# Assignment-03
# Introduction to ML and Data Preprocessing

| | |
|---|---|
| **Submission Deadline:** | **10PM 9th November 2025** |
| **Submission Link:** | https://forms.office.com/r/XuSfVAqRc4 |
| **Instructions** | 1- All answers must be written on A4 sheets.<br>2- Only handwritten submissions will be accepted.<br>3- Write your Name and Roll No. at the top of every page. |

**Q1.** Explain what is meant by *Machine Learning*. How does it differ from traditional programming?
*(Answer in about 150–200 words.)*

---

**Q2. Differentiate between** Supervised**,** Unsupervised**, and** Reinforcement Learning **with suitable** examples **for each.**
***(Tabular format preferred.)***

---

**Q3. Describe any** two real-world case studies **where Machine Learning has been effectively used. Discuss the** problem**,** data used**, and** ML model applied**.**

---

**Q4. Why is data preprocessing important in ML pipelines? Explain the main steps involved** — data cleaning, normalization, and transformation — **with examples.**

---

**Q5. Explain** feature selection **and** dimensionality reduction**.**
**Give one real-world example of when dimensionality reduction (e.g., PCA) can be helpful.**

---

**Q6. Write a Python program using** pandas **and** scikit-learn **to:**

- Load a dataset (e.g., sklearn.datasets.load_wine())
- Handle missing values (if any)

- Normalize the features
- Apply **PCA** to reduce the data to 2 dimensions
  Display the transformed dataset.

*(Include code and output screenshot if possible.)*

---

**Q7. What is** regression analysis**? Differentiate between** Linear **and** Polynomial Regression**. Give one scenario where each would be appropriate.**

---

**Q8. Using scikit-learn, implement** Linear Regression **on a simple dataset (e.g., predicting house prices or student marks).**
**Display:**

- Coefficients
- Intercept
- Predicted vs actual values (plot using matplotlib)

---

**Q9. Explain the need for** regularization **in regression. Compare** Ridge **and** Lasso **regression in terms of:**

- Mathematical idea
- Impact on model coefficients
- Use cases

---

**Q10. Implement Polynomial Regression using scikit-learn on a small dataset (you may generate synthetic data).**
**Tasks:**

- Create sample data (e.g., using np.linspace and a quadratic relationship).
- Fit both Linear and Polynomial Regression (degree = 2 or 3) models.
- Plot both regression curves on the same graph for comparison.
- Comment on how polynomial regression improves the model's fit.

---