

## 构成

- Main

主函数，通过配置文件配置文件目录，控制两个Mapper-Reducer的操作以及临时文件的存储和交换。

- FormalPossibility

通过一个Mapper-Reducer结构匹配每一行标签和词汇，统计输入情感词汇的数目，输出到HDFS上面的临时文件。其中Mapper完成匹配，Reducer完成统计任务。

- FPReader

通过输入配置文件目录，读取临时文件，解析成对于词数，情感数，词汇情感结合数目的统计。

- NBPredictor

在Setup中初始化FPReader，读取临时文件次数统计

在Mapper里面使用平滑后贝叶斯公式进行情感预测，输出行号与预测结果

在Reducer里面通过行号进行排序，并输出

## 编译执行过程

Train.txt, test.txt放在java文件相同目录。

```
javac -d . *.java # build class files
jar -cvf NB.jar *.class # pack jar
rm *.class # clear class
hadoop jar NB.jar Main /NBinput train.txt text.txt result # run
cat result/part-r-* # show result
```