

---

## Final Project for Machine Learning Course

Team Members :

1. Chey Sreylin : [cheysreylin20@kit.edu.kh](mailto:cheysreylin20@kit.edu.kh)
2. Oun Vikreth : [unvikreth@gmail.com](mailto:unvikreth@gmail.com)
3. Chhoun Sophanon : [chhounsophanon14@gmail.com](mailto:chhounsophanon14@gmail.com)
4. Sari PuthiPonareay : [ponareay171819.ss@gmail.com](mailto:ponareay171819.ss@gmail.com)

# Bitcoin Prediction using Machine Learning Algorithms

29<sup>th</sup> August 2022

## OVERVIEW

Bitcoin is the best-known cryptocurrency. It is a decentralized digital currency that you can buy, sell and exchange directly, without an intermediary like a bank. As more and more people become interested in bitcoin, they started trading and buying during the stock market hours. However, trading crypto is not easy at all, sometimes the price is down and sometimes it's up which can cause more risks to traders losing their money and other properties. That's why before you start trading, you have to clearly research and know well about it. Some people even have their own different techniques for trading and evaluating the stock price market. In that, machine learning is one of the best and extremely powerful techniques in optimizing the decision making process of humans before they start investing.

## PROJECT OBJECTIVE

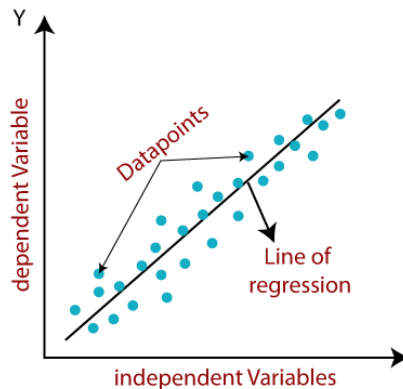
1. Using and testing different algorithms in machine learning to predict the stock price market of bitcoin.
2. Evaluate on the different models to find out which model provides the most accuracy results.

## ALGORITHMS CHOSEN

To predict and find out the most accurate result, here we will use different models in machine learning algorithms : Linear Regression, Support Vector Regression, Decision Tree, Random Forest, and Multilayer Perceptron algorithms.

## 1. Linear Regression Algorithm

Linear Regression is the supervised Machine Learning model in which the model finds the best fit linear line between the independent and dependent variable.



Hypothesis function of Linear Regression

$$y = \theta_1 + \theta_2 \cdot x$$

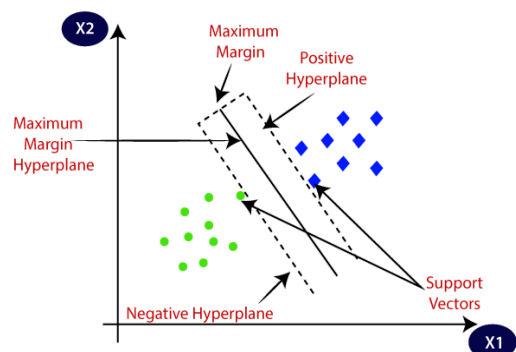
[Image source](#)

## 2. Support Vector Regression Algorithm

Support Vector Regression is a supervised learning algorithm that is used to predict discrete values. Support Vector Regression uses the same principle as the SVMs. The basic idea behind SVR is to find the best fit line. In SVR, the best fit line is the hyperplane that has the maximum number of points.

[Image source](#)

- Kernel: A kernel helps us find a hyperplane in the higher dimensional space without increasing the computational cost.
- Hyperplane: There can be multiple lines/decision boundaries to segregate the classes in n-dimensional space, but we need to find out the best decision boundary that helps to classify the data points.
- Support Vectors: The data points or vectors that are the closest to the hyperplane and which affect the position of the hyperplane are termed as Support Vector.



### 3. Decision Tree Regression Algorithm

Decision tree regression is a supervised learning algorithm that is used to observe features of an object and trains a model in the structure of a tree to predict data in the future to produce meaningful continuous output.

- Information Gain: Information gain is the measurement of changes in entropy after the segmentation of a dataset based on an attribute.

