Bloom's Level	Cognitive Process	Question (5 Marks)	Expected Focus in Answer
L1 – Remember	Recall facts & basic concepts	**Q1. Define supervised and unsupervised learning in machine learning. Give one real-world example of each.**	Definitions + 1 example each. (Short recall of definitions)
L2 – Understand	Explain ideas or concepts	**Q2. Explain why overfitting occurs in machine learning models and how cross-validation helps to prevent it.**	Describe concept of overfitting + brief explanation of cross-validation's role.
L3 – Apply	Use information in new situations	**Q3. A dataset contains heights and weights of students. Describe the steps you would take to train a linear regression model to predict weight from height.**	Outline data preprocessing, splitting, training, and prediction steps.
L4 – Analyze	Draw connections among ideas	**Q4. Compare and contrast decision trees and support vector machines in terms of interpretability, computational cost, and suitability for different data types.**	Identify similarities/differences and draw insights.
L5 – Evaluate	Justify a decision or course of action	**Q5. You are given a dataset with a large number of features. Argue whether feature selection or dimensionality reduction (like PCA) would be more appropriate, providing reasons for your choice.**	Provide reasoned judgment on which technique to use and why.
L6 – Create	Produce new or original work	**Q6. Propose a new application of reinforcement learning in an everyday setting (outside of gaming). Describe how the agent, environment, actions, and rewards would be defined.**	Develop an original example and describe components.

Q1 — Evaluation Answer Key & Perfect Baseline Answer

1) Marking rubric (total 5 marks)

- **Definition of supervised learning — 2 marks**
  Full credit (2): Clearly states it uses **labeled** data and learns a mapping from inputs → outputs to perform prediction (mentions classification/regression is a plus).
  Partial (1): Mentions labels or prediction but unclear on mapping idea.
  Zero (0): Wrong or missing.

- **Example of supervised learning — 1 mark**
  Full credit (1): A correct, brief real-world example (e.g., email spam classification, house-price regression, medical diagnosis).
  Partial (0.5): Example given but ambiguous or not clearly supervised.
  Zero (0): Wrong or missing.

- **Definition of unsupervised learning — 1.5 marks**
  Full credit (1.5): States it uses **unlabeled** data to discover structure/patterns (clustering, dimensionality reduction, association).
  Partial (0.75): Mentions unlabeled data or discovering structure but vague.
  Zero (0): Wrong or missing.

- **Example of unsupervised learning — 0.5 marks**
  Full credit (0.5): A correct, brief real-world example (e.g., customer segmentation, anomaly detection, topic modelling).
  Partial (0.25): Vague example.
  Zero (0): Wrong or missing.

---

2) Keywords/phrases to look for (useful for quick marking)

- **Supervised:** “labeled data”, “input → output mapping”, “prediction”, “classification”, “regression”.

- **Unsupervised:** “unlabeled data”, “discover structure/patterns”, “clustering”, “dimensionality reduction (PCA)”, “anomaly detection”.

3) Common student mistakes (for partial-credit guidance)

- Saying “supervised = training” without mentioning labels or outputs (award partial).
- Calling K-means supervised (it’s unsupervised) — treat as incorrect example.
- Mixing up goals: e.g., treating clustering as prediction — partial or zero depending on explanation.

---

4) Perfect answer (baseline — ~5-mark length)

Answer: Supervised learning is a class of machine learning methods that learn a mapping from input features to target outputs using **labeled** training data; its objective is to predict labels for new inputs (common tasks: classification and regression). *Example:* Email spam detection, where messages (features) are labeled “spam” or “not spam” and the model learns to predict that label for new emails.
Unsupervised learning operates on **unlabeled** data and seeks to **discover structure or patterns** such as clusters, associations, or lower-dimensional representations. *Example:* Customer segmentation using purchase histories (clustering) to group similar customers for targeted marketing.

If you want, I can: (a) convert this rubric into a short grading checklist you can print, or (b) produce 3 model answers at different quality levels (full, partial, poor) to help calibrate marking. Which would help you most?

Q2 — Evaluation Answer Key & Perfect Baseline Answer

1) Marking rubric (total 5 marks)

- **Definition + cause of overfitting — 2 marks**
  Full credit (2): Clear statement that overfitting occurs when a model learns the training data **too well**, capturing noise or idiosyncrasies instead of the underlying pattern, resulting in poor generalization to new data. Mentions typical causes: excessive model complexity relative to data, small training set, noisy labels, or lack of regularization.
  Partial (1): Vague definition (says “model fits data” but not noise/generalization) or

mentions only one cause.
Zero (0): Incorrect or missing.

- **Diagnosis / evidence of overfitting — 0.5 marks**
  Full credit (0.5): Notes the characteristic sign: low training error but substantially higher validation/test error (or widening train–validation error gap).
  Partial (0.25): Mentions error but not the contrast between train and validation.
  Zero (0): Missing.

- **Explanation of cross-validation (mechanics) — 1.25 marks**
  Full credit (1.25): Describes k-fold cross-validation (split data into k folds, train on k−1 folds and validate on the held-out fold, repeat k times), or an equivalent scheme (leave-one-out, stratified). Explains that it produces multiple validation scores and an aggregate estimate of generalization performance.
  Partial (0.6): Mentions splitting into train/validation or folds but lacks repetition/aggregation detail.
  Zero (0): Missing/incorrect.

- **How cross-validation helps prevent overfitting (practical role) — 1.25 marks**
  Full credit (1.25): Explains that CV gives a more reliable estimate of generalization used for model selection and hyperparameter tuning, helps choose simpler/less overfit models, and reduces the chance of selecting a model that only performs well on one particular split. Mentions that averaging across folds reduces variance of the estimate.
  Partial (0.6): States that CV is used for model selection but not how it prevents overfitting.
  Zero (0): Missing.

---

2) Keywords / phrases to look for

- **Overfitting:** “noise”, “poor generalization”, “high variance”, “model complexity”, “training vs validation error”.

- **Cross-validation:** “k-fold”, “folds”, “train on k−1 / validate on 1”, “aggregate/average validation score”, “model selection / hyperparameter tuning”, “reduces variance of estimate”.

---

3) Common student mistakes (grading tips)

- Saying “overfitting = high error” without specifying *on which data* — should be high on unseen/validation but low on training.

- Describing cross-validation as just “train/validation split” without mentioning repeated folds and aggregation — give partial credit.

- Claiming cross-validation *directly* reduces training overfitting (it doesn’t change the trained model unless used for selection/regularization or to pick hyperparameters) — award partial credit but expect explanation that CV is a method for evaluation and selection.

- Confusing cross-validation with regularization (they are complementary).

---

4) Perfect answer (baseline — ~5-mark length)

Answer (baseline):
Overfitting happens when a model fits the training data too closely, learning noise or idiosyncratic patterns instead of the true underlying relationship; as a result the model shows low training error but performs poorly on new, unseen data. Common causes are excessive model complexity relative to the amount of data, noisy or mislabeled examples, and insufficient regularization. Cross-validation (for example, k-fold CV) mitigates this risk by repeatedly splitting the dataset into k folds, training on k−1 folds and validating on the held-out fold, then averaging the validation scores across folds. That averaged estimate gives a more reliable measure of generalization and is used when selecting models and tuning hyperparameters; by choosing the model/hyperparameter configuration that performs best under cross-validation, we reduce the likelihood of picking a model that is overfitted to a single train/validation split. (Optional note: CV increases computational cost but provides a robust selection criterion.)