Assignment -1

Sec-A

Q1. Define Artificial Intelligence (AI) and list its key application domains. Provide examples.

Sol:- Artificial Intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think and learn like human. These system can perform tasks that typically require human intelligence such as visual perception, speech recognition, decision making, and language translation.

Key Application Domains:

1. Healthcare: AI is used for diagnosing diseases, personalised treatement plans, robotic surgeries and many more. for eg:- IBM, watson Health.

2. Finance: AI helps in fraud detection, algorithmic trading, and personalised banking. Eg. AI driven chat-bots in customer service.

3. Transportation:- Self driving cars and traffic management systems. Eg:- Tesla, Waymo.

4. Retail:- AI is used for inventory management, personalized shopping experience. Eg. Amazon's recommendation system.

Page No.
Date : / /

5. Manufacturing: Predictive maintainance, quality control and automation of product lines. for eg Siemens uses AI for smart manufacturing

Q2. Explain the different levels of models in AI. How do they influence AI techniques?

Ans. AI models can be categorized into three levels:-

1. Narrow AI (weak AI):- Design to perform a narrow task. It does not posses general intelligence. for eg:- facial recognition or internet searches

2. General AI (Strong AI): Has the ability to understand, learn, and apply knowledge across a wide range of tasks, similar to human intelligence. This level is still theoretical and not yet achieved.

3. Superintelligent AI: Surpasses human intelligence in all aspects, including creativity, problem-solving and emotional intelligence. This is a hypothetical concepts and remains a topic of speculation and debate.

These level influence AI techniques by determing the complexity and scope of the tasks that AI systems can handle. Narrow AI uses specific algorithms tailored to particular tasks, while General AI would require more advanced, adaptive algorithms capable of solving problems across various domains.

Q3 What is an intelligent Agent? Discuss the characterstics of learning agents with egs.

Answers: An intelligent agent is an autonomous entity that observes and act upon an environment to achieve specific goals. It can percieves its Environment through sensors and act upon it using ~~sensors~~ actuators.

Characterstics of Learning Agents:-

→ Performance Element: Executes actions based on the agents knowledge.

→ Learning Element: Improves the agent's performance by learning from experiences

→ Critic: Provides feedback on the agent's action to guide learning.

→ Problem Generator:
Suggests actions that lead to new experiences and learning opportunities.

Eg:- A self-driving car is a learning agent. It uses sensors to percieve the environment, makes decision based on it's knowledge, learns from driving experiences, and improves its performance over ~~over~~ over time.

Page No.
Date : / /

Q4. Highlight the advantages and limitations of AI. Provide examples of its impact in real-world scenarios.

Answer Advantages:

→ AI can process large amounts of data quickly and accurately.
→ Automates repetitive and mundane tasks, freeing up human resources for more complex activities.
→ Provides data-driven insights and predictions, enhancing decision-making processes.
→ Offers personalized experiences in various domains, such as healthcare and retail.

Limitations:
→ AI lacks the ability to think creatively and innovatively like humans.
→ Issues related to privacy, bias, and job displacement
→ AI requires vast amounts of data to function efficiently.
→ Developing and maintaining AI systems can be expensive and complex.

Impact in real world scenarios :-

→ Healthcare :-

AI powered diagnostic tools can detect diseases like cancer at an early stage, improving patient outcomes.

→ Finance :-

AI algorithms can detect frauds in real-time, reducing financial losses.

Page No
Date : / /

Q5. Describe any two AI techniques and their role in solving real-world problems.

Answer: Machine Learning (ML):

→ Role: Enables systems to learn from data and improve their performance over time without being explicitly programmed.
→ Example: Predictive analytics in healthcare to forecast disease outbreaks.

Natural Language Processing (NLP):

→ Role: ~~Enables systems to learn from data and improve their performance over time~~

→ Role: Allows machines to understand, interpret and generate human language.
→ Example: Chatbots like Chat-Gpt

Sec B

Q1. What is propositional logic? Illustrate with an example.

Answer: Propositional Logic is a branch of logic that deals with propositions, which can either be either be true or false. It uses logical connectives like AND, OR, NOT, and IMPLIES to form complex expressions

Eg.- let p represents "It is raining".
let q represents "I will carry an umbrella".
Then p→q : If it is raining then I will carry an umbrella

Page No.
Date : / /

Q2. Define predicate logic and differentiate it from propositional logic with examples.

Answer: Predicate logic also known as First-Order Logic, extends propositional logic by dealing with predicates and quantifiers. It allows for more expressive statements about objects and their properties.

eg.
$p \rightarrow q$
$\forall x (P(x) \rightarrow Q(x))$
In predicate logic, P(x) could represent "x is a bird" and Q(x) could represent "x can fly". This statement $\forall x (P(x) \rightarrow Q(x))$ means "For all x, if x is a bird, then x can fly".

Q3. Convert the following propositional formula into Conjuctive Normal Form (CNF):

Answer: To convert propositional formula into CNF:
→ Eliminate bi-conditional ($\leftrightarrow$) and implication ($\rightarrow$) operators.
→ Move NOT operators inward using De Morgan's laws.
→ Distribute AND over OR to achieve the CNF.

