

# 开题报告

## 项目背景

俗语云“未见其人，先闻其实”，说明每个人的声音会像指纹那样，是有固定的“声纹”，而即使不认识的人，也能通过声音判别出是男是女，说明是性别的声纹是有固定的模式判别，而此项目则是让机器能辩雌雄。

## 问题描述

性别语音识别是一个分类问题。通过对音频信号提取出有可能关联的特征，训练出识别模型。

## 输入数据

数据集包含 3168 个样本，其中 50%位男性，50%位女星。数据集包含以下特征：

meanfreq: 频率平均值 (in kHz)

sd: 频率标准差

median: 频率中位数 (in kHz)

Q25: 频率第一四分位数 (in kHz)

Q75: 频率第三四分位数 (in kHz)

IQR: 频率四分位数间距 (in kHz)

skew: [频谱偏度](#)

kurt: [频谱峰度](#)

sp.ent: 频谱熵

sfm: [频谱平坦度](#)

mode: 频率众数

centroid: [频谱质心](#)

peakf: 峰值频率

meanfun: 平均基音频率

minfun: 最小基音频率

maxfun: 最大基音频率

meandom: 平均主频

mindom: 最小主频

maxdom: 最大主频

dfrange: 主频范围

modindx: 累积相邻两帧绝对基频频差除以频率范围

label: 男性或者女性

## 解决办法

问题为二分类问题，对数据进行数据探索，分别使用随机森林、XGBoost 模型，比较得分。

## 基准模型

决策树模型作为基准模型。

## 评估指标

分类问题，评估指标使用 ROC 曲线和 P-R 曲线进行模型的评估。

## 设计大纲

数据探索->模型训练->得分比较