# Lab 2 Supervise 实验报告

何春望 PB17000075

实现了k近邻、支持向量机、决策树三种模型。支持向量机实现了线性核、二次核和高斯核；决策树同样使用二等级制预测第三期成绩。各模型参数为k=7, C=0.1, sigma=0.5，两组数据多次测试平均结果分别如下：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| mat | kNN | linear SVM | poly SVM | rbf SVM | Decision Tree |
| 是 | 90.38% | 94.69% | 83.39% | 78.86% | 88.10% |
| 否 | 76.31% | 80.37% | 71.36% | 80.85% | 70.37% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| por | kNN | linear SVM | poly SVM | rbf SVM | Decision Tree |
| 是 | 94.47% | 95.37% | 93.26% | 91.98% | 92.74% |
| 否 | 91.67% | 91.62% | 88.71% | 92.17% | 87.06% |

可以看出，kNN和决策树和线性SVM均能很好的预测结果，这是由于数据本身线性性质决定的。同样因此，多项式核和高斯核的SVM模型分类效果反而略有下降。

另外，por数据集的预测结果普遍优于mat数据集，这可能是由于科目性质不同造成的。而考虑前两阶段成绩时的预测效果普遍高于不考虑的，这是由第三期成绩与前两期强相关决定的。