MS4424: ADVANCED PREDICTIVE ANALYTICS

Effective Term
Semester A 2022/23

Part I Course Overview

Course Title
Advanced Predictive Analytics

Subject Code
MS - Management Sciences
Course Number
4424

Academic Unit
Management Sciences (MS)

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
MS3251 Analytics Using SAS or IS2240 Python Programming for Business

Equivalent Courses
Nil

Exclusive Courses
MS4224 Enterprise Data Mining

Part II Course Details

Abstract

Data mining plays a very important role in business – sales, marketing, and customer support. It is being used to discover implicit and useful knowledge from vast datasets. The course covers concepts fundamental to the understanding and applications of advanced data mining methods to business problems.

In recent years, there is an explosive growth of data mining techniques, applications and computer capabilities. This course focuses on advanced data mining techniques and applications utilizing computing software such as R, Python or SAS. It is an extension of MS4224 Enterprise Data Mining and MS4252 Big Data Analytics covering data mining techniques used in predictive analytics. According to SAS Institute (the leader in business analytics software and services, and the largest independent vendor in the business intelligence market.), predictive analytics is "the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data. The goal is to go beyond knowing what has happened to providing a best assessment of what will happen in the future." (https://www.sas.com/en_hk/insights/analytics/predictive-analytics.html). It is thought that the topics of this course together with the topics of other courses in the Business Analysis major provide students a comprehensive training in business analysis, and give them a competitive edge.

This course is different from MS4212 Predictive Analytics and Forecasting in that it analyses more granular data without a natural temporal ordering whereas MS4212 analyses time series data with a natural temporal ordering. Students are expected to have some programming language knowledge which is either learned from pre-cursor courses or by themselves.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Define and formulate real-world problems in scientific terms and plan for a rational and knowledge-based solution using advanced data mining techniques;	10	x	x	x
2	Collect, integrate and manage the relevant data and information for analysis;	10	x	x	x
3	Illustrate and apply advanced data mining techniques with the use of appropriate computing software	40	x	x	x
4	Evaluate critically the use of advanced data mining techniques in current studies and assess their appropriateness, accuracy and limitations in solving real-world problems;	30	x	x	x
5	Present effectively (oral, written) the methods, results, interpretations and strategies of the project to real-world problems.	10	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.

Teaching and Learning Activities (TLAs)

	TLAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lectures	Explain the concepts and knowledge in advanced data mining	1, 2, 3, 4, 5	
2	Computer Lab	Hands-on experience with the techniques and problem solving activities based on real world datasets using computer software such as R, Python or SAS. During the laboratory sessions, the instructor can identify problems encountered by students and provide assistance. The laboratory sessions consolidate and supplement what the students learn in lectures. The students will also have opportunities to work together and help each other.	1, 2, 3, 4, 5	
3	Final Project	Students work on a final project to solve real world problems using advanced techniques learnt in the course.	1, 2, 3, 4, 5	

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Assignment Assignment is designed to assess students' understanding of the concepts and techniques learnt in the course.	1, 2, 3, 4	10	The exact weighting for each component shall be determined by the Course Leader and to be announced to students at the beginning of the course.
2	Group project Students work in groups to apply advanced predictive analytic techniques to solve business problems and give verbal and written presentation of the problem and the findings.	1, 2, 3, 4, 5	30	

3	Mid-term Test A mid-term test is given to assess students' professional knowledge of the concepts, the techniques and the applications they have learned in the past weeks.	1, 2, 3, 4	20	

Continuous Assessment (%)

60

Examination (%)

40

Examination Duration (Hours)

3

Additional Information for ATs

Examination
The examination is designed to assess students' professional knowledge of the concepts, the techniques and the applications they have learned in the whole course.

Assessment Rubrics (AR)

Assessment Task

Assignment

Criterion

Assignment is designed to assess students' understanding of the concepts and techniques learnt in the course.

Excellent (A+, A, A-)

Strong evidence of knowing how to apply the key concepts and techniques.

