BitCoin Price Prediction

Final Project Presentation

Presented By: Chey Sreylin, Vekret, Ponareay, Sophanon

Presentation Agenda

1. Introduction

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 we came up with some algorithms in machine learning to predict the future price.

2. Machine Learning Algorithms Chosen

1. Linear Regression
2. Random forest
3. Support Vector Regression
4. Decision Tree
5. Multi layers perceptrons

1. Linear Regression

Linear regression is a machine learning algorithms based on supervised learning. It's use with numerical data to predict the dependent variable(y) based on the given independent variable(x) and produce the output as continue value.

Hypothesis Function for Linear Regression

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

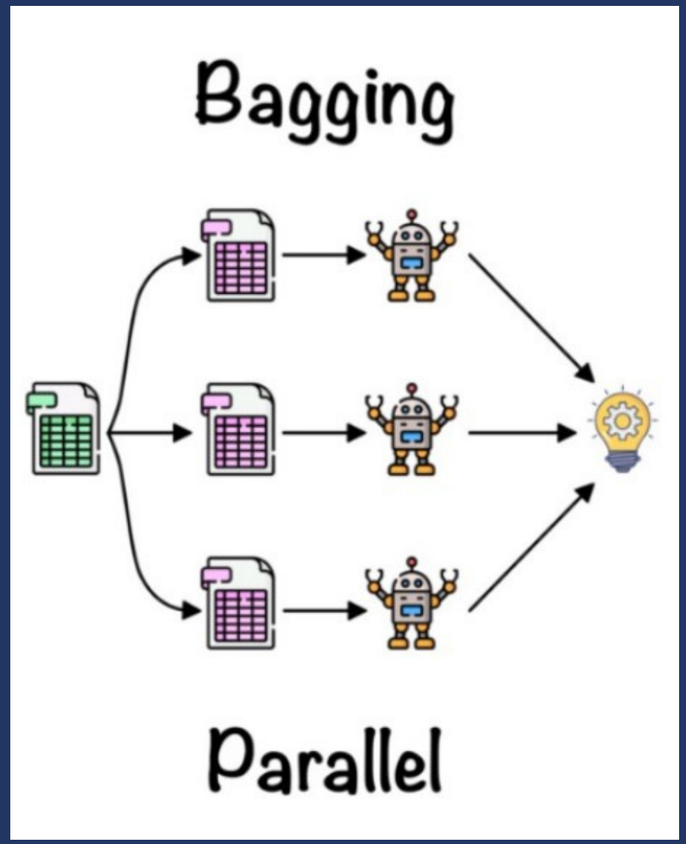
y = dependent variable (output)
x = independent variable (input)
Q1 = Intercept
Q2 = Slopt (Coefficient)

2. Random Forest Algorithms

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression.

