Vivekanand Education Society's Institute Of Technology
Address: Hashu Adwani Memorial Complex, Collector's Colony, Chembur, Mumbai, Maharashtra 400074

Department of Artificial Intelligence and Data Science

Sentiment Analysis of Stocks via News Headlines

17, Swayam Gaikwad
24, Khyati Hegde
45, Anjali Parwani

Agenda

- Introduction and Importance
- Techniques
- Architecture
- Dataset
- Web Scraping
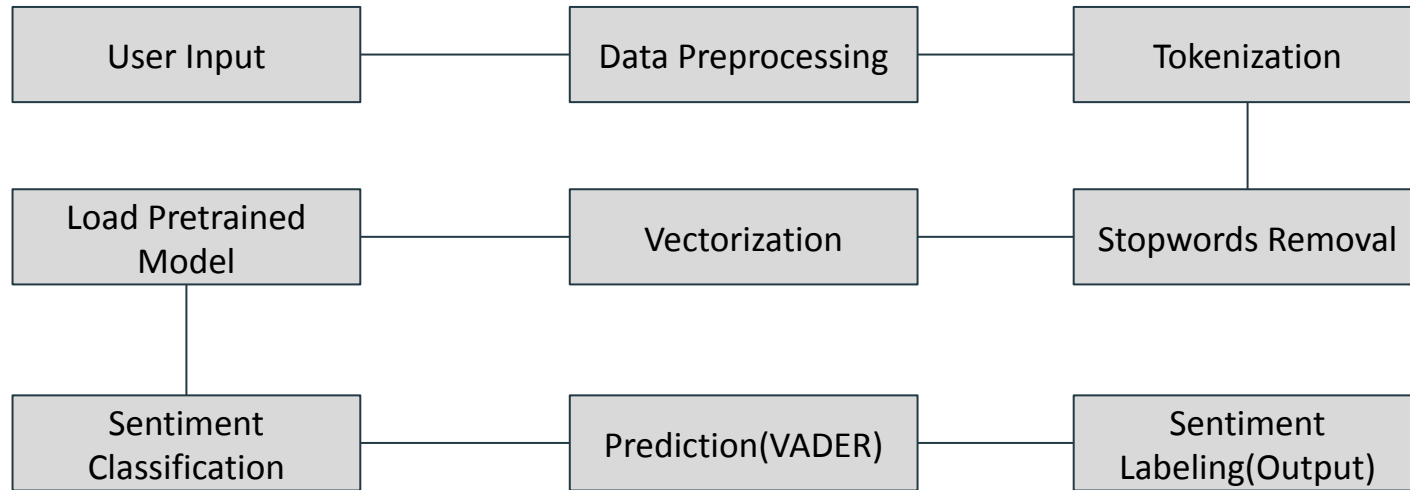- Algorithms/Libraries Used
- Output
- Visualizations

Introduction And Importance

- The rapid dissemination of information in today's digital age, accurately assessing the emotional tone of market-related news is crucial for informed trading decisions.

- This project uses sentiment analysis to understand how the market reacts to company-specific news, aiming to provide insights that may influence stock performance predictions.

- Accurate sentiment prediction helps identify market trends, manage risk, and personalize investment strategies, improving the overall decision-making process in financial markets.

Techniques

- **Rule-Based Systems and Keyword Matching:** Early approaches to sentiment analysis primarily relied on predefined rules and keyword matching to categorize news as positive, negative, or neutral.

- **Lexicon-Based Approach (VADER):** In this project, we leveraged the VADER (Valence Aware Dictionary and Sentiment Reasoner) model, a lexicon-based sentiment analysis tool. VADER is specifically designed for short texts, such as news headlines, and is sensitive to the nuances of financial language.

Architecture

User Input	Data Preprocessing	Tokenization
Load Pretrained Model	Vectorization	Stopwords Removal
Sentiment Classification	Prediction(VADER)	Sentiment Labeling(Output)

Dataset

- **News Articles**: The dataset consists of financial news articles collected from various sources, providing insights into market trends and company performance.(HTML Files)

- **Categories**: The news articles are categorized into topics such as stock performance, earnings, market predictions, and macroeconomic news.

- **Labels**: Each article is labeled based on sentiment (positive, negative, neutral) for classification purposes.

Web Scraping

Finviz Scraping :

- **Source**: Financial data was scraped from **Finviz**, a stock screener website that provides real-time market analysis, stock charts, and financial news.
- **Use**: The data gathered includes stock-related news, which is utilized for further sentiment analysis in predicting market movement.

