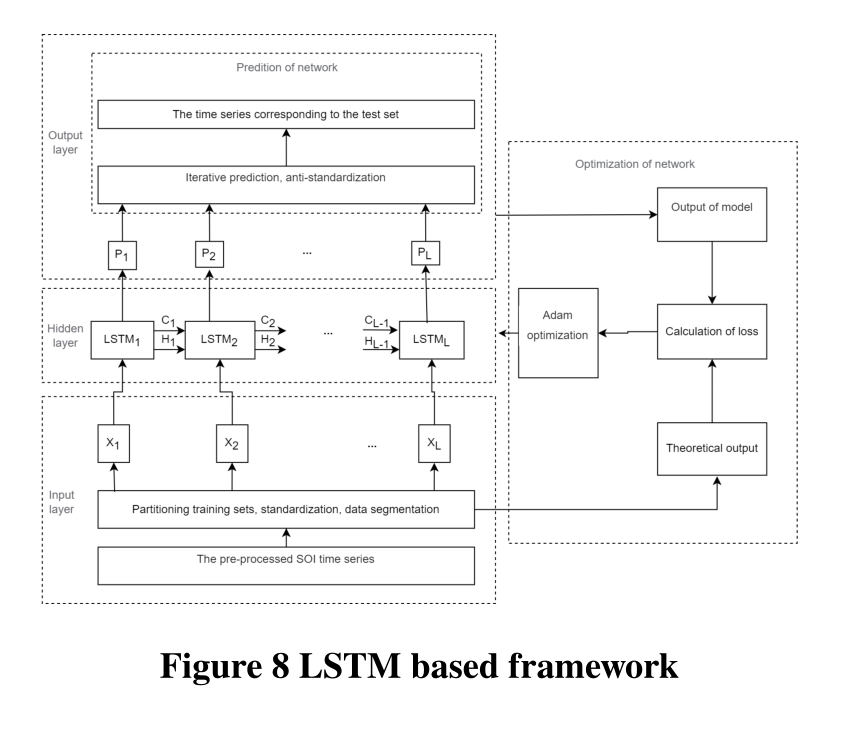
Long Short-Term Memory(LSTM) is a temporal Recurrent Neural Network (RNN), which

is suitable for processing and predicting important events with long intervals and delays in

time series.

LSTM改进了RNN大量重复神经网络模块的链式形式，将隐藏层的RNN细胞替换为LSTM细胞，其关键是它的状态。它们只用一些次要的线性交互就能贯穿整个链式结构，避免了长期依赖性问题。LSTM中有一些控制门，通过决定信息通过的多种方式，使得信息能很容易地以不变的形式从中流过，具备了长期记忆能力。

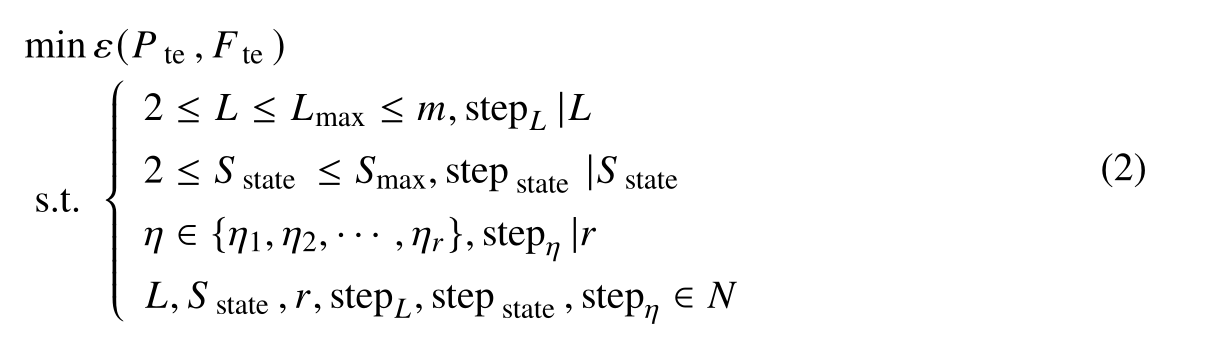
LSTM improves the chain form of a large number of repetitive neural network modules of RNN, replacing the hidden layer of RNN cells with LSTM cells, whose key is its state. The cells use only minor linear interactions throughout the chain structure, avoiding long-term dependency problems. There are some control gates in LSTM, which allow information to easily flow through in the same form by determining the various ways in which it can pass through, providing long-term memory.



The overall framework of the LSTM prediction model constructed in this paper is shown in the figure, including five functional modules: input layer, hidden layer, output layer, network training and network prediction. The input layer is responsible for the preliminary processing of the original number of players to meet the network input requirements; the hidden layer uses LSTM cells to build a single layer of circulating god meridian, with a **sliding window** calculating the statistics in one window to predict the next value and then moving forward; the output layer provides prediction results; the Adam optimization algorithm is used for network training; the network prediction is made point by point by iteration.

Parameter optimization of LSTM prediction model is based on multi-layer grid search. Among all the parameters involved, the segmentation window length L, state vector size Sstate and learning rate μ are the most critical.

We adopt the grid search method to optimize the three parameters, which is based on the highest prediction accuracy at all test points in the test set, that is, the lowest prediction error



First, the number of random seeds and the number of training steps were fixed according to the preset value range of three parameters. Then, the values of the three parameters were traversed respectively, and the LSTM was trained and predicted in the innermost layer, and the corresponding model parameters and model accuracy were saved. Finally, all the saved results are sorted according to the prediction accuracy from high to low, and the first parameter combination is the preferred model parameter.

According to the optimal parameters, the number of hidden units and training times are

determined to be 200 and 3000 respectively, the L is 5, the Sstate is …? and the error of the test set is minimum