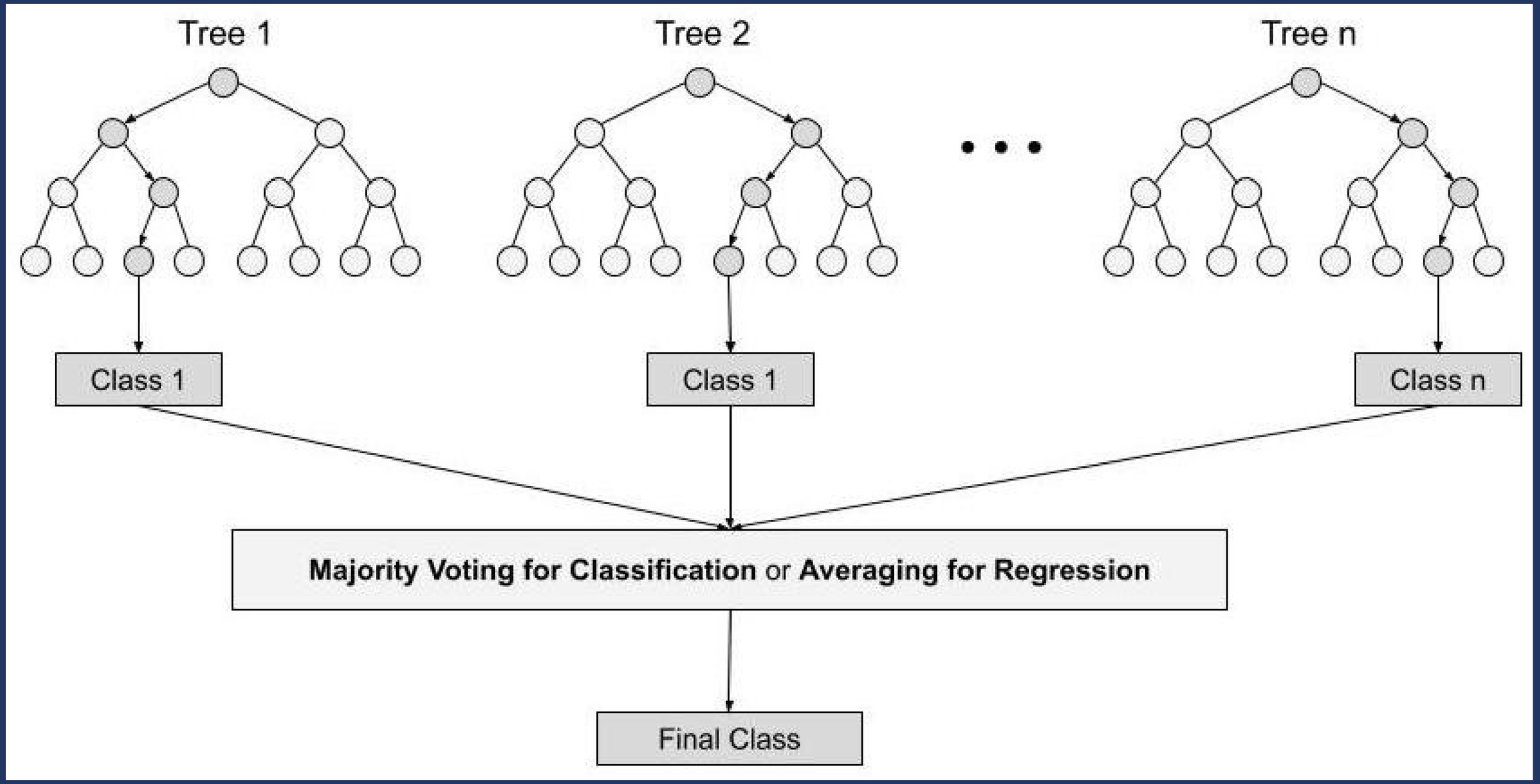
2.1. How Random Forest Work?

- **Random forest** works on the Bagging principle. Now let's dive in and understand bagging in detail.
- **Bagging**: It creates a different training subset from sample training data with replacement & the final output is based on majority voting.



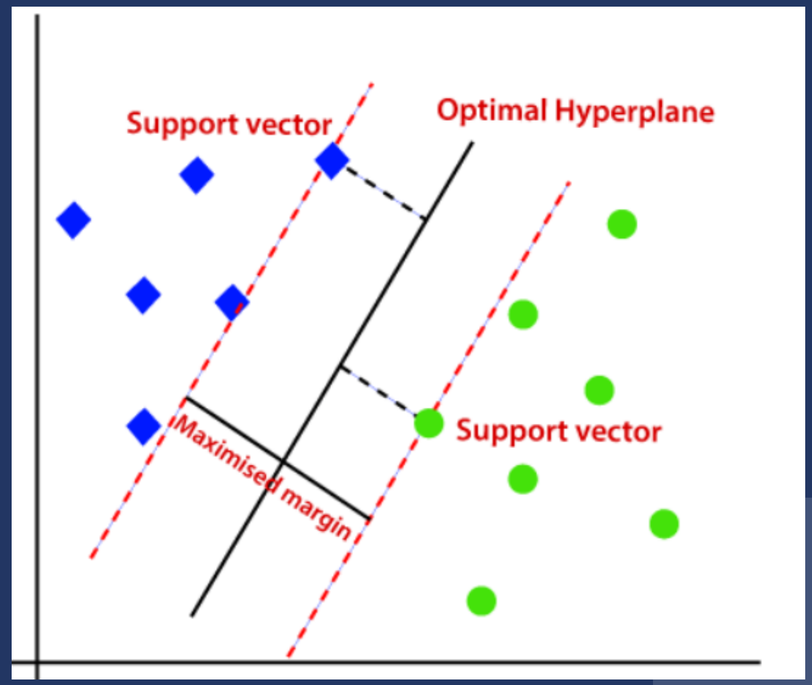
Steps involved in random forest algorithm:

- Step 1: In Random forest n number of random records are taken from the data set having k number of records.
- Step 2: Individual decision trees are constructed for each sample.
- Step 3: Each decision tree will generate an output.
- Step 4: Final output is considered based on Majority Voting or Averaging for Classification and regression respectively.


