1. Define concept of classification.

   A) Classification refers to the process of categorizing objects, data or information into different classes or groups based on their characteristics, properties, or attributes. The goal of classification is to identify patterns in the data and assign them to predefined classes or categories, such as in supervised learning. This process involves training a model on a labeled dataset, where each instance has a known class or category, and then using the model to classify new, unseen data.

   Classification is a fundamental task in machine learning and is used in a wide range of applications, such as image and speech recognition, fraud detection, sentiment analysis, and spam filtering. Some popular classification algorithms include decision trees, logistic regression, support vector machines, and neural networks. The choice of algorithm depends on the nature of the data and the problem being solved.

2. How you will design a machine learning system? Steps of developing a machine learning system.
   A)Designing a machine learning system involves several steps, including:

   - Problem formulation: The first step is to define the problem and set the goals of the machine learning system. This involves understanding the business problem or task that needs to be solved and the data that will be used to train the model.

   - Data collection and preparation: The second step is to gather and preprocess the data. This involves selecting the appropriate data sources, cleaning and filtering the data, and transforming it into a format suitable for machine learning algorithms.

   - Feature engineering: The third step is to extract or create relevant features from the data. This involves selecting the most informative attributes that capture the patterns in the data.

   - Model selection: The fourth step is to choose the appropriate machine learning algorithm(s) that will be used to train the model. This involves considering the trade-offs between accuracy, interpretability, and computational complexity.

   - Model training: The fifth step is to train the selected model(s) on the prepared data. This involves splitting the data into training and validation sets, selecting hyperparameters, and optimizing the model performance.

   - Model evaluation: The sixth step is to evaluate the trained model(s) using appropriate metrics, such as accuracy, precision, recall, F1-score, or area under the ROC curve. This step helps to assess the model's performance and identify areas for improvement.

   - Model deployment: The final step is to deploy the trained model into a production environment, where it can make predictions on new, unseen data.

This involves integrating the model with the application or system, monitoring its performance, and updating it over time.

3. What are real life applications of machine learning?
A)Machine learning has numerous real-life applications across various industries, some of which include:
    - Image and Speech Recognition: Machine learning is used in image and speech recognition to identify patterns and features in images and sounds. Applications include facial recognition, voice recognition, and object detection.
    - Natural Language Processing (NLP): NLP is a subfield of machine learning that deals with the interaction between computers and humans in natural language. Applications include sentiment analysis, chatbots, and machine translation.
    - Fraud Detection: Machine learning is used to detect fraudulent activities in financial transactions, insurance claims, and credit card purchases.
    - Healthcare: Machine learning is used in healthcare to predict disease outbreaks, diagnose diseases, and develop personalized treatment plans.
    - Predictive Maintenance: Machine learning is used to predict when machines or equipment will fail, enabling preventive maintenance and minimizing downtime.
    - Recommendation Systems: Machine learning is used to create personalized recommendations for products, services, and content based on user behavior and preferences.
    - Autonomous Vehicles: Machine learning is used in autonomous vehicles to recognize objects and obstacles, plan routes, and make decisions in real-time.
    - Retail: Machine learning is used to analyze customer data, optimize pricing strategies, and forecast demand for products.
    - Energy Management: Machine learning is used to optimize energy consumption, predict energy usage patterns, and identify opportunities for energy savings.

Overall, machine learning is a versatile technology that has applications in almost every industry, from finance to healthcare, to transportation, to marketing, and beyond.

4. List and explain issues in machine learning

   A)

1.4 ISSUES IN MACHINE LEARNING

1. Which algorithm we have to select to learn general target functions from specific training dataset? What should be the settings for particular algorithms, so as to converge to the desired function, given sufficient training data? Which algorithms perform best for which type of problems and representations?
2. How much training data is sufficient? What should be the general amount of datathat can be found to relate the confidence in learned hypotheses to the amount training experience and the character of the learner's hypothesis space?
3. Prior knowledge held by the learner is used at which time and manner to guide the process of generalizing from examples? If we have approximately correct knowledge, will it helpful even when it is only approximately correct?
4. What is the best strategy for choosing a useful next training experience, and how does the choice of this strategy after the complexity of the learning problem?
5. To reduce the task of learning to one or more function approximation problems, what will be the best approach? What specific functions should the system attempt to learn? Can this process itself be automated?
6. To improve the knowledge representation and to learn the target function, how can the learner automatically alter its representation?

