Sentiment Analysis of Economic conditions based on Global News Reports

Course: BIA 660 C
Instructor: Jingyi Sun

Submitted by: Group 6
Mayank Raj
Akshit Walia
Kartik Ahluwalia

TABLE OF CONTENTS

INTRODUCTION

The state of the global economy has a significant impact on how people, groups, and nations make financial decisions. Yet, deciphering market sentiments from news headlines can be challenging, especially for those who are unfamiliar with economic ideas. This frequently results in hasty conclusions that are based more on feelings than on logic and research.

We describe a project on sentiment analysis of economic situations based on international news reports to address this problem. The objective of this project is to develop a sentiment analysis system for specific set of industries by utilizing the massive volume of online news stories. Our goal is to create a platform that can effectively gauge market sentiment and offer insightful information for making decisions.

We acknowledge that economics is a complicated discipline, and sentiment analysis in this area has peculiar difficulties. Yet, we think it is possible to draw insightful conclusions from news broadcasts and offer a useful resource for those making financial decisions if the correct methodology and approach are used. We anticipate that our work will advance the field of sentiment research and provide a workable approach to comprehending the prospects for the global economy.

LITERATURE REVIEW

This project involves knowledge and direction taken from six different research papers.

The first segment of papers is used to positively validate the feasibility of our project. This segment includes Kang (2017) and Da (2015).

Kang (2017) exudes the positive possibility that the proposed project is feasible in the first place. It proposes a method for predicting stock prices using financial news articles. The proposed method uses sentiment analysis and machine learning techniques to extract sentiment features from news articles and predict stock prices. The method was evaluated using a dataset of financial news articles and corresponding stock prices from the Korean stock market. The results show that the proposed method outperforms traditional time series models in terms of prediction accuracy.

Da (2015) revisits the relationship between investor sentiment and stock returns using a machine learning approach. The authors use sentiment scores derived from news articles and tweets to predict daily stock returns. They find that sentiment-based trading strategies significantly outperform traditional buy-and-hold strategies, and that sentiment is a stronger predictor of stock returns during periods of high uncertainty.

The second segment of the papers are used to study the extent of impact of the project itself.
This segment includes Chen (2016), Azevedo (2019) and Gholampour (2018).

Chen (2016) investigates the impact of news sentiment on stock market volatility using sentiment analysis of news articles. The authors construct a sentiment index based on the frequency of positive and negative words in news articles and find that news sentiment has a significant impact on stock market volatility. They also find evidence of asymmetric effects, with negative news having a greater impact on volatility than positive news.

Azevedo (2019) examines the relationship between news sentiment and economic policy uncertainty in Europe. The authors use sentiment analysis of news articles to construct a news sentiment index and find that news sentiment has a significant negative impact on economic policy uncertainty. They also find evidence of a feedback loop between news sentiment and economic policy uncertainty, with increased economic policy uncertainty leading to more negative news sentiment.

Gholampour (2018) investigates the usefulness of sentiment analysis of Twitter data in predicting financial market movements. The study collects a large dataset of tweets related to the financial market and applies various sentiment analysis techniques to classify the tweets as positive, negative, or neutral. The sentiment scores are then used to predict the movements of the S&P 500 and the Dow Jones Industrial Average. The results show that sentiment analysis can be an effective tool for predicting short-term movements in the financial market, especially when used in combination with other machine learning techniques. The study also finds that the use of Twitter data for sentiment analysis can provide valuable insights for investors and financial analysts. Overall, the findings of this study suggest that sentiment analysis of Twitter data has the potential to improve the accuracy of financial market predictions.

The final segment of our literature analysis includes a Giglio (2016) which connects the economic policy uncertainty with the stock market, which could be a tangible outcome or use case for the proposed project, with much greater impact in the world.

Giglio (2016) uses sentiment analysis of news articles to investigate the impact of economic policy uncertainty on the stock market. The authors construct an economic policy uncertainty index based on the frequency of words related to uncertainty in news articles and find that economic policy uncertainty has a significant negative impact on stock market returns. They also find that the impact of economic policy uncertainty varies across different sectors and regions.

