Presented By : Kaivalya Anand Pandey

Improving Sentiment Analysis for Stock trend Prediction

Scientific Advisor : 1. PhD Evgeny N. Pavlovskiy
2.Alexey Stankus

Novosibirsk State University

December 29, 2020

Introduction

- Stock Trend Prediction is to predict whether the stocks will go in high or low.
- The well-known efficient-market hypothesis (EMH) suggests that stock prices reflect all currently available information and any price changes based on the newly revealed relevant information.
- However, due to the implicit correlations between daily events and their effect on the stock prices, finding relevant information that contribute to the change of price for each individual stock are difficult.

Current State of Art

- Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network by Huicheng Liu.
- The model is used as Attention-based LSTM (At-LSTM).
- News Datasets are taken by Reuters News and Bloomberg News from 2006 to 2013 made available by Ding.
- Stock Prices are taken from Yahoo Finance of same duration as news.

Methodolody used in SOTA

- Bag of Words Model and Word Embedding. It is also known as the vector space model.
- Model used - Long Short Term Memory RNN.
- Optimisation algorithm - Adadelta Optimization Algorithm.
- Loss Function - Cross entropy loss

$$J(w) = -\frac{1}{N} H(p_n, q_n) = -\frac{1}{N} \sum_{i=1}^{N} [y_n * \log \tilde{y}_n + (1 - y_n) * \log(1 - \tilde{y}_n)]$$

Model Architecture

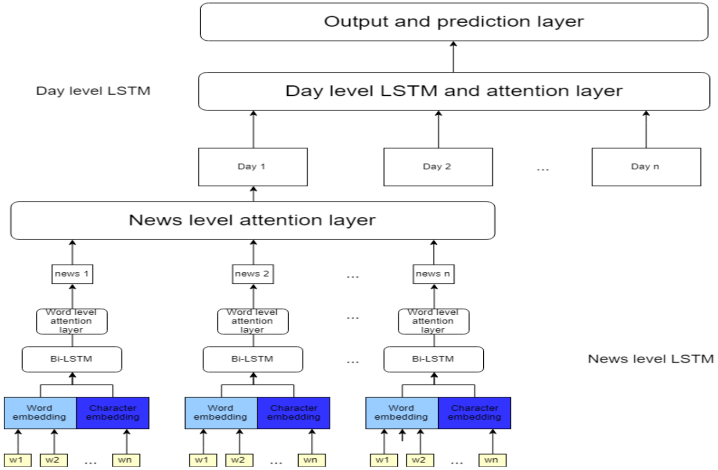

Figure 3: Model Structure

Model Architecture

- Model Input: Entire news title content as our models input and use LSTM-based encoder to encode it to a distributional representation to tackle with sparsity problem
- Word Embedding: The embedding layer takes a sequence of sentences as input, this sequence corresponds to a set of titles of news articles. These embedding are unique vectors of continuous values with length $w = (w1, ..., wl)$ and $wi \in Rm$ for each word in the training corpus, m is the word level embedding dimension.
- Character Embedding: The character composition feeds all characters of each word into a Convolutional Neural Network (CNN) with max-pooling to obtain representations $c = (c1,...,cl)$ and $cn \in Rn$ for each word in the training corpus, n is the character composition dimension.

Model Architecture

- Finally each word is represented as a concatenation of word-level embedding and character-composition vector ei = [wi; ci]. A matrices es $\in$ Rk$*$(m+n) can be used to represent a news after the embedding layer, where k is the length of the news.
- News Level Bi-LSTM and Self-Attention Layer : words and their context in the news title are fed into a Bi-LSTM based sentence encoder to perform distributional representation.
- Bidirectional LSTM (Bi-LSTM) is a variant of LSTM which shows better result then uni-direction LSTM in recent NLP tasks as they can understand context better.

Model Architecture

- Word level Self-attention layer

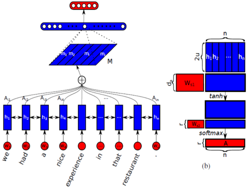

Figure 4: Self attention mechanism

Model Architecture

- News level Self-attention layer: Not all news contributes equally to predicting the stock trend. Hence, in order to reward the news that offers critical information, multi-hop self-attention on top of the encoding layer is used to aggregate the news weighted by an assigned attention value.
- Day level Bi-LSTM and self-attention layer : news published at different dates contribute to the stock trend unequally, self-attention mechanism is used to reward the dates that contribute most to the stock trend prediction.
- Output and Prediction Layer: The last stage of At-LSTM model is a traditional fully connected layer with softmax as activation function whose output is the probability distribution over labels.

Rsults Comparison

S&P 500 index prediction Experimental Results		
Model	Average Accuracy	Max Accuracy
SVM	56.38%	–
Bag-At-LSTM	61.93%	63.06%
WEB-At-LSTM	62.51%	64.42%
Ab-At-LSTM	60.6%	61.93%
Doc-At-LSTM	59.96%	60.6%
Tech-At-LSTM	62.51%	64.42%
CNN-LSTM	61.36%	63.06%
E-NN	58.83%	–
EB-CNN	64.21%	–
KGEB-CNN	**66.93%**	–
At-LSTM	63.06%	**65.53%**

Future Work

- My future work will be to try to increase accuracy of above mentioned results by trying to extract some useful information from news text and trying to improve the model architecture.
- I will also try to include any useful approach which might increase our accuracy from Knowledge Graph Event Embedding(KGEB) which is more powerful to model content in news title than sequence embedding used in this work.

THANK YOU

Questions?