Flood Risk Prediction – Machine Learning Hands-on Activity

■■ **Duration:** 45 Minutes

■ Objective:

The objective of this activity is to apply supervised machine learning algorithms to predict the risk of flooding based on environmental and meteorological features. Students will utilize models such as Logistic Regression, Decision Tree, Random Forest, SVM, Ridge, and Lasso Regression to analyze real-world data and make predictions.

■ Dataset:

You will use the publicly available dataset from Kaggle to perform this activity:
Flood Prediction Dataset – Kaggle

■ Problem Statement:

Floods are one of the most frequent and devastating natural disasters, often resulting from heavy rainfall, river overflow, or inadequate drainage systems. The ability to predict flood occurrence based on environmental conditions can significantly aid disaster preparedness and response efforts. In this activity, you are tasked with building and evaluating machine learning models that predict whether a flood is likely to occur given a set of meteorological and hydrological features.

■ Activity Tasks (45 Minutes):

- 1. Load the dataset and explore its structure (columns, data types, missing values, etc.).
- 2. Perform basic data preprocessing (handle missing values, encode categorical variables if any, and split the data into train/test sets).
- 3. Train at least three machine learning models from the following list: Logistic Regression, Decision Tree, Random Forest, SVM, Ridge Regression, Lasso Regression.
- 4. Evaluate each model using accuracy and F1-score. Present the results clearly.
- 5. Identify the most important features influencing flood occurrence using feature importance or model coefficients.

- 6. (Bonus) Interpret the results and briefly discuss how such a model could be useful in real-world disaster management.

■ Deliverables:

At the end of the activity, each student must submit a short Jupyter notebook (.ipynb) or Python script (.py) containing the following:
• Data preprocessing steps
• Model training and evaluation
• Feature importance analysis
• Final conclusions