**SVM算法和K-means算法比较**

**问题：**对不同算法进行数据分类比较

分类技术在很多领域都有应用，例如可以通过客户分类构造一个分类模型来对银行贷款进行风险评估；当前的市场营销中很重要的一个特点是强调客户细分。客户类别 分析的功能也在于此，采用数据挖掘中的分类技术，可以将客户分成不同的类别，比如呼叫中心设计时可以分为：呼叫频繁的客户、偶然大量呼叫的客户、稳定呼叫 的客户、其他，帮助呼叫中心寻找出这些不同种类客户之间的特征，这样的分类模型可以让用户了解不同行为类别客户的分布特征；其他分类应用如文献检索和搜索 引擎中的自动文本分类技术；安全领域有基于分类技术的入侵检测等等。所以对数据分类并挖掘是一个值得研究的问题。

# 数据挖掘的任务：对于相同的数据集，使用不同的分类规则，比较分类的结果更加的精准和有效。

# 数据集：iris.data.set标准数据集

<http://kdd.ics.uci.edu/summary.data.type.html>

**算法及实验平台：**K-means分类算法、SVM

使用数据集，用以上2种不同的分类方法来进行实验，从而比较分类算法的性能。实验平台是MATLAB和C++。

**评估和比较：**准确率、错误率

拟从实验结果来评价算法的性能。

**完成计划时间表：**

|  |  |
| --- | --- |
| 时间 | 进度 |
| 2016.4.1-2016.4.15 | 完成数据库的下载，查找好相关资料 |
| 2016.4.16-2016.5.15 | 完成k-means,SVM代码编写，写进度报告1 |
| 2016.5.16-2016.5.31 | 做测试实验，记录实验结果，写进度报告2 |
| 最后两周 | 做最后的实验结果的分析，完成最后的报告 |