

**Department of Mathematical and Computational Sciences
National Institute of Technology Karnataka, Surathkal, Mangalore**

**First Semester MTech (CDS) 2021 – 2022
MA-722 FOUNDATIONS OF MACHINE LEARNING ALGORITHMS
PROJECT WORK**

Total Marks: 40

Instructions for Project work:

- Collect a dataset containing at least 200 records
- Dataset can be collected from any of the following sectors: Healthcare, Retail, Agriculture, Manufacturing, Sports, Insurance, Financial services etc. or from government datasets (data.gov.in, population census, COVID-19 etc.)
- Prepare a project report containing the following:
 - An explanation of the dataset
 - Identify a problem or goal (Predictive or Descriptive)
(For example: In case of retail industry, grouping or classifying the customers based on various demographic parameters, identifying customers for banking services, predicting the production rate, prediction for sports/ agricultural data, grouping the patients based on the type of disease OR find out which type of blood tests are common)
 - Apply suitable machine learning algorithms (one or many) for the dataset and depict the results
 - Explain the outcome of the algorithm with the decision / observation identified for the defined problem/goal
 - Can apply multiple cases on the algorithms (for example: Test using different values of k in case of knn algorithm, different support and confidence values for apriori algorithm) for refining the results
 - Explain the results using plots, confusion matrix, accuracy, misclassification error.
- Submit the pdf of the project report along with the dataset files
- The project report directly downloaded from any internet source will lead to rejection of the project work
- Repetition of projects with the colleagues will lead to rejection of the project work
- Submit the report in moodle on or before 10th January, 2022.
