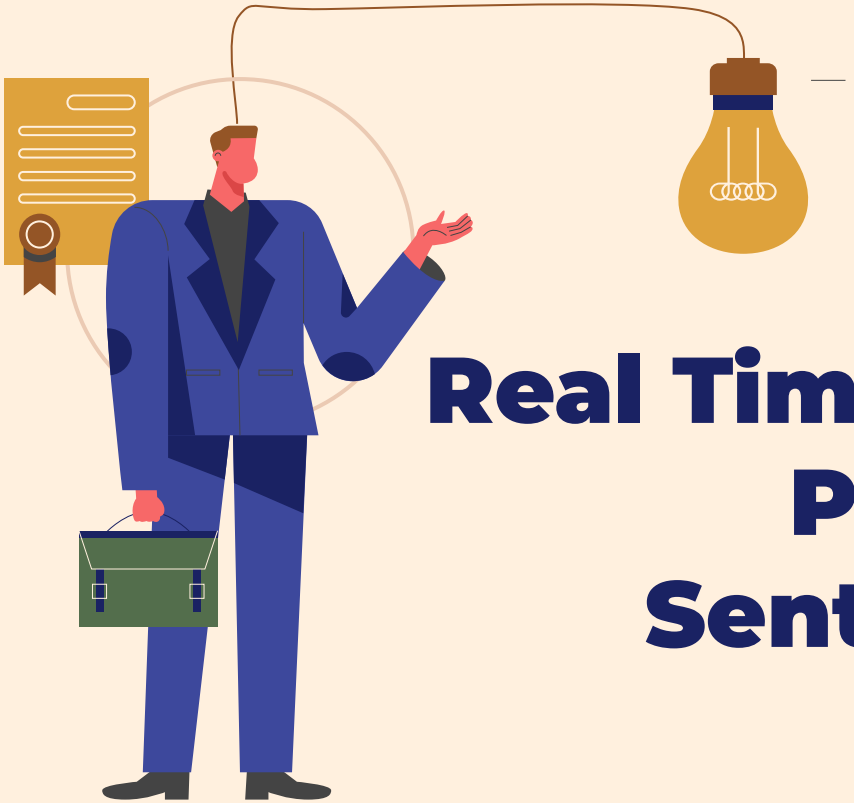
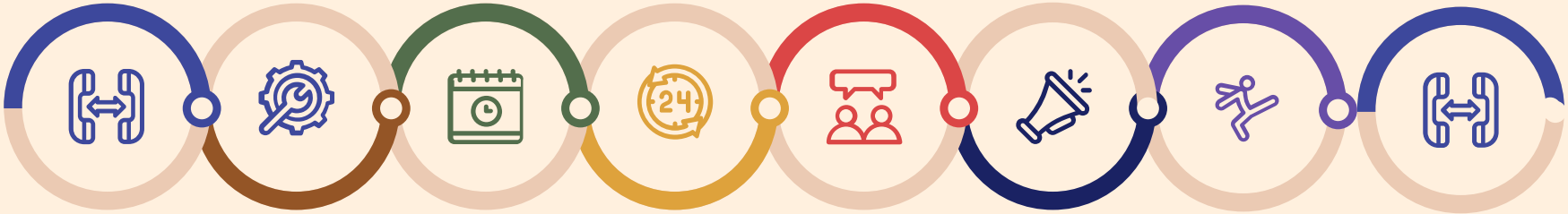


Real Time Stock Market Prediction using Sentiment Analysis

Group: ERROR 429 - G2 Team 3

Neo Jia Ying, Wong Wei Ling, Tay Yu Liang, Nor Aisyah, Ng Si Ying

Table of Contents



01
Our Client

Tiger Brokers

- Leading global online broker
- Regulated by Monetary Authority of Singapore
- Provides clients with a secure and reliable trading platform
- Easy access to global markets with low commissions
- Relatively new as compared to others brokerage companies
    - Interactive Brokers (IBKR), TD Ameritrade (TDA SG)

Situation Analysis

- ***Tiger brokers main customer segment: Retail Investors***
  - Non-professional investors who typically trades with smaller amounts than institutional investors
  - Have ***limited knowledge*** on investing stocks
    - Have to individually search for stock price predictions from multiple sources
    - Manually compare them before making a decision on the stock(s) to invest in
    - ***Time-consuming*** *and* ***inefficient***

Business Problem

Tiger Brokers wants to increase their **market share** by having **more unique selling points** that would **appeal to retail investors** to use Tiger Brokers.

Proposed Analytical Solution

Analyse Movement of Stock Prices

Use Twitter Sentiments, Machine Learning and Deep Learning models to **predict** and **forecast** stock prices to **analyse** stock price movement

02
Techniques & Methodology

Solution Architecture

Scrape **tweets related to a stock**

Data Cleaning using **Pandas**

Perform **Sentiment Analysis** on Scrapped Twitter dataset using **chosen Sentiment Classifier**

Get **respective stock prices** from **Yahoo Finance**

Train and **test models (Supervised, Reinforcement, Deep learning)**

Evaluate models using **RMSE**

Forecast using the **best model**

Dashboard for **visualisation** of **actual**, **predicted** and **forecasted** prices

Evaluate Sentiment Classifiers Performance **(Accuracy & F-score)**

Test Python Libraries **Sentiment Classifiers' performance** (Flair, Vader, Textblob) using Kaggle dataset

Obtain dataset with **labelled** sentiments from **Kaggle**

Scraping Tweets

01 Tweepy

```
consumer_key    = 'aQe6gaW[redacted]
consumer_secret = 'ngJEZ5c[redacted]
access_token    = '121942[redacted]
access_token_secret   = 'oXl0gX[redacted]

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth)
```

02 Twint

```
!pip3 install --user --upgrade git+https://github.com/twintproject/twint.git@origin/master#egg=twint
import twint
import nest_asyncio
c = twint.Config()
c.Search = "GameStop"
c.Since = '2021-03-10'
c.Until = '2021-03-10'
nest_asyncio.apply()
c.Store_csv = True
c.Output = "/content/drive/MyDrive/FA PROJECT 2021/gme_twintoutput_1day.csv"
twint.run.Search(c)
```

03 Snscrape

```
import snscrape.modules.twitter as sntwitter
import pandas as pd

tweets_list = []

for i,tweet in enumerate(sntwitter.TwitterSearchScraper('NVAX lang:en since:2021-03-20 until:2021-03-21').get_items()):
    tweets_list.append([tweet.date, tweet.content])

tweets_df = pd.DataFrame(tweets_list, columns=['Datetime', 'Text'])
tweets_df = tweets_df[::-1]
```