Tree 1
Tree 2
Tree n
Class 1
Class 1
Class n
Majority Voting for Classification or Averaging for Regression
Final Class

3. Support Vector Regression

Support Vector Regression is a component of Support Vector Machines which is a types of supervised learning that use to predict continues value, unlike SVC for discrete value.



Kernel

Kernel is the function for converting a lower dimensional data set to a higher dimensional dataset .It's help to search for a hyperplane in higher dimensional space while reducing the computing cost.

- Types of kernel used in SVR

1. Linear

$$y = wx + b$$

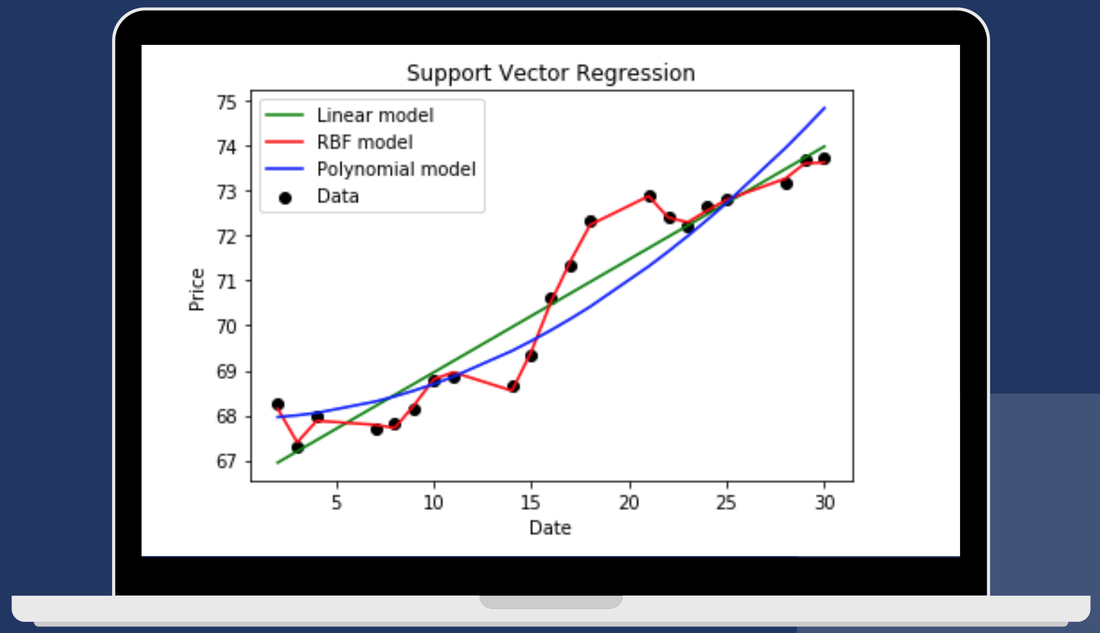
2. Polynomial

$$f(X1, X2) = (X1^T . X2 + 1)^d$$

3. Radial Basis Function (RBF)

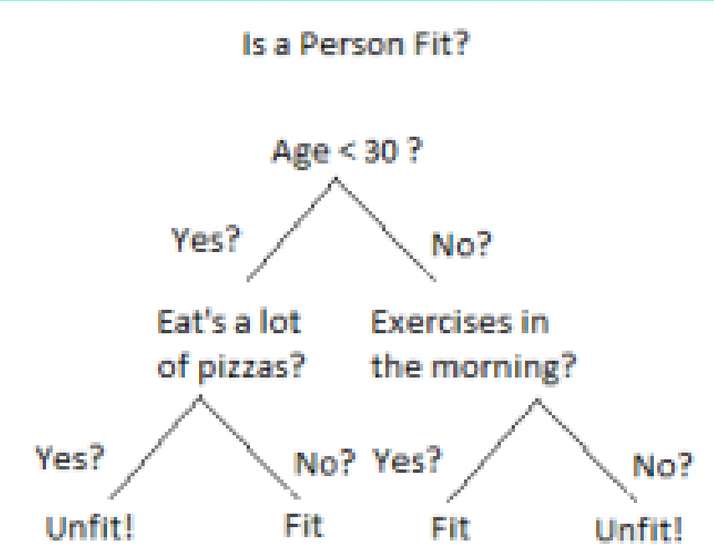
Gaussian Radial Basis function

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$$



4. Decision Tree Algorithms

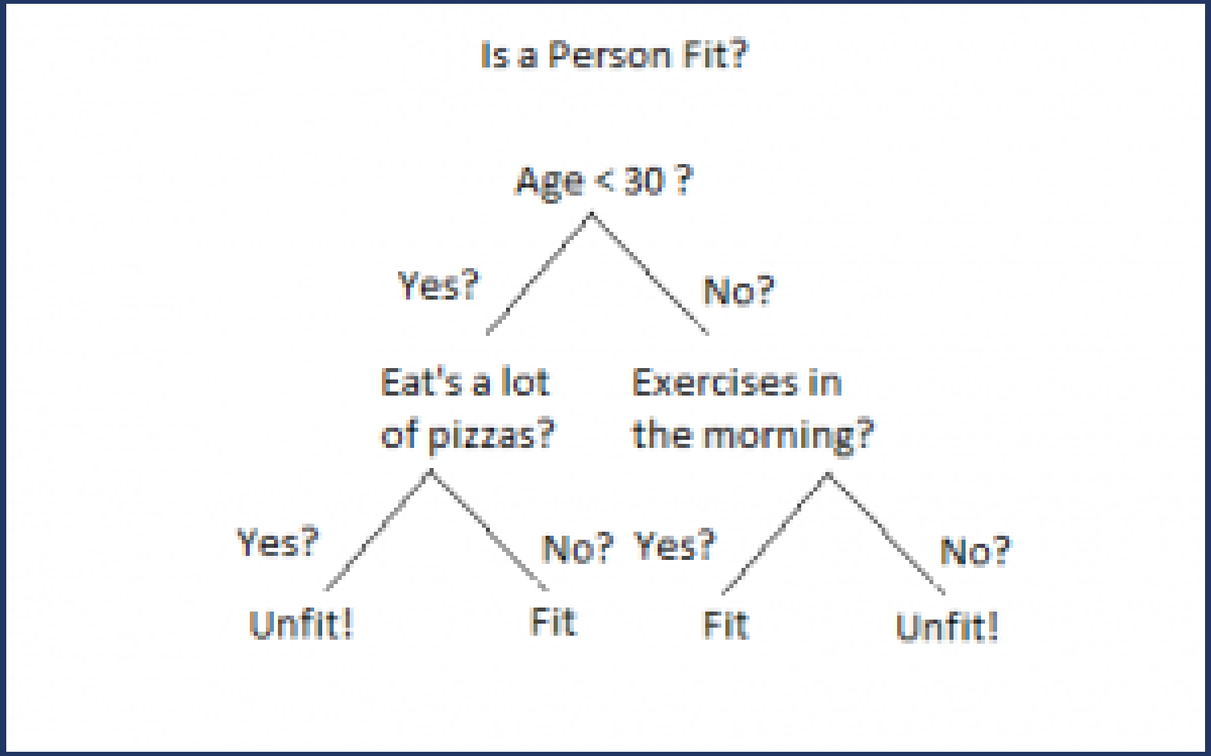
Decision Trees are a type of Supervised Machine Learning (that is you explain what the input is and what the corresponding output is in the training data) where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves.



