

**Data Mining
Information Systems Department
Faculty of Computers and Information
Cairo University**



Assignment 3

Classification

Instructions:

1. Assignment should be done individually; copies will be graded to zero.
2. Your program should include a user friendly interface.
3. You must send your assignment on Monday 24 December to a.adel@fci-cu.edu.eg , e.osama@fci-cu.edu.eg .
4. Discussion will be held with Eng. Amany Adel & Eng. Eslam Osama at the week starting from 22 December.
5. No late submissions will be allowed.
6. Total grade is 5 marks.

- You will be given a dataset for **Car Evaluation**, The Car Evaluation Database contains examples with the structural information directly relates cars like (buying price, price of the maintenance, number of doors, capacity in terms of persons to carry, the size of luggage boot, estimated safety of the car) to evaluate the car level.
- For this assignment you are being asked to apply the Bayesian classifier that correctly recognizes evaluation level of a car based on a set of 6 features.
- The class label is car acceptability which is the last column in the provided comma separated file.

Requirements:

1. Divide the data set into 2 subsets, 1st one will be 75% of the data and call it "Training Set", 2nd set will be 25% of the data, and call it "Testing set"
2. Apply the Bayesian classifier, to build a classifier (model) from the first set "Training set".
3. Apply the Bayesian classifier you built in step-1 on the second set "Testing Set" to calculate the accuracy of the classifier.
4. Use the k-nearest algorithm to get the 5-nearest classes for each record in the "Testing Set", use majority voting to classify such records and calculate the accuracy of this classifier.
5. Compare the 2 classifiers Bayesian and K-Nearest.

Data Description:

Feature	Values Range	Description
Feature 1	vhigh, high, med, low.	Buying price
Feature 2	vhigh, high, med, low.	Maintenance Price
Feature 3	2, 3, 4, 5more.	Number of doors
Feature 4	2, 4, more.	Capacity in terms of persons to carry
Feature 5	small, med, big.	the size of luggage boot
Feature 6	low, med, high.	Estimated safety of the car
Class	unacc, acc, good, vgood	car acceptability