**Bitcoin Price Prediction Using Machine Learning**

**Course Instructor:**

**Dr. Tahseen A. Jilani**

**Group Project :**

**Group Members :**

**Syed Fahad Shooshtary**

**Waseem Abbass**

**Mustafa Abbass Lakho**

**Waqas Khan**

**Class: Final Year**

**Course: Data Mining**

**ABSTRACT:**

In this research word we have use **Decision Tree Regression** and **Linear Regression** algorithm to find Bitcoin price prediction. To built a better understanding of its worth influence and a view of this good invention, we first give a brief overview of Bitcoin again economics. After that, we define the database, including data from stock market indices, sentiment, block chain and Coin market cap.

**KEYWORDS:**

Bitcoin, Crypto Currency, , Block chain, Decision Tree Regression, Linear Regression, Predictor, Future trading, Artificial Intelligence, Intelligence, trading, learning, Decision, Algorithm, Block, chain, problem, solving, project, assignment, important, decentralized, authority, github, repositories, search, Wikipedia, working, work, excel, research, google, yahoo, paper, columns, variables dependent, independent, machine learn and Prediction

**INTRODUCTION:**

Bitcoin is a digital currency that was 1st introduced in January 2009. It is the most power full cryptocurrency in the world, and it is traded on more than 40 exchanges worldwide, accepting more than 30 different currencies. Bitcoin, as a currency, presents a novel potential for price forecasting due to its extreme volatility, which is significantly higher than that of traditional currencies.

With the introduction of Bitcoin ten years ago, the global economy has become more flexible and responsive, albeit in small numbers. Bitcoin was first proposed as a program that addressed the Double Spend problem (Nakamoto & Shah, 2017), a common problem with Digital Cash systems. However, the influence in subsequent years was enormous. The concept of "Bitcoin" underpins everything from Distributed Ledger Technologies (DLT) to Intelligent Agreements to Cryptocurrencies. This was discovered. During a distinct power division, intuitive motive was mixed in. Machine learning, on the other hand, is still employed, and data is taken into consideration nowadays, thanks to a significant rise in hardware efficiency. As a result, we have a tendency to forecast Bitcoin's value, despite the fact that this dynamic exists not only on Bitcoin exchanges but also on financial markets in general.

Bitcoins are produced as a result of the mining process. They can be exchanged for a variety of goods and services. Bitcoin has been chastised for its use in unlawful transactions, the high quantity of electricity (and thus carbon impact) required for mining, price volatility, and exchange theft. It has been described as a speculative bubble by some investors and economists at different times. Others have utilize it as a kind of investment, despite the fact that various regulatory bodies have issued bitcoin investor alerts. Despite domestic and international resistance, El Salvador became the first country to embrace Bitcoin as legal cash in September 2021.

In a white paper issued on October 31, 2008, the term "bitcoin" was defined. It's made up of several elements. It's a combination of the words "bit" and "coin." No standard for uppercase of bitcoin; some sites use bitcoin, capitalized, to refer to the tech and network, while others use bitcoin, lowercase, to refer to the unit of account. The Wall Street Journal is a newspaper published in the United States. The Oxford English Dictionary and the Chronicle of Higher Education In all circumstances, lowercase bitcoin is recommended.

Transactions

A Forth-like programming language is used to define transactions. One or more inputs and one or more outputs make up a transaction. When a user transmits bitcoins, he or she creates an output that lists each address and the amount of bitcoin being sent to that address. Each input must relate to a previously unspent output in the blockchain to avoid duplicate spending.

In a cash transaction, using several inputs is equivalent to using multiple coins. Users can transmit bitcoins to several recipients in one transaction since transactions can have numerous outputs. The sum of inputs (payment coins) can exceed the anticipated sum of payments, just as it can in a cash transaction. An additional output is employed in this situation, which returns the change to the payer. The transaction fee is made up of any input satoshis not included in the transaction outputs.

Despite the fact that transaction costs are optional, miners can pick and choose which transactions to process, prioritizing those that pay larger fees. Miners can choose transactions based on the fee paid in relation to their storage size, rather than the total cost paid.

These charges are usually expressed as satoshis per byte. The number of inputs used to produce the transaction, as well as the number of outputs, determine the transaction's size.

Initially, the blockchain's blocks were limited to 32 megabytes in size. Satoshi Nakamoto introduced the one megabyte block size restriction in 2010. The one megabyte block size limit eventually caused issues with transaction processing, including as increased transaction fees and delayed transaction processing. Lightning Network, according to Andreas Antonopoulos, is a potential scalability solution and a second-layer routing network.

