DJIA-Direction-Prediction

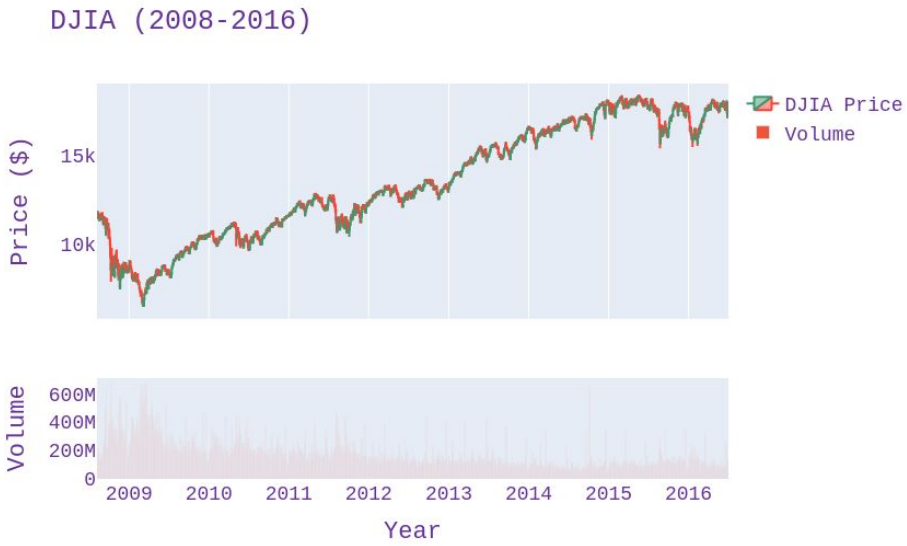
David Pacheco Aznar
Computational Learning, Computational Mathematics and Data Analysis, 2022

Link to GitHub Repo:
https://github.com/Davidpazn/DJIA-Direction-Prediction

Index

- Objective
- Preprocessing
- Models
- Results
  - Test
  - Stacking Ensemble?
  - Out of Bag
- Conclusions and Future Work

