

# Task 11

## Machine Learning

Upload .py or Ipython extension file on GitHub public repo "100DaysofBytewise" and share the link in the submission form by 22 July 2024.

### 1. Predicting Diabetes Onset Using Logistic Regression

**Dataset:**

- Diabetes Dataset

**Preprocessing Steps:**

- Handle missing values if any.
- Standardize features.
- Encode categorical variables if any.

**Task:**

- Implement logistic regression to predict diabetes onset and evaluate the model using accuracy, precision, and recall.

### 2. Classifying Iris Species Using Decision Trees

**Dataset:**

- Iris Dataset

**Preprocessing Steps:**

- Handle missing values if any.
- Standardize features.
- Encode categorical variables if any.

**Task:**

- Implement a decision tree classifier to classify iris species and evaluate the model using confusion matrix and accuracy.

### 3. Predicting Titanic Survival Using Logistic Regression

**Dataset:**

- Titanic Dataset

**Preprocessing Steps:**

- Handle missing values (e.g., fill missing ages with median).
- Encode categorical variables (e.g., one-hot encoding for embarked and gender).
- Standardize numerical features.

**Task:**

- Implement logistic regression to predict survival on the Titanic and evaluate the model using ROC-AUC.

### 4. Classifying Emails as Spam Using Decision Trees

**Dataset:**

- Spam Email Dataset

**Preprocessing Steps:**

- Handle missing values if any.

- Standardize features.
- Encode categorical variables if present.

**Task:**

- Implement a decision tree classifier to classify emails as spam or not and evaluate the model using precision, recall, and F1-score.

## 5. Predicting Customer Satisfaction Using Logistic Regression

**Dataset:**

- Customer Satisfaction Dataset

**Preprocessing Steps:**

- Handle missing values (e.g., fill missing values with median).
- Encode categorical variables (e.g., one-hot encoding for region).
- Standardize numerical features.

**Task:**

- Implement logistic regression to predict customer satisfaction and evaluate the model using accuracy and confusion matrix.