# 概述

支持向量机(support vector machines ：SVM）是一种二分类模型。它是定义在特征空间上的、间隔最大的线性分类器。

间隔最大使得支持向量机有别于感知机。

如果数据集是线性可分的，那么感知机获得的模型可能有很多个，而支持向量机选择的是间隔最大的那一个。

支持向量机还支持核技巧，从而使它成为实质上的非线性分类器。

SVM的分类思想本质上和线性回归LR分类方法类似，就是求出一组权重系数，在线性表示之后可以分类。我们先使用一组trainging set来训练SVM中的权重系数，然后可以对testingset进行分类。

SVM就是先训练出一个分割超平面separation hyperplane, 然后该平面就是分类的决策边界，分在平面两边的就是两类。显然，经典的SVM算法只适用于两类分类问题，当然，经过改进之后，SVM也可以适用于多类分类问题。

我们希望找到离分隔超平面最近的点，确保它们离分隔面的距离尽可能远。这里点到分隔面的距离被称为间隔margin，我们希望这个margin尽可能的大。支持向量support vector就是离分隔超平面最近的那些点，我们要最大化支持向量到分隔面的距离。

# 线性可分支持向量机

# 线性支持向量机

# 非线性支持向量机

# 支持向量回归

# SVDD

# 序列最小最优化方法

# 特点

## 优点

有严格的数学理论支持，可解释性强。

能找出对任务至关重要的关键样本（即：支持向量）。

采用核技巧之后，可以处理非线性分类/回归任务。

## 缺点

训练时间长。当采用SMO 算法时，由于每次都需要挑选一对参数，因此时间复杂度为O(N2)，其中N为α的长度，也就是训练样本的数量。

当采用核技巧时，如果需要存储核矩阵，则空间复杂度为O(N2)。

模型预测时，预测时间与支持向量的个数成正比。当支持向量的数量较大时，预测计算复杂度较高。

# 应用

因此支持向量机目前只适合小批量样本的任务，无法适应百万甚至上亿样本的任务。