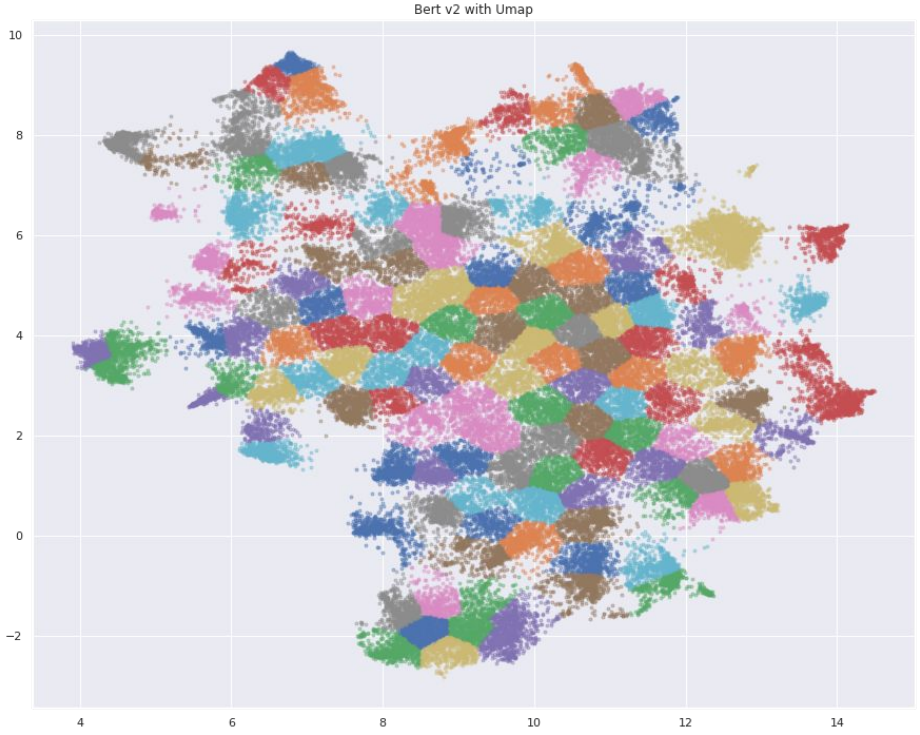
Objective

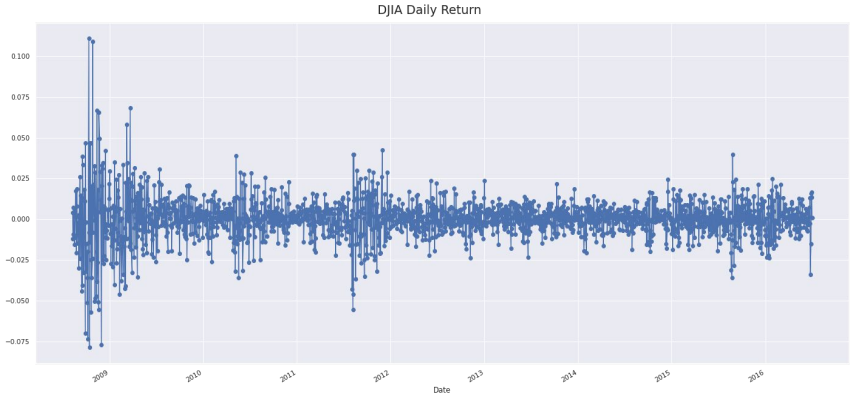


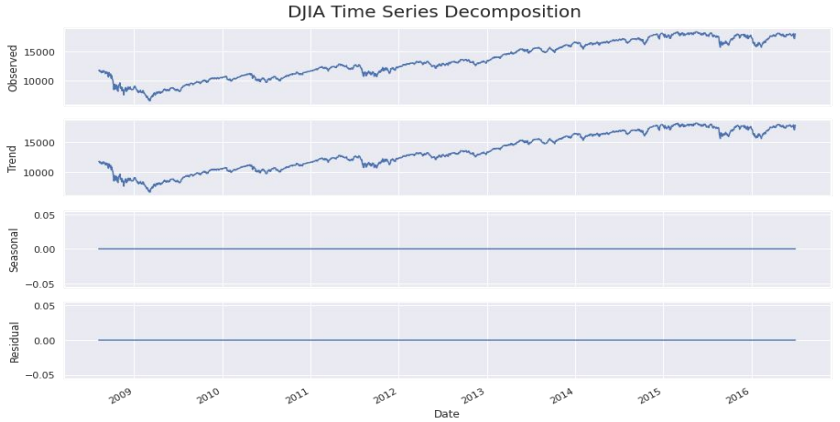
reddit

Dataset from:
https://www.kaggle.com/aaron7sun/stocknews

Preprocessing





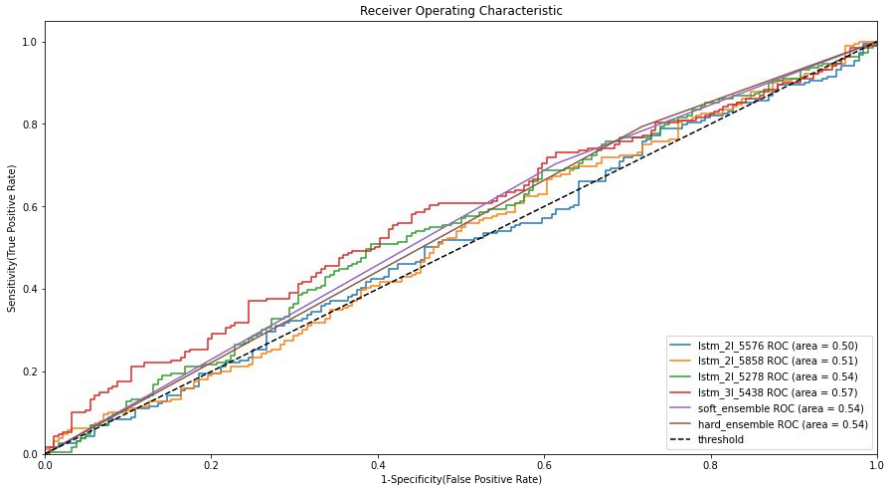


