Project report

Computer Configuration:

MacBook Pro 2022 Apple M2
RAM: 8 GB
MACOS: 14.1.2

Coding environment:

Jupyter notebook
Google colab

Data Source:

Yahoo finance

Pipe line:

data process.ipython
1,download data from yfinance
2,check missing values and trading days
3,check outliers and anomalies

Feature engineering.ipython
1,create function to calculate variance, correlation, return and volume
2,add feature to the ticker for different time windows

Feature selection.ipython
1,use BorutaShap to choose the 15 most important features for predicting the target.

Wavenet.ipython
Construct wavenet methods to predict the target.
TabNet.ipython
Construct TabNet methods to predict the target.

UNet.ipython
Construct UNet methods to predict the target.
LGB.ipython
Construct LightGBM methods to predict the target.

MLP.ipython

Construct MLP methods to predict the target.

1dCNN.ipython

Construct 1 dimensional CNN methods to predict the target.

According to the performance on GOOG:

MLP>UNET>TabNet>1dCNN>WaveNet>LGB

So we continue with MLP on other 4 tickers and get the rmse for each ticker:

GOOG:0.087
JPM:0.073
XOM:0.077
AMZN:0.084
GME: 0.155