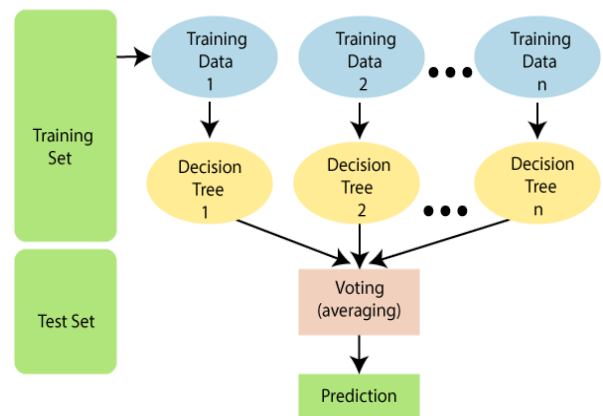
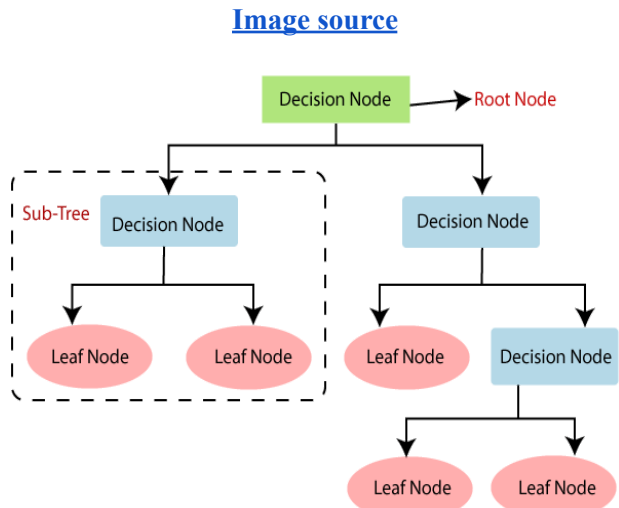
$$IG(S, A) = H(S) - \sum_{i=0}^n P(x) * H(x)$$

- Entropy : Entropy is a metric to measure the impurity in a given attribute. It specifies randomness in data. Entropy can be calculated as:

$$H(S) = \sum_{x \in X} p(x) \log_2 \frac{1}{p(x)}$$

### 4. Random Forest Regression Algorithm

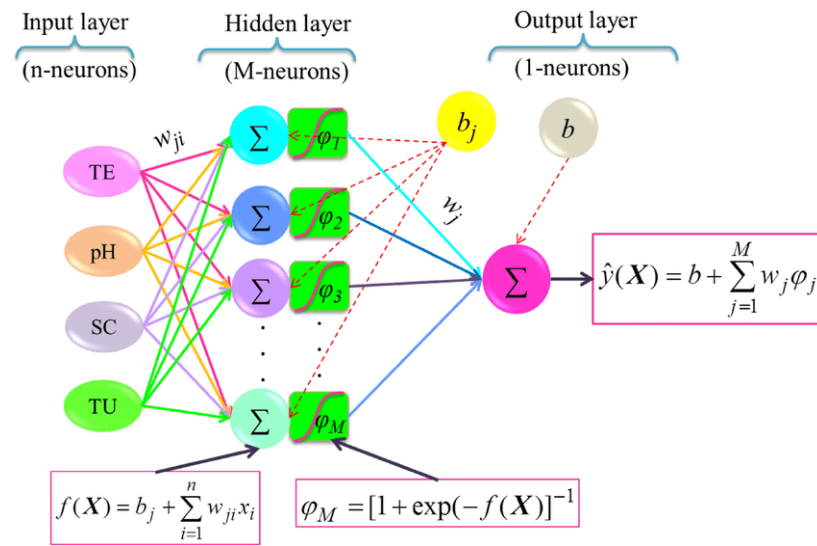
Random Forest Regression algorithms are a class of Machine Learning algorithms that use the combination of multiple random decision trees each trained on a subset of data. The use of multiple trees gives stability to the algorithm and reduces variance. The random forest regression algorithm is a commonly used model due to its ability to work well for large and most kinds of data.



[Image source](#)

## 5. Multilayer Perceptron Algorithm

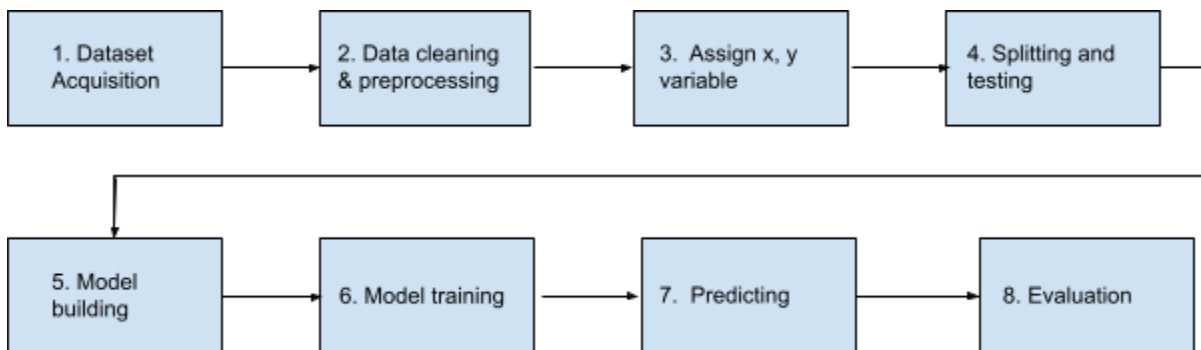
A multilayer perceptron (MLP) is a feedforward artificial neural network that generates a set of outputs from a set of inputs. An MLP is characterized by several layers of input nodes connected as a directed graph between the input and output layers. MLP uses backpropagation for training the network.



