Gated Recurrent Units and Deep Convolutional Neural Networks for Stock Price Prediction

Sarthak Jindal[1]

Indraprastha Institute of Information Technology, Delhi
sarthak15169@iiitd.ac.in
2015169

Abstract. This paper proposes a novel hybrid approach for stock price prediction. Stock price prediction implies telling whether a particular stock price will go up or down on a certain day. The model proposed in this paper uses Gated Recurrent Units to take into account the sequential nature of days, which is not handled by CNNs. However, the CNN is used to capture relationships between long term and short term stock patterns. Finally we have a simple feedforward neural network which learns the weights to be given to the CNN prediction and the weight to be given to the GRU prediction.

Keywords: Gated Recurrent Units · Convolutional Neural Network · Stock Price Prediction.

1 Introduction

The problem of stock price prediction has been of interest since long. Not only does it provide scope for research, there is a lot of profit that can be generated by coming up with an accurate model to solve the problem. However, given the random nature of stock price changes it is hard to predict whether a stock price will go up or down on any given day. Handwritten rules are therefore not a good solution to the problem. Deep learning can come to our rescue here by learning hidden features in data.

So, in this paper a deep learning model is proposed for stock price prediction. The model uses Gated Recurrent Units which work on sequential input and predict the next item in the sequence. This is ideal for our use case because what we effectively want to do is to predict the stock price on a given day provided that we have the stock price on previous days. But the prediction made by the GRU is only one part of the model. We use a deep CNN to make another prediction. CNNs have been used as our second model choice because even though CNNs don't work on sequential data, they can be used to capture relationships between stock price data over several days. This is done by pre-processing the data in the form of a matrix. We use news events for building the matrix where each row of the matrix represents all the news events that occurred on that day. This CNN approach for event-driven stock price prediction has been borrowed from previous work in IJCAI 2015[1].

Finally, we need to assign weights to the predictions made by the GRU and the CNN. However, instead of assigning weights manually, we learn those weights. To learn these weights, we use a standard feed-forward neural network.

2 Motivation

The major motivation for solving the stock price prediction problem is to make profit. In simple words, if one can tell whether a particular stock price is going to go up or down on a particular day then he can make the decision which is the most profitable. Profit can be earned not only by individuals but also by organisations through a model which solves this problem.

Another motivation is that this problem belongs to a class of general prediction problems. Solving this problem can help to reach to solutions for other similar prediction problems. For example, weather prediction is a problem which is very similar to stock price prediction problem. The input data consists of the weather information of days of let's say the past month and the desired output is the prediction for the weather on the current day. It can easily be seen how a model for stock price prediction can be transferred to the weather prediction problem.

3 Literature Review

One of the best works for stock price prediction was accepted at IJACI 2015 titled 'Deep learning for event-driven stock price prediction' [1]. The authors' major contributions were proposing a neural tensor network for learning event embeddings and then proposing a deep convolutional neural network to capture the effect of long term events on stock prices.

Another one of the more analytic and comparison based works was accepted recently at Expert Systems with Applications(Elsevier 2017) title 'Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies' [2]. The authors have used unsupervised feature extraction methods like Principal Component Analysis, Autoencoder, Restricted Boltzmann Machine etc along with covariance estimation to analyse how deep learning models can be made better for stock price prediction.

4 Methodology

4.1 Data Pre-Processing for GRU

Firstly, we need to pre-process the data in the form of sequences. We can choose any fixed length for input sequences, say past 30 day stock prices. Then for each sequence of 30 day stock prices we need to give the model a desired target for training. This target is simply the stock price on the 31st day.

