
Stock Prediction Using Twitter Sentiment Analysis



ByteRabbit



ByteRabbit



UNIVERSITY OF SAN FRANCISCO
CHANGE THE WORLD FROM HERE



Bing



Cynthia



Shumin



ChatGPT VS. Microsoft Stock Prices



chatgpt microsoft stock price



News



Shopping



Images



Videos



Maps



Books

About 38,400,000 results (0.43 seconds)



Forbes

<https://www.forbes.com/sites/qai/2023/03/16/m...>

Microsoft Stock Rallies As Next Generation ChatGPT-4 Is ...

4 days ago — Microsoft is a major equity partner of OpenAI, the company who has created ChatGPT, and **their stock price has gained 6.5% so far this week.**



The Coin Republic

<https://www.thecoinrepublic.com/2023/02/11/micr...>

Microsoft Stock Prediction: Will MSFT Reach The \$10 Trilli

Feb 11, 2023 — MSFT stock price is currently on a halt and can see a new rise soon
Prediction: Will MSFT Reach The **\$10 Trillion** Mark With ChatGpt?



Business Insider

<https://markets.businessinsider.com/News/stocks>

Microsoft Stock Price Could **Soar 20%** on ChatGPT Bet

Jan 23, 2023 — Microsoft's huge bet on ChatGPT is a 'game changing AI investment' that could boost the stock by **20%**, Wedbush says.



Investors.com

<https://www.investors.com/news/technology/ai-st...>

AI Stocks: Microsoft Stock Rises On Artificial Intelligence Plans

Feb 7, 2023 — OpenAI is the organization behind text generator **ChatGPT** and image ... **lives rates Microsoft stock** as outperform with a 12-month price target ...



