

# 2024—2025学年第 1学期 教 学 进 度 表

课程名称	人工智能导论I			课程性质	必修
总 学 时	48 （其中讲课：32 学时；实验（习）：16 学时 ）				
系 部	计算机学院计算机科学技术系			专业年级	2021年级计科(5)班, 计科(6)班
主讲教师	应龙			辅导教师	
周次	讲 课	实验（习）、习题课、测验			
学时数	章节主要内容（包括重点、难点）	学时数	内 容		
1	第1章 人工智能基础： 内容：人工智能的概念和哲学基础，不同版本的人工智能和 Beneficial machines，人工智能的基础，人工智能的发展历程，人工智能主要流派，人工智能的现状。Agent-environment 模型。 重点：人工智能主要流派，人工智能的数学基础（可计算理论），Agent-environment 模型	人工智能的起源			
		可计算载体			
		智能计算方法			
2	第1章 人工智能基础（续）， 第2章 问题求解与搜索： 内容：好的行为—理性的概念，The Nature of Environments, Agents 的结构，Agent 程序的各组件如何工作，课程总体框架。问题求解与形式化描述，搜索算法框架（树搜索与图搜索）。 难点：Agents 的结构，Agent程序的各组件如何工作，问题求解与形式化描述。	理性的概念			
		智能体的结构			
2	第2章 问题求解与搜索（续）： 内容：搜索策略，A*算法（评价函数和启发式函数，一致性，可容性，信息性），A*搜索算法的完备性，树搜索A*算法的最优性，图搜索A*算法的最优性，启发式函数设计的主要方法。 难点：A*搜索算法的完备性，图搜索A*算法的最优性，启发式函数设计的主要方法。	问题求解形式化：状态空间模型			
		搜索算法框架：树搜索，图搜索			
		A*搜索算法：评价函数和启发式函数，可容性，一致性，A*搜索算法的完备性，A*搜索算法的最优性			
3	第2章 问题求解与搜索（续）， 第3章 对抗搜索： 内容：局部搜索和最优化问题，连续空间中的局部搜索方法，遗传算法（算法框架，主要操作，运作原理）。多智能体环境，对抗搜索。 难点：局部搜索和最优化问题，遗传算法。多智能体环境。	启发式函数设计的主要方法			
		遗传算法（算法框架）			
4	第3章 对抗搜索（续）： 内容：最小最大搜索，博弈论基本概念，囚徒困境，博弈的均衡，纳什均衡的求解，博弈的分类，策梅洛定理，打破囚徒困境。Alpha-Beta 剪枝原理。 难点：最小最大搜索，博弈的均衡，纳什均衡的求解，Alpha-Beta 剪枝原理。	对抗搜索：最小最大搜索	2		
		博弈论的基本概念			
		博弈的均衡：纳什均衡，囚徒困境			
5	国庆放假				
6	第3章 对抗搜索（续）， 第4章 知识与推理： 内容：Alpha-Beta 剪枝算法，Alpha-Beta 时间复杂度分析，启发式Alpha-Beta 搜索算法，评估函数，蒙特卡洛树搜索算法的主要思想，多臂老虎机模型，上限置信区间算法，算法框架，提前模拟终止，选择策略，模拟策略。知识表示概念，思考快与慢。 难点：Alpha-Beta 时间复杂度分析，多臂老虎机模型，上限置信区间算法，选择策略，思考快与慢。	Alpha-Beta 剪枝：Alpha-Beta 剪枝原理，剪枝算法	2		
		Monte Carlo 树搜索：主要思想，上限置信区间算法，算法步骤			
7	第4章 知识与推理（续）， 第5章 不确定性推理与因果推理： 内容：知识表示概述，推理概念和概述，语义网络，知识图谱推理归纳学习FOIL算法，路径排序推理算法(PRA)，基于分布式表示的知识推理。贝叶斯网络的概念。 难点：语义网络，归纳学习 FOIL算法，路径排序推理算法(PRA)，基于分布式表示的知识推理。	知识表示的基本概念			
		推理的基本概念			
		语义网络：知识图谱的概念			
		知识图谱推理：FOIL算法			
		知识图谱推理：路径排序推理			
		知识图谱推理：分布式表示推理			