DATA COLLECTION

The project utilizes **NEWSDATA.IO**, a platform that gives users access to structured news data from a variety of sources. Users of the platform can search and get news articles using the platform's API according to a variety of criteria, including date range, keyword, language, and geography. It categorizes articles into different topics, such as business, finance, politics, technology, and sports, making it easy for users to find relevant content.

Users can scrape articles using constraints like Date Published, Keyword, Category, and Country. The scraped data is stored in a data frame in the format. The dataset was created using different keywords and categories for various sectors such as banking, technology, energy and economy to get a diverse set of articles .

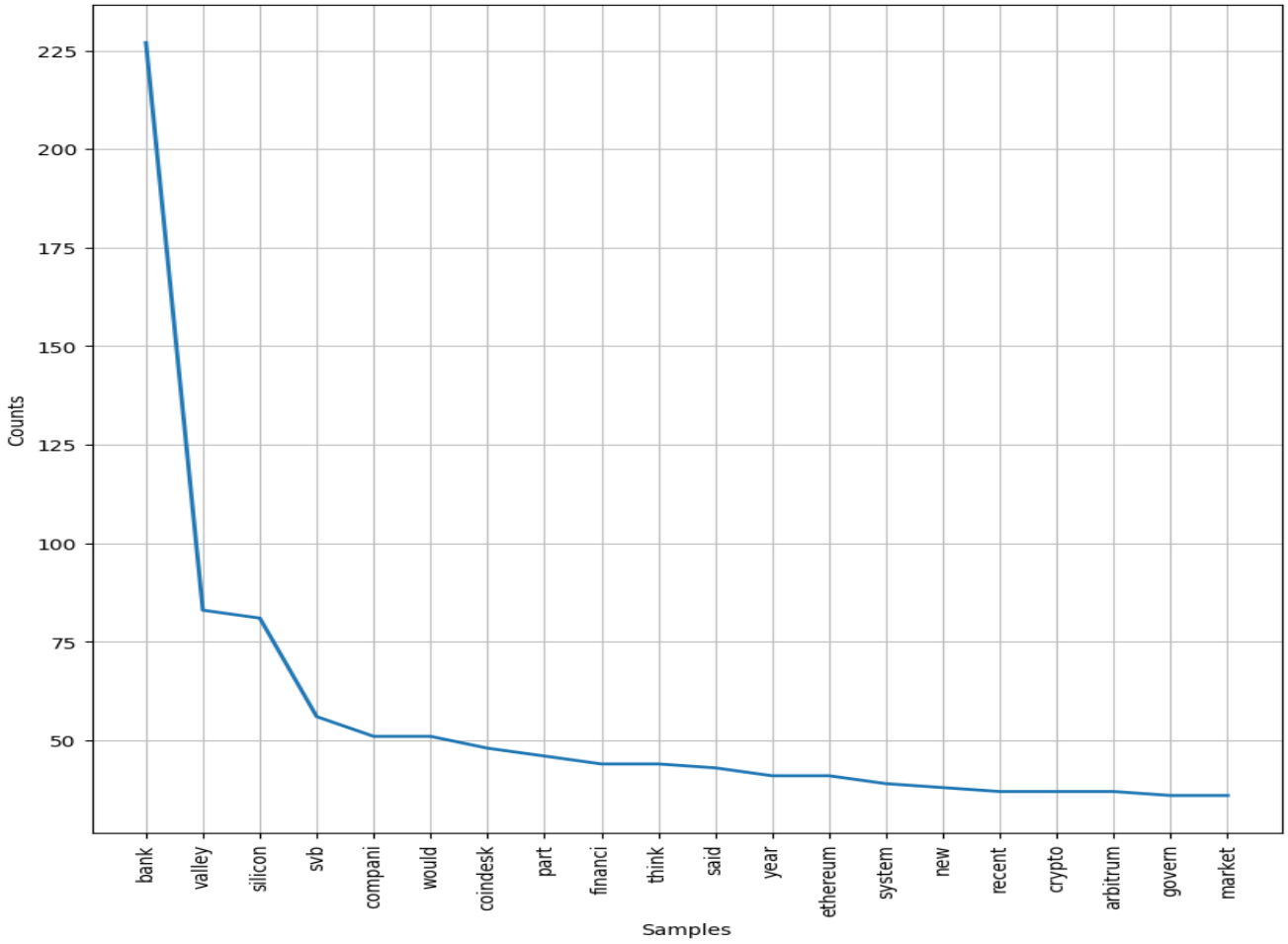
Date of Publication	Title	Country	Category	Content

