**赛题说明：**

赛题为2分类，垃圾短信和非垃圾短信区分，隶属自然语言处理

**文件说明：**

reademe：赛题说明文件

train：训练数据集和标签。

test：测试数据集 共200条

test-lable:测试数据集结果

**数据格式：**

txt

ham：非垃圾短信

spam：垃圾短信

**数据样例：**

ham Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...

spam Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entry question(std txt rate)T&C's apply 08452810075over18's

前一个为正确类别，后一个为短信内容（不带附件）