

Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction

**A Seminar Report
(Term Paper)**

Submitted by

ADITHYAN B ASOK

CUAVCMF002



**DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF CALICUT**

2022

Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction

Adithyan b asok
MSc Computer Science
Department of Computer Science
University of Calicut, India
adithyanbasok@gmail.com

Abstract—Stock market plays an important role in the economic development. Due to the complex volatility of the stock market, the research and prediction on the change of the stock price, can avoid the risk for the investors. The traditional time series model ARIMA can not describe the nonlinearity, and can not achieve satisfactory results in the stock prediction. As neural networks are with strong nonlinear generalization ability, this paper proposes an attention-based CNN-LSTM and XGBoost hybrid model to predict the stock price. The model constructed in this paper integrates the time series model, the Convolutional Neural Networks with Attention mechanism, the Long Short-Term Memory network, and XGBoost regressor in a non-linear relationship, and improves the prediction accuracy. The model can fully mine the historical information of the stock market in multiple periods. The stock data is first pre-processed through ARIMA. Then, the deep learning architecture formed in pretraining-finetuning framework is adopted. The pre-training model is the Attention-based CNN-LSTM model based on sequence-to-sequence framework. The model first uses convolution to extract the deep features of the original stock data, and then uses the Long Short-Term Memory networks to mine the long-term time series features. Finally, the XGBoost model is adopted for fine-tuning. The results show that the hybrid model is more effective and the prediction accuracy is relatively high, which can help investors or institutions to make decisions and achieve the purpose of expanding return and avoiding risk. Source code is available at <https://github.com/zshicode/Attention-CLX-stock-prediction>.

Index Terms—*Attention mechanism, Convolutional Neural Networks, Long Short-Term Memory, XGBoost, stock prediction*

I. INTRODUCTION

STOCK market plays an important role in the economic development. Due to the high return characteristics of stocks, the stock market has attracted more and more attention from institutions and investors. However, due to the complex volatility of the stock market, sometimes it will bring huge loss to institutions or investors. Considering the risk of the stock market, the research and prediction on the change of the stock price can avoid the risk for the investors.

The traditional time series model ARIMA can not describe the nonlinear time series, and needs to satisfy many pre conditions before modeling, and can not achieve remarkable results in the stock forecasting. In recent years, with the rapid development of artificial intelligence theory and technology, more and more researchers apply artificial intelligence method to the financial

market. On the other hand, the sequence modeling problem, focusing on natural language sequences, protein sequences, stock price sequences, and so on, is important in the field of artificial intelligence research [8], [13]. The most representative artificial intelligence method is neural networks, which are with strong nonlinear generalization ability.

Recurrent Neural Network (RNN) was adopted for analyzing sequential data via neural network architecture, and Long Short-Term Memory (LSTM) model is the most commonly used RNN. LSTM introduced gate mechanism in RNN, which can be seen as simulation for human memory, that human can remember useful information and forget useless information [6]. Attention Mechanism [7], [16] can be seen as simulation for human attention, that human can pay attention to useful information and ignore useless information. Attention-based Convolutional Neural Networks (ACNN) are widely used for sequence modeling [4], [10]. Combining Attention-based Convolutional Neural Networks and Long Short-Term Memory, is a self-attention based sequence-to-sequence (seq2seq) [15] model to encode and decode sequential data. This model can solve long-term dependency problem in LSTM, hence, it can better model long sequences. LSTM can capture particular long-distance correspondence that fits the structure of LSTM itself, while ACNN can capture both local and global correspondence. Therefore, this architecture is more flexible and robust.

Transformer [16] is the most successful sequential learning self-attention based model. Experiments on natural language processing demonstrates that Transformer can better model long sequences. Bidirectional Encoder Representation Transformer (BERT) with pretraining [2] can perform better than the basic Transformer. Pretraining is a method to significantly improve the performance of Transformer (BERT).

This paper proposes a hybrid deep learning model to predict the stock price. Different from the traditional hybrid prediction model, the proposed model integrates the time series model ARIMA and the neural networks in a non-linear relationship, which combines the advantages of the two vanilla models, and improves the prediction accuracy. The stock data is first preprocessed through ARIMA. The stock sequence is put into neural networks (NN) or XGBoost after preprocessing via ARIMA($p=2, q=0, d=1$). Then, the deep learning architecture formed in pretraining-finetuning framework [2], [5] is adopted. The pre-training model is the Attention-based CNN-LSTM