

# SYE8012: DATA MINING AND STATISTICAL MODELING

---

## Effective Term

Semester A 2025/26

## Part I Course Overview

### Course Title

Data Mining and Statistical Modeling

### Subject Code

SYE - Systems Engineering

### Course Number

8012

### Academic Unit

Systems Engineering (SYE)

### College/School

College of Engineering (EG)

### Course Duration

One Semester

### Credit Units

3

### Level

R8 - Research Degree

### Medium of Instruction

English

### Medium of Assessment

English

### Prerequisites

Nil

### Precursors

Basic Probability and Statistics

### Equivalent Courses

SEEM8012 Data Mining and Statistical Modeling(offered until 2021/22)/ ADSE8012 Data Mining and Statistical Modeling (offered until 2023/24)

### Exclusive Courses

Nil

## Part II Course Details

### Abstract

This course focuses on data mining tools and techniques that are useful for a wide range of applications in manufacturing, service, logistics, health and medical, financial and banking, etc. We discuss four basic data mining operation steps: business objective identification, data preparation, knowledge discovery, and consolidation/implementation. We cover both supervised learning and unsupervised learning methods and algorithms, including regression, classification, forecasting, clustering, association rules, and market basket analysis etc. The methods will be illustrated with case studies in credit card fraud detection, telecommunication, express mail service, inventory management, customer relationship management, and bioinformatics.

### Course Intended Learning Outcomes (CILOs)

| CILOs |                                                                                                                         | Weighting (if app.) | DEC-A1 | DEC-A2 | DEC-A3 |
|-------|-------------------------------------------------------------------------------------------------------------------------|---------------------|--------|--------|--------|
| 1     | Recognize basic statistical learning, data mining, machine learning, and knowledge discovery and potential applications | 15                  | x      |        |        |
| 2     | Familiarize the operational steps on data mining and knowledge discovery                                                | 15                  | x      |        |        |
| 3     | Recognize and apply supervised learning methods and algorithms and their applications.                                  | 20                  | x      |        |        |
| 4     | Recognize and apply unsupervised learning methods and algorithms and their applications.                                | 20                  | x      |        |        |
| 5     | Demonstrate how data mining methods and algorithms can be applied to real life problems in various applications         | 30                  | x      | x      |        |

#### A1: Attitude

Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

#### A2: Ability

Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

#### A3: Accomplishments

Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

### Learning and Teaching Activities (LTAs)

| LTAs |         | Brief Description                                    | CILO No.      | Hours/week (if applicable) |
|------|---------|------------------------------------------------------|---------------|----------------------------|
| 1    | Lecture | - large class activity<br>- questions and discussion | 1, 2, 3, 4, 5 | 39 hours/sem               |

### Assessment Tasks / Activities (ATs)

|   | ATs                   | CILO No.      | Weighting (%) | Remarks ("- for nil entry) | Allow Use of GenAI? |
|---|-----------------------|---------------|---------------|----------------------------|---------------------|
| 1 | Group Work            | 1, 2, 3, 4, 5 | 40            | -                          | No                  |
| 2 | Individual Coursework | 1, 2, 3, 4    | 25            | -                          | No                  |
| 3 | Test                  | 2, 3, 4, 5    | 35            | -                          | No                  |

**Continuous Assessment (%)**

100

**Examination (%)**

0

**Minimum Continuous Assessment Passing Requirement (%)**

30

**Minimum Examination Passing Requirement (%)**

0

**Assessment Rubrics (AR)****Assessment Task**

Group Work (for students admitted before Semester A 2022/23 and in Semester A 2024/25 &amp; thereafter)

**Criterion**

Application of class materials and teamwork

**Excellent**

(A+, A, A-) High

**Good**

(B+, B, B-) Significant

**Fair**

(C+, C, C-) Moderate

**Marginal**

(D) Basic

**Failure**

(F) Not even reaching marginal levels

**Assessment Task**

Individual Coursework (for students admitted before Semester A 2022/23 and in Semester A 2024/25 &amp; thereafter)

**Criterion**

Application of class materials

**Excellent**

(A+, A, A-) High

**Good**

(B+, B, B-) Significant

**Fair**

(C+, C, C-) Moderate

**Marginal**

(D) Basic

**Failure**

(F) Not even reaching marginal levels

---

**Assessment Task**

Test (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

**Criterion**

Understanding of class materials

**Excellent**

(A+, A, A-) High

**Good**

(B+, B, B-) Significant

**Fair**

(C+, C, C-) Moderate

**Marginal**

(D) Basic

**Failure**

(F) Not even reaching marginal levels

---

**Assessment Task**

Group Work (for students admitted from Semester A 2022/23 to Summer Term 2024)

**Criterion**

Application of class materials and teamwork

**Excellent**

(A+, A, A-) Excellent

**Good**

(B+, B) Good

**Marginal**

(B-, C+, C) Marginal

**Failure**

(F) Failure

---

**Assessment Task**

Individual Coursework (for students admitted from Semester A 2022/23 to Summer Term 2024)

**Criterion**

Application of class materials

**Excellent**

(A+, A, A-) Excellent

**Good**

(B+, B) Good

**Marginal**

(B-, C+, C) Marginal

**Failure**

(F) Failure

---

**Assessment Task**

Test (for students admitted from Semester A 2022/23 to Summer Term 2024)

**Criterion**

Understanding of class materials

**Excellent**

(A+, A, A-) Excellent

**Good**

(B+, B) Good

**Marginal**

(B-, C+, C) Marginal

**Failure**

(F) Failure

---

## Part III Other Information

**Keyword Syllabus**

- Introduction to Data Mining
- Data Processing and Data Preparation
- Supervised Learning Methods
- Linear Methods for Prediction/Regression
- Linear Methods for Classification
- Model Assessment and Inferences
- Tree Models and Related Methods
- Neural Networks and SVM
- Forecasting and Time Series Modeling
- Unsupervised Learning Methods
- Clustering and Association Methods
- Data Mining Case Studies

**Reading List**

**Compulsory Readings**

| Title |                                                                                    |
|-------|------------------------------------------------------------------------------------|
| 1     | The Elements of Statistical Learning by Hastie, Tibshirani, and Friedman, Springer |
| 2     | Lecture notes                                                                      |

**Additional Readings**

| Title |     |
|-------|-----|
| 1     | NIL |