

Level: 1st year “Computer Science”
Module: Introduction to Artificial Intelligence

Date: 14/05/2026
Duration: 1h30m

Exam n°2

Typical Correction

Comprehension questions **(9 points)**

Complete the following statements

1. Artificial intelligence (AI) **is a field of computer science that allows machines (algorithms) to simulate human intelligence.** **(0.5 p)**
2. What is the difference between AI, machine learning, and deep learning? **(1 p)**
AI: wide field of intelligent systems.
Machine Learning: a subset of AI where systems learn from data.
Deep Learning: a subset of machine learning using neural networks with many layers.
3. Based on the given data, there are two primary categories of machine learning algorithms: **(1 p)**
 - a. **In Supervised learning, the data are labeled.**
 - b. **In Unsupervised learning, the data are unlabeled.**
4. Data can be of many types; it can **be structured or unstructured.** **(0.5 p)**
5. In machine learning, more data always improves **learning** and leads to better **performance.** **(0.5 p)**
6. Data preprocessing is a crucial step that enhances accuracy and reliability. Explain at least two techniques used in this process. **(0.5 p)**
Data Cleaning: remove errors and missing values. Normalization: scale values.
7. AI improves creativity by **generating ideas quickly.** **(0.5 p)**
8. AI scheduling mainly helps to **optimize time.** **(0.5 p)**
9. The Todoist tool is mainly used for **task organization.** **(0.5 p)**
10. Brainstorming is **a group problem-solving technique** that requires **free thinking.** **(0.5 p)**
11. A plagiarism detector identifies **copied human works (if text matches existing, human-authored content),** while an AI detector **identifies machine-generated text.** **(0.5 p)**
12. Bias is one of the most significant problems in AI; it can emerge particularly during data collection, where the dataset may appear **imbalanced.** **(0.5 p)**
13. What is the difference between Elicit and Scholarcy tools? **(0.5 p)**
Elicit: used to find and analyze research papers via a request entered by the user.
Scholarcy: used to summarize and simplify documents.
14. When using academic databases or AI-powered search tools, an intelligent query means **formulating a clear, organized, and precise search request to obtain relevant and high-quality results.** **(0.5 p)**
15. K-Nearest Neighbors (KNN) is a simple supervised machine-learning algorithm. Explain the workflow of this algorithm. **(1 p)**
KNN classifies data points (samples) based on their k closest neighbors in the dataset. The idea is that similar points are near each other. In a classification task, the algorithm uses distance metrics, such as Euclidean distance, to find neighbors and a majority vote of its neighbors' labels to determine the appropriate label.

Case study analysis

(3 points)

A hospital uses an Artificial Intelligence system that analyzes **X-ray images** to detect tumors (exist “1”, not-exist “0”).

Questions

1. Identify the AI domain
Computer Vision, because the system analyzes images to detect tumors. (0.5 p)
2. Is this Weak or Strong AI? Justify
Weak AI, because it performs a specific task (tumor detection). (1 p)
3. Is this Symbolic or Connectionist AI?
Connectionist AI. Uses deep learning to learn from a large image dataset. (1 p)
4. What type of learning is used?
Supervised learning. (0.5 p)

Practical Questions

(8 points)

1. What are the advantages of using Google Colab compared to installing Python locally? Give 2 advantages **(0.5 p)**
 - No installation needed.
 - Free access to GPU/TPU.
2. What is the difference between: Runtime RAM and Storage in Google Drive? **(0.5 p)**
 - RAM (Runtime memory): Temporary, erased when the session ends.
 - Google Drive: Permanent storage, files stay saved.
3. Where does Colab execute your code? **(0.5 p)**
On remote cloud servers provided by Google.
4. What command is used to install modules (libraries) in Python? **(0.5 p)**
pip install <package_name> .
5. What does the following plot represent? Confusion matrix (0.5 p)

		Predicted Values		
		Setosa	Versicolor	Virginica
Actual Values	Setosa	16 <small>(cell 1)</small>	0 <small>(cell 2)</small>	0 <small>(cell 3)</small>
	Versicolor	0 <small>(cell 4)</small>	17 <small>(cell 5)</small>	1 <small>(cell 6)</small>
	Virginica	0 <small>(cell 7)</small>	0 <small>(cell 8)</small>	11 <small>(cell 9)</small>

6. Which cells represent true positives and true negatives for the Setosa class? **(0.5 p)**

TP: the value of cell 1 is the TP value, which is 16. TN= (cell 5 + cell 6 + cell 8 + cell 9)=17 + 1 +0 + 11=29

7. Complete with the required libraries: **(0.75 p)**
 - a. The library that provides tools for modeling and machine learning algorithms is the scikit-learn (sklearn) library.

- b. The library used for creating static, animated, and interactive visualizations is the **matplotlib library**.
- c. The library used for data analysis, manipulation, and cleaning is the **pandas library**.
- 8. To build a reliable predictive model, write the corresponding Python program using the KNN classifier and the Iris dataset. Outline the steps involved in a standard machine-learning pipeline.

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
# Load data /Data collection
data = load_iris()
X = data.data
y = data.target
# Split data into training and testing datasets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
# Build the model
model = KNeighborsClassifier(n_neighbors=3)
#Train the model
model.fit(X_train, y_train)
# Prediction & evaluation
from sklearn import metrics
y_pred=model.predict(X_test)
print ("Accuracy",metrics.accuracy_score(y_test,y_pred))
```

(0.75 p)

(0.5 p)

(0.5 p)

(0.5 p)

(0.5 p)

(0.5 p)

(0.5 p)

(0.5 p)