

- [report.pdf](#)为实验报告
- prediction_2.txt为2-gram预测结果
- prediction_3.txt为3-gram预测结果
- prediction_LSTM.txt为LSTM预测结果
- src为实验代码文件以及中间数据文件
 - word_clean文件夹为1000个txt去噪之后以空格分隔的文本
 - word_train文件夹为1000个转换为序号的文本，每行为一个list便于读入
 - All_words_dict.txt为所有出现的词以及对于的序号
 - questions_num.txt为questions.txt转换为对应序号文件
 - answer_num.txt为answer.txt转换为对应序号文件
 - list.txt为新闻链接的txt文件
 - get_news.py为获取新闻数据的程序
 - clean.py为对数据进行预处理的程序
 - n-gram.py为n-gram程序文件以及其中有测试预测正确率的函数
 - lstm_data_process.py为进行LSTM模型构建前的数据处理程序
 - LSTM.py为LSTM模型构建以及训练程序