The Hashemite University

الجامعـــــة الهاشميـــــة

Deanship of Academic Development and International Outreach

عمادة التطوير الأكاديمي والتواصل الدولي

Syllabus: Data Mining (2010042312)

Second Semester 2021-2022

COURSE INFORMATION	
Course Name: Data Mining **Semester:** Second **Department:** Department of Information Technology **Faculty:** Prince Al-Hussein Bin Abdullah II Faculty for Information Technology	**Course Code: 2010042312** **Section:** 1 **Core Curriculum**:
Day(s) and Time(s): Sun, Tue, Thu 11:00-12:00 **Classroom:** e.g. IT206	**Credit Hours:** 3 **Prerequisites**: 151002240

COURSE DESCRIPTION

Data mining or knowledge discovery from databases (KDD) is one of the most active areas of research in databases. It is at the intersection of database systems, statistics, AI/machine learning, and data visualization. In this course, we will introduce the concepts of data mining and present data mining algorithms and applications. Topics include association rule mining, sequential pattern mining, classification models, clustering, data visualization, mining complex types of data (text mining, multimedia mining, Web mining), data mining languages, data mining applications and new trends. The practical part includes applications and exercises using a data mining tool such as WEKA.

DELIVERY METHODS

The course will be delivered through a combination of active learning strategies. These will include:

- Class lectures: Class lectures will expose students to the knowledge required by this course
- Collaborative learning through small groups for the purpose of solving the course's project.
- E-learning resources: e-reading assignments and practice quizzes through Model and Microsoft Team

FACULTY INFORMATION

Name	**Ahmad Aloqaily**
Academic Title:	**Associate Professor**
Office Location:	**IT224**

Telephone Number:	**Ext. 5083**
Email Address:	**aloqaily@hu.edu.jo**
Office Hours:	**Sun, Tue and Thu** 10:00-11:00 **Mon and Wed** 11:00-12:00 *Please send an e-mail (aloqaily@hu.edu.jo) to meet at any other time.*

REFERENCES AND LEARNING RESOURCES

Required Textbook: Jiawei Han, Micheline Kamber and Jian Pei, Data Mining -- Concepts and Techniques, Morgan Kaufmann, Third Edition, 2013.

Useful Web Resources:

Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

STUDENT LEARNING OUTCOMES MATRIX*

Core Curriculum Learning Outcomes	Program Learning Outcomes	Course Student Learning Outcomes	Assessment Method
CC-LO-1: Maintaining excellence in the educational process, especially applied education.	PLO-9: Develop the required skills to work on teams of people from diverse backgrounds in developing Data-driven and AI-based solutions.	Be able to understand the concepts, strategies, and methodologies related to the design and construction of data mining.	• Exams • assignments
		Be able to determine an appropriate mining strategy for given large dataset.	• Exams assignments
		Be able to understand the concepts, strategies, and methodologies related to the design and construction of data mining.	• Exams • assignments
.CC-LO-3: Preparing the competitiveness of graduates at a level capable of supplying the Jordanian and international market with qualified scientific competencies.	PLO-2: Demonstrate proficiency in different AI algorithms and techniques.	Be able to apply appropriate mining techniques to extract unexpected patterns and new rules that are "hidden" in large databases.	• Exams • assignments
	PLO-3: The ability to develop and assess Data-driven and AI-based solutions.	Be able to obtain knowledge of current data mining applications.	• Exams • assignments
	PLO-1: Demonstrate proficiency in different data analytics algorithms and techniques.	Be able to comprehend several data pre-processing methods.	• Exams • assignments

ACADEMIC SUPPORT

It is Hashemite University policy to provide educational opportunities that ensure fair, appropriate and reasonable accommodation to students who have disabilities that may affect their ability to participate in course activities or meet course requirements. Students with disabilities are encouraged to contact their Instructor to ensure that their individual needs are met. The University through its Special Need section will exert all efforts to accommodate individual needs.

Special Needs Section:

Tel: 053903333 EXT 5023/4583

Location: (https://hu.edu.jo/facnew/index.aspx?typ=68&unitid=70000000)

Email: (huniv@hu.edu.jo)

COURSE REGULATIONS

Participation

Class participation and attendance are important elements of every student's learning experience at Hashemite University, and the student is expected to attend all classes. A student should not miss more than 15% of the classes during a semester. *Those exceeding this limit of 15% will receive a failing grade regardless of their performance*. It is a student's responsibility to monitor the frequency of their absences. **Attendance record begins on the first day of class irrespective of the period allotted to drop/add and late registration. It is a student's responsibility to sign in; failure to do so will result in a non-attendance being recorded.**

In exceptional cases, the student, with the instructor's prior permission, could be exempted from attending a class provided that the number of such occasions does not exceed the limit allowed by the University. The instructor will determine the acceptability of an absence for being absent. A student who misses more than 25% of classes and has a valid excuse for being absent will be allowed to withdraw from the course.

Plagiarism

Plagiarism is considered a serious academic offense and can result in your work losing marks or being failed. HU expects its students to adopt and abide by the highest standards of conduct in their interaction with their professors, peers, and the wider University community. As such, a student is expected not to engage in behaviors that compromise his/her integrity as well as that of Hashemite University.