Scraping Stock Prices

```python
import yfinance as yf

data = yf.download(  # or pdr.get_data_yahoo(...
        # tickers list or string as well
        tickers = "NVAX",

        # use "period" instead of start/end
        # valid periods: 1d,5d,1mo,3mo,6mo,1y,2y,5y,10y,ytd,max
        # (optional, default is '1mo')
        period = "60d",

        # fetch data by interval (including intraday if period < 60 days)
        # valid intervals: 1m,2m,5m,15m,30m,60m,90m,1h,1d,5d,1wk,1mo,3mo
        # (optional, default is '1d')
        interval = "30m",

        # group by ticker (to access via data['SPY'])
        # (optional, default is 'column')
        group_by = 'ticker',

        # adjust all OHLC automatically
        # (optional, default is False)
        auto_adjust = True,

        # download pre/post regular market hours data
        # (optional, default is False)
        prepost = False,

        # use threads for mass downloading? (True/False/Integer)
        # (optional, default is True)
        threads = True,

        # proxy URL scheme use use when downloading?
        # (optional, default is None)
        proxy = None
    )
```

01 **yfinance**

Deciding on a Sentiment Classifier

- 3 sentiment classifiers: **Vader, Textblob and Flair**
- Tested classifiers accuracy with a labelled Kaggle Dataset (tweets with sentiment labelled as shown in the image)
- First column - Sentiment of particular tweet
  - 0 → Negative
  - 2 → Neutral
  - 4 → Positive
- Classifiers accuracy result will be discussed in 'Preliminary Results' section

0		1467810369	Mon Apr 06 22:19:45	NO_QUERY	_TheSpecialOne_	@switchfoot http://t
0		1467810672	Mon Apr 06 22:19:49	NO_QUERY	scotthamilton	is upset that he can't
0		1467810917	Mon Apr 06 22:19:53	NO_QUERY	mattycus	@Kenichan I dived m
0		1467811184	Mon Apr 06 22:19:57	NO_QUERY	ElleCTF	my whole body feels
0		1467811193	Mon Apr 06 22:19:57	NO_QUERY	Karoli	@nationwideclass nc
0		1467811372	Mon Apr 06 22:20:00	NO_QUERY	joy_wolf	@Kwesidei not the w

Machine Learning Regression Model (Supervised Learning)

Linear regression

Minimise residual sum of squares between the actual data points and predicted values by linear approximation

Random Forest

Constructs **decision trees when training** and outputs the **mean of classes** as prediction of all trees

Adaboost

Fits dataset regressor and **fits additional copies of regressor** on same dataset with **weights of instances adjusted** according to error of current prediction

Machine Learning Regression Model (Reinforcement Learning)

ARIMA

Uses regression analysis to identify **strength of a dependent variable compared to other changing variables**. Uses **moving average** of observation and calculates residual error and applies it to the lagged observations

SARIMAX

Add-on of ARIMA which **supports time-series data** with **seasonal component.** SARIMAX can include **exogenous variable** which is the **compound score** from **Vader**.

Deep Learning Models

Long short-term memory (LSTM)

Recurrent neural network (RNN) architecture. Uses **previous information** to **affect** the **next information**. Use LSTM to decide sequence prediction problem as dataset is mainly **time dependent** and **ordered by time**.

Time-series forecasting tends to look at **seasonality trends**, hence we use LSTM to see if stock price will increase or decrease at a 30 minutes interval

03

Dataset

Labelled Kaggle Dataset

- Columns:
  - target: the polarity of the tweet (0 = negative, 2 = neutral, 4 = positive)
  - ids: The id of the tweet
  - date: the date of the tweet
  - flag: The query
  - user: the user that tweeted
  - text: the text of the tweet
- Used for **testing the accuracy** of Flair, Vader & Textblob

0	1467810369	Mon Apr 06 22:19:45	NO_QUERY	_TheSpecialOne_	@switchfoot http://t
0	1467810672	Mon Apr 06 22:19:49	NO_QUERY	scotthamilton	is upset that he can't
0	1467810917	Mon Apr 06 22:19:53	NO_QUERY	mattycus	@Kenichan I dived m
0	1467811184	Mon Apr 06 22:19:57	NO_QUERY	ElleCTF	my whole body feels
0	1467811193	Mon Apr 06 22:19:57	NO_QUERY	Karoli	@nationwideclass nc
0	1467811372	Mon Apr 06 22:20:00	NO_QUERY	joy_wolf	@Kwesidei not the w

Scrapped Tweets

- Scrapped tweets regarding desired stock using ***Snscraper***
- Eg. Tweets relating to **'PFE'**
- 2 columns (Datetime, text)
- **23770** tweets from 30 December 2020 to 26 March 2021

	Datetime	Text
99	2020-12-30 09:33:47	@DalhiMackle @RNAiAnalyst 1) $NVAX scale up is...
100	2020-12-30 09:39:42	@sciencescanner I'm not sure this is advice a ...
101	2020-12-30 09:41:01	The kind of news you love to start the day. Be...
102	2020-12-30 10:04:22	@iangtobin @sciencescanner Hardly anyone will ...
103	2020-12-30 10:12:40	@alabamawins17 It's hard to tell, MRNA has alr...
...	...	...
23765	2021-03-26 23:30:00	CHECKING IN ON .@JIMCRAMER PREDICTIONS\n\n#STO...
23766	2021-03-26 23:45:22	The Value Stock Rotation Isn't Over—Not by a L...
23767	2021-03-26 23:47:41	【Get rich in 100 days】\nDay 18☺\n\nDOW🚹NAS🚹...
23768	2021-03-26 23:48:27	$nvax $mrna $pfe another article on the subjec...
23769	2021-03-26 23:54:54	$PFE\nI've been waiting on this strong move ab...

