EECS 6895 Project Final Presentation

Stock Performance Prediction and Recommendation with Deep Learning Analysis

Jiaqing Chen (jc5657)
Wannuo Sun (ws2591)

TABLE OF CONTENTS

Background

Background

- An era of growing economy and financial inflation
- People always eager to earn money from stock
- Widely used big data & AI

Challenges

- Stock price is influenced by multiple factors
- Stock is usually random walk
- Volume of data is significantly large

GOALS

The goal of our system is to comprehensively analyze stock and predict the stock price from both short-term and long-term view for a certain number of companies. With the help of the cutting-edge deep learning models, it could serve as a tool for investors to make their investment decisions.

Timeline Progress

Milestone 1

Literature Survey & Project Proposal

Milestone 2

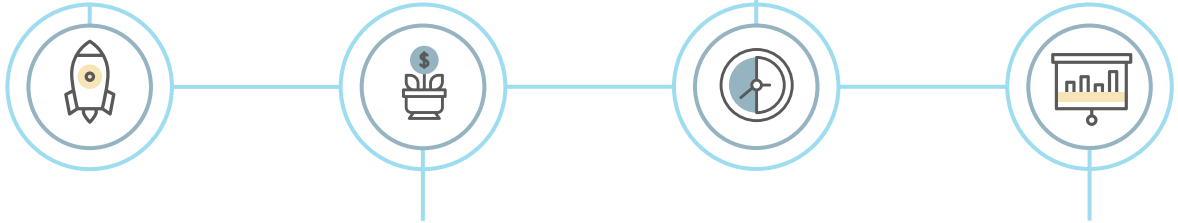
Data Collection & Basic model construction

Milestone 3

Twitter Collection & Model Enhancement & Front-end

Final report

Enlarge dataset & System Scale & Analysis Automation



Literature Survey

Financial analysis

Technical analysis: use pure historical data to predict stock market

Fundamental analysis: combine with financial factors like tweets and news

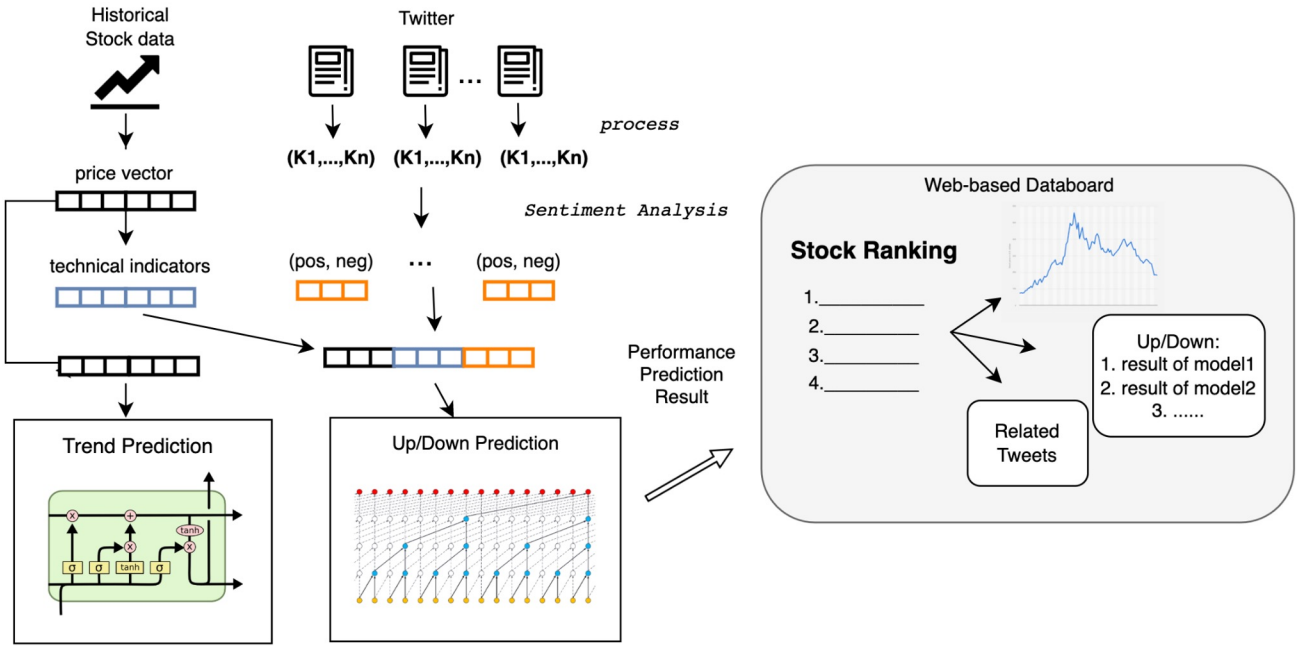
Time-series Deep Learning model

LSTM: -dependency reflection
-prediction accuracy
-smaller model update cycle
TCN: up-to-date genetic architecture

