



K. N. Toosi University of Technology

In the name of God
Artificial Intelligence

Faculty of mechanical
engineering

Instructors:
Dr. Esmail Najafi
Javad Khoramdel

Assignment 1

Due date: 99/01/23

1. In Fuel Consumption dataset:

- a) Plot Cylinder vs the Emission, to see how linear their relation is.
- b) Try to use a multiple linear regression with the same dataset but this time use __FUEL CONSUMPTION in CITY__ and __FUEL CONSUMPTION in HWY__ instead of FUELCONSUMPTION_COMB. Does it result in better accuracy?
- c) Try to use a polynomial regression with the dataset but this time with degree three (cubic). Does it result in better accuracy?
- d) Try to use the following parameters for K in K-fold cross validation and evaluate the model by using mean absolute error, maximum absolute error and R2 score. Then plot each measure vs K.

K=2, 3, 5, 10, 20

2. In House Sales Prediction:

- a) Assume that due to high computational cost, we want to choose key features and neglect the others. In order to achieve that:

- Use a linear regression model and calculate accuracy for each feature in dataset.
- Use multiple linear regression model and each time, leave one feature out and train the model on other features.

Use R2 score metric for comparison and write the outputs in a table. Which parameter has more effect on accuracy?

- b) Apply polynomial regression of degree 3 and repeat the previous step.

Good Luck