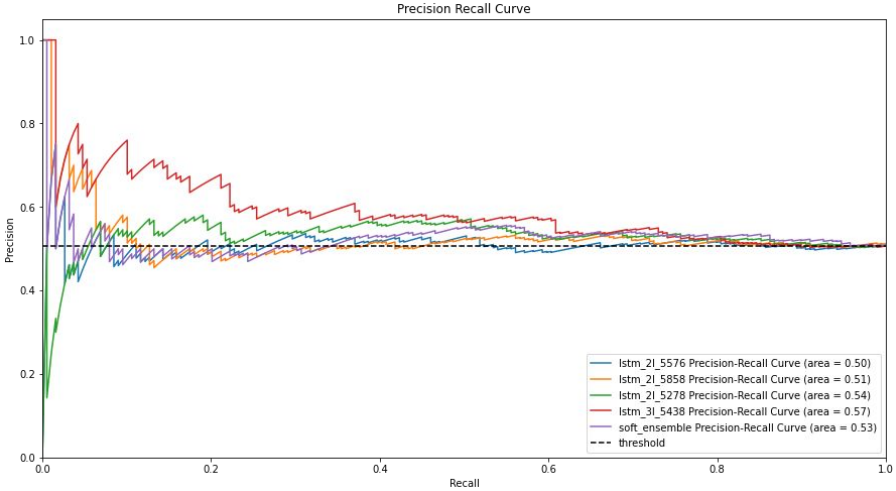
Top Performing Tuned Models

Model	Time (5 Folds CV)	AUC (Tuned) - CV
2 Layer LSTM	1 min 15s	Up to 58.68
3 Layer LSTM	3 min 20s	Up to 57.28
Soft Voting Ensemble: LGBM[1] + RF[2] + XGB	20s	Up to 56.60
Hard Voting Ensemble: LGBM[1] + RF[2] + XGB	21s	Up to 58.25