1. **Classification Trees**
2. **Regression Trees**

1. **Classification Trees: (Yes/No types)**

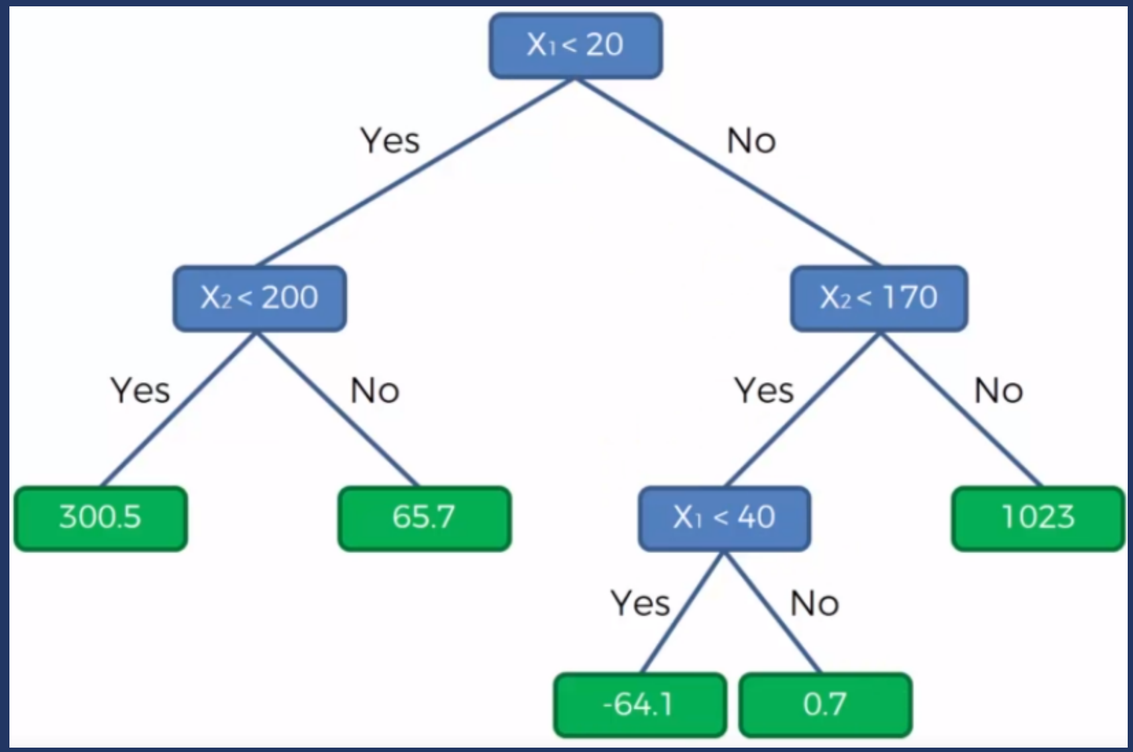
Classification Tree is a type of machine learning algorithm used for classifying remotely sensed and ancillary data in support of land cover mapping and analysis



2. **Regression Tress: (Continuous Data types)**

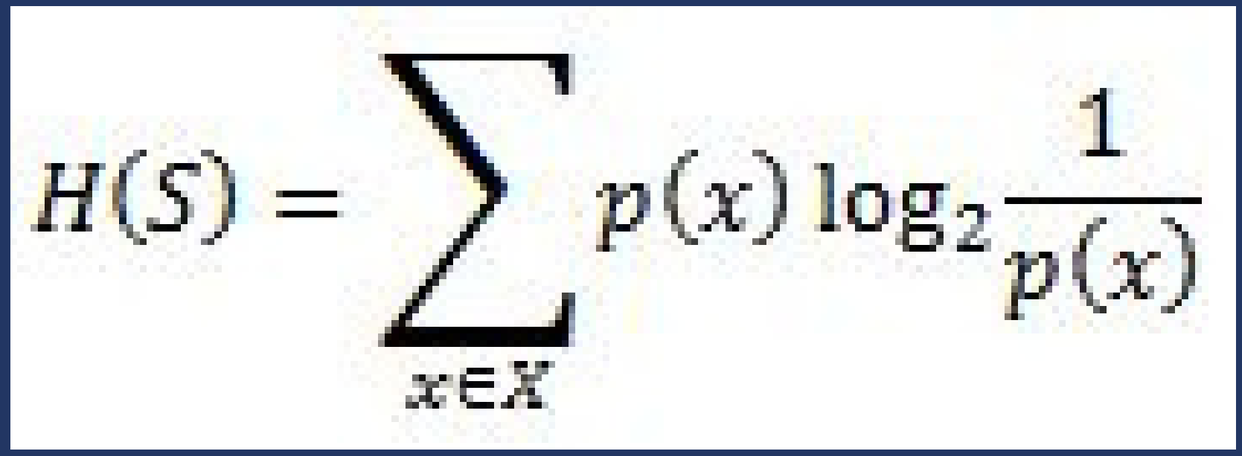
A regression tree is basically a decision tree that is used for the task of regression which can be used to predict continuous valued outputs instead of discrete outputs.

