Machine learning NanoDegree

Capstone Proposal

Name: kelvin

E-mail: Kelvinwj98@gmail.com

17/09/2021

Domain background

Nowdays cryptocurrency has become one of hottest topic ,it rivals traditional currency since cryptocurrency has advantage as being decentralized,in a decentralized world, game-theoritical incentives take over the role of a general, governing authority.

And now a lot of people has trying to invest in Cryptocurrency (Bitcoin,Ethereum and BNB),but the problem is commonly people don’t know how to valuing a coin,according to www.forbes.com/ 90% trader lose their money since it takes a lot of effort to mastering how to invest and trade.The goal of this project is to attempt to predict the price of coin in the next few days using machine learning algorithm to help people that trying to making money from cryptocurrency.

Problem Statement

The problem is timeseries problem.

The idea is to train the model to behave like an technical analyst,we want the model to take the input from historical data of coin price and make a prediction about the price in the future using RNN-LSTM.

Datasets and inputs

The datasets are provided by CoinMarketCap on Kaggle competition website.the datasets contains 2991 datas and 10 columns.

Input data fields:

- SNo- The id of training set.
- Name- The name of coin.
- Symbol- The abbreviation that contains 3 character from the coin (ex: Bitcoin-BTC , Ethereum-ETH).
- Date- The date of each dataFrame.
- High- The highest price at which a stock traded during a period.
- Low- The lowest price of the coin on that date.
- Open- Starting period of trading on a securities exchange.
- Close- End period of trading on a securities exchange.
- Volume- Volume is the amount of an asset or security that changes hands over some period of time, often over the course of a day.Marketcap- total dollar market value of a coin’s.

Solution statement

The solution we provide for this problem is to apply RNN-LSTM for Multivariate Timeseries forecasting.

First, I will try to process the dataset and get the visualization of the data to get some insight.

Then I will try to extract features from datasets and delete useless features such as Coin Name,Symbol and No.

Benchmark model

For this problem, the benchmark model will be using RNN-LSTM , and compare it with other multivariate timeseries forecasting algorithms.

Evaluation metrics

Prediction results are evaluated on Threshold metric(Accuracy)

Project design

I will first check the data features.see if there is features that wont be used for processing , check for nan values and that see how many features in dataset if there is too many I will perform dimension reduction to make it simplier for machine learning to train.

I will train the model using RNN and use a hyperparameter tuning to make the model automatically tuning the hyperparameter after done training with certain hyperparameter value.

And then I will take a look at training job with has the best accuracy value.

Reference:

1,https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistory

2, https://www.investopedia.com/terms/c/cryptocurrency.asp