



**Knowledge Network of
Indian Institute of Technology Gandhinagar
Under TEQIP-II Initiative**

**Summer School on
Data Mining and Machine Learning**

The summer course on “Data Mining and Machine Learning” will give an introduction to participants about the fundamental principles involved in mining patterns from data, from both the statistical and algorithmic viewpoints. The course will be taught in an activity oriented manner, with significant hands-on exercises involving the use of machine learning libraries to analyze data. The course will refresh the necessary mathematical background, assuming only a knowledge of basic probability and linear algebra, as well as basic programming skills (familiarity with Matlab or python is fine). This school will provide faculty members and students interested in these topics a chance to learn about the concepts behind these topics, acting either as their introduction to this, or as a up to date review. Participation in this summer school is invited through registration.

Topics to be covered

Review and introduction: linear algebra and probability review, models and cost functions; *regression:* linear regression, logistic regression, overfitting, regularization; *supervised methods:* decision trees, naive Bayes, kNN, support vector machines, bagging, random forest; *unsupervised methods:* clustering, dimensionality reduction; frequent patterns: association rules; *machine learning system design:* evaluation, bias-variance tradeoff, handling skewed data, building a spam classifier; *applications:* machine learning in vision, search engines, recommendation etc.

The following format will be followed in the summer school:

The day comprises of 2 hrs of Instruction + 2 hrs of hands-on activities and problem solving

Presenters: Instructors:

1. Prof. Anirban Dasgupta
 2. Prof. Dinesh Garg
 3. Prof. Kamal Karlapalem
 4. Prof. Shanmuganathan Raman
-