Machine Learning in Intraday Stock Trading

Motivation

The ability to precisely predict the price movement of stocks is the key to profitability in trading. In light of the increasing availability of financial data, prediction of price movement in the financial market with machine learning has become a topic of interests for both investors and researchers.

Our group aims to predict whether the price in the next minute will go up or down, using the time series data of stock price, technical-analysis indicators, and trading volume.

Data & Features

Data

SPDR S&P 500 trust (NYSE: SPY) 1-minute interval data from March 1st until May 24th 2019.

Source: IEX Trading

Data points are labeled as positive if the price increases in the next time interval, and negative if the price decreases.
Training set: 9381 positive & 9337 negative examples
Test set: 1045 positive & 1091 negative examples

Features

Original Features: Closing Price, Trading Volume, Date, and Minute Mark

Augmented Features: Simple Moving Average (SMA), Exponential Moving Average (EMA), Crossovers, and other technical Indicators with 5, 10, 12, 20, 26, 50, 100, 200 days lookback window)

Models

Baseline Model: Logistic Regression with and without regularization

Support Vector Machine: SVC with Linear, Polynomial (degree 3), Sigmoid, and Radial Basis Function kernel. The variables are tuned by adjusting the cost of constraint violation as well as the constant of regularization term in Lagrange formulation.

Neural Network model: Single layer Convolutional Neural Network with linear layer and cross entropy loss.

RNN models: Single-layer LSTM, Multi-layer LSTM, and Multi-layer GRU with 128 hidden units in each layer with RELU activation function. Regularization includes early stopping and dropping out parameters. Hyperparameter tuning is performed by using grid search method.

Results

Models	Accuracy (Training Set)	Accuracy (Test Set)	AUC (Test Set)	Profit/Hour (Test Set)
Baseline (Logistic)	0.4988	0.4899	0.4876	$-0.11
SVM (Linear)	0.5354	0.5341	0.5355	$0.68
SVM (Polynomial)	**0.5452**	**0.5449**	**0.5449**	**$1.61**
SVM (RBF)	0.5433	0.5384	0.5393	$0.79
SVM (Sigmoid)	0.5024	0.4983	0.5021	$-0.69
GRU	0.5141	0.5096	0.4928	$0.51
LSTM (Single-Layer)	0.5011	0.4983	0.4989	$-0.20
LSTM (Multi-Layer)	0.5127	0.5110	0.4817	$0.47
CNN	0.5130	0.4889	0.5000	$-0.46

Table1: Summary of results from different models

Note: Training set has 18,718 data points, while test set has 2,136 data points

Discussion

- From our observations of the results, we find out that higher accuracy does not necessarily translate to higher economic profits.
- For instance, our GRU model outperforms the multi-layer LSTM in terms of accuracy, but it generates less profit.
- Our correct predictions could correspond to time with small changes (less profit); some incorrect predictions might correspond to time of large changes (huge loss).
- At this point, we weren't able to control for that. Thus, the accuracy of model predictions doesn't reflect actual profits.

Future Work

In the future we hope to modify our models to takes into account the magnitude of profit or loss. This includes multi-class classification that accounts for magnitude of price movements or even regression models predicting next-minute price. We also hope to expand our models to incorporate data from other stocks.

Reference

[1] Chen, Sheng & He, Hongxiang. (2018). Stock Prediction Using Convolutional Neural Network. IOP Conference Series: Materials Science and Engineering. 435. 012026. 10.1088/1757-899X/435/1/012026.

[2] Hegazy, Osman, Soliman, Omar S. & Salam, Mustafa A (2013) Machine Learning Model for Stock Market Prediction. Faculty of Computers and Informatics, Cairo University, and Higher Technological Institute (H.T.I), 10th of Ramadan City, Egypt

[3] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). doi:10.1109/icacci.2017.8126078