Plagiarism includes the following examples and it applies to all student assignments or submitted work:

- **Use of the work, ideas, images or words of someone else without his/her permission or reference to them.**
- **Use of someone else's wording, name, phrase, sentence, paragraph or essay without using quotation marks.**
- **Misrepresentation of the sources that were used.**

The instructor has the right to fail the coursework or deduct marks where plagiarism is detected

Late or Missed Assignments

In all cases of assessment, students who fail to attend an exam, class project or deliver a presentation on the scheduled date without prior permission, and/or are unable to provide a medical note, will automatically receive a fail grade for this part of the assessment.

- Submitting a term paper on time is a key part of the assessment process. Students who fail to submit their work by the deadline specified will automatically receive a 10% penalty.

Assignments handed in more than 24 hours late will receive a further 10% penalty. Each subsequent 24 hours will result in a further 10% penalty.

- In cases where a student misses an assessment on account of a medical reason or with prior permission; in line with University regulations an incomplete grade for the specific assessment will be awarded and an alternative assessment or extension can be arranged.

Student Complaints Policy

Students at Hashemite University have the right to pursue complaints related to faculty, staff, and other students. The nature of the complaints may be either academic or non-academic. For more information about the policy and processes related to this policy, you may refer to the students' handbook.

COURSE ASSESSMENT

Course Calendar and Assessment

Students will be graded through the following means of assessment and their final grade will be calculated from the forms of assessment as listed below with their grade weighting taken into account. The criteria for grading are listed at the end of the syllabus

Assessment	Grade Weighting	Deadline Assessment
Mid-term Exam	30%	TBA
Assigmnets	30%	29-5-2022
Final Exam	40%	TBA

Description of Exams

Test questions will predominately come from the material presented in the lectures. Semester exams will be conducted during the regularly scheduled lecture period. The exam will consist of a combination of multiple-choice, short answer, match, true and false and/or descriptive questions.

Homework: Will be given for each chapter, while the chapter is in progress you are supposed to work on them continuously and submit them in the next lecture when I finish the chapter.

You are also expected to work on in-chapter examples, self-tests and a representative number of end-of-chapter problems. The answers to self-tests and end-of-chapter exercises are given at the end of the book.

Quizzes: Unannounced quizzes will be given during or/and at the end of each chapter based upon the previous lectures. It will enforce that you come prepared for the class.

No make-up exams, homework or quizzes will be given. Only documented absences will be considered as per HU guidelines.

Grades are not negotiable and are awarded according to the following criteria*:

Letter Grade	Description	Grade Points
A+	Excellent	4.00
A		3.75
A-		3.50
B+	Very Good	3.25
B		3.00
B-		2.75
C+	Good	2.50
C		2.25
C-		2.00
D+	Pass	1.75
D	Pass	1.50
F	Fail	0.00
I	Incomplete	-

WEEKLY LECTURE SCHEDULE AND CONTENT DISTRIBUTION

"Lecture hours and weeks are approximate and may change as needed"

Chapter 1	**Introduction to Data Mining and KDD**	**Week 1**	**3 hours**
1.1	**Why Data Mining?**		
1.2	**What Is Data Mining?**		
1.3	**A Multi-Dimensional View of Data Mining**		
1.4	**What Kinds of Data Can Be Mined?**		
1.5	**What Kinds of Patterns Can Be Mined?**		
1.6	**What Kinds of Technologies Are Used?**		
1.7	**What Kinds of Applications Are Targeted?**		
1.8	**Major Issues in Data Mining**		
1.9	**Why Data Mining?**		
Chapter 2	**Getting to Know Your Data**	**Week 2-3**	**6 hours**
2.1	**Data Objects and Attribute Types**		
2.2	**Basic Statistical Descriptions of Data**		
2.3	**Data Visualization**		

2.4 Measuring Data Similarity and Dissimilarity			
Chapter 3	**Data Preprocessing**	**Week 4-6**	**9 hours**
3.1 Data Preprocessing: An Overview			
3.2 Data Quality			
3.3 Major Tasks in Data Preprocessing			
3.4 Data Cleaning			
3.5 Data Integration			
3.6 Data Reduction			
3.7 Data Transformation and Data Discretization			
Chapter 6	**Mining Association Rules**	**Week 7-9**	**9 hours**
6.1 Basic Concepts			
6.2 Frequent Item set Mining Methods			
6.3 Which Patterns Are Interesting?—Pattern Evaluation Methods			
Chapter 8	**Data Classification Techniques**	**Week 10-12**	**9 hours**
8.1 Classification: Basic Concepts			
8.2 Decision Tree Induction			
8.3 Bayes Classification Methods			
8.4 Rule-Based Classification			
8.5 Model Evaluation and Selection			
8.6 Lazy Learners (or Learning from Your Neighbors)			
Chapter 10	**Cluster Analysis**	**Week 13-14**	**6 hours**
10.1 Cluster Analysis: Basic Concepts			
10.2 Partitioning Methods			
10.3 Hierarchical Methods			
10.4 Density-Based Methods			
10.5 Grid-Based Methods			
10.6 Evaluation of Clustering			
Review		***Week 15***	
University Exams		***Week 16***	