QSkill Internship

- **Domain:** Artificial Intelligence & Machine Learning
- **Starting Date: 10th January, 2026**
- **Ending Date: 10th February, 2026**
- **Deadline for task submission: 10th February, 2026**
- **Do any 2 tasks from the given below 3 tasks.**

1. Iris Flower Classification

Objective: Classify iris flowers into three species (Setosa, Versicolor, Virginica) based on measurements of their petals and sepals.

- **Dataset:** The classic Iris dataset from UCI Repository or scikit-learn.
- **Steps:**
  - Load the dataset and explore it visually (scatter plots or histograms).
  - Split the data into training/test sets.
  - Preprocess if needed (usually, it's already clean).
  - Train a simple classifier (e.g., Logistic Regression, K-Nearest Neighbors, Decision Tree).
  - Evaluate with accuracy, precision, or confusion matrix.
- **Skills gained:** Numeric data analysis, classification modeling, evaluating results.

2. Spam Mail Detector

Objective: Build a classifier that distinguishes between spam and non-spam (ham) emails using textual data.

- **Dataset:** Public datasets like the SMS Spam Collection (UCI) or Enron Email Dataset.
- **Steps:**
  - Load the messages and labels (spam or ham).
  - Preprocess the text (lowercasing, remove stopwords, tokenization).
  - Convert text into numeric features (Bag of Words or TF-IDF).
  - Split into train/test sets.
  - Train a simple model (Naive Bayes, Logistic Regression).
  - Measure performance with accuracy, precision, or F1 score.
- **Skills gained:** Text preprocessing, feature extraction, basic NLP, classification.

3. House Price Prediction

Objective: Predict the price of a house based on features such as size, location, and number of bedrooms.

- **Dataset:** Boston Housing dataset or any basic housing dataset with numeric features.
- **Steps:**
  - Load the dataset and explore data distributions.
  - Handle missing data and preprocess inputs (normalization).
  - Split into train/test sets.
  - Train a regression model (Linear Regression is classic for beginners).
  - Evaluate predictions using metrics like Mean Squared Error (MSE).
- **Skills gained:** Handling tabular data, regression, feature engineering, basic metrics.