1   What is supervised learning, and how does it differ from unsupervised learning?

Mathematical Component: Write the general form of a supervised learning model. Explain the role of the loss function $L(y, \hat{y})$ in training the model.

2   Explain the assumptions underlying linear regression. How would you check if these assumptions hold for your dataset?

Mathematical Component: Given a dataset (X, y), write the equation for the linear regression model and derive the normal equations for finding the regression coefficients β.

3   Describe the difference between logistic regression and decision tree classifiers. When would you choose one over the other?

Mathematical Component: For logistic regression, write the logistic function $\sigma(z) = \frac{1}{1+e^{-z}}$ and derive the gradient of the log-likelihood function with respect to the weights.

4   How does a Support Vector Machine (SVM) classify data? Explain the concept of the kernel trick.

Mathematical Component: Write the optimization problem for the linear SVM and describe the constraints.

5   What is unsupervised learning? Provide two real-world applications.

Mathematical Component: Given a dataset $\{x_1, x_2, \ldots, x_n\}$, write the objective function for the K-means clustering algorithm.

6   Explain the difference between K-means and hierarchical clustering. What are the advantages and disadvantages of each?

Mathematical Component: Describe how the Euclidean distance metric is used in the K-means algorithm to update cluster centroids.

7   Discuss the difference between correlation and causation. Provide an example of a situation where correlation does not imply causation.

Mathematical Component: Given two random variables X and Y, write the formula for the Pearson correlation coefficient and describe what it measures.

8   Describe the difference between a probability density function (pdf) and a cumulative distribution function (cdf).

Mathematical Component: Given a continuous random variable X with pdf $f(x)$, write the cdf $F(x)$.