Ownership

Bitcoin addresses are associated with bitcoins on the blockchain.

Picking a random valid private key and computing the accompanying bitcoin address is all it takes to create a bitcoin address. This calculation can be completed in a fraction of a second. Computing the private key of a particular bitcoin address, on the other hand, is essentially impossible. Users can share or make public a bitcoin address without jeopardising the private key associated with it. Furthermore, the number of viable private keys is so large that it's quite unlikely that someone will compute a key pair that's already in use and has money. Due to the large number of valid private keys, brute force cannot be used to compromise a private key. The owner must know the appropriate private key in order to spend their bitcoins. The owner must be aware of the relevant private key and sign the transaction digitally. The public key is used by the network to verify the signature; the private key is never released.

The bitcoin network will not accept any other evidence of ownership if the private key is lost.

The coins are practically lost because they are no longer useful. For example, in 2013, one user claimed to have unintentionally thrown a hard disc with his private key, resulting in the loss of 7,500 bitcoins, valued at $7.5 million at the time. About 20% of all bitcoins are thought to have been lost, with a market worth of almost $20 billion at July 2018 prices.

The private key must be kept secret to preserve bitcoin security. If a third party has access to the private key, such as through a data breach, the third party can use it to steal any bitcoins linked with it. Around 980,000 bitcoins were stolen from cryptocurrency exchanges as of December 2017.

In terms of ownership, 0.5 percent of bitcoin wallets control 87 percent of all bitcoins ever produced as of March 16, 2018.

Mining is a type of data storage that makes advantage of computer computing power. By continuously combining newly broadcast transactions into a block, which is subsequently broadcast to the network and verified by recipient nodes, miners keep the blockchain consistent, complete, and unalterable. The previous block's SHA-256 cryptographic hash is included in each block. As a result, it is linked to the preceding block, and the blockchain is given its name. A new block must include a proof-of-work to be accepted by the rest of the network (PoW). Miners must find a nonce (once-used number) that is numerically smaller than the network's difficulty target when the block content and the nonce are hashed together. This proof is simple to verify for any node in the network, but it takes a long time to generate because, in order to generate a secure cryptographic hash, miners must try many different nonce values (usually in ascending natural numbers: 0, 1, 2, 3,) before finding one that is less difficult than the difficulty target. Block hashes have a lower difficulty target than SHA-256 hashes since the difficulty target is so small. The amount of work required to generate a block can be modified by altering this difficulty target. Nodes deterministically alter the difficulty target based on the recent pace of block generation every 2,016 blocks (about 14 days given roughly 10 minutes per block), with the goal of keeping the average duration between new blocks at ten minutes. As a result, the system adapts to the overall amount of mining power on the network automatically. To generate a block hash smaller than the difficulty target, it takes on average 79 sextillion (79 thousand billion billion) attempts as of September 2021. Large-scale computations are incredibly expensive and require specialised technology.

The proof-of-work system, together with block chaining, makes blockchain alterations extremely difficult. In order for one block's modifications to be accepted, an attacker must modify all following blocks. As new blocks are mined on a regular basis, the challenge of changing a block becomes more difficult as time passes and the number of succeeding blocks (also known as confirmations of the given block) grows.

A Mining pool is frequently used to pool computing resources in order to reduce volatility in miner income. Individual mining rigs are frequently required to wait significant periods of time for a block of transactions to be confirmed and money to be received. Every time a participating server solves a block, all participating miners get compensated. This payment is based on the amount of work that each miner put in to assist find that block.

In order for one block's modifications to be accepted, an attacker must modify all following blocks. As new blocks are mined on a regular basis, the challenge of changing a block becomes more difficult as time passes and the number of succeeding blocks (also known as confirmations of the given block) grows.

A Mining pool is frequently used to pool computing resources in order to reduce volatility in miner income. Individual mining rigs are frequently required to wait significant periods of time for a block of transactions to be confirmed and money to be received. Every time a participating server solves a block, all participating miners get compensated. This payment is based on the amount of work that each miner put in to assist find that block.

Supply :

The successful miner who discovers a new block is granted permission by the rest of the network to keep all transaction fees from transactions included in the block, as well as a pre-determined reward of newly minted bitcoins. This reward is presently 6.25 newly minted bitcoins every block as of May 11, 2020. This transaction is responsible for the creation of all bitcoins in existence. The reward will eventually be rounded down to zero, and the maximum of 21 million bitcoins will be reached.