5. Calculate eigen vector of a given matrix

   A = 1 2 -3

   2 4 -6

   -1 -2 3
6. What are the performance measures to analyze quality of model?

A) The performance of a machine learning model can be evaluated using several performance measures. The choice of performance measure depends on the problem type and the goals of the model. Some common performance measures include:

- Accuracy: The percentage of correctly classified instances in the test set.
- Precision: The ratio of true positives to the total number of predicted positives. Precision measures the model's ability to identify only the relevant instances.
- Recall: The ratio of true positives to the total number of actual positives. Recall measures the model's ability to find all relevant instances.
- F1-score: The harmonic mean of precision and recall. F1-score is a balanced measure that combines precision and recall.
- Mean Squared Error (MSE): The average squared difference between the predicted and actual values. MSE is used for regression problems.
- Root Mean Squared Error (RMSE): The square root of the mean squared error. RMSE is used for regression problems.

- Confusion Matrix: A matrix that summarizes the number of true positives, false positives, true negatives, and false negatives. Confusion matrix is used to calculate other performance measures such as accuracy, precision, recall, and F1-score.

7. Explain overfitting and underfitting of model.
   A) Overfitting and underfitting are two common problems that can occur when building machine learning models.
   - Overfitting occurs when a model is too complex and fits the training data too closely, to the point that it starts to capture the noise or random fluctuations in the data. This can lead to poor generalization performance on new data, because the model has essentially memorized the training data and is not able to make accurate predictions on unseen data. Overfitting is more likely to occur when the model has a large number of parameters relative to the size of the training data, or when the training data is noisy or contains outliers.
   - Underfitting occurs when a model is too simple and does not capture the underlying patterns in the data. This can also lead to poor generalization performance, because the model is not able to make accurate predictions even on the training data. Underfitting is more likely to occur when the model is too simple relative to the complexity of the underlying patterns in the data, or when the training data is too limited or poorly representative of the problem domain.
8. Calculate SVD of a given matrix A = 1 0 1

   -2 1 0

9. Diagonalize the given matrix A as A= XDX-1(inv)
   A = 1 1 1
   1 1 1
   1 1 1.
10. Explain support vector machine.
    A) Support Vector Machine(SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems as well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane. It becomes difficult to imagine when the number of features exceeds three.

    Select the best line or the best hyperplane that segregates our data points :

    choose the hyperplane whose distance from it to the nearest data point on each side is maximized. If such a hyperplane exists it is known as the maximum-margin hyperplane/hard margin.

    SVM is robust to outliers ie The SVM algorithm has the characteristics to ignore the outlier and finds the best hyperplane that maximizes the margin

11. What is regularized regression?
    A) Regularized regression is a type of linear regression that is used to prevent overfitting in models with a large number of features. In traditional linear regression, the goal is to minimize the sum of squared errors between the predicted and actual values.
    However, in situations where there are many features, the model may become too complex and overfit the data, leading to poor generalization performance on new data.

    Regularized regression introduces a penalty term to the loss function that encourages the model to have smaller coefficients, effectively shrinking the coefficients towards zero. This penalty term helps to reduce the complexity of the model and prevent overfitting. The two most common types of regularized regression are Ridge Regression and Lasso Regression.

    it is important to choose the right type of regularization and the appropriate value of the regularization parameter to achieve the best results.

12. Explain norm of a vector.
    A)

3.1 Norms 71

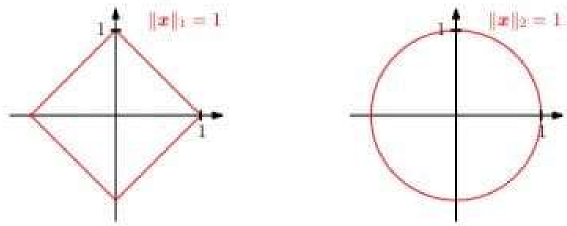

Figure 3.3 For different norms, the red lines indicate the set of vectors with norm 1. Left: Manhattan norm; Right: Euclidean distance.

3.1 Norms

When we think of geometric vectors, i.e., directed line segments that start at the origin, then intuitively the length of a vector is the distance of the "end" of this directed line segment from the origin. In the following, we will discuss the notion of the length of vectors using the concept of a norm.

