

# **Assignment: Anomaly Detection and Time Series**

## **Q1. What is Dimensionality Reduction? Why is it important?**

Ans. It reduces the number of features while keeping useful information. Helps avoid overfitting, improves model speed, and supports visualisation.

## **Q2. Name and briefly describe 3 dimensionality reduction techniques.**

Ans. PCA: Reduces features using maximum variance.

T-SNE: Non-linear reduction for visualising high-dimensional data.

Autoencoders: Neural-network-based feature compression.

## **Q3. What is clustering? Mention 3 clustering algorithms.**

Ans. It is grouping similar data points without labels.

Algorithms: K-Means, Hierarchical, DBSCAN.

## **Q4. Define anomaly detection and its significance.**

Ans. Identifying unusual data points that deviate from normal patterns. Important for fraud detection, system faults, and security.

## **Q5. List and describe 3 anomaly detection techniques.**

Ans. Statistical: Uses probability distribution (e.g., z-score).

Proximity-based: Uses distance or density (kNN, LOF).

Model-based: Uses ML models like One-Class SVM, Autoencoders.

## **Q6. What is Time Series Analysis? Name 2 components.**

Ans. Study of time-based data for patterns and forecasting.

Components: Trend and Seasonality.

## **Q7. Difference between Seasonality and Cyclic behaviour.**

Ans. Seasonality repeats at fixed time intervals.

Cyclic patterns are irregular and depend on external factors like the economy.

## **Q8. Python code for K-Means clustering (short):**

Ans.

```
from sklearn.cluster import KMeans  
km = KMeans(n_clusters=4).fit(X)  
print(km.labels_, km.cluster_centers_)
```

## **Q9. What is Inheritance in OOP? Give an example.**

Ans. It allows a class to gain features of another class.

Example: `class Dog(Animal): pass`

**Q10. How is Time Series used for anomaly detection?**

Ans. By forecasting expected values and marking large deviations as anomalies.