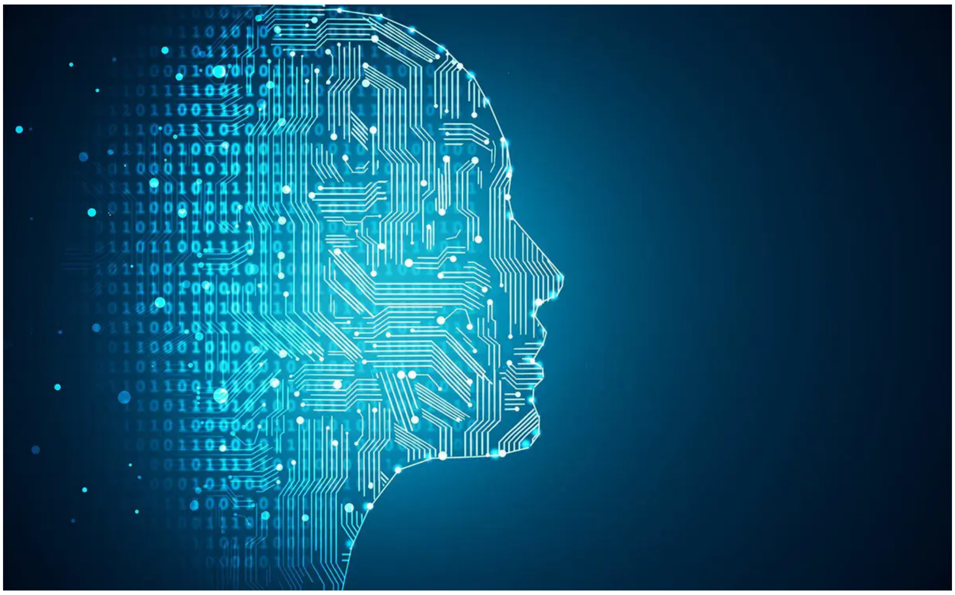
烧脑的 SVM 推导

zidea 关注



机器学习基础

什么是 SVM 算法

- 二元线性分类问题(简单)
  - 可分问题
  - 什么样线性方程是最好线性的方程，离这条子线最近那些点离这条线最远，这也是 SVM 的目标
  - 有很多判别线
  - 支持向量与我们直线最近那些点(向量)就是支持向量
- 回忆解析几何，点到直线的距离
- 点 (x,y) 到 $Ax+By+C=0$ 的距离

$$\frac{|Ax+By+C|}{\sqrt{A^2+B^2}}$$

- 扩展到 n 维空间 $\theta^T x_b = 0 \Rightarrow w^T + b = 0$

$$\frac{|w^T+b|}{||w||} ||w|| = \sqrt{w_1^2 + w_2^2 \cdots w_i^2}$$

我们了解到了如何在 n 维空间进行求解点到线或平面的距离后，我么知道所有点到平面的距离都应该大于支持向量到平面距离，然后接下来我们再尝试用数学方式把这些思想表达出来。

这里对于分类问题使用 1 和 -1 表示两类事物，而非 0 和 1。

$$\frac{w^T x^{(i)}+b}{||w||} \leq -d \quad \forall y^{(i)} = -1$$

通过公式不难看出对于任意样本点 $y^i = 1$ 都满足 $\frac{w^T x^{(i)}+b}{||w||} \geq d$

对等式两边分别除以 d 就得到下面不等式

$$\begin{cases} \frac{w^T x^{(i)}+b}{||w||d} \geq & \forall y^{(i)} = 1 \\ \frac{w^T x^{(i)}+b}{||w||d} \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

这里 $||w||$ 是 n 维的向量的模是一个数字，d 也是数，我们可以对 w 和截距 b 同时除以一个数。转换成下面方程

$$\begin{cases} w_d^T x^{(i)} + b_d \geq 1 & \forall y^{(i)} = 1 \\ w_d^T x^{(i)} + b_d \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

那么我们这里方程中有两个未知数 W_d 和 b_d 需要我们求解，这样我们就可以使用 w 和 d 直接进行替换。但是现在使用 w 和 d 和之前 w 和 d 差一个系数关系。

我们在进一步进行推导出，这样我们将两个不等式合并表示为一个不等式。也就是说明我们所有点都要满足这个不等式关系。

$$y^{(i)}(w^T x^{(i)} + b) \geq 1$$

推导到现在我们发现决策边界线可以表达为

$$W_d^T + b = 0$$

而其上下两侧的支持向量的直线可以用 $W_d^T + b = 1$ 和 $W_d^T + b = -1$

对于任意支撑支持向量

$$\max \frac{|w^T x + b|}{||w||} \Rightarrow \max \frac{1}{||w||} \Rightarrow \min ||w|| \Rightarrow \min \frac{1}{2}||w||^2$$

经过一些列推导我们得到最小值，求取最小值也就是我们问题变为可以优化的问题。不过这一切是建立在满足以下不等式基础上

$$s.t. y^{(i)}(w^T x_i + b) \geq 1$$