- **Entrophy**
- **Information Gain**



2. **Regression Tress:** **(Continuous Data types)**

Entrophy : Entropy, also called as Shannon Entropy is denoted by H(S) for a finite set S, is the measure of the amount of uncertainty or randomness in data.



2. **Regression Tress:** **(Continuous Data types)**

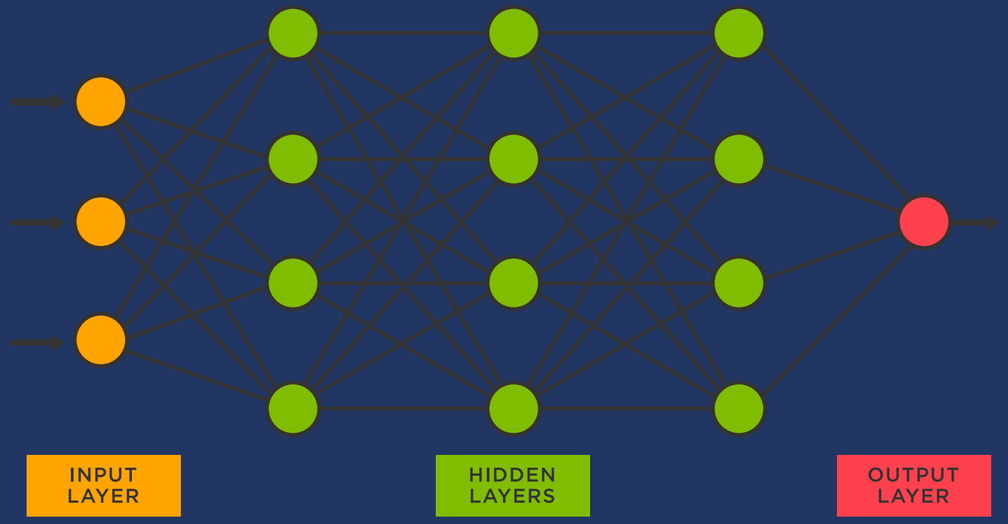
- **Information Gain :** Information gain is also called as Kullback-Leibler divergence denoted by IG(S,A) for a set S is the effective change in entropy after deciding on a particular attribute A. It measures the relative change in entropy with respect to the independent variables.

$$IG(S,A) = H(S) - H(S,A)$$

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

5. Multi layers perceptron

Multi Layers perceptron or MLP is a fully connected class of **feedforward artificial neural network** (ANN). MLP consists of three layers of nodes: an input layer, a hidden layer, and an output layer. Each node is a neuron that uses a nonlinear activation function except for an input layer.



Evaluation

ALGORITHM	ACCURACY	MAE	R2
MLP Regressor	0.9983	431.95	0.9983
Linear Regression	0.9992	291.63	0.9992
SVR	0.9377	3631.73	0.9377
Random Forest	0.9985	411.85	0.9985
Decision Tree	0.9975	568.86	0.9975