8	2	第5章 不确定性推理与因果推理（续）： 内容：贝叶斯网络条件独立性关系，贝叶斯网络精确推理（枚举法，Variable Elimination, Belief propagation），贝叶斯网络近似推理，贝叶斯网络学习，辛普森悖论，从关联到因果推理，因果推理的主要模型，条件独立性，D-分离和D-链接，do算子，引入干预的操纵图模型。 难点：贝叶斯网络精确推理 Belief propagation，贝叶斯网络近似推理，贝叶斯网络学习，D-分离和D-链接，干预的因果效应。	贝叶斯网络：概念，条件独立关系		2	黑白棋AI算法
			贝叶斯网络：D-分离和D-链接			
			贝叶斯网络：精确推理，近似推理			
			贝叶斯网络：结构学习			
			因果推理的概念，因果图，干预的概念 干预的后门准则：后门准则的条件，调整公式			
9	2	第5章 不确定性推理与因果推理（续）， 第6章 机器学习： 内容：反事实推理，机器学习的基本概念，kmeans聚类，机器学习的重要元素，经验风险和期望风险，归纳偏好，没有免费午餐定理，结构风险最小化，计算学习理论（PAC Learning，泛化误差上界，假设空间复杂度）。 难点：经验风险和期望风险，没有免费午餐定理，结构风险最小化，泛化误差上界	机器学习的基本概念：样本空间，特征空间，标记空间，概念，假设空间			
			kmeans 聚类：目标函数，超参优化			
			机器学习的重要元素：策略，模型，算法；目标函数			
			机器学习的基本概念：训练集，验证集，测试集，经验风险，期望风险，过拟合，欠拟合			
			归纳偏好，没有免费午餐定理，结构风险最小化			
10	2	第6章 机器学习（续）， 第7章 深度学习： 内容：支持向量机的原理与形式化，原问题与对偶问题，KKT条件，支持向量机的对偶形式，核函数与核方法。人工神经网络拓扑结构。 难点：支持向量机的原理与形式化，KKT条件，核函数与核方法。人工神经网络拓扑结构。	机器学习的基本概念：机器学习的分类：监督，无监督，强化学习，半监督，自监督		2	kmeans聚类
			SVM: soft margin, hinge loss, 核函数			
11	2	第7章 深度学习（续）： 内容：元素方式的反向传播算法，矩阵方式的反向传播算法，反向传播算法的一般形式和计算图，神经网络激活函数（Sigmoid, Softmax），常用损失函数（最小均方误差，交叉熵）。 难点：矩阵方式的反向传播算法，反向传播算法的一般形式和计算图，Softmax激活函数，交叉熵损失函数。	神经网络的基本部件			
			神经网络的拓扑结构			
			矩阵方式的反向传播算法			
12	2	第7章 深度学习（续）： 内容：循环神经网络随时间反向传播，实时循环学习，梯度爆炸与梯度消失，多层神经网络优化改进，神经网络正则化（Dropout, L1范数）。 难点：实时循环学习，梯度爆炸与梯度消失，多层神经网络优化改进。	循环神经网络BP算法：随时间反向传播，实时循环学习		2	基于特征脸的分类识别
			梯度爆炸与梯度消失，多层神经网络优化改进 神经网络正则化：Dropout, L1范数			
13	2	第7章 深度学习（续）， 第8章 强化学习： 内容：神经网络中的归一化，神经网络结构的发展，判别式方法与生成式方法概念，判别式方法与生成式方法概念，GAN（优化的目标函数，JS divergence, Wasserstein GAN），强化学习问题的定义，贝尔曼方程。 难点：神经网络正则化，神经网络中的归一化，Wasserstein GAN，贝尔曼方程。	神经网络中的归一化：BN, LN			
			卷积，卷积神经网络			
			判别式方法与生成式方法概念			
			GAN: 优化的目标函数，JS divergence, Wasserstein distance 概念，Wasserstein GAN			
14	2	第8章强化学习（续）： 内容：策略改进定理，策略评估方法，Q-Learning 算法，Deep Q-Learning 算法，策略梯度定理，REINFORCE 算法，使用基准函数的 REINFORCE 算法。 难点：策略优化定理，Deep Q-Learning 算法，策略梯度定理，使用基准函数的 REINFORCE 算法。	马尔可夫决策过程		2	基于CNN的MNIST手写字符识别
			策略改进定理			
			Q-Learning 算法，Deep Q-Learning 算法			
			策略梯度定理，REINFORCE 算法			
			使用基准函数的 REINFORCE 算法			
15	2	第8章 强化学习（续）， 第9章 人工智能博弈与安全： 内容：Advantage Actor-Critic 算法，RLHF, Imitation Learning的概念；反事实遗憾最小化算法，安全子博弈，数据安全与模型安全，对模型的攻击（白盒攻击，黑盒攻击，Non-Targeted and Targeted Attacks），对抗样本（PGD攻击，GAN生成）。 难点：RLHF，反事实遗憾最小化算法，对抗样本（PGD攻击，GAN生成）。	Advantage Actor-Critic 算法			
			反事实遗憾最小化算法			
			安全子博弈：概念			
			数据安全与模型安全：概念，对模型的攻击：白盒攻击，黑盒攻击			
16			对抗样本，对抗训练		4	Adversarial Examples for Classification

领导签字：

年 月 日