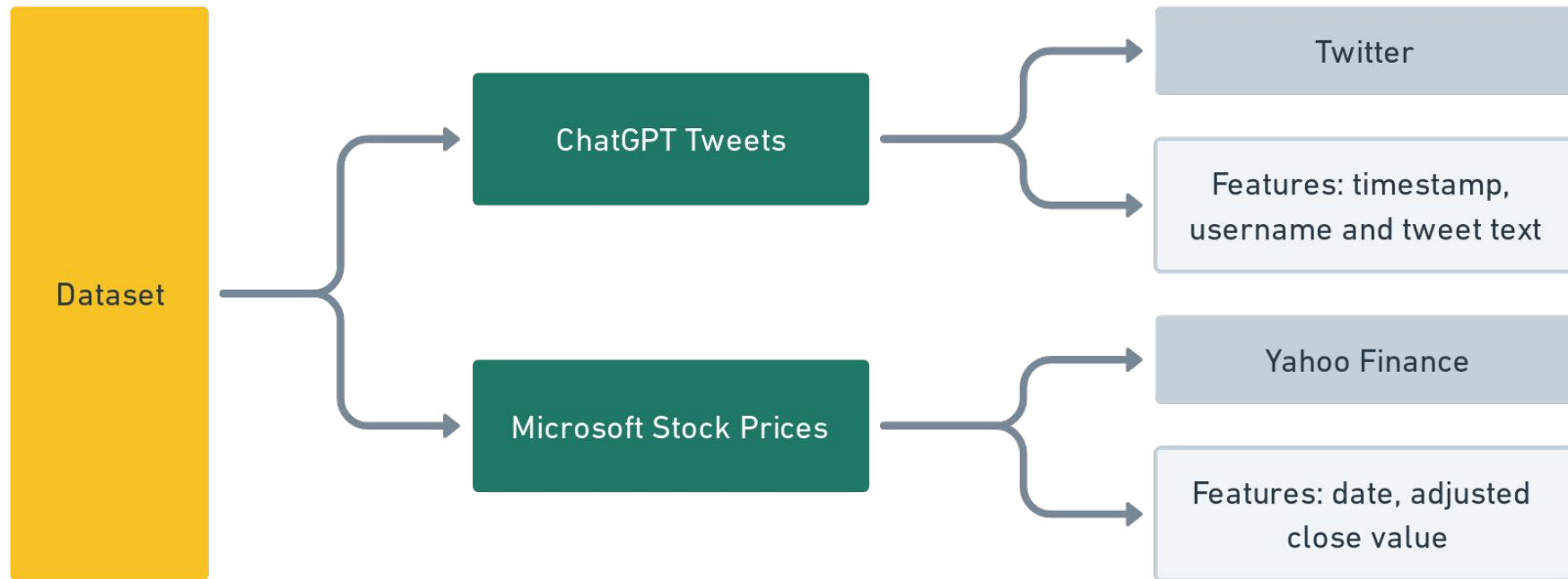
The Motley Fool

<https://www.fool.com/investing/2023/02/16/why...>

Why ChatGPT Caused Microsoft Stock to Stumble Today

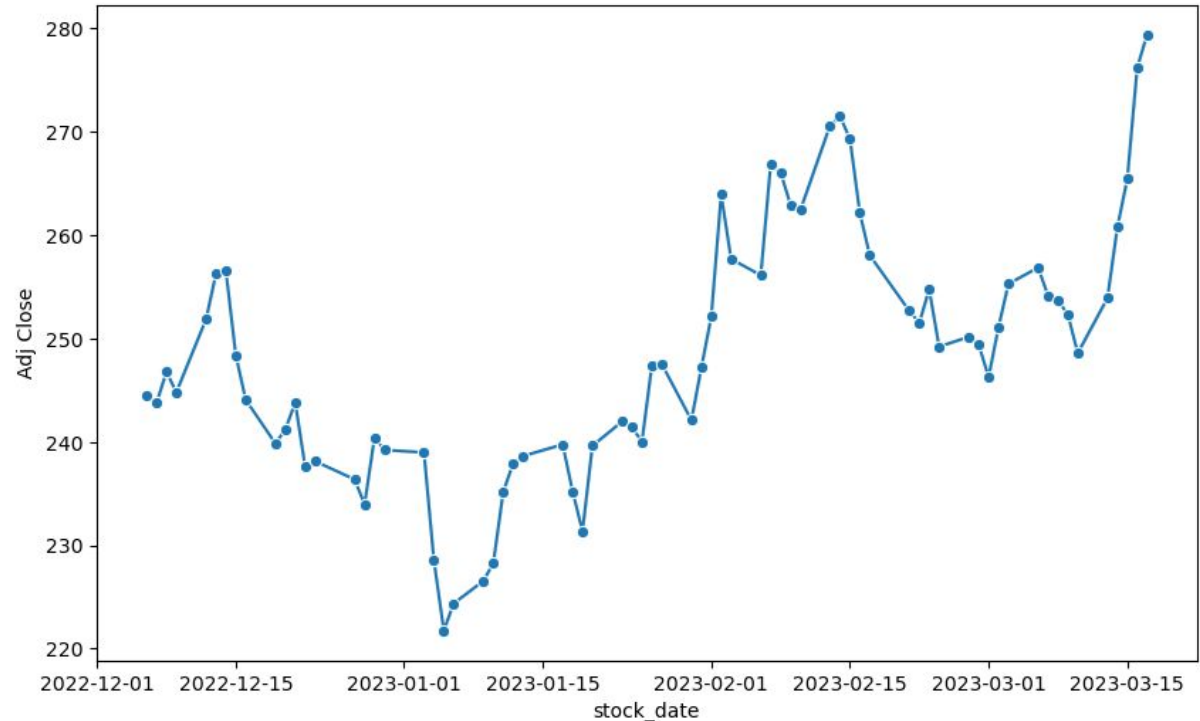
Feb 16, 2023 — Shares of Microsoft (**MSFT -2.58%**) slipped in early trading on the Nasdaq Thursday **falling 2%** through 9:45 a.m. ET. Multiple reports are now ...

