

《机器学习》教案

适用专业： 机器人、人工智能

教学年度：

教学班级：

授 课 教师		课程名称	
授 课 地点		授课时间	第 14 次
章 节 题目	决策树		
教 学 目的	模型加速能力是重要的手段也是推向模型进化的主要原因		
教 学 重点	剪枝的可能性和本质		
教 学 难点	理解剪枝的可能性和本质		
教 学 方法	启发式，案例		
教 具	多媒体，程序运行结果，黑板板书		

案例设计	用例子讲解计算过程
教学内容	<ul style="list-style-type: none"> ● 决策树预剪枝 <ol style="list-style-type: none"> 1. 决策树的参数包括：树的深度，叶节点内的最小样本数、和信息增益阈值； 2. 预剪枝观点是等到决策树完全生长完成再剪枝不如预先完成；预剪枝能减少计算量； ● 决策树悲观剪枝 <ol style="list-style-type: none"> 1. 悲观剪枝的条件是我们没有验证集的情况下，我们可利用置信区间的概念进行剪枝； 2. 剪枝前后，决策树的期望错误率不超过预定义的一个阈值； 3. 期望错误率的建模过程用二项式分布，即用二项式分布来描述落入一个节点内样本点数量； 4. 二项式分布难以计算可以用高斯分布进行近似； 5. 能用悲观剪枝进行决策树剪枝代码实现； ● 决策树代价复杂度剪枝 <ol style="list-style-type: none"> 1. 复杂度剪枝是既能描述分类准确性又能描述决策树复杂度的指标； 2. 准确性的描述用分类错误率而决策树复杂度用叶结点数量； 3. 利用最优子决策树序列的嵌套性贪心的计算复杂度剪枝的指标；
教学过程	<ol style="list-style-type: none"> 1. 讲解剪枝与过拟合的关系 2. 讲解剪枝的两种范式
教学总结	

		课程名称	
--	--	------	--

授 课 教师			
授 课 地点		授课时间	第 15 次
章 节 题目	集成学习		
教 学 目的	集成学习的分类，学习器的优劣和泛化误差		
教 学 重点	集成学习的动机，什么是泛化误差		
教 学 难点	如何将多个弱学习器层次的累叠在一起，掌握偏差-方差分解过程中的假设和推导过程；		
教 学 方法	启发式，案例		
教 具	多媒体，程序运行结果，黑板板书		
案 例 设计	深度神经网络和集成学习的实验分析		
教 学 内容	<ul style="list-style-type: none"> ● 集成学习的分类 <ol style="list-style-type: none"> 1. 集成学习的动机是将能力弱的学习器进行组合后获得能力更强的强分学习器； 2. 将多个弱学习器加法叠加在一起-> bagging 和 boosting； 3. 成法集成学习-> 无法将弱能力的学习器组合后变强； 4. 将多个弱学习器层次的累叠在一起-> 深度神经网络； 5. 集成学习对弱学习器必须有一定的限制条件，包括，易于训练和并行化，弱学习器之间应该有某种互补性，必须在一定原则下进行集成。 ● 学习器的优劣和泛化误差 		

《机器学习》教案

	1.
--	----

	<p>泛化误差是指利用不同的训练集合（可能是独立同分布）对模型进行多次训练后，对同一测试集合误差的期望；</p> <p>2. 掌握偏差-方差分解过程中的假设和推导过程；</p> <p>3. 弱学习器用偏差和方差进行刻画；</p>
教学过程	<p>1. 分析集成的动机</p> <p>2. 提出集成学习的学习器叠加</p> <p>3. 讲解神经网络</p> <p>4. 讲解弱学习器</p> <p>5. 分析泛化误差的影响</p> <p>6. 理解偏差-方差分解过程</p>
教学总结	

授课教师		课程名称	
授课地点		授课时间	第 16 次
章节题目	集成学习		
教学目的	掌握偏差和方差是看待模型的重要手段		
教学重点	模型选择与交叉验证		
	掌握模型选择的依据		

