

Using a novel Skip Gram, Knowledge Graph and Long Short-Term Memory (SKG-LSTM) model for Stock Price Prediction



By Mr Luk Tsz Kit, Mr Liege Cheung (UIUC), Dr Adela Lau, and Mr Rogers Chan (ImpactInvest Limited)
 (https://dslab.saas.hku.hk/cgi-bin/application1.cgi/)

Financial news bring significant returns and volatility on stock market. A good example is the GameStop and AMC stock prices going up and down sharply in 2021 which were associated to social media and news discussions. This project aims to propose a novel methodology which uses financial news, ontology and knowledge graph to predict stock price movement. The Data Science Lab collected the stock price data and news from Yahoo finance. An Ontology based Named Entity Recognition (O-NER) is used to train the N-gram model for the news analysis and stock price movement prediction. A hybrid model of combining Skip-Gram, Knowledge Graph and Long Short-Term Memory (SKG-LSTM) is used to analyse the sentiment of the news. There are six components in the SKG-LSTM.

Data pre-processing model: It uses NLP techniques to extract the keywords from news title.

Graphical representation for news title: It converts the extracted words to graphical format based on the sentence structure.

Skip-gram model: It trains a model to embedded the extracted words into vectors based on similarity measurement.

Grouping nodes: It groups the nodes that have closed distance in vector representation.

Knowledge graph embedding (KGE): It embeds the sentence structure knowledge graph into a feature vector.

Stack long short-term memory (LSTM): It predicts the stock price movement based on the feature vector.

The results showed that the SKG-LSTM achieves a 69% accuracy rate. SKG-LSTM used sentence structure to construct KG for stock prediction that brings significant improvement comparing with the baseline model of “NER + SVM”, which is commonly used with an accuracy rate of 54% only.

“News information is unstructured. The ontology can convert the unstructured news into knowledge for improving the stock price prediction,” said Mr Dominic Wu, Partner of Vision Real Capital.

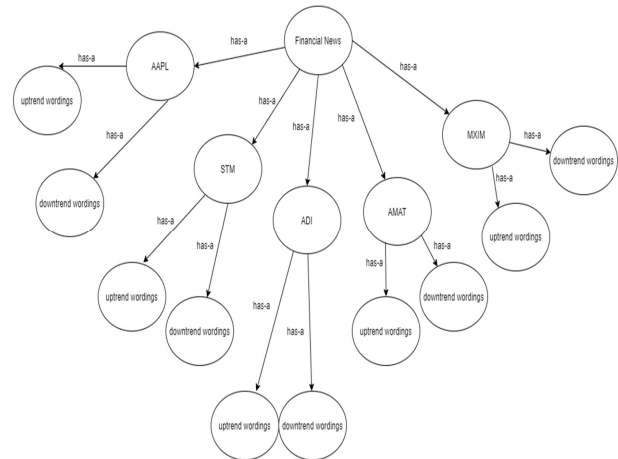


Figure 1 Ontology of Financial News

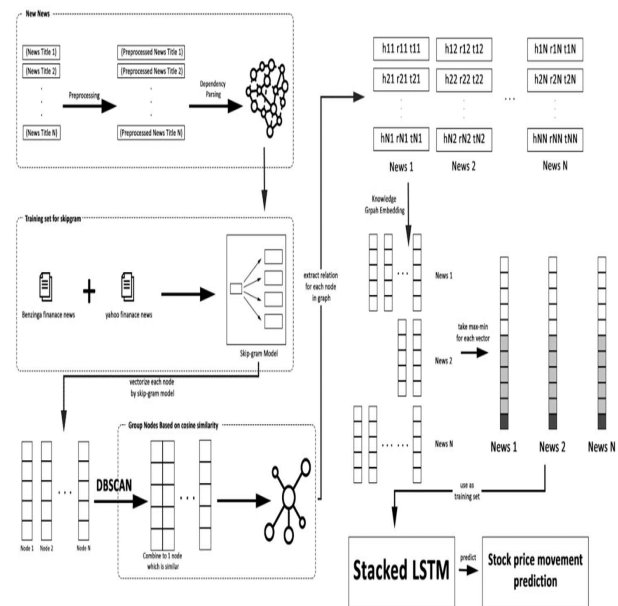


Figure 2 An overview of the SKG-LSTM model

“Other than news, social media is another recent method for stock price prediction. The Data Science Lab always research some new methods for solving existing business problems with our business partners. The developed SKG-LSTM model can be applied to some other areas such as COVID19 preventive health policy control, mental health risk monitoring, etc.,” said Dr Adela Lau, Deputy Director of HKU SAAS Data Science Lab.