ARTIFICIAL INTELLIGENCE
PAPER 1
(THEORY)

Maximum Marks: 70
Time Allotted: Three Hours
Reading Time: Additional Fifteen Minutes

Instructions to Candidates

1. You are allowed **additional fifteen minutes** for **only** reading the question paper.
2. You must **NOT** start writing during the reading time.
3. This question paper has **7 printed pages**.
4. It is divided into **two parts** and has **9 questions** in all.
5. **Part I** is compulsory and has **two** questions.
6. **Part II** is divided into **seven questions**. Answer **any five** questions.
7. While attempting **Multiple Choice Questions** in Part I, you are required to **write only ONE option as the answer**.
8. **Each question in Part II** has **three sub parts. Any five** questions have to be attempted.
9. The intended marks for questions are given in brackets [ ].

Instruction to Supervising Examiner

1. Kindly read **aloud** the Instructions given above to all the candidates present in the examination hall.

Note: The Specimen Question Paper in the subject provides a realistic format of the Board Examination Question Paper and should be used as a practice tool. The questions for the Board Examination can be set from any part of the syllabus, though the format of the Board Examination Question Paper will remain the same as that of the Specimen Question Paper. The weightage allocated to various topics, as given in the syllabus, will be strictly adhered to.

PART I (20 MARKS)

*Answer **all** questions.*

While answering questions in this Part, indicate briefly your working and reasoning, wherever required.

Question 1

(i) Which of the following is **NOT** a commonly used application of Artificial Intelligence? **[1]** **(Recall)**

(a) Email Spam Filtering

(b) Online Shopping Recommendations

(c) Manual Typewriting

(d) Voice Assistants

(ii) Given below are two statements marked Assertion and Reason. Read the statements carefully and choose the correct option. **[1]**

Assertion: Data cleaning improves the quality of data for analysis.

Reason: Data cleaning helps in generating random missing values for analysis.

(Analysis)

(a) Both Assertion and Reason are true and Reason is the correct explanation for the Assertion.

(b) Both Assertion and Reason are true but Reason is not the correct explanation for the Assertion.

(c) Assertion is true and Reason is false.

(d) Both Assertion and Reason are false.

(iii) Which stage in an AI project involves understanding the business issue and defining objectives? **[1]** **(Recall)**

(a) Data Exploration

(b) Problem Scoping

(c) Modelling

(d) Evaluation

(iv) Which of the following is an example of a Natural Language Processing (NLP) application? **[1]** **(Recall)**

(a) Image recognition

(b) Crop monitoring

(c) Language translation

(d) Transport management system

(v) Given below are two statements marked Assertion and Reason. Read the statements carefully and choose the correct option. **[1]**

Assertion: The drop duplicates() method in Pandas is used to clean a dataset by removing redundant entries.

Reason: Duplicated rows can lead to skewed or inaccurate results during data analysis. **(Analysis)**

(a) Both Assertion and Reason are true and Reason is the correct explanation for the Assertion.

(b) Both Assertion and Reason are true but Reason is not the correct explanation for the Assertion.

(c) Assertion is true and Reason is false.

(d) Both Assertion and Reason are false.

(vi) Given below are two statements marked Assertion and Reason. Read the statements carefully and choose the correct option. **[1]**

Assertion: AI systems can unintentionally amplify existing social biases in training data.

Reason: Bias in AI arises only when developers deliberately add it. **(Analysis)**

(a) Both Assertion and Reason are true and Reason is the correct explanation for the Assertion.

(b) Both Assertion and Reason are true but Reason is not the correct explanation for the Assertion.

(c) Assertion is true and Reason is false.

(d) Both Assertion and Reason are false.

(vii) Which of the following operations is valid for two matrices of the same order? **[1]** **(Recall)**

(a) Finding inverse

(b) Scalar product

(c) Determinant

(d) Addition

(viii) Which of the following measures is least affected by extreme values in a dataset? **[1]** **(Recall)**

(a) Mean

(b) Median

(c) Mode

(d) Variance

(ix) Define *artificial intelligence*. Mention *one* application of AI in Transportation. **[1]** **(Recall)**

(x) What is the primary goal of data visualisation in AI? **(Recall)** **[1]**

Question 2

(i) What is the difference between *deterministic* and *probabilistic systems*? Give *one* example of each. **[2]** **(Understanding)**

(ii) Write a Python command using Matplotlib to display a line graph with x = [1, 2, 3] and y = [3, 2, 1]. **[2]** **(Create)**

(iii) Mention *two* chart types and briefly state when each is most suitable. **[2]** **(Understanding)**

(iv) Differentiate between dependent and independent variables in the context of regression analysis with *one* example. **[2]** **(Understanding)**

(v) List *two* ethical responsibilities developers must consider while creating AI systems. **[2]** **(Recall)**

PART II (50 MARKS)

*Answer **any five** questions.*

The answers in this section should consist of the programs in either python environment or any program environment with python as the base.

