



# Novel Approaches to Sentiment Analysis for Stock Prediction

AGRIM RAWAT  
JEWEL BENNY  
KNV KARTHIKEYA  
MUDIT GAUR



# Introduction

- In the world of finance, stock market forecasts have been a crucial and contentious topic.
- Many people believe and think that it's nothing but a gamble and the stocks and their valuation can't be predicted.
- Through fundamental analysis, technical analysis, and, more recently, machine learning, others have tried to forecast the market.
- A technique like machine learning may be well suited for such an application given that the stock market is primarily a quantitative one.



# Motivation and Aim

**01**

The current generation of machine learning models has concentrated on sentiment or technical analysis as a single feature.

**02**

However, as the stock market also largely relies on investor emotions and fundamental firm data, which cannot be represented by a straightforward numerical indicator.

**03**

Hence a model is proposed where the model takes in both stock financial data and news information, which we encode into a fixed-length vector.

**04**

We saw a lot of promise in the related work utilising sentiment to forecast stock movement. We were worried that a single mood value wouldn't adequately convey the nuanced nature of the news and corporate information, though.



# Proposed Model

- In order to more effectively process news data, a model encodes the data into real vectors before feeding the entire vector into a classification model.
- By merging the trade data for the stock and the data from the daily news that is relevant to the ticker, one may predict the stock's movement the next day.
- We divide our model into 2 subparts, one being representing the news we from our sources as a fixed-length real scalar or vector and then combining this information with trading information, technical indicators, and macro data, to make the prediction.
- The stock movement prediction model needs the text representation model to transform news sentences and excerpts into real-value fixed-length scalar or vector representations.



# Data Used for Prediction

- Headlines and Snippets related to the Stock market scrapped from Google news and other sources.
- Daily extraction of trading-related data, such as price and volume, from each stock ticker using the Yahoo Finance API.
- Computing technical indicators including CCI, RSI and EVM from trading data.

# Workflow and Timeline



# Stock Movement Prediction and Various Classifiers Used

**Logistic Regression** is used as a baseline to map the features to stock movement.

**Random Forest** is used for classification and regression, which builds a large number of decision trees during training time and produces the class that represents the majority vote of the individual leaves.

**Support Vector Machines** are mentioned in previous research to be effective in stock prediction applications. We can regulate the model by using different costs.

**Convolutional Neural Networks** have been widely used for image processing. We thought it might be effective to do convolutions over the sentence embeddings because of their structure.

**Recurrent Neural Networks** are proven to be effective in dealing with sequential data, with the output being dependent on the previous computations. The hidden layer captures information about what has been calculated so far.



