STAMFORD UNIVERSITY BANGLADESH
Faculty of Engineering
Department of Computer Science and Engineering

Course Outline

Part A: Course Information

Program	:	Master of Science in Computer Science and Engineering (MCE)
Course Code	:	MCSE 642
Course Title	:	Data Mining
Course Type	:	Core Course
Credit Value	:	3.0 Cr. Hr.
Contact Hours	:	
Year/Level/Semester/Term	:	Summer 2022 Semester
Academic Session	:	Summer 2022
Prerequisite Course	:	None
Total Marks	:	100

Course Teacher	:	X
Class Schedule	:	X
Counseling Hour	:	X

Rationale of the Course :

This subject provides a theoretical and minorly practical introduction to knowledge discovery and data mining. A knowledge discovery process includes data preprocessing; machine learning techniques to learn hidden knowledge/patterns of data; and reporting & visualization of the resulting knowledge. The course will cover relevant topics of these subject area, which would enable students to gain practical knowledge to train a machine to learn, analyze and represent the knowledge from data.

Course Objectives :

This course is intended to help students to:

1. Understand basic applications, concepts, and techniques of data mining and knowledge discovery.
2. Learn different data pre-processing techniques, e.g., data selection, data summarization, and data cleaning.
3. Learn different machine learning algorithms, like classification, clustering etc. to extract hidden knowledge from data.
4. Report and visualize the extracted knowledge.
5. Generate and analyse results based on recent performance evaluation techniques.
6. Solve real world data mining problems using available tools and software.

Course Learning Outcomes :

Students who complete *Data Mining* course will be able to:

1. Understand and remember basic and advanced concepts of data mining and knowledge discovery.
2. Apply and analyze different techniques of data preprocessing, data visualization, and learning algorithms.
3. Evaluate the results of a mining algorithm through various performance evaluation and visualization techniques.
4. Create analytical solution for a real-world data mining problem, and implement using available data mining toolkits.

STAMFORD UNIVERSITY BANGLADESH
Faculty of Engineering
Department of Computer Science and Engineering

Mapping Course Learning Outcomes (CLOs) with Program Learning Outcomes (PLOs)

CLOs	PLOs																	
	Fundamental Skills (FS)						Social Skills (SS)						Thinking Skills (TS)		Personal Skills (PS)			
	FS1	FS2	FS3	FS4	FS5	FS6	SS1	SS2	SS3	SS4	SS5	SS6	TS1	TS2	PS1	PS2	PS3	PS4
CLO1	2	1					2			1			1		1		1	1
CLO2	2	2	1			1	1		1	1			1		1		1	1
CLO3	2	2	1	2		1	2			1			1		1		2	1
CLO4	2	2	1	2	1	1	1			1	1		2	1	1		2	2

**1 = Moderate, 2 = Strong, 3 = Very Strong

Mapping of Course Outcomes (CLOs) with the Program Outcomes (PLOs) Using Bloom's Taxonomy:												
No.	Course Learning Outcome (Upon completion of the course, the students will be able to)	Program Outcome										
		1	2	3	4	5	6	7	8	9	10	11	12
CLO1	Understand and remember basic and advanced concepts of data mining and knowledge discovery.	Y	Y								Y		Y
CLO2	Apply and analyze different techniques of data preprocessing, data visualization, and learning algorithms.	Y	Y	Y	Y	Y				Y	Y	Y	Y
CLO3	Evaluate the results of a mining algorithm through various performance evaluation and visualization techniques.	Y	Y			Y				Y	Y	Y	
CLO4	Create solution for a real-world data mining problem using available data mining toolkits.	Y	Y	Y	Y	Y	Y			Y	Y	Y	Y
Y=Yes													

Part B: Course Plan

Week Plan	Topic	Teaching-Learning Strategy	Assessment Strategy	Corresponding CLOs
Week 1	Course Orientation			
	Introduction to data, data mining, and knowledge discovery	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Class Test (Short Questions) • Oral Presentation • Viva-voce • Open-ended questions	1
Week 2	Data Preprocessing, summarization, visualization techniques	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs, Short Questions) • Group Work (Group discussion, Group assignment) • Class Test (Short Questions) • Oral Presentation • Viva-voce • Open-ended questions	1, 2

Week Plan	Topic	• Teaching-Learning Strategy	• Assessment Strategy	Corresponding CLOs
Week 3	Rule Based Mining, Association and Correlation Mining	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs, Short Questions) • Group Work (Group discussion, Group assignment) • Class Test (Short Questions) • Oral Presentation • Viva-voce • Open-ended questions	1, 2
Week 4	Classification: Basic Concepts and Techniques, K Nearest Neighbor Classifier	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs, Short Questions) • Group Work (Group discussion, Group assignment) • Class Test (Short Questions) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 5	Decision Tree: ID3 algorithm	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 6	Decision Tree: C4.5 and CART algorithm	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 7	Ensemble Learning: Bagging and Boosting	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 8	Artificial Neural Network: Basics	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce Open-ended questions	1,2,4

