

A BERT-Based Model for Financial Social Media Sentiment Analysis

Authors : Josiel Delgadillo, Johnson Kinyua, Charles Mutigwe

Abstract : The purpose of sentiment analysis is to determine the sentiment strength (e.g., positive, negative, neutral) from a textual source for good decision-making. Natural language processing in domains such as financial markets requires knowledge of domain ontology, and pre-trained language models, such as BERT, have made significant breakthroughs in various NLP tasks by training on large-scale un-labeled generic corpora such as Wikipedia. However, sentiment analysis is a strong domain-dependent task. The rapid growth of social media has given users a platform to share their experiences and views about products, services, and processes, including financial markets. StockTwits and Twitter are social networks that allow the public to express their sentiments in real time. Hence, leveraging the success of unsupervised pre-training and a large amount of financial text available on social media platforms could potentially benefit a wide range of financial applications. This work is focused on sentiment analysis using social media text on platforms such as StockTwits and Twitter. To meet this need, SkyBERT, a domain-specific language model pre-trained and fine-tuned on financial corpora, has been developed. The results show that SkyBERT outperforms current state-of-the-art models in financial sentiment analysis. Extensive experimental results demonstrate the effectiveness and robustness of SkyBERT.

Keywords : BERT, financial markets, Twitter, sentiment analysis

Conference Title : ICASA 2022 : International Conference on Applications of Sentiment Analysis

Conference Location : Cairo, Egypt

Conference Dates : December 15-16, 2022