Machine Learning Self-Study Worksheet

Worksheet: 2025003

Level: Easy
Topics Covered: Supervised Learning

Part 1: Conceptual Questions

<table>
<tr><th>Q. No.</th><th>Question</th><th>Points</th></tr>
<tr><td>1.</td><td>Match learning tasks with appropriate algorithms:
Match each task with the most suitable supervised learning algorithm from the list below. Justify.
Algorithms: Logistic Regression, Linear Regression, Decision Tree, KNN

• Predicting a student’s test score based on hours studied. → _____
• Classifying tumor as benign or malignant based on cell features. → _____</td><td>5</td></tr>
<tr><td>2.</td><td>Complete the following table comparing the two main types of supervised learning:

<table><tr><th>Aspect</th><th>Regression</th><th>Classification</th></tr><tr><td>Output Type</td><td></td><td></td></tr><tr><td>Performance Metrics</td><td></td><td></td></tr><tr><td>Common Algorithms</td><td></td><td></td></tr></table></td><td>5</td></tr>
</table>

Part 2: Numerical Questions

<table>
<tr><th>Q. No.</th><th>Question</th><th>Points</th></tr>
<tr><td>3.</td><td>You are given the following actual and predicted values from a regression model:

<table><tr><th>Data Point</th><th>Actual (Y)</th><th>Predicted (Ŷ)</th></tr><tr><td>1</td><td>3</td><td>2.5</td></tr><tr><td>2</td><td>5</td><td>5.5</td></tr><tr><td>3</td><td>2</td><td>3.0</td></tr><tr><td>4</td><td>7</td><td>6.5</td></tr></table>• Calculate the Mean Squared Error (MSE).</td><td>5</td></tr>
</table>

4.	Given the data points: • Class A: (1, 1), (1, 2), (2, 2) • Class B: (6, 6), (7, 5), (7, 6) Use **Euclidean distance** and determine the class of the point (3, 3) using **1-nearest neighbor** and **3-nearest neighbors**. → Report both class predictions.	5

Part 3: Coding Exercise

Q. No.	**Complete & Run the Code** (Use your own Python IDE or Google Colab.) Complete the KNN classification code for k = 3 and predict the class of a new point:	**Points**

5.	from sklearn.tree import _____
	from sklearn.model_selection import train_test_split
	from sklearn.metrics import classification_report
	
	# Sample data
	X = [[1], [2], [3], [4], [5], [6]]
	y = [0, 0, 0, 1, 1, 1]
	
	X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
	
	# Create the model
	model = DecisionTreeClassifier()
	model._____(X_train, y_train)
	
	# Predict
	y_pred = model.predict(X_test)
	print(classification_report(y_test, y_pred))
	print("Accuracy:", accuracy_score(y_test, y_pred))
		5

Notes & Disclaimer

- For quick notes or queries, visit: https://www.instagram.com/kparse_code/ or https://www.youtube.com/channel/UCGuhk1P1OA0X_7Ar2-xadEw .

- **Related Resources**:
  - Google Colab – to run Python code online without installing anything
  - StatQuest: https://www.youtube.com/@statquest/playlists
  - 3Blue1Brown: https://www.youtube.com/@3blue1brown/playlists
  - Numerical: https://www.youtube.com/@MaheshHuddar/playlists

- **Want more worksheets like this?**
  Try using **AI prompt** like:
  "Create a beginner-level machine learning worksheet with 2 conceptual questions(with sub-questions), 2 numerical, and 1 code-completion exercise on [topic(s)]"

- **How to Use This Worksheet**:
  - These worksheets are designed for **self-study and concept reinforcement**.
  - Try to solve without looking up the answers first.
  - Use a notebook or Python IDE to experiment — code is the best teacher!
  - Feel free to modify or extend questions as your understanding grows.

- **Disclaimer**:
  This worksheet is for self study & **educational purposes only**. Accuracy of content may vary depending on updates to libraries or definitions. Always refer to official documentation or textbooks for exam-level prep.