

1. Explain what eigenvalues and eigenvectors are. How are they applied in Principal Component Analysis (PCA)?

Eigenvectors: are directions in the data.

Eigenvalues: show how important each direction is.

In PCA : we use eigenvectors to find the most important directions in the data.

Eigenvalues help us choose which directions keep more information.

PCA helps to reduce the number of features and keep the important data.

2. What is the difference between Grid Search and Random Search in hyperparameter tuning?

- Grid Search try all possible values of hyperparameters.
It is accurate but very slow.
- Random Search try random values.
It is faster and sometimes gives good results.
- Both methods are used to find the best parameters for the model.

3. What is the difference between correlation and causation? Why is it important to distinguish them?

Correlation means two variables change together.

Causation means one variable causes the other.

It is important to know the difference because correlation does not always mean cause.

Wrong understanding can lead to wrong decisions.

4. How can we evaluate the performance of classification and regression models? Mention key metrics for each.

For classification models, we use metrics like Accuracy, Precision, Recall, and F1-score.

They help us know how good the model is.

For regression models, we use MAE, MSE, RMSE, and R².

They measure how close the predictions are to the real values.

5. Explain Gradient Descent. How does it work to minimize a loss function, and what are the differences between Batch, Stochastic, and Mini-batch Gradient Descent?

Gradient Descent is an algorithm used to minimize the loss function.

It works by changing model parameters step by step to reduce error.

Batch Gradient Descent uses all data.

Stochastic Gradient Descent uses one sample.

Mini-batch Gradient Descent uses part of the data.