CSE 587 : Data Intensive Computing

# Project Phase 3

**Authors:**
**Yahsika Nihalani 50425015**
yashikav@buffalo.edu
**Ashutosh Shailesh Bhawsar 50416025**
abhawsar@buffalo.edu

Instructor: **Prof. Eric Mikida**

Date Last Edited: December 3, 2022

**The framework used to build the web app running on localhost: Streamlit.**

# 1 Models used from phase 2

## 1.1 Random Forest Classifier:

- In Phase 2, we developed various models for predicting and classifying the bitcoin stock price.
- Among all the models, Random Forest Classifier, we can achieve around 64-65% accuracy using the golden/death cross identification and prediction of calls. This model can be considered moderately effective among all other models in the current market.
- Since we output the trading signals, the investors will be aware of some highly volatile conditions beforehand and can avoid huge losses and, in inverse conditions, be highly profitable.
- We used the Random Forest Classifier model to classify whether the user should Buy or Sell the stocks for the particular date he selects.

**Dataset sample:**

| | Timestamp | Open | Close | High | Low | Volume_BTC | Volume_Currency | Weighted_Price | Change_percentage | SMA10 | SMA20 | SMA50 | Trade_signal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3374 | 2021-03-27 | 55,081.2600 | 55,839.4200 | 56,686.1500 | 53,948.3500 | 1.8200 | 100,884.7300 | 55,193.3600 | 1.3800 | 55,420.7140 | 56,204.7735 | 51,692.5584 | 1.0000 |
| 3375 | 2021-03-28 | 55,817.8500 | 55,790.9200 | 56,573.0400 | 54,677.5100 | 1.4500 | 80,632.1200 | 55,832.9600 | -0.0900 | 55,235.0010 | 56,372.9295 | 52,022.7348 | 1.0000 |
| 3376 | 2021-03-29 | 55,790.2800 | 57,600.1000 | 58,402.6800 | 54,892.4200 | 3.7300 | 213,754.5600 | 56,913.9900 | 3.2400 | 55,187.8110 | 56,506.4835 | 52,397.5690 | 1.0000 |
| 3377 | 2021-03-30 | 57,623.6600 | 58,760.5900 | 59,388.6600 | 57,011.0000 | 2.3600 | 138,231.2400 | 58,346.9100 | 2.0100 | 55,250.0310 | 56,651.8985 | 52,644.4518 | 1.0000 |
| 3378 | 2021-03-31 | 58,767.7500 | 58,778.1800 | 58,778.1800 | 58,755.9700 | 2.7100 | 159,417.7500 | 58,764.3500 | 0.0300 | 55,389.6700 | 56,702.2030 | 52,889.9114 | 1.0000 |

**Figure 1:** Cleaned and preprocessed Dataset after EDA

### 1.1.1 Tuning of hyperparameters:

- For the Random Forest Classifier, the first parameter that we tuned was the SMA(Simple moving average). The model was tested with an SMA value of 10, 50, and 100.
- It was observed that the SMA value 10 gave the best accuracy and results, followed by SMA 50 and 100.
- So in phase 3, the SMA of 10 days was used.

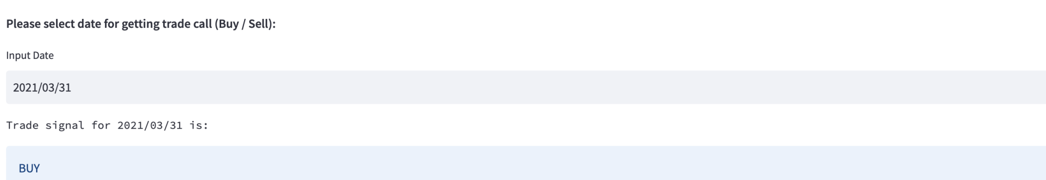


**Figure 2:** Random Forest Classifier input by user and output trade call in GUI

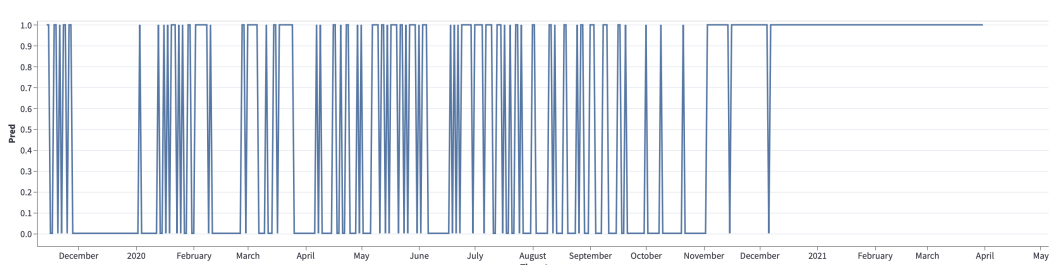


**Figure 3:** Random Forest Classifier model output interactive overview in GUI

## 1.2 FB Prophet model:

- The next model used from phase 2 was the FB Prophet Model.
- The reason for choosing this model to fit the bitcoin data is its ability to predict seasonality and trend. Along with the price prediction, it also provides upper and lower bounds of the forecast that can give us a range of predictions, and we can cut our losses if the market price goes the other way.
- We have three inputs for this model from the user on web UI: investment date (start date), liquidation date (end-date), and the amount of investment in BTC. We compare the initial and later portfolio values with the expected high and low values so that our model can be used as an effective decision-support system for maximizing the gains from the crypto market.

