

## Research Article

# A CNN-LSTM-Based Model to Forecast Stock Prices

Wenjie Lu,<sup>1,2</sup> Jiazheng Li,<sup>3</sup> Yifan Li,<sup>3</sup> Aijun Sun,<sup>1</sup> and Jingyang Wang<sup>3</sup>

<sup>1</sup>Business School, Jiangsu Second Normal University, Nanjing 210000, China

<sup>2</sup>School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang 050018, China

<sup>3</sup>School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China

Correspondence should be addressed to Aijun Sun; saj1970@jssnu.edu.cn

DW/hW \$ \$ A UfaTWd \$ " - DW[eW \$ @ahW\_ TWd \$ " - 3UWbfW %@ahW\_ TWd \$ " - BgT1eZW \$ & @ahW\_ TWd \$ " \$ "

Academic Editor: Abd E. I.-Baset Hassanien

Copyright © 2020 Wenjie Lu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Stock price data have the characteristics of time series. At the same time, based on machine learning long short-term memory (LSTM) which has the advantages of analyzing relationships among time series data through its memory function, we propose a forecasting method of stock price based on CNN-LSTM. In the meanwhile, we use MLP, CNN, RNN, LSTM, CNN-RNN, and other forecasting models to predict the stock price one by one. Moreover, the forecasting results of these models are analyzed and compared. The data utilized in this research concern the daily stock prices from July 1, 1991, to August 31, 2020, including 7127 trading days. In terms of historical data, we choose eight features, including opening price, highest price, lowest price, closing price, volume, turnover, ups and downs, and change. Firstly, we adopt CNN to efficiently extract features from the data, which are the items of the previous 10 days. And then, we adopt LSTM to predict the stock price with the extracted feature data. According to the experimental results, the CNN-LSTM can provide a reliable stock price forecasting with the highest prediction accuracy. This forecasting method not only provides a new research idea for stock price forecasting but also provides practical experience for scholars to study financial time series data.

## 1. Introduction

The change trend of the stock price has always been identified as a very important problem in the economic field [1]. Stock prices are affected by various internal and external factors, such as domestic and foreign economic environment, international situation, industry prospect, financial data of listed companies, and stock market operation. Thus, the forecasting method also has different emphasis [2, 3].

The traditional analysis method is based on economics and finance, which mainly uses the fundamental analysis method and the technical analysis method. On the one hand, the fundamental analysis method pays more attention to the intrinsic value of stocks and qualitatively analyzes the external factors that affect the stock, such as interest rate, exchange rate, inflation, industrial policy, finance of listed companies, international relations, and other economic and political factors. On the other hand, the technical analysis method mainly focuses on the direction of stock price, trading volume, and investors' psychological expectation,

which primarily focuses on analyzing the stock index trajectory of individual stocks or the whole market by using K-line chart and other tools. At present, traditional fundamental analysis and technical analysis are still the most commonly employed methods for many organizations and individual investors [4, 5].

The accuracy of the traditional fundamental analysis method is difficult to be convincing. The reason is not only that the influencing factors are in a long-term cycle, but also the forecasting results are more dependent on the professional quality of analysts. As a financial time series, stock data have the characteristics of random walk [6]. Based on statistics and probability theory, some scholars use time series linear forecasting model to predict the short-term stock price with a large number of long-term data, such as vector autoregression (VAR) [7], Bayesian vector autoregression (BVAR) model [8], autoregressive integrated moving average mode (ARIMA) [9], and generalized autoregressive conditional heteroskedasticity model (GARCH) [10]. However, the accuracy of using time series



















