**Bitcoin Price Prediction Using Machine Learning**

**Course Instructor:**

**Dr. Tahseen A. Jilani**

**Group Project :**

**Group Members :**

**Syed Fahad Shooshtary**

**Waseem Abbass**

**Mustafa Abbass Lakho**

**Nasir Ali Khan**

**Waqas Khan**

**Class: Final Year**

**Course: Data Mining**

**ABSTRACT:**

In this research paper we use **Decision Tree Regression** and **Linear Regression** algorithm to find Bitcoin price prediction. To develop a better understanding of its price influence and a common 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 Coinmarketcap.

**KEYWORDS:**

Bitcoin, Crypto Currency, Machine Learning, Blockchain, Decision Tree Regression, Linear Regression, Prediction

**INTRODUCTION:**

Bitcoin is a cryptocurrency that was created in January 2009. It is the world’s most valuable cryptocurrency and is traded on over 40 exchanges around the world, accepting over 30 different currencies. As a currency, Bitcoin offers a new opportunity for price forecasting as it has high volatility, which is much higher compared to traditional currencies.

With the appearance of Bitcoin 10 years ago the globe economist, albeit in small numbers, is flexible and responsive. Bitcoin introduced itself as a program that solved the Double Spend problem (Nakamoto & Shah, 2017), a preferred issue with Digital Cash systems. However, the impact in the coming years was great. Distributed Ledger Technologies (DLT), Intelligent Agreements, Cryptocurrencies, etc. it's all supported by the thought of "Bitcoin". This was identified, during a separate power division mixed with intuitive motive. On the opposite side of the spectrum, and data is taken into account nowadays, over time with a major increase in hardware efficiency, Machine learning continues to be used. As a result, we tend to predict the worth of Bitcoin, while the dynamic isn't not only on Bitcoin exchanges but also on finance markets generally.

**LITERATURE SURVEY:**

We have all considered where bitcoin costs will be one year, two years, five years or even 10 years from now. It's really difficult to anticipate however each and every one of us loves to do it. Tremendous measures of benefits can be made by purchasing and selling bitcoins, whenever done accurately.. It has been proven to be a fortune for many people in the past and is still making them a lot of money today. But this doesn’t come without its downside too. If not thought of and calculated properly, you can lose a lot of money too. You should have an incredible comprehension of how and precisely why bitcoin costs change (organic market, guidelines, news, and so forth), which implies you should realize how individuals make their bitcoin predictions. Considering these things (supply and demand, regulations, news, etc.), one must also think about the technology of bitcoin and its progress. This aside, we now have to deal with the technical parts using various algorithms and technologies which can predict precise bitcoin prices. Although we came across various models which are currently present like Decision Tree Regressor Model (DTRM), Linear Regression Model (LRM), etc. with machine learning and Decision Tree concepts. Normally a time series is a sequence of numbers along time. This is due to the fact that this being a time series data set, the overall data sets should be split into two parts: inputs and outputs. Since it can very easily handle multiple input forecasting problems.

In the approach which we are following, the DTRM will use the previous data to predict bitcoin prices 10 days ahead of its closing price.  In the approach used by us, we implement Decision Tree Regressor Model (DTRM) and a Linear Regression Model (LRM).

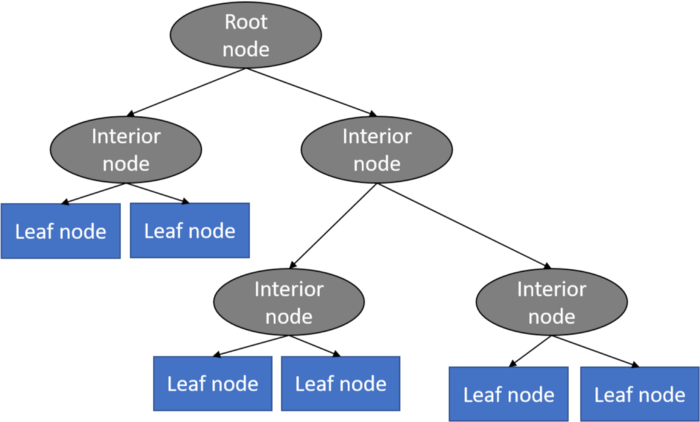
**Technologies Used:**

* **Decision Tree Regression:**

Decision Tree Analysis is a general, predictive modelling tool that has applications spanning a number of different areas. In general, decision trees are constructed via an algorithmic approach that identifies ways to split a data set based on different conditions. It is one of the most widely used and practical methods for supervised learning. Decision Trees are a non-parametric supervised learning method used for both classification and regression tasks. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

