Module 2: Machine Learning — 11 Lessons (Paragraph Edition)

With Key Terms and Lesson Quick■Check MCQs

Lesson 1 — Types of Learning

Machine learning problems fall into three families. In supervised learning, models learn a mapping from inputs to known labels. In unsupervised learning, algorithms discover structure without labels. In reinforcement learning, an agent interacts with an environment and learns a policy that maximizes long■term reward.

Key terms: supervised, unsupervised, reinforcement

Lesson 2 — The ML Data Pipeline

The standard workflow turns raw information into reliable predictions: dataset, preprocessing, training, and prediction (inference) on unseen data.

Key terms: dataset, preprocessing, inference

Lesson 3 — Linear Regression

Linear regression models the relationship between a numeric target and features using a straight line. Training selects slope and intercept that minimize mean squared error between true and predicted values.

Key terms: slope, intercept, MSE

Lesson 4 — Classification

Classification assigns inputs to discrete classes. k■nearest neighbors (k■NN) finds the k closest labeled neighbors and returns the majority label. Decision trees can capture nonlinear rules.

Key terms: classification, k■NN, boundary

Lesson 5 — Overfitting vs Underfitting

Underfitting occurs when a model is too simple, yielding high errors. Overfitting happens when it memorizes noise, achieving low training error but poor test performance.

Key terms: bias, variance, generalization

Lesson 6 — Train/Validation/Test & Cross■Validation

Split data into train, validation, and test sets. Hyperparameters are tuned on validation (or via k■fold cross■validation), and test is used only once for unbiased evaluation.

Key terms: validation, cross■validation, holdout

Lesson 7 — Evaluation Metrics

Regression metrics include MAE, MSE, and R^2. Classification metrics include accuracy, precision, recall, F1, and confusion matrix.

Key terms: precision, recall, F1

Lesson 8 — Feature Engineering & Preprocessing

Feature engineering includes scaling, encoding, and imputing. For text, bag■of■words and TF■IDF create count■based representations. Preprocessing must avoid leakage.

Key terms: scaling, encoding, imputation

Lesson 9 — Model Selection & Hyperparameter Tuning

Model selection chooses configurations that perform best on validation data. Hyperparameters are tuned via grid/random search with cross■validation.

Key terms: hyperparameters, cross■validation, search

Lesson 10 — Clustering & Dimensionality Reduction

K■means clustering assigns points to centroids iteratively. PCA reduces dimensionality by projecting onto directions of maximum variance.

Key terms: k■means, PCA, clusters

Lesson 11 — JavaScript Project: Mini ML Lab

Firstly, build an interactive linear regression demo where the algorithm fits a line to user■plotted points. Secondly, create a k■NN classification playground to visualize decision boundaries. Finally, add a k■means clustering explorer where users generate random points and watch centroids update step by step.

Key terms: regression, classification, clustering

Lesson Quick■Check MCQs

Lesson 1 — Types of Learning

Q: Which task is most suitable for reinforcement learning?

A. Grouping similar samples without labels

B. Training an agent to maximize long■term reward through trial and error

C. Predicting prices from labeled examples

D. Compressing features into fewer dimensions

Answer: Training an agent to maximize long■term reward through trial and error

Lesson 2 — The ML Data Pipeline

Q: What is the correct order of the pipeline described in the lesson?

A. Preprocessing → Dataset → Training → Prediction

B. Dataset → Training → Preprocessing → Prediction

C. Dataset → Preprocessing → Training → Prediction

D. Training → Dataset → Prediction → Preprocessing

Answer: Dataset → Preprocessing → Training → Prediction

Lesson 3 — Linear Regression

Q: Which quantity is minimized during training to fit the line?

A. Cross■entropy

B. Mean squared error

C. Hinge loss

D. Kullback–Leibler divergence

Answer: Mean squared error

Lesson 4 — Classification

Q: Increasing k in k■NN typically has what effect on the decision boundary?

A. Makes it more jagged (higher variance)

B. Smooths it (lower variance)

C. Always increases training accuracy without test impact

D. Guarantees underfitting on every dataset

Answer: Smooths it (lower variance)

Lesson 5 — Overfitting vs Underfitting

Q: Which pattern indicates overfitting as defined in the lesson?

A. Low training error and low test error

B. High training error and high test error

C. Low training error but high test error

D. High training error but low test error

Answer: Low training error but high test error

Lesson 6 — Train/Validation/Test & Cross■Validation

Q: According to the lesson, when should the held■out test set be used?

A. For selecting hyperparameters

B. Only once at the end to estimate generalization

C. For computing scaling parameters

D. To augment the training data when data is scarce

Answer: Only once at the end to estimate generalization

Lesson 7 — Evaluation Metrics

Q: Which metric combines precision and recall into a single number?

A. Accuracy

B. ROC■AUC

C. F1 score

D. R^2

Answer: F1 score

Lesson 8 — Feature Engineering & Preprocessing

Q: Which practice can cause data leakage as described in the lesson?

A. Fitting a scaler on the training set then applying to validation/test

B. One■hot encoding categorical variables

C. Imputing missing values using training■set statistics only

D. Fitting a scaler on the entire dataset before splitting

Answer: Fitting a scaler on the entire dataset before splitting

Lesson 9 — Model Selection & Hyperparameter Tuning

Q: Which is a hyperparameter according to the lesson?

A. The learned weight w in linear regression

B. The number of neighbors k in k■NN

C. The predicted probability for a test sample

D. The residual error for a training point

Answer: The number of neighbors k in k■NN

Lesson 10 — Clustering & Dimensionality Reduction

Q: What is the primary function of PCA as defined in the lesson?

A. Assign points to nearest centroids iteratively

B. Project data onto directions of maximum variance

C. Train a deep neural network

D. Guarantee higher classification accuracy

Answer: Project data onto directions of maximum variance

Lesson 11 — JavaScript Project: Mini ML Lab

Q: Which algorithms are demonstrated in the Mini ML Lab project?

A. Decision trees, SVM, Naive Bayes

B. Linear regression, k■NN, k■means

C. Logistic regression, Random Forest, PCA

D. Q■learning, Monte Carlo search, CNNs

Answer: Linear regression, k■NN, k■means