Q4. Explain the unification algorithm and its significance in predicate logic. Provide an example.

Answer: The unification Algorithm is used in predicate logic to make two logical expressions identical by finding a substitution for their variables.

Eg. Expression $P(f(x), y)$ and $P(z, g(a))$-
unification $[z \rightarrow f(x), y \rightarrow g(a)]$

Page No
Date : / /

Sec-C

Q1. Define machine learning and explain the significance of datasets in machine learning. How are datasets handled effectively?

Answer: Machine Learning (ML) is a subset of artificial intelligence that involves the use of algorithms and statistical models to enable computers to learn from and make predictions or decisions based on data.

Significance of Datasets:

Datasets are crucial in machine learning as they provide the examples from which the model learns. The quality and quantity of the data directly impact the performance of the machine learning model.

→ **Handling Datasets Effectively:**

→ **Data cleaning:** Removing or correcting errors and inconsistencies in the data.

→ **Data transformation:** Converting data into a suitable format for analysis.

→ **Data Augmentation:** Increasing the diversity of the training data without collecting new data.

→ **Data splitting:** Dividing the data into training, validation and test sets.

Page No.
Date : / /

Q2. Explaining the process of dividing a dataset into training, testing and validation sets. Why is this division important?

Answer Process of Dividing a Dataset into Training, Testing and Validation sets

1. Understand the Dataset:

→ Analyze the dataset to understand its size, distribution and type of data. (Ensure data is preprocessed before splitting).

2. Determine the Splitting Ratio.

→ 70 → 80% Training set
→ 10 → 20% Validation set
→ 10 → 30% Test set

3. Randomly Shuffle the Data:

→ Shuffle the dataset to ensure that samples are distributed randomly. This avoids biases from ordered data

4. Split the Data:

→ Use Splitting techniques:
↳ Manual Splitting: Dividing the data based on predifined indices.

5. Assign the sets:

→ To train → Training set

→ To tune → Validation sets

→ To test → Test sets

Importance of Dividing the Dataset

→ Avoid Overfitting:

→ without a validation set, the model may memorize the training data instead of learning general patterns, resulting in poor performance on new data.

→ Hyperparameter Tuning:-

↳ The validation set allows for testing different configurations without biasing the final evaluation.

→ Performance Evaluation:-

→ The testing set provides an unbiased assessment of the model's performance after training is complete, ensuring it generalizes well to new data.

→ Model Comparison:

→ Seperate testing data helps compare multiple models under the same conditions.

→ Generalization:

→ Dividing the data helps in creating robust models that work well in real-world scenerios, not just on the data they were trained on.

Page No.
Date : / /

Q3. Describe cross-validation and its role in improving machine learning models.

Answer: Cross-validation is a technique used to evaluate the performance of a machine learning model by dividing the data into multiple subsets and training/testing the model on different combinations of these subsets.

Role: It helps in improving the model by providing a more accurate estimate of its performance and reducing the risk of overfitting.

Q4. Compare PCA, LDA and ICA as dimensionality reduction techniques. Provides applications for each.

Answer: Principle component Analysis (PCA):

Reduces dimensionality by transforming the data into a set of orthogonal components that capture the maximum variance.
Eg. Image compression.

Linear Discriminant Analysis (LDA):

Reduces dimensionality by finding linear combinations of features that best seperate different classes.
Eg. Face recognition

Independent component Analysis (ICA):

Reduces dimensionality by finding components that are statistically independent.
Eg. Blind source seperation

Page No.
Date : / /

Q5 What are feature sets in machine learning? Discuss their importance in building predictive

Answer: Feature Sets: A collection of ~~measuring~~ measurable properties or characteristics used as input to a machine learning model.

Importance:

→ Predictive Power: Good features improve the model's ability to make accurate predictions.

→ Model Performance: The quality and relevance of features directly impact the model's performance.

Q6. Provide step-by-step instructions to perform Principal Component Analysis (PCA) on a given dataset.

→ On next page

Page No
Date : / /

→ Algorithm for PCA

S1:- Standardize the data :- Applying standard deviation in the dataset.

S2:- Compute the covarence matrix :- Calculate the covarence matrix to understand the relationships blw diff. variables.

Note:- If your datasets have n features then your covarence matrix will be of n x n order

S3:- Calculate Eigen values and Eigen vectors :- Sort the eigen values in descending order. The eigen vector associated with first eigen value is the first principle ~~order~~ component. Select the top 'k' eigen vectors to form a new subspace

S4:- Project data onto the Principal comp.: Multiply the original dataset by the selected eigen vectors (principal comp) to obtained the transformed data in new k-D space.

