PRACTICE PAPER–3

Time: 2 Hours **Maximum Marks: 50**

General Instructions:

1. Please read the instructions carefully.
2. This Question Paper consists of **21 questions** in two sections: **Section A & Section B**.
3. Section A has Objective type questions whereas Section B contains Subjective type questions.
4. **Out of the given (5 + 16 =) 21 questions, a candidate has to answer (5 + 10 =) 15 questions in the allotted (maximum) time of 2 hours.**
5. All questions of a particular section must be attempted in the correct order.
6. **SECTION A—OBJECTIVE TYPE QUESTIONS (24 MARKS):**
   (a) This section has 5 questions.
   (b) Marks allotted are mentioned against each question/part.
   (c) There is no negative marking.
   (d) Do as per the instructions given.
7. **SECTION B—SUBJECTIVE TYPE QUESTIONS (26 MARKS):**
   (a) This section has 16 questions.
   (b) A candidate has to do 10 questions.
   (c) Do as per the instructions given.
   (d) Marks allotted are mentioned against each question/part.

SECTION A: OBJECTIVE TYPE QUESTIONS

1. Answer any 4 out of the given 6 questions on Employability Skills. **(1 × 4 = 4 marks)**

(a) Which type of verbal communication involves sharing information using spoken words over the phone?
   (i) Written Communication (ii) Oral Communication
   (iii) Visual Communication (iv) Non-verbal Communication

(b) While solving a difficult problem, you try different ways to reach the answer and do not give up even if you fail at first. Which self-management skill are you demonstrating?
   (i) Adaptability (ii) Goal Setting
   (iii) Perseverance (iv) Physical Fitness

(c) Which of the following is not a recommended method to improve emotional intelligence?
   (i) Ignoring your emotions when making decisions
   (ii) Practising yoga and meditation to stay calm
   (iii) Observing and reflecting on your own behaviour
   (iv) Thinking rationally before reacting to situations

(d) Which of these is not a typical way to select multiple files at once in File Explorer?
   (i) Holding down the Ctrl key and clicking each file
   (ii) Dragging the mouse over multiple files
   (iii) Pressing Shift and selecting the first and last file
   (iv) Double-clicking on any file

(e) After completing his education, Aman took a loan to open his first bakery. He is now working hard to attract customers and make his business stable. Which stage of the entrepreneurial process is Aman in?
   (i) Grow (ii) Enter
   (iii) Survive (iv) Retire

(f) Planting trees in your neighbourhood to improve air quality supports which Sustainable Development Goal?
   (i) SDG 7 – Affordable and Clean Energy (ii) SDG 13 – Climate Action
   (iii) SDG 3 – Good Health and Well-being (iv) SDG 5 – Gender Equality

2. Answer any 5 out of the given 6 questions. **(1 × 5 = 5 marks)**

(a) An AI system that analyzes numerical data and finds patterns in it works in the ..... domain

(b) **Assertion (A):** Value-based frameworks in AI ethics provide guidance by focusing on fundamental ethical principles and values.

Reason (R): These frameworks reflect different moral philosophies guiding ethical reasoning and help assess the moral worth of actions.

(i) Both A and R are correct and R is the correct explanation of A.
(ii) Both A and R are correct but R is not the correct explanation of A.
(iii) A is correct but R is incorrect.
(iv) A is incorrect but R is correct.

(c) A pizza delivery chatbot can take orders only when customers type exactly the phrases it knows (*for example*, 'Order a large cheese burst pizza').

If a customer types the same request in a different way (like 'Can I get a big bursting cheese pizza?'), the chatbot fails to understand and respond correctly.

What is the most likely reason for this issue?

(i) The chatbot is too slow.
(ii) The chatbot uses fixed rules that cannot handle variations in language.
(iii) The customer is asking incorrect questions.
(iv) The chatbot was trained on images instead of text.

(d) To check the performance of a machine learning algorithm, data is divided into ..... and testing sets.

(e) ..... involves techniques like resizing, noise reduction and adjusting brightness to make images ready for further analysis; it is a subset of ..... .

(i) Computer Vision; Image Processing
(ii) Machine Learning; Artificial Intelligence
(iii) Image Processing; Computer Vision
(iv) Object Detection; Image Processing

(f) State True or False:

All sentences conveying a similar meaning will always be processed identically by an NLP model, regardless of word order or phrasing differences.

3. Answer any 5 out of the given 6 questions. **(1 × 5 = 5 marks)**

(a) A hospital wants to use an AI program to scan patient X-rays and automatically highlight areas that might show signs of pneumonia, helping radiologists prioritize their work. Which AI domain is essential for this tool?

