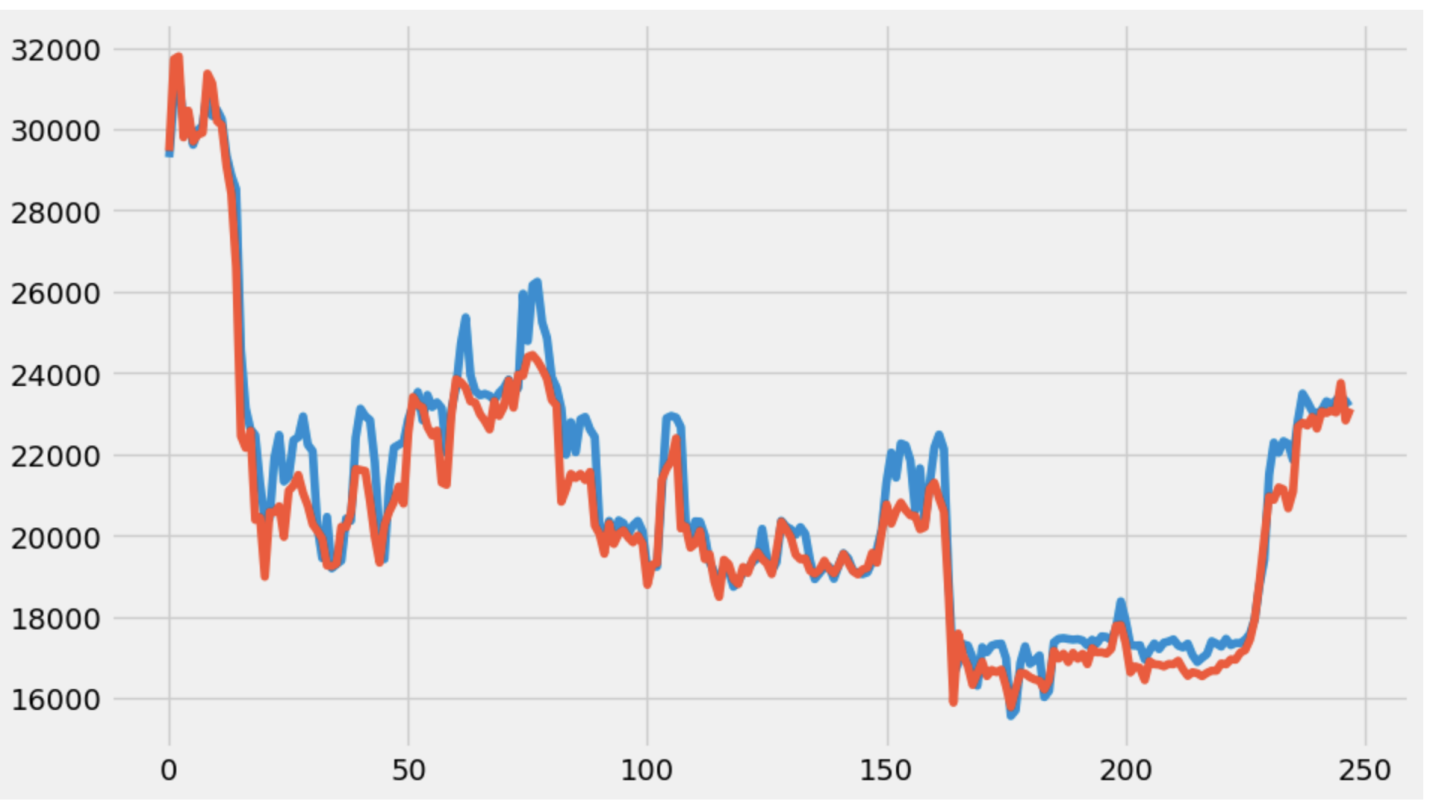
**Figure 8:** Calculating the Gini index & K value of bitcoin stable market

**4.3.4. Evaluating Results**

Linear Regression is one of the easiest Machine Learning Algorithms. Linear Regression model is developed in the field of statistics and is used for finding relationships. So, in statistical modeling, regression analysis is a set of statistical processes that are used for estimating the relationship among the input and the output variables. The relationships were estimated between the one or more dependent variables and the single independent variable. Finally, a function over the independent variables is estimated which is known as a regression function that is used for future predictions.