Week Plan	Topic	• Teaching-Learning Strategy	• Assessment Strategy	Corresponding CLOs
Week 9	Artificial Neural Network: Backpropagation	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce Open-ended questions	1,2,4
Week 10	Performance Evaluation Techniques	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	3
Week 11	Clustering: Introduction and Basic concepts	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 12	Clustering: Advanced	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 13	Dimensionality Reduction Techniques: Concepts	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2
Week 14	Dimensionality Reduction Techniques: Continued Review	• Direct Instruction (Class lecture) • Indirect Instruction (Reflective discussion, Inquiry, Case study) • Interactive Instruction (Discussion) • Experiential Learning (Demonstrate through online/software implementation) • Independent Learning (Journals, Reports)	• Close-ended questions (Matching, F/B, MCQs) • Group Work (Group discussion, Group assignment) • Assignment (Project-based, Concept mapping etc.) • Oral Presentation • Viva-voce • Open-ended questions	1,2,3,4

Part C: Assessment and Evaluation

The assessment and evaluation strategies for the course are given as follows:

Assessment Strategies	
Class Participation:	Students' individual in-class responses, attention, and sense of discipline, morality will be adjudged on the basis of 5 (five) marks.
Class Test/Quiz:	Students will sit for only 1 (one) class test/quiz during the semester. The test/quiz will be taken before midterm. Class test/quiz marks will be assessed in 5 (five). **No makeup class test will be taken.** Students are strongly recommended not to miss any test.
Group Work:	The students will have to form groups consisting of maximum 4 members. There will be 1 assignment consisting of 5 (five) marks. The topics or case studies will be given as assignments in groups during the class which they have to prepare at home and will submit on or before the due date.
Assignment:	**No late submission of assignments will be accepted.**
Oral Presentation:	Students, in groups, will have to present the report of their assignments. Oral presentations of the students will be assessed in 5 (five) marks. **No late presentation will be accepted.**
Viva-vocé	Students will have to appear for viva-vocé before their Midterm (5 marks) and Final examination (5 marks).
Midterm Exam:	Midterm exam will be held according to the Academic Calendar published by the university. Midterm assessment marks will be 30 (thirty).
Final Exam:	Final exam will be held according to the Academic Calendar published by the university. Final assessment marks will be 40 (forty). Course contents learnt before Midterm examination will be included in the syllabus of Final Examination.
Make-up Procedure:	No late submission and/or make-up assignment/presentation/quiz will be allowed without prior permission and adequate and reasonable proof of absence.

STAMFORD UNIVERSITY BANGLADESH
Faculty of Engineering
Department of Computer Science and Engineering
Part C: Assessment and Evaluation (continued…)

Marks Distribution: CIE & SMEE

Assessment Type	Assessment Criteria	Bloom's Taxonomy								Marks Distribution
		Cognitive						Affective	Psycho-motor	
		Remember	Understand	Apply	Analyze	Evaluate	Create	A1- A5	P1-P5	
Continuous Internal Evaluation (CIE)	Class Participation	5						A2	P3, P4	5%
	Assignments/ Short projects			10				A2, A3		10%
	Class Test, Viva-vocé, Oral Presentation		10	5						15%
Semester Mid and End Examination (SMEE)	Midterm Examination		15	15						30%
	Final Examination		10	10	20					40%
	Total	5	35	40	20					100%

Attendance and Class Performance Marks Distribution:

** A student is expected to attend all the classes in each course. Maximum 20% absence in a course in one semester may be exempted under emergency. It is the responsibility of the student to keep the course teacher informed regarding absence from classes.

** A student may be dropped from a course for absence from three consecutive classes without sufficient reasons and evidence.

Attendance (%)	Marks
90-100	5
85-89	4.5
80-84	4.0
75-79	3.5
70-74	3.0
65-69	2.5
60-64	2.0
55-59	1.5
50-54	1.0
45-49	0.5
Below 45%	0

Grading System:
Letter grades and grade points are used to evaluate the performance of a student in the given course as follow:

Marks Range	Letter Grade	Grade Point
80% and above	A+	4.00
75% to less than	A	3.75
70% to less than	A-	3.50
65% to less than	B+	3.25
60% to less than	B	3.00
55% to less than	B-	2.75
50% to less than	C+	2.50
45% to less than	C	2.25
40% to less than	D	2.00
Less than 40%	F	0.00

Code of Conduct

1. It is mandatory for all the students to participate in the class regularly and maintain proper discipline in the class.
2. If a student fails to attend any class test, term exams, or final examination, he/she will get a zero in that class test, term, or final examination.
3. Adopting unfair means in the exams will be considered as a serious crime and the student shall be placed to the university disciplinary committee.
4. All the assignments, class test and exam copies should be neat and clear and demonstrate professionalism.
5. No student is allowed to duplicate other student's work directly or with minor change.
6. Plagiarism is strictly restricted. One need to provide a reference while using someone else's words, ideas, or research in assignments/exams.

Part D: Learning Materials

Recommended Readings:	
Textbooks:	1.Krzysztof J. Cios, Witold Pedrycz, Roman W. Swiniarski and Lukasz A. Kurgan, Data Mining: A Knowledge Discovery Approach, published by Springer. 2. Alex A. Freitas, Data Mining and Knowledge Discovery with Evolutionary Algorithms, published by Springer.
Reference Books:	1. Daniel T. Larose, Discovering Knowledge in Data: An Introduction to Data Mining, published by Wiley-Interscience
Supplementary Learning Materials:	• Lecture Sheets • Listening Materials from YouTube and other online materials • Research articles