Data



Data - Microsoft Stock Price

- **Source**
 - [Yahoo Finance](#)
- **Time Range**
 - From 12/01/2022
 - To 03/20/2023



Data - ChatGPT Tweets

- **Source**

- [Twitter](#)

- **Method/Library**

- snsrape

- **Time Range**

- 12/01/2022 - 03/19/2023

Unnamed: 0		Date	User	Tweet
0	0	2023-03-19 23:59:54+00:00	Julia	Oh my god, the South Park episode about ChatGP..
1	1	2023-03-19 23:59:36+00:00	Sej	Bing powered by #chatgpt is a new way of learn...
2	2	2023-03-19 23:59:04+00:00	theha	yall need to start using chatGPT to improve yo...
3	3	2023-03-19 23:58:52+00:00	jamie	Chat GPT and Twitter Blue stops me from saying...
4	4	2023-03-19 23:58:49+00:00	fifthoe	chat gpt is insane it's gonna change how ppl d...
...
305455	4102	2022-12-01 00:14:08+00:00	Suave	#ChatGPT from @OpenAI looks extremely promisin...
305456	4103	2022-12-01 00:08:07+00:00	Timoth	Me: Henry is a fish that got eaten by a cat ca...
305457	4104	2022-12-01 00:05:27+00:00	Ted_Und	Long day of meetings, then you open Twitter an...
305458	4105	2022-12-01 00:02:43+00:00	Timot	Me: Can you give me a sensationalist headline ...
305459	4106	2022-12-01 00:00:28+00:00	b	chatgpt is really cool but man saying shit to ...