	pubDate	title	country	category	content
0	2023-03-16 16:45:00	Blockchain Association seeks info from Fed, FD...	united states of america	business	The United States-based crypto advocacy group ...
1	2023-03-16 15:00:00	A North Sea Hydrogen Pipeline Network Is Possible	australia	business	The North Sea has the potential to become the ...
2	2023-03-16 12:37:05	Why it's hard for the US to cut or even contro...	united states of america	business	President Joe Biden's 2024 proposed budget inc...
3	2023-03-16 12:10:23	US regulators avoided a banking crisis by swif...	united states of america	business	U.S. regulators' swift reaction to the collaps...
4	2023-03-16 10:10:00	What happens to Silicon Valley without Silicon...	united states of america	business	Could Silicon Valley Bank's failure become a c...
5	2023-03-16 10:00:19	Meta Announces 10,000 Layoffs in 'Year of Effi...	australia	business	Meta, the social media company, will continue ...
6	2023-03-16 10:00:00	Finally — Some Good News For The Economy Thank...	united states of america	business	A ray of sunshine! The U.S. and Britain may so...
7	2023-03-16 09:00:20	Golf at 3 p.m. Thursday? Sure, It's the Aftern...	united states of america	business	None
8	2023-03-16 04:26:57	US credit crunch means it's time to buy gold a...	united states of america	business	None
9	2023-03-15 22:30:17	Michael Burry Compares Current Banking Turmoil...	australia	business	Michael Burry Compares Current Banking Turmoil...

Exploratory Data Analysis

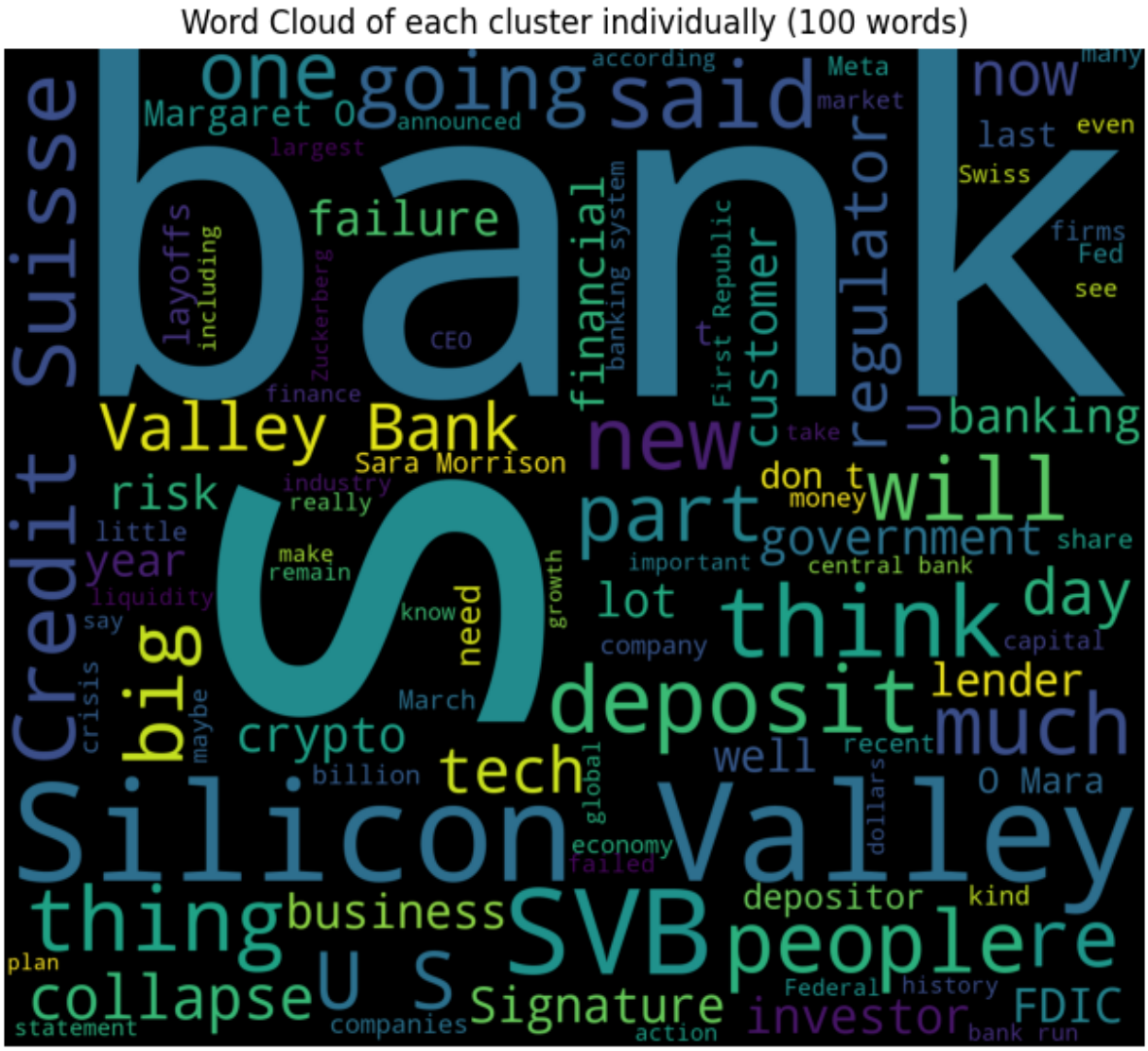
The text data had to be pre-processed to be used. Stop-words were removed using the ‘nltk’ library. The text was converted to lower-case. Numbers, links and punctuations were removed from the text corpus. Finally stemming was performed over the entire corpus.