教学难点	
教学方法	启发式，案例
教具	多媒体，程序运行结果，黑板板书
案例设计	Boosting 讲解和 randomforst 的实验分析
教学内容	<ul style="list-style-type: none"> ● 模型选择与交叉验证 <ol style="list-style-type: none"> 1. 训练集、验证集都属于训练集，而测试集和训练集构成全部数据；这三个集合之间互相没有交集； 2. 因为对数据集的一次固定（随机）划分，验证集或测试集里面仍然存在偏好（例如，某些特殊的数据出现在测试集里面，而训练集少甚至没有）； 3. K 折交叉验证是加强版本的交叉验证，是一种让样本都有一次机会成为验证集，是样本某个指标平均意义下对模型的选择； 4. 理解高偏差，高方差，低偏差，低方差及其组合在：1）随着样本数增加；2）随着模型复杂度变化，这两个因素下的性能（误差）的表现； 5. 我们通常认为模型参数变化后，具体的模型也会变化。狭义的理解，模型选择是对具体模型参数的选择。广义的理解，模型选择是对任意模型的选择。
教学过程	<ol style="list-style-type: none"> 1. 分析集成的可能性 2. 提出集成学习的几种范式 3. Bagging 的偏差和方差的推导和假设 4. 讲解交叉验证的方法
教学总结	

授课教师		课程名称	
授课地点		授课时间	第 17 次
章节题目	集成学习		
教学目的	<p>偏差和方差是看待模型的重要手段</p> <p>平均集成，可加性集成和叠加性集成的关系</p>		
教学重点	<p>泛化性的分析和理解</p> <p>拟合，过拟合，欠拟合之间的关系</p> <p>模型选择与交叉</p>		
教学难点	Bagging 的偏差和方差的推导过程		
教学方法	启发式，案例		
教具	多媒体，程序运行结果，黑板板书		
案例设计	Boosting 讲解和 randomforst 的实验分析		
教学内容	<ul style="list-style-type: none"> ● Bagging <ol style="list-style-type: none"> 1. Bagging 就是对具有低偏差但高方差学习器进行集成； 2. 低偏差是指每个学习器分类能力强，容易会存在过拟合现象，高方差是指每个学习器预测当训练集变化时候，学习器的预测变化也很大。决策树是典型的低偏差但高方差的学习器； 3. Bagging 的方差和偏差的推导过程所用的数学技巧； 4. Bagging 为了让强分类器的方差减少，让弱学习器权重变小，相关性也变小； 5. Bagging 为了让强学习器的偏差减少，我们需要所有弱学习的偏差都小，并且希望偏差差异越小越好。 6. Bagging 的包外数据及其比例的推导过程； 7. 		

	<p>证明有放回的采样方法获得的样本子集与全样本集合属于同一分布。</p> <p>8. 理解 SVM 作为弱学习器无法获得有效的 bagging 结果, 而决策树的 bagging 能获得有效的 bagging 结果。</p>
教学过程	<ol style="list-style-type: none"> 1. 分析集成的可能性 2. 提出集成学习的几种范式 3. Bagging 的偏差和方差的推导和假设 4. 随机森林是 bagging 下的一个特殊例子
教学总结	

授课教师		课程名称	
授课地点		授课时间	第 18 次
章节题目	集成学习		
教学目的	<p>什么是随机森林</p> <p>随机森林的动机和特性</p>		
教学重点	<p>如何进行随机的特征抽取和样本抽取</p> <p>袋外数据产生过程</p>		
教学难点	随机森林决策树之间的相关性		
教学方法	启发式, 案例		
教具	多媒体, 程序运行结果, 黑板板书		
案例设计	随机化过程和袋外数据产生过程		
教学内容	<ul style="list-style-type: none"> ● 随机森林 <ol style="list-style-type: none"> 1. 		

