Week 1

Question 1

Which of the following best describes supervised learning?

- A. The model learns from labeled data
- B. The model finds hidden patterns in data without labels
- C. The model interacts with an environment to maximize rewards
- D. The model clusters similar data points together

答案正确性: 正确（A） 考点: 监督学习的核心定义（基于标记数据学习） 解析: 选项 B 是无监督学习，C 是强化学习，D 是聚类（无监督），符合 PPT 中监督学习 “labeled data with defined output” 的定义。

Question 2

Which of the following is NOT a type of supervised learning?

- A. Classification
- B. Regression
- C. Clustering
- D. Decision Trees

答案正确性: 正确（C） 考点: 监督学习的子类型区分 解析: 监督学习包括回归（B）和分类（A），决策树（D）是分类 / 回归算法；聚类（C）是无监督学习的子类型，故为正确答案。

Question 3

Which type of machine learning is primarily used for finding patterns in data without pre-existing labels?

- A. Supervised Learning
- B. Unsupervised Learning
- C. Reinforcement Learning
- D. Semi-Supervised Learning

答案正确性: 正确（B） 考点: 无监督学习的核心定义 解析: 无监督学习的核心是 “无标记数据中发现模式”，与选项 B 完全匹配；A 需标记数据，C 靠奖励机制，D 是半监督（结合标记与无标记数据）。

Question 4

Which of the following is a common challenge in supervised learning?

- A. The need for large amounts of labeled data
- B. The inability to handle structured data
- C. Difficulty in identifying patterns in high-dimensional data
- D. The lack of any need for feature engineering

答案正确性: 正确（A） 考点: 监督学习的常见挑战 解析: 监督学习依赖大量标记数据（PPT 提到 "labeling data is labour intensive"）；B 错误（监督学习可处理结构化数据），C 是无监督学习挑战，D 错误（部分监督任务需特征工程）。

Question 5

Labelling data is know to be labour intensive... However supervised learning requires a lot of data... How do we deal with this problem? Where there is a lot of data, but only a small subset is labelled.

答案:Semi-Supervised Learning 考点: 监督学习标记数据不足的解决方案 解析: 半监督学习结合少量标记数据和大量无标记数据，完美解决 "数据多但标记少" 的问题，符合 PPT 中提出的该场景痛点。

Week 2

Question 1

Which of the following methods is typically used for balancing exploration and exploitation?

- A. Gradient Descent
- B. Epsilon Greedy Strategy
- C. Back propagation
- D. Feature Selection

答案正确性: 正确（B） 考点: 强化学习的探索 - 利用平衡策略 解析: ε- 贪心策略（B）是平衡探索与利用的核心方法；A 是深度学习优化算法，C 是神经网络权重更新方法，D 是特征工程步骤，均与题意无关。

Question 2

A robot is trained using RL to navigate a maze. If it receives a +10 reward for reaching the goal and a -1 reward for each step taken, what type of behavior is the robot likely to develop?

- A. It will take the longest possible route to explore the environment
- B. It will learn the shortest path to the goal
- C. It will avoid reaching the goal to keep exploring
- D. It will take random actions indefinitely

答案正确性: 正确（B） 考点: 强化学习的奖励机制对行为的影响 解析: 到达目标获 + 10 奖励，每步 - 1 惩罚，智能体为最大化累积奖励，会学习最短路径（减少惩罚步数）；A、C、D 均违背奖励机制的导向。

Question 3

Which of the following is NOT a key component of an RL system?

- A. Environment
- B. Reward Signal
- C. Labeled Data
- D. Cost Function

答案正确性: 正确（CD） 考点: 强化学习的核心组件 解析: RL 的关键组件是 Agent、Environment、State、Action、Reward（PPT 明确说明）；C（Labeled Data）是监督学习的核心，RL 无标记数据；D（Cost Function）是监督 / 深度学习的概念，RL 核心是 Reward Function，故 CD 均不属于 RL 关键组件。

Question 4

Which RL algorithm estimates the value of state–action pairs and uses them to make decisions?

- A. Policy Gradient
- B. Q Learning
- C. Evolution Strategies
- D. Principal Component Analysis

答案正确性: 正确（B） 考点: 强化学习算法的核心功能 解析: Q 学习（B）的核心是估计状态 - 行动对（S-A）的价值（Q 值），并基于 Q 值决策；A 是基于策略的算法，C 是进化算法，D 是无监督降维算法。

Question 5

In RL, what does an “action” represent?

