**SVM的基本原理**可以简述如下：

1.基本思想：ＳＶＭ的学习目的主要为求解能够正确划分训练数据集并且几何间隔最大的分离超平面。SVM的基本思想是尝试找到一个使不同类别样本到达超平面距离最大的一个位置。

2.支持向量：决定决策边界的数据叫做支持向量。会影响超平面的位置和宽度。

3.线性可分情况：当存在一个超平面完全将两类样本分开时称为线性可分问题。SVM的目标是找到一个使不同类别样本到达超平面距离最大的一个位置。而线性不可分的问题一定要转换为线性可分来求解。

4.间隔最大化：为求间隔最大化，SVM设立出优化目标。首先要从无数分类中找出最优分类，最大化间隔。这个最优可以通过求解一个二次规划，即非线性规划问题来得到。

5.转化为对偶问题：为了求解二次规划问题，将其转化为对偶问题来求解。对偶函数是凹函数，对凹函数求最大相当于凸函数求最小，也即对偶问题是凸优化问题

6.核函数：核函数是一系列函数的统称，可以将数据映射到高维空间中，解决其在某些维度中线性不可分的问题。常用的核函数有线性核、多项式核、高斯核以及sigmoid核。

7.软间隔：通过对每一个样本点引入一个松弛变量，允许一些样本点落在超平面的另外一侧。对于那些不满足约束条件的样本点，使得函数间隔加上松弛变量之后大于等于1

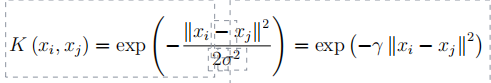
8.多类别分类：对于多类别分类问题，可以使用一对多（依次把某个类别的样本归为一类,其他剩余的样本归为另一类）或一对一（在任意两类样本之间设计一个SVM，因此k个类别的样本就需要设计k(k-1)/2个SVM）的方法来组合多个二分类SVM模型。

**SVM效果分析**

**我们将通过校赛A题来第三问对SVM来进行效果分析。根据题目的要求，我们选择建立SVM多分类模型。**

1.数据预处理：对数据进行预处理，包括数据清洗、缺失值处理、特征选择等，确保数据的质量和可用性。

2.数据划分：将数据集划分为训练集和测试集。因为A题数据给的并不是很多，所以我们确定训练集和测试集的比为8：2

3.模型训练：使用训练集对SVM模型进行训练，调整模型的超参数，如核函数类型、正则化参数等。而对于核函数的选择，本文选择高斯核函数进行求解，其定义公式如下

对于高斯核函数，其可以反映出样本两点之间的相似度大小。当σ确定后，若两点之间距离越小，则相似度趋近于1；若距离越大，则相似度趋近于0。

4.模型评估：使用测试集对训练好的SVM模型进行评估，可以使用一些常见的评估指标，如准确率、精确率、召回率、F1值等。将测试集导入训练完后的模型，我们计算了其准确率为94.74%。

5.**结果可视化**：可以通过绘制ROC曲线、混淆矩阵等方式对模型效果进行可视化展示。