

《机器学习与自然语言处理初级班》课程安排

课程	主要理论内容	实训内容	课时数	总课时数	上课方式
机器学习与自然语言处理初级班	1.机器学习入门，朴素贝叶斯分类，K 近邻分类	1.使用朴素贝叶斯方法进行文本分类	3 理论+5 实训	48	每周六
	2.线性感知机，线性支持向量机	2.使用 SVM-Light 进行文本分类	3 理论+5 实训		
	3.K-means 聚类及层次聚类，混合模型基础	3.使用 K 均值法对文本聚类	3 理论+5 实训		
	4.线性回归与逻辑回归，梯度下降法	4.使用逻辑回归技术预估广告点击率	3 理论+5 实训		
	5.马尔科夫模型与语言模型，隐马模型基础	5.隐马模型维特比算法实现中文分词，语言模型构建训练	3 理论+5 实训		
	6.网页排序基础，pagerank，learning to rank 基础 (pointwise，pairwise)	6.使用 pair-wise 方法实现网页排序	3 理论+5 实训		

课程目的：教会大家机器学习算法是什么？怎么用？强调知道每个技术是什么，会用该技术解决常见的简单问题

培养目标：培养初级机器学习工程师、初级数据挖掘工程师

适用人群：掌握一门编程语言，熟悉数据结构及常见简单算法，有一定的编程经验，有志于成为机器学习工程师、数据挖掘工程师的人

网址：www.peileyuan.com

《机器学习与自然语言处理高级班》课程安排

课程	主要理论内容	实训内容	课时数	总课时数	上课方式
机器学习与自然语言处理高级班	1. 主体模型 (Topic modeling), PLSA , LDA	1. 利用 EM 算法训练 PLSA	3 理论+5 实训	48	每周日
	2. 模型融合, 决策树, BOOSTING	2. 使用 Adaboost 算法进行文本分类	3 理论+5 实训		
	3. 隐马模型的前向后向算法, CRF, 最大熵模型 ME	3. 使用 ME 模型实现文本分类	3 理论+5 实训		
	4. 关联规则挖掘基础, Apriori 算法, FP-Growth 算法	4. 使用 Apriori 算法挖掘演员合作关系	3 理论+5 实训		
	5. L-BFGS, 序列最小优化算法 (SMO)	5. L-BFGS 训练 LR 模型	3 理论+5 实训		
	6. Learning to rank 深入, listwise、pairwise (不同 loss)	6. 多种 pairwise 方法的比较	3 理论+5 实训		

课程目的：教会大家这些机器学习算法或模型的原理及如何优化调整，深入理解算法的原理，然后在此基础上，创新地运用该技术解决更复杂的问题

培养目标：培养中高级机器学习工程师、中高级数据挖掘工程师

适用人群：有机器学习技术基础，掌握机器学习与自然语言处理基础班内容，有一定的使用经验，有志于成为更高层次机器学习、数据挖掘方面人才者