- A. The immediate reward received by the agent
- B. A choice made by the agent that affects the environment
- C. The function mapping states to values
- D. The final outcome of an episode
-

答案正确性: 正确（B） 考点: 强化学习中 “行动（Action）” 的定义 解析: 行动是智能体的选择，会影响环境状态（PPT 定义）；A 是 Reward，C 是 Value Function，D 是 Episode 的结果，均不符合。

Question 6

What is the key difference between on-policy and off-policy learning?

- A. Off-policy methods use past experiences, while on-policy methods only use current experiences
- B. On-policy methods always perform better than off-policy methods
- C. Off-policy methods do not require exploration
- D. On-policy methods use a separate network for learning

答案正确性: 正确（A） 考点: 离策略（Off-policy）与在策略（On-policy）的区别 解析: 离策略（如 Q 学习）可利用过去的经验数据，在策略（如 SARSA）仅使用当前策略产生的经验；B 错误（无绝对优劣），C 错误（离策略仍需探索），D 错误（与网络数量无关）。

Question 7

What is the goal of an agent in Reinforcement Learning (RL)?

- A. Minimise immediate rewards
- B. Maximise cumulative rewards
- C. Minimise the number of actions taken
- D. Determine Exploration–Exploitation Tradeoff

答案正确性: 正确（BD） 考点: 强化学习智能体的目标与核心权衡 解析: 智能体的核心目标是最大化累积奖励（B）；D（探索 - 利用权衡）是 RL 的核心问题，也是智能体需解决的关键；A 错误（应最大化奖励），C 错误（步数不是核心目标）。 Week 3 例题解析

Week 3

Question 1

Why has there been a resurgence of deep learning in recent years?

- A. Advancements in hardware (GPUs)
- B. Deep learning consumes less resources than traditional machine learning
- C. Explosion of data
- D. Deep learning automates feature selection

答案正确性: 正确（A, C, D） 考点: 深度学习复兴的原因 解析: PPT 提到深度学习复兴源于硬件进步（GPU，A）、数据爆炸（C）、自动特征提取（D）；B 错误（深度学习比传统 ML 消耗更多资源）。

Question 2

Which of the following best describes a neural network activation function（AF）?

- A. AF determines the learning rate
- B. AF initializes weights and biases
- C. AF introduces non-linearity
- D. AF determines the batch size

答案正确性: 正确（C） 考点: 激活函数的核心作用 解析: 激活函数的核心是引入非线性（PPT 明确说明），使网络能学习复杂模式；A 是学习率的作用，B 是权重初始化的作用，D 是批量大小的定义，均不符合。

Question 3

What is backpropagation in neural networks?

- A. BP initialises weights
- B. BP computes gradients to update weights

- C. BP normalises input data
- D. BP reduces network complexity

答案正确性: 正确（B） 考点: 反向传播（Backpropagation）的功能 解析: 反向传播通过链式法则计算损失函数对权重 / 偏置的梯度，用于更新参数（PPT 定义）； A 是权重初始化，C 是数据归一化，D 是网络简化，均不符合。

Question 4

What is the discount factor in deep learning?

- A. a parameter to stabilise gradient descent
- B. a parameter to discount future rewards
- C. a parameter to regularise training
- D. a parameter used for deep reinforcement learning

答案正确性: 正确（D） 考点: 折扣因子的应用场景 解析: 折扣因子（γ）是强化学习的核心参数（PPT Chapter2 第 17 页），而深度学习包含深度强化学习（DL ⊇ Deep RL），故折扣因子在深度学习中的应用场景是深度强化学习（D）； B 描述的是折扣因子的功能，但题目问的是 “在深度学习中” 的定义，故 D 更准确。

Question 4

How to address overfitting in deep learning

- A. Increase training data
- B. Increase the number of features
- C. Deactivate some neurons (dropout)
- D. Decrease the learning rate

答案正确性: 正确（A, C） 考点: 深度学习过拟合的解决方案 解析: 增加训练数据（A）可减少过拟合，Dropout（停用神经元，C）是常用正则化方法；B 错误（增加特征会加剧过拟合），D 错误（学习率与过拟合无关，影响收敛速度）。