(i) Natural Language Processing
(ii) Recommender Systems
(iii) Computer Vision
(iv) Statistical Modelling

(b)

Fruit	Color	Price
Apple	Red	INR 20
Orange	Orange	INR 10
Banana	Yellow	INR 5
Grape	Purple	INR 20
Kiwi	Green	INR 20

In the given fruit dataset, what do we call the entire entry for one fruit (*e.g.*, 'Orange, Orange, ₹ 10')?

(i) Feature
(ii) Model
(iii) Instance
(iv) Label

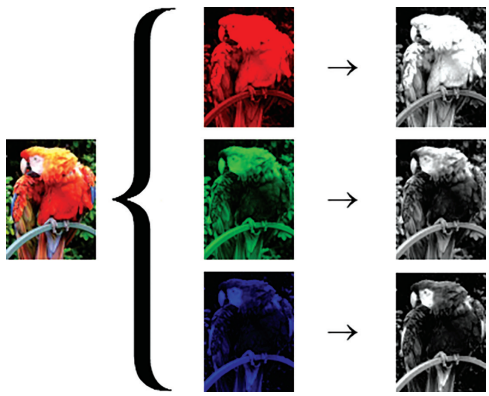
(c) An AI model is designed to detect a rare disease in patients. It was tested on 500 patients. Among them, 100 patients actually had the disease. The model correctly identified 85 of the patients who had the disease but missed 15 cases. What is the recall of the model?

(i) 0.90 (ii) 0.85
(iii) 0.95 (iv) 0.80

(d) **Assertion (A):** An AI model that achieves 99% accuracy on its training data but only 60% accuracy on its test data is considered 'overfitting'.
Reason (R): Overfitting occurs when a model learns the noise and details of the training data too well, leading to poor performance when it encounters new, unseen data.

(i) Both A and R are correct and R is the correct explanation of A.
(ii) Both A and R are correct but R is not the correct explanation of A.
(iii) A is correct but R is incorrect.
(iv) A is incorrect but R is correct.

(e)



What do we call this type of image created in the last phase of the above transformation pipeline?

(i) Grayscale images (ii) RGB images
(iii) 3-color images (iv) None of these

(f) Count and write the number of **stopwords** from this list that appear in the following sentence: 'Computer Vision is a domain of AI that trains machines to interpret and understand the visual world.'

4. Answer any 5 out of the given 6 questions. (1 × 5 = 5 marks)

(a) What is the primary function of the AI Project Cycle in developing a model?

(i) It provides a structured sequence of stages to guide the process.
(ii) It automatically collects and cleans all the necessary data.
(iii) It writes the programming code for the model.
(iv) It guarantees the model will be 100% accurate.

(b) **Statement 1:** Precision is the preferred metric to use when the primary goal is to minimize False Negatives (*e.g.*, in medical screening).
Statement 2: Recall is the preferred metric to use when the primary goal is to minimize False Positives (*e.g.*, in a spam filter).

(i) Both Statement 1 and Statement 2 are correct.
(ii) Both Statement 1 and Statement 2 are incorrect.
(iii) Only Statement 1 is correct.
(iv) Only Statement 2 is correct.

(c) A company uses an AI model to select candidates for a marketing job. Four attributes are considered, each with a specific weight:

- Has a marketing certification ($x_1 = 1$, $w_1 = 0.5$)
- Prior sales experience ($x_2 = 1$, $w_2 = 0.4$)
- Knowledge of analytics tools ($x_3 = 0$, $w_3 = 0.3$)
- Excellent presentation skills ($x_4 = 1$, $w_4 = 0.6$)

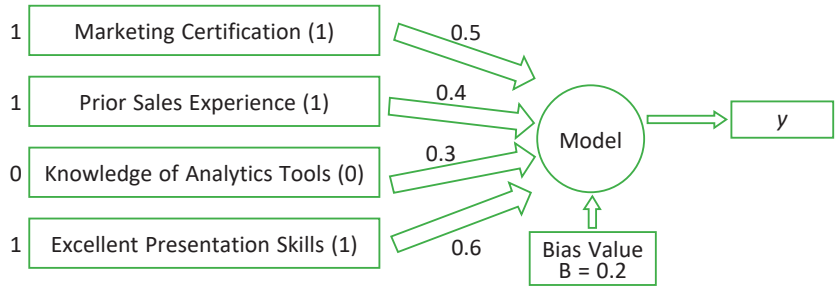
Bias value, b = 0.2

Formula: $y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + (1 \times b)$

What will be the value of *y* for this scenario?