Fungible: Bitcoin is pseudonymous, which means that funds are attached to bitcoin addresses rather than real-world individuals. The owners of bitcoin addresses aren't revealed, but all transactions on the blockchain are open to the public. Furthermore, "idioms of use" (e.g., transactions that spend coins from many inputs signal that the inputs may have a similar owner) and matching public transaction data with known information on owners of certain addresses can be used to link transactions to individuals and corporations. Furthermore, bitcoin exchanges that sell bitcoins for traditional currencies may be compelled by law to acquire personal data. Each transaction can be assigned a unique bitcoin address to increase financial anonymity.

Wallets and comparable software treat all bitcoins as equals in theory, ensuring a fundamental level of trust. Researchers have pointed out that each bitcoin's history is recorded and publicly accessible on the blockchain ledger, and that some users may refuse to accept bitcoins from problematic transactions, jeopardising the currency's fungibility. Mt. Gox froze, suspended the accounts of anyone who deposited bitcoins that had recently been stolen in 2012.

Rather than being used in merchant transactions, the vast majority of bitcoin transactions take place on a cryptocurrency exchange. Bitcoin use in a retail setting is problematic due to ten-minute delays in processing payments through the blockchain. Prices are rarely quoted in bitcoin units, and many trades need one or two conversions into traditional currencies. Payment service providers may make the conversions for merchants who accept bitcoin payments.

Only three of the top 500 U.S. online stores accepted bitcoin in 2017 and 2018, down from five in 2016. High transaction fees and long transaction times are contributing to this reduction, which is attributable to bitcoin's scaling concerns.

According to Bloomberg, in June 2018, the top 17 crypto merchant-processing providers handled $69 million, down from $411 million in September 2017. According to Nicholas Weaver, a researcher reported by Bloomberg, Bitcoin is "not truly usable" for retail transactions because to high prices and the inability to manage chargebacks. According to economist Kim, high price fluctuation and transaction costs make paying for minor retail items with bitcoin problematic. Bitcoin is still utilised for large-ticket purchases on sites like Overstock.com, as well as cross-border payments to freelancers and other merchants.

Bitcoin Core is a free and open-source programme that acts as a bitcoin node (a group of computers that make up the bitcoin network) and provides a bitcoin wallet that completely verifies payments. It is regarded as the standard implementation of bitcoin. Satoshi Nakamoto first released the software under the name "Bitcoin," but it was later renamed "Bitcoin Core" to separate it from the network. The Satoshi client is another name for it.

The cryptography library libsecp256k1 is also maintained by the project. Furthermore, a cryptocurrency wallet is included by default, which can be used to transfer payments. Bitcoins can be sent and received using this wallet. It does not make it easier to buy or sell bitcoin. It enables users to create QR codes that may be used to receive payment.

This distributed ledger, which has grown to over 235 terabytes in size as of January 2019, must be downloaded or synchronised before the client can fully participate. Although, because it is possible to run in pruning mode, the entire blockchain is not required at once. Bitcoin Core includes bitcoin, a command-line-based daemon with a JSON-RPC interface. It also gives users access to testnet, a global testing environment that mimics the bitcoin main network by using a different blockchain and employing worthless "test bitcoins." Regtest, or Regression Test Mode, produces a private blockchain that can be used for local testing. Finally, bitcoin-cli is supplied, which is a basic software that allows users to send RPC commands to bitcoin.

Hard-coded checkpoints in the client are exclusively used to prevent Denial of Service attacks against nodes that are syncing the chain for the first time. As a result, the checkpoints featured are from a few years ago. Satoshi Nakamoto introduced a one megabyte block size restriction in 2010. The maximum network bandwidth was thus limited to around three transactions per second. Since then, network bandwidth has gradually increased as block sizes have been increased and wallet behaviour has improved. Satoshi Nakamoto included a network alert system as a way of informing users of critical bitcoin news. It was decommissioned in November 2016. It was no longer relevant because bitcoin news is already extensively distributed.

Bitcoin Core features a Forth-inspired scripting language for defining transactions and specifying parameters. Script Pub Key is used to "freeze" transactions based on a set of circumstances in the future. To achieve these conditions or "unlock" a transaction, Script Sign is used. OP Codes are used to conduct operations on the data. There are two stacks in use: main and alt. Looping is not permitted.. It was decommissioned in November 2016. It was no longer relevant because bitcoin news is already extensively distributed.

