

## Machine Learning Lab (AI332L)

**Class:** BS Artificial Intelligence  
**Instructor:** Dr. M. Ehatisham-ul-Haq

**Semester:** 5<sup>th</sup> (Fall 2023)  
**Email:** ehtisham@mail.au.edu.pk

### | Lab 09 |

## Problem Solving in Machine Learning using Python

### Lab Objective:

In this lab tutorial, we will apply the techniques learned in the previous labs for problem-solving in Python.

### Setting up Python with Google Colab

1. Go to [Google Colab](https://colab.research.google.com/).
2. Create a new notebook by clicking on **File** > **New Notebook**.
3. You're now ready to start writing and executing Python code in the notebook!

### Lab Task(s):

**Task 01:** Consider a scenario where we are predicting a student's exam score ( $s$ ) based on the number of hours ( $h$ ) they studied and the temperature ( $t$ ) outside, as shown in the table.

- a) Is it possible to claim mathematically that higher outside temperatures impact a student's performance? If so, how, and what is the confidence level in this claim? Conversely, if we believe the claim is incorrect, what confidence level supports this view?
- b) Predict a student's exam score if the temperature is 40 and 45 degrees on two different days, given that the weight value relevant to the number of study hours is fixed as 2.

$h$	5	8	3	7	6	4	9	2	10	6
$t$	25	20	30	15	18	28	22	35	19	25
$s$	60	75	50	70	65	55	80	45	85	68

**Task 02:** The HR department of an institute wants to predict employee salary and eligibility for promotion based on factors like years of experience, job performance, and department. The data given in the table (on the right side) can be used to build predictive models for HR decision-making. However, it is important to preprocess the data before model building. Your task is to use your intuition and apply the most suitable data imputation, encoding, and scaling techniques (where applicable) on the given data to get it ready for model building.

Experience (in years)	Performance	Department	Salary	Promotion
5	Excellent	Sales	60000	Eligible
3	Good	Marketing	45000	Ineligible
?	Excellent	Finance	80000	Eligible
2	?	HR	35000	Ineligible
4	Good	?	55000	Ineligible
6	Excellent	Engineering	70000	Eligible
8	Excellent	Marketing	85000	Eligible
1	Poor	Finance	30000	Ineligible
9	Excellent	HR	90000	Eligible
5	Good	Engineering	60000	Ineligible

**Task 03:** Use the dataset below to learn and construct a machine learning Classifier that predicts if the students “Pass” or “Fail” machine learning course, based on their study hours per day, assignments grade (High, Medium, or Low), whether or not they attended the lectures with at least 50% attendance (T or F), and their “Satisfactory” or “Unsatisfactory” grade in the quizzes (S or U). Classify a random student with features (L, T, S) as “Pass” or “Fail” using the trained classifier.

Study Hours per Day	Assignments Grade	Lectures Attended	Quiz Grade	Status
06	L	F	U	Fail
10	L	T	S	Pass
07	M	F	U	Fail
12	M	T	S	Pass
09	H	F	U	Pass
10	H	T	S	Pass