**MODEL 3 Overview**

For the distribution of the report of the attempt times of a word, we divide the factors which are affecting it into three types: letter frequency, word frequency and letter repetition times, a total of seven parameters (the letter frequency of each letter should be taken into account), and the target output results are also seven (1,2,3,4,5,6,X). On this basis, we set up BP neural network, optimized by error feedback, and predict the distribution of future report results.

为什么加入重复次数（重复次数考虑的意义）

A new consideration in this model is the number of letters repeated as a parameter. Through our own experience and searching for relevant strategy data, we've noticed that, unless we can't recall another word or are absolutely certain, when considering possible words, players tend to choose words with fewer repetitions (or even zero) in order to eliminate more misinformation. As a result, it often takes more attempts to find the answer when the letter is repeated.

具体建立

BPNN system is applied to our study, and its structure is as follows:

Input layer -- Each input layer of prediction data contains 7 neurons, each representing a parameter describing the attributes of the data (frequency of the first, second, third, fourth, and fifth letters, number of letter repeats, or word frequency).

Hidden Layer -- This layer has one neuron for each case in the training dataset. The neuron stores the values of the predictor variables for the case along with the target value. In this case, there are 10 neurons in this layer.

Output layer -- Through the calculation and fitting of the data, the calculated (1,2,3,4,5,6, X) is derived as the prediction result.

**/picture/**

In the process of establishing the model, we adjusted the parameters and found that the prediction results were better when the number of overfittings was set to 10 and the number of hidden layer neurons was set to 100.

**/picture/(不同数值跑出结果对比（如果有/有必要的话）)**