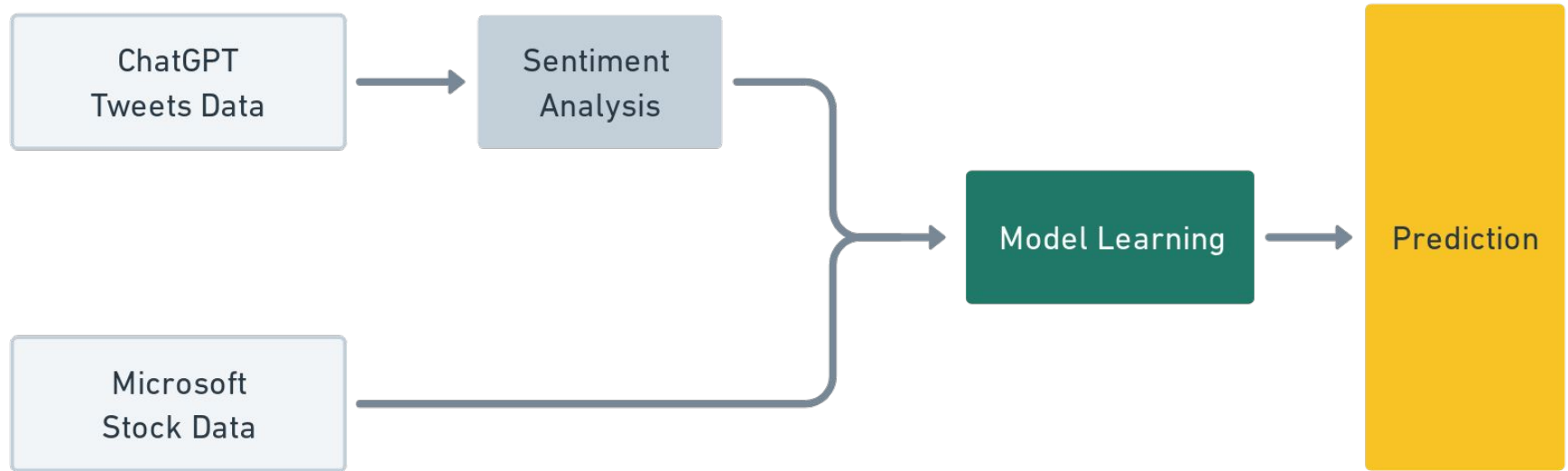
305460 rows x 4 columns

Data - Sentimental Analysis Results and Stock Price

	Clean_Tweet	Negative	Neutral	Positive	Sentiment	tweet_date	Adj Close
0	Oh my god, the South Park episode about ChatGP...	0.002614	0.010037	0.987349	1	2023-03-19	272.230011
1	Bing powered by #chatgpt is a new way of learn...	0.001523	0.061577	0.936899	1	2023-03-19	272.230011
2	yall need to start using chatGPT to improve yo...	0.044139	0.216546	0.739315	1	2023-03-19	272.230011
3	Chat GPT and Twitter Blue stops me from saying...	0.103467	0.325155	0.571378	1	2023-03-19	272.230011
4	chat gpt is insane it's gonna change how ppl d...	0.116454	0.427531	0.456015	1	2023-03-19	272.230011
...
296378	#ChatGPT from @user looks extremely promising!...	0.001137	0.006888	0.991975	1	2022-12-01	254.382843
296379	Me: Henry is a fish that got eaten by a cat ca...	0.310093	0.618012	0.071895	0	2022-12-01	254.382843
296380	Long day of meetings, then you open Twitter an...	0.661903	0.283301	0.054796	-1	2022-12-01	254.382843
296381	Me: Can you give me a sensationalist headline ...	0.245476	0.606909	0.147615	0	2022-12-01	254.382843
296382	chatgpt is really cool but man saying shit to ...	0.643117	0.273827	0.083056	-1	2022-12-01	254.382843

296383 rows x 7 columns

Data






Sentiment Analysis - NLP Models

- **NLTK** (Natural Language Toolkit)
 - Pro: most well-know and full NLP library.
 - Con: only focus on plain texts without analyzing of contextual semantics.
- **BERT** (Pre-training of Deep Bidirectional Transformers for Language Understanding)
 - Pro: able to analyze contextual semantics.
 - Con: very compute-intensive, trouble with scaling.
- **RoBERTa** (A Robustly Optimized BERT Pre-training Approach)
 - Pro: trained on a larger dataset than BERT therefore more robust.
 - Con: longer training time.

Sentiment Analysis - RoBERTa


cardiffnlp/twitter-roberta-base-sentiment-latest
like 134

Text Classification
PyTorch
TensorFlow
Transformers
tweet_eval
English
roberta
arxiv:2202.03829

Model card
Files
Community 9

Edit model card

Twitter-roBERTa-base for Sentiment Analysis - UPDATED (2022)


This is a RoBERTa-base model trained on ~124M tweets from January 2018 to December 2021, and finetuned for sentiment analysis with the TweetEval benchmark. The original Twitter-based RoBERTa model can be found [here](#) and the original reference paper is [TweetEval](#). This model is suitable for English.

- Reference Paper: [TimeLMs paper](#).
- Git Repo: [TimeLMs official repository](#).

Labels: 0 -> Negative; 1 -> Neutral; 2 -> Positive

This sentiment analysis model has been integrated into [TweetNLP](#). You can access the demo [here](#).

Downloads last month
1,742,716



Hosted inference API

Text Classification Examples

Hey @Bing and @Cynthia, check out ChatGPT at <https://chat.openai.com!> 😊

Compute

Computation time on Intel Xeon 3rd Gen Scalable cpu: 12.860 s

positive	0.913
neutral	0.085
negative	0.002

JSON Output

Maximize

Sentimental Analysis - Step 1

- tweet = "Hey @Bing and @Cynthia, check out ChatGPT at <https://chat.openai.com>! 😊"

"Hey, check out ChatGPT !"