Q-

x	y
2	4
0	0
-1	-3

Sol:- mean of $x = 0.33 = \bar{x}$

mean of $y = 0.33 = \bar{y}$

Now

x	y	$x - \bar{x}$	$y - \bar{y}$
2	4	1.67	3.67
0	0	-0.33	-0.33
-1	-3	-1.33	-3.33

Page No.
Date : / /

Now,

Covariance matrix = $\begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix}$

Formula:-

$$cov(X,Y) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$$

$$cov(x,x) = \frac{1}{3-1} \sum_{i=1}^{2} (x_i - \bar{x})(x_i - \bar{x})$$

$$cov(x,x) = \frac{1}{2}\left[(1.67)^2 + (-0.33)^2 + (-1.33)^2\right]$$

$$= 1.67$$

$$cov(x,y) = \frac{1}{2}\left[(1.67 \times 3.67) + (-0.33 \times -0.33) + (-1.33 \times -3.33)\right]$$

$$= 3.67$$

$$cov(y,x) = cov(x,y) = 3.67$$

$$cov(y,y) = \frac{1}{2}\left[(3.67)^2 + (-0.33)^2 + (-3.33)^2\right]$$

$$= 8.33$$

$\therefore$ cov matrix = $\begin{bmatrix} 1.67 & 3.67 \\ 3.67 & 8.33 \end{bmatrix}$

Now,

Calculating eigen values,

$\lambda_1 = 10$

$\lambda_2 = 0$

for λ_1 eigen vector $v_1 = \begin{bmatrix} 0.38 \\ 0.92 \end{bmatrix}$

for λ_2, $v_2 = \begin{bmatrix} 0.92 \\ -0.38 \end{bmatrix}$

Now

Projected data = Orig. data $\times$

for first data point (1.67, 3.6

$\rightarrow (1.67 \times 0.38) + (3.67 \times 0.92)$

$= 4.011$

for second data point

$= -0.429$

for 3rd

$= -3.569$

$\therefore$ final data = $\begin{bmatrix} 4.011 \\ -0.429 \\ -3.569 \end{bmatrix}$

Page No.
Date : / /

Q7 What are the limitations of machine learning? Provide examples of challenges in its implementation.

Answer:

→ Data Dependency :- Requires large amounts of high-quality data.

→ Overfitting :- models may perform well on training data but poorly on new data.

→ Interpretability :- Complex models can be difficult to interpret

→ Bias and Fairness :- Models can inherit biases present in the training data.

Challenges:

→ Data Privacy :-
Ensuring data privacy and security.

→ Scalability :-
Handling large-scale data efficiently.

Q8. Explain how feature scaling impacts the performance of machine learning algorithms.

Answer Feature scaling standardizes the range of features to ensure they contribute equally to the model.

Impact:

→ Improves Convergence: Helps gradient-based algos converge faster

→ Prevents Bias: Ensure no single feature dominates the learning process.

Page No.
Date : / /

Q9. Determine if the following set of clause is satisfiable: $(\neg A \vee B)$, $(\neg B \vee C)$, $(\neg(C \vee D))$, $(\neg D)$

Sol:- Given clause:

$$(\neg A \vee B), (\neg B \vee C), (\neg C \vee D), (\neg D)$$

To determine satisfiability, we can use the resolution method

→ Resolve $(\neg(C \vee D))$ and $(\neg D)$ to get $\neg C$

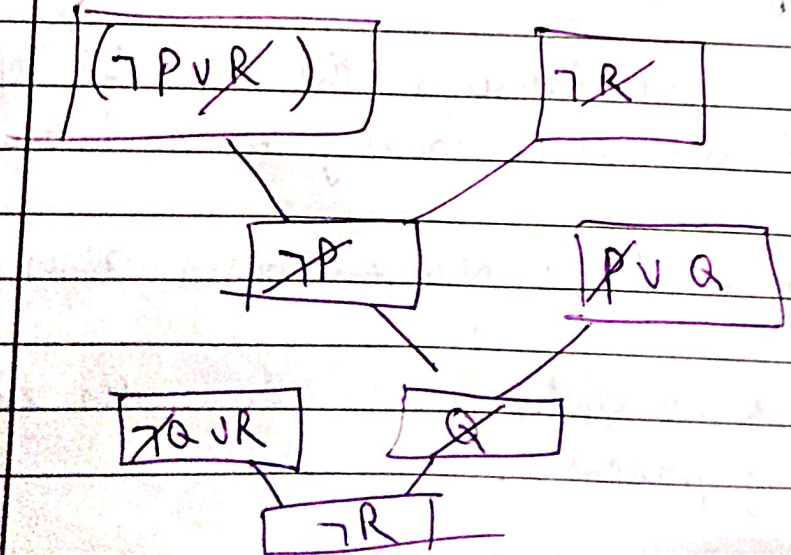
→ Resolve $(\neg B \vee C)$ and $\neg C$ to get $\neg B$

→ Resolve $(\neg A \vee B)$ and $\neg B$ to get $\neg A$

Since we do not derive a contradiction, the set of clauses is satisfiable.

Q10. Prove by resolution that the following statements entail
R: $(P \vee Q)$, $(\neg P \vee R)$, $(\neg Q \vee \neg R)$.

Sol:- let's assume $\neg R \rightarrow$ true



Page No.
Date : / /

Here we have derived $\neg R$ from assumption, which is a contradiction since $\neg R$ is already assumed in knowledge base.

$\therefore R \rightarrow$ True

H.P.