Deep Learning for event-based stock price dependencies across multiple market sectors.

1. Introduction

All modern businesses have value chains so diverse and widespread that their performance is highly dependent on the performance of businesses in other sectors. These dependencies could be stemming from the complementary or supplementary nature of products and services or the supply chain relationship usually in form of client-vendor one. Stock prices are usually a good representation of perception of business in market. While not always a direct measure of true performance of the business they are sensitive to the cross-market events. Stock prices of businesses in a sector usually move together. Studying the cross-sector sensitivity helps build models to predict stock prices based on observed sensitivity. This can help add another dimension to the existing price prediction algorithms. The key challenge is to extract meaningful events from historical time series data of multiple companies and build models that can identify relationship between stock prices and events. The problem at hand can be broken into two major components, extraction of events from time series data and event-based predictions.

2. Literature Survey

One of the ways in which event detection in Time Series is “change-point detection” where in dynamic phenomenon which have significant behavioral changes over time are identified [2]. The focus of this study is to apply data mining techniques to identify change points in a raw time series data where it is difficult to define threshold of change. The problem can be defined as identifying pre-defined number of piecewise segmented models between change points. Maximum likelihood estimate of the changepoints are computed based on the number of timepoints in a segment by a likelihood criteria function. Overall this approach predicts a change point based on the change in underlying parameters of data in a segment using model selection techniques.

Change point detection methods can be classified into two categories, real time detection which focuses on instantaneous changes and retrospective detection that allow for longer reaction times[2][3]. We focus on retrospective detection given the nature of our problem and the dataset provided. One such method involves tracking changing data adaptively and gradually forgetting the effect of past data using autoregressive modeling and detecting outliers relative to this trained model[4].

Event based prediction in the stock market deals with anticipating the fluctuation of stock prices as a result of certain events taking place like earnings announcements, large scale litigations or the release of a new product.

There has been quite a bit of work done in event-based prediction. Xiao Ding et al. discuss a method of using events extracted from news text to make predictions [5]. These events are represented as dense vectors (represented using event embeddings) which are then sent to a deep convolutional neural network to model the short-term and long-term influences of events on the stock price movements. Event embeddings are like templates for similar types of events, for example: If Apple and Google release a product, these events would have the same embedding.

The input to the model is a chronologically ordered sequence of event embeddings, where the events in one day are averaged together. The model understands the events as long-term, mid-term and short-term based on the time-span of the event. The different durations of events are extracted from the data by using a sliding window that combines 'l' neighboring events for long-term and 'm' neighboring events for mid-term. The model eventually learns the effects of these durations. The model takes in this time series aspect as well as a number of important local features (determined by the pooling layer of the neural network) to produce a binary output (+1 for an increase in stock prices and -1 for a decrease). The results for this study showed that using deep convolutional networks and event embeddings increase predictions by 6% over the S&P 500 index prediction.

3. Proposal

We aim to provide a method to model price changes in stock prices of companies in one sector based on events detected in the stock price changes in another sector. Essentially, we want to see if the price of the stock of an organization in one sector is in any way dependent on the events happening in unrelated sectors.

We take 5 different sectors in the market. For each sector we pick a random set of 10 companies that we take as representatives of that sector. We define Output Sector as the sector whose prices we want to predict. We define the Input Sector as the one for which we track the events. We also propose to take a window of time T for which we capture event E in the Input Sector, where each event E is characterized by m features. Therefore, at any given time T we capture m features for each of the 10 companies, giving us 10m features at time T for the Input Sector. We also calculate the price Index of the output sector, which is the weighted average of the traded stock prices in the time T of all the companies. This gives us out target variable PI(S), defined as the price index of the Output Sector. In the end we will have the following table:

PI(Output Sector)	1	2	…	10m
P1	e_1	e_2		e_{10m}
..				
P(t)				

The next step is to build a simple regression model, that models the prices of the Output Sector on the event features from the Input Sector. Since there are 5 sectors, and we model the price of each sector with every other sector, we have 5 x 5 = 25 different models that try to predict the dependence of the Output Sector (eg. S1) with itself and the 4 other Input Sectors (S2 to S5). In the end we can model each sector on every other sector, where we give the weights and biases of each model we create.

	S1	S2	S3	S4	S5
S1	W11, b11	W12, b12	W13, b13	…	…

S4					
S5	…	…	…	…	W55, b55

4. References

[1] Valery Guralnik, Jaideep Srivastava, Event Detection from Time Series Data”, KDD '99 Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining - August 18, 1999

[2] M. Basseville, I.V. Nikiforov, Detection of abrupt changes: theory and application Prentice-Hall, Inc., Upper Saddle River, NJ, USA (1993)

[3] Adams, R. P., & MacKay, D. J. C. (2007). Bayesian online changepoint detection. Technical Report. arXiv:0710.3742v1 [stat.ML]

[4] J. Takeuchi, K. YamanishiA unifying framework for detecting outliers and change points from non-stationary time series data, IEEE Transactions on Knowledge and Data Engineering, 18 (4) (2006), pp. 482-492

[5] Xiao Ding, Yue Zhang , Ting Liu , Junwen Duan, “Deep Learning for Event-Driven Stock Prediction”, Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015), July 2015