Web Scraping

Tesla Web Scraping

- **Source**: Specific web scraping was done on **Tesla-related news** from financial news sites and company blogs to collect sentiment around Tesla's stock performance.
- **Use**: This data helps in understanding how news surrounding Tesla impacts its stock price, providing a case study for the model's prediction capability.
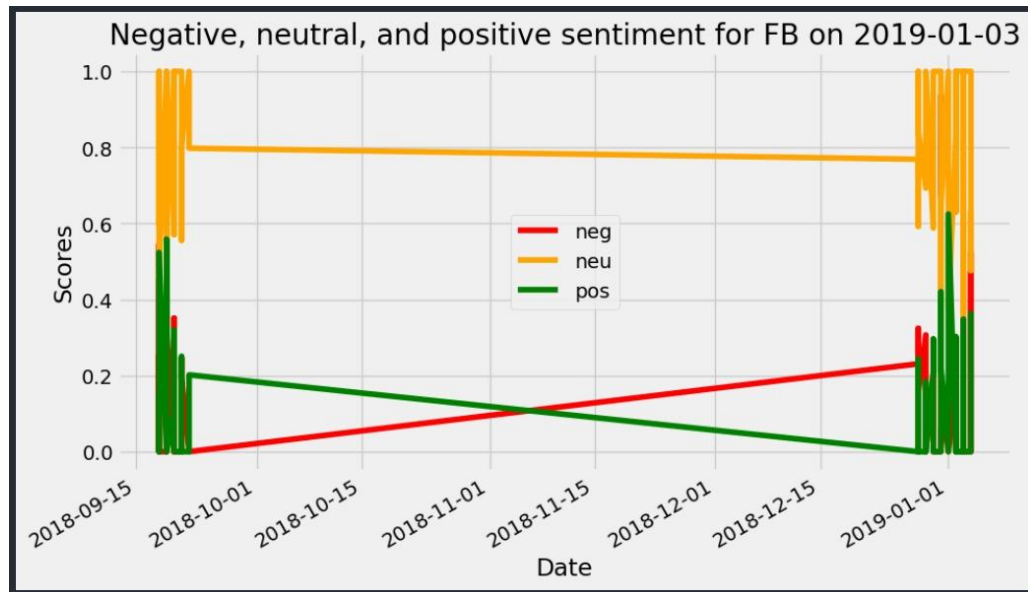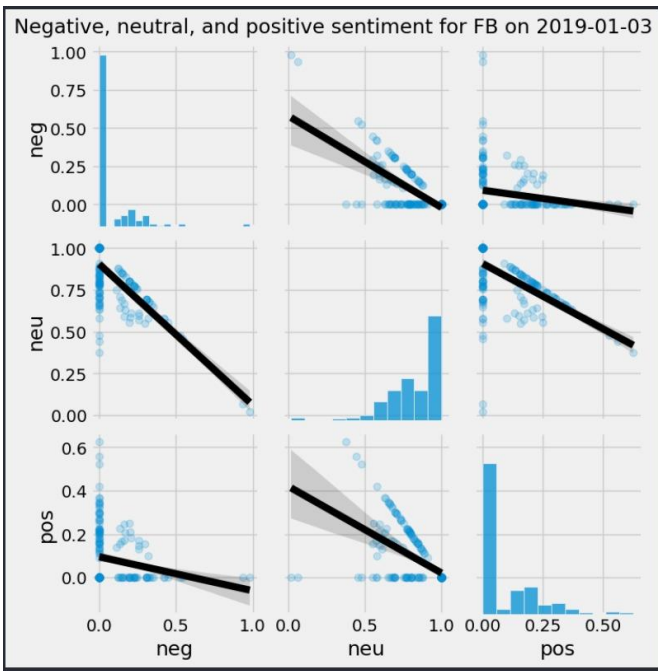
Algorithms/Libraries Used

- **BeautifulSoup**: For web scraping and parsing HTML content.
- **Count Vectorization**: Transforms text into numerical representation for document-term matrix.
- **Multinomial Naive Bayes**: A probabilistic classifier used for sentiment classification.
- **VADER Sentiment Analysis**: Pretrained sentiment analysis model tailored for financial headlines.
- **Randomized Search CV**: For hyperparameter tuning of models.
- **Matplotlib and Seaborn**: For data visualization, particularly sentiment scores over time.

Output

time	headline	neg	neu	pos	compound
06:00:00	Why Internet Censorship Doesnt Work and Never ...	0.000	1.000	0.000	0.0000
08:04:00	3 Big Stock Charts for Thursday: Medtronic, Fa...	0.000	1.000	0.000	0.0000
09:07:00	The Zacks Analyst Blog Highlights: Facebook, W...	0.000	0.846	0.154	0.2500
09:12:00	Why The FAANGs Are Crucial To A 2019 Market Rally	0.000	1.000	0.000	0.0000
09:34:00	Apple warning stoking fears of slowdown in Chi...	0.426	0.574	0.000	-0.6369
09:48:00	Stock Market News For Jan 3, 2019	0.000	1.000	0.000	0.0000
09:58:00	Facebook Falls 3%	0.981	0.019	0.000	-0.9993
10:08:00	Mark Zuckerberg Halts Stock Sales as Facebook ...	0.000	0.784	0.216	0.2960
10:21:00	Facebook Under Fire for Collecting Data From A...	0.231	0.769	0.000	-0.3400
10:42:00	Why Netflix Stock Will Rise Back Above $300 In...	0.000	1.000	0.000	0.0000
12:25:00	3 Great Reasons You Should Be Bullish on Twili...	0.000	0.661	0.339	0.6249
13:31:00	Take the Money And Run From Twilio Stock	0.000	1.000	0.000	0.0000
13:36:00	What's Next For Apple (AAPL) After It Slashed ...	0.147	0.853	0.000	-0.2263
15:14:00	Zuckerberg Paused His Selling of Facebook Shar...	0.000	0.662	0.338	0.6249
15:26:00	Is Facebook Stock the Best FANG Stock to Buy?	0.000	0.656	0.344	0.6369
16:10:00	Google Stock Upgraded, Viewed As 'Most Defensi...	0.000	0.878	0.122	0.1010
17:24:00	Facebook (FB) Reportedly Testing Dark Mode on ...	0.000	1.000	0.000	0.0000
17:45:00	Investors Are Focused Too Closely on Facebook ...	0.000	0.652	0.348	0.5859
22:59:00	[$$] Facebook Begins New Year in Fixer-Upper Mode	0.000	1.000	0.000	0.0000

Visualizations

Visualizations