Part I   Course Overview

Course Title:	Data Mining and Statistical Modeling
Course Code:	SEEM8012
Course Duration:	One Semester
Credit Units:	3
Level:	R8
Proposed Area: *(for GE courses only)*	☐ Arts and Humanities ☐ Study of Societies, Social and Business Organisations ☐ Science and Technology
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	Nil
Precursors: *(Course Code and Title)*	Basic Probability and Statistics
Equivalent Courses: *(Course Code and Title)*	MEEM8012 Data Mining and Statistical Modeling
Exclusive Courses: *(Course Code and Title)*	Nil

Part II   Course Details

1. **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 fault detection, telecommunication, express mail service, inventory management, customer relationship management, and bioinformatics.

2. **Course Intended Learning Outcomes (CILOs)**

No.	CILOs#	Weighting* (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Recognize basic statistical learning, data mining, machine learning, and knowledge discovery and potential applications	15%	✓		
2.	Familiarize the operational steps on data mining and knowledge discovery	15%	✓		
3.	Recognize and apply supervised learning methods and algorithms and their applications.	20%	✓		
4.	Recognize and apply unsupervised learning methods and algorithms and their applications.	20%	✓		
5.	Demonstrate how data mining methods and algorithms can be applied to real life problems in various applications	30%	✓	✓	
		100%			

* If weighting is assigned to CILOs, they should add up to 100%.
 # Please specify the alignment of CILOs to the Gateway Education Programme Intended Learning outcomes (PILOs) in Section A of Annex.

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 self-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.

3. Teaching and Learning Activities (TLAs)

TLA	Brief Description	CILO No.					Hours/week (if applicable)
		1	2	3	4	5	
Lecture	- large class activity - questions and discussion	✓	✓	✓	✓	✓	39 hours/sem

4. Assessment Tasks/Activities (ATs)

Assessment Tasks/Activities	CILO No.					Weighting*	Remarks
	1	2	3	4	5		
Continuous Assessment: __100__ %							
Group Work	✓	✓	✓	✓	✓	40%	
Individual Coursework	✓	✓	✓	✓		25%	
Test		✓	✓	✓	✓	35%	
Examination: __0__ % (duration: , if applicable)							

The weightings should add up to 100%. | 100% |

5. Assessment Rubrics

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Group Work	Application of class materials and teamwork	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Individual Coursework	Application of class materials	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Test	Understanding of class materials	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III   Other Information (more details can be provided separately in the teaching plan)

1.   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

2.   Reading List
2.1   Compulsory Readings

1.	The Elements of Statistical Learning by Hastie, Tibshirani, and Friedman, Springer
2.	Lecture note

2.2   Additional Readings

   NIL