Answer

1) Factor analysis is a statical method used to identify underlying relationships or latent factor that explain the correlation among a set observed variable.

It reduce data dimensionality steps:

1) data collection and preparation.
2) correlation matrix
3) Factor extraction
4) factor rotation
5) Interpretation and validation.

2) Eigenvectors are special vectors associated with a square matrix that, when multiplied by the matrix, result in a scalar multiple of themselves. Their properties are critical in linear algebra and applications like PCA

Properties:

1) linear independence
2) Direction preservation
3) orthogonality
4) non-zero
5) Basis formation

clustering is an unsupervised learning technique that group similar data points into clusters based on their features without prior labels. that principle aims to maximize intra-cluster similarity and minimize inter-cluster similarity

Key principle:-

1) similarity-Based grouping
2) Distance minimization
3) No predefined Labels
4) feature Relevance
5) Iterative optimization

High dimensional datasets often referred as curse of dimensionality

key problems :-

1) increased sparsity
2) computational complexity
3) overfitting risk
4) feature redundancy
5) distance metric issues

5) overfitting :-

occurs when a ML models learns the training data too well, including noise, leading to high accuracy on training data but poor on test data

underfitting :-

poor performance on both training and test data.

prevention methods

1) Regularization
2) Cross validation
3) Dimensionality reduction
4) increase Training data.