大家晚上好，今天我分享的主题是AdaBoost——自适应提升算法。

首先我将介绍一下集成学习与AdaBoost这两个概念

相信大家对机器学习都有一定的了解。简而言之，机器学习，是指人们收集到很多数据，利用计算机系统来建立一个模型，并利用这个模型去处理新的数据的一个过程。

而集成学习是一种机器学习范式，体现了「团结就是力量」、「群体智慧」的思想。在生活中，假设我们随机去问很多人一个很复杂的问题，然后把他们的答案合并起来。通常情况下会发现这个合并的答案比一个人给出的答案要更好。这就叫做群体智慧。

使用机器学习的术语，我们可以这样描述集成学习的过程。就传统而言，我们会用一个学习器去解决问题，即单模型。而集成学习是将多个单模型（弱分类器）组合成一个强分类器，这个强分类器能取所有单模型之所长，达到相对的最优性能。而且在组合时，我们可以用同样的学习器，例如都使用神经网络，决策树，或者是SVM支持向量机这些单模型组成，这就叫同质的学习器。有时候也可以使用不同的学习器，那就叫异质的学习器。

总之，用多个相同或不同单模型组合成集成模型，来解决一个问题，这一组学习器就叫做集成。因此，这个技术就叫做集成学习。

同时，按集成中个体的生成方式来说，我们可以把集成学习分成两大类，一类叫序列化方法。在这类方法里面，每一个新的个体，它的生成和前一个个体，或者说和以往的个体都是有联系的。代表是AdaBoost算法。而还有一种叫并行化的方法，所有的新个体它可以同时生成。代表是Bagging算法和Random Forests随机森林。，

AdaBoost于起源1989年Kearns卡恩斯和Vallant瓦兰特提出的一个设想“Weakly Learnable” ?= “Strongly Learnable”。其中Weakly Learnable是指一个算法模型，其正确率接近0.5，只比随机猜测的正确率高一点。Strongly Learnable指的是在一个多项式时间内能够得到一个算法模型，该模型的误差小于0.5，且获得该模型的概率较大。因此他们提出，我们是否能将一个比随机猜测准确率只高一点的弱学习算法，通过Boost转化为一个强学习算法？

1990年的Schapire夏皮尔证明了这个问题是可实现的，并且提出了如何实现的数学过程。而1993年Freund弗罗因德提出了可以通过投票方式，将几个弱分类器算法集合。而在1995年，这两位计算机科学家一起提出了AdaBoost算法。

AdaBoost的全称为Adaptive Boost，可以翻译为自适应提升算法。它是一种通过改变训练样本权重来学习多个弱分类器并线性组合成强分类器的Boosting算法。

它的自适应体现在前一个基本分类器分错的样本会得到加强，加权后的全体样本再次被用来训练下一个基本分类器。同时在每一轮中加入一个新的弱分类器，直到得到某个预定的足够小的错误率或达到预定的最大迭代次数。

接着介绍一下AdaBoost分类算法的具体流程

给定训练集，其中元素可看成二维向量。且其中的Y取自-1或者1，即这是一个二元分类

首先初始化训练数据的权值分布。假设如果有𝑁个样本且均匀分布，则每一个训练样本最开始都被赋予权值1/𝑁。

然后是进行多轮迭代，训练弱分类器。训练过程中，先训练弱分类器。通过计算得出的，可以实现动态调整权重系数。

如果某个样本点已经被准确的分类，那么在构造下一个训练集中，他的权值会被降低。相反，如果某个样本点没有被准确分类，那么它的权值就会得到提高，使其在最终的分类函数中起着较大的决定作用。权值更新过的样本集会被用于训练下一个分类器，整个训练过程如此迭代的进行下去。

最后采用加权平均法结合策略，将多个弱分类器组合成强分类器。

对于AdaBoost多元分类算法，其实原理和二元分类类似，最主要区别在弱分类器的系数上。如AdaBoost SAMME算法，它的弱分类器的系数如下所示。

这是一个AdaBoost流程示意图：XXX

然后第三部分是关于AdaBoost的分析与应用

首先AdaBoost是简单且有效的算法，可应用于几乎所有的领域。AdaBoost算法最成功的应用之一是机器视觉里的目标检测问题，如人脸检测、行人检测和车辆检测。在深度卷积神经网络用于此问题之前，AdaBoost算法在视觉目标检测领域的实际应用上一直处于主导地位。

在2001年Viola和Jones设计了一种人脸检测算法，称为VJ框架。VJ框架是人脸检测历史上有里程碑意义的一个成果，奠定了AdaBoost目标检测框架的基础。

框架中用多个AdaBoost分类器合作完成对候选图像的分类，这些分类器组成一个流水线。在这些AdaBoost分类器中，前面的分类器很简单，包含的弱分类器很少，可以快速排除掉大量非人脸图像，如右边图示中，仅通过黑白的对比区分。但也可能会把一些不是人脸的图像判定为人脸。如果一个候选图像通过了第一级分类器的筛选，则送入下一级分类器中进行判定，否则丢弃掉，以此类推。如果一个检测窗口通过了所有的分类器，则认为是人脸。

其次，AdaBoost有着极强的生命力。它不是简单一个算法，是一个General Boosting的过程。图片是其一般性过程。由此衍生出了不同的算法，发展成为一个算法族，共性是源自AdaBoost。其中有XGBoost, GBDT, FilterBoost, GnetleBoost等

AdaBoost有很好的理论保证。Freund和Schapire在提出AdaBoost算法的同时，给出了训练集误差范围，还给出了泛化误差（测试集上误差）的范围

AdaBoost还有其他优点，如PPT上所罗列出来的，由于时间原因，在此不一一赘述。但是AdaBoost自身也有缺点，如对异常样本敏感，异常样本可能会在迭代过程中获得较高的权重值，最终影响模型效果。但是它仍因其有点大于缺点，被用于模式识别、计算机视觉领域中二分类和多分类场景

最后是AdaBoost的一个实践案例