Stock Market Prediction with Deep Learning: A Character-based Neural Language Model for Event-based Trading

Leonardo dos Santos Pinheiro
Macquarie University
Capital Markets CRC
lpinheiro@cmcrc.com

Mark Dras
Macquarie University
mark.dras@mq.edu.au

Abstract

In the last few years, machine learning has become a very popular tool for analyzing financial text data, with many promising results in stock price forecasting from financial news, a development with implications for the Efficient Markets Hypothesis (EMH) that underpins much economic theory. In this work, we explore recurrent neural networks with character-level language model pre-training for both intraday and interday stock market forecasting. In terms of predicting directional changes in the Standard & Poor's 500 index, both for individual companies and the overall index, we show that this technique is competitive with other state-of-the-art approaches.

1 Introduction

Predicting stock market behavior is an area of strong appeal for both academic researchers and industry practitioners alike, as it is both a challenging task and could lead to increased profits. Predicting stock market behavior from the arrival of new information is an even more interesting area, as economists frequently test it to challenge the Efficient Market Hypothesis (EMH) (Malkiel, 2003): a strict form of the EMH holds that any news is incorporated into prices without delay, while other interpretations hold that incorporation takes place over time.

In practice, the analysis of text data such as news announcements and commentary on events is one major source of market information and is widely used and analyzed by investors (Oberlechner and Hocking, 2004).

Financial news conveys novel information to broad market participants and a fast reaction to the release of new information is an important component of trading strategies (Leinweber and Sisk, 2011).

But despite the great interest, attempts to forecast stock prices from unstructured text data have had limited success and there seems to be much room for improvement. This can be in great part attributed to the difficulty involved in extracting the relevant information from the text. So far most approaches to analyzing financial text data are based on bag-of-words, noun phrase and/or named entity feature extraction combined with manual feature selection, but the capacity of these methods to extract meaningful information from the data is limited as much information about the structure of text is lost in the process.

In recent years, the trend for extracting features from text data has shifted away from manual feature engineering and there has been a resurgence of interest in neural networks due to their power for learning useful representations directly from data (Bengio et al., 2013). Even though deep learning has had great success in learning representations from text data (e.g. Mikolov et al. (2013a), Mikolov et al. (2013b) and Kiros et al. (2015)), successful applications of deep learning in textual analysis of financial news have been few, even though it has been demonstrated that its application to event-driven stock prediction is a promising area of research (Ding et al., 2015).

Finding the most informative representation of the data in a text classification problem is still an open area of research. In the last few years a range of different neural networks architectures have been proposed for text classification, each one with strong results on differ-