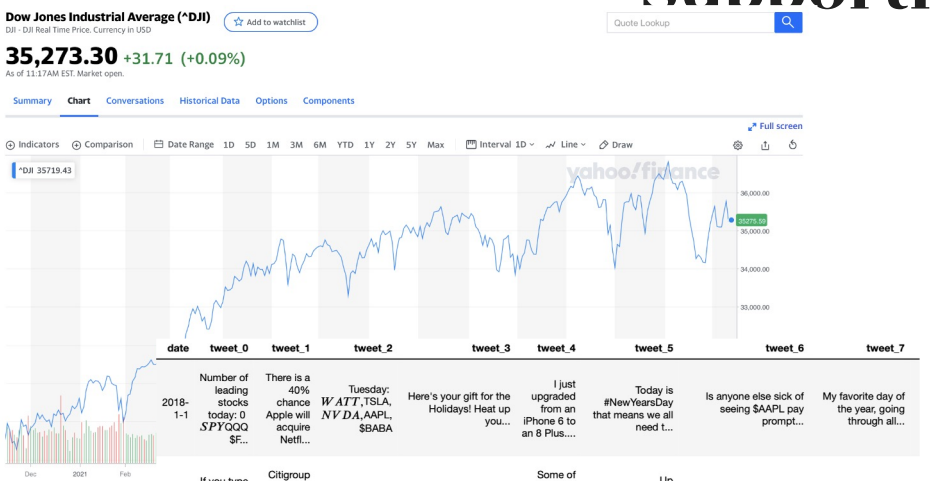
Sentiment in Behavior Economics

The nature of the market trend is impacted by sentiment analysis in the stock market.
Twitter: rich information about the companies

System Design



Supporting Dataset



date	tweet_0	tweet_1	tweet_2	tweet_3	tweet_4	tweet_5	tweet_6	tweet_7
2018-1-1	Number of leading stocks today: 0 $SPY $QQQ $F...	There is a 40% chance Apple will acquire Netfl...	Tuesday: $WATT, $TSLA, $NVDA, AAPL, $BABA	Here's your gift for the Holidays! Heat up you...	I just upgraded from an iPhone 6 to an 8 Plus....	Today is #NewYearsDay that means we all need t...	Is anyone else sick of seeing $AAPL pay prompt...	My favorite day of the year, going through all...
2018-1-2	If you type "Sell" in Google Finance, it'll br...	Citigroup comments that there was a "40% chanc...	"Nay, we don't need patents...or a strong pate...	180 for AAPL? This month it's possible more ...	Some of my "best performing #STOCKS TODAY 1/2:...	Up Moves:\n+3.7% $XME\nFXI\n+3% + 3.5 $EWZ\n+...	$SIEB $YY $WUBA $X $NFLX $FSLR $AAPL $FB $ERX ...	Number of leading stocks today: 74 $SPY $QQQ $...
2018-1-3	U can make ur iPhone 6 basically new again by ...	$LITE best way to play surging 3D sensing mark...	The Nasdaq Composite closed above 7,000 for th...	$BABA $TQ 200-250 $MU 90 $AAPL 220 $NVDA $3...	Long in $BAC $AAPL $CSX $JPM $KHC $KO #WallStreet	In long on $BAC $AAPL $KHC $KO #WallStreet	Number of leading stocks today: 76 $SPY $QQQ $...	Learning to trade profitably and comfortably i...
2018-1-4	$AAPL Says all IOS devices affected. All MAC ...	$AAPL Headlines hitting known meltdown issues ...	#AAPL is undervalued. Senseless debates genera...	If AAPL doesn't hit a $1Trillion MKTcap, I wil...	AAPL store holiday sales broke some kind of re...	$AAPL news addd 6 cents to QQQ post market #Lucky	Guys, do you also think $AAPL should buy $NFLX...	Over half a billion users visit the iOS App St...

