**人工智能实验（3）**

**基于支持向量机的嘴唇运动识别**

**任务：**检测视频中的每个人、每一帧是否在说话。

**原始数据：**75条视频，每个视频中有若干人，他们有些在说话，有些不在说话。

**我们处理后的数据：**training.data包含训练数据。它来自我们的项目，用于检测视频中的人是否说话。特点如下：

1. 基于面部特征检测从原始图像中获取嘴部区域M。
2. 计算上一帧的嘴部区域和当前帧之间的光流（optical flow），并生成描述嘴部运动的分数S。
3. 计算表示张口程度的参数V。
4. 对于第i帧，我们还要计算其前一帧和后一帧的S和V。
5. 因此，我们生成一个6维特征向量为X=[Si-1 Si Si+1 Vi-1 Vi Vi+1]。
6. 标签在每一行的末尾，其中+1代表说话，-1代表不说话。

traning.data数据示例：

0.65307 0.9135 0.99926 0.18748 0.14696 0.16105 -1

0.95969 2.0065 3.3661 0.27832 0.25881 0.18512 1

0.9135 0.99926 1.362 0.14696 0.16105 0.19826 -1

2.0065 3.3661 2.8797 0.25881 0.18512 0.18681 1

0.99926 1.362 1.5848 0.16105 0.19826 0.34915 -1

3.3661 2.8797 2.8833 0.18512 0.18681 0.24276 1

1.362 1.5848 1.4066 0.19826 0.34915 0.30785 -1

2.8797 2.8833 1.7312 0.18681 0.24276 0.24073 1

。。。

**作业要求：**

1. 使用training.data，构建SVM分类器。
2. 打开testingNoAnswer.data文件，利用（1）中的分类器，预测每一行的类别是+1还是-1。

testingNoAnswer.data数据示例

0.60796 1.1404 0.45322 0.37019 0.36954 0.37602

0.69282 0.51609 0.25678 0.25269 0.25603 0.24682

1.1404 0.45322 1.526 0.36954 0.37602 0.39615

0.51609 0.25678 0.21723 0.25603 0.24682 0.23132

0.45322 1.526 1.5283 0.37602 0.39615 0.49313

0.25678 0.21723 0.2792 0.24682 0.23132 0.23871

1.526 1.5283 2.5519 0.39615 0.49313 0.46897

1. 将预测结果写入到submit.txt中，其格式请参考给出的submit sample.txt
2. 请提交你写的PYTHON程序，需要加上注释让助教能看懂。
3. 请提交你的预测结果submit.txt。
4. 将（4）和（5）的材料打包提交。

**评分标准：**

我们会根据你程序的清晰性打出程序分，我们会根据你的预测结果打出测试分，最后你的总分为60%程序分+40%测试分。

注意事项：

（1）这是我们研究中遇到的真实问题，这个研究发表在：

Mai Xu, Yufan Liu, **Haoji (Roland) Hu**, Fei He, Find who to look at: Turning from action to saliency, IEEE Transactions on Image Processing 27 (9), 4529-4544, 2018.

有兴趣的同学可以看一下，了解更多的内容。

（2）如果有感兴趣的同学想自己从头做，可以找我要75个视频自己做。