

Circle the correct answers
(Wrong will be penalized)

1. In supervised learning, the training dataset generally contains:

- A. Only input vectors
- B. Random unlabeled vectors
- C. Input-output pairs**
- D. Only predicted labels

2. Which learning paradigm mainly uses unlabeled data?

- A. Supervised learning
- B. Unsupervised learning**
- C. Reinforcement learning
- D. Online learning

3. In machine learning, a feature usually represents:

- A. A compilation error
- B. An evaluation metric
- C. A hidden neuron
- D. An input attribute**

4. Which dataset split is commonly used to estimate final generalization performance?

- A. Training set
- B. Validation set
- C. Test set**
- D. Support set

5. Predicting a house price is mainly considered:

- A. Binary classification
- B. Multi-label learning
- C. Regression**
- D. Clustering

6. Which prediction type produces discrete values?

- A. Classification**
- B. Regression
- C. Generation
- D. Interpolation
- E. Compression

7. One-hot encoding is mainly used to:

- A. Reduce model accuracy
- B. Normalize gradients
- C. Represent categorical values numerically**
- D. Remove labels
- E. Encode continuous values

Correct = +0.5

Wrong = -0.25

8. Which feature type naturally contains an order relation?

- A. Ordinal**
- B. Binary
- C. Nominal
- D. Boolean
- E. Continuous

9. In K-NN, the parameter K refers to:

- A. Number of datasets
- B. Number of hidden layers
- C. Number of epochs
- D. Number of nearest neighbors**

10. K-NN classification usually assigns:

- A. The smallest feature value
- B. The majority class among neighbors**
- C. A random category
- D. The average distance

11. The Euclidean distance mainly measures:

- A. Geometric separation between vectors**
- B. Average probability error
- C. Similarity through labels only
- D. Dataset imbalance
- E. Difference between actual and predicted labels

12. Overfitting usually means:

- A. High training error and low test error
- B. Low training error and high test error**
- C. High training error and high test error
- D. Zero validation data
- E. Very small datasets only

13. Underfitting generally occurs when:

- A. The model fails to capture patterns**
- B. The model memorizes all data
- C. The dataset is normalized
- D. The optimizer converges perfectly

14. Which loss is commonly used in linear regression?

- A. Binary-Cross Entropy
- B. Categorical entropy
- C. Mean Squared Error**
- D. Dice loss

15. In a confusion matrix, accuracy corresponds approximately to:

- A. Incorrect predictions / total predictions
- B. Correct predictions / total predictions**
- C. False positives / false negatives
- D. Training loss / epochs

16. Which statement best describes reinforcement learning?

- A. Model learns only from static labels
- B. Model never updates parameters
- C. Model ignores environments
- D. Model receives rewards from interactions**

17. Which Python library is used to store a K-NN trained model?

- A. pandas
- B. numpy
- C. matplotlib
- D. joblib**
- E. tensorflow

18. In the classification training, the labels are stored in:

- A. X
- B. y**
- C. label
- D. probs
- E. sample

19. What does train_test_split() mainly do?

- A. Converts labels to strings
- B. Removes outliers
- C. Train and test the model
- D. Divides data into subsets**
- E. Computes accuracy

20. In the statement df.drop(columns=["class"]), the result mainly contains:

- A. Features only**
- B. Labels only
- C. Features and labels
- D. Predictions only

21. Which sklearn method trains the K-NN classifier?

- A. predict()
- B. compile()
- C. train()
- D. score()
- E. fit()**

22. In the sklearn linear regression, model.coef_ mainly stores:

- A. Training labels
- B. Learned weights**
- C. Learned biases
- D. Validation samples
- E. Activation outputs

23. Which metric is computed using mean_squared_error?

- A. L1 Loss
- B. L2 Loss**
- C. mean absolute error
- B. Accuracy
- D. Recall

24. The perceptron output becomes 1 when:

- A. The weighted sum exceeds the threshold**
- B. All weights are negative
- C. Inputs are always binary zero
- D. The loss is maximal

25. In ANN terminology, synapses are analogous to:

- A. Labels
- B. Distances
- C. Weights**
- D. Biases
- E. Epochs

Exercise (7.5 pts): We aim to solve a color classification problem using an artificial neural network (Simple Perceptron Model). There is a training sample consisting of 4 examples divided into two color classes: RED (represented by 1) and BLUE (represented by 0). Each example is described by its RGB representation (see Table 1). The goal is to train the perceptron model using an Error-Correction algorithm in order to determine the class of each example.

