Program: ECE
Degree Level: Master

Faculty/College: TGGS

Course 090245340

Principles of Data Mining

King Mongkut's University of Technology North Bangkok
The Sirindhorn International Thai-German Graduate School of Engineering
Electrical and Computer Engineering Program

Section 1: General Information

1. **Course code and course title**

   090245340 Principles of Data Mining

2. **Total credits**

   3 credits ☐ (2-2-5) ☑ (3-0-6) ☐ (3-0-9) ☐ (2-3-7)

3. **Curriculum and course category**:

   Curriculum: *Master of Engineering in Electrical and Computer Engineering*

   Course category: Required Courses

   ☐ Core Course ☐ Specific Core Course
   ☐ Industrial Internship ☐ Master Thesis

   Elective Courses

   ☐ General Elective ☐ Specific Elective ☐ Other Elective

4. **Course coordinator/ Instructors**

   Course Coordinator: _____

   Instructor(s): Yodsawalai Chodpathumwan

5. **Semester/ year of study**

   ☐ Semester 1 (Aug. to Dec.) ☑ Semester 2 (Jan. to May) Academic Year: 2021

6. **Pre-requisite** (**if any**)

   ☑ No ☐ Yes, please provide: .....

7. **Co-requisites** (**if any**)

   ☑ No ☐ Yes, please provide: .....

8. **Venue of study**

   Lecture Day/Time: Tuesday 9.00-12.00

   ☐ On-site: Lecture Room No.:……..... ……. Floor:……..…

   ☑ TGGS, KMUTNB ☐ Faculty of Engineering, CU ☐ RWTH

   ☑ On-line*: Teaching Media: ☑ Microsoft Teams ☐ Google Meet
   ☐ Zoom ☐ Webex
   ☐ Other (specify) …… ……..…

Remark: * *During COVID-19, the teaching can be on-site and/or on-line according to TGGS Policy*.

9. Information for quality assurance in education

This course shows evidence of:

- Integration of research or creative activities with instruction; use of research-based learning management; knowledge management practices for learning improvement
- Integration of academic services and course implementation

10. Date of latest revision:

July 2020

Section 2: Course Description and Implementation

1. Course Description (*As written in the Official Approved Curriculum*)

Principles and algorithms of data mining. Data cleaning and integration. Descriptive and predictive mining. Frequent, sequential and structured pattern mining. Clustering. Outlier analysis and fraud detection. Other research topics in data mining.

2. Number of hours per semester

Lecture	Practice	Self-study
45 hours/ semester (*3 hours/week**)	*30 hours* (*2 hours/week**)	*75 hours/ semester* (*5 hours/week**)

Remark: * *Based on 15 weeks of lecture*

Course Category: ☑ Lecture ☐ Practice ☐ Laboratory

Course Evaluation: ☑ A-F ☐ S/U ☐ P

3. Number of hours per week for academic guidance to individual students

☐ 1. Giving academic advice (minimally number hour per week) during the office hour

☑ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ .....

The student can arrange the time via email for the meeting date/time.

☐ 2. Adopting information technology-based academic advising

☑ Email: yodsawalai.c@tggs.kmutnb.ac.th

☐ Phone:

☐ Communication Apps: Line ID:

(Please notify the lecturer when adding the line.)

☐ Meeting Online: The platform will be informed to students upon the request.

☐ Other (specify) .....

☐ 3. .....

Program: ECE
Degree Level: Master

Faculty/College: TGGS

4. Course Learning Outcomes (CLOs): Students should be able to:

CLO 1. Explain and analyze basic principles for data preparation and management including data cleaning, data transformation and data warehousing in the context of data mining

CLO 2. Explain the basic principles of frequent pattern mining, and apply the mining method for effective data mining

CLO 3. Demonstrate proficiency in theoretical principals in the context of data mining

CLO 4. Identify, analyze and modify existing methods in data mining in various context

CLO 5. Design and apply data mining methods to solve problems in real-world context and communicate result

5. The mapping between Expected Learning Outcomes (ELOs) from the curriculum and Course Learning Outcomes (CLOs)

Table 5.1 ELOs-CLOs Consistency *(for a subject-specific course/ a specific curriculum)*

ELOs/CLOs consistency	CLO 1	CLO 2	CLO 3	CLO 4	CLO 5
ELO1					
ELO2	✓	✓	✓	✓	
ELO3			✓	✓	
ELO4					
ELO5					
ELO6					✓
ELO7				✓	✓
ELO8				✓	✓
ELO9					
ELO10					

Remark: All ELOs and ELOs for the course (highlighted row) are as written in the Official Approved Curriculum.

Table 5.2 Mapping desirable characteristics of KMUTNB graduates and CLOs *(for non-specific courses designed for various curriculums)*

Consistency between desirable characteristics of KMUTNB Graduates- CLOs	CLO 1	CLO 2	CLO 3	CLO 4
1. **Professional credentials with critical thinking skills**			✓	✓
2. **Integrity and social responsibility**				

Consistency between desirable characteristics of KMUTNB Graduates- CLOs	CLO 1	CLO 2	CLO 3	CLO 4
3. **Innovative and technopreneur mindset**				✓
4. **Global Competence**			✓	

Section 3: Student Improvement in relation to Course Learning Outcomes (CLOs)

Organizing learning to develop skills/ knowledge; evaluation of CLOs in accordance with the ones identified in Section 2.4

Course Learning Outcomes (CLOs)	Teaching Methods compliant with CLOs	Evaluation Methods compliant with CLOs
CLO 1	• Lecture* • Case studies	• Exam****
CLO 2	• Lecture* • Case studies	• Exam****
CLO 3	• Lecture* • Case studies	• Exam****
CLO 4	• Lecture* • Case studies	• Exam**** • Presentation • Literature survey
CLO 5	• Case studies • Individual or group project	• Presentation

*Remark: * Lecture on the concept of the topic is introduced with basic or fundamental definitions, visualization and correlations. For the complicated equation, the derivation from the basic laws can be shown to students. So, the students do not memorize the equations but understand the basic concept and basic equation. The lecturer will introduce the advanced and new concepts, technologies, and findings to students from publications such as journals and websites and from the research and industrial experiences.*

*** Active learning by asking questions related to the topic in the lecture and encouraging the students to response to the questions. If the students cannot response with answers, then the lecturer will give some guidance until the students can response.*

**** Quiz in the closed-book format on the basic concepts and equations with simple problem solving to evaluate their learning. The solution will be given to students after grading, so they can identify their mistakes and