(i) 1.70 (ii) 1.37

(iii) 1.18 (iv) 1.45



(d) What is the primary purpose of splitting your dataset into a 'training set' and a 'testing set'?

(i) To have a smaller dataset to train on, which makes the model build faster.

(ii) To have an unseen set of data to evaluate the model's ability to generalize to new examples.

(iii) To mix the data together so the model can see all examples twice.

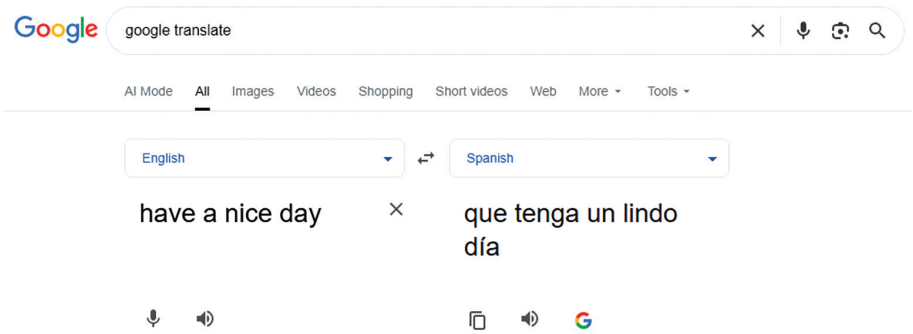
(iv) To prove that the model has a high F1 Score.

(e) In a Convolutional Neural Network, a small matrix of numbers (like a 3×3 grid) is used to scan over an image to detect specific features like edges or corners. What is this small matrix called?

(i) Kernel (ii) Pixel

(iii) Resolution (iv) Image Layer

(f) Identify the application of NLP shown in the given picture:



(i) Autogenerated captions (ii) Language Translation

(iii) Text Classification (iv) Keyword Extraction

5. Answer any 5 out of the given 6 questions. **(1 × 5 = 5 marks)**

(a) An AI model is built to screen resumes for a software developer job. It is trained on the company's past 10 years of hiring data, where 85% of successful hires were male. Which factor might *unfairly* cause the model to rank a highly qualified female applicant lower?

(i) Her years of programming experience

(ii) The model's learned statistical bias from the training data

(iii) The university she graduated from

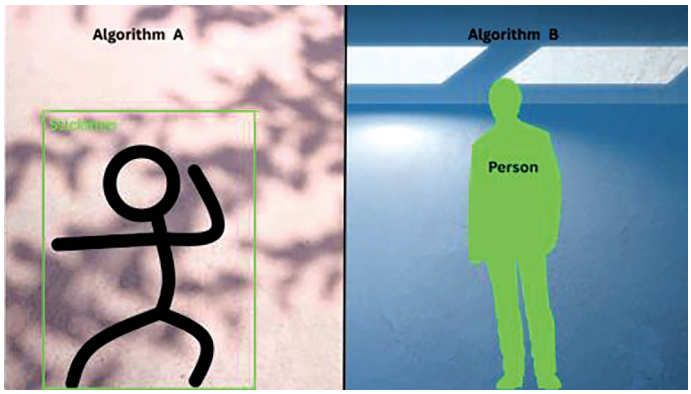
(iv) Her list of technical skills

(b) An AI model predicted a student would score 85 on a test, but the student's actual score was 90. What is the absolute error rate of this prediction (relative to the actual score), rounded to three decimal places?

(i) 0.059 (ii) 0.050

(iii) 0.944 (iv) 0.056

(c) Which computer vision task is shown in Algorithm A?



(i) Image Classification
(ii) Object Detection
(iii) Image Segmentation
(iv) Feature Extraction

(d) ..... is the outcome of the model correctly predicting the negative class as negative.

(e) Which NLP technique involves counting the frequency of each word in a document to represent it, but disregards grammatical structure and word order?

(i) TF-IDF
(ii) Bag-of-Words
(iii) Semantic Analysis
(iv) Text Normalization

(f) What are the stemmed and lemmatized forms of the word 'studies'?

(i) Stemmed: study, Lemmatized: study
(ii) Stemmed: studi, Lemmatized: studi
(iii) Stemmed: studi, Lemmatized: study
(iv) Stemmed: study, Lemmatized: studi

SECTION B – SUBJECTIVE TYPE QUESTIONS

Answer any 3 out of the given 5 questions on Employability Skills in 20–30 words each. (2 × 3 = 6 marks)

6. What are phrases? Give an example.

7. Explain the importance of identifying sources of stress in effective stress management.

8. 'Phishing is a fraudulent attempt to obtain sensitive information, such as usernames or passwords, by disguising as a trustworthy entity.' Explain two common signs indicating that an email might be a phishing attempt.