	<p>在特征层面和样本层面进行随机的特征抽取和样本抽取，随机化能降低随机森林中决策树的相关性；</p> <p>2. 从袋外数据产生过程理解随机森林中决策树之间仍然有相关性。</p>
教学过程	<p>1、理解随机森林是 bagging 下的一个特殊例子</p> <p>2、理解随机森林决策树如何产生相关性和如何降低</p>
教学总结	

授课教师		课程名称	
授课地点		授课时间	第 19 次
章节题目	无监督学习		
教学目的	<p>无监督学习的目的</p> <p>K-mean 等聚类算法的假设</p>		
教学重点	<p>K 均值聚类算法</p> <p>非凹非凸函数的计算过程怎么保证解的稳定性</p> <p>模型选择的常用方法和思考维度</p>		
教学难点	<p>期望最大化算法</p> <p>K 均值聚类算法的目标函数</p> <p>K 均值聚类算法是基于聚类是等方差的高斯分布</p> <p>模型选择方法</p>		
教学方法	启发式，案例		
教具	多媒体，程序运行结果，黑板板书		
案例设计	非凹非凸如何被优化到局部最小		

<p>教学内容</p>	<ul style="list-style-type: none"> ● 无监督学习的动机 <ol style="list-style-type: none"> 1、从大量无标记的样本里面挖掘出可能性的数据内含的规律； 2、数据生成应该受到潜在因素的控制，这些潜在因素中最重要的因素之一是聚集性； 3、聚集性是规律重现的直观认识。 ● K 均值聚类算法 <ol style="list-style-type: none"> 1、K 均值聚类算法是依靠计算样本点的距离判断样本之间的聚集性； 2、K 均值聚类算法的目标函数是非凹非凸，只有固定一个变量优化另一个变量的情况下目标函数才能被优化到局部最小； 3、K 均值聚类算法的优化是基于梯度下降算法进行优化； 4、K 均值聚类算法是基于聚类是等方差的高斯分布； 5、K 均值聚类算法目标函数的只能得到局部最优，基于初始化能影响解这一假设，K 均值++是基于聚类中心相互远离这样观测设计的顺序初始化方法； 6、BIC 和 AIC 是模型选择中常用的两种方法，AIC 是拟合精度和模型参数量来选择模型，而 BIC 是考虑样本数量和拟合精度之间的均衡。 7、拟合精度，数据量，模型参数是模型选择的思考维度。
<p>教学过程</p>	<ol style="list-style-type: none"> 1、理解无监督学习的动机的由来 2、理解 K 均值聚类算法的依据 3、理解 K 均值聚类算法目标函数的优化 4.了解模型选择的思考维度
<p>教学总结</p>	

<p>授课教师</p>		<p>课程名称</p>	
<p>授课地点</p>		<p>授课时间</p>	<p>第 20 次</p>

章节题目	无监督学习
教学目的	从概率的角度理解模型 高斯混合模型的期望最大化算法
教学重点	基于混合高斯模型的聚类 期望最大化算法 非凹非凸函数的计算过程
教学难点	期望最大化算法的本质 目标函数数学求解
教学方法	启发式，案例
教具	多媒体，程序运行结果，黑板板书
案例设计	高斯混合模型在不同参数下的性能
教学内容	<ul style="list-style-type: none"> ● 基于混合高斯模型的聚类 <ol style="list-style-type: none"> 1.对 K 均值聚类的等方差高斯分布转化为任意的多元高斯分布； 2.高斯混合模型是有限参数的概率模型，与之对应的是非参数的 Parzen 窗； 3.高斯混合模型优化中的混合系数是隐含变量，期望是对隐含变量的积 4.形成隐含变量分布下的目标函数的期望；最大化是对目标函数的期望的优化，获得变量的解； 5.期望最大化算法本质是对复杂分布在 KL 散度度量下的逐步逼近； 6.数学上，如果我们能判断函数的凹凸性，我们还可以利用 Jessian 不等式对目标函数转换为目标函数的上下界进行求解。
教学过程	<ol style="list-style-type: none"> 1.理解 K 均值聚类的等方差的高斯分布转化 2.理解高斯混合模型优化的参数期望等细节 3.理解理解期望最大化算法的本质