Yahoo! Finance

Historical stock data (2018-2022): price & volumes

Calculate technical indicators

Twitter:

sntwitter.TwitterSearchScraper API (2018-2022): 60 tweets per day

Extract Sentiment

Methodology - Long-term Trend Prediction

Goal: Predict next 15 day stock price (For a long-term general trend)

Model: LSTM (3-layer with dropout)
Predict on: Previous 30-day open price

Roll-over prediction: involve new prediction in next-day prediction

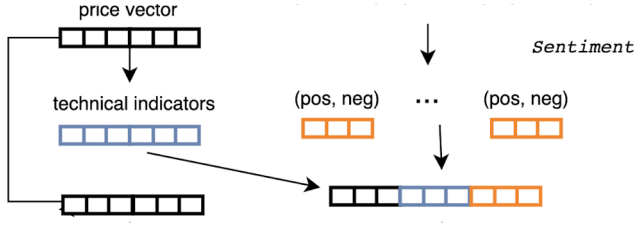
Performance:

- Observe Prediction Trend vs. Actual Trend

Methodology - Next-day Up/Down Prediction

Data encoder: history driven + knowledge driven

- **stock price vector:** open, close, ...
- **+ technical indicators:** Exponential Moving Average (EMA), Average Direction Index (ADX) …
- **+ sentiment factors**
  - VADAR model
    - Compound score >= 0.05 --> positive
    - Compound score <= -0.05 --> negative
  - Positive proportion = #pos / (#pos + #neg)
  - Negative proportion = #neg / (#pos + #neg)



Methodology - Next-day Up/Down Prediction

Goal: Predict whether the stock will go up or go down in the next day (For a short-term quick judgement)

Model: Self-made ensemble model
Model involvement: traditional machine learning models (ex. SVM), LSTM, TCN model
Voting decision: voted based on models with acc>75%;

Class definition:

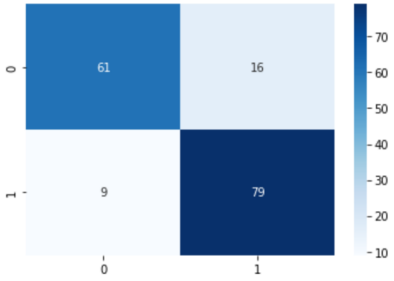
$$y_i = \begin{cases} 1, & p_{i+n} \leq p_i \times (1+\alpha) \\ 0, & p_{i+n} > p_i \times (1+\alpha) \end{cases},$$

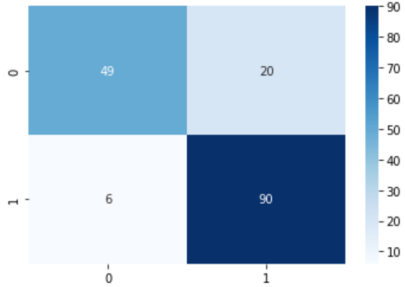
Performance: Different stock model has different prediction power due to stock's predictability

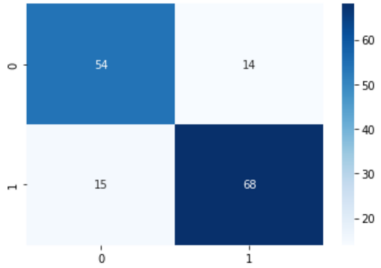
- Accuracy score: ~80% accuracy (at worst >60%)
- Confusion Matrix

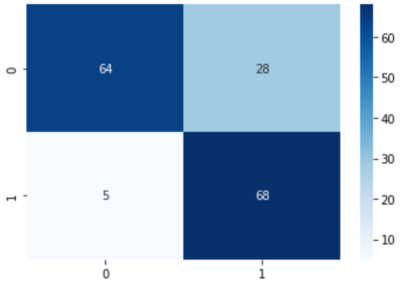
Methodology - Next-day Up/Down Prediction

Model	Accuracy
KNN	0.7784
SVM	0.8352
Random Forest	**0.8466**
Logistic Regression	0.8125
Naive Bayes	0.7613
LSTM	0.8182
TCN	**0.8636**









