

# Stock Sentiment Analysis

Stock Sentiment Analysis Using TOP 25  
News Headlines

Presented By-Dev Dashora

25,423 Pageviews  
5.32 Pages/Visit

Traffic Sources Overview

Direct Traffic	3,097.00 (40.49%)
Search Engines	2,910.00 (38.04%)
Referring Sites	1,642.00 (21.47%)



Visitors Overview



Email-devdashora.work@gmail.com

A black street sign for Wall Street is mounted on a pole. The sign is rectangular with a black background and white text. It reads "WALL ST" in large, bold, sans-serif letters. Above the "WALL ST" text, there is a smaller section with an arrow pointing left and the numbers "9-15". To the right of the main text, there is a small, square inset image showing a person walking past a classical building with columns.

# Content

*Project Objective*

*About Dataset*

*Machine Learning Algorithm*

*Evaluation & Results*

*Project Overview*

*Contact ME*



# Project Objective

---

This project aims to perform stock sentiment analysis using the top 25 news headlines from a dataset. By analyzing the sentiment of these headlines, the model can determine the overall market sentiment, aiding in investment decision-making. This project effectively uses a Random Forest Classifier to analyze stock sentiment based on news headlines.

# About Dataset

## *Source and Collection:*

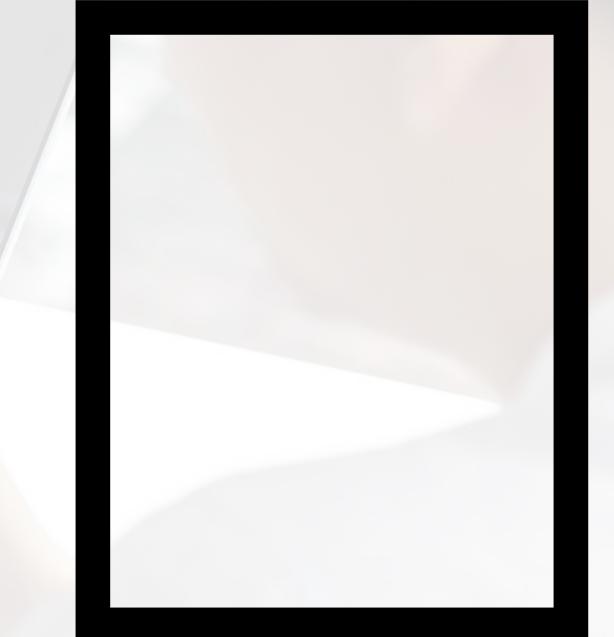
The dataset consists of the top 25 news headlines on stock market developments, gathered from leading financial news websites.

## *Scope and Timeframe:*

The dataset covers a variety of sectors and companies, with headlines spanning a specific period to capture relevant market events and trends.

## *Purpose and Analysis:*

The dataset is used to perform sentiment analysis, providing insights into the market's emotional response to news, which can influence stock price movements and investment strategies.



# Machine Learning Algorithm Used

## Bag of Words (BoW) with Bigrams:



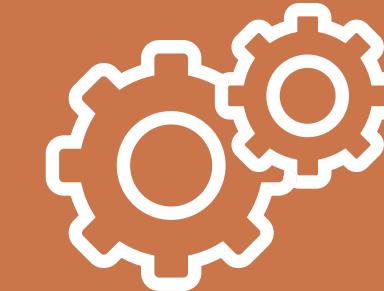
converts text into numerical features by counting the frequency of word pairs (bigrams)

## Random Forest Classifier:



combines multiple decision trees to enhance classification accuracy and generalization by averaging their predictions.

## CountVectorizer:



A feature extraction tool that converts a collection of text documents into a matrix of token counts, capturing the frequency of each word in the text.

# Evaluation & Results

[[139 47]
[ 9 183]]
0.8518518518518519
precision
0 0.94
1 0.80
recall
0 0.75
1 0.95
f1-score
0 0.83
1 0.87
support
0 186
1 192
accuracy 0.85 378
macro avg 0.87 0.85 378
weighted avg 0.87 0.85 0.85 378

The model achieves an accuracy of approximately 84% on the test dataset. Evaluation metrics such as confusion matrix, precision, recall, and F1-score indicate strong performance, especially in identifying both positive and negative sentiments accurately.

# Project Overview

---

- **Feature Extraction:** Utilized Bag of Words with bigrams via CountVectorizer to convert text data into numerical features.
- **Model Training:** Employed a Random Forest Classifier with 200 trees and entropy for information gain to classify sentiment from news headlines.
- **Accuracy:** The model achieved approximately 84% accuracy on the test dataset.
- **Evaluation:** Assessed model performance using confusion matrix, precision, recall, and F1-score to ensure effective sentiment classification.
- **Insights:** The project effectively identifies both positive and negative sentiments in stock-related news, providing valuable insights for market sentiment analysis.

# Thank You



*devdashora.work@gmail.com*

*hydrogen012*

*devdashora*