

L5 SOD&NIT (A&B)

MACHINE LEARNING HOMEWORK

To be submitted on Monday

1. Explain the main steps involved in the **Machine Learning Life Cycle** and why each step is important.
2. Compare and contrast **Supervised, Unsupervised, and Semi-Supervised Learning**. Provide a real-world example for each.
3. What are the advantages and disadvantages of using **Machine Learning** in real-world applications? Provide examples.
4. How does **Reinforcement Learning** differ from other types of machine learning? Give an example of its application.
5. Describe the role of **Python** and popular libraries like **TensorFlow** and **Scikit-learn** in the machine learning workflow.
6. Why is it important to prepare a **Machine Learning Environment** before starting a project? What tools are typically used?
7. Explain the difference between **Artificial Intelligence, Machine Learning, and Deep Learning** using a comparison table.
8. Define the following terms: **Data, Information, Dataset, Data Warehouse, and Big Data**. How do they relate to machine learning?
9. What are the **6 V's of Big Data**? Explain each with an example.
10. Compare **Structured, Semi-Structured, and Unstructured Data**. Give an example of each type.
11. Describe at least three **sources of data** used in machine learning and how they can be collected.
12. What are some common **methods for gathering a machine learning dataset**? Include both manual and automated approaches.

13. How can **public datasets** like those from Kaggle or UCI be useful in machine learning projects?
14. Explain the process of **loading a dataset using Pandas and Scikit-learn**. Provide a code example for each.
15. What is the purpose of **data visualization** in machine learning? Name three tools used for visualization and their key features.
16. Compare **Matplotlib, Seaborn, and Plotly**. When would you choose one over the others?
17. Describe the use cases for the following types of plots: **Scatter Plot, Line Plot, Bar Chart, Histogram, Box Plot, and Heatmap**.
18. What are some **best practices** to follow when creating data visualizations? Why are they important?
19. How can you **interpret patterns and trends** from a line plot? Provide an example.
20. Explain how **correlation and causation** differ. How can visualizations help identify correlation?