

Pattern Recognition and Machine Learning 2022 Winter Semester

Lab - 1 Assignment

Early Bird Submission Deadline: Jan 23, 6:00 PM (Full point)
Late Submission Deadline: Jan 25, 2022, 12:00 Midnight (20% penalty)
Final deadline: Jan 27, 2022, 12:00 Midnight (10% penalty)

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

- Convert file data to list **(1 Marks)**
- Convert User Input to a Number **(1 Marks)**
- Convert String to Datetime in Python **(1 Marks)**
- How to call external commands in Python? **(1 Marks)**
- How to count the occurrences of a list item? **(1 Marks)**
- How to flatten lists in Python? **(1 Marks)**
- How to merge dictionaries in Python? **(1 Marks)**
- Remove duplicate items from a list in Python? **(1 Marks)**
- 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)**

- Display first row of first matrix
- Display second column of second matrix
- Perform matrix multiplication
- Perform element-wise multiplication
- 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). **(15 Marks)**

The marks distribution according to the tasks are as follows:

- 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)**
- Write a function to handle the missing values in the dataset (e.g., any NA, NaN values). **(2 Marks)**
- Write a function to reduce noise (any error in the feature) in individual attributes **(2 marks)**
- Write a function to encode all the categorical features in the dataset according to the type of variable jointly. **(5 Marks)**
- Write a function to normalize / scale the features either individually or jointly. **(2 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]. **(3 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](#). **(10 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