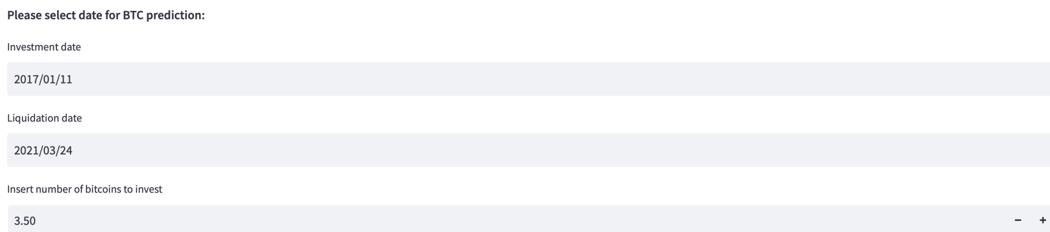


**Figure 4:** FB Prophet model inputs by user in GUI

### 1.2.1 Tuning of hyperparameters:

- For the FB Prophet model, the trend and seasonality are determined by the model itself, so the only thing we could do to get the best results is to change the training and testing data split.
- Our data is pre-processed to give daily prices, and the model determines trends for daily, weekly, and yearly intervals.
- We tested the RMSE and fit of the model for our data and found the 85:15 split gives the best results and most accurate forecasts.

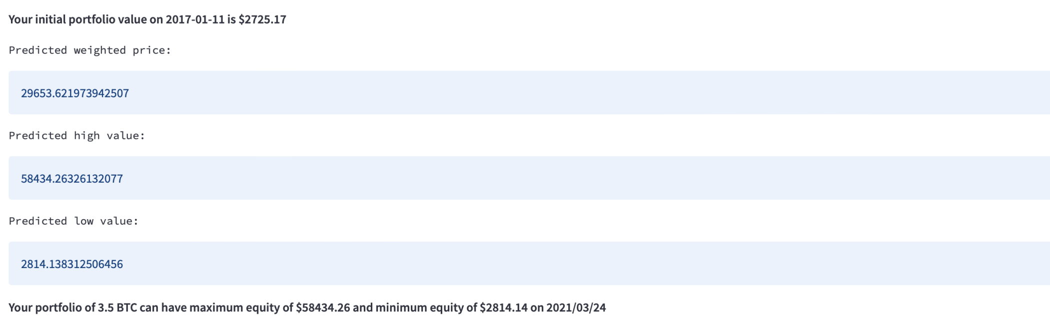


**Figure 5:** Prophet model outputs in GUI


**Figure 6:** Prophet model output interactive graph in GUI

# 2 Mapping results to the problem statement

## 2.1 Analyze the bitcoin data and realize trends in price, volume, and weighted average using time series analysis

We used different data modeling and exploration techniques to analyze data, realize trends and patterns in the data and get higher-order insights through different visualizations during the EDA phase. We also preprocessed the data suitable for determining the future price and devising profitable strategies.

## 2.2 Design suitable prediction algorithms/models to determine the price of bitcoin

We demonstrated five different statistical and machine learning models to help a trader/investor to make accurate trading calls and help in determining the future price of bitcoin. Linear regression showed us the overall trend of bitcoin, KNN gave us a good accuracy for the predicted price, LSTM showed us a better fit for training data and made us aware / warned us about the outliers such as unprecedented events happening around the world affecting crypto price.

## 2.3 Develop strategies to help an investor with trading decisions

Using Random Forest classification, we got the actual trading calls on a daily basis, and the FB prophet model showed us the forecasted upper and lower bounds necessary for a trader to cut losses and take profits at appropriate levels. We developed the moving average indicator from scratch and incorporated it in the random forest model to give the BUY or SELL calls for the desired date.

## 2.4 Expose the visualizations and build interactive UI for users

We used Streamlit for building interactive UI for getting inputs from a user and embedded two models: Random Forest Classifier and FB Prophet, for exposing the visualizations of graphs and relevant data as output. The GUI also prints out proper messages as output for each model, ensuring a user-friendly application.

# 3 Insights gained by the user

- The user can use our application to get insights into the trading patterns and output accordingly. The user can input the date, and the Random forest predicts whether the user can BUY or SELL the cryptocurrency on that particular date.

- The application also contains a graph that shows the trade signals with X-axis containing the timestamp and Y axis containing the predicted values. '0' indicates selling the cryptocurrency, and '1' indicates buying the cryptocurrency. This gives users a good insight and accordingly makes trading strategies to maximize profits.

- Using the FB prophet model, the user can input the investment date, the Liquidation date, which is the end date of the portfolio, and the number of bitcoins to invest. As an output user can see the Initial Portfolio value, the predicted weighted price, the predicted high value, and the predicted low value.

- The application also contains graphs that contain dates on the x-axis dates, and the y-axis contains the price of 1 bitcoin in USD. The user can check the graph, view predicted and actual values for a particular date, and devise a trading strategy accordingly.

## 4 Future scope and extension

- The future work for this application/project would be to accept an input file from the user containing the cryptocurrency historical data in the prescribed format (essential columns present). Then the data preprocessing, EDA, machine learning models implementation, and visualizations in GUI would take place in the exact same manner as that of the current implementation of BTC.
- The models can also be modified to consider external factors such as news and major events, which can potentially shift the price of cryptocurrency markets and help in better prediction in times such as recession or pandemic.
- The other models from phase 2 that are not included in phase 3 can be used to incorporate into the GUI, and the user can thus gain more in-depth insights about the market.

## 5 REFERENCES

1. https://docs.streamlit.io/library/api-reference/
2. https://medium.com/analytics-vidhya/how-well-can-machine-learning-models-predict-the-price-of-bitcoin-f036fdecdc03
3. Stock Trading With Random Forests, Trend Detection Tests and Force Index Volume Indicators