CSE523 Machine Learning

Project: Crypto Forecasting

Weekly Project Report: March 30, 2022

Team: CryptoPolice

Sameep Vani	AU1940049
Kashvi Gandhi	AU1940175
Khushi Shah	AU1920171
Kavya Patel	AU1940144

- Tasks performed this week:
    1. Made timestamp atomic by splitting it into date and time
    2. Implementation of Lag features
    3. Finding Correlation using heatmaps between all the features of the dataset for each Crypto Currency
- Outcomes of the tasks performed

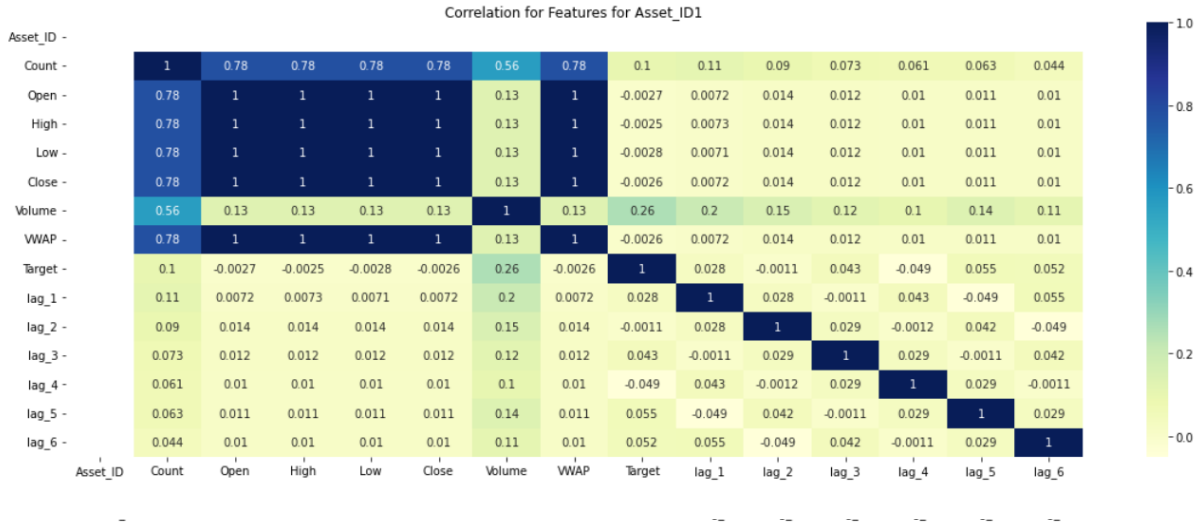
1. Making timestamp atomic.

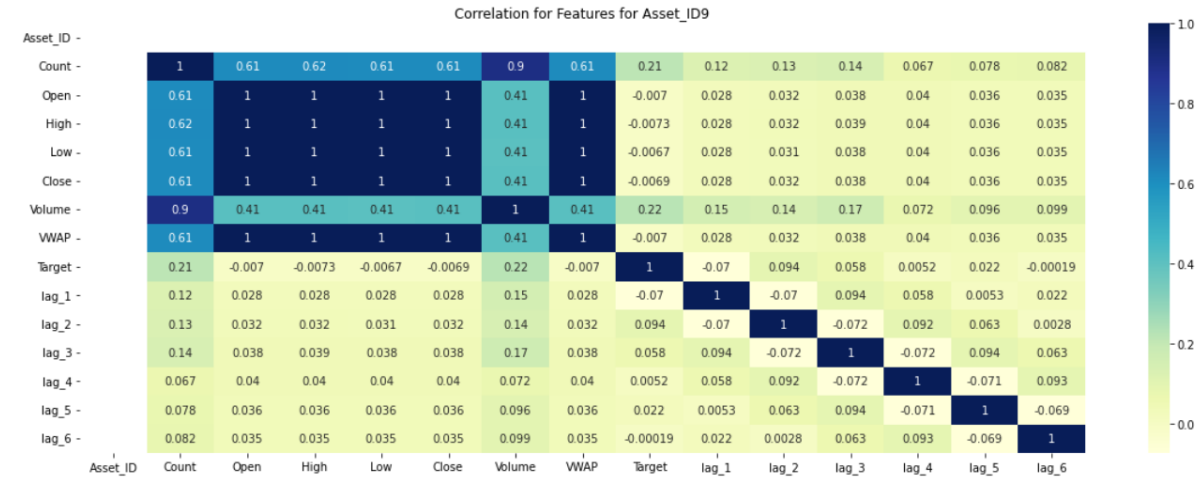
	timestamp	Asset_ID	Count	Open	High	Low	Close	Volume	VWAP	Target	new_date	new_time
0	2018-01-01 00:01:00	2	40.0	2376.5800	2399.5000	2357.1400	2374.5900	19.233005	2373.116392	-0.004218	2018-01-01	00:01:00
1	2018-01-01 00:01:00	0	5.0	8.5300	8.5300	8.5300	8.5300	78.380000	8.530000	-0.014399	2018-01-01	00:01:00
2	2018-01-01 00:01:00	1	229.0	13835.1940	14013.8000	13666.1100	13850.1760	31.550062	13827.062093	-0.014643	2018-01-01	00:01:00
3	2018-01-01 00:01:00	5	32.0	7.6596	7.6596	7.6567	7.6576	6626.713370	7.657713	-0.013922	2018-01-01	00:01:00
4	2018-01-01 00:01:00	7	5.0	25.9200	25.9200	25.8740	25.8770	121.087310	25.891363	-0.008264	2018-01-01	00:01:00

2. Implementing Lag Features.

	new_date	Asset_ID	lag_1	lag_2	lag_3	lag_4	lag_5	lag_6	Target
0	2018-01-01	0	NaN	NaN	NaN	NaN	NaN	NaN	0.000148
1	2018-01-02	0	0.000148	NaN	NaN	NaN	NaN	NaN	0.000393
2	2018-01-03	0	0.000393	0.000148	NaN	NaN	NaN	NaN	0.000549
3	2018-01-04	0	0.000549	0.000393	0.000148	NaN	NaN	NaN	0.000006
4	2018-01-05	0	0.000006	0.000549	0.000393	0.000148	NaN	NaN	0.005618

3. Finding Correlation between the features of the dataset using heatmaps of the seaborn library for each cryptocurrency





- Tasks to be performed next week:
    1. We will implement different models like Linear regression, Decision Tree, KNN Regression, Random Forest Regression, XGBoost, LGBM, ARIMA, SARIMA, etc.
    2. Moreover, we will complete the hyperparameter optimization using the 'optuna' library for a best-fitting model and submit the predictions.