教学总结	
------	--

授课教师		课程名称	
授课地点		授课时间	第 21 次
章节题目	无监督学习		
教学目的	理解期望最大化算法的动机和应用和难点 DBSCAN 对噪声的处理本质		
教学重点	高斯混合模型的期望最大化算法 噪声下基于密度的空间聚类算法		
教学难点	期望最大化算法 KL 散度的由来		
教学方法	启发式，案例		
教具	多媒体，程序运行结果，黑板板书		
案例设计	高斯混合模型在不同参数下的性能		
教学内容	<ul style="list-style-type: none"> ● 期望最大化算法 <ol style="list-style-type: none"> 1、KL 散度的由来？ 2、期望最大化算法主要解决具有隐变量的模型； 3、期望最大化算法的难度在于期望的求解。 ● 噪声下基于密度的空间聚类算法 <ol style="list-style-type: none"> 1、DBSCAN 是利用 Parzen 窗的思想，让每个样本作为概率密度估计的计算点，通过定义连通性将两个样本点进行合并，再确定阈值区分不同的聚类； 2、DBSCAN 对噪声的处理本质上是一个概率上基于人工定义的分类问题； 		

教学过程	<ol style="list-style-type: none"> 1、理解期望最大化算法的动机目标 2、理解 DBSCAN 的基本思想 3、理解 DBSCAN 处理噪声的本质
教学总结	

授 课 教 师		课程名称	
授 课 地 点		授课时间	第 22 次
章 节 题 目	降维分析		
教 学 目 的	让学生了解降维分析的动机，掌握主成分分析方法，理解主成分分析中对数据进行归一化的作用。		
教 学 重 点	主成分分析方法		
教 学 难 点	主成分分析如何保留高维数据的“本质”信息；主成分分析的算法流程；		
教 学 方 法	启发式，案例		
教 具	多媒体，程序运行结果，黑板板书		
案 例 设 计	用主成分分析将高维矩阵变为低维		
	<ul style="list-style-type: none"> ● 降维的动机 		

<p>教 学 内 容</p>	<ol style="list-style-type: none"> 1. 高维向量维度间可能存在相关性，因此可以用一个变量来代替向量中的多个变量； 2. 降维是多个变量整合为一个变量，而特征选择是找符合特性（如，判别力）； 3. 降维的变量的可解释性是根据数据建模的特性所赋予。 <ul style="list-style-type: none"> ● 主成分分析 <ol style="list-style-type: none"> 1. 理解主成分分析中投影矩阵是正交基是约束解空间性质的手段，让低维特征互相不相关，获得紧致的特征； 2. 主成分分析解的特性是让投影空间中的解尽量不相关； 3. 主成分分析剪掉均值预先处理的动机本质来源于投影空间中的解尽量不相关的多解性； 4. 主成分分析解体现了压缩-扩张矛盾的过程保证解的稳定性； 5. 主成分分析解的求解过程利用投影向量的正交性等价于特征值和特征向量的求解过程；
<p>教 学 过 程</p>	<ol style="list-style-type: none"> 1. 理解高维度数据如何用低维空间表示 2. 理解降维分析的动机 3. 讲解主成分分析目标函数即其简化形式 4. 讲解主成分分析目标函数的优化 5. 主成分分析中对数据进行归一化的作用
<p>教 学 总 结</p>	

<p>授 课 教 师</p>		<p>课程名称</p>	
<p>授 课 地 点</p>		<p>授课时间</p>	<p>第 23 次</p>
	<p>降维分析</p>		

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：

<https://d.book118.com/938130025121006072>