26

It is a supervised machine learning algorithm and In linear regression the machine learning model will find the relationship between one or more dependent variables(x) with respect to the single independent variable(y) over a regression function. The linear regression is mostly used in statistics for finding the relationships between the variables and also used for forecasting purposes.The representation of a linear regression is so simple where it combines a specific set of input values to an output value based on an equation as such both input and output values are numeric. Fig-6 : Random forest regression with and 93% accuracy



**Figure 9:**RandomForest Regressor Model with 93% Accuracy

27

****

**Figure 10:** Linear Regression accuracy with 99%

28

**5. Conclusion**

Bit coin is the most volatile entity in the world. Already Existing machine learning models are there but all these models are not giving much accuracy in predicting the bit coin price. Hence, we proposed a machine learning model which uses inequality measures as well as statistical methods to predict the bit coin price. From the experimental analysis, we observed that our proposed model is predicting the bit coin price more accurately than already existing models.

To evaluate inequality within the Bit coin ecosystem, researchers might study and develop more thorough measurements. Additional aspects like transaction volume, network participation, and mining concentration may need to be taken into account in this. We are planning to apply another crypto currency such as ethereum, doge coin and also wanted to compare how the inequality measures will be for different currencies. The dynamics and possible effects on the market can be better understood by analyzing the long-term patterns in inequality in the Bit coin ecosystem. Researchers can look at how wealth distribution changes over time, determine the underlying causes of these changes, and research any relationships between patterns of inequality and market activity.

29

**References**

[1]. M. Amjad and D. Shah, "Trading Bit coin and Online Time Series Prediction," in NIPS 2016 Time Series Workshop, 2017.

[2] L. Kristoufek, Bit coin meets Google trends and Wikipedia: quantifying the relationship between phenomena of the internet era, Sci. Rep. 3 (2013) 3415.

[3] S. McNally, J. Roche and S. Caton, "Predicting the price of Bitcoin using machine learning," in 26th Euro micro International Conference on Parallel, Distributed and Network-based Processing (PDP), 2018

[4] H. Jang and J. Lee, "An Empirical Study on Modeling and Prediction of Bit coin Prices with Bayesian Neural Networks based on Block chain Information," in IEEE Early Access Articles, 2017.

[5] <https://arxiv.org/ftp/arxiv/papers/2006/2006.14473.pdf>

[6]<https://towardsdatascience.com/using-machine-learning-to-predict-future-bitcoin-prices-6637e7bfa58f>

[7]<https://www.analyticsvidhya.com/blog/2021/05/bitcoin-price-prediction-using-recurrent-neural-networks-and-lstm/>

[8] Zheshi Chen, Chunhong Li, Wenjun Sun, *Bitcoin price prediction using machine learning: An approach to sample dimension engineering, Journal of Computational and Applied Mathematics*, Volume 365,2020,112395, ISSN 0377-0427,<https://doi.org/10.1016/j.cam.2019.112395>.

[9] [Muniye, Temesgen & Rout, Minakhi & Mohanty, Lipika & Satapathy, Suresh. (2020). *Bitcoin Price Prediction and Analysis Using Deep Learning Models*. 10.1007/978-981-15-5397-4\_63.

[10] MVM, , AP, , BJ, , MK, & , RS (2021). Bitcoin price prediction using machine learning. *International Journal of Engineering Technologies and Management Research*, *8*(5), 20. doi: 10.29121/ijetmr.v8.i5.2021.953

[11] Farris, Frank. (2010). *The Gini Index and Measures of Inequality. American Mathematical Monthly*. 117. 851-864. 10.4169/000298910X523344.

30

[12] Alqassem, Israa & Rahwan, Iyad & Svetinovic, Davor. (2018). *The Anti-Social System Properties: Bitcoin Network Data Analysis.* IEEE Transactions on Systems, Man, and Cybernetics: Systems. PP. 1-11. 10.1109/TSMC.2018.2883678.

[13] Park, Ji-Won & Kim, Chae. (2020). *Getting to feasible income equality.*

[14] [https://www.investopedia.com/terms/l/lorenz-curve.asp\](https://www.investopedia.com/terms/l/lorenz-curve.asp%5C)

[15] <https://doi.org/10.1016/j.cam.2019.112395>

[16] <https://fred.stlouisfed.org/categories/94>

[17] [A, Ghosh, B. Chakrabarti,Physica A, vol-573, pages-125944(2021)]. [S. Manna, S. Biswas, B. Chakrabarti, arxiv: 2111.11230]

[18]<https://towardsdatascience.com/improve-linear-regression-for-time-series-forecasting-e36f3c3e3534>

[19]<https://in.investing.com/crypto/bitcoin>

[20]<https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html>

31