

COURSE SYLLABUS

DATA MINING

Course code: 220109

1. General information

<i>Course type</i>	<i>Number of credits</i>	<i>Number of learning periods</i>
General <input type="checkbox"/>	Theory: 02 Exercise:	Theory: 30 Exercise:
Basic <input type="checkbox"/>		
Specialized <input checked="" type="checkbox"/>	Practice: 01	Practice: 30
Required <input checked="" type="checkbox"/>		
Elective <input type="checkbox"/>		

Learners:

Level	Bachelor
Discipline	Information Technology

Course requirements:

Prerequisites	Database
Parallels	None
Other requirements	None

2. Learning resources

Books	[1] Jiawei Han, Micheline Kamber (2011). <i>Data mining: Concepts and Techniques 3rd Edition</i> . Morgan Kaufmann Publishers. [2] Zhao, Y. (2012). <i>R and data mining: Examples and case studies</i> . Academic Press.
References	[3] Jamie MacLennan, ZhaoHui Tang & Bogdan Crivat (2011), <i>Data Mining with Microsoft SQL Server 2008</i> , Wiley Publishing. [4] Đỗ Phúc (2009). <i>Giáo Trình Khai Thác Dữ Liệu</i> . NXB Đại học Quốc gia TP.HCM.

Other learning materials	[5] http://www.kdnuggets.com/ [6] http://www.rdatamining.com/
--------------------------	---

3. Course description

The course provides students basic/specialized knowledge on data mining and the use of data mining in the knowledge economy. The course content includes multiples data mining strategies and algorithms which can be applied in business. The course also aims to provide opportunities to practice professional skills including selection and implementation of suitable data mining strategies for specific business context. Additionally, the course develops students' appropriate awareness and attitudes on the roles of data mining in modern social life in conjunction with important soft skills such as group working and report presentation.

4. Course learning outcomes (CLOs)

After finishing the course, students will be able to:

		Satisfy LOs of the program	Satisfy LOs of the ABET
❖ Topic 1: Disciplinary Knowledge and Reasoning			B.1.1
L1.	Present general concepts of data mining and the pipeline of data mining	1.2.1 1.2.2	B.1.2 B.1.4
L2.	Understand data characteristics and data preprocessing techniques	1.3.3 1.3.7	B.1.5 B.1.6
L3.	Present concepts of data warehouse and OLAP		
L4.	Apply mining strategies of frequent patterns and association rules		
L5.	Build and train classification and regression models for data mining		
L6.	Build clustering models for data mining		
L7.	Present applications and techniques of text mining and web mining		
L8.	Use data mining tools to solve basic problems		
❖ Topic 2: Personal and Professional Skills and Attributes			

L9.	Proactive Vision and Intention in Life	2.5.3	
❖ Topic 3: Interpersonal Skills: Teamwork and Communication			
L10.	Forming Effective Teams	3.1.1	
L11.	Technical and Multidisciplinary Teaming	3.1.5	
❖ Topic 4: Conceiving, Designing, Implementing and Operating Systems in The Enterprise, Societal and Environmental Context – The Innovation Process			
L12.	Understanding Needs and Setting Goals	4.2.1	
L13.	Defining Function, Concept and Architecture	4.2.2	
L14.	Modeling of System and Ensuring Goals Can Be Met	4.2.3	
L15.	Development Project Management	4.2.4	

5. Course content

Course content	CLOs	Number of learning periods		
		Theory	Practice	Others
Chapter 1. Introduction to Data mining	L1	5	0	
1.1. Why data mining?				
1.2. What is data mining?				
1.3. What kinds of data can be mined?				
1.4. What kinds of patterns can be mined?				
1.5. Which technologies are used?				
1.6. Major applications of data mining				
1.7. Issues of data mining				
1.8. A brief history of data mining				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9(T)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			

<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>	L12 (T)			
Chapter 2. Getting know data and data preprocessing	L2	5	0	
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 of data				
2.5. Data preprocessing: an overview				
2.6. Data cleaning				
2.7. Data integration				
2.8. Data reduction				
2.9. Data transformation and data discretization				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9 (U)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			
<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>				
Chapter 3. Data warehouse and OLAP	L3	2	3	
3.1. Data warehouse concepts				
3.2. Data cube and OLAP				
3.3. Design and application of data warehouse				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9 (U)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			
<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>				

