

COURSE SYLLABUS

CSC14004 – DATA MINING AND APPLICATIONS

1. GENERAL INFORMATION

Course name:	Data Mining and Application
Course name (in Vietnamese):	Khai thác dữ liệu và ứng dụng
Course ID:	CSC14004
Knowledge block:	Specialized knowledge
Number of credits:	4
Credit hours for theory:	45
Credit hours for practice:	30
Credit hours for self-study:	90
Prerequisite:	
Prior-course:	AI, Data Base.
Instructors:	

2. COURSE DESCRIPTION

The course is designed to provide students with the basics of the field of Data Mining and its applications to other sciences. This course covers general contents related to the process of discovering knowledge from data and the insights related to common techniques in data mining such as common set mining and conclusions. knowledge assembly, classification, clustering and evaluation.

3. COURSE GOALS

At the end of the course, students are able to:

ID	Description	Program LOs
G1	Work at an individual and team level to present a sample of scientific reports and present subject-related content	2.2, 2.3.1, 2.3.2

G2	Know and explain English terms in the field of data mining	2.4.3, 2.4.5
G3	Explain basic concepts, terminology in the field of data mining	1.4
G4	Understand the meaning and application of data mining problems to real world problems	1.3.6, 1.4
G5	Understand and explain how basic algorithms work in data mining	1.4, 4.1.2
G6	Build up some components of a small-scale data mining system	5.1.3, 5.2.2, 5.3.2, 6.1.1, 6.1.2
G7	Use software tools	1.3.6

4. COURSE OUTCOMES

CO	Description	I/T/U
G1.1	Team formation, organization, operation and management	U
G1.2	Participating in group discussions and debates on subject topics	U
G1.3	Analyzing, synthesizing, and writing technical documents in a given form by individual or team collaboration	T, U
G1.4	Presentation on topics related to subject content	U
G2.1	Know, understand the specialized English terminology of their subject	I
G2.2	Read English language materials related to fake articles	I
G3.1	Explain the basics of data mining, components of the knowledge discovery process	I, T

G4.1	Understand the importance of data mining in practical applications as well as other scientific fields	I, T
G4.2	Selecting the right algorithm for problems using data mining	I, T
G5.1	Describing and illustrating the implementation process of algorithms in the subject	I, T, U
G5.2	Compare the same algorithms that solve a task in the data mining system	I, T, U
G5.3	Selecting, manually implementing algorithms in knowledge discovery into small data set	I, T, U
G5.4	Analysis and evaluation of knowledge obtained from data mining process	I, T, U
G6.1	Build some basic components for a simple data mining system	I, T, U
G7.1	Use a number of tools to support pre-processing, analysis, data mining tasks (e.g. Excel, Weka, RapidMiner,...)	I, T, U

5. TEACHING PLAN

ID	Topic	Course outcomes	Teaching/Learning Activities (samples)
1	Introducing Data Mining	G1.1, G1.2, G2.1, G3.1, G4.1	Lecturing Q&A, Group discussion QZ1: Quiz 1 (Chapter 1)
2	Preprocessing	G2.1, G3.1, G4.1, G4.2, G5.1, G5.2, G5.3	Lecturing Demonstration, Q&A QZ2: Quiz 2

3	Mining frequency pattern and association rules	G1.2, G2.1, G3.1, G4.2, G5.1, G5.2, G5.3, G5.4	Lecturing Demonstration, discussion
4	Mining frequency pattern and association rules	G2.1, G5.2, G6.1, G7.1	Lecturing Demonstration, discussion
5	Mining frequency pattern and association rules	G2.1, G5.1, G5.2, G6.1, G7.1,	Lecturing Demonstration
6	Classification	G1.2, G2.1, G3.1, G4.2, G5.1, G5.2, G5.3, G5.4	Question & answer Case study and discussion
7	Classification	G1.2, G2.1, G3.1, G4.2, G5.1, G5.2, G5.3, G5.4	Lecturing Demonstration
8	Clustering	G1.2, G2.1, G3.1, G4.2, G5.1, G5.2, G5.3, G5.4	Lecturing Q&A, discussion
9	Clustering	G2.1, G3.1, G4.2, G5.1, G5.2, G5.3, G5.4	Lecturing
10	Seminar topic	G1.2, G2.2, G3.1, G4.1, G5.2	Case study, discussion Demonstration

11	Seminar topic	G1.2, G2.2, G3.1, G4.1, G5.2	Lecturing Q&A, Discussion Project submitted
----	---------------	------------------------------	---

For the practical laboratory work, there are 10 weeks which cover similar topics as it goes in the theory class. Each week, teaching assistants will explain and demonstrate key ideas on the corresponding topic and ask students to do their lab exercises either on computer in the lab or at home. All the lab work submitted will be graded. There would be a final exam for lab work.

6. ASSESSMENTS

ID	Topic	Description	Course outcomes	Ratio (%)
A1	Assignments			40%
A11	Quizzes: QZ1, QZ2, QZ3, and QZ4.	Small quizzes in class for each topic	G1.1, G1.2, G2.1, G2.2, G4.1, G5.2	5%
A12	Homework: HW1, HW2, and HW3	HW1, HW3: reading comprehension and writing reports in English HW2, HW3: practicing based on knowledge taught in class	G1.3, G1.4, G3.1, G4.2, G5.1	5%
A13	Weekly labwork: LW1–LW10		G1.2, G2.1, G3.1, G4.1, G5.2, G5.3, G6.1, G7.1	30%
A3	Exams			60%

A32	Midterm exam	Closed book exam. Describe the understanding of different topics, analyze & program to solve problems	G2.1, G2.2, G3.1, G4.1	20%
A33	Final exam	Closed book exam. Describe the understanding of different topics, analyze & program to solve problems	G3.1, G4.1, G5.1, G5.2, G5.3, G5.4	40%

7.

RESOURCES

Textbooks

- **Data Mining: Concepts & Technique**, 4th Edition, Jiawei Han, Jian Pei, Hanghang Tong, The Morgan Kaufmann Series in Data Management Systems, 2022
- **Mining of Massive Datasets**, 3rd ed. Jure Leskovec, Anand Rajaraman, Effrey David Ullman. Stanford University, California. February 2020.
- **Introduction to Data Mining**, 2nd edition, P-N. Tan, M. Steinbach, V. Kumar, Pearson. (January 4, 2018)
- **Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data**, Second Edition, B. Liu, Springer, 2011

Others

- Visual Studio .NET
- Java
- Microsoft SQL, 2008

8. GENERAL REGULATIONS & POLICIES

- All students are responsible for reading and following strictly the regulations and policies of the school and university.
- Students who are absent for more than 3 theory sessions are not allowed to take the exams.



-
- For any kind of cheating and plagiarism, students will be graded 0 for the course. The incident is then submitted to the school and university for further review.
 - Students are encouraged to form study groups to discuss on the topics. However, individual work must be done and submitted on your own.