## 垃圾邮件分析

已知文件SMSSpamCollection，数据格式如下：

ham Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...  
ham Ok lar... Joking wif u oni...  
ham U dun say so early hor... U c already then say...

每条记录的第一个单词是ham(非垃圾)和spam（垃圾），用于表明该条记录是否为垃圾消息。

请通过贝叶斯算法，根据学习源数据SMSSpamCollection，实现后期输入一封邮件，可以判断该条消息是否为垃圾消息。

性能要求：结果输出时间小于1秒。

将该项功能结合到SSM项目中

提交：

- 项目代码

- 数据库sql

- 文档（参考文档模板）

文档中体现：

* 算法介绍
* Map reduce 阶段的实现思路
* 架构图
* 部署说明
* 业务流程图