Pattern Recognition and Machine Learning

Lab - 1 Assignment

Early Bird Submission Deadline: Jan 17, 6:00 PM
Late Submission Deadline: Jan 19, 2022, 12:00 Midnight (20% penalty)

Very Important: upload .py file and .ipynb link in the given google form otherwise your submission will not be considered.

Submit your RollNumber.py file link here: python file link
Submit your ipynb file link here: Colab link

Colab file for your reference: Demo link

Problem 1 (Basic of Python): Perform the following operations in python. (**10 Marks**)

a. Convert file data to list (**1 Marks**)
b. Convert User Input to a Number (**1 Marks**)
c. Convert String to Datetime in Python (**1 Marks**)
d. How to call external commands in Python? (**1 Marks**)
e. How to count the occurrences of a list item? (**1 Marks**)
f. How to flatten lists in Python? (**1 Marks**)
g. How to merge dictionaries in Python? (**1 Marks**)
h. Remove duplicate items from a list in Python? (**1 Marks**)
i. Write a Python script to check whether a given key already exists in a dictionary. (**2 Marks**)

Problem 2 (Numpy): Using numpy create two matrices of same size of your choice, fill the non-zero values into these two matrices. Now perform following: (**10 Marks)**

(a) Display first row of first matrix
(b) Display second column of second matrix
(c) Perform matrix multiplication
(d) Perform element-wise multiplication
(e) Perform dot product between each column of first matrix and each column of second matrix

Problem 3 (Pandas): A csv file has been provided to you at this link. The given dataset is related to cars and contains 26 columns.In the given dataset, “Price” is the target variable (i.e., the output). (**10 Marks**)

The marks distribution according to the tasks are as follows:

i) Assign a type to each of the following features (a) Model, (b)Type, (c) Max. Price and (d)Airbags from the following: ordinal/nominal/ratio/interval scale. (**1 Marks**)

ii) Write a function to handle the missing values in the dataset (e.g., any NA, NaN values). (**1 Marks**)
iii) Write a function to reduce noise (any error in the feature) in individual attributes (**2 marks**)
iv) Write a function to encode all the categorical features in the dataset according to the type of variable jointly. (**3 Marks**)
v) Write a function to normalize / scale the features either individually or jointly. (**1 Marks**)
vi) Write a function to create a random split of the data into train, validation and test sets in the ratio of [70:20:10]. (**2 Marks**)

Problem 4 (Plotting): Plot following functions: (**5 Marks**)

a. y = 5x + 4 where x ranges from [-10, 10].
b. y = ln(x) where x > 10 and x < 100.
c. y = x^2 where x ranges from [-10, 10].

x	0	1	2	3	4
y	2	3	4	5	6

Table 1: Data

Problem 5 (Evaluation Matrix): download colab file and dataset from given link. (**15 Marks**)
Some task we already performed on given data in same colab file, Now your task is to perform following operation from inbuilt and scratch:

1. Average Accuracy and Class-Wise Accuracy
2. Precision
3. Recall
4. F1-Score
5. Sensitivity
6. Specificity