|  |
| --- |
| 西南民族大学学生实验报告  教学单位：计科学院 实验室名称：BS226 实验时间：2019年10月27日 |
| 实验项目名称：实验4—SVM 算法 实验成绩： 教师签名： |
| 实验项目报告内容（1、实验目的；2、实验主要过程与结果；3、分析讨论）。  **一、实验目的**  学会使用支持向量机进行两类手写数字的识别。  **二、实验主要过程与结果**   1. SMO高效优化算法    1. 应用简化版SMO算法处理小规模数据集   定义loadDataSet()函数，打开文件并对其进行逐行解析，从而得到每行的类标签和整个数据矩阵，定义selectJrand(i, m)函数，若输入的函数值不等于输入值i，函数就会进行随机选择；定义clipAlpha(aj, H, L)函数，调整大于H或小于L的alpha值。函数的代码如图1所示。运行命令，测试函数，如图2所示。    图 1    图 2  定义简化版的SMO算法smoSimple(dataMatIn, classLabels, C, toler, maxIter)，程序清单如图3、图4、图5所示，运行效果如图6、图7、图8、图9所示。    图 3    图 4    图 5    图 6    图 7    图 8    图 9   1. 利用完整Platt SMO算法加速优化   构建一个仅包含init方法的opStruct类，将数值保存在数据结构中，函数清单如图10所示；构建calcEk(oS, k)函数，计算E值并返回，函数清单如图11所示；构建selectJ(i, oS, Ei)函数，选择第二个alpha（内循环的alpha）的值，函数清单如图12所示；构建updataEk(oS, k)函数，计算误差并存入缓存，函数清单如图13所示。    图 10    图 11    图 12    图 13  构建innerL(i, oS)函数，用自己的数据结构在oS传递，并用selectJ()来选择第二个alpha的值，在alpha值改变时更新Ecache，函数代码如图14、15所示。    图 14    图 15  构建函数smoP(dataMatIn, classLabels, C, toler, maxIter, kTup=('lin', 0))，程序清单如图16所示。在控制台中部分运行结果如图17所示。    图 16    图 17  定义calcWs(alphas, dataArr, classLabels)函数计算w，函数代码如图18所示。执行结果如图19、图20所示。    图 18    图 19   1. 在复杂数据上应用核函数    1. 径向基核函数   定义核转换函数，函数清单如图20所示，修改optStruct类，引入一个新的变量kTup，程序清单如图21所示。对innerL()和calcEk()进行修改。    图 20    图 21   * 1. 在测试中使用核函数   构建一个测试函数，利用核函数进行分类，程序清单如图22所示。执行结果如图23所示。    图 22    图 23   1. 示例：手写识别问题回顾   构建loadImages(dirName)函数和testDigits(kTup=('rbf', 10))函数，函数清单如图24、图25所示。执行结果（部分）如图26所示。    图 24    图 25    图 26  **三、分析讨论**   1. 在本质上，支持向量机是一个二类分类器，分类结果不是+1就是-1。 2. SMO算法每次只优化两个alpha值来加快SVM的训练速度，而核方法将数据从一个低维空间映射到一个高维空间，将一个在低维空间中的非线性问题转换成高维空间下的线性问题进行求解。 |

注：实验报告的内容及格式可由学院根据学科专业特点确定；全校各专业必须使用学校统一封面。