"Hey check out ChatGPT"

@Bing @Cynthia

@user @user



<https://chat.openai.com>

http



Sentimental Analysis - Step 2

```
tweet = "Hey @user and @user check out ChatGPT at http 😊"
```

- **Convert the tweet into pytorch tensor**
- {
 66, 20013, 534, 10311, 23, 2054, 17841, 27969, 2]]),

 'attention_mask': tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]])
}



Sentimental Analysis

'input_ids', 'attention_mask'

- **Pass the tensor to the model to do the sentiment analysis**
- `SequenceClassifierOutput(`
 `loss=None,`
 `logits=tensor([[-3.7680, 1.2519, 2.6841]],`
 `grad_fn=<AddmmBackward0>), hidden_states=None,`
 `attentions=None`
 `)`

Sentimental Analysis

```
tensor = [-3.7680, 1.2519, 2.6841]
```

- **Convert the tensor into probability using softmax and add labels**

- Negative: 0.00127156,

Neutral: 0.19250846

Positive: 0.80622

Positive

tweet = "Hey @Bing and
@Cynthia, check out ChatGPT
at <https://chat.openai.com!>



Sentimental Analysis

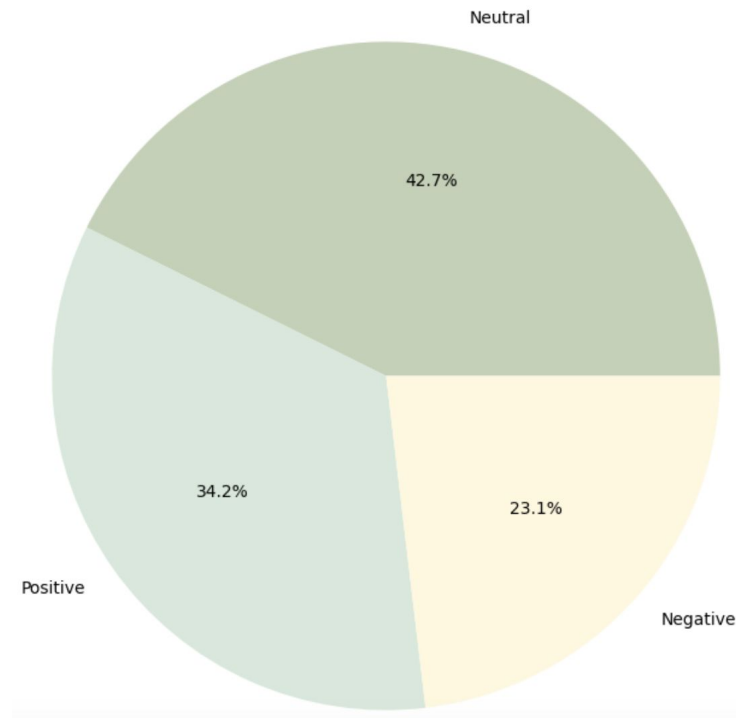
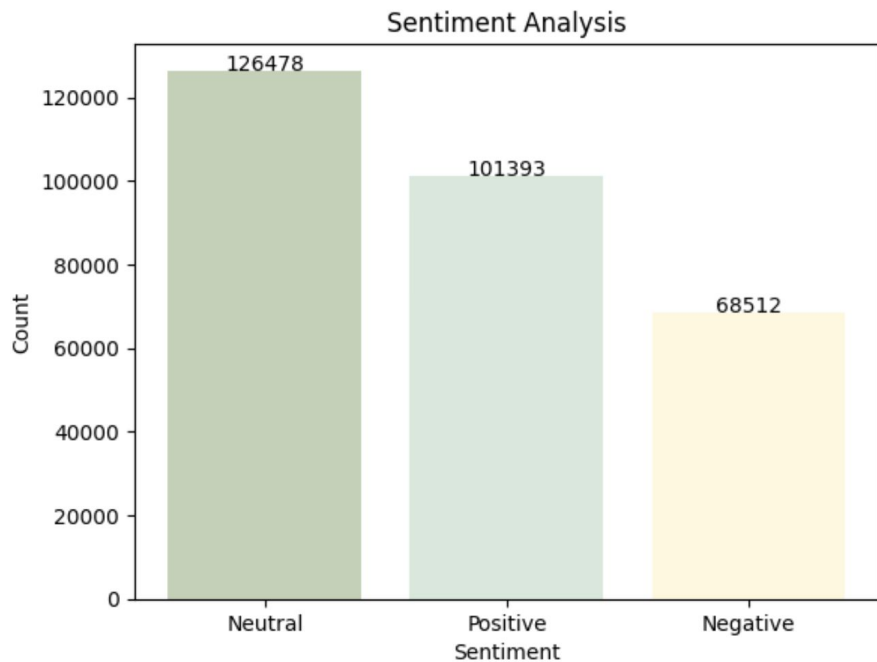
Unnamed: 0	Date	Tweet	Clean_Tweet	Negative	Neutral	Positive	Sentiment
0	2023-03-19 23:59:54+00:00	Oh my god, the South Park episode about ChatGP..	Oh my god, the South Park episode about ChatGP..	0.002614	0.010037	0.987349	1
1	2023-03-19 23:59:36+00:00	Bing powered by #chatgpt is a new way of learn...	Bing powered by #chatgpt is a new way of learn...	0.001523	0.061577	0.936899	1
2	2023-03-19 23:59:04+00:00	yall need to start using chatGPT to improve yo...	yall need to start using chatGPT to improve yo...	0.044139	0.216546	0.739315	1
3	2023-03-19 23:58:52+00:00	Chat GPT and Twitter Blue stops me from saying...	Chat GPT and Twitter Blue stops me from saying...	0.103467	0.325155	0.571378	1
4	2023-03-19 23:58:49+00:00	Chat GPT and Twitter Blue stops me from saying...	Chat GPT and Twitter Blue stops me from saying...	0.116454	0.427531	0.456015	1
...
296378	2022-12-01 00:14:08+00:00	#ChatGPT from @OpenAI looks extremely promisin...	#ChatGPT from @user looks extremely promising!...	0.001137	0.006888	0.991975	1
296379	2022-12-01 00:08:07+00:00	Me: Henry is a fish that got eaten by a cat ca...	Me: Henry is a fish that got eaten by a cat ca...	0.310093	0.618012	0.071895	0
296380	2022-12-01 00:05:27+00:00	Long day of meetings, then you open Twitter an...	Long day of meetings, then you open Twitter an...	0.661903	0.283301	0.054796	-1
296381	2022-12-01 00:02:43+00:00	Me: Can you give me a sensationalist headline ...	Me: Can you give me a sensationalist headline ...	0.245476	0.606909	0.147615	0
296382	2022-12-01 00:00:28+00:00	chatgpt is really cool but man saying shit to ...	chatgpt is really cool but man saying shit to ...	0.643117	0.273827	0.083056	-1

