1) Differentiate between classification and regression?

2) Give general form of a confusion matrix for the binary classification

3)List out the different measures used for distance calculation

4) List some applications of machine learning in industry.

5)What is an error function in the context of a neural network?

6)What is inductive bias in decision tree learning?

7)Differentiate between k-means and hierarchical clustering

8)Define Bayes theorem

9)Write the purpose of kernel in SVM Algorithm?

10)What is reinforcement learning?

Answers :

1. Classification and regression are two different types of supervised learning in machine learning. In classification, the goal is to predict a categorical or discrete target variable. The input data is labeled with a specific class or category, and the algorithm learns to identify patterns and make predictions based on those patterns. In regression, the goal is to predict a continuous target variable. The input data is labeled with a numerical value, and the algorithm learns to identify patterns and make predictions based on those patterns.
2. The general form of a confusion matrix for binary classification is as follows:

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

3. Some of the measures used for distance calculation include Euclidean distance, Manhattan distance, Minkowski distance, Mahalanobis distance, and cosine similarity.
4. Machine learning has many applications in industry, including:

- Predictive maintenance in manufacturing
- Fraud detection in finance
- Personalized marketing in e-commerce
- Customer service automation in telecommunications
- Image recognition in healthcare
- Quality control in food production

- Supply chain optimization in logistics
- Energy consumption optimization in utilities
- Sentiment analysis in social media

5. In the context of a neural network, an error function (also called a loss function or objective function) is a measure of how well the network is performing on a specific task. It calculates the difference between the network's predicted output and the actual output, and this difference is used to update the network's parameters during training.
6. Inductive bias in decision tree learning refers to the assumptions or biases that the algorithm makes about the underlying distribution of the data. This bias affects how the decision tree is constructed and which attributes are selected as the root node and subsequent branches. The choice of inductive bias can affect the accuracy and interpretability of the resulting decision tree.
7. K-means and hierarchical clustering are two different clustering algorithms. K-means is a partition-based clustering algorithm that assigns each data point to the nearest centroid and iteratively refines the centroids until convergence. Hierarchical clustering is a hierarchical-based clustering algorithm that groups data points into a hierarchy of nested clusters. It can be either agglomerative, where each data point starts as its own cluster and clusters are successively merged, or divisive, where all data points start in the same cluster and clusters are successively split.
8. Bayes theorem is a mathematical formula that describes the probability of an event occurring based on prior knowledge of conditions that might be related to the event. It is named after the 18th-century English mathematician Thomas Bayes. The formula is:

$$P(A|B) = (P(B|A) * P(A)) / P(B)$$

where $P(A|B)$ is the conditional probability of A given B, $P(B|A)$ is the conditional probability of B given A, $P(A)$ is the prior probability of A, and $P(B)$ is the prior probability of B.

9. The purpose of a kernel in the SVM (Support Vector Machine) algorithm is to transform the input data into a higher-dimensional feature space where the classes can be more easily separated. Kernels allow SVMs to handle non-linearly separable data by transforming the input features into a new space where the data is more easily separable using a linear boundary.
10. Reinforcement learning is a type of machine learning where an agent learns to make decisions by interacting with an environment. The agent takes actions based on its current state, and the environment provides feedback in the form of rewards or punishments. The goal of reinforcement learning is to learn a policy (a mapping of states to actions)

a. List one problem that could be solved with machine learning.
b. Define Baye's theorem.
c. List out a problem, which can be best solved by Naïve Baye's classifier.
d. Define hypothesis space.
e. List one major difference between classification and regression.
f. Write one major difference between supervised and unsupervised learning.
g. How support vectors will be chosen in SVM?
h. Mcculloch-pitt neuron is same as perceptron?
i. Are KNN and K-Means belongs to same category of machine learning?
j. List out evolutionary computing methodologies.

a. Fraud detection in financial transactions.

b. Bayes' theorem is a mathematical formula that describes the probability of an event occurring based on prior knowledge of conditions that might be related to the event.

c. Email spam detection is a problem that can be best solved by Naïve Bayes classifier.

d. Hypothesis space is the set of all possible hypotheses that can be learned by a machine learning algorithm.

e. Classification predicts a categorical outcome while regression predicts a continuous outcome.

f. Supervised learning uses labeled data while unsupervised learning uses unlabeled data.

g. Support vectors are chosen in SVM based on their distance from the decision boundary.

h. McCulloch-Pitt neuron is a simplified model of a biological neuron and is not the same as a perceptron.

i. KNN and K-Means do not belong to the same category of machine learning. KNN is a supervised learning algorithm while K-Means is an unsupervised learning algorithm.

j. Evolutionary computing methodologies include genetic algorithms, particle swarm optimization, and evolutionary strategies

10 x 1 = 10M

a. Outline any two issues of Machine Learning
b. Define the inductive learning hypothesis.
c. Explain about single point crossover operation.
d. Model the program tree representation in genetic programming.
e. Outline the characteristics of Backpropagation algorithm.
f. List the features of Bayesian learning method.
g. Define Baye's theorem.
h. Illustrate Naïve Baye's Classifier.
i. Determine the formula for Euclidean distance between two features.
j. Define locally weighted linear regression.

a. Two issues of Machine Learning are:

Bias and fairness: Machine learning models can be biased towards certain groups or outcomes based on the data used to train them, which can result in unfair or discriminatory decisions.