Technologies & Implementation

System Architecture

Web App

API

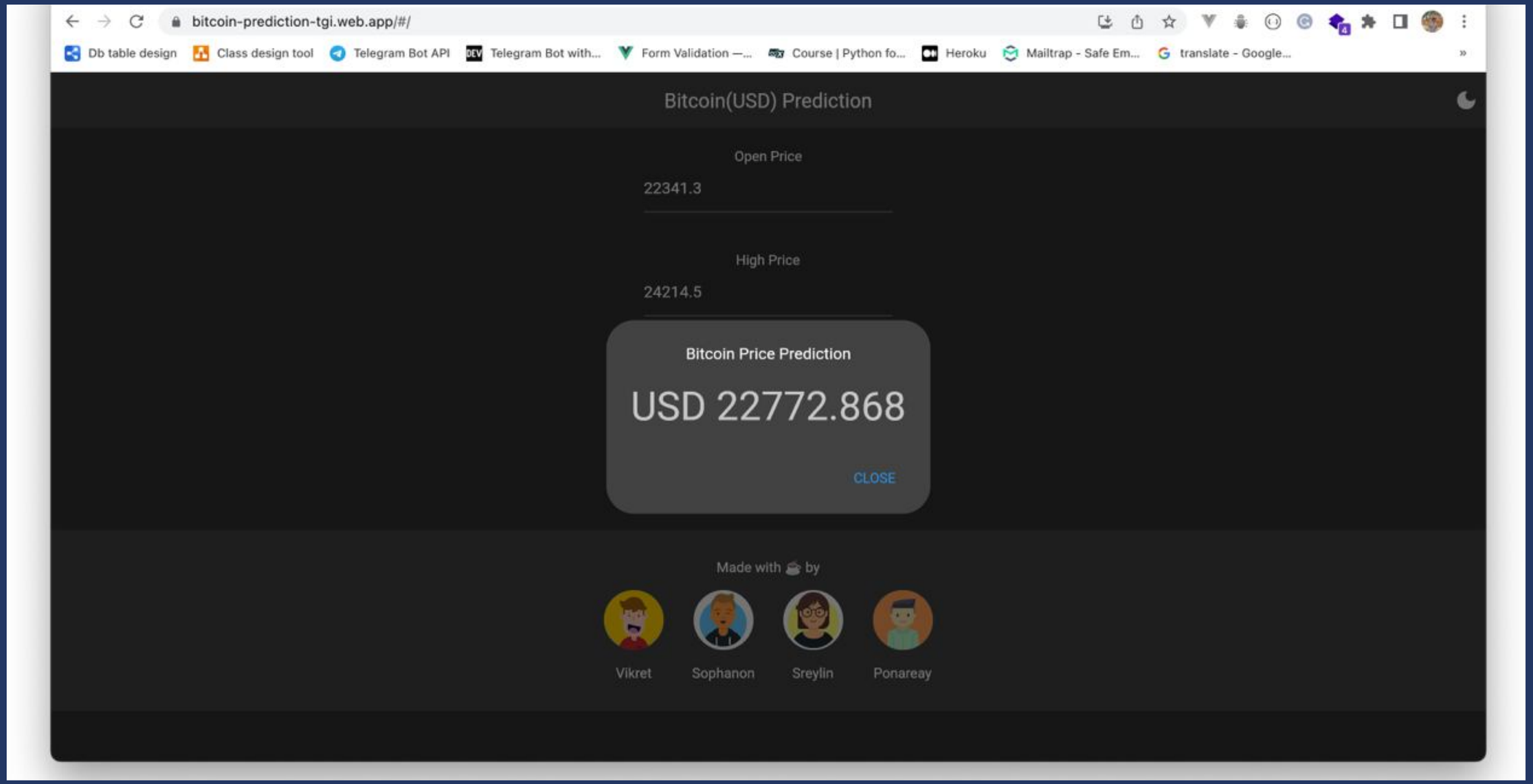
Server

Technologies

Flutter

Tool

Deployment


bitcoin-prediction-tgi.web.app/#/
Db table design
Class design tool
Telegram Bot API
Telegram Bot with...
Form Validation —...
Course | Python fo...
Heroku
Mailtrap - Safe Em...
translate - Google...
Bitcoin(USD) Prediction
Open Price
22341.3
High Price
24214.5
Bitcoin Price Prediction
USD 22772.868
CLOSE
Made with by
Vikret
Sophanon
Sreylin
Ponareay