8 hours !!!



296383 rows x 8 columns

Sentimental Analysis



Why a Neural Network Model?

- Neural networks have been considered to be a very effective learning algorithm for decoding nonlinear time series data, and financial markets often follow nonlinear trends.

Layers

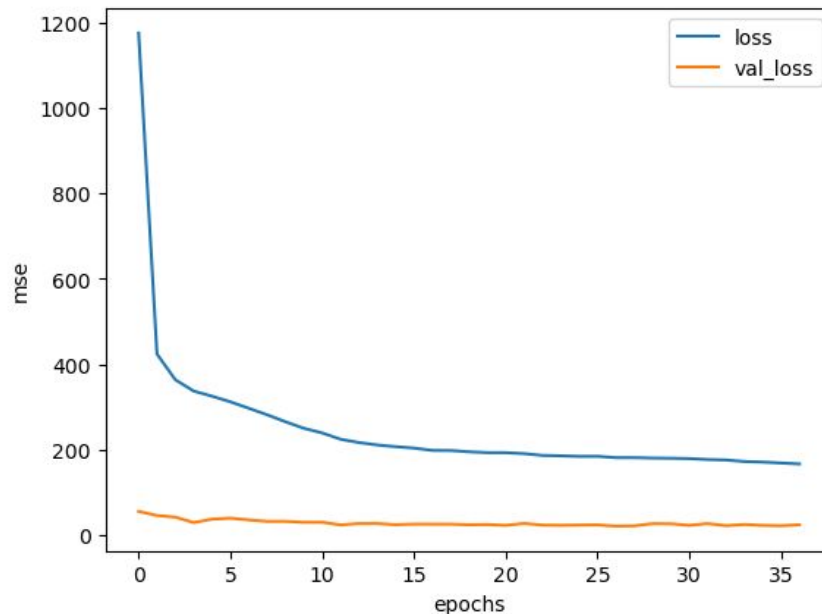
Input: Positive, Neutral and Negative, and the stock price

