**Long Short-Term Memory Network**

Long Short-Term Memory Network (LSTM) is a Neural Network to deal with sequence input which can fits our scenario of series data. It has been successfully applied to many sequence learning problems and has been considered as one of the latest methods to deal with series forecasting problems.

LSTM is an improved method of simple Recurrent Neural Network (RNN) that can solve the problem that RNN cannot handle long-distance dependencies of one sequence. The hidden layer at each moment in its structure contains a memory cell, and each memory contains multiple storage units. Each storage unit contains a memory unit and 3 gates (input gate, forget gate and output gate). The input gate decides whether to let new input, the forget gate deletes unimportant information, and the output gate decides what information to output.

There are 4 inputs and 1 output in the LSTM unit. The 4 inputs are the control signals of the three gates, the input gate, forget gate, output gate with input data. After inputting the data and hidden information from previous unit , forget gate , input gate and output gate will be computed using the inner parameters in gates (equation 1, 2, and 3). 𝑠𝑖𝑔𝑚𝑜𝑖𝑑 function is used to regulate the opening and closing degree of the gate between 0 and 1. The last retained data is stored in the memory unit, and it will be updated to through the forget gate and input gate. The element wise product between and and element wise product between and standard RNN cell, shown in equation (4).

After obtaining , the value of c is updated in the memory unit, and at the same time, it passes through a tanh activation function, multiply by the output gate to get the hidden unit , shown in equation (5).

Multiple LSTM units can be connected to form a whole network comprehensively and we can have our ultimate LSTM Neural Network.

**Experiment Result**

Figure 1 demonstrates the development of training loss. In the first 10 epochs, the training loss drops significantly. It becomes steadier in the next 20 epochs, and we believe that it meets its convergence at last 20 epochs.

Figure 2 illustrates the plot between coefficient and epochs. The coefficient increases rapidly in the first 30 epochs, and it fluctuates in the next 40 epochs. Finally, it becomes steady in last 30 epochs.

**图片包含 图形用户界面

描述已自动生成**

Figure 1 Training loss - Epoch Plot

**图表, 折线图

描述已自动生成**

Figure 2 Coefficient - Epoch Plot