Definition 3.1 (Norm). A *norm* on a vector space V is a function (norm)

$$\|\cdot\| : V \to \mathbb{R}, \quad (3.1)$$
$$\boldsymbol{x} \mapsto \|\boldsymbol{x}\|, \quad (3.2)$$

which assigns each vector $\boldsymbol{x}$ its *length* $\|\boldsymbol{x}\| \in \mathbb{R}$, such that for all $\lambda \in \mathbb{R}$ and $\boldsymbol{x}, \boldsymbol{y} \in V$ the following hold: (length)

- *Absolutely homogeneous:* $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$ (absolutely homogeneous)
- *Triangle inequality:* $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (triangle inequality)
- *Positive definite:* $\|\boldsymbol{x}\| \geqslant 0$ and $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \boldsymbol{0}$ (positive definite)

Figure 3.2 Triangle inequality.

In geometric terms, the triangle inequality states that for any triangle, the sum of the lengths of any two sides must be greater than or equal to the length of the remaining side; see Figure 3.2 for an illustration. Definition 3.1 is in terms of a general vector space V (Section 2.4), but in this book we will only consider a finite-dimensional vector space $\mathbb{R}^n$. Recall that for a vector $\boldsymbol{x} \in \mathbb{R}^n$ we denote the elements of the vector using a subscript, that is, x_i is the i^{th} element of the vector $\boldsymbol{x}$.

13. Explain supervised machine learning.
    A) In this type of learning we use data which is comprises of input and corresponding output. For every instance of data we can have input X and corresponding output "Y". From this ML system will build model so that given an observation 'X', for new observation "X" it will try to find out what is corresponding "Y". In supervised learning training data is labelled with the correct answers, e.g. "spam" or "ham." Two most important types of supervised learning are classification (where the outputs are discrete labels, as in spam filtering) and regression (where the outputsare real-valued).

14. Explain unsupervised machine learning

A) In unsupervised learning you are only given input 'X', there is no label to the data and given the data or different data points, you may want to form clusters or want to find some pattern. Two important unsupervised learning tasks dimension reduction and clustering.

15. Find vectors that are orthogonal to [1,2,3]. Explain why we can have infinite number of such vectors.

A)

better method

Qb 15) orthogonal to [1, 2, 3]

Unknown vector = (x, y, z)

dot product formula:-

$[1,2,3](x,y,z] = 1x+2y+3z = 0$

~~for~~ lets $z = t$ and $y = s$

$1x + 2s + 3t = 0$ — (1)

$1x = -2s-3t$

$x = -2s-3t$

So any vector in the form of $(-2s-3t, s, t]$ is orthogonal to $[1, 2, 3]$

to show that any value in s and t are orthogonal we'll put $s = 1$ and $t = -1$

So in eqn (1)

$1x + 2s + 3t = 0$

$1(-2(1)-3(-1)) + 2(1) + 3(-1) = 0$

$1(-1) + 2(1) + 3(-1) = 0$

$-1+2-3 = 0$

$2-1-3 = 0$

$2-2 = 0$

$0 = 0$

∴ Since dot product is zero we have confirmed that $[-1, 1, -1]$ is orthogonal to $[1, 2, 3]$

- We can have an infinite number of such vectors because there are infinitely many choices for x and y.
- For any fixed value of z and y, we can compute the corresponding value of x using the above equation, and obtain a vector that is orthogonal to [1, 2, 3].
- This means that there are infinitely many planes that are orthogonal to [1, 2, 3], and each plane contains infinitely many vectors that are orthogonal to [1, 2, 3].

16. Explain least squares method for supervised machine learning technique.

A) The least squares method is a common technique used in supervised machine learning for linear regression problems. In this method, the goal is to find the line (for simple linear regression) or hyperplane (for multiple linear regression) that best fits the given set of data points.

The least squares method minimizes the sum of squared errors between the predicted values and the actual values in the training data. Mathematically, this can be expressed as:

$\min \Sigma (y - \hat{y})^2$

where y is the actual value of the target variable, ŷ is the predicted value of the target variable, and the sum is taken over all training examples.

To solve this optimization problem, the least squares method uses the normal equation, which gives the values of the model parameters that minimize the sum of squared errors. The normal equation is given by:

$\theta = (X^T X)^{-1} X^T y$

where θ is the vector of model parameters (including the intercept term), X is the design matrix that contains the input features, y is the vector of target values, and $(X^T X)^{-1}$ is the inverse of the matrix product of the transpose of X and X.

Once the model parameters are estimated using the least squares method, the model can be used to make predictions on new data by simply computing the dot product between the input features and the model parameters.

The least squares method is a simple and widely used technique for linear regression problems. However, it may not be suitable for more complex nonlinear regression problems, where other techniques such as polynomial regression or kernel regression may be more appropriate.