SOCIAL MEDIA ANALYZE

Social Media Sentiment Analysis

Using VADER & TextBlob with Python (PRAW)

Project Overview

Understanding Online Sentiments

Social media platforms like Reddit are vast reservoirs of public opinion and discourse. Analyzing this data can provide invaluable insights into user emotions, attitudes, and trending topics. This project focuses on leveraging powerful Natural Language Processing (NLP) tools to extract and interpret sentiment from Reddit data.

Rich Data Source

Tapping into social media for opinions and discussions.

Emotional Insights

Understanding the underlying sentiment in user-generated content.

NLP Application

Utilizing VADER and TextBlob for comprehensive sentiment analysis.

How It Works

Project Methodology

01

Data Collection

*Reddit posts and comments are programmatically fetched using the **PRAW (Python Reddit API Wrapper)** library, targeting specific subreddits or keywords to gather relevant data.*

02

Sentiment Analysis

Two distinct NLP libraries are employed:

- ***VADER (Valence Aware Dictionary and sEntiment Reasoner)**: Provides a compound score, along with positive, negative, and neutral percentages, ideal for social media text.*
- ***TextBlob**: Calculates both polarity (ranging from -1 to +1) and subjectivity (ranging from 0 to 1), offering a nuanced perspective.*

03

Data Storage & Export

*All extracted data, including the original text and calculated sentiment scores, are organized and stored efficiently in a **CSV (Comma Separated Values)** file format. This structured output facilitates easy access for subsequent analysis, visualization, and integration with other data tools.*

Powering the Analysis

Key Tools & Libraries

Python

The primary programming language, offering extensive libraries for data science.

PRAW

Simplifies interaction with the Reddit API for data extraction.

Pandas

Essential for efficient data manipulation and analysis.

NLTK

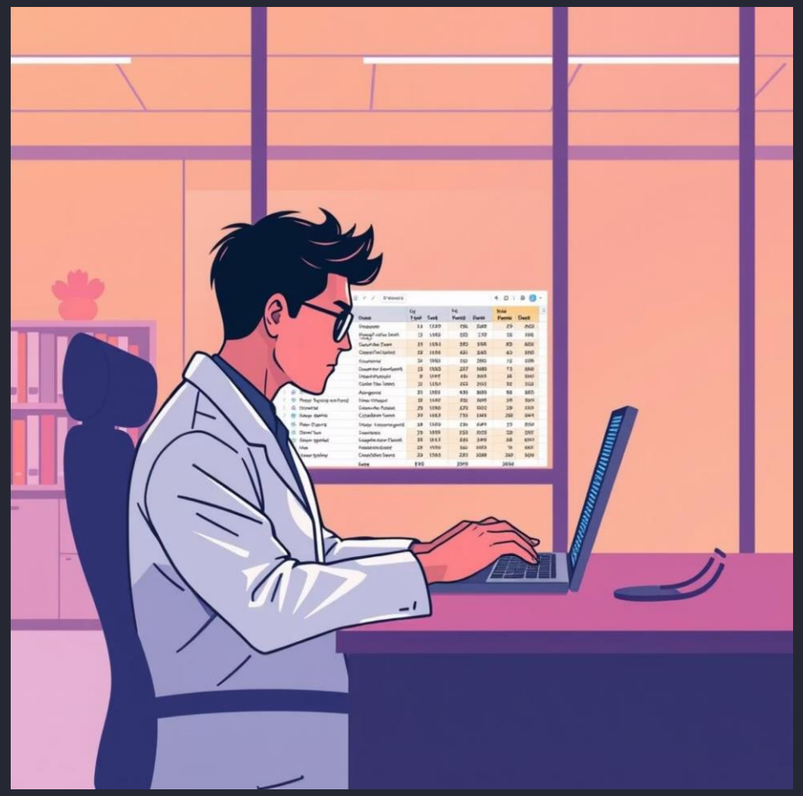
Provides the VADER lexicon and tools for natural language processing.

TextBlob

Offers a simple API for common NLP tasks, including sentiment analysis.

Findings from the Analysis

Results & Output



*Our analysis focused on posts from the **r/MachineLearning** subreddit, a vibrant community for AI and ML enthusiasts. The project successfully extracted and processed a significant volume of data, generating comprehensive sentiment metrics for each piece of content.*

- ***VADER Scores**: Each post received a **compound score** (normalized sum of all lexicon ratings), along with individual **positive, negative,** and **neutral percentages**.*
- ***TextBlob Scores**: We obtained **polarity scores** (indicating the emotional leanings from -1 to +1) and **subjectivity scores** (measuring the extent of personal opinion versus factual information, from 0 to 1).*

*All raw data and sentiment scores were meticulously saved in a **.csv file**, making it ready for advanced visualization tools (e.g., Tableau, Power BI) and deeper statistical exploration.*

Key Takeaways & Future Directions

Conclusion & Next Steps

This project demonstrates the power of sentiment analysis in extracting meaningful insights from unstructured social media data. By combining VADER and TextBlob, we achieve a more robust and reliable understanding of online discussions.

Actionable Insights

Sentiment analysis provides a clear window into public opinion, vital for brand monitoring, trend detection, and market research.

Enhanced Reliability

Integrating multiple sentiment models mitigates individual biases and improves the overall accuracy of emotional classification.

Future Scope

Next steps include deploying advanced machine learning models (e.g., deep learning) for more granular sentiment classification and integrating real-time data streams.