CS571 - ARTIFICIAL INTELLIGENCE LAB

ASSIGNMENT-7: Linear Regression

Date: September 18, 2023

Deadline: September 24, 2023

Total Credit: 30

- Markings will be based on the correctness and soundness of the outputs.
- Marks will be deducted in case of plagiarism.
- Proper indentation and appropriate comments are mandatory.
- *All code needs to be submitted in '.py' format.* Even if you code it in '.IPYNB' format, download it in '.py' format and then submit
- You should zip all the required files and name the zip file as:
    - <roll_no>_assignment_<#>.zip, eg. 1501cs11_assignment_01.zip.
- Upload your assignment (the zip file) in the following link: https://www.dropbox.com/request/y0ACUAAKE7cOudmDvFJO

Problem Statement:

The newspaper companies circulate special Sunday editions along with their daily editions. However, all these companies want to investigate the feasibility of these Sunday editions for the next week based on the sales from current week sales. The objective is to predict which newspaper companies should stop their Sunday edition if the predicted Sunday circulation should be at least more than 30% of their daily sales.

- For example if the no. of daily newspapers is 100 then that newspaper company can circulate their Sunday edition if the predicted no. of Sunday newspapers is more than 130
- Since the sale of daily newspapers for the next week is unknown, consider three cases; minimum, maximum and average no. of newspapers sold from the entire dataset.
- The task is to create a simple statistical model to predict the companies that should not circulate their Sunday edition

Dataset:
The dataset can be downloaded from the following link:
https://www.dropbox.com/scl/fi/f4gpd8shciyxf78oh6kej/newspaper_data.tsv?rlkey=njcbg1kfb2kbxqsk05lqry6n7&dl=0

Note: Code should be written from scratch and existing tools can not be used

Documents to submit:

- Model code
- Final parameter values of the linear regression model
- Names of companies who should not circulate their Sunday edition
- Plot the data and the predicted regression curve that fits the data

For any queries regarding this assignment, contact:

Ramakrishna Appicharla (ramakrishnaappicharla@gmail.com),
Arpan Phukan (arpanphukan@gmail.com) and,
Sandeep Kumar (sandeep.kumar82945@gmail.com)