

Due	Mar 6 by 23:59	Points	10	Submitting	a file upload	Attempts	0
Allowed Attempts	1	Available	Feb 15 at 0:00 - Mar 6 at 23:59 20 days				

Instructions for Assignment Evaluation

1. Please follow the naming convention as `<Group no>_<Dataset name>.ipynb`.

Eg – for group 1 with a weather dataset your notebooks should be named as -
Group1_WeatherDataset.ipynb.

2. Inside each jupyter notebook, you are required to mention your name, Group details and the Assignment dataset you will be working on.

3. Organize your code in separate sections for each task. Add comments to make the code readable.

4. Deep Learning Models are strictly not allowed. You are encouraged to learn classical Machine learning techniques and experience their behavior.

5. Notebooks without output shall not be considered for evaluation

Prepare a jupyter notebook (recommended - Google Colab) to build, train and evaluate a Machine Learning model on the given dataset. Please read the instructions carefully.

Dataset - <https://www.kaggle.com/karnikakapoor/rain-prediction-ann/data> 

1. Import Libraries/Dataset

- a. Download the dataset
- b. Import the required libraries

2. Data Visualization and Exploration

- a. Print at least 5 rows for sanity check to identify all the features present in the dataset and if the target matches with them.
- b. Print the description and shape of the dataset.
- c. Provide appropriate visualization to get an insight about the dataset.
- d. Try exploring the data and see what insights can be drawn from the dataset.

3. Data Pre-processing and cleaning

- a. Do the appropriate preprocessing of the data like identifying NULL or Missing Values if any, handling of outliers if present in the dataset, skewed data etc. Apply appropriate feature engineering techniques for them.
- b. Apply the feature transformation techniques like Standardization, Normalization, etc. You are free to apply the appropriate transformations depending upon the structure and the complexity of your dataset.
- c. Do the correlational analysis on the dataset. Provide a visualization for the same.

4. Data Preparation

- a. Do the final feature selection and extract them into Column X and the class label into Column into Y.
- b. Split the dataset into training and test sets.

5. Model Building

- a. Develop ANN Model
- b. Train the model and print the training accuracy and loss values.

6. Performance Evaluation

- a. Print the confusion matrix. Provide appropriate analysis for the same.
- b. Do the prediction for the test data and display the results for the inference.