Good (B+, B, B-)

Evidence of knowing how to apply the key concepts and techniques.

Fair (C+, C, C-)

Some evidence of knowing how to apply the key concepts and techniques.

Marginal (D)

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

Failure (F)

Little or no evidence of familiarity with the subject matter.

Assessment Task

Group Project

Criterion

Students work in teams to apply advanced techniques to solve business problems and give verbal and written presentation of the problem and the findings.

Excellent (A+, A, A-)

Strong evidence of original thinking; good organization, capacity to analyse and synthesize; superior grasp of subject matter; evidence of extensive knowledge base.

Good (B+, B, B-)

Evidence of grasp of subject, some evidence of critical capacity and analytic ability; reasonable understanding of issues; evidence of familiarity with literature.

Fair (C+, C, C-)

Some evidence of grasp of subject, little evidence of critical capacity and analytic ability; reasonable understanding of issues.

Marginal (D)

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

Failure (F)

Little evidence of familiarity with the subject matter; weakness in critical and analytic skills; limited or irrelevant use of literature

---

Assessment Task

Mid-term Test

Criterion

A test is given to assess students' professional knowledge of the concepts, the techniques and the applications they have learned in the past weeks.

Excellent (A+, A, A-)

Strong evidence of showing familiarity with the subject matter and analytic ability

Good (B+, B, B-)

Good evidence of showing familiarity with the subject matter and analytic ability

Fair (C+, C, C-)

Fair evidence of showing familiarity with the subject matter and analytic ability .

Marginal (D)

Marginal evidence of showing familiarity with the subject matter and analytic ability

Failure (F)

Not even reaching marginal levels

Assessment Task

Examination

Criterion

The examination is designed to assess students' professional knowledge of the concepts, the techniques and the applications they have learned in the whole course

Excellent (A+, A, A-)

Strong evidence of original thinking; good organization, capacity to analyse and synthesize; superior grasp of subject matter; evidence of extensive knowledge base.

Good (B+, B, B-)

Evidence of grasp of subject, some evidence of critical capacity and analytic ability; reasonable understanding of issues; evidence of familiarity with literature.

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; limited or irrelevant use of literature.

Part III Other Information

Keyword Syllabus

Introduction to Advanced Data Mining
What is advanced data mining? What can advanced data mining do? Advanced Data mining terminology.

Data Mining Process
Data preparation. Defining a study. Reading the data & Building a model. Understanding your model. Prediction. Overview of computer software (R, Python or SAS).

Advanced Data Mining Tools and Technologies
Memory-Based Reasoning. Automatic Cluster Detection. Decision trees and random forests. Neural Nets. Genetic Algorithms. Link Analysis. Statistical foundations of Naïve Bayes classification and support vector machine.

Lasso regression, ridge regression, and elastic net regularization.
Predicting behaviour, lifetime and performance: FRM analysis, lifetime value model, A/B testing. Statistical foundations of recommendation systems: Content-based Filtering, Neighborhood-based Collaborative Filtering, Model-based Collaborative Filtering.

Reading List

Compulsory Readings

	Title
1	Leventhal B, Predictive Analytics for Marketers : Using Data Mining for Business Advantage, Kogan Page, 2018.
2	Strickland, Jeffrey S. Predictive Analytics Using R, Lulu Inc, 2014.
3	Forte, R M. Mastering predictive analytics with R : master the craft of predictive modeling by developing strategy, intuition, and a solid foundation in essential concepts, Packt Publishing, 2015.

Additional Readings

	Title
1	Michael Berry, Gordon Linoff, Data mining techniques: For Marketing, Sales, and Customer Relationship Management, 2nd Edition, John Wiley & Sons, 2011.
2	Jiawei Han and Micheline Kamber, Data mining: Concepts and techniques, 3/e, Elsevier Publishers, 2012.

3	Miller, Thomas W. Marketing Data Science: Modelling Techniques in Predictive Analytics with R and Python, Pearson Education, 2015.
4	Xu G, Zong Y, Yang Z, Applied Data Mining, CRC Press, 2013.