Collective Prediction

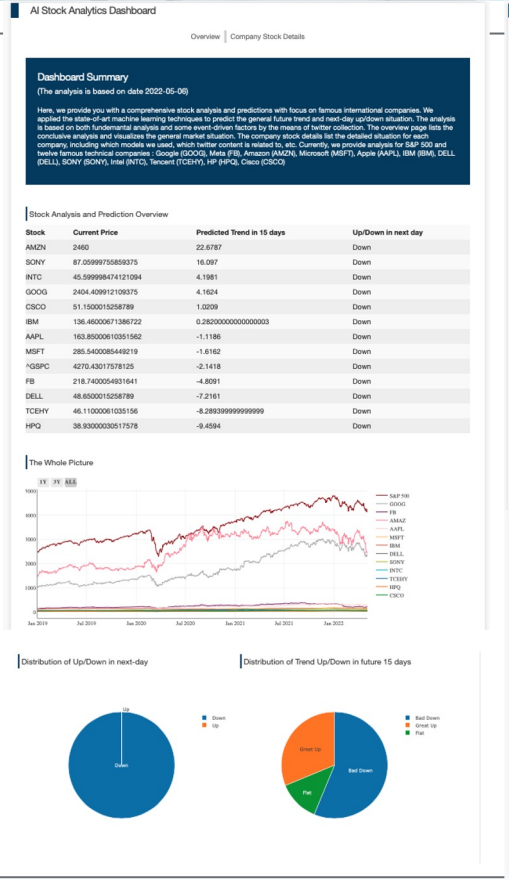
Historical-driven trend prediction

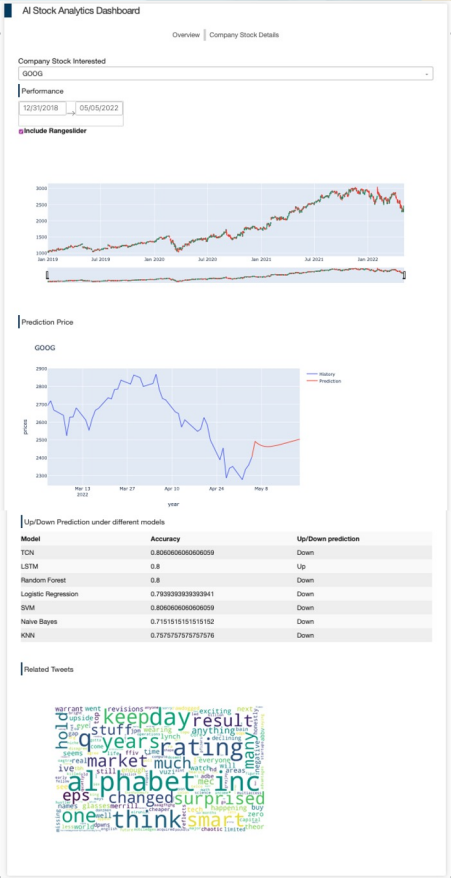
- Period prediction => General trend for a 15-day window
- Investment plan

Up/Down class prediction

- Next-day prediction => Immediate up or down tomorrow
- Investment decision today

Stock	Current Price	Predicted Trend in 15 days	Up/Down in next day
AMZN	2297	22.6787	Up
SONY	84.43000030517578	16.097	Down
INTC	44.4900016784668	4.1981	Down
GOOG	2310.3798828125	4.1624	Up
CSCO	49.1500015258789	1.0209	Up
IBM	135.47000122070312	0.28200000000000003	Up
AAPL	156.00999450683594	-1.1186	Up
MSFT	274.80499267578125	-1.6162	Up
^GSPC	4128.169921875	-2.1418	Up
FB	207.33999633789062	-4.8091	Down
DELL	46.75	-7.2161	Down
TCEHY	44.400001525878906	-8.289399999999999	Down
HPQ	37.349998474121094	-9.4594	Down





Front-end Dashboard

Current analysis is for 12 technical companies:
Google, Meta, Amazon, Apple, Microsoft, IBM, Dell, Intel, Tencent, Cisco, Sony, HP

Two pages: Overview & Company stock details

Installation: (GitHub will be public soon)

clone https://github.com/amayaqing/stock-prediction-with-deep-learning-analysis.git

$ python3 app.py

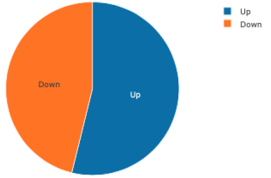
Automatically update to today's prediction when opening the dashboard

Overview

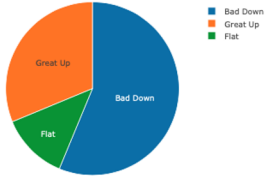
Stock Analysis and Prediction Overview

Stock	Current Price	Predicted Trend in 15 days	Up/Down in next day
AMZN	2297	22.6787	Up
SONY	84.43000030517578	16.097	Down
INTC	44.4900016784668	4.1981	Down
GOOG	2310.3798828125	4.1624	Up
CSCO	49.1500015258789	1.0209	Up
IBM	135.47000122070312	0.28200000000000003	Up
AAPL	156.00999450683594	-1.1186	Up
MSFT	274.80499267578125	-1.6162	Up
^GSPC	4128.169921875	-2.1418	Up
FB	207.33999633789062	-4.8091	Down
DELL	46.75	-7.2161	Down
TCEHY	44.400001525878906	-8.289399999999999	Down
HPQ	37.349998474121094	-9.4594	Down

