PRACTICE PAPER–5

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 barrier affects effective communication when people use technical jargon unfamiliar to the listener?
   (i) Emotional Barrier (ii) Language Barrier
   (iii) Physical Barrier (iv) Cultural Barrier

(b) When a classmate suggests a new way to study and you are willing to try it instead of insisting on your usual method, which self-management skill are you displaying?
   (i) Adaptability (ii) Time Management
   (iii) Confidence (iv) Stress Management

(c) Which of the following is not recommended while setting personal goals?
   (i) Setting unrealistic and vague goals
   (ii) Breaking larger goals into smaller, achievable tasks
   (iii) Monitoring progress regularly
   (iv) Being specific about your objectives

(d) What happens when you press Ctrl+C after selecting a file in File Explorer?
   (i) The file is deleted permanently.
   (ii) The file is copied to the clipboard.
   (iii) The file is moved to another folder.
   (iv) The file is renamed automatically.

(e) Rakesh owned a chain of mobile stores for decades. After handing over his business to his children, he no longer manages daily operations and spends most of his time on personal activities. Which stage of the process is Rakesh in?
   (i) Retire (ii) Enter
   (iii) Survive (iv) Grow

(f) Promoting equal pay and equal opportunities for men and women at the workplace is part of which Sustainable Development Goal?
   (i) SDG 3 – Good Health and Well-being
   (ii) SDG 5 – Gender Equality
   (iii) SDG 8 – Decent Work and Economic Growth
   (iv) SDG 13 – Climate Action

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

(a) Models which detect patterns and trends within numerical data in spreadsheets or databases use the ..... domain of AI.

(b) **Assertion (A):** Virtue-based framework focuses on the character and intentions of individuals involved in decision-making.

Reason (R): Virtue-based framework evaluates whether actions align with principles such as honesty, compassion and integrity.

(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 machine learning model is designed to recognize handwritten digits (0-9). The model was trained exclusively on a dataset of neat, standard handwriting. When users submit digits written in a highly stylized or messy cursive script, the model frequently gives the wrong answer. What is the primary technical limitation demonstrated by this model's failure?

(i) The model was given too much data during its initial training.
(ii) The model cannot understand numbers written after noon.
(iii) The model lacks the ability to generalize and accurately interpret data outside its training set distribution.
(iv) The programmer forgot to include the number '8' in the training set.

(d) ..... is the process of assessing how well an AI model performs on new, unseen data.

(e) Tasks like correcting image tone, cropping and sharpening are performed under ....., which is classified as a subset of ..... .

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

(f) State True or False:

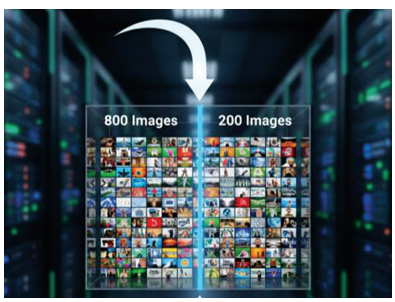
'Text Normalization' involves processes like converting all text to lowercase, removing punctuation and stemming words to their root form.

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

(a) A city wants to optimize its bus routes by predicting rider demand. Planners are analyzing historical data, including time of the day, day of the week, weather conditions (temperature, rainfall) and census data from different neighbourhoods, to forecast how many people are likely to be at a specific bus stop. Which AI domain is primarily used to analyze these variables and make such a prediction?

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

(b) This diagram shows a dataset being split to build and evaluate an AI model. What are the '800 Images' and '200 Images' sets called?



(i) Features and Labels
(ii) Training Data and Testing Data
(iii) Supervised and Unsupervised Data
(iv) Accurate and Inaccurate Data

(c) An AI model is screening 500 patients for a specific disease. In this group, 100 patients actually have the disease (and 400 are healthy). The model correctly identifies 80 of the sick patients, but it also incorrectly flags 40 healthy patients as being sick.
What is the **recall** of the model?
(i) 0.67
(ii) 0.80
(iii) 0.88
(iv) 0.84

(d) **Assertion (A):** In a spam email filter, achieving high **precision** is the primary goal.
Reason (R): High precision minimizes the number of legitimate emails (False Positives) that are incorrectly marked as spam, ensuring important emails are not lost.
(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) Identify the application of Computer Vision from the given picture:

(i) Facial Recognition
(ii) CV in Retail
(iii) Search by Image
(iv) Face Filters

(f) Count and write the number of tokens in the sentence: 'Text Normalization is a crucial step in Natural Language Processing that involves converting text to lowercase and removing punctuation.'

