Maria Eusse Henao and Paula Lopez Burgos

DS-340

October 22, 2023

Final Project Proposal

Collect financial news articles to analyses which companies have been talked about the most and in what sentiment to predict which company would be the best to invest in.

**Methods.** Using an API ( ) to get up do date financial news articles. We will use natural language processing to do sentiment analysis. First, we will need to process the data before feeding it to the neural network; tokenization, lowercasing, removing stop words, punctuation, stemming. We will represent words in a vector format with pre-trained word embeddings like BERT. The architecture will be Recurrent Neural Network (RNN) or a Transformer based model (we could fine-tune pre-trained transformer models).

**Existing resources.** UsingBERT (Bidirectional Encoder Representations from Transformers)from the transformers library to tokenize and use the base training model.

What’s new. We are going to use the … API. Moreover, we need to adapt the code to our number of classes on our sentiment analysis. We will fine-tune our model by using TensorFlow, and will use Adam as our optimizer, the categorical cross entropy function as our loss function. And we could use the SoftMax function to get the probability of each sentiment.

**Plan.**

1. API implementation
2. Process text (tokenize) and create sentiment classification.
3. Train the neural network and tune learning rate and loss function to find best performance.
4. Evaluate performance of the model (accuracy) with F1-score