Time Series Forecasting with Recurrent Neural Networks

Bachelorarbeit

zur Erlangung des akademischen Grades „Bachelor of Science (B. Sc.)“
im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik,
Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät
der Leibniz Universität Hannover

vorgelegt von

Simon Fallnich
[redacted]

Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 25. August 2017

Abstract

Artificial neural networks are increasingly used in the analysis and forecasting of financial markets. While classic architectures are not suitable for modeling temporal dependencies, recurrent neural networks and especially *Long Short-Term Memory* (LSTM) networks can learn to deal with tasks involving long time lags. The present thesis provides an overview of the structure, functioning and learning algorithms of neural networks and describes the principle of LSTM. Using an LSTM network, forecasts are made for the realized volatility of two different indices and compared with the results of a linear HAR-RV-J model. It is found that in environments with moderate volatility, the LSTM network provides the best results and requires the least amount of input data. In environments with high and strongly fluctuating volatility, however, the results of the LSTM network are slightly worse compared to those of the linear model. In any case, suitable configuration of the network parameters and appropriate transformation of the input values are crucial.

Keywords: Time Series, Forecasting, Recurrent Neural Networks, Realized Volatility, LSTM, HAR-RV-J, Bitcoin, S&P500

Contents

1 Introduction

Study the past if you would define the future.
— Confucius

Predicting financial markets has always been significant. While every market participant makes at least implicit assumptions about future developments, the entire business models of investment banks and hedge funds are based on the correct assessment of price and risk. Key figures like volatility — which plays an important role in asset pricing and risk management — are traditionally predicted with regressive models such as GARCH (see Bollerslev 1986). However, as machine learning methods and artificial neural networks in particular have gained momentum in several domains over the last decades, they are also increasingly applied in the forecasting of (financial) time series. Especially recurrent neural networks are frequently used due to their ability to model temporal dependencies.

Previous papers on forecasting with recurrent neural networks (e.g., Dunis and Huang 2002; Bildirici and Ersin 2009) typically focus on one or more related assets or indices and consistently report good results. Therefore, this thesis aims to assess the predictive performance of recurrent neural networks in different environments and also compare it with the results of a parametric model. For this purpose, forecasts of the *realized volatility*, a measure proposed by Andersen and Bollerslev (1998) that approximates the daily volatility with intraday returns, are made for two different indices. The first selected index is the CoinDesk Bitcoin Price Index, which represents the average Bitcoin price of major exchanges.[1] The price, which is determined by supply and demand on dedicated exchanges, is characterized by strong fluctuations and a high volatility. The second index is the popular Standard & Poor's 500 index, which includes 500 of the largest companies listed in the Unites States and, in contrast, has moderate volatility values. The forecasts

1 Bitcoin is a peer-to-peer system for decentralized transactions introduced by one or more anonymous authors under the pseudonym Nakamoto (2008), which exploits a cryptographic hash function (a function that is hard to invert) to store transactions (practically) unchangeable in a public, distributed database, referred to as *blockchain*, thus eliminating the need for a trusted third party.

are made with a Long Short-Term Memory network, a recurrent neural network with special processing units, and a linear HAR-RV-J model, which is commonly applied to predict realized volatility.

The thesis is structured as follows:

Chapter 2 describes artificial neural networks in detail, their functioning and components as well as different learning algorithms and the Long Short-Term Memory cell in particular.

Chapter 3 gives an overview of the conducted experiments, the data used and the implementation of the Long Short-Term Memory and HAR-RV-J models.

Chapter 4 describes the results of the experiments, discusses them, gives recommendation for action and specifies the limitations of the research.

Chapter 5 summarizes the findings and mentions approaches for further research.

2 Conclusion

The aim of this thesis is to provide a differentiated assessment of the performance of recurrent neural networks in the forecasting of time series. For this purpose, the fundamental properties and learning algorithms of neural networks, especially of recurrent neural networks, which use their feedback connections to model temporal dependencies, are discussed. Furthermore, the Long Short-Term Memory (LSTM) memory cell is presented, which overcomes the problems that occur in the training of conventional recurrent architectures. Using a LSTM network, forecasts are made for the realized volatility of two indices with different volatility characteristics.

Compared to those of a linear HAR-RV-J model, the predictions of the network are among the best results and require the least amount of input data in environments with moderate volatility (Standard & Poor's 500). Both models clearly outperform trivial predictions. However, in environments with high and strongly fluctuating volatility (Bitcoin Price Index), the linear model produces more robust results with a logarithmic transformation of the input data. In both cases, a suitable choice of the network and training parameters as well as the appropriate transformation of input values are of great importance.

Especially forecasting in environments with extreme conditions holds potential for further research. A promising approach is the combination of the predictions of several models, so-called *ensemble learning*, which is not considered, since the evaluation of the individual performance of recurrent networks should be paramount. Another interesting topic is the automatic determination of parameters and input values, referred to as *meta* or *feature learning*, which might bring advantages in comparison to the manual approach used in the experiments. Finally, it would also be interesting to consider entirely different time series, since in the present case, the results suggest that the LSTM network, although it provides good results, cannot fully exploit its essential strength — to learn dependencies over long periods of time.

As it is expected that the use of neural networks will continue to increase in the future, a stronger cooperation between different research domains should be encouraged in order to allow new developments to benefit from the observations of specific applications.