Capstone Project

BitPred: A Bitcoin Price Predictor

Kanishka Pandey — E18CSE083

Department of Computer Science, Bennett University, Greater Noida

Abstract

Having learnt the basics of data science and machine learning in my third year, I already had experience with most of the widely used python libraries and interfaces in analyst workflows. Drawing on the knowledge I accumulated in more than three years of working on NLP, image and video processing, sentimental analysis semester projects and my interest in emerging technologies and frameworks, I decided to research the principles and working of blockchain technology and how it is being used to develop a completely anonymous currency that has real-life applications despite its vastly fluctuating perceived value to create a deep learning model on python that could be used to predict the prices of the most popular cryptocurrencies based on past trends and provide useful insights for its users and investors.

Introduction

I aim to capitalize on the cryptocurrency trend by using a variety of efficient deep learning algorithms to try to the predict the price of bitcoin which will then be extended to predict for a variety of currencies in circulation today. It primarily uses convolutional and recurrent neural network models to implement long short-term memory cells and gated recurrent units that will then predict the value that was observed for a single bitcoin on a certain date. After extensive training and hyperparameter tuning, the models in this project could be used to try to predict future values of various blockchain-based currencies anywhere from 5-20 minutes to days into the future.

This project is trying to be highly accessible, such that it could be used both by early investors in crypto as well as new users that are just now jumping on the bandwagon. Everyone from individuals looking to get into decentralized finance to large corporations wanting to diversify their investment portfolios can use these models to automate large amounts of data collection, draw inferences on past trends, and maybe even predict the near future prices to avoid sudden valuation drops and maximize their profits.

From l to r: dashboard wireframe, actual homepage, homepage wireframe

Proposed Method

As the goal is to create an interactive UI, The application utilizes a python flask server which when combined with some HTML/XML scripting and bokeh interactive plots creates a visually appealing web application which asks for user time period preferences and computes prediction in real-time, showing results returned by three models: SARIMAX, Facebook Prophet, and my custom LSTM network. The demo dataset was populated using the yfinance library which was then value-checked against the public datasets maintained by crypto exchanges like Poloniex before saving the .csv file.
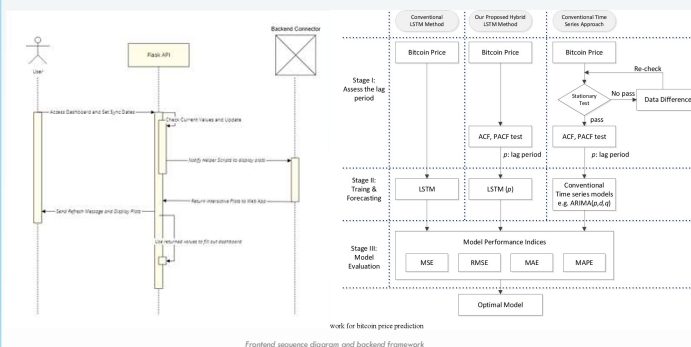

Frontend sequence diagram and backend framework

As this dataset represents a time series, the first overview focuses on seasonality present in the trends. As this did not result in any useful information right away, we now move on to a more technical statistical method in the form of time series decomposition. BitPred automatically analyses prices between any time period that the user selects and separates the data trend into 3 components: Trend component — changes along with the data in the long term. The period here updates at a fixed frequency which shows how the perceived value of the currency changes over time. Seasonal — This shows if the data follows patterns which are affected by seasonal factors. These factors may be yearly, weekly, daily or even hourly based on the sample scale and volatility. Residual — In short, everything left after isolating the above two components are recorded here. As the trends followed by the crypto market are largely dependent on the public interest in these technologies, the data will always have some random noise which may be analyzed but cannot be accurately predicted. Applying Loess decomposition on data from 2017–2021, we start to see some clear patterns that our models can easily utilize.
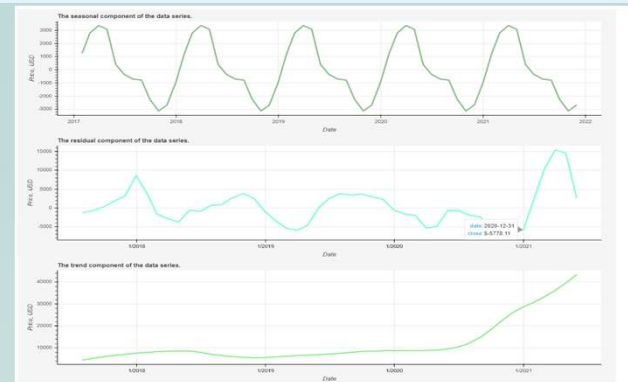
Experimental Results and Discussion

The initial model predictions seem to be quite promising and error values seem to fall under expected parameters. LSTMs with different number of features were tested and multiple validation parameters were recorded with different time periods. Underlying trends were visualized and stored for later scrutiny. The confidence intervals drawn over predicted prices helped quantify the precision of the results.
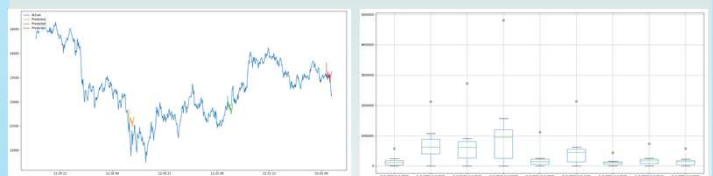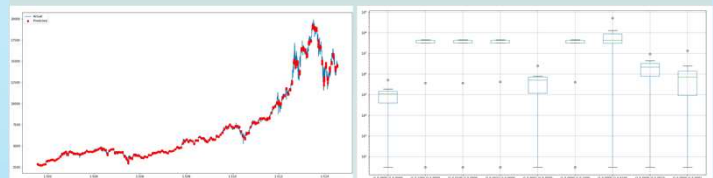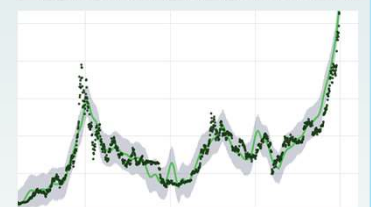

Root Mean Squared Error

ARIMA: 508.94965228374423

Prophet: 1786.1781454204304

LSTM: 424.92938066897835

Conclusions

This project would never have been made without the support of the CSE department and my mentors Dr. Deepak Garg and Dr. Suneet Kumar Gupta. While the UI part works well, all the three final models have their own pros and cons which can be analyzed and minimized by aggregating the results. The SARIMAX model has reliable forecasts but does not do well with limited data. Prophet offers a fast, automated solution but is less accurate as the hyperparameter tuning options are limited. My custom model tries to incorporate multiple use cases with a small number of parameters but is prone to overfitting and utilizes a lot of resources. The project has a lot of future scope and I hope to implement a more robust solution with options for predicting multiple financial assets.

References

- Currency Analyzer - https://github.com/rebeccabernie/CurrencyAnalyser

- Kriptomist - https://github.com/flowcoin/kriptomist

- BitVision - https://github.com/shobrook/BitVision

- An Analysis of the Viability and Volatility of Cryptocurrency - https://github.com/rebeccabernie/CurrencyAnalyser/blob/master/Dissertation.pdf

- LSTM regression using TensorFlow - https://mourafiq.com/2016/05/15/predicting-sequences-using-rnn-in-tensorflow.html

- Comparison between Seq2Seq RNNs and ARIMA models for Cryptocurrency prediction - https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1258222&dswid=-4305

- A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions - https://www.sciencedirect.com/science/article/abs/pii/S2214212620307535