Chapter 4. Mining frequent patterns and association rules	L4	5	7	
4.1. Basic concepts				
4.2. Frequent itemset mining				
4.3. Apriori algorithm				
4.4. FP-Growth algorithm				
4.5. Frequent patterns' interestingness evaluation				
4.6. Mining multilevel and multidimensional association rules				
4.7. Constraint based association rules				
4.8. Implementation association rule mining				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9 (U)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			
<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>	L12(T) <input type="checkbox"/> L15(T)			
Chapter 5. Classification and Regression	L5	6	10	
5.1. Classification concepts				
5.2. Decision tree based classification				
5.3. Bayes classification methods				
5.4. Rule based classification				
5.5. Linear Regression and Logistic Regression				
5.6. Model Evaluation and Selection				
5.7. Techniques to improve classification accuracy				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9(U)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			

<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>	L12(T) <input type="checkbox"/> L15(T)			
Chapter 6. Clustering	L6	4	5	
6.1. Clustering concepts				
6.2. Partition based clustering				
6.3. Hierarchical clustering				
6.4. Density based clustering				
6.5. Grid based clustering				
6.6. Clustering model evaluation				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9(U)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			
<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>	L12(T) <input type="checkbox"/> L15(T)			
Chapter 7. Text mining and web data mining	L7	3	5	
7.1. Text mining				
7.2. Web data mining				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	L9(U)			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	L10(U), L11(U)			
<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>	L12(T) <input type="checkbox"/> L15(T)			
Summary of skills in course level				
<input type="checkbox"/> <i>Personal and Professional Skills and Attributes</i>	T			
<input type="checkbox"/> <i>Interpersonal Skills: Teamwork and Communication</i>	U			
<input type="checkbox"/> <i>CDIO in the enterprise, societal and environmental context</i>	T			

6. Teaching and learning methods

ID	Teaching method/technique		Description
M1.	Lecturing	<input checked="" type="checkbox"/>	
M2.	Questions – Answers	<input checked="" type="checkbox"/>	
M3.	Group-based Learning	<input checked="" type="checkbox"/>	
M4.	Problem-based Learning	<input checked="" type="checkbox"/>	
M5.	Project-based Learning	<input type="checkbox"/>	
M6.	Case studies	<input checked="" type="checkbox"/>	
M7.	Role play	<input type="checkbox"/>	
M8.	Demo	<input checked="" type="checkbox"/>	
M9.	Simulations	<input type="checkbox"/>	
M10.	Debate	<input type="checkbox"/>	
M11.	Game	<input type="checkbox"/>	
M12.	Brainstorming	<input type="checkbox"/>	
M13.	Think-Pair-Share	<input type="checkbox"/>	

7. Course assessment

ID	Assessment activity		Quantity	Weight	LOs assessed
T1.	Text-based midterm exam	<input type="checkbox"/>			
T2.	Text-based final exam	<input checked="" type="checkbox"/>		50%	L4, L5, L6, L7
T3.	Practice midterm exam	<input type="checkbox"/>			
T4.	Practice final exam	<input type="checkbox"/>			
T5.	Report	<input checked="" type="checkbox"/>	10	20%	L4, L5, L6
T6.	In-class exercises	<input checked="" type="checkbox"/>	5	20%	L4, L5, L6
T7.	Homework assignments	<input checked="" type="checkbox"/>	5	10%	L1, L2, L3

T8.	Question – Answer	<input type="checkbox"/>			
T9.	Term Project	<input type="checkbox"/>			
T10.	Final Exam	<input type="checkbox"/>			
Overall score formula		$((T5*2 + T6*2+T7)/5 + T2)/2$			

8. Course requirements and expectations

8.1. Requirements on attendance

- Students are responsible for attending all classes. In case of absence due to force majeure circumstances, there must be sufficient and reasonable evidence.
- Students who do not attend more than 20% of the class sections, whether for reason or not, are deemed not to have completed the course and must re-enroll in the following semester.

8.2. Requirements and expectations on student behaviors

- Students must show their respects for teachers and other learners.
- Students must be on time. Students who are late more than five minutes will not be allowed to attend the class.
- Students should not make noise and interfere with others in the learning process.
- Students should not eat, chew gum, and use devices such as cell phones, music players during class hours.
- Laptops and tablets can only be used in class for the purpose of learning.
- Students who violate the above principles will be asked to leave the class and considered absent from the class.

8.3. Requirements on learning issues

Issues related to applying for score reservation, scoring complaints, scoring, exam disciplines are done according to the Learning Regulation of Tra Vinh University.

9. Tentative course instructor

Hà Thị Thúy Vi

DEAN

DEPARTMENT HEAD

LECTURER

Hà Thị Thúy Vi