

Computer Vision

LAB 5

1. In the attachment, Folder Q1 contains a dataset for fruit classification. This dataset includes a variety of features, but we are specifically interested in using mass, width, height, and color score (columns 4, 5, 6, and 7) to identify the type of fruit. There are four types of fruits in this dataset: orange, mandarin, apple, and lemon. Use the K-Nearest Neighbors (KNN) algorithm for this classification task. Try using odd values of k from 1 to 20 as the number of neighbors and determine the best performing value. Additionally, examine the effect of feature normalization on the model's performance. Analyze whether normalization has any impact on the results and report your experiments.
2. In Folder Q2, there is a dataset named data.csv, which contains two columns:
 - The first column represents the population of a city
 - The second column indicates the profitability of opening a shop in that city

Train a regression model to estimate the profitability based on the city population, aiming to minimize prediction error.

Evaluate the model performance using MAE (Mean Absolute Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error).

Note: Some population values are missing in the dataset. Make sure to handle these missing values appropriately during preprocessing.