Overfitting and underfitting: Machine learning models can sometimes perform poorly on new or unseen data if they are either too complex (overfitting) or too simple (underfitting).

b. Inductive learning hypothesis is a set of assumptions made by a machine learning algorithm about the target function to be learned based on a finite set of training examples.

c. Single point crossover operation is a genetic operator used in genetic algorithms to combine genetic information from two parent individuals by selecting a random crossover point and swapping the genetic material before and after the point to create two new offspring individuals.

d. Program tree representation in genetic programming involves representing computer programs as trees, where each node represents a function or operation, and the leaves represent input variables or constants.

e. Backpropagation algorithm is a supervised learning algorithm used for training artificial neural networks. It involves computing the error gradient of the output layer and then propagating it backwards through the network to adjust the weights of the connections between neurons.

f. Features of Bayesian learning method include:

Bayesian learning involves updating prior beliefs based on new data.

Bayesian models are flexible and can handle uncertainty and missing data.

Bayesian models can be used for both regression and classification tasks.

Bayesian models can be used to estimate model parameters and make predictions.

g. Bayes' theorem is a mathematical formula that describes the probability of an event occurring based on prior knowledge of conditions that might be related to the event.

h. Naïve Bayes Classifier is a probabilistic algorithm used for classification tasks that assumes all features are independent of each other. It involves calculating the conditional probability of each class given the observed features and selecting the class with the highest probability as the predicted class.

i. The formula for Euclidean distance between two features x and y is:

sqrt(sum(xi-yi)^2)

j. Locally weighted linear regression is a non-parametric algorithm used for supervised learning tasks that involves fitting a linear regression model to a subset of the training data based on their similarity to a given test point. The weights assigned to each training point are based on their distance from the test point.

---

a. Compare inductive and deductive learning task.

b. What is machine learning?

c. How can you avoid overfitting?

d. What is perceptron?

e. What are Bayesian networks?

f. Why instance based learning algorithm sometimes referred as Lazy

learning algorithm?

g. Why is naive Bayes so ‘naive’?

h. What is replication in genetic algorithms?

i. Compare lazy and eager learning.

j. Define kernel.

==========

a. Inductive learning is a type of machine learning where the model learns from examples and generalizes the knowledge to make predictions on new data. Deductive learning is a type of machine

learning where the model is based on pre-defined rules and logical reasoning. Inductive learning is data-driven, while deductive learning is knowledge-driven.

b. Machine learning is the process of training a computer to perform a task by providing it with examples or data. It involves using algorithms and statistical models to enable the computer to improve its performance on a specific task as it receives more data.

c. Overfitting occurs when a model is too complex and fits the training data too closely, resulting in poor performance on new, unseen data. To avoid overfitting, techniques like regularization, early stopping, and cross-validation can be used.

d. Perceptron is a type of neural network that consists of a single layer of neurons with binary outputs. It is used for binary classification tasks, where the model learns to separate data points into two classes based on a linear decision boundary.

e. Bayesian networks are probabilistic graphical models that represent the relationships between variables using a directed acyclic graph. They are used to make predictions and perform inference on complex systems with uncertain or incomplete information.

f. Instance-based learning algorithm is sometimes referred to as a lazy learning algorithm because it does not require a training phase to learn a model. Instead, it stores the training instances and uses them to classify new data points based on their similarity to the stored instances.

g. Naive Bayes is called 'naive' because it makes the assumption that the features are conditionally independent given the class label, even though this assumption may not hold in reality. However, this simplifying assumption allows for efficient and effective classification.

h. Replication in genetic algorithms refers to the process of creating multiple copies of a parent chromosome with small mutations. This helps to maintain genetic diversity in the population and prevent premature convergence.

i. Lazy learning algorithms do not construct a model from the training data, but instead, they store the training instances and use them to make predictions on new data. Eager learning algorithms, on the other hand, construct a model from the training data and use it to make predictions.

j. A kernel is a function that transforms the input data into a higher-dimensional space where it is easier to separate the data points with a linear decision boundary. Kernel functions are used in support vector machines to perform non-linear classification or regression tasks

a. Show the final design of the checkers learning program.

b. List any three issues in Machine Learning.

c. Define version space.

d. What is the purpose of backpropagation algorithm?

e. List the set of functions that can be represented by feedforward

networks.

f. What is the posterior probability of Bayes theorem?

g. Represent Bayes optimal classification.

h. Define Gibbs algorithm.

i. List any two disadvantages of instance-based approaches.

j. List the three key properties shared by k-nearest neighbor.

Answer :

b. Three issues in Machine Learning are overfitting, underfitting, and data bias.

c. Version space is the set of all hypotheses that are consistent with the observed training data in machine learning.

d. The purpose of the backpropagation algorithm is to adjust the weights of a neural network in order to minimize the difference between the network's predicted output and the actual output.

e. Feedforward networks can represent a wide range of functions, including linear functions, nonlinear functions, and functions that map one input to multiple outputs or multiple inputs to one output.

f. The posterior probability of Bayes theorem is the probability of a hypothesis H given some observed data D, P(H|D).

g. Bayes optimal classification is a probabilistic approach to classification that selects the hypothesis with the highest posterior probability.

h. Gibbs algorithm is a sampling technique used to estimate the joint probability distribution of a set of variables.

i. Two disadvantages of instance-based approaches are their sensitivity to noise and their high computational complexity.

j. The three key properties shared by k-nearest neighbor are locality, similarity, and laziness.