



# DATA MINING (Week 5)

# DSBA CURRICULUM DESIGN

## FOUNDATIONS

Data Science Using  
Python

Statistical Methods  
for Decision Making

## CORE COURSES

Advanced  
Statistics

Data Mining  
(Week-5/6)

Predictive Modelling

Machine Learning

Time Series  
Forecasting

Data Visualization

## DOMAIN APPLICATIONS

Financial Risk  
Analytics

Web & Social Media  
Analytics

Marketing Retail  
Analytics



# LEARNING OBJECTIVE OF THIS MODULE

- Clustering
- CART & Model Performance Measures
- Random Forest
- Neural Network

# LEARNING OBJECTIVES OF THIS SESSION

- CART
- RANDOM FOREST
- NEURAL NETWORKS
- COMPARISON BETWEEN THE ACCURACY OBTAINED FROM DIFFERENT MODELS

## CASE STUDY - Diabetes

Management of hyperglycemia in hospitalized patients has a significant bearing on outcome, in terms of both morbidity and mortality. However, there are few national assessments of diabetes care during hospitalization which could serve as a baseline for change. This analysis of a large clinical database was undertaken to provide such an assessment and to find future directions which might lead to improvements in patient safety.

Based on the given data, build different models to predict whether the patient is suffering from Diabetes or not? Secondly Compare the model outputs of different models.



**ANY QUESTIONS**





**HAPPY LEARNING**