Scrapped Stock Prices

- Scrapped stock prices regarding desired stock using ***yfinance***
- Eg. Stock prices of ‘PFE’

Datetime	Open	High	Low	Close	Volume
2020-12-30 09:30:00	37.174999	37.217999	36.869999	37.099998	2945949
2020-12-30 10:00:00	37.099998	37.240002	37.009998	37.009998	1901644
2020-12-30 10:30:00	37.009998	37.139999	37.005001	37.090000	1440046
2020-12-30 11:00:00	37.080002	37.180000	37.070000	37.079899	996011
2020-12-30 11:30:00	37.075001	37.160000	37.020000	37.044998	1033359
...	...	...	...	...	...
2021-03-26 13:30:00	35.950001	36.080002	35.950001	36.064999	2182093
2021-03-26 14:00:00	36.063099	36.099998	36.009998	36.095001	1614327
2021-03-26 14:30:00	36.090000	36.130001	36.030102	36.099998	2005702
2021-03-26 15:00:00	36.095001	36.125000	36.020000	36.040001	1353892
2021-03-26 15:30:00	36.030102	36.290001	36.029999	36.250000	4236630

04
Data Cleaning and Preprocessing

Scrapped tweets

- For **Kaggle** dataset (for testing sentiment classifiers performance)
  - Drop irrelevant columns (e.g. Id, Date, Flag, User)
- For both **Kaggle** dataset and **Scrapped Tweets** dataset
  - Removing noise that will affect the analysis
    - @ sign
    - Username
    - Retweet
    - # hashtag
    - Http links
    - Emojis

Datetime	Text	processed_tweet
2020-12-30 00:02:19+00:00	Sellers Weaken the Stock Market's Hold On Over...	sellers weaken the stock markets hold on overb...
2020-12-30 00:03:41+00:00	$DECN #Covid19 Testing Is Pretty Much Old News...	decn covid19 testing is pretty much old news t...
2020-12-30 00:10:08+00:00	@PersimmonTI $tril \nToo bad Jan's reply wasn...	tril too bad jans reply wasnt ill call pfe and...
2020-12-30 00:13:01+00:00	#Epilepsy Market Carries Exceptional Growth $9...	epilepsy market carries exceptional growth b b...
2020-12-30 00:22:01+00:00	@megtirrell Makes sense, but simultaneously 1a...	makes sense but simultaneouslya should be vacc...
...	...	...
2021-03-26 23:30:00+00:00	CHECKING IN ON .@JIMCRAMER PREDICTIONS\n\n#STO...	checking in on predictions stockstobuy cost ...
2021-03-26 23:45:22+00:00	The Value Stock Rotation Isn't Over—Not by a L...	the value stock rotation isnt overnot by a lon...
2021-03-26 23:47:41+00:00	【Get rich in 100 days】\nDay 18☺\n\nDOW🇯NAS🇯...	get rich in days day downasmy pf i never cut m...
2021-03-26 23:48:27+00:00	$nvax $mrna $pfe another article on the subjec...	nvax mrna pfe another article on the subject t...
2021-03-26 23:54:54+00:00	$PFE\nI've been waiting on this strong move ab...	pfe ive been waiting on this strong move above...

Stock Prices

- Join yfinance's close stock prices into dataframe with tweets using 'Datetime'

	Datetime	Text	processed_tweet	Close
99	2020-12-30 09:33:47	@DalhiMackle @RNAiAnalyst 1) $NVAX scale up is...	nvax scale up issue which caused a month dela...	37.099998
100	2020-12-30 09:39:42	@sciencescanner I'm not sure this is advice a ...	im not sure this is advice a doctor or pharmac...	37.099998
101	2020-12-30 09:41:01	The kind of news you love to start the day. Be...	the kind of news you love to start the day bet...	37.099998
102	2020-12-30 10:04:22	@iangtobin @sciencescanner Hardly anyone will ...	hardly anyone will be better off waiting espe...	37.009998
103	2020-12-30 10:12:40	@alabamawins17 It's hard to tell, MRNA has alr...	its hard to tell mrna has already had a huge c...	37.009998
...	...	...	...	...
23765	2021-03-26 23:30:00	CHECKING IN ON .@JIMCRAMER PREDICTIONS\n\n#STO...	checking in on predictions stockstobuy cost ...	36.250000
23766	2021-03-26 23:45:22	The Value Stock Rotation Isn't Over—Not by a L...	the value stock rotation isnt overnot by a lon...	36.250000
23767	2021-03-26 23:47:41	【Get rich in 100 days】\nDay 18☺\n\nDOW🚹NAS🚹...	get rich in days day downasmy pf i never cut m...	36.250000
23768	2021-03-26 23:48:27	$nvax $mrna $pfe another article on the subjec...	nvax mrna pfe another article on the subject t...	36.250000
23769	2021-03-26 23:54:54	$PFE\nI've been waiting on this strong move ab...	pfe ive been waiting on this strong move above...	36.250000

Sentiment Analysis

- Performed sentiment analysis

	Datetime	Text	processed_tweet	Close	Comp	Negative	Neutral	Positive
0	2020-12-30 09:33:47	@DalhiMackle @RNAiAnalyst 1) $NVAX scale up is...	nvax scale up issue which caused a month dela...	37.099998	0.5994	0.080	0.765	0.155
1	2020-12-30 09:39:42	@sciencescanner I'm not sure this is advice a ...	im not sure this is advice a doctor or pharmac...	37.099998	0.6639	0.042	0.820	0.138
2	2020-12-30 09:41:01	The kind of news you love to start the day. Be...	the kind of news you love to start the day bet...	37.099998	0.7430	0.000	0.867	0.133
3	2020-12-30 10:04:22	@iangtobin @sciencescanner Hardly anyone will ...	hardly anyone will be better off waiting espe...	37.009998	0.4019	0.047	0.849	0.104
4	2020-12-30 10:12:40	@alabamawins17 It's hard to tell, MRNA has alr...	its hard to tell mrna has already had a huge c...	37.009998	0.2263	0.059	0.844	0.097
...	...	...	...	...	...	...	...	...
6438	2021-03-26 15:00:07	S&P 500 Index Update 📈\n\nCharts are relea...	sampp index update charts are released earlie...	36.040001	0.5719	0.091	0.731	0.177
6439	2021-03-26 15:04:51	$PFE Has A Poor #Technical Analysis Score (TA ...	pfe has a poor technical analysis score ta sco...	36.040001	-0.7351	0.205	0.795	0.000
6440	2021-03-26 15:23:12	$BNTX: BioNTech and Pfizer (PFE) receive EMA a...	bntx biontech and pfizer pfe receive ema appro...	36.040001	0.4767	0.000	0.807	0.193
6441	2021-03-26 15:24:13	Today Top Flow in S&P 500 #SP500, Buy Flow...	today top flow in sampp sp500 buy flow amp sel...	36.040001	0.2023	0.000	0.951	0.049
6442	2021-03-26 15:27:44	#stockstobuy \n\n🔥 $FDBL 🔥\n\n*Friendable Reta...	stockstobuy fdbl friendable retains sheppard...	36.040001	0.6249	0.000	0.859	0.141

