SEEM4113: DATA MINING

Effective Term
Summer Term 2023

Part I Course Overview

Course Title
Data Mining

Subject Code
SEEM - Systems Engineering and Engineering Management
Course Number
4113

Academic Unit
Systems Engineering (SYE)

College/School
College of Engineering (EG)

Course Duration
One Semester

Credit Units
3

Level
B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
MA2172 Applied Statistics for Science and Engineering or equivalent course

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract
This course aims at providing students the knowledge of classical data mining algorithms including algorithms in clustering and classification as well as training students to interpret and communicate data mining results. Students will learn to use

software packages such as R, Weka and Statistica to apply data mining algorithms, develop data-driven models, and conduct analyses.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Understand the knowledge of data and methods in data pre-processing	20	x	x	
2	Familiarize classical data mining algorithms in clustering and classification	30	x	x	
3	Apply software packages such as R, Weka and Statistica to apply data mining algorithms to conduct analyses	30	x	x	
4	Demonstrate the ability of interpreting data mining analyses results	20	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.

Teaching and Learning Activities (TLAs)

	TLAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lectures	Lectures	1, 2, 3, 4	26 hours/semester
2	Group Activities	In-class exercises, in-class Q&A, class projects and discussions	1, 2, 3, 4	13 hours/semester
3	Consultation hour	1 hour per week will be scheduled as consultation hour for clearing doubts of students who can meet the teaching staff on an individual or small group basis in his/her office.	1, 2, 3, 4	13 hours/semester

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Assignments, projects, testStudents will be assessed their understanding of concepts and techniques learned in class, reading materials and their ability to apply these concepts, techniques and subject-related knowledge.	1, 2, 3, 4	50	

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Assessment Rubrics (AR)

Assessment Task

Coursework

Criterion

Tutorial exercises, assignments and project, test

Excellent (A+, A, A-)

Strong evidence of capacity to analyse and synthesize; superior grasp of subject matter.

Good (B+, B, B-)

Evidence of grasp of subject, some evidence of critical capacity and analytic ability.

Fair (C+, C, C-)

Student who is profiting from the university experience; understanding of the subject; ability to develop solutions to simple problems in the material.

Marginal (D)

Sufficient familiarity with the subject matter to enable the student to progress without repeating the course.

Failure (F)

Little evidence of familiarity with the subject matter; weakness in critical and analytic skills.

Assessment Task

Examination

Criterion

2-hr final examination (either open or closed book based on instructor’ s discretion)

Excellent (A+, A, A-)

Strong evidence of capacity to analyse and synthesize; superior grasp of subject matter.

Good (B+, B, B-)

Evidence of grasp of subject, some evidence of critical capacity and analytic ability.

Fair (C+, C, C-)

Student who is profiting from the university experience; understanding of the subject; ability to develop solutions to simple problems in the material.

Marginal (D)

Sufficient familiarity with the subject matter to enable the student to progress without repeating the course.

Failure (F)

Little evidence of familiarity with the subject matter; weakness in critical and analytic skills.

Part III Other Information

Keyword Syllabus

Data mining. Data. Data types. Clustering. Classification. K-means. Hierarchical clustering. Association Rules. Classification and Regression Trees. Entropy. Information Gain. K-nearest neighbours. Support Vector Machine. Neural Networks. Software packages such as R, Weka and Statistica.

Reading List

Compulsory Readings

	Title
1	Lecture notes and slides provided by the instructor.

Additional Readings

	Title
1	P. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. Addison-Wesley.