Each program should be written using variable description / mnemonic codes so that the logic of the program is clearly depicted.

Question 3

(i) Explain the role of AI in improving healthcare and agriculture. Mention *two* specific uses for each domain. **[4]** **(Application & Analysis)**

(ii) Write a Python function check_even_odd(n) that takes a number as input and prints whether it is even or odd. Also, call the function with n = 7. **[3]** **(Create)**

(iii) Differentiate between *Machine Learning (ML)* and *Deep Learning (DL)*. Mention *one* real-world application of each. **[3]** **(Understanding)**

Question 4

(i) Compare Matplotlib and Seaborn libraries in terms of features, ease of use, and application. Which one would you recommend for statistical plots? **[4]** **(Analysis)**

(ii) Explain the advantages of using data visualisation in the field of AI. **[3]** **(Understanding)**

(iii) Write a Python program using Matplotlib to create a bar chart showing the number of students interested in three domains: AI (30), Robotics (20), IoT (15) **[3]** **(Create)**

Question 5

(i) As a data analyst for the company Shopping Paradise, you are analysing data of online customer purchases. The data from two tables is given below:

Table A: Customer Info

ID	NAME	CITY
1	AMAN	DELHI
2	BEENA	MUMBAI
3	CHITRA	CHENNAI

Table B: Purchases

ID	PRODUCT	PRICE
2	LAPTOP	55000
1	KEYBOARD	1500
3	MOUSE	600

Answer the following:

(a) Identify the type of join needed to combine customer details with their purchases. **(Application)** **[1]**

(b) Perform the join and write the resulting table. **(Application)** **[1]**

(c) Explain how this is an application of **Set Theory** in AI data preparation. **(Analysis)** **[2]**

(ii) Consider the matrices and perform the below: **(Application)**

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

(a) Perform A + B **[1]**

(b) Find the transpose of matrix A **[1]**

(c) Determine the product A × B **[1]**

(iii) A dataset records the scores of 5 students in a test: 55, 60, 65, 70, 95. **(Application)**

(a) Calculate the mean score. **[1]**

(b) Identify the median. **[1]**

(c) Is the mode defined for this dataset? Justify your answer. **[1]**

Question 6

(i) You downloaded a CSV file from Kaggle containing missing values, duplicate entries, and inconsistent column formatting.

Write a Python program using Pandas to: **(Create)**

(a) Load the file. **[1]**

(b) Remove duplicates. **[1]**

(c) Fill the missing values with the column mean. **[1]**

(d) Convert a column called City to title case. **[1]**

(ii) What are the *three* common data cleaning tasks performed before analysing a dataset? Illustrate each briefly. **[3]** **(Understanding)**

(iii) What is the purpose of a Least Squares Regression Line in Linear Regression? Explain how it is determined. **[3]** **(Understanding)**

Question 7

(i) The table below shows the number of hours studied and marks obtained by students: **[4]**

Hours Studied (X)	Marks (Y)
2	50
3	60
5	80

Using the method of least squares, find the linear regression equation of Y on X. **(Application)**

(ii) List and briefly describe *any three* types of data models used in data modelling. **[3]** **(Understanding)**

(iii) Explain with an example how poor ethical considerations in AI design can negatively impact society. **[3]** **(Evaluate)**

Question 8

(i) You are working with a Kaggle dataset on housing prices. Write a Python code to: **(Create)**

(a) Remove duplicate rows **[1]**

(b) Handle missing values in the 'Price' column **[1]**

(c) Standardize the 'Area' column using z-score normalization **[2]**

(ii) Differentiate between Narrow AI, General AI, and Superintelligent AI with one real-world or hypothetical example for each. **[3]** **(Understanding)**

(iii) Write a Python program to plot a bar chart showing the number of AI projects completed by four students: A: 3, B: 5, C: 2, D: 4 **[3]** **(Create)**

Question 9

(i) You are given a dataset containing the number of study hours and corresponding student scores.

