DEPARTMENT OF COMPUTER & INFORMATION SYSTEMS ENGINEERING
BACHELORS IN COMPUTER SYSTEMS ENGINEERING

Course Code and Title: CS-324 Machine Learning
Complex Engineering Problem
TE Batch 2019, Spring Semester 2022

TERM PROJECT

Course Learning Outcome
CLO 2: Investigate techniques and principles for implementing machine learning systems.
(Taxonomy level C4, PLO-4 Investigation).

Complex problem-solving attributes (CPA) covered (as per PEC - OBA manual – 2019)
- **CPA-1 Depth of analysis required:** Have no obvious solution and require abstract thinking, originality in analysis to formulate suitable models.
- **CPA-2 Level of interaction:** Require resolution of significant problems arising from interactions between wide ranging or conflicting technical, engineering or other issues.
- **CPA-3 Familiarity:** Can extend beyond previous experiences by applying principles-based approaches.

Term Project Title:
Prediction of Cumulative Grade Point Average

Description
Cumulative Grade Point Average (CGPA) refers to the overall Grade Point Average (GPA), obtained by dividing the total Grade Points (GPs) earned in all courses attempted by the total degree-credit hours in all attempted courses.

- You are required to develop a machine learning system to predict final CGPA of a student at the end of fourth year given GPs of the courses obtained in initial years (up to first, second or third year).
- The dataset to be used is attached with this file with name The_Grades_Dataset.csv.
- You are required to develop and compare several models as per the following description:
  - Model 1: predict final CGPA based on GPs of first year only.
  - Model 2: predict final CGPA based on GPs of first two years.
  - Model 3: predict final CGPA based on GPs of first three years.

Minimum Required Features
1. Perform all necessary data preprocessing steps on the dataset.
2. Implement at least two models from the given models.
3. For each model, implement at least two different machine learning algorithms of your choice.
4. Also develop a simple interface to allow the user to test your system by providing data for any student and getting predicted final CGPA in return.
5. Prepare a report containing all the relevant details and analysis as described in the following section Deliverables.

- You can add any additional features to the project to make it distinct and stand out.
- Students can work in groups of 3 at maximum.
- Your assignment will be graded on the attached rubrics (see Google Classroom).

Deliverables

- Code on Jupyter notebook
  - Develop the code in a single Jupyter notebook.
  - Add a markdown cell at the top showing roll numbers and names of all group members.
  - Save the notebook with the following naming convention: CS19XXX, where XXX represents roll number of any one group member in 3 digits. For example: CS19032.py
  - Add comments generously.
- Prepare a report organized as follows:
  - Details of the data preprocessing steps applied.
  - Details of the models and machine algorithm chosen for implementation.
  - Details of any distinguishing features (if any).
  - Tabular or graphical comparison of all the models.
  - Comments on the performance of the implemented machine learning system, including issues like underfitting and overfitting, suggesting any techniques for improvement.
  - Attach the provided rubric on **top** of your report, with names and roll numbers filled in. This will serve as the title page of your report; no need to attach any other title page.
  - Save the report as pdf file, with the following naming convention: CS19XXX, where XXX represents roll number of any one group member in 3 digits. For example: CS19032.py.

Submission

- Online submission:
  - Submit both the files (notebook and report) with the specified naming convention on google classroom latest by the end of **Week 15**.
  - Any one of the group members can do the submission.
- Bring the notebook (softcopy) and report (hardcopy) during final practical exam in **Week 16**.