Final Data Used for Analysis

- Columns:
  - Close prices
  - Compound score
    - Normalised between -1 and 1

Datetime	Close	Comp
2020-12-30 09:30:00	37.099998	0.668767
2020-12-30 10:00:00	37.009998	0.143571
2020-12-30 10:30:00	37.090000	0.317000
2020-12-30 11:00:00	37.079899	0.047514
2020-12-30 11:30:00	37.044998	0.317800
...	...	...
2021-03-26 13:00:00	35.950001	0.415055
2021-03-26 13:30:00	36.064999	0.270407
2021-03-26 14:00:00	36.095001	0.342650
2021-03-26 14:30:00	36.099998	0.091308
2021-03-26 15:00:00	36.040001	0.228140

05

Preliminary Results

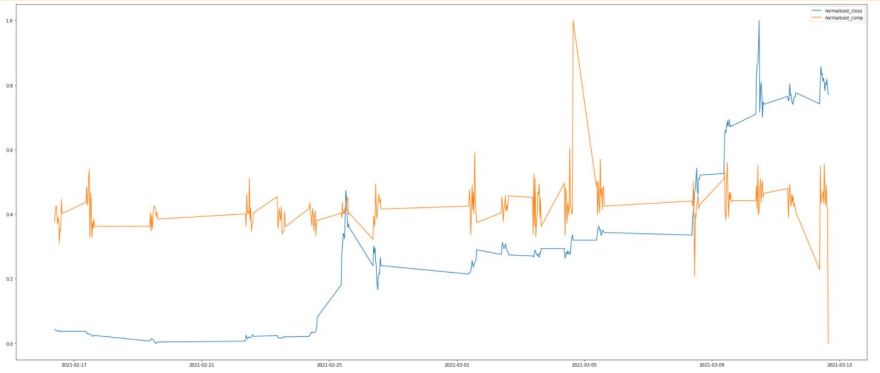
Choosing the best sentiment classifier

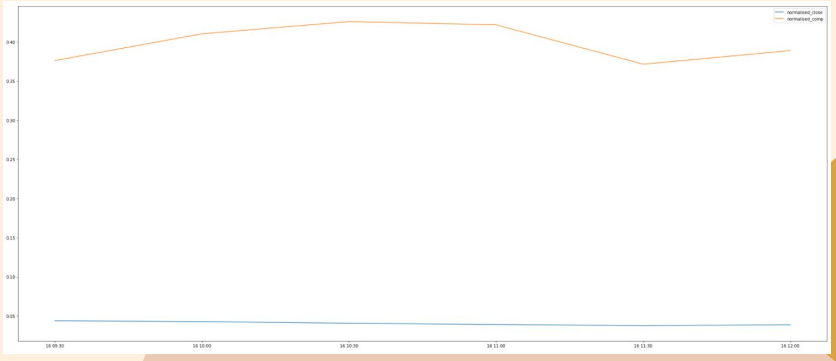
	Vader	Textblob	Flair
Accuracy	65%	56%	50%

- Works well with emojis & slangs
- Yield good results when used on Twitter data

Correlation Analysis

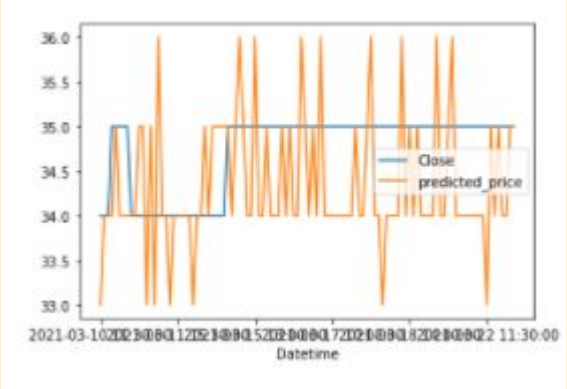
- Performed correlation analysis on 4 stocks (GameStop, SuperCom, Pfizer and AMC)
  - Gamestop: 0.22
  - Supercom: -0.0956
  - Pfizer: 0.0918
  - AMC: -0.0243
- **Shifted** Close prices **upwards** as we feel that the effect of the sentiments of tweets will be at a later timing/days
  - Gamestop: only a period of time where the correlation is **0.685**
  - Supercom, Pfizer, AMC: **No time period** where correlation is **>=0.5**
- We still decided to move on with our prediction as **correlation does not imply causation**





