|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| AGRAGANYA | | | |  | |
| Course Name | Higher Diploma | | | | |
| Module Name | Machine Learning | | | | |
| Batch | 1 | | | | |
| Learning Outcomes Covered (Mention according to the Module Descriptor) | LO1. How to use Jupyter Notebook and Python packages – Numpy, Matplotlib, Scikit-learn for data visualization and machine learning.  Program in Python using the latest Python.  LO4. Preprocess data, clean data, and analyze large data.  LO5. Training machine learning model. | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
|  | | | | |
| Assessment | CW No | 1 | | | | |
| Assessment Mode | **Individual | ~~Group~~** | | **Group (if it is group mode only)** | | |
| **Group Size** | | **Grouping Criteria** |
|  | | - |
| Assessment Type | **~~Practical Test | Report | Software | Presentation | VIVA | MCQ~~** | | | | |
| **If other specify** | **Python Code** | | | |
| Hand in Date | Time | 2025/03/04 @ 11.00 a.m | | | | |
| Hand out Date | Time | 2025/03/04@ 12.30 p.m | | | | |
| Submission Details (Format and Location) | Upload to the LMS on or before 2025/03/04 before 12.40 p.m | | | | |
| Plagiarism Criteria | Do not copy from Chat GPT | | | | |
| Assessment | CW Description  Follow the instructions and complete the program.  (This assessment will carry 10% of your total marks)  Login to LMS and download the folder containing the Assignment 1 with the relevant datasets.   1. Consider the Agriculture Dataset. (40 Marks) 2. Examine the dataset and identify the variable. Train an appropriate model for the data set. Use cross validation approach as the model evaluation technique. 3. Evaluate the performance of the linear regression model using appropriate metrics. 4. Include appropriate plotting in your model. 5. Obtain some predictions and check the accuracy. 6. Consider the ‘Fuel Consumption’ Dataset. (60 Marks) 7. Convert the Fuel Consumption column into a binary outcome by creating a new column ‘High\_Fuel’ where a sale is classified as 'High' if it exceeds the mean value of Fuel Consumption. Train a logistic regression model to predict ‘High\_Fuel’  * Evaluate the performance of the logistic regression model using appropriate metrics. * Include appropriate plotting in your model. * Obtain the predicted probabilities for class 1. * Obtain the confusion matrix and relevant calculations. * Obtain some predictions and check the accuracy. * Check whether the class is balance or not. If not conduct the logistic regression again after balancing classes. | | | | | |
|  | | | | | |
|  | | | | | |