1. 路透社21578数据中，4400条数据为训练集，1600条数据为测试集。
2. 之前的数据选择了2000维的变量，数据集中的变量是按照词频进行排序的，1000维之后的变量为十分稀疏的矩阵，在优化后的模型中输入变量为1000。优化后的模型大小为1000-700-100-10。
3. test\_example\_DBN是主文件，可以直接执行。
4. 文件minFunc是一个函数集合，通过它调用了lbfgs。