Hidden: We follow suggestions to avoid underfitting and overfitting

Output: stock price of the next day.

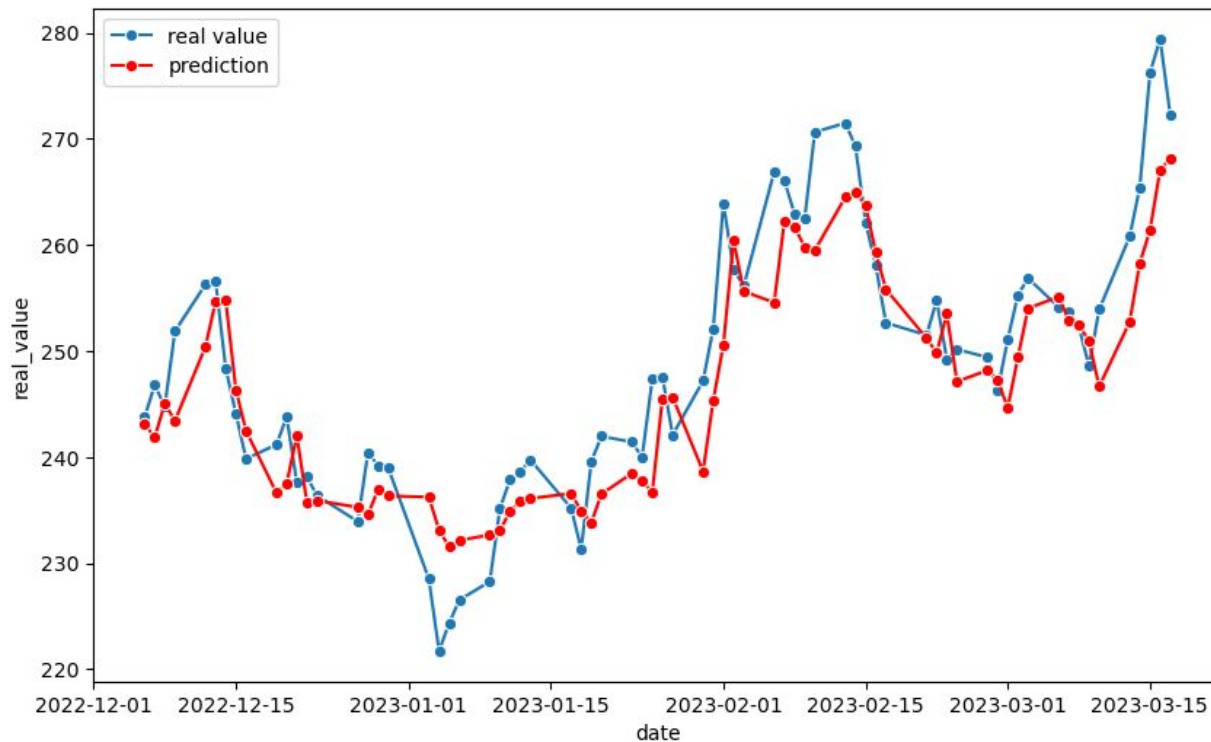
Loss Function Plot

- **Loss Function:**
 - MSE
- **Dropout Layers:**
 - 50%
- **Early Stopping:**
 - Epoch = 50
 - Patience = 10
 - It stopped at epoch = 37



Results

RMSE = 4.89



Thank you!



ByteRabbit