Application of Regularized GRU-LSTM Model in Stock Price Prediction

Introduction

The stock market is a highly complex nonlinear movement system, and its fluctuation law is affected by many factors, so the prediction of the stock price index is a very challenging task. In spite of high risk, some investors and institutions still choose the stock market to invest. Therefore, the stock price index prediction has attracted the attention of both private and institution investors. Recurrent Neural Network (RNN) have been proved to be one of the most forceful models for processing sequential data, it can recognize complex nonlinear relationships which are difficult to capture using traditional forecasting models. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are the two most satisfactory RNN structures. LSTM network has a longer memory capacity for preserving and processing the previous information, then, for large data, the LSTM network may derive better results. However, GRU is much faster than LSTM since it has fewer parameters. In this paper, LSTM and GRU are combined, and a new Regularized GRU-LSTM network model is proposed with better performance.

Regularized GRU-LSTM Model

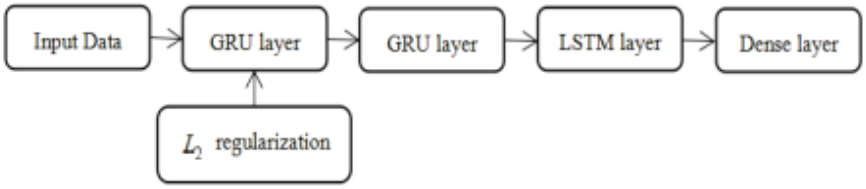
Regularization is a form of regression which can reduce the complexity and instability of the model during the learning process, thereby reducing the risk of model overfitting. The general form of regularization can be described as follows.

$$\min_{\omega}\{\sum_{t=1}^{T} l(y_t, f(x_t, \omega)) + \sum_{i=1}^{m} \lambda_i \rho_i(\omega)\}$$

In which, $l(\cdot,\cdot)$ represents the loss function in the model, T represents the number of samples, $\lambda\rho(\omega)$ is the regularized item, λ is to control the equilibrium relationship between the regular term and the loss function, thus called the regularization tunable parameter. $\rho(\omega)$ has many different expression forms, among which the L1 norm and the L2 norm are commonly used. In general, different norm penalties will cause different model generalization directions. In our model, we selected L2 norm as the regularized item and Mean Absolute Error(mae) as the loss function, the objective function is as follows.

$$\min_{\omega}\{\frac{1}{T}\sum_{i=1}^{T} | f(x_i, \omega) - y_i | + \lambda_2 \| \omega \|_2^2\}$$

The Regularized GRU-LSTM model consists of two GRU layers, one LSTM layer, and finally a dense output layer as shown in Figure below. In each layer, we select the default activation function. Since the LSTM layers are nonlinear, only one dense layer is needed to accumulate their outputs.



Conclusion

A new regularized GRU-LSTM model is proposed and compared it with the GRU and LSTM network models. The experimental results showed that this model had a better performance in the short-term forecast of the stock closing price. It has not only high prediction accuracy, but also high convergence speed. However, the stock market is a frequently changing dynamic system and the prediction of the stock market is a very complex issue, we should not only take into account the problem of the prediction method itself, but also recognize some characteristics of the stocks, and the network system must be trained many times according to the new data produced by the stock market to adapt to the complicated circumstances. Although the regularized GRU-LSTM model has higher prediction accuracy and convergence speed, its stability is sometimes relatively poor. Therefore, a future enhancement is to find an optimization method to improve the stability of the model.