	R	G	B
RED	255	0	0
	248	80	68
BLUE	0	0	255
	67	15	210

Table 1

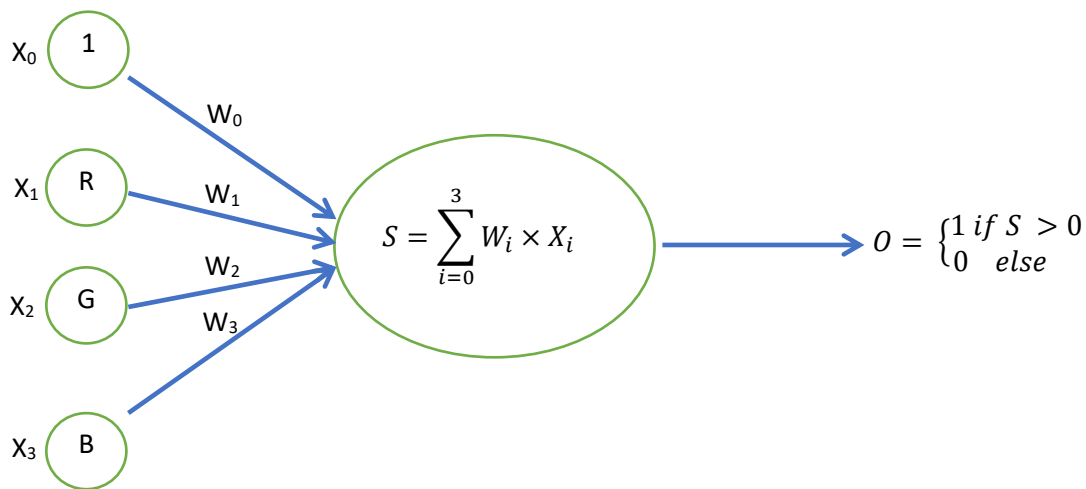
- 1) List the different parameters used by the model.
- 2) Draw the model architecture.
- 3) Illustrate the training process using a table.
 - Stopping criterion: Number of epochs = 2.
- 4) Comment on the results.

Solution :

Model Parameters : (1.5 pt)

- Input vector : [1,R,G,B, Class], $X_0 = 1, X_1 = R, X_2 = G, X_3 = B$, Class = 1 if Red, Class = 0 if Blue (0.25)
- Initial weights : $W_0 = -1, W_1 = 2, W_2 = 1, W_3 = -2$ (Random real values) (0.25)
- The summation $S = \sum_{i=0}^3 W_i \times X_i$ (0.25)
- The output $O = \begin{cases} 1 & \text{if } S > 0 \\ 0 & \text{else} \end{cases}$ (0.25)
- Weights update : $W_i = W_i + (C - O) \times X_i$ (0.25)
- Training sample: (1.255.0.0,1), (1.284.80.68,1), (1.0.0.255,0), (1.67.15.210,0) (0.25)

Model architecture : (02 pts)



Training process : (03 pts)

Iterations		W_0	W_1	W_2	W_3	X_0	X_1	X_2	X_3	$\sum w_i \cdot x_i$	O	C	W_0	W_1	W_2	W_3
Epoch 1	Example 1	-1	2	1	-2	1	255	0	0	509	1	1	-1	2	1	-2
	Example 2	-1	2	1	-2	1	284	80	68	475	1	1	-1	2	1	-2
	Example 3	-1	2	1	-2	1	0	0	255	-511	0	0	-1	2	1	-2
	Example 4	-1	2	1	-2	1	67	15	210	-272	0	0	-1	2	1	-2
Epoch 2	Example 1	-1	2	1	-2	1	255	0	0	509	1	1	-1	2	1	-2
	Example 2	-1	2	1	-2	1	284	80	68	475	1	1	-1	2	1	-2
	Example 3	-1	2	1	-2	1	0	0	255	-511	0	0	-1	2	1	-2
	Example 4	-1	2	1	-2	1	67	15	210	-272	0	0	-1	2	1	-2

Comment on the results: (01 pt)

Weight stabilization occurs immediately during the first iteration*, indicating that the perceptron effectively converges without difficulty to classify the training sample, which further suggests an optimal initialization of the weight vector.

* Different weights lead to different results