Bitcoin Core features a Forth-inspired scripting language for defining transactions and specifying parameters. Script Pub Key is used to "freeze" transactions based on a set of circumstances in the future. To achieve these conditions or "unlock" a transaction, Script Sign is used. OP Codes are used to conduct operations on the data. There are two stacks in use: main and alt. Looping is not permitted.

**LITERATURE SURVEY:**

We've all speculated on where bitcoin prices will be in a year, two years, five years, or even ten years. Anticipation is really difficult, but it is something that each of us enjoys doing. Purchasing and selling bitcoins can provide enormous benefits if done correctly. Many individuals have found it to be a fortune in the past, and it is still making them a lot of money today. This, however, is not without its drawbacks. You can lose a lot of money if you don't think about it and calculate it properly. You should have a solid understanding of how and why bitcoin prices fluctuate (organic market, guidelines, news, and so on), which means you should understand how people make bitcoin forecasts. In addition to these factors (supply and demand, laws, news, and so on), one must consider bitcoin's technology and advancement. Apart from that, we now have to deal with the technical aspects, which require the use of numerous algorithms and technologies that can accurately estimate bitcoin prices. Although we came across several models utilizing machine learning and Decision Tree ideas, such as the Decision Tree Regression Model (DTRM), Linear Regression Model (LRM), and so on. Normally, a time series is a set of numbers that progresses over time. This is because the whole data set should be split into two parts: inputs and outputs, as this is a time series data collection. Because it can readily tackle predicting difficulties with various inputs.

The DTRM will use historical data to forecast bitcoin prices 10 days ahead of its closing price in the approach we're taking. We use a Decision Tree Regressor Model (DTRM) and a Linear Regression Model in our method (LRM).

**Technologies Used:**

* **Decision Tree Regression:**

Decision Tree Analysis is a general-purpose predictive modelling tool with applications in a variety of fields. In general, decision trees are built using an algorithm that determines multiple ways to segment a data set based on certain conditions. It is one of the most popular and practical supervised learning algorithms. Decision Trees are a supervised non-parametric learning method that may be utilized for both classification and regression applications. The goal is to learn simple decision rules from data attributes to develop a model that predicts the value of a target variable.

**Decision Tree Algorithm**

In the shape of a tree structure, a decision tree constructs regression or classification models. It incrementally cuts down a dataset into smaller and smaller sections while also developing an associated decision tree. A tree with decision nodes and leaf nodes is the end result. A decision node (for example, Outlook) can have two or more branches (for example, Sunny, Overcast, and Rainy), each of which represents a value for the attribute being checked. A decision on the numerical aim is represented as a leaf node (e.g., Hours Played). The root node is the topmost decision node in a tree that corresponds to the best predictor. Both category and numerical data can be handled by decision trees.

It is considered to be a component of artificial intelligence. Machine learning algorithms create a model based on training data to make predictions or judgments without having to be explicitly programmed to do so. Machine learning algorithms are utilized in a wide range of applications, including medicine, email filtering, speech recognition, and computer vision, where developing traditional algorithms to do the required tasks is difficult or impossible.

However, not all machine learning is statistical learning. A subset of machine learning is strongly related to computational statistics, which focuses on making predictions using computers. The discipline of machine learning benefits from the study of mathematical optimization since it provides tools, theory, and application domains. Data mining is a similar branch of research that focuses on unsupervised learning for exploratory data analysis. Data and neural networks are used in some machine learning implementations to replicate the functioning of a biological brain.

Learning algorithms are based on the assumption that methods, algorithms, and conclusions that have worked in the past would likely continue to work in the future. These inferences can be clear, such as "because the sun has been rising every morning for the last 10,000 days, it will most likely rise again tomorrow morning." "X percent of families include geographically different species with colour variants, thus there's a Y percent possibility that undiscovered black swans exist," for example.

Without being expressly designed, machine learning programmes may complete tasks. It entails computers learning from data in order to do specific jobs. It is possible to write algorithms that inform the machine how to perform all steps required to solve the problem at hand for basic jobs entrusted to computers; no learning is required on the computer's behalf. It can be difficult for a human to manually build the algorithms required for increasingly complicated tasks. In practise, assisting the computer in developing its own algorithm rather than having human programmers explain each required step can prove to be more productive.

Machine learning is a discipline that uses a variety of ways to train computers how to complete tasks for which no entirely suitable solution exists. When there are a large number of possible replies, one strategy is to classify some of the correct answers as valid. The computer can then utilise this as training data to refine the algorithm(s) it uses to determine right answers. The MNIST dataset of handwritten digits, for example, has frequently been used to train a system for the task of digital character recognition.

