Data Mining Task

The task is to extract patterns using historical cryptocurrency data and apply these patterns to predict future prices. This will be accomplished using a type of Recurrent Neural Network(RNN) called Long Short-term Memory(LSTM) to look at historical market data from two cryptocurrencies in particular, Ethereum(ETH) and Bitcoin(BTC).

Dataset

The dataset consists of the previous 3 months history of Bitcoin and Ethereum prices pulled from Coinbase Pro API calls using the Historic_Crypto python library. This data comes in as a Pandas Data frame consisting of 5 columns which are time, low price, high price, open price, close price, and volume. These values are updated and captured every 5 minutes for a total of 288 data points for each day over the 3-month period. To avoid the use of excessive API calls, I will save the previous 3 months' worth of datapoints as a csv file and manipulate the data through accessing the csv file instead. This will consist of two csv files, one for Bitcoin history and one for Ethereum history.

Methodology

The datasets can be partitioned into distinct training and testing sets to build the model. From here, I will perform data normalization to help the model more accurately extract relevant features from the data values. This will be done using the python Sklearn preprocessing library which has built in functions available to process the data before extracting features. Next the processed data is used to build an LSTM model using the python Keras library. After building the model, I will perform some hyperparameter optimization techniques with the hope of improving accuracy of the model. This will include learning rate, batch size, number of layers, and the optimizer algorithm. This model is then tested using the testing data and the results plotted using the python library matlab.pyplot to display the results in a readable format.

Final Product

The primary outcome of this project is to build an LSTM model that performs time series forecasting which can be used to predict the future prices of Bitcoin and Ethereum using historical market data. I will then plot this data visually with graphs to display the results to the user in an easy-to-understand plot. The success of this project will be measured by looking at some metrics such as accuracy, precision, and recall. The goal of this project is to build an LSTM model that will help semi-accurately predict the future prices of Ethereum and Bitcoin.