9. Karan starts a mobile app to help local farmers sell their produce directly to customers. He invests his own savings to build the app and hires two people. He is aware that a similar app failed last year, but he believes his marketing plan is better.
From the paragraph, identify and write the two entrepreneurial functions or characteristics being described.

10. Describe any two characteristics of green consumers.

Answer any 4 out of the given 6 questions in 20–30 words each. (2 × 4 = 8 marks)

11. What is the key purpose of the modelling and evaluation stage of the AI Project Cycle?

12. Identify the type of deep learning model for each description:

(a) This model is primarily used for tasks involving sequential data, like language translation or predicting stock market prices. Its key feature is an internal 'memory' or loop that allows it to retain information from previous steps.

(b) This model is the 'eyes' of AI, primarily used for image-related tasks. It uses special layers with 'filters' (or kernels) to automatically detect features like edges, corners and textures.

13. How does a **Rule-Based AI Model** differ from a **Learning-Based AI Model**?

14. What is 'overfitting' in machine learning and why is it important to use a separate 'test set' to detect it?

15. Define the term 'image features' in the context of Computer Vision. Give one example of a simple feature that a computer might be trained to detect.

16. Identify the stage of NLP and explain:
Input: "The quick brown fox jumps."
Output: "['The', 'quick', 'brown', 'fox', 'jumps', '.']"

Answer any 3 out of the given 5 questions in 50–80 words each. **(4 × 3 = 12 marks)**

17. Read the case study below and answer the following questions:
A new AI system is developed to detect skin cancer from images of moles. The AI was trained on a massive dataset of 100,000 images, all collected from dermatology clinics in Australia and Europe. The model shows 99% accuracy in testing. However, when the model is deployed in hospitals worldwide, it performs very poorly on patients with darker skin, frequently failing to identify cancerous moles that human doctors can spot.
(a) Identify the specific source of the bias in this AI model.
(b) Mention two **bioethics principles** that this model violates and explain how they apply.

18. (a) What is the name of the learning model inspired by the human brain, which uses multiple layers of interconnected nodes to learn complex patterns?
(b) Name the three fundamental types of layers that make up a basic version of this model.
(c) Explain the function of each of these three layers.

19. Identify the name of the application of Machine Learning (ML) or Deep Learning (DL) being used in the following scenarios:
(a) You use a banking app that monitors your spending. It learns your typical purchase patterns (*e.g.*, groceries, gas, rent). One day, it detects a large, unusual transaction from a different country and sends you an alert, flagging it as a potential risk.
(b) A company analyzes thousands of customer reviews for their new product. The AI reads each comment and automatically assigns it a score, such as 'Positive,' 'Negative,' or 'Neutral,' to understand public opinion.
(c) An online streaming service analyzes the shows you watch and compares your viewing history to millions of other users. It then suggests a new show for you, based on what 'users with similar taste' also enjoyed.
(d) You are dictating an email to your phone. AI listens to your voice and converts your spoken words into text that appears on the screen.

20. Read the following paragraph and answer the questions that follow:
An AI model for a bank is tested on 1,000 transactions to detect fraud. 'Fraud' is the positive class. The model correctly identified 40 transactions as fraud. It also correctly identified 940 transactions as legitimate (not fraud). However, the model incorrectly flagged 10 legitimate transactions as fraud. It also missed 10 fraudulent transactions, labelling them as legitimate.
(a) Draw the confusion matrix based on the above information.
(b) Calculate the **Precision** of this fraud detection model. Show your working.
(c) Write the total number of False Positives, *i.e.*, legitimate transactions flagged as fraud.

21. Read the following documents and answer the questions that follow:
- **Document 1:** 'The cat chased the dog.'
- **Document 2:** 'The dog ran.'
- **Document 3:** 'The cat and the dog are friends.'

After performing basic text pre-processing (removing punctuation, converting to lowercase and tokenizing), the documents become:
- **Document 1:** [the, cat, chased, the, dog]
- **Document 2:** [the, dog, ran]
- **Document 3:** [the, cat, and, the, dog, are, friends]

Questions:
(a) Create the dictionary (vocabulary) of unique words from all three documents.
(b) Construct the document vector for **Document 1** using the dictionary.
(c) Explain the main limitation of this model regarding the meaning of words (semantics)?
(d) Why might 'stopwords' (like 'the', 'and', 'are') be removed before creating the dictionary?