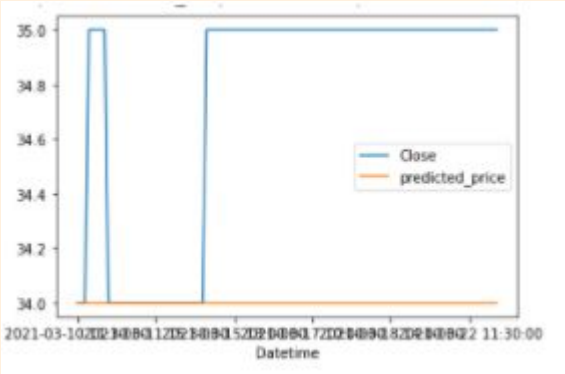
Supervised Learning Regressors Result

Random Forest



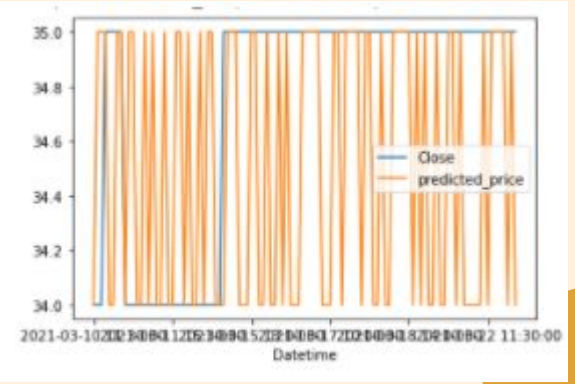
RMSE: 0.915

Linear Regression



RMSE: 0.640

Adaboost



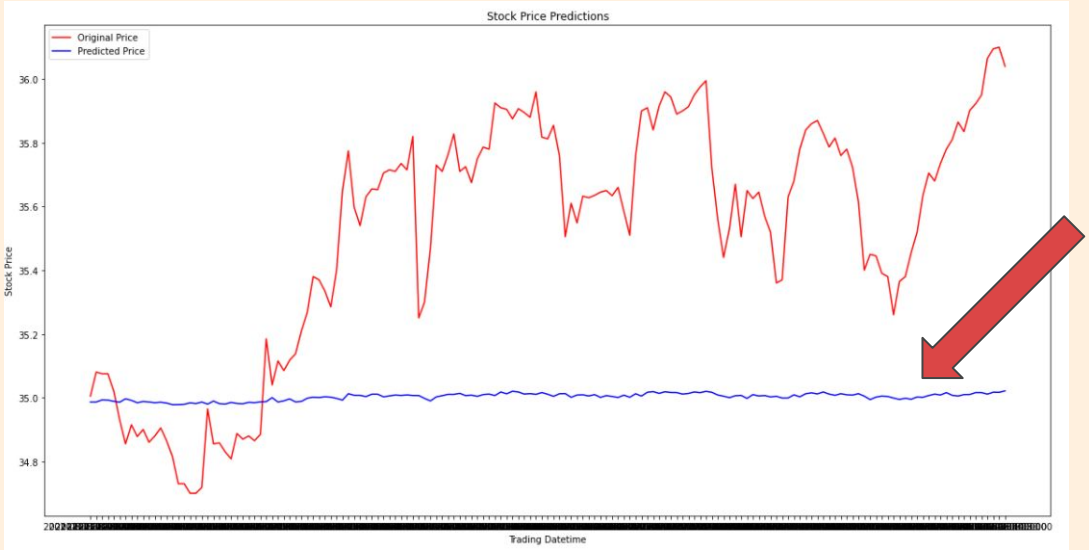
RMSE: 0.568

Reinforcement Learning ARIMA/SARIMAX

Datetime	Pred_Close	Actual_Close
2021-03-05 11:30:00	34.127029	34.180099
2021-03-05 12:00:00	34.130745	34.145000
2021-03-05 12:30:00	34.127200	34.195000
2021-03-05 13:00:00	34.128812	34.264999
2021-03-05 13:30:00	34.135996	34.340000
...	...	...
2021-03-23 12:30:00	34.164293	35.625000
2021-03-23 13:00:00	34.160342	35.645000
2021-03-23 13:30:00	34.161567	35.570000
2021-03-23 14:00:00	34.157918	35.520000
2021-03-23 14:30:00	34.159647	35.360001

159 rows × 2 columns

- **RMSE: 0.640**
- Stock price prediction is not very accurate

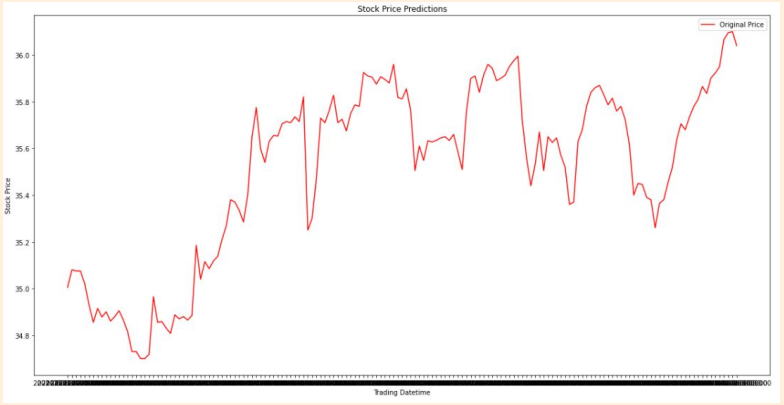


Reinforcement Learning ARIMA/SARIMAX

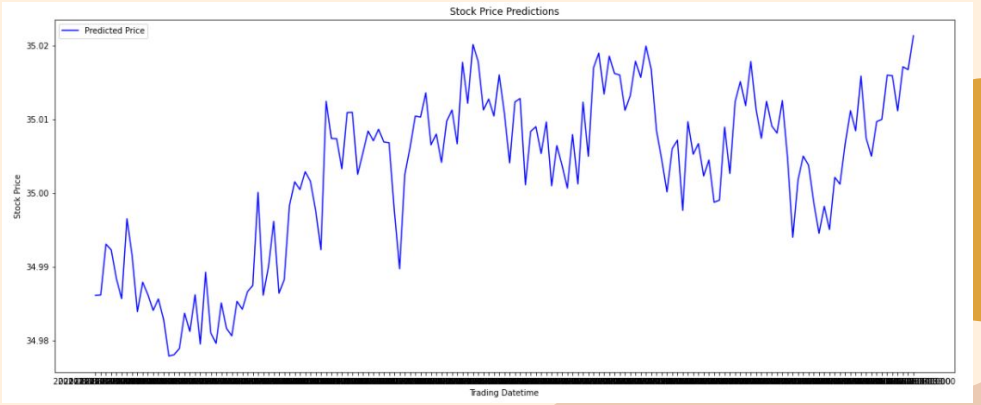
However, when analysed **individually**

- **Predicted** stock price **movement** is quite **consistent** with the **actual** stock price movement