[Image source](#)

## PROCESS OF DATA PREDICTION

- All algorithms that we have chosen can work in the same flows of prediction. The Working process can be explained in the below steps and diagram :



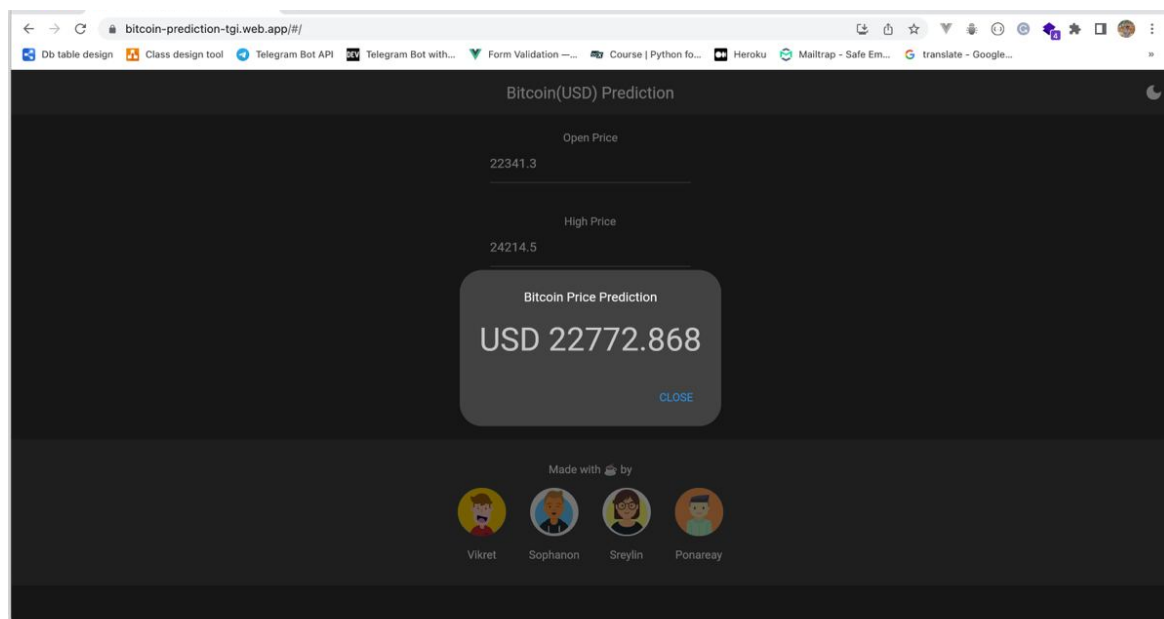
## EXPERIMENTAL AND EVALUATION ON THE RESULT

To predict and find out the most accurate result, here we will use different models in machine learning algorithms : Linear Regression, Support Vector Regression, Decision Tree, Random Forest, and Multilayer Perceptron algorithms.

ALGORITHMS	ACCURACY SCORE	R2 SCORE
Linear Regression	0.9992	0.9992
Support Vector Regression	0.9377	0.9377
Decision tree	0.9975	0.9975
RandomForest	0.9985	0.9985
Multilayers Perceptron	0.9983	0.9983

## DEPLOYMENT

- Here is the final product after the deployment



## CONCLUSION

Based on the evaluation, all the algorithms that we used, provide very good accuracy results. Every model used can be applied in different cases and adjust their performance into different environments and dataset that we use. So that, we can say machine learning plays an important role in the stock price market, it helps you to discover more of the future movement value of stock price and financial trading. The more accurate result of prediction will help all those investors make more profit.

## REFERENCES

1. <https://www.javatpoint.com/linear-regression-in-machine-learning>
2. <https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/>
3. <https://www.javatpoint.com/machine-learning-random-forest-algorithm>
4. <https://medium.com/analytics-vidhya/regression-trees-decision-tree-for-regression-machine-learning-e4d7525d8047>
5. <https://www.xoriant.com/blog/decision-trees-for-classification-a-machine-learning-algorithm#:~:text=Introduction%20Decision%20Trees%20are%20a,namely%20decision%20nodes%20and%20leaves>