**Data mining:**

While machine learning and data mining both use similar methodologies and have significant overlap, machine learning focuses on prediction based on known qualities learnt from training data, whereas data mining focuses on the finding of (previously) unknown properties in the data. Machine learning uses data mining methods as "unsupervised learning" or as a preprocessing step to increase learner accuracy; data mining, on the other hand, uses data mining methods as "unsupervised learning" or as a preprocessing step to improve learner accuracy.

The underlying assumptions they operate with cause a lot of the confusion between these two academic communities (which frequently have distinct conferences and journals, with ECML PKDD being a notable exception). The ability to reproduce known knowledge is commonly measured in machine learning, whereas the fundamental aim in knowledge discovery and data mining (KDD) is the finding of previously undiscovered knowledge. When compared to existing knowledge, an uninformed (unsupervised) technique will easily beat other supervised methods, however supervised methods cannot be used in a typical KDD task due to the lack of training data.

**Learning under supervision**

Supervised learning algorithms create a mathematical model of a set of data that includes both the inputs and the outputs that are sought. The information is referred to as training data, and it consists of a collection of training instances. Each training example has one or more inputs and a supervisory signal as the desired output. Each training sample is represented by an array or vector, sometimes referred to as a feature vector, and the training data is represented by a matrix in the mathematical model. Supervised learning techniques develop a function that may be used to predict the output associated with fresh inputs by iteratively optimizing an objective function. The algorithm will be able to accurately estimate the output for inputs that were not part of the training data if it uses an optimum function. An algorithm that learns to complete a task is one that increases the accuracy of its outputs or predictions over time.

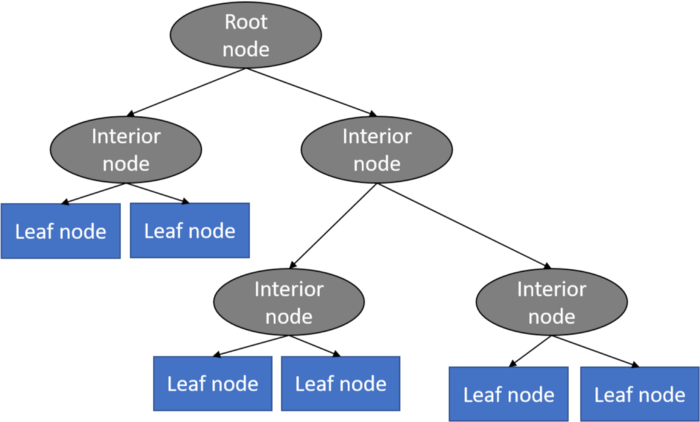
Active learning, classification, and regression are examples of supervised learning algorithms. When the outputs are limited to a small set of values, classification techniques are employed, and regression algorithms are used when the outputs can have any numerical value within a range. An incoming email, for example, would be the input to a classification algorithm that filters emails, and the output would be the name of the folder to file the email in. The purpose of similarity learning, which is closely connected to regression and classification, is to learn from examples using a similarity function that quantifies how similar or related two items are. Ranking, recommendation systems, visual identification tracking, face verification, and speaker verification are some of the uses.

**Unsupervised learning algorithms**

Unsupervised learning algorithms use a collection of data with only inputs and look for structure in the data, such as grouping or clustering of data points. As a result, the algorithms learn from unlabeled, unclassified, and uncategorized test data. Unsupervised learning algorithms discover commonalities in the data and react depending on the existence or lack of such commonalities in each new piece of data, rather than responding to feedback.

The field of density estimation in statistics, such as calculating the probability density function, is a key application of unsupervised learning. Unsupervised learning, on the other hand, comprises various domains that need summarizing and explaining data aspects.

Cluster analysis divides a set of observations into subsets (called clusters) so that observations within the same cluster are comparable based on one or more predetermined criteria, while observations from different clusters are distinct. Different clustering approaches make different assumptions about the structure of the data, which is commonly characterized by some similarity metric and evaluated, for example, by internal compactness, or the similarity between cluster members, and separation, or the difference between clusters. Estimated density and graph connectedness are used in other approaches.