Actual



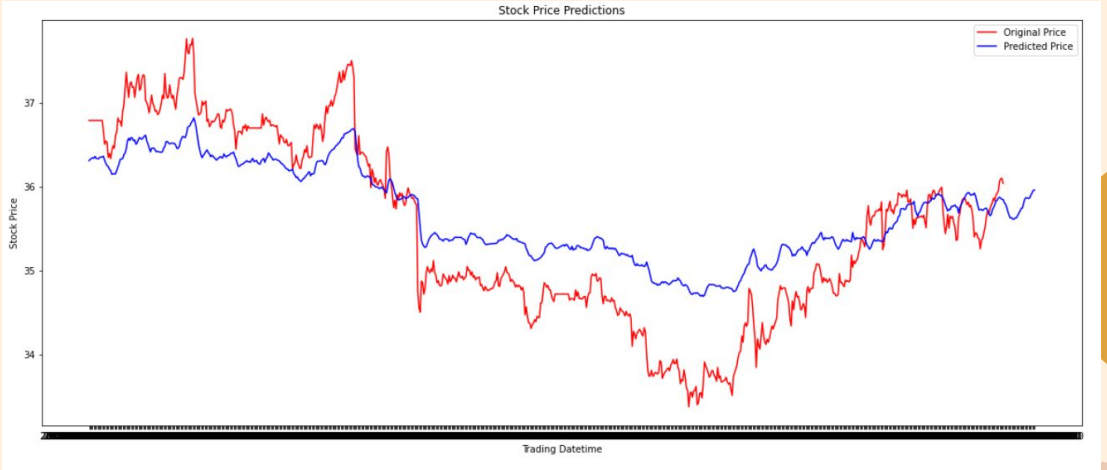
Predicted



Deep Learning
LSTM

	Close	Comp	Predicted_Close
2021-01-01 09:30:00	36.790001	0.265300	36.310955
2021-01-01 10:00:00	36.790001	0.245100	36.325344
2021-01-01 10:30:00	36.790001	0.257640	36.337463
2021-01-01 11:00:00	36.790001	0.113250	36.344875
2021-01-01 11:30:00	36.790001	0.501367	36.339699
...	...	...	...
2021-03-30 13:30:00	NaN	NaN	35.866657
2021-03-30 14:00:00	NaN	NaN	35.901096
2021-03-30 14:30:00	NaN	NaN	35.926502
2021-03-30 15:00:00	NaN	NaN	35.957150
2021-03-30 15:30:00	NaN	NaN	35.960175

- **RMSE: 0.590**
- Price prediction more accurate



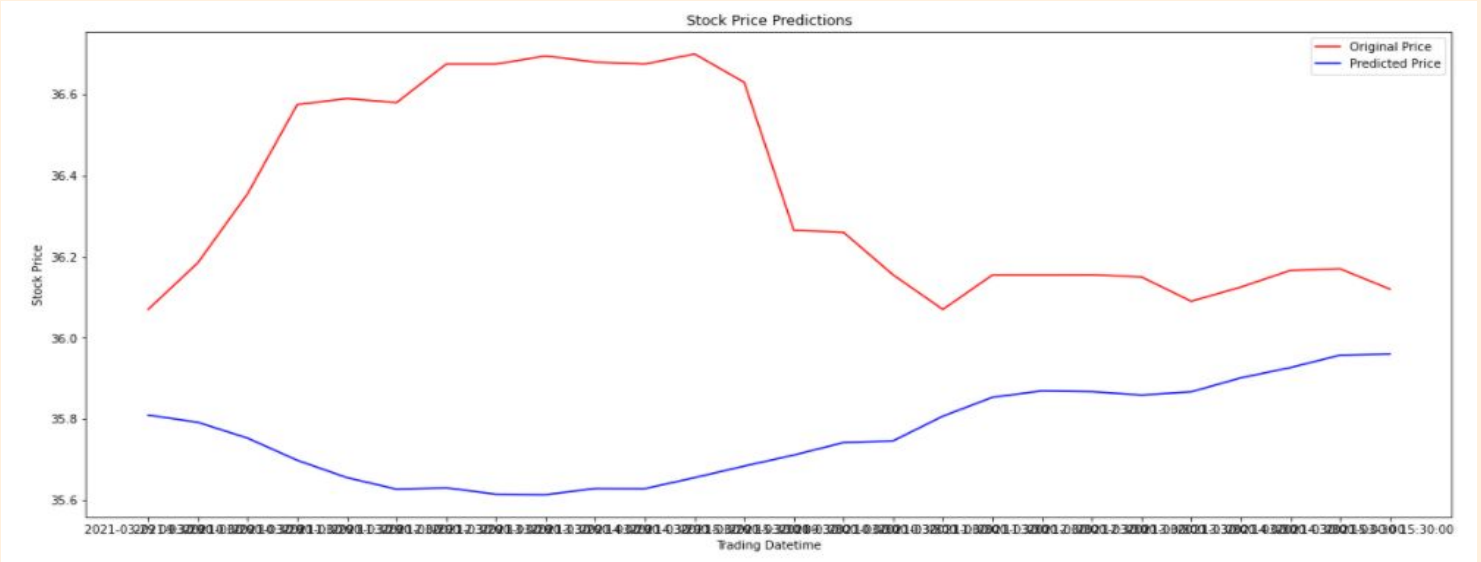
Models Result

	Random Forest	Linear Regression	Adaboost	SARIMAX	LSTM
RMSE	0.915	0.640	0.568	0.640	0.590

06

Final Results

Measuring the outcome



- RMSE of forecasted prices and actual prices: 0.680
- Price movement is only partially accurate (second half)
- Stock market is volatile
- Difficult to make a very accurate prediction

07

Final Solution (POC)

