

CSDN 学院人工智能工程师课程大纲

第一阶段 机器学习原理及推荐系统实现				
时间	主题	理论	实战案例	课程目标
Week1	机器学习简介	<ol style="list-style-type: none"> 1. 机器学习定义； 2. 机器学习行业应用举例； 3. 机器学习任务：监督学习（分类、回归）、非监督学习（聚类、降维）、半监督学习、迁移学习、强化学习； 4. 机器学习算法的组成部分：目标函数（损失函数+正则）、优化方法； 5. 模型评估和模型选择：模型复杂度、过拟合、交叉验证、超参数空间、网格搜索... 	房价预测案例 数据集探索：单特征分布模拟及可视化、离群点检测、多特征相关性分析及可视化（NumPy、Pandas、Matplotlib, seaborn） 实现模型评估和模型选择：交叉验证、网格搜索（scikit-learn）	熟悉机器学习领域的常用术语，了解机器学习在 AI 中的地位 学习环境配置：常用软件、环境配置及机器学习库 anaconda: Python、科学计算包（NumPy, SciPy, Pandas）、数据可视化工具包（Matplotlib, seaborn）、机器学习库（scikit-learn） 学会用机器学习工具包从头到尾用线性回归解决一个实际问题
Week2	Logistic 回归分析、神经网络、SVM	<ol style="list-style-type: none"> 1. 分类算法的损失函数：logistic 损失、Hingloss 损失、 2. 优化算法：IRLS（梯度下降、牛顿法）、BP 算法、SMO（序列最小最优化算法） 3. 正则化：L1/ L2 4. 复习模型评估 5. 其他：最小间隔、核方法、支持向量回归 	电商商品分类案例 用 Logistic 回归、神经网络和 SVM 等分类器实现商品分类 比较不同模型以及不同参数下 SVM（不同正则参数和核函数）的性能，体会各模型的特点	理解分类任算法（Logistic 回归、神经网络、SVM）原理，复习数据集探索，并学会在 scikit-learn 框架下采用各分类算法分类具体任务。
Week3	决策树模型	<ol style="list-style-type: none"> 1. 损失函数：信息增益、Gini 系数 	电商商品分类案例	学习 Boosting 集成思想及基于树的

	(CART)、基于树的集成学习算法(随机森林、GBDT)	<ol style="list-style-type: none"> 划分: 穷举搜索、近似搜索 正则: L2/L1 预防过拟合: 预剪枝及后剪枝 Bagging 原理 Boosting 原理 流行的 GBDT 工具: XGBoost 和 LightGBM 	XGBoost 在实际案例上的参数调优	集成算法
Week4	聚类、降维、矩阵分解	<ol style="list-style-type: none"> 主成分分析 (PCA) 独立成分分析 (ICA) 非负矩阵分解 (NMF) 隐因子模型 (LFM) KMeans 聚类和混合高斯模型 GMM(EM 算法) 吸引力传播聚类算法 (Affinity Propagation 聚类算法) 	人脸图像特征提取: PCA、ICA、NMF 电商用户聚类案例	学习用降维技术对高维特征进行降维
Week5	特征工程、模型融合& 推荐系统实现	<ol style="list-style-type: none"> 数据预处理: 缺失值处理 特征编码: 标签编码、Dummy (One hot) 编码、后验均值编码 文本特征提取 特征组合 特征选择 协同过滤 基于内容的过滤 FFM & LFM 排序学习 模型融合: Blending、Stacking 	商品推荐案例 复习数据探索、数据离群点检测和处理 数据预处理: 缺失值处理 特征编码 组合各种特征工程技术和机器学习算法实现推荐系统	学会常用数据预处理方法及特征编码方法 学习特征工程的一般处理原则 实现一个实际的推荐系统

第二阶段 深度学习原理及实战项目强化训练

时间	主题	理论	实战案例	课程目标
Week1	神经网络入门	<ol style="list-style-type: none"> 1. 神经网络历史与现状 2. 神经网络的分类：全连接、卷积、循环 3. 神经网络的应用：图像、语音、自然语言处理 4. 神经网络的计算：权重、损失和梯度 5. 神经网络的优化：前向/反向传播和梯度下降 6. 全局最优、局部最优和鞍点 7. 正则化、归一化 	LeNet 与传统神经网络对比 包含知识点—— 损失函数、L1/L2 正则、梯度下降/ 随机梯度下降/动量随机梯度下降	熟悉神经网络领域的常用术语，了解很精网络在 AI 环境中的位置。
	机器学习环境配置	<ol style="list-style-type: none"> 1. 学习环境配置：常用软件、环境配置及机器学习库 anaconda: Python、Python 科学计算包 (NumPy, SciPy, Pandas)、Python 数据可视化工具包 (Matplotlib, seaborn) 神经网络框架 Tensorflow 2. 简单神经网络实现手写数字识别 	Mnist 手写数字识别 包含知识点—— Mnist 数据集探索 Tensorflow 基础概念 计算图 session	学会用 tensorflow 解一个实际问题。
Week2	神经网络基础	<ol style="list-style-type: none"> 1. 多层神经网络结构：输入、输出、隐 	手写数字识别 Mnist	使用不同结构的神经网络结构验

		层和激活函数 2. 前向计算和损失 向量化计算和 one-hot 编码 sigmoid、softmax 及交叉熵 3. 反向传播及迭代优化 梯度下降及动量 4. 过拟合与欠拟合 5. 正则化、批正则化及 Selu 6. Dropout	验证码识别 包含内容—— 使用简单神经网络实现手写数字识别	证网络结构对效果的影响
	卷积神经网络	1. 卷积 padding、stride、kernel 和 channel 局部相关性 感受野 感受野的计算 2. 池化 max_pooling、average_pooling global_average_pooling 3. 局部网络连接	LeNet 数据集 Mnist 用 LeNet 实现手写数字识别 局部相关性 权值共享	了解卷积神经网络的相关概念和基础知识
Week3	卷积神经网络分类任务	1. 图像分类介绍与实现 2. imagenet 数据集与预训练模型 3. Inception 网络 4. ResNet 网络 5. 细粒度分类	Flowers 数据集 基于 imagenet 预训练模型的迁移	学习图像分类任务目前主要模型算法
	卷积神经网络	1. 检测任务介绍与实现	PascalVOC 数据集	学习检测任务目前主要模型算法

	检测任务	<ol style="list-style-type: none"> 特征提取 区域建议 区域合并 R-CNN Fast/Faster R-CNN SSD YOLO 	使用 Tensorflow 训练一个检测模型	
Week4	卷积神经网络分割任务	<ol style="list-style-type: none"> 分割任务简介 反卷积 (deconv/transpose-conv) FCN 	COCO 数据集 学习语义分割模型	学习主流分割模型
Week5	循环神经网络	<ol style="list-style-type: none"> RNN 基本原理 门限循环单元 (GRU) 长短期记忆单元 (LSTM) 词向量提取: Word2Vec 编码器-解码器结构 注意力机制模型: Attention Model 图片标注 (Image Captioning) 图片问答 (Visual Question Answering) 	写诗机器人	学习循环神经网络的原理及应用

第三阶段 四个工业级实战项目（可选）及成果展示

项目名称	项目内容
自然语言处理：文本分类	根据企业的注册、投资及经营范围等相关信息，对企业进行分类，为企业的估值提供参考。
广告点击率预测（CTR）	预测用户浏览给定网页的广告点击率，提高广告投放精准度。
车辆检测及型号识别	用深度学习方法从图片中检测车辆并识别其型号。
看图说话机器人	用计算机视觉和深度学习方法分析图片内容，并对图片自动生成文字描述。