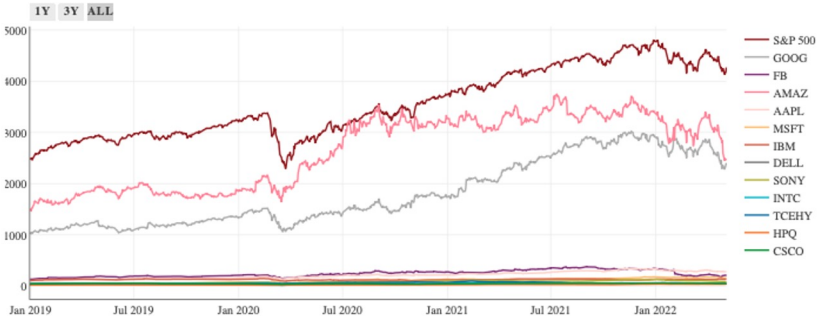
Distribution of Up/Down in next-day



Distribution of Trend Up/Down in future 15 days



The Whole Picture



Company Stock Interested

TCEHY

Up/Down Prediction under different models

Model	Accuracy	Up/Down prediction
TCN	0.8060606060606059	Down
LSTM	0.8	Up
Random Forest	0.8	Down
Logistic Regression	0.7939393939393941	Down
SVM	0.8060606060606059	Down
Naive Bayes	0.7151515151515152	Down
KNN		

Prediction Price

Company Stock Details

Select the company stock interested

Related Tweets

System is easily scalable!

Six Company stock

["GOOG", "FB", "AMZN", "MSFT", "AAPL", "IBM"]

Twelve company stocks

["GOOG", "FB", "AMZN", "MSFT", "AAPL", "IBM", 'DELL', 'SONY', 'INTC', 'TCEHY', 'HPQ', 'CSCO']

Future: More...

This system serves as a start

Conclusion

Conclusion

- Build a system that comprehensively analyze and predict the stock price from both short-term and long-term view
- Historical-driven & Knowledge driven
- Easily scalable system --> serve for future extension

Discussion and Possible Future enhancement

- More twitter could be better
- More twitter involvement if any better way
- More element-driven, ex. company correlation, financial report.....

References

[1] X . Zhou, Z. Pan, G. Hu, S. Tang and C. Zhao, "Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets," Mathematical Problems in Engineering, pp. 1-11, 2018.

[2] T. Mikolov, K. Chen, G. Corrado and J. Dean, "Efficient Estimation of Word Representations in Vector Space," ArXiv:1301.3781, 2013.

[3] J. Pennington, R. Socher and C. D. Manning, "GloVe: Global Vectors for Word Representation," in Empirical Methods in Natural Language Processing (EMNLP), 2014.

[4] S. Deng, N. Zhang, W. Zhang, J. Chen, J. Z. Pan and H. Chen, "Knowledge-Driven Stock Trend Prediction and Explanation via Temporal Convolutional Network," Companion Proceedings of The 2019 World Wide Web Conference, 2019.

[5] X . Ding, Y. Zhang, T. Liu and J. Duan, "Deep Learning for Event-Driven Stock Prediction," in Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.

[6] S . Bai, J. Z. Kolter and V. Koltun, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling," ArXiv:1803.01271, 2018.

[7] J . Liu, Z. Lu and W. Du, "Combining Enterprise Knowledge Graph and News Sentiment Analysis for Stock Price Prediction," in Hawaii International Conference on System Sciences, 2019.

[8] J. Bollen, H. Mao and X. Zeng, "Twitter mood predicts the stock market," Journal of Computational Science, vol. 2, pp. 1-8, 2011.

[9] W . Xu, W. Liu, C. Xu, J. Bian, J. Yin and T. Liu, "REST: Relational Event-driven Stock Trend Forecasting," ArXiv:2102.07372, 2021.

[10] T. Lin, T. Guo. and K. Aberer, "Hybrid neural networks for learning the trend in time series." In Proceedings of the twenty-sixth international joint conference on artificial intelligence, no. CONF, pp. 2273-2279. 2017.

[11] J. Bollen, H. Mao and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science,* vol. 2, pp. 1-8, 2011.

[12] A. A. Ariyo, A. Adewumi and C. Ayo, "Stock Price Prediction Using the ARIMA Model," in UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, 2014.

THANKS!