1. OPTUNA: Bayes optimization
2. Skopt: Bayes optimization

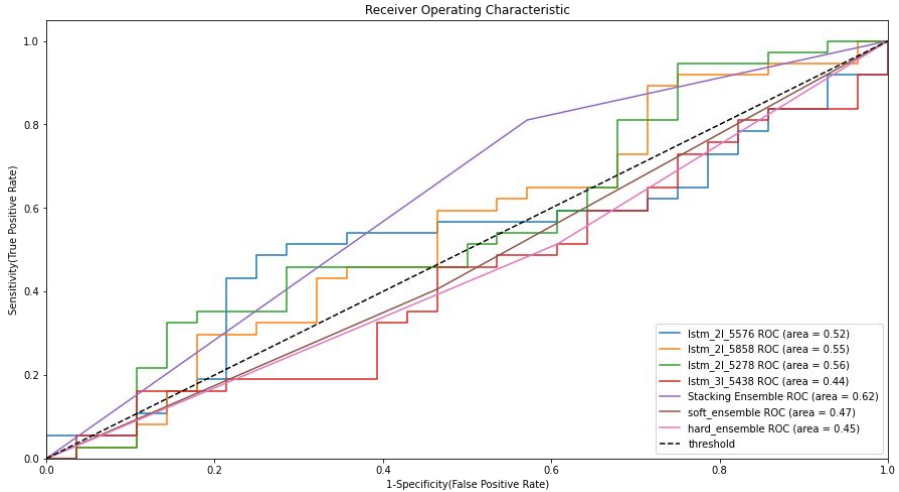
Results: Test

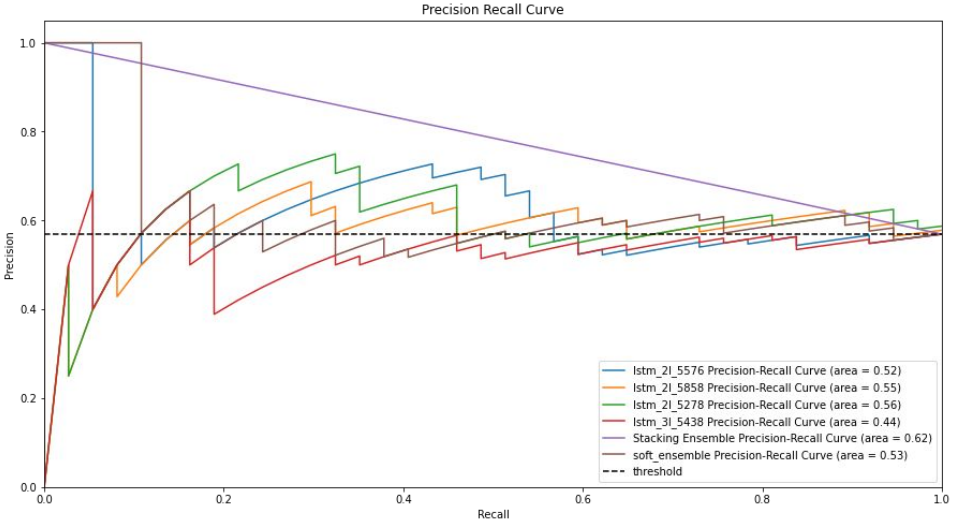


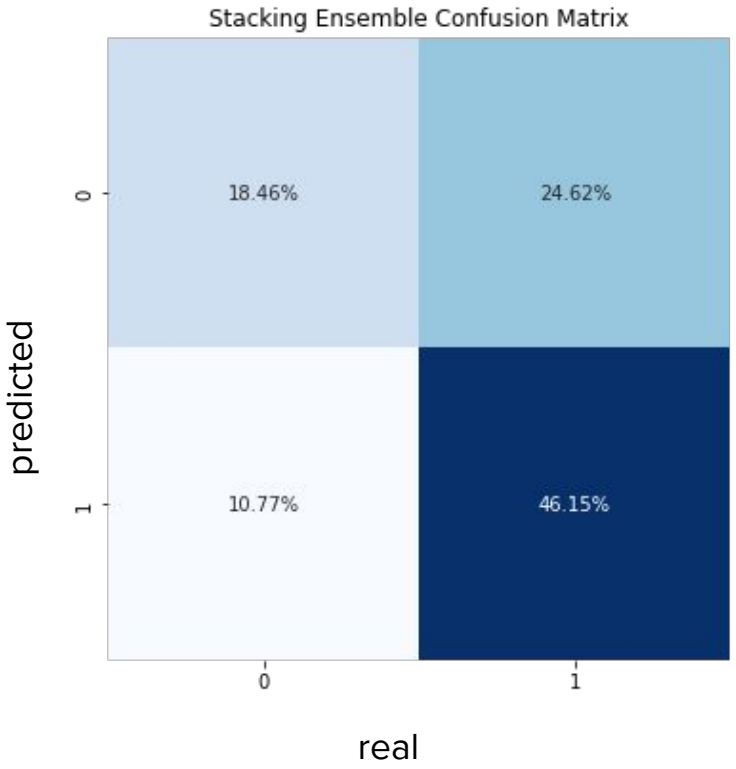


Stacking Ensemble of top 2 performing LSTMs on test-set?

Results: Out Of Bag







Stacking Ensemble

	precision	recall	f1-score	support
0	0.63	0.43	0.51	28
1	0.65	0.81	0.72	37
accuracy			0.65	65
macro avg	0.64	0.62	0.62	65
weighted avg	0.64	0.65	0.63	65

Conclusions and Future Work

- Stacking Ensemble performed best
- Lack Of Time Series Data
- Might be better idea to add industry-specific news instead of geopolitical ones
- Use NN to join probabilities of stacked LSTMs