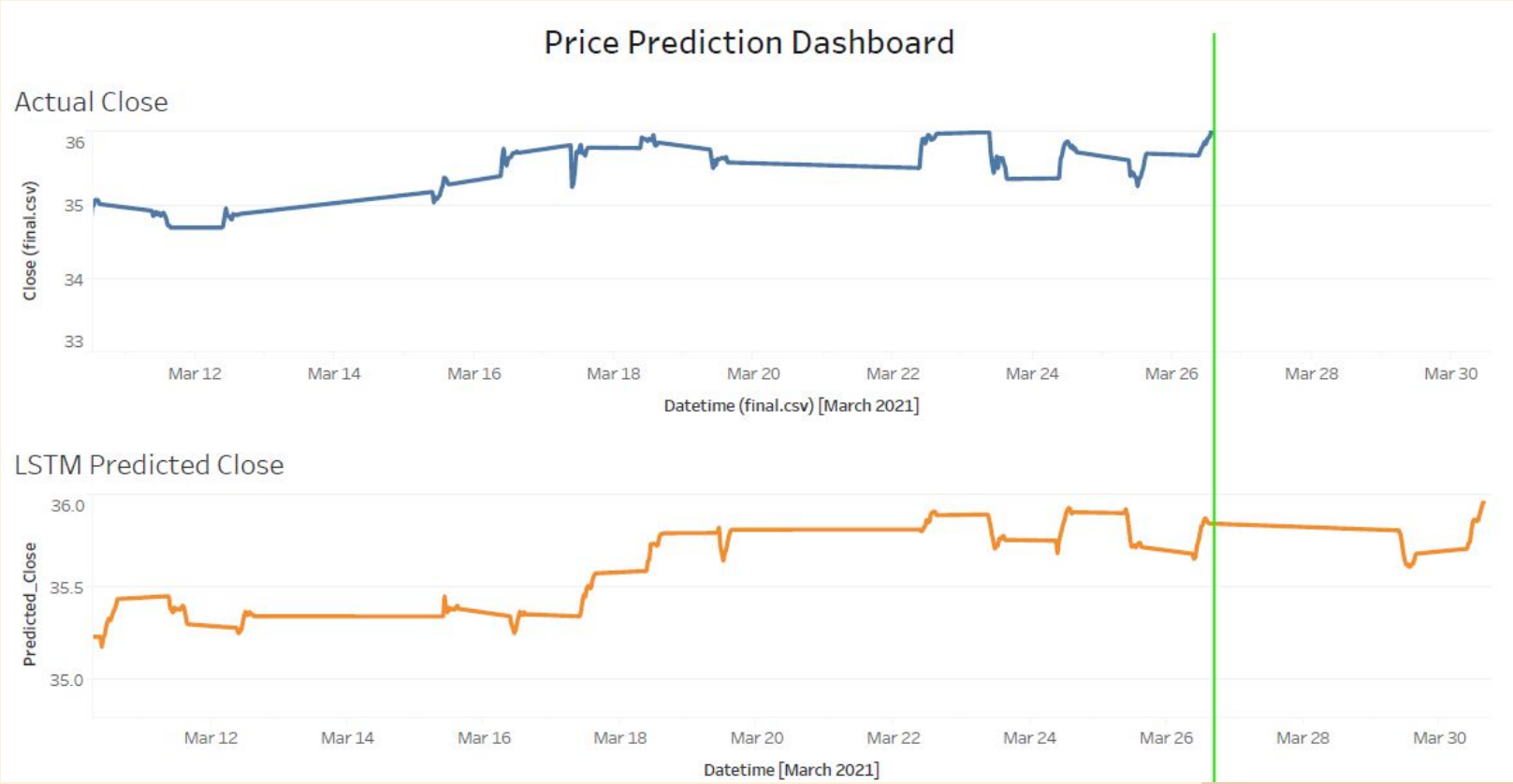
Competitive Advantage

- Allow retail investors to understand how **sentiments** of the general public affects the stock prices
- Provides **forecasting** of stock prices for up to 3 days
- **Unique selling point:**
  - Provides an **alternative metric (i.e. our proposed indicator)** for Retail Investors to time their **entry and exit** using a **benchmark**

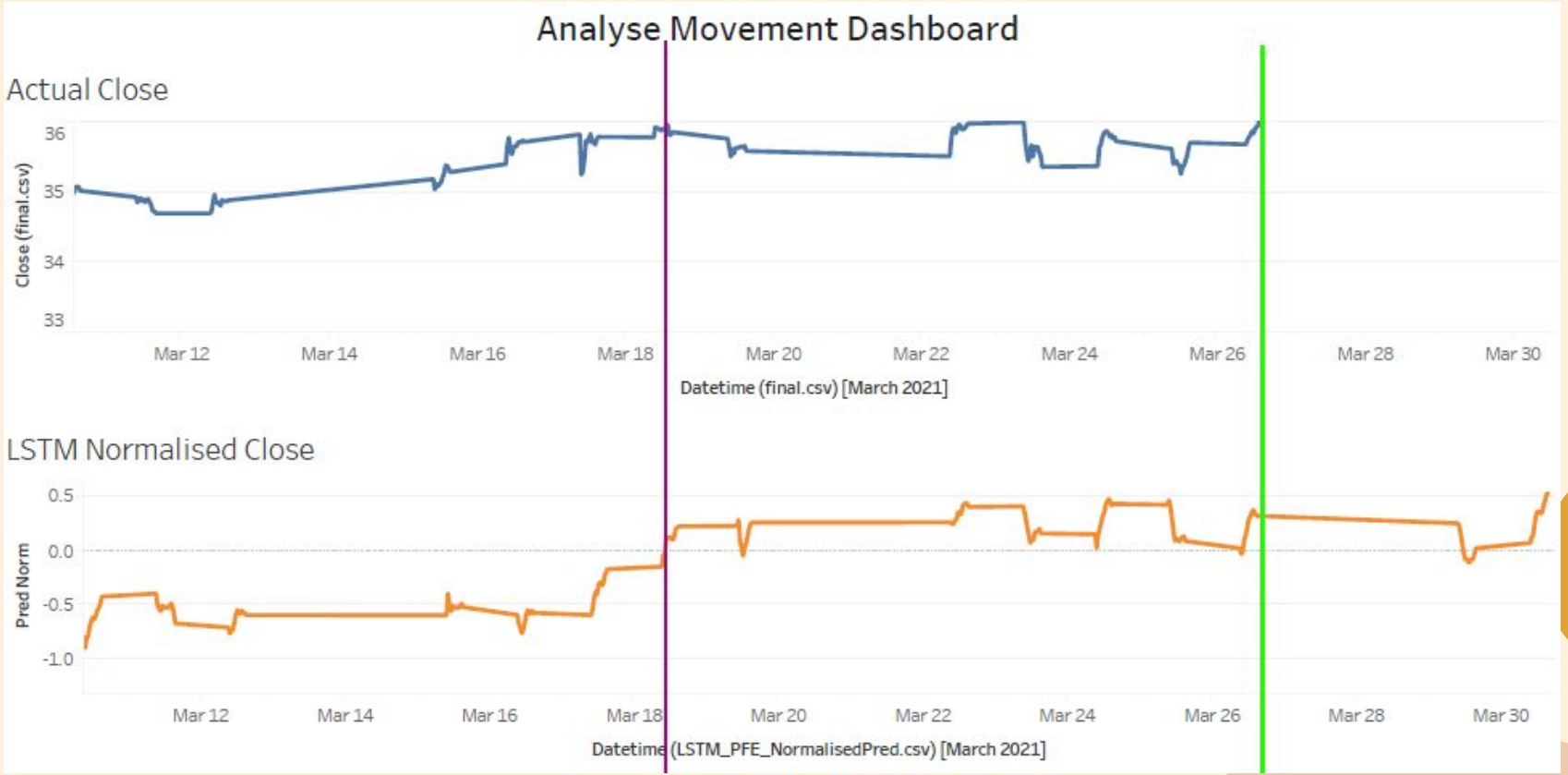
RSI vs Our Proposed Indicator

Relative Strength Index (RSI)	Our Proposed Indicator
● Upper bound: 100 ● Lower bound: 0 ● Stock prices cross the 70 mark ○ Stock is becoming overbought or undervalued 	● Upper bound and lower bound is not fixed, depending on the spread of the data ● Follows a standard normal distribution ○ Mean: 0 ○ SD: 1 ● We recommend using 0 mark ○ If stock prices crosses 0, with a positive gradient ○ Indicates that stock price will more likely have an increasing trend

