**DL Final Assignmet - Crypto Prediction - Classification**

**Built With -**

* Python - 3x
  + Data science packages
  + Dash
* Jupyter Notebooks / google colabs

**Installations required**

Please look at the **requirements.txt**

**Code Usage**

Overall end-to-end process has been divided into 5 steps.

**Classification**

(one or 2 line description what thenotebook does)

* 0\_preprocess\_data.ipynb
  + The script is developed to pool the data together and generate hourly/daily/weekly data
* 1\_Classification\_Hourly/Daily\_EDA.ipynb
  + In this script, we tried to explore partial correlation and seasonality of our target variable. After deciding to use daily data, there are additional EDA for only daily data
* 2\_model\_building\_tuning\_F3.ipynb
  + In this script, we only use columns provided in original data and time variables to build dummy classifier, random forest and LGBM
* 3\_model\_building\_tuning\_F13.ipynb
  + In this script, we generate lag, technical analysis and market features and apply dummy, random forest, XGB and LSTM classifier
* 4\_model\_building\_LSTM\_F13.ipynb
  + In this script, we use new features, roll the data and apply LSTM
* 5\_model\_Testing\_F13.ipynb
  + This script is developed to plot the comparison between different models

**Data Files**

<https://drive.google.com/drive/folders/1pa_hEBqyljZhN8mueUV2iam4Qr0ULNH4>

All coin minute data are pooled, and each coin is one hot encoded.

**Feature Sets :**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Feature Set** | **Frequency of data** | **Basic Features** | **Time Features** | **Technical Financial Metrics** | **Comments** |
| F1 | Hourly | - | Lag/ Mov Average for 5 periods , Market value & weightage | - | Calculating the Hourly Time lags |
| F2 | Hourly | - | Lag/ Mov Average for 12 periods  , Market value & weightage | - | Calculating the Hourly Time lags |
| F3 | Hourly | All of Basic Features |  | - | Building the model with the features present In the dataset |
| F4 | Hourly | All of Basic Features | Volatility for 14 \*24 periods &  , Market value & weightage | - | Checking for daily lags & volatility |
| F5 | Hourly | All of Basic Features | Mov Average/ Volatility on weekly basis &  , Market value & weightage | - | Checking for weekly  lags & volatility |
| F6 | Hourly | High', 'Low', 'Volume',date | Volatility on hourly, weekly & Hourly&  , Market value & weightage | - |  |
| F7 | Daily | High', 'Low', 'Volume',date | Volatility / Lag for 6 daily lags &  , Market value & weightage | - |  |
| F8 | Daily | High', 'Low', 'Volume',date | Volatility / Lag for 30 &  , Market value & weightage | - |  |
| F9 | Daily | High', 'Low', 'Volume',date | Mov Average/ Volatility on weekly basis &  , Market value & weightage | - |  |
| F10 | Daily | High', 'Low', 'Volume',date | Volatility on hourly, weekly & Hourly &  , Market value & weightage |  |  |
| F11 | Daily | High', 'Low', 'Volume',date | Lag/volatility for 30 days &  , Market value & weightage | Volaitily, trend & Momentum |  |
| F12 | Daily | High', 'Low', 'Volume',date | Lag for 30 days &  , Market value & weightage | Volaitily, trend & Momentum |  |
| F13 | Daily | High', 'Low', 'Volume',date | Lags for weekly, monthly , daily | Volaitily, trend & Momentum | This feature set is Best performer |
| **F14** | **Daily** | **High', 'Low', 'Volume',date** | **Lags 1day -> Rolling window 28 days** | **Volaitily, trend & Momentum** | **Used for LSTM** |