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

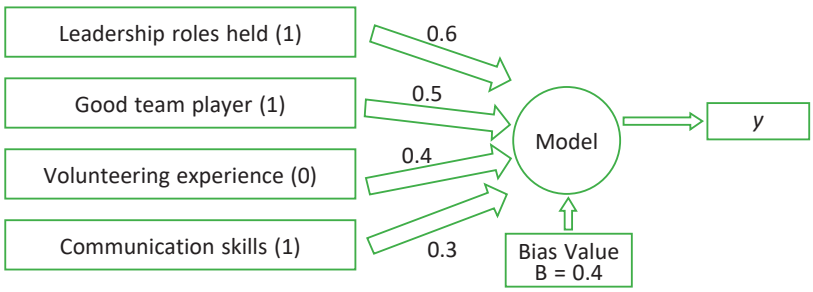
(a) In the 'Problem Scoping' stage, what is the main purpose of using the '4Ws Problem Canvas'?
(i) To build the final AI model
(ii) To collect and acquire all the data
(iii) To understand the problem deeply and define its context
(iv) To test the model's accuracy

(b) **Statement 1:** A 'Confusion Matrix' is a table used to evaluate the performance of a regression model.
Statement 2: 'Accuracy' is calculated by dividing the total number of correct predictions (both true positives and true negatives) by the total number of all predictions.
(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) To select students for a leadership camp, four criteria with assigned weights are assessed:
- Leadership roles held ($x_1 = 1$, $w_1 = 0.6$)
- Good team player ($x_2 = 1$, $w_2 = 0.5$)
- Volunteering experience ($x_3 = 0$, $w_3 = 0.4$)
- Communication skills ($x_4 = 1$, $w_4 = 0.3$)

Bias value, $b = 0.4$
$y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + (1 \times b)$

What is *y* for this applicant?

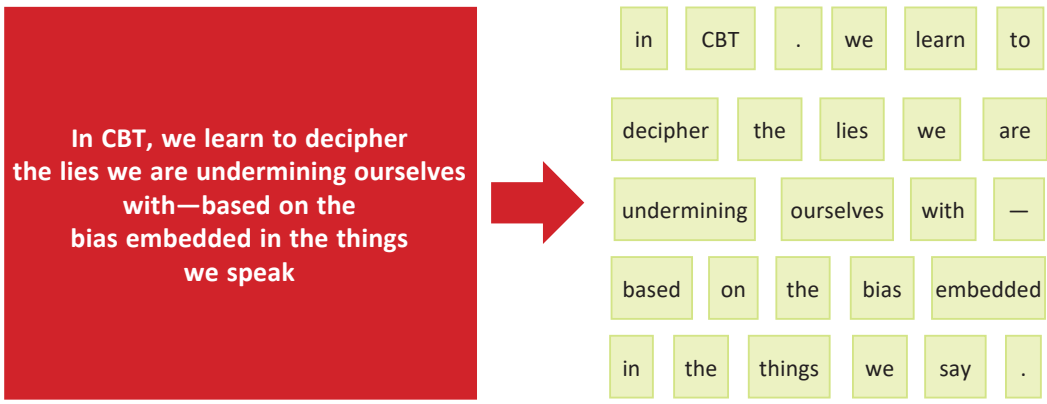


(i) 1.8 (ii) 1.4
(iii) 1.7 (iv) 2.0

(d) Why is it important to evaluate a model using metrics like Precision, Recall and F1 Score, instead of just using Accuracy?
(i) Because Accuracy is too complicated to calculate
(ii) Because these metrics provide a better understanding of *how* the model is wrong (*e.g.*, False Positives vs False Negatives), which is critical in many real-world problems.
(iii) Because Precision and Recall always give a higher score than Accuracy
(iv) Because these metrics are only used for regression, not classification

(e) The measurement of an image's dimensions in terms of its total width and height (such as 1920×1080) is known as its:
(i) Color Depth (ii) Pixel Value
(iii) Resolution (iv) Kernel Size

(f) Which text normalization technique is represented?



(i) Sentence Segmentation (ii) Tokenization
(iii) Stemming (iv) Lemmatization

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

(a) A computer vision model is trained to identify 'a person in a professional setting.' The training dataset was gathered online and 90% of the images were of people wearing Western-style suits. What is a likely *biased* outcome?
(i) The model will be very good at identifying all professionals.
(ii) The model may fail to recognize a person in a traditional Indian kurta as a 'professional.'
(iii) The model will be able to read the person's name tag.
(iv) The model will run faster because the data was not diverse.

(b) An AI model predicted that a website would receive 2,200 visitors on a particular day. The actual number of visitors was 2,500. What is the error percentage of this prediction, relative to the actual value?
(i) 13.6% (ii) 88.0%
(iii) 10.0% (iv) 12.0%

(c) Which computer vision task is shown in the given image?