**Algorithm:**

It is a tree-structured classifier with three types of nodes. The Root Nodeis the initial node which represents the entire sample and may get split further into further nodes. The InteriorNodesrepresent the features of a data set and the branches represent the decision rules. Finally, the Leaf Nodes represent the outcome. This algorithm is very useful for solving decision-related problems. This algorithm is very useful for solving decision-related problems.



* **Linear Regression:**

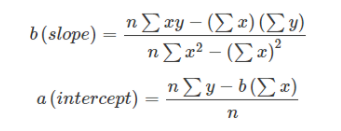
**Linear Regression** is a machine learning algorithm based on **supervised learning**. It performs a **regression task**. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables, they are considering and the number of independent variables being used. Linear regression is used to predict a quantitative response Y from the predictor variable X.

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**

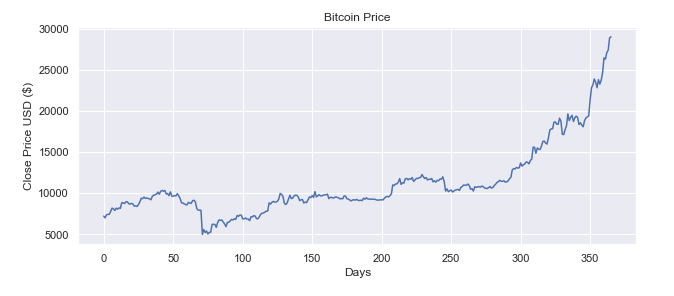
**Data Description:**

The data is collected from **Yahoo Finance**

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

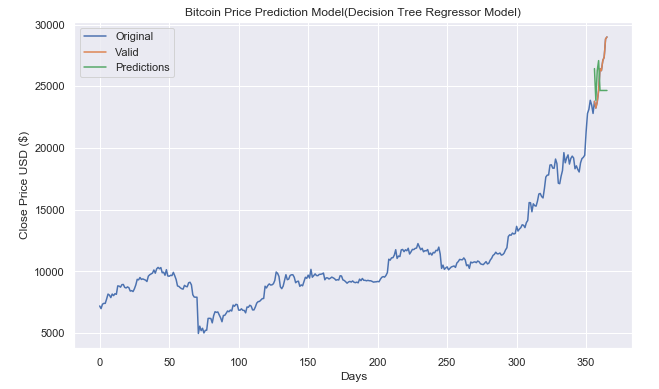
It consists of various attributes that are taken into predicting Bitcoin closing price.

****

****

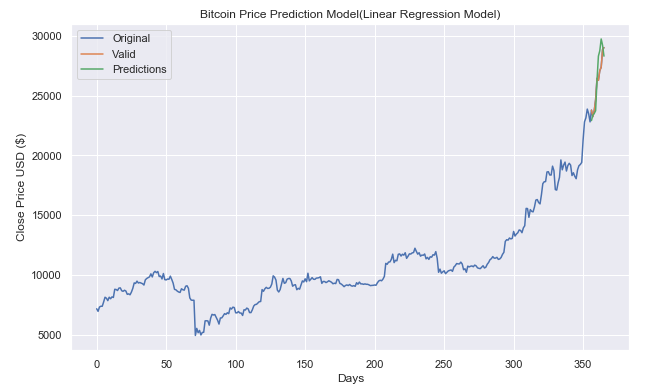
**Predict the price of Bitcoin using DTRM (Decision Tree Regressor Model):**



****

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

****

****

**Conclusion:**

In this project 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>