MS4224: ENTERPRISE DATA MINING

Effective Term
Semester B 2024/25

Part I Course Overview

Course Title
Enterprise Data Mining

Subject Code
MS - Department of Decision Analytics and Operations
Course Number
4224

Academic Unit
Department of Decision Analytics and Operations (DAOS)

College/School
College of Business (CB)

Course Duration
One Semester

Credit Units
3

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

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
MS3252 Regression Analysis

Precursors
CB2200 Business Statistics or equivalent
MS3251 Analytics Using SAS or equivalent

Equivalent Courses
Nil

Exclusive Courses
MS4424 Data Mining and Modelling

Part II Course Details

Abstract

This undergraduate course introduces students to the fundamental concepts and techniques of data mining in business applications. It focuses on developing analytical skills to identify, formalize, and solve real-world problems using business intelligence. Students will learn to extract valuable insights from large datasets, make data-driven decisions, and apply various data mining algorithms to business scenarios. The curriculum covers key topics such as data preprocessing, classification, clustering, and association rule mining. Through hands-on projects and case studies, students will gain practical experience in applying these techniques to real business challenges. By the end of the course, participants will be well-prepared for positions in data modeling and business analytics in both commercial and government sectors, equipping them with skills applicable in local and global environments.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Discuss the history, nature role and relevant concepts of data mining.	10		x	
2	Evaluate a wide range of emerging and newly-adopted methodologies and technologies to facilitate the knowledge discovery.	40		x	x
3	Discuss the pros and cons of various processes, methodologies in knowledge discovery critically.	20		x	
4	Perform data mining tasks proficiently through using SAS Enterprise Miner software	30	x	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	Interactive Lecture	Students will actively participate in the interactive lectures to develop a comprehensive understanding of the fundamental concepts and method of data mining.	1, 2, 3	
2	Tutorial	Students will apply the learned concepts, techniques and SAS Enterpriser Miner skills on exercise questions	1, 2, 3, 4	

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Mid-term Test The test will assess the students' basic understanding of the material learnt in the first half of the course.	1, 2, 3	20	
2	Group Project Group projects on selected topics to enhance students' ability to model real-world problem and apply relevant data mining tools.	1, 2, 3, 4	20	
3	Individual presentation and Q&A Students will deliver presentations on selected topics; Students are required to answer the data mining questions during the Q&A session	1, 2, 3, 4	20	

Continuous Assessment (%)

60

Examination (%)

40

Examination Duration (Hours)

3

Additional Information for ATs

Written Examination The exam will assess the students' understanding of the material learnt in the course and their ability to apply subject related knowledge.

Assessment Rubrics (AR)

Assessment Task

Mid-term Test

Criterion

1.1 ABILITY to EXPLAIN the key concepts and fundamental knowledge of data mining

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 project

Criterion
2.1 CAPACITY for COLLABORATING with students to carry out problem-based activities based on real world problems.

2.2 ABILITY to EXPLAIN in DETAIL and with ACCURACY methods in analysing the relationship between business and sustainability solutions.

2.3 CAPACITY for SELF-DIRECTED LEARNING to find solutions to the problems and make recommendations for implementing the solutions

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 presentation and Q&A

Criterion
3.1 ABILITY to UNDERSTAND the knowledge of big data and social network analysis

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
Examination

Criterion
4.1 ABILITY to EXPLAIN the key concepts and fundamental knowledge of data mining

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

Part III Other Information

Keyword Syllabus

Data Preparation for Analysis
Summary Statistics; Data Visualization; OLAP and Multidimensional Data Analysis; Data Warehouse; Star Schema; Analysis Subject; Single view of the customer; Subject-oriented data; Data preparation for data mining; Dimension reduction; Integrity of data; Statistical ethics; Responsibilities to data; Potential misconduct;

Concepts of Data mining - SEMMA processes of Data mining; Supervised and unsupervised Statistical Learning;

Predictive Modelling - Logistic Regression; Decision Tree; Artificial Intelligence; Neural Network; Ordinal and Multinomial Logit Model; Tree-based Regression model; Learning Algorithm;

Pattern Recognition/Customer Behaviour Recognition - Similarity Measures; Clustering Analysis, RFM Analysis; Bayesian statistics, Association Analysis; Market Basket Analysis; Apriori Algorithm; Multi-level association rules; Sequential Pattern Mining;

Model validation - Goodness of fit; Model tuning; Model assessment and implementation

Business Analytics/Intelligence - Knowledge Discovery, Credit Scoring, Credit model development; Reject Inference;

Reading List

Compulsory Readings

	Title
1	Tan, P.N., Steinbach, M. and Kumar, V., Introduction to Data Mining. Pearson, 2014.
2	Thomas, L., Crook, J. and Edelman, D., Credit Scoring and its Applications, Second Edition, 2017. SIAM
3	Forster Provost and Tom Fawcett, 2013. Data Science for Business. O' Reilly Media, Inc.

Additional Readings

	Title
1	Paolo Giudici, Applied Data Mining: Statistical Methods for Business and Industry, John Wiley & Sons, 2003
2	Matignon, Randall. Data Mining Using SAS Enterprise Miner. Second Edition. Wiley, 2007
3	Cerrito, Patricia, Introduction to Data Mining Using SAS Enterprise Miner. SAS Institute, 2007
4	Michael Berry, & Gordon Linoff, Data mining techniques: For marketing, sales, and customer support, John Wiley & Sons, 2004
5	Patricia B. Cerrito, Introduction to Data Mining Using SAS Enterprise Miner, SAS Institute, 2006.
6	Michael Berry, & Gordon Linoff, Mastering Data Mining, John Wiley & Sons, 2000. Jiawei Han, & Micheline Kamber, Data mining: Concepts and techniques, Morgan Kaufmann Pub., 2000
7	Bart Baesens, Analytics in a BIG DATA WORLD – The essential guide to data science and its applications. WILEY, 2014
8	Bart Baesens, Credit Risk Modeling Using SAS, SAS Institute, 2011