(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 positive.

(e) Which NLP model calculates a numerical score to measure how important a word is to a specific document, relative to a larger collection of documents (a corpus)?

(i) TF-IDF
(ii) Script Bot
(iii) Bag-of-Words
(iv) Smart Bot

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

(i) Stemmed: bet, Lemmatized: good
(ii) Stemmed: better, Lemmatized: good
(iii) Stemmed: bet, Lemmatized: better
(iv) Stemmed: better, Lemmatized: better

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 compound sentences? Give an example.

7. List and describe the key traits that help an individual become independent, especially in the context of developing self-management skills.

8. "A 'weak password' is a password that is easy for a person or computer program to guess, such as '123456' or 'password'." List two rules you should follow to create a 'strong password' that is difficult to guess.

9. Sameer runs a small courier service in his city. He realizes he is losing business to larger companies. He decides to buy three electric scooters to offer faster, eco-friendly deliveries for small packages. He uses his business earnings to pay for the scooters, hoping this new service will attract more clients.
Identify and write the two characteristics of an entrepreneur that Sameer is displaying.

10. Discuss the concept of sustainable development and its significance in addressing global environmental challenges?

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

11. What is the primary relationship between the **Modelling** stage and the **Evaluation** stage in the AI project lifecycle?

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

(a) This model is the technology behind facial recognition on your phone and object detection in self-driving cars. It is specifically designed to analyze 2D pixel data.

(b) This model is the most basic 'deep learning' model, forming the foundation for others. It consists of an input layer, one or more hidden layers and an output layer, but has no specialized 'memory' or 'convolutional' layers.

13. How is **Clustering** (Unsupervised Learning) different from **Classification** (Supervised Learning)?

14. What is the purpose of a **Train-Test split** and how does this process help a developer evaluate a model's ability to generalize?

15. What is a kernel (or filter) in the context of a Convolutional Neural Network (CNN)? What is its primary purpose?

16. Identify the stage of NLP and explain:
Input: "My dog's name is Max. He is very playful."
Output: "The word 'He' is understood to refer to 'Max' from the previous sentence."

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 questions that follow:
A hospital builds an AI system to predict which patients with pneumonia are at the highest risk of death, to prioritize them for ICU admission. The system learns from historical data. It discovers that patients with asthma who get pneumonia almost always recover. Based on this, the AI system learns that 'asthma = low risk.' However, the *reason* they recover is that the hospital *always* sends asthmatic patients to the ICU immediately on admission as a policy.
(a) Explain the flawed 'logic' the AI learned from the data.
(b) Mention two bioethics principles that this AI model would violate if it were used to make decisions. Explain how.

18. (a) What is the name of the learning model that trains on a labelled dataset, where each data point is tagged with a correct output or answer?
(b) Name the two main categories of problems this learning model solves.
(c) Explain each category briefly and provide one example of each.

19. Identify the name of the application of Machine Learning (ML) or Deep Learning (DL) being used in the following scenarios:
(a) A real estate website uses AI to predict the price of a house. It bases its prediction on features like the house's square footage, the number of bedrooms and the neighbourhood's location.
(b) A news app uses AI to read a long, 10-paragraph news article and automatically generates a short, two-sentence paragraph that captures the main points.
(c) You use a website to translate a sentence from English to French. The AI model analyzes the meaning of the English sentence and constructs a new, grammatically correct sentence in French.
(d) A website has a pop-up window where you can ask questions. It understands your intent and provides relevant answers from its knowledge base.

20. Read the following paragraph and answer the questions that follow:
An AI system was used to predict which of the 400 students would receive a scholarship. 'Will Get Scholarship' is the positive class. The model correctly predicted 70 students who received the scholarship. It also correctly predicted 280 students who did not receive it. However, the model predicted 30 students would get the scholarship, but they did not. It also predicted 20 students would not get the scholarship, but they actually did.
(a) Draw the confusion matrix for this scholarship model.
(b) Calculate the **Recall** of this model. Show your working.
(c) Write the total number of wrong predictions made by the model.

21. Read the following documents and answer the questions that follow:
- **Document 1:** 'The model is good.'
- **Document 2:** 'The model is not good.'

After performing basic text pre-processing (removing punctuation, converting to lowercase and tokenizing), the documents become:
- **Document 1:** [the, model, is, good]
- **Document 2:** [the, model, is, not, good]

Questions:
(a) Create the dictionary (vocabulary) of unique words from both documents.
(b) Construct the document vector for **Document 1** and **Document 2** using the dictionary.
(c) Look at the two vectors you created. What problem does this example highlight about the Bag-of-Words model?
(d) How does this model lose the 'syntactic structure' (grammar) of the original sentences?