

# Financial News Sentiment Analyzer for Market Insights

## Team members:

- a. Deepan Tripathy, MS in CS (002692015)
- b. Ayantan Dandapath, MS in CS (002712017)

## Project Idea:

Financial News Sentiment Analyzer utilizes a web crawler that collects financial news articles from various sources such as financial news websites, business journals, and blogs. The crawler focuses on extracting articles related to specific companies or financial markets. Once the articles are collected, a sentiment analysis algorithm analyzes the tone and sentiment expressed in each article classifying as positive, negative or neutral. The sentiment analysis can help gauge market sentiment towards specific companies or financial instruments.

## Basic Algorithm:

- a. Define dataset:
  - Identify a list of reputable financial news websites, blogs, and journals and determine the criteria for selecting articles, such as specific keywords, company names, or market-related terms.
  - Utilize web scraping tools or libraries like Scrapy, to navigate through the target websites.
- b. Design a Web Crawler:
  - Parse HTML content like article titles, publication dates, authors, and article bodies.
  - Utilize techniques like headless browsing with tools like Selenium to handle websites that load content dynamically using JavaScript.
  - Extract financial metrics and market-related data mentioned in the articles like stock symbols, market indices, or specific financial ratios.
- c. Sentiment Analysis Algorithm:
  - Data preprocessing like cleaning the article text by removing stop words, punctuation, and special characters. Normalize the text by converting it to lowercase and applying techniques like word stemming or lemmatization.
  - Feature extraction by data vectorization by using bag-of-words, TF-IDF, or word embeddings
  - Sentiment classification by using Naive Bayes, XgBoost, KNN or SVM
- d. Result Visualization:
  - Generate graphs that display sentiment trends over time through seaborn, matplotlib or tableau, word clouds to highlight frequently mentioned terms in positive or negative articles.
- e. Alerting System Algorithm:

Determine the thresholds for sentiment scores that trigger alerts and develop a system that can send alerts to users when the specified sentiment thresholds are met like email triggering system.

Timeline:

Dataset definition	1st July - 3rd July
Design crawler	3rd July - 5th July
Implement and Test crawler	5th July - 7th July
Re-work crawler and improve	7th July - 10th July
Implement data preprocessing	10th July-13th July
Design sentiment analysis algorithm	14th July - 16th July
Unit test 1	17th July - 20th July
Re-work based on unit test 1	20th July - 22nd July
Final test run	23rd July - 26th July
Project report preparation and submission	26th July

Purpose:

Purpose of the project is to provide users with actionable insights and a comprehensive understanding of market sentiment derived from the analysis of financial news articles.