PW SKILLS

Assignment – 5

ML

1. What is clustering in machine learning?
2. Explain the difference between supervised and unsupervised clustering.
3. What are the key applications of clustering algorithms?
4. Describe the K-means clustering algorithm.
5. What are the main advantages and disadvantages of K-means clustering?
6. How does hierarchical clustering work?
7. What are the different linkage criteria used in hierarchical clustering?
8. Explain the concept of DBSCAN clustering.
9. What are the parameters involved in DBSCAN clustering?
10. Describe the process of evaluating clustering algorithms.
11. What is the silhouette score, and how is it calculated?
12. Discuss the challenges of clustering high-dimensional data.
13. Explain the concept of density-based clustering.
14. How does Gaussian Mixture Model (GMM) clustering differ from K-means?
15. What are the limitations of traditional clustering algorithms?
16. Discuss the applications of spectral clustering.
17. Explain the concept of affinity propagation.
18. How do you handle categorical variables in clustering?
19. Describe the elbow method for determining the optimal number of clusters.
20. What are some emerging trends in clustering research?
21. What is anomaly detection, and why is it important?
22. Discuss the types of anomalies encountered in anomaly detection.
23. Explain the difference between supervised and unsupervised anomaly detection techniques.
24. Describe the Isolation Forest algorithm for anomaly detection.
25. How does One-Class SVM work in anomaly detection?
26. Discuss the challenges of anomaly detection in high-dimensional data.
27. Explain the concept of novelty detection.
28. What are some real-world applications of anomaly detection?

29. Describe the Local Outlier Factor (LOF) algorithm.
30. How do you evaluate the performance of an anomaly detection model?
31. Discuss the role of feature engineering in anomaly detection.
32. What are the limitations of traditional anomaly detection methods?
33. Explain the concept of ensemble methods in anomaly detection.
34. How does autoencoder-based anomaly detection work?
35. What are some approaches for handling imbalanced data in anomaly detection?
36. Describe the concept of semi-supervised anomaly detection.
37. Discuss the trade-offs between false positives and false negatives in anomaly detection.
38. How do you interpret the results of an anomaly detection model?
39. What are some open research challenges in anomaly detection?
40. Explain the concept of contextual anomaly detection.
41. What is time series analysis, and what are its key components?
42. Discuss the difference between univariate and multivariate time series analysis.
43. Describe the process of time series decomposition.
44. What are the main components of a time series decomposition?
45. Explain the concept of stationarity in time series data.
46. How do you test for stationarity in a time series?
47. Discuss the autoregressive integrated moving average (ARIMA) model.
48. What are the parameters of the ARIMA model?
49. Describe the seasonal autoregressive integrated moving average (SARIMA) model.
50. How do you choose the appropriate lag order in an ARIMA model?
51. Explain the concept of differencing in time series analysis.
52. What is the Box-Jenkins methodology?
53. Discuss the role of ACF and PACF plots in identifying ARIMA parameters.
54. How do you handle missing values in time series data?
55. Describe the concept of exponential smoothing.
56. What is the Holt-Winters method, and when is it used?

57. **Discuss the challenges of forecasting long-term trends in time series data.**

58. **Explain the concept of seasonality in time series analysis.**

59. **How do you evaluate the performance of a time series forecasting model?**

60. **What are some advanced techniques for time series forecasting?**