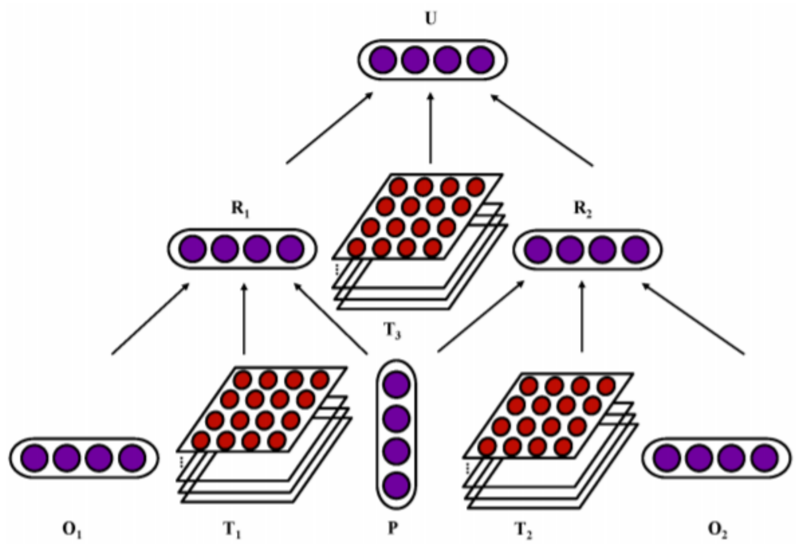

Fig. 1. A figure showing the working of the neural tensor network[1].

4.2 Data Pre-Processing for CNN

The idea of event embeddings has been borrowed from recent work in IJCAI[1]. To give an example, consider an event 'Microsoft sues Barnes and Noble'. We need to represent this event so that it can be used as input for a CNN. One option is to simple use a bag of words represntation i.e. Microsoft, sues, Barnes, and, Noble. However, this representation does not capture the fact that Microsoft is the actor, 'sue' is the action and Barnes and Noble is the object or defendant. What we need is a more structured representation.

So the authors in the paper[1] have suggested the use of event-embeddings. These are dense vectors which represent an event.

4.3 Model

GRU The GRU component of the model comprises of alternate GRU and dropout layers. GRU layer are used for predicting the stock price. Whereas the dropout layers are used for generalisation purpose. In simple words, the dropout layers cancels out the effect of some nodes so that the model does not overfit.

CNN The CNN component of the model takes as input a matrix of event embeddings. The model consists of multiple convolutional layers alternating with

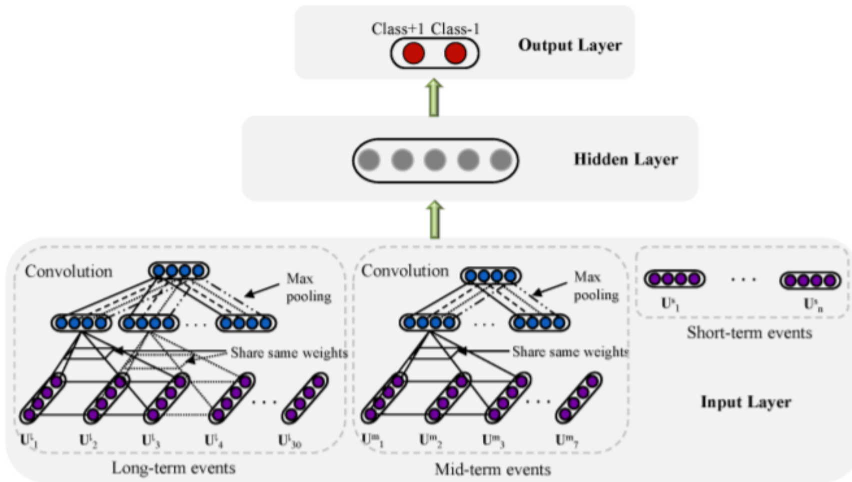

Fig. 2. A figure showing the working of the deep convolutional neural network[1].

max pooling layers. The convolutional layers are used to capture the relationships between events within the window being considered. Then max pooling layers are used to choose the globally most representative features.

Standard Feed-Forward Neural Network The predictions of the GRU and the CNN are used for making the final prediction. However, the best weights for GRU and CNN network are learnt from a simple neural network instead of manual assignment.

References

1. Author, Xiao Ding, Author, Yue Zhang, Author, Ting Liu, Author, Junwen Duan: Deep Learning for Event-Driven Stock Prediction . Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)
2. Author, Eunsuk Chong, Author, Chulwoo Hanb, Author, Frank C.Park: Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications. Elsevier. 15 October 2017.