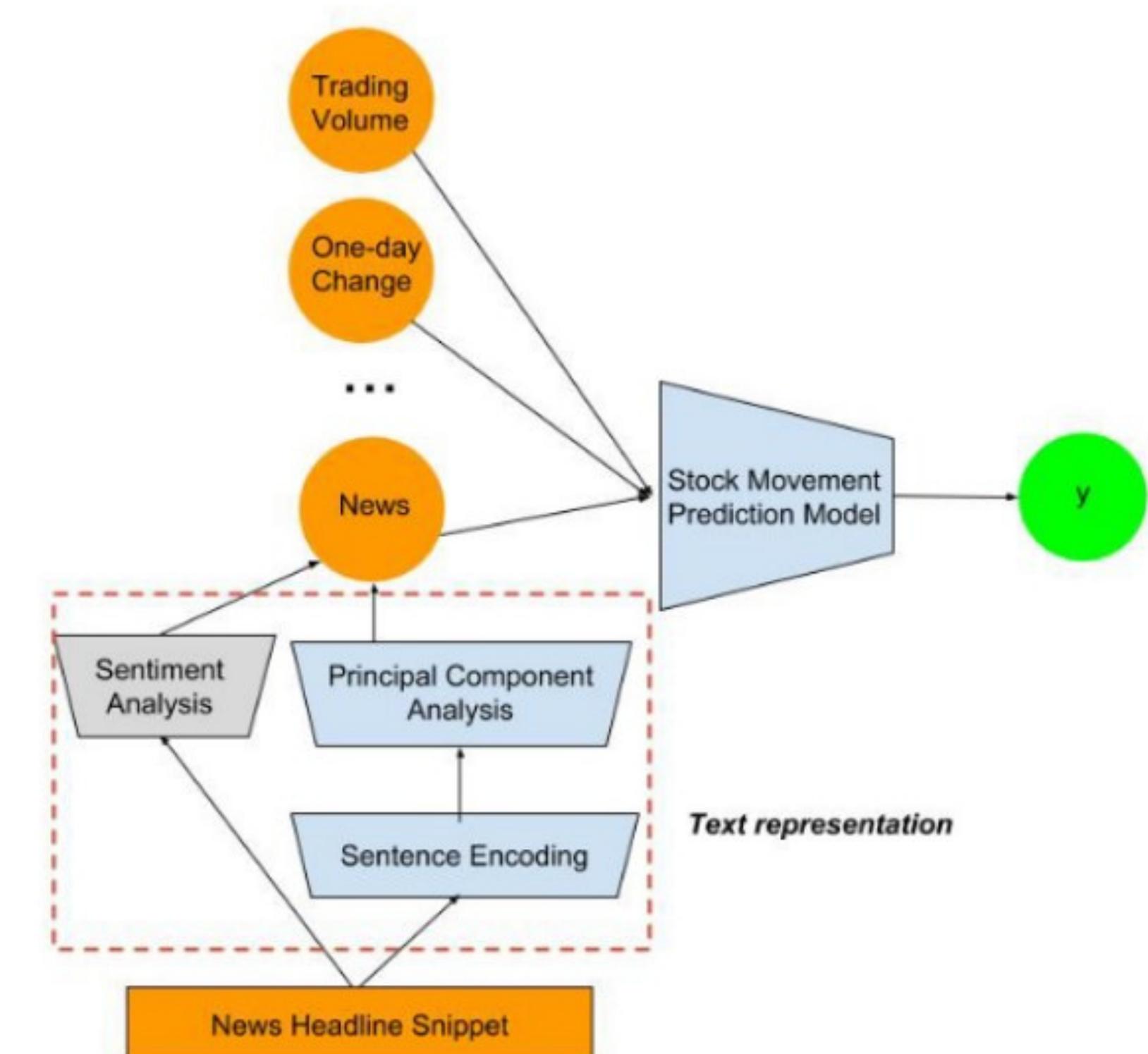
# Work Done Currently

- Made python script for web scraping from Google news using Beautiful Soup library.
- The following information is scraped from news articles -
  - Date or time when news article was released
  - Ticker with company name
  - News article URL
  - News article headline
  - News article source
  - News article snippets