You plan to apply **Linear Regression** to predict scores.

Before that, answer the following: **(Analysis)**

(a) What type of data modelling is suitable for this problem and why? **[1]** **(Application)**

(b) Define dependent and independent variables in this context. **(Analysis)** **[1]**

(c) Explain why bias in data collection must be avoided when training this model. **[2]** **(Understanding)**

(ii) Describe the role of data acquisition and data exploration in an AI project. Why are these steps critical? **[3]** **(Understanding)**

(iii) Explain how environmental concerns should be addressed in AI development. Provide *one* example. **[3]** **(Evaluate)**

ARTIFICIAL INTELLIGENCE
PAPER 1
(THEORY)
ANSWER KEY

PART I (20 MARKS)

Question 1

(i) (c) or Manual Typewriting **[1]**

(ii) (c) or Assertion is true and Reason is false. **[1]**

(iii) (b) or Problem Scoping **[1]**

(iv) (c) or Language translation **[1]**

(v) (a) or Both Assertion and Reason are true and Reason is the correct explanation for the Assertion. **[1]**

(vi) (c) or Assertion is true and Reason is false. **[1]**

(vii) (d) or Addition **[1]**

(viii) (b) or Median **[1]**

(ix) Artificial Intelligence is the ability of machines to mimic human intelligence and perform tasks such as learning, reasoning, and problem-solving. One application in transportation is AI driven traffic Management Systems. **[1]**

(x) To represent data graphically for easier understanding of patterns, trends, and insights. **[1]**

Question 2

(i) **Deterministic systems** always produce the same output for a given input. Example: A calculator performing 2 + 2. **[2]**

Probabilistic systems may produce different outcomes for the same input due to uncertainty. Example: A weather prediction model.

(ii) import matplotlib.pyplot as plt **[2]**

```
plt.plot([1, 2, 3], [3, 2, 1])
plt.show()
```

(iii) Bar Chart - Best for comparing values across categories. **[2]**
Scatter Plot- Useful to identify relationships or correlations between two variables.

(iv) Dependent Variable: The outcome or variable being predicted (e.g., house price). Independent Variable: The predictor or input variable (e.g., square footage). Example: In predicting a student's score based on hours studied: **[2]**

Hours studied → Independent variable

Exam score → Dependent variable

(v) Ensuring the privacy and security of user data. Minimizing algorithmic bias to promote fairness and equity. **[2]**

PART II (50 Marks)

Question 3

(i) In Healthcare: **[4]**

1. Diagnosis Enhancement – AI helps doctors detect diseases early using image-based diagnosis like X-rays and MRIs.

2. Personalized Treatment Plans – AI analyses patient data to recommend personalized treatment options.

In Agriculture:

1. Precision Farming – AI tools monitor soil health, weather, and crop conditions to optimize farming.

2. Yield Prediction – AI models forecast crop output, helping farmers plan harvest and sales.

(ii) **[3]**

```
def check_even_odd(n):
        if n % 2 == 0:
print("Even") else:
      print("Odd")
check_even_odd(7)
```

(iii) Machine Learning (ML) is a subset of AI where algorithms learn from data and make predictions or decisions. Example: Email spam filtering. Deep Learning (DL) is a subset of ML that uses neural networks with many layers to model complex patterns. **[3]**

Example: Facial recognition in images.

Question 4

(i) Matplotlib: Low-level, flexible, great for custom plots but requires more code **[4]**

Seaborn: Built on top of Matplotlib, provides high-level, attractive, and informative statistical graphics

Ease of use: Seaborn has simpler syntax for complex visualizations

Recommendation: Seaborn is better for statistical plots like boxplots and heatmaps due to built-in features and aesthetics

(ii) ○ Helps in understanding large and complex datasets **[3]**
○ Identifies trends, outliers, and relationships
○ Aids in communicating results clearly to stakeholders

(iii) **[3]**

```
import matplotlib.pyplot as plt
domains = ['AI', 'Robotics', 'IoT']
students = [30, 20, 15]
plt.bar(domains, students)
plt.title('Student Interests')
plt.ylabel('Number of Students')
plt.show()
```

Question 5

(i) (a) Type of Join: Inner Join on the column ID. **[1]**

(b) Resulting Table (after join): **[1]**

ID	Name	City	Product	Price
1	Aman	Delhi	Keyboard	1500
2	Beena	Mumbai	Laptop	55000
3	Chitra	Chennai	Mouse	600

