

Final Project Proposal – NLP for Financial News Sentiment & Summarization

Problem Description:

Financial news plays a crucial role in shaping market dynamics and investor sentiment. This project aims to first **analyze the sentiment** (positive, negative, neutral) of financial news articles and then **predict stock prices** based on the sentiment analysis. By assessing how the news influences market sentiment, we aim to predict the subsequent stock price movement. The project will consist of two main stages:

1. **Sentiment Analysis:** Classifying the sentiment of articles using various NLP models such as **Recurrent Neural Networks (RNNs)**, **encoder-decoder architectures**, and **Transformer-based models** like **BART** and **T5**.
2. **Stock Price Prediction:** Based on the sentiment analysis, predicting the movement of stock prices (up/down/neutral) using machine learning models that integrate both **sentiment** and **historical price data**.

Data Sources:

We will collect financial news articles using public APIs like **NewsAPI** and RSS feeds from reliable financial platforms like **Yahoo Finance** and **Investing.com**. Additionally, stock price data will be gathered from **Yahoo Finance API** or other financial data providers to combine with sentiment analysis for price prediction.

Anticipated Challenges :

- Collecting and preparing sufficient clean data for both the sentiment analysis and stock price prediction tasks.
- Handling domain-specific vocabulary in financial news.
- Managing long input sequences, particularly in articles with extensive details.
- Predicting stock prices accurately from sentiment analysis, given the complexities of market behavior.
- Evaluating the quality of sentiment analysis and determining its correlation with stock price movements.

Initial Roles:

- **Baptiste:** Responsible for collecting, cleaning, and preparing the datasets from the APIs, and handling stock price data for the prediction task.
- **Thibault:** Will implement and fine-tune NLP models for sentiment classification, then use sentiment data for stock price prediction.
- **Both:** Will create a baseline comparison model and evaluate the performance of different models, comparing sentiment-based predictions with a benchmark.