

Assignment – 1

ML

- 1. Define Artificial Intelligence (AI).**
- 2. Explain the differences between Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Data Science (DS).**
- 3. How does AI differ from traditional software development?**
- 4. Provide examples of AI, ML, DL, and DS applications.**
- 5. Discuss the importance of AI, ML, DL, and DS in today's world.**
- 6. What is Supervised Learning?**
- 7. Provide examples of Supervised Learning algorithms.**
- 8. Explain the process of Supervised Learning.**
- 9. What are the characteristics of Unsupervised Learning?**
- 10. Give examples of Unsupervised Learning algorithms.**
- 11. Describe Semi-Supervised Learning and its significance.**
- 12. Explain Reinforcement Learning and its applications.**
- 13. How does Reinforcement Learning differ from Supervised and Unsupervised Learning?**
- 14. What is the purpose of the Train-Test-Validation split in machine learning?**
- 15. Explain the significance of the training set.**
- 16. How do you determine the size of the training, testing, and validation sets?**
- 17. What are the consequences of improper Train-Test-Validation splits?**
- 18. Discuss the trade-offs in selecting appropriate split ratios.**
- 19. Define model performance in machine learning.**
- 20. How do you measure the performance of a machine learning model?**
- 21. What is overfitting and why is it problematic?**
- 22. Provide techniques to address overfitting.**
- 23. Explain underfitting and its implications.**
- 24. How can you prevent underfitting in machine learning models?**
- 25. Discuss the balance between bias and variance in model performance.**
- 26. What are the common techniques to handle missing data?**
- 27. Explain the implications of ignoring missing data.**
- 28. Discuss the pros and cons of imputation methods.**

29. How does missing data affect model performance?
30. Define imbalanced data in the context of machine learning.
31. Discuss the challenges posed by imbalanced data.
32. What techniques can be used to address imbalanced data?
33. Explain the process of up-sampling and down-sampling.
34. When would you use up-sampling versus down-sampling?
35. What is SMOTE and how does it work?
36. Explain the role of SMOTE in handling imbalanced data.
37. Discuss the advantages and limitations of SMOTE.
38. Provide examples of scenarios where SMOTE is beneficial.
39. Define data interpolation and its purpose.
40. What are the common methods of data interpolation?
41. Discuss the implications of using data interpolation in machine learning.
42. What are outliers in a dataset?
43. Explain the impact of outliers on machine learning models.
44. Discuss techniques for identifying outliers.
45. How can outliers be handled in a dataset?
46. Compare and contrast Filter, Wrapper, and Embedded methods for feature selection.
47. Provide examples of algorithms associated with each method.
48. Discuss the advantages and disadvantages of each feature selection method.
49. Explain the concept of feature scaling.
50. Describe the process of standardization.
51. How does mean normalization differ from standardization?
52. Discuss the advantages and disadvantages of Min-Max scaling.
53. What is the purpose of unit vector scaling?
54. Define Principle Component Analysis (PCA).
55. Explain the steps involved in PCA.
56. Discuss the significance of eigenvalues and eigenvectors in PCA.
57. How does PCA help in dimensionality reduction?
58. Define data encoding and its importance in machine learning.
59. Explain Nominal Encoding and provide an example.

60. Discuss the process of One Hot Encoding.
61. How do you handle multiple categories in One Hot Encoding?
62. Explain Mean Encoding and its advantages.
63. Provide examples of Ordinal Encoding and Label Encoding.
64. What is Target Guided Ordinal Encoding and how is it used?
65. Define covariance and its significance in statistics.
66. Explain the process of correlation check.
67. What is the Pearson Correlation Coefficient?
68. How does Spearman's Rank Correlation differ from Pearson's Correlation?
69. Discuss the importance of Variance Inflation Factor (VIF) in feature selection.
70. Define feature selection and its purpose.
71. Explain the process of Recursive Feature Elimination.
72. How does Backward Elimination work?
73. Discuss the advantages and limitations of Forward Elimination.
74. What is feature engineering and why is it important?
75. Discuss the steps involved in feature engineering.
76. Provide examples of feature engineering techniques.
77. How does feature selection differ from feature engineering?
78. Explain the importance of feature selection in machine learning pipelines.
79. Discuss the impact of feature selection on model performance.
80. How do you determine which features to include in a machine-learning model?