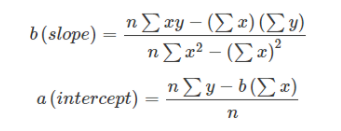
* **Linear Regression:**

Mathematically, we can write a linear regression equation :

we can write a linear regression equation as:

**y = a + bx**

Where a and b given by the formulas:



Here x and y are two variables on regression line.

b = Slope of the line

a = y-intercept of the line

y = Dependent variable from dataset

**Machine Learning Pipeline:**

**Libraries used:**

In the machine learning backend program, **Pandas** is mainly used for-related activities. **Numpy** is used for matrix/vector performance and for keeping data and training sets, **Scikit-learn** (also known as sklearn) is used to make min-max standardization. Finally, **matplotlib** and **sea born** is used to display charts.

**Data Range:**

**01/01/2020 to 12/31/2020**

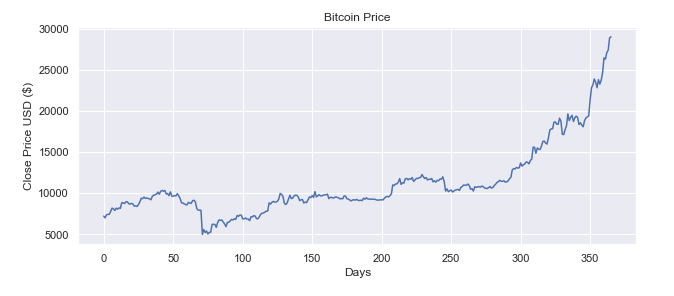
**Data Description:**

The data is collected from **Yahoo Finance**

<https://finance.yahoo.com/quote/BTC-USD?p=BTC-USD&ncid=stockrec>

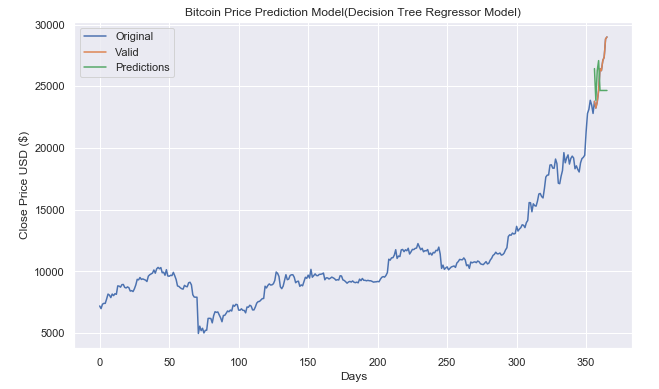
It consists of various attributes that are taken into predicting Bitcoin closing price.

****

****

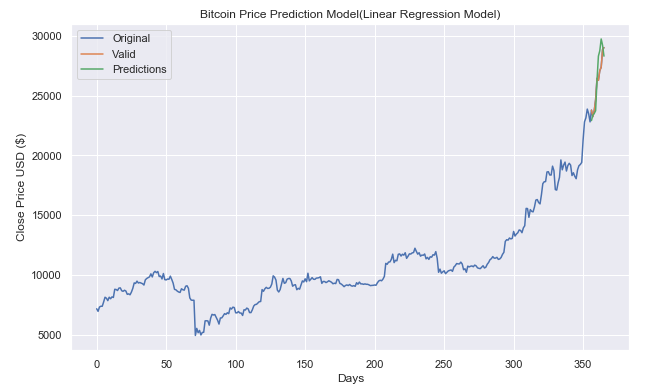
**Predict the price of Bitcoin using DTRM (Decision Tree Regressor Model):**



****

**Predict the price of Bitcoin using LRM (Linear Regression Model):**

****

****

**Conclusion:**

In this project report Decision Tree Regression model gave the 89.37% accuracy rate while the Linear Regression Model gave 93.67% accuracy rate.

Also train another dataset gave the same accuracy rate which will in the form of array which further converted into new column of our dataset.

It is inclusive of all the parameters needed to evaluate the prediction of Bitcoin. The model is trained to produce results with satisfactory accuracy, after which it produces accurate results as to whether a trader should trade or not without any manual work.

**References:**

<https://www.granthaalayahpublication.org/ijetmr-ojms/index.php/ijetmr/article/view/IJETMR21_A05_2586/771>

<https://www.hackerearth.com/practice/machine-learning/machine-learning-algorithms/ml-decision-tree/tutorial/>

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

<https://towardsdatascience.com/machine-learning-basics-decision-tree-regression-1d73ea003fda>

<https://www.geeksforgeeks.org/ml-linear-regression/#:~:text=Linear+Regression+is+a+machine,relationship+between+variables+and+forecasting>.

<https://www.springboard.com/blog/ai-machine-learning/what-is-linear-regression/>

<https://finance.yahoo.com/quote/BTC-USD?p=BTC-USD&ncid=stockrec>