Applications

Stock Price prediction

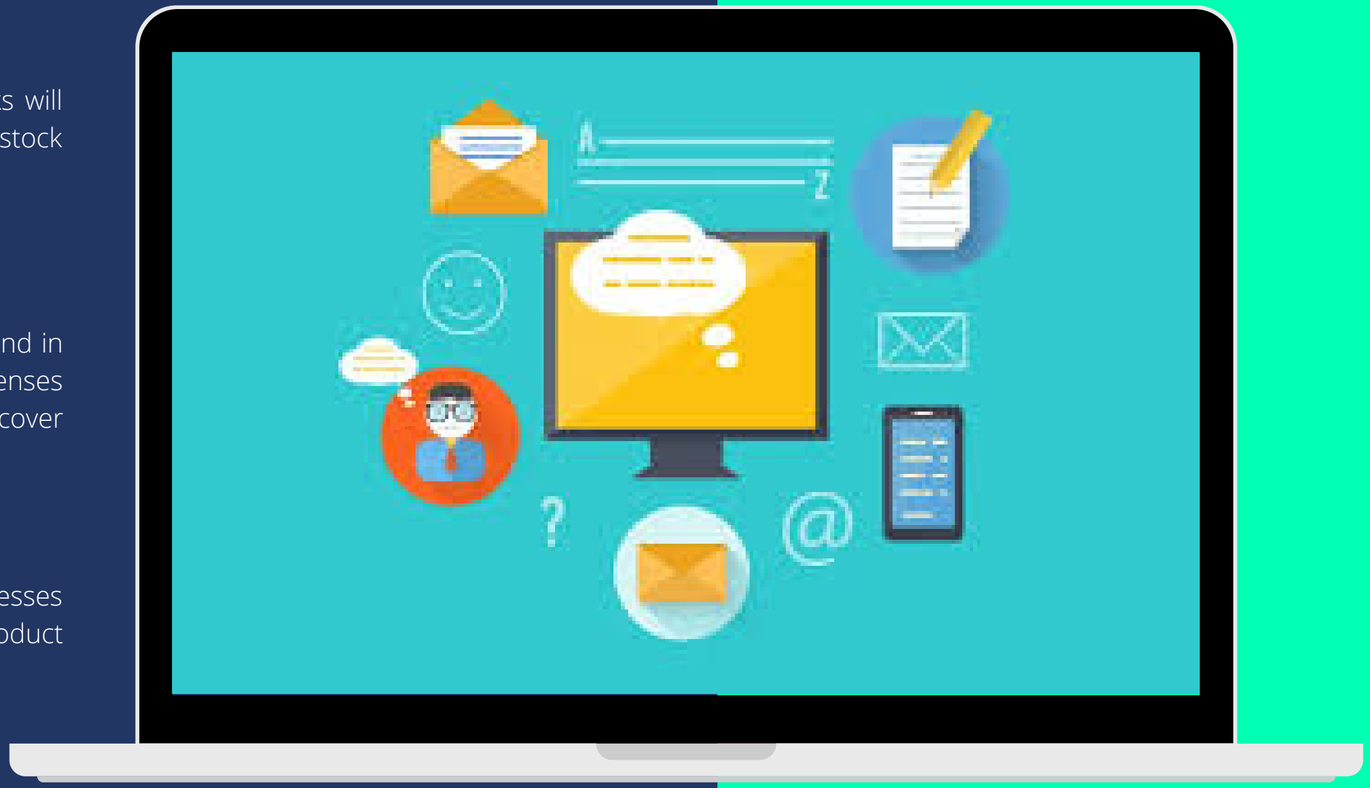
A lot of businesses use these models to predict how stocks will perform in the future. This is done by analyzing past data on stock prices and trends to identify patterns.

Cash forecasting

Predicting forecast to see how much cash they'll have on hand in the future. This is important for things like managing expenses and ensuring that there is enough cash on hand to cover unexpected costs.

Analysing survey data

Also can be used to analyze survey data. This can help businesses understand things like customer satisfaction and product preferences.



Conclusion

Result

PIC	Algorithm	Progress
Sophanon	MLP Regressor	100%
Sreylin	Linear Regression	100%
	SVR	100%
Vikreth	Random Forest	100%
Ponareay	Decision Tree	100%

Result

PIC	Task	Progress
Sreylin	Data Preprocessing	100%
Sophanon	Frontend	100%
	API	100%

@reallygreatsite

Demo


@reallygreatsite
Thank You
Any Questions ?
10%
20%
30%
40%
50%
70%