<p style="text-align:center">**Day 10 Training Report**</p>

4 July 2025

Supervised Learning Mini Project — Housing Price Prediction / Spam Classification / Iris Dataset

On **Day 10**, students applied all the **supervised learning concepts** learned so far in a **mini project**. The goal was to **combine theory with practical implementation** on a real dataset.

---

1. Project Objective

Students could choose one of the following datasets/projects:

1. **Housing Price Prediction** (Regression) – Predict house prices based on features like area, bedrooms, and location.
2. **Spam Classification** (Classification) – Identify whether an email is spam or not.
3. **Iris Dataset Classification** (Classification) – Predict the species of iris flowers based on petal and sepal measurements.

---

2. Steps Followed in the Project

1. **Data Loading:** Imported datasets using Pandas.
2. **Data Cleaning:** Handled missing values, duplicates, and outliers.
3. **Data Preprocessing:** Scaled features if required and performed encoding for categorical data.
4. **Model Selection:** Chose a suitable algorithm:
   - Regression → Linear Regression
   - Classification → Logistic Regression, k-NN
5. **Model Training & Testing:** Split data into training and test sets, trained the model, and made predictions.
6. **Evaluation:** Used **accuracy, confusion matrix, precision, recall**, or **RMSE** depending on the task.
7. **Visualization:** Visualized results using **Matplotlib** or **Seaborn**.

---

3. Example Workflow (Iris Dataset)

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score

iris = load_iris()
```

```
X_train, X_test, y_train, y_test = train_test_split(iris.data, iris.target, test_size=0.2,
random_state=42)

model = KNeighborsClassifier(n_neighbors=3)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

print("Accuracy:", accuracy_score(y_test, y_pred))
```

Outcome:

- Predicted species of iris flowers accurately.
- Evaluated model performance using metrics and visualized predictions.