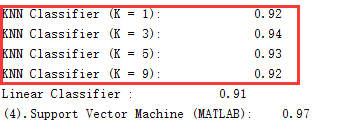
## 实验结果分析

1. KNN

在使用KNN算法对数据进行二分类时，使用了实验二中KNN分类的函数。可能由于算法不够优化，而这次的训练集和样本集数据量庞大，导致运行速度很慢，需要1-2分钟的时间完成分类，分类结果的准确度如下：

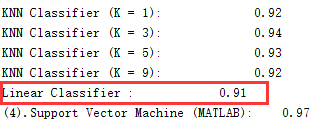


分别取K值1,3,5,9进行分类，后面对应着分类的准确率，可见当K=3时，分类的准确率最高。

因此，K的选取对分类准确率存在的较大影响，虽然在本次实验中并没有较为明显的体现，但是K值的选取是要适中的，不宜过小，也不宜过大。K值过小会造成某些点的噪声过大，K值过大则会造成无关点对分类结果的影响。

1. 线性分类

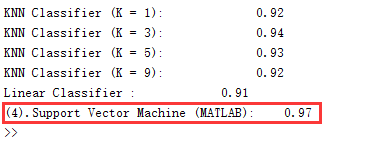
线性分类主要使用了Fisher 线性分类器，直接调用matlab库函数实现，分类准确率如下：



从结果可以看出，分类的准确率之哟偶0.91，比KNN的每一个结果都要低。

我们知道，线性分类器的原理是构造一个超平面，位于平面两侧的数据分别为不同的两类。这样的方法比较简单快速，但是也同样意味着超平面的选取十分重要，而一旦选取，分类结果将不可变化，而超平面也不可避免地会对一些距离超平面十分近的数据点进行错误分类，因此准确率较低。

1. SVM分类



SVM分类可以直接调用matlab库函数实现。

从实验结果来看，SVM分类的精度高达97%。

SVM分类的理论依据是：找到一个分类面，使两类的点到此面的最小距离最大，相较于超平面可能会将靠近超平面的点错误分类，大大提高了分类的准确率。