The frequency distribution of the token was plotted.

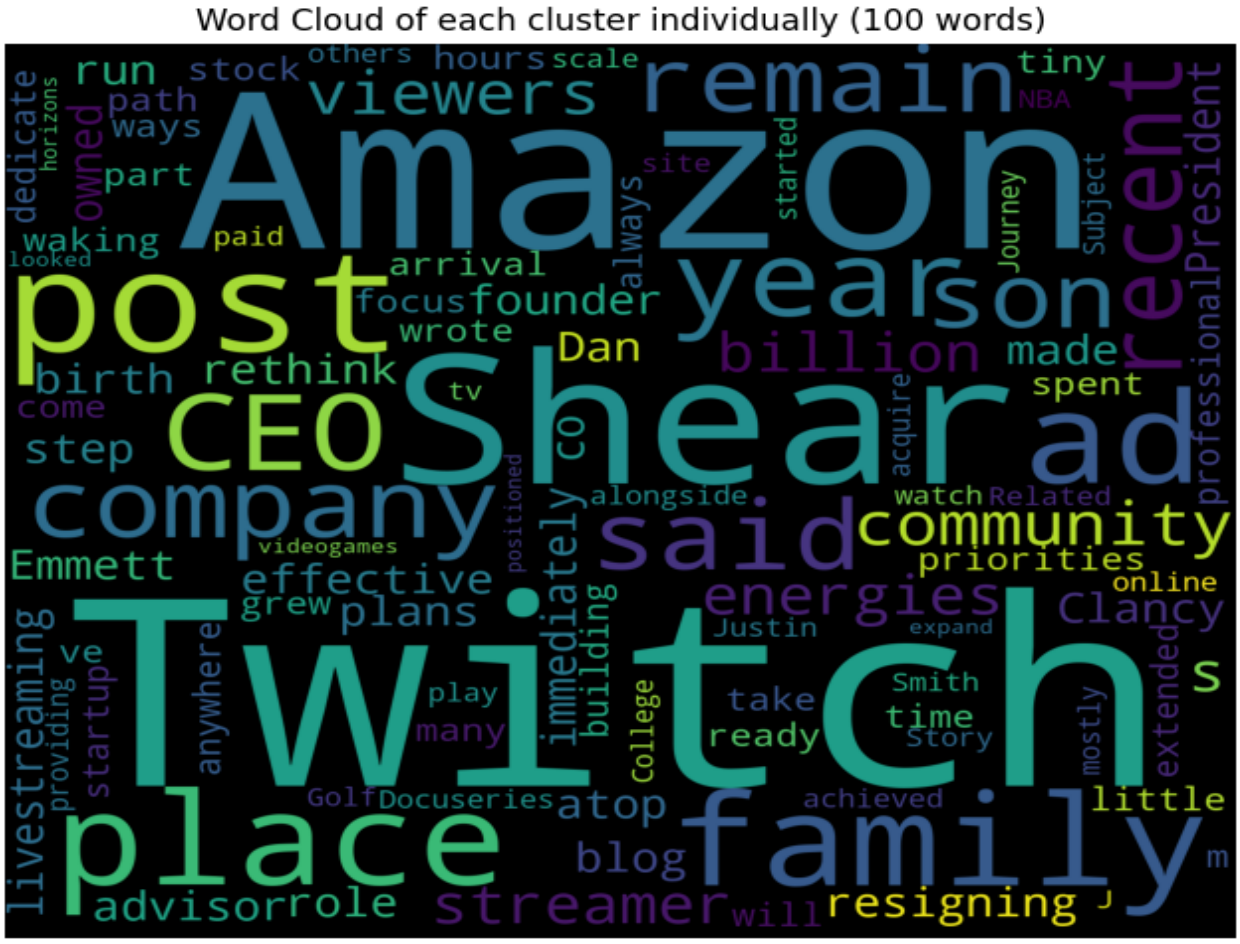


The tfidf vector for the corpus was created and clustering was