(c) This operation links two sets (Customer Info and Purchases) by a common attribute (ID), a classic application of Set Theory and Relational Algebra. Joins are used in AI to combine datasets for training models or generating features from multiple tables. **[2]**

(ii) (a) **[1]**

$$A + B = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

(b) **[1]**

$$A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

(c) **[1]**

$$AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

(iii) Dataset: 55, 60, 65, 70, 95

(a) Mean = (55 + 60 + 65 + 70 + 95) / 5 = 69. **[1]**

(b) Median = 65 (middle value of ordered list) **[1]**

(c) Mode = Not defined / No mode. All values are unique, so no value repeats. **[1]**

Question 6

(i) import pandas as pd

(a) df = pd.read_csv('dataset.csv') **[1]**

(b) df = df.drop_duplicates() **[1]**

(c) df = df.fillna(df.mean(numeric_only=True)) **[1]**

(d) df['City'] = df['City'].str.title() **[1]**

(ii) Removing Duplicates – Eliminate repeated rows using drop_duplicates() **[3]**

Handling Missing Values – Fill or drop rows with NaN using fillna() or dropna()

Correcting Inconsistent Data – Standardize formats or fix spelling using mapping or string functions

(iii) The Least Squares Regression Line minimizes the sum of the squared differences between observed and predicted values. It is determined by calculating the slope and intercept that minimize the total squared residuals (errors) using statistical formulas. **[3]**

Question 7

(i) The linear regression equation is of the form Y=a+bX, where: **[4]**

- b is the slope of the line.
- a is the y-intercept.

The formulas to calculate b and a are:

slope (b)

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b = \frac{3(680) - (10)(190)}{3(38) - (10)^2} = \frac{2040 - 1900}{114 - 100} = \frac{140}{14} = 10$$

intercept (a)

$$a = \frac{\sum Y - b\sum X}{n}$$

$$a = \frac{190 - 10 \cdot 10}{3} = \frac{190 - 100}{3} = \frac{90}{3} = 30$$

The linear regression equation.

Substitute the values of a=30 and b=10 into the equation Y=a+bX:

Y=30+10X

The linear regression equation of Y on X is **Y=30+10X**.

(ii) Relational Model – Represents data in tables (relations) with rows and columns. Dimensional Model – Optimized for data warehousing, involving fact and dimension tables.
Entity-Relational Model – Uses entities, attributes, and relationships to design a conceptual schema of data. **[3]**

(iii) If a facial recognition system is trained primarily on lighter-skinned faces, it may misidentify or fail to recognize individuals with darker skin tones. This can lead to discrimination, such as wrongful accusations or denial of services, thereby causing social harm and undermining public trust in AI. **[3]**

Question 8

(i)
```
import pandas as pd
from scipy.stats import zscore
df = pd.read_csv("housing.csv")
```
(a) `df = df.drop_duplicates()` **[1]**

(b) `df['Price'] = df['Price'].fillna(df['Price'].mean())` **[1]**

(c) `df['Area'] = zscore(df['Area'])` **[2]**

(ii) ○ Narrow AI: Google Maps (specific task) **[3]**
○ General AI: A robot with human-like reasoning (still theoretical)
○ Superintelligent AI: Exceeds human intelligence (hypothetical future)

(iii) **[3]**
```
import matplotlib.pyplot as plt
names = ['A', 'B', 'C', 'D']
projects = [3, 5, 2, 4]
plt.bar(names, projects)
plt.xlabel('Students')
plt.ylabel('AI Projects')
plt.title('Projects Completed')
plt.show()
```

Question 9

(i) (a) Relational data modelling is suitable, as the data can be stored in a structured format (rows and columns) with a clear relationship between attributes (study hours ↔ scores) **[1]**

(b) Independent variable: Study hours (input), Dependent variable: Scores (output). **[1]**

(c) Bias can lead to unfair or inaccurate predictions. For example, if the dataset only includes high-performing students, the model will overestimate scores for students who study less, making it unreliable and ethically problematic. **[2]**

(ii) Data Acquisition: Gathering relevant data from sources (sensors, web, CSV files). **[3]**
Data Exploration: Understanding patterns, distributions, and anomalies in data.
These steps ensure data quality and guide the modelling direction.

(iii) AI systems should be designed to optimize energy efficiency, use sustainable **[3]**
hardware, and minimize carbon footprint.
Example: Training smaller ML models instead of massive ones to reduce energy consumption in data centers.