# 灵敏度分析

# Sensitivity analysis

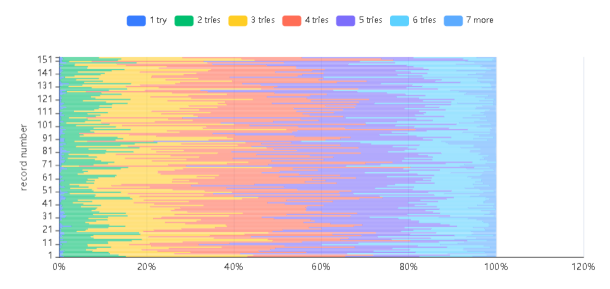
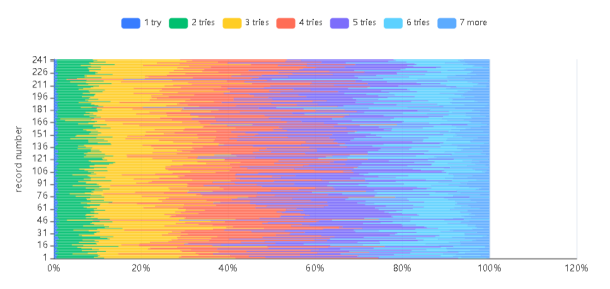
在问题二中，我们运用了神经网络以及两种不同内置算法来对比预测给定方案词的百分比分布。

诸多因素如训练集比例的划分，隐藏神经元的层数和数量都会影响巡检结果和预测结果。  
我们调整二者进行结果观测，观察这些参数对模型的影响。

In Problem 2, we applied a neural network and two different built-in algorithms to predict the percentage distribution of words for a given scheme in comparison.

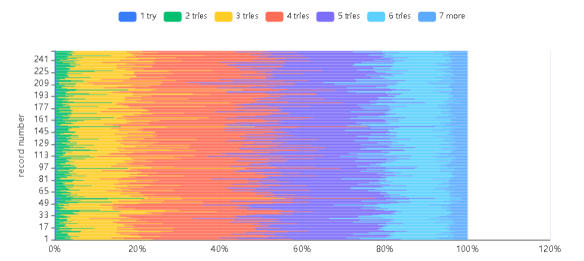
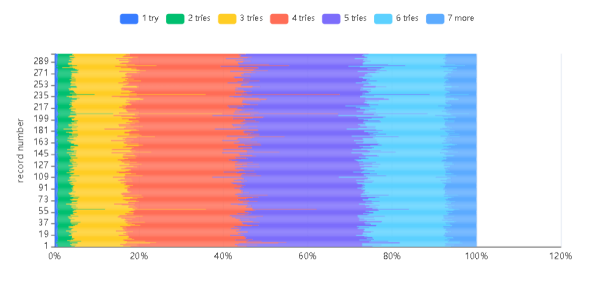
Many factors such as the division of training set proportions, the number of layers and the number of hidden neurons affect the inspection results and the prediction results.

We adjusted both for result observation and observed the effect of these parameters on the model.

LM划分数据集划分50：25：25 BR划分50：50

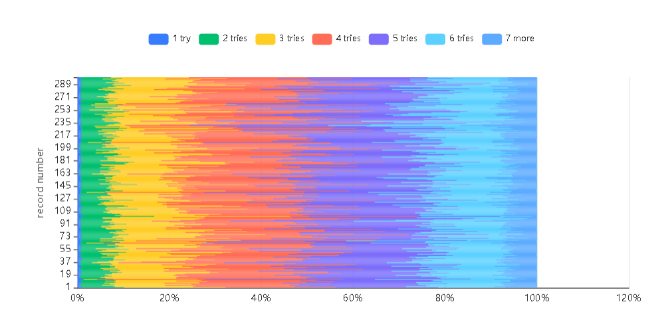
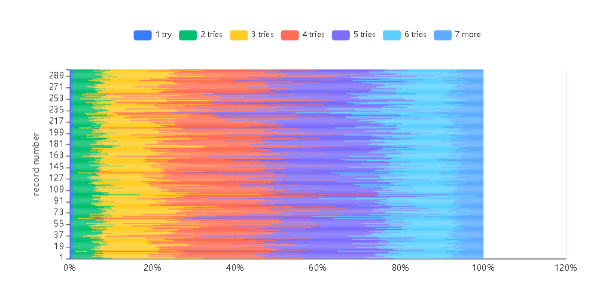
LM division data set division 50:25:25 BR division 50:50

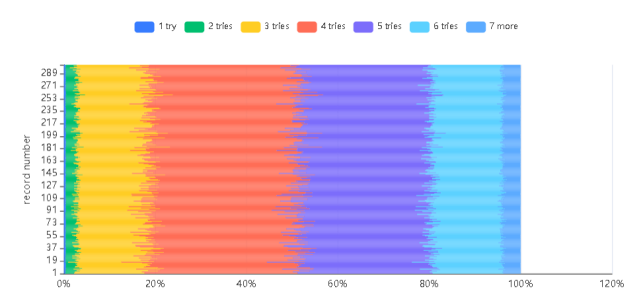
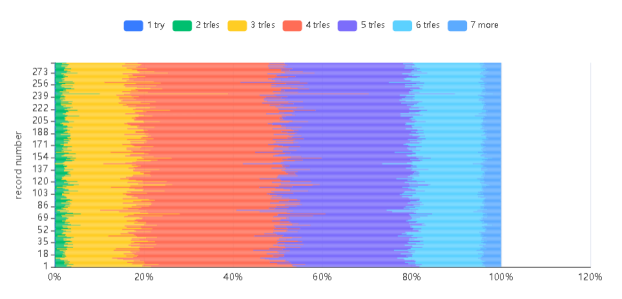
LM division data set division 90:5:5 BR division 90:10

通过上述堆积条形图，可以看出当减小训练集到50%时,两种算法的性能均有所下降。我们认为其原因是因为训练集样本的局限性导致神经网络训练效果不理想。而增大训练集比例到90%时，得到的结果和task2中得到的结果非常接近，而且比较稳定。这佐证了上述想法。并且说明了我们task数据集划分的合理性。

The above stacked bar chart shows that the performance of both algorithms decreases when the training set is reduced to 50%. We believe that the reason for this is the limitation of the training set samples which leads to the unsatisfactory training of the neural network. When the training set is increased to 90%, the results are very similar to those obtained in task2 and are more stable. This confirms the above idea. And it shows the reasonableness of our task data set division.

10个隐藏神经元LM 10个隐藏神经元BR

10 hidden neurons LM 10 hidden neurons BR



2个隐藏神经元LM 2个隐藏神经元BR

2 hidden neurons LM 2 hidden neurons BR

从上述堆积条形图，可以看出增加隐藏神经元到10个时，和减少到两个时。神经网络的预测性能均比较稳定。隐藏神经元数量在2-10个间具有相当的稳定性。

From the stacked bar graph above, it can be seen that when increasing the hidden neurons to 10, and when decreasing to two. The prediction performance of the neural network is relatively stable both. The number of hidden neurons is fairly stable between 2 and 10.