# 1、简要介绍下SVM

SVM，全称是support vector machine，中文名叫支持向量机。SVM是一个面向数据的分类算法，它的目标是为确定一个分类超平面，从而将不同的数据分隔开。

个人理解：寻找分类之间最小距离的最大值，支持向量决定最小距离，寻找让最小距离尽可能大的分类方式，是寻找让间隙最大的模型，这样的模型容错率高、鲁棒性好。

扩展：支持向量机学习方法包括构建由简至繁的模型：线性可分支持向量机、线性支持向量机及非线性支持向量机。当训练数据线性可分时，通过硬间隔最大化，学习一个线性的分类器，即线性可分支持向量机，又称为硬间隔支持向量机；当训练数据近似线性可分时，通过软间隔最大化，也学习一个线性的分类器，即线性支持向量机，又称为软间隔支持向量机；当训练数据线性不可分时，通过使用核技巧及软间隔最大化，学习非线性支持向量机。

[支持向量机通俗导论（理解SVM的三层境界）](https://www.cnblogs.com/v-July-v/archive/2012/06/01/2539022.html) [机器学习之深入理解SVM](http://blog.csdn.net/sinat_35512245/article/details/54984251)

# 2、介绍下tensorflow的计算图

Tensorflow是一个通过计算图的形式来表述计算的编程系统，计算图也叫数据流图，可以把计算图看做是一种有向图，Tensorflow中的每一个计算都是计算图上的一个节点，而节点之间的边描述了计算之间的依赖关系。

TensorFlow一般分为两部分，构造部分和执行部分。构造部分包含计算流图；执行部分通过Session来执行图中的计算

Tensorflow Python库有一个默认图（default graph），节点构造器（op构造器）可以增加创建源节点（source op），源节点输出传递给其他节点（op）做运算

# 3、GBDT和XGBoost的区别

XGBoost类似于GBDT的优化版，不论是精度还是效率上都有了提升。与GBDT相比，具体的优点有：

1.损失函数是用泰勒展式二项逼近，而不是像GBDT里的就是一阶导数；

2.对树的结构进行了正则化约束，防止模型过度复杂，降低了过拟合的可能性；

3.节点分裂的方式不同，GBDT是用的基尼系数，XGBoost是经过优化推导后的。

知识点链接：[集成学习总结](https://xijunlee.github.io/2017/06/03/%E9%9B%86%E6%88%90%E5%AD%A6%E4%B9%A0%E6%80%BB%E7%BB%93/)

# 4、用欧氏距离，不用曼哈顿距离？