TEAM 1

# The Model Design





# Outputs

Date	Ticke Url	headline	source	snippet
0 4 days ago	AMZ <a href="https://inc42.com/buzz/c">https://inc42.com/buzz/c</a>	Disney+ Hotstarâ€™s ARPU Declines Due To Lower Per-Subscribers	inc42	It had a 50% market share in Subscription Video on Demand
1 5 days ago	AMZ <a href="https://www.prnewswire.com">https://www.prnewswire.com</a>	iRobot Reports Third-Quarter 2022 Financial Results	PR Newswire	The company's inventory balance was \$419 million, or 191
2 2 days ago	AMZ <a href="https://capital.com/robinhood">https://capital.com/robinhood</a>	Robinhood Shareholders   Who Owns The Most Shares of	Capital.com	As of 1 October 2022, the venture capital firm had increased
3 4 days ago	AMZ <a href="https://www.nasdaq.com">https://www.nasdaq.com</a>	Disney (DIS) Q4 Earnings Miss Estimates, Revenues Rise	Nasdaq	Disney+, as of Oct 1, 2022, had 164.2 million paid subscribers
4 3 days ago	AMZ <a href="https://www.marketscreener.com">https://www.marketscreener.com</a>	IROBOT CORP Management's Discussion and Analysis of	MarketScreener	... statements concerning our pending acquisition by Amazon,
5 3 days ago	AMZ <a href="https://www.fool.com">https://www.fool.com</a>	Beyond Meat, Inc. (BYND) Q3 2022 Earnings Call Transcript	The Motley Fool	... 10-Q for the quarter ended October 1, 2022, to be filed
6 4 days ago	AMZ <a href="https://www.marketscreener.com">https://www.marketscreener.com</a>	Transcript : Celsius Holdings, Inc., Q3 2022 Earnings Call, No	MarketScreener	Amazon sales, another nontrack channel for the third quarter
7 4 days ago	AMZ <a href="https://www.yahoo.com">https://www.yahoo.com</a>	Capri Holdings Cuts 2023 Outlook After Q2 Results Face Imp	Yahoo Finance	Shares for Capri Holdings were down more than 4% by
8 4 days ago	AMZ <a href="https://www.tipranks.com">https://www.tipranks.com</a>	Microsoft Accused By EU of Unfair Practices	TipRanks	... on October 1, 2022, and certain other practices. It's worth
9 4 days ago	AMZ <a href="https://www.marketscreener.com">https://www.marketscreener.com</a>	MASIMO CORP Management's Discussion and Analysis of	Market Screener	approximately \$9.7 million . Provision for Income Taxes. Our
0 4 days ago	AMZ <a href="https://finance.yahoo.com">https://finance.yahoo.com</a>	ENTRÃ‰%E RESOURCES ANNOUNCES THIRD QUARTER 2022	Yahoo Finance	... level at 298 metres below ground level as at October 2,
1 4 days ago	AMZ <a href="https://www.marketscreener.com">https://www.marketscreener.com</a>	Transcript : Celsius Holdings, Inc., Q3 2022 Earnings Call, No	MarketScreener	Amazon sales, another nontrack channel for the third quarter
2 1 month ago	AMZ <a href="https://www.livemint.com">https://www.livemint.com</a>	Amazon app quiz October 2, 2022: Check answers to win to Mint	Mint	Amazon app quiz October 2, 2022: Check answers to win
3 1 month ago	AMZ <a href="https://seekingalpha.com">https://seekingalpha.com</a>	Get Paid To Buy Amazon 10% Lower (NASDAQ:AMZN)	Seeking Alpha	Amazon.com, Inc. (NASDAQ:AMZN) is one of the stocks that
4 1 month ago	AMZ <a href="https://www.fool.com">https://www.fool.com</a>	ir Will Amazon Be Worth More Than Apple by 2	The Motley Fool	Motley Fool Returns. Motley Fool Stock Advisor. Market-
5 1 month ago	AMZ <a href="https://seekingalpha.com">https://seekingalpha.com</a>	Amazon's pricey 'Rings of Power' dominates streaming	Seeking Alpha	Amazon Prime Video (NASDAQ:AMZN) spent heavily on its
6 1 month ago	AMZ <a href="https://in.investing.com">https://in.investing.com</a>	Google Pixel Watch was available for pre-order on Amazon	Investing.com	IANS; Stock Market News; 2022-10-02 10:45. Google Pixel
7 1 day ago	AMZ <a href="https://www.mitchellrep.com">https://www.mitchellrep.com</a>	Nov. 7 Regular Meeting Minutes - Mitchell Republic   News	Mitchell Daily Republic	PAYROLL OCTOBER 2, 2022 â€“ OCTOBER 15, 2022: City
8 1 month ago	AMZ <a href="https://www.businessstoc">https://www.businessstoc</a>	ONDC wants to give the small seller a platform beyond	Business Today	ONDC is an ambitious initiative by the government to
9 4 weeks ago	AMZ <a href="https://www.wordonfire.com">https://www.wordonfire.com</a>	Gospel Reflection - Monday, October 17,	Word on Fire	Bishop Barron is a #1 Amazon bestselling author and has
0 2 weeks ago	AMZ <a href="https://finance.yahoo.com">https://finance.yahoo.com</a>	JOHN HANCOCK PREMIUM DIVIDEND FUND NOTICE TO	Yahoo Finance	... declared on October 3, 2022, and payable on October 31,
1 1 month ago	AMZ <a href="https://www.fool.com">https://www.fool.com</a>	ir Down \$600 Billion, Is Amazon Stock a	The Motley Fool	Amazon (AMZN 4.30%) has seen its stock price tumble in 2022
2 3 weeks ago	AMZ <a href="https://www.team-bhp.com">https://www.team-bhp.com</a>	Mods on my brand-new VW Virtus GT: K&N air filter, LI	Team-BHP	Bulb Type: H8; Stock Wattage: 70W (total) ... part but from
3 1 month ago	AMZ <a href="https://zeenews.india.com">https://zeenews.india.com</a>	Amazon app quiz today, October 3, 2022: To win Rs 500,	Zee News	Here are all the five questions and its answer for October 3,
4 1 month ago	AMZ <a href="https://www.bloomberg.com">https://www.bloomberg.com</a>	European Stocks News Updates: October 3,	Bloomberg.com	European Stocks Gain as Energy Shares Rally on Oil, Yields
5 1 month ago	AMZ <a href="https://blog.ted.com">https://blog.ted.com</a>	Climate challenges and solutions: The talks from the TED	TED Blog	When and where: Monday, October 3, 2022, at the Barbican



# References

01

**Novel Approaches to Sentiment Analysis for Stock Prediction**  
**Chris Wang, Yilun Xu, Qingyang Wang – cs229stockprediction**

02

**Bollen et al. (2010) Twitter Mood Predicts the Stock Market. Journal of Computational Science.**

03

**Madge, S., Bhatt, S. (2015). Predicting Stock Price Direction using Support Vector Machines. Independent Work Report Spring**

04

**Beautiful Soup Documentation –**  
**<https://www.crummy.com/software/BeautifulSoup/bs4/doc/#>**



TEAM 1

# Thank You