Price Prediction & Forecast Dashboard



Price Movement Indicator



How to ensure successful implementation

Implementation

- Obtain sufficient data to train model
- Tune hyperparameters for LSTM to achieve lower RMSE

08

Conclusion

Potential Challenges & Mitigation Strategies

1) Scraping Tweets:
    a) Had the **issue of tweet limits** when using Tweepy API & Twint
    b) Used **SNScraper** instead

2) Number of days the prices can be forecasted highly dependent on the number of tweets scrapped
    a) **Popular** stocks: Can forecast a **longer period** (7 days).
    b) **Unpopular** stocks: Can forecast maximum **3 days**.
    c) Standardized our model to forecast up to 3 days

Assumptions

1) Users will **not** have **spelling errors** when inputting the stock names to scrape tweets and stock prices
2) Users will **want** to use the model with the **highest accuracy** of sentiment classifier and predictive model
3) Users will not make their investment **decision solely based on our solution**
4) Our models are used for **short term investments**
5) Users are allowed to **choose their own benchmark** on the normalised predicted close value to enter/exit instead of the recommended one.

Learning Points

1) Should scrape data with **small-cap stocks for a longer time period** as it **sped up the efficiency** of project process instead of large-cap stocks for a shorter time period
2) Basic machine learning regression models does not work well with time-series data (stocks data) which make predictions inaccurate
    a) **Time-series forecasting model** works better (LSTM, SARIMAX)
3) **No one-size-fits-all** python library that can **scrape** specific time-period data for stock prices

Moving Forward....

Future Work

- Test model on different stocks to ensure that it also provides predictions on large-cap stocks
- Validate our indicator by using longer period of time
- Improve the sentiment classifier model (current accuracy: 65%)

References

365 Data Science. (2021, January 24). What is a sarimax model? Retrieved from https://365datascience.com/tutorials/python-tutorials/sarimax/

Bakshi, C. (2020, June 09). Random forest regression. Retrieved from https://levelup.gitconnected.com/random-forest-regression-209c0f354c84

Chen, J. (2020, September 16). Autoregressive integrated moving average (arima). Retrieved from https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp

Difference between samples, time steps and features in neural network. (n.d.). Cross Validated. Retrieved from https://stats.stackexchange.com/questions/264546/difference-between-samples-time-steps-and-features-in-neural-network#:~:text=TimeSteps%20are%20ticks%20of%20time,a%20second%20for%20signal%20processing

Epoch. (2019, May 17). DeepAI. Retrieved from https://deepai.org/machine-learning-glossary-and-terms/epoch

Fernando, J. (2020, November 17). Relative strength index (rsi). Retrieved April 01, 2021, from https://www.investopedia.com/terms/r/rsi.asp

Grace, K. (2008, December 8). Assessing the fit of regression models. (2020, January 16). The Analysis Factor. https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/

Hong, J. (2020, May 21). 3 types of sequence prediction problems. Retrieved from https://jinglescode.github.io/2020/05/21/three-types-sequence-prediction-problems/

Hutto, C., & Gilbert, E. (n.d.). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Retrieved from http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf

Jones, R. (2020, June 3). Sentiment Analysis with the flair NLP library. Medium. Retrieved from https://rileymjones.medium.com/sentiment-anaylsis-with-the-flair-nlp-library-cfe830bfd0f4

References

Kaggle. (2017, September 13). Sentiment140 dataset with 1.6 million tweets. Retrieved from https://www.kaggle.com/kazanova/sentiment140

Khandelwal, R. (2020, November 11). Time series prediction using Sarimax. Retrieved from https://medium.datadriveninvestor.com/time-series-prediction-using-sarimax-a6604f258c56

Mayank, M. (2020, November 14). Guide to custom recurrent modeling in Keras. Medium. Retrieved from https://towardsdatascience.com/guide-to-custom-recurrent-modeling-in-keras-29027e3f8465

Packt. (n.d.). Computing regression accuracy. Retrieved from https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781789808452/1/ch01lvl1sec21/computing-regression-accuracy

Papers with code. (n.d.) Tanh activation explained. The latest in Machine Learning | Papers With Code. Retrieved from https://paperswithcode.com/method/tanh-activation

Phi, M. (2020, June 28). Illustrated guide to recurrent neural networks. Retrieved from https://towardsdatascience.com/illustrated-guide-to-recurrent-neural-networks-79e5eb8049c9

Shah, P. (2020, November 06). My absolute go-to for sentiment analysis - textblob. Retrieved from https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524

Shahul, ES. (2021, March 19). Sentiment analysis in Python: TextBlob vs Vader sentiment vs flair vs building it from scratch. neptune.ai. https://neptune.ai/blog/sentiment-analysis-python-textblob-vs-vader-vs-flair

Sklearn AdaBoost. (n.d.). Sklearn.ensemble.adaboostregressor¶. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostRegressor.html

References

Sklearn LinearRegression. (n.d.). Sklearn.linear_model.LinearRegression¶. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

Smith, T. (n.d.). Pmdarima.arima.auto_arima¶. Retrieved from https://alkaline-ml.com/pmdarima/modules/generated/pmdarima.arima.auto_arima.html

Tigerbrokers. (n.d.). Tiger brokers. Retrieved from https://www.tigerbrokers.com.sg/

Twint. (n.d.). PyPI. Retrieved from https://pypi.org/project/twint/

What is batch size in neural network? (n.d.). Cross Validated. Retrieved from https://stats.stackexchange.com/questions/153531/what-is-batch-size-in-neural-network

What is the definition of LSTM timestep? (n.d.) Retrieved from https://www.quora.com/What-is-the-definition-of-LSTM-timestep

Wu, S. (2020, June 14). What are the best metrics to evaluate your regression model? Medium. https://towardsdatascience.com/what-are-the-best-metrics-to-evaluate-your-regression-model-418ca481755b

Zhao, H., Sinha, A., Huang, Z., & Deng, S. (2018, September 19). Can social media sentiment affect stock market performance? Retrieved from https://blogs.lse.ac.uk/businessreview/2018/09/19/microblogging-sentiment-and-stock-returns/

THANK YOU!