

# ML Project

## Team Members

Aya Elsheikh 20220096

Judy Mohamed 20220127

Bassmala Hossam 20220108

Logyne Ammar 20220354

Mennatullah Haitham 20220508

Sara Emad 20220197

## Overview

---

This document outlines the details of our machine learning project focused on two different datasets: one for flower species classification and another for house price prediction.

This project implements various algorithms, including logistic, KNN (K-Nearest Neighbors) as classifiers, KNN and multiple linear as regressors.

# Datasets

---

## Flower Species Classification

- Source: [Flowers Recognition Dataset](#)
- Description: This dataset contains images of different flower species, specifically 5 classes: Daisy, Dandelion, Rose, Sunflower, and Tulip.
- Labels: The labels correspond to the species of the flowers.
- Missing Values: No missing values are present in the dataset, as it consists of images organized in directories by class.
- test size : 20%

## House Prices Prediction

- Source: [House Prices Regression Dataset](#)
- Description: This dataset includes various features related to houses, such as square footage, number of bedrooms, and location, aimed at predicting house prices.
- Labels: The target variable is the house price.
- Missing Values: The dataset contains some missing values in certain features, which were handled using imputation techniques.
- test size : 30%

# Applied Algorithms

---

## Flower Species Recognition

### Logistic Regression

- To classify the flower species based on extracted features from the images.
- Features were extracted using image processing techniques, and logistic regression was applied to predict the flower species.

### K-Nearest Neighbors (KNN)

- To classify the flower species using the KNN algorithm.
- The classifier was trained on the same feature set used for logistic regression, with hyperparameter tuning for optimal performance.

## House Price Prediction

Multiple Linear Regression

- This model was built using the relevant features, and the coefficients were estimated through ordinary least squares.

KNN Regression

- This model used the distance to the nearest neighbors to predict the target variable.

# House price prediction

Algorithm	MSE	MAE	R2
Multiple Linear Regression	67306371797.759	168580.418	0.5097
KNN	39544138199.646	96741.654	0.7384

# Flower species recognition

Algorithm	Accuracy	Loss curve
Logistic	57%	1.2860
KNN	48.50%	5.3872