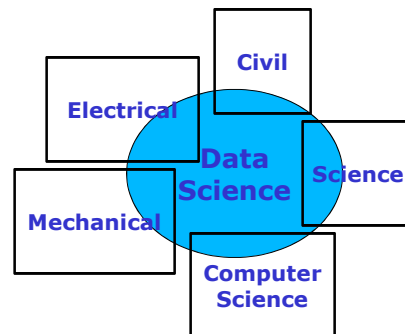# Data Science – III
(Data Processing and Modeling)

Data Science

- Multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insight from structured and unstructured data

- Data is available in every discipline

Civil

Electrical

Data Science

Science

Mechanical
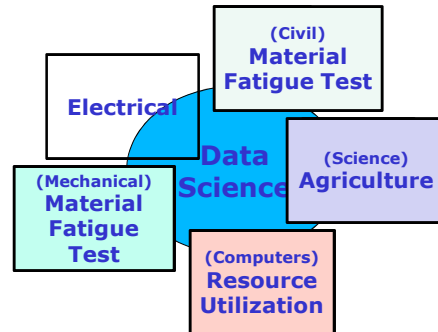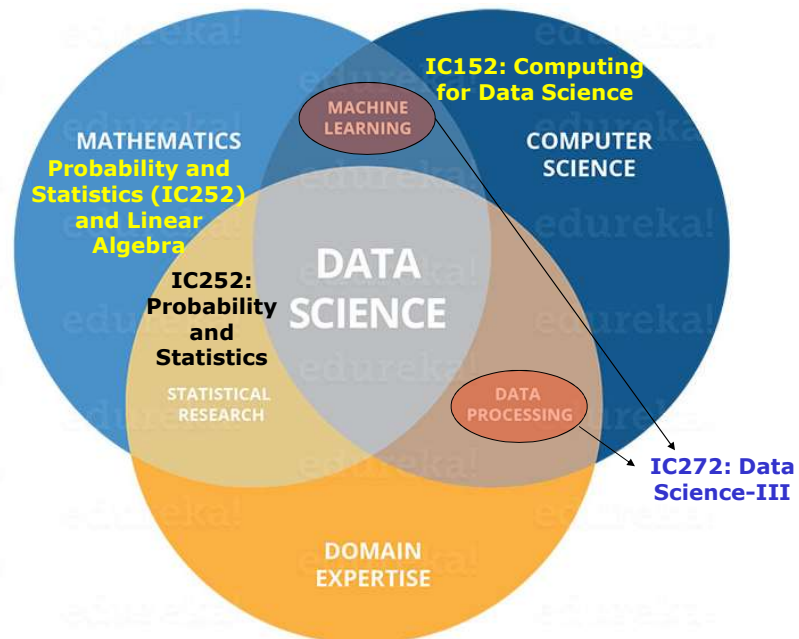
Computer Science

2

Data Science

- Multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insight from structured and unstructured data
- Data is available in every discipline

- Concept to unify statistics, data analysis, machine learning and their related methods to understand and analyse the actual phenomena with data
  - Includes theories from mathematics, statistics and computer science

Electrical

(Civil) Material Fatigue Test

Data Science

(Mechanical) Material Fatigue Test

(Science) Agriculture

(Computers) Resource Utilization

3

Data Science

MATHEMATICS
Probability and Statistics (IC252) and Linear Algebra

MACHINE LEARNING

IC152: Computing for Data Science

COMPUTER SCIENCE

IC252: Probability and Statistics

DATA SCIENCE

STATISTICAL RESEARCH

DATA PROCESSING

IC272: Data Science-III

DOMAIN EXPERTISE

4

Data Science – III: Big Picture

Data Collection

Data Modeling
(Machine
Learning) → Inference

Data Preprocessing

Database →

| Data Cleaning and Cleansing | → | Feature Representation |

5