performed for preliminary analysis. Eventually the word cloud was formed for all the clusters separately.

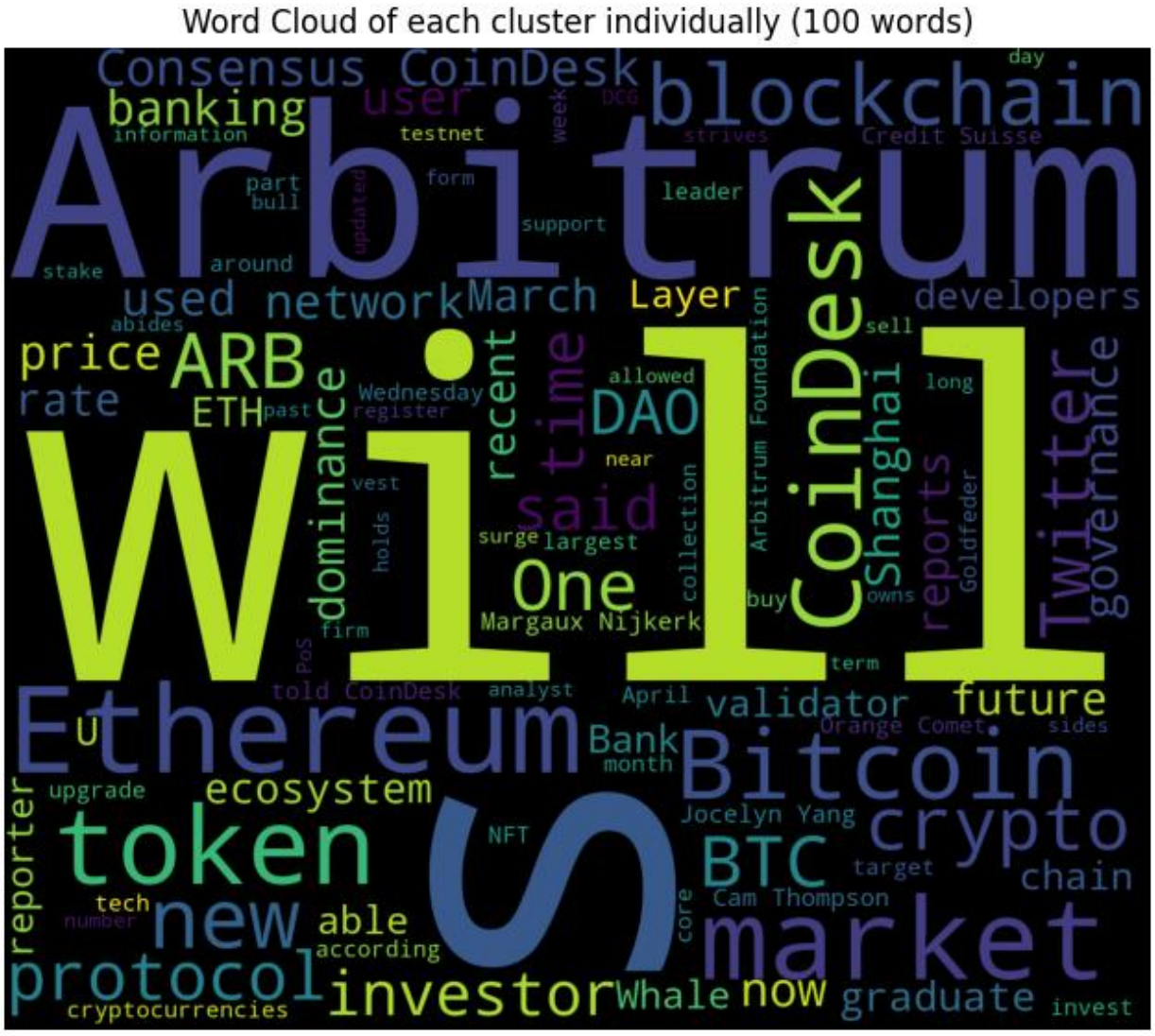
Cluster 1:



Cluster 2:


Word Cloud of each cluster individually (100 words)
Amazon
Twitch
Shear
post
place
family
company
CEO
said
remain
viewers
years
recent
billion
community
energies
plans
effective
streamer
advisorrole
resigning
livestreaming
run
stock
others
hours
scale
tiny
path
ways
owned
dedicate
horizons
part
waking
paid
looked
arrival
focus
founder
wrote
Dan
birth
rethink
come
step
tv
co
made
spent
acquire
alongside
watch
Related
professional
President
Journey
Subject
started
site
always
Emmett
videogames
positioned
immediately
building
priorities
online
Clancy
Justin
expand
grew
anywhere
ve
providing
startup
play
many
College
take
ready
Smith
time
Story
mostly
extended
s
Golf
Docuseries
atop
achieved
little
blog
m
will
J

Cluster 3:


Word Cloud of each cluster individually (100 words)
Arbitrum
Will
Ethereum
Bitcoin
market
token
blockchain
CoinDesk
Consensus CoinDesk
banking
user
used network
March
Layer
developers
price
ARB
rate
ETH
dominance
recent
time
DAO
said
One
Shanghai
Twitter
governance
reports
Margaux Nijkerk
future
validator
Bank
ecosystem
crypto
BTC
chain
new
able
protocol
investor
Whale
now
graduate
reporter
Jocelyn Yang
Cam Thompson
Credit Suisse
Arbitrum Foundation
Goldfeder
Orange Comet
testnet
information
leader
support
around
stake
abides
Wednesday
register
past
allowed
near
vest
holds
surge
largest
collection
buy
owns
sell
long
term
firm
PoS
told CoinDesk
analyst
April
month
upgrade
NFT
target
core
tech
number
according
cryptocurrencies
invest
sides
day
week
DCG
strives
updated
form
part
bull

References:

1. Kang, E., Hyun, K., & Lee, S. (2017). Stock Market Prediction Using Financial News Articles.
2. Journal of Information Processing Systems, 13(4), Pages 862-870.

3. Da, Z., Engelberg, J., & Gao, P. (2015). Investor Sentiment and Stock Returns Revisited: A Machine Learning Approach. Journal of Empirical Finance, Pages 32, 35-55.

4. Chen, A. S., & Wu, Y. (2016). The Effect of News Sentiment on Stock Market Volatility. Journal of Banking and Finance, 69, Pages 1-11.

5. Azevedo, J., Pereira, L., & Nunes, L. C. (2019). News Sentiment and European Economic Policy Uncertainty. Research in International Business and Finance, 48, Pages 449-463.

6. Giglio, S., Kelly, B. T., & Pruitt, S. (2016). Economic Policy Uncertainty and the Stock Market: Evidence from News Aggregators. Journal of Financial Economics, 120(3), Pages 441-463.

7. Gholampour, V., Ghiassi, M., & Safaei, M. (2018). Predicting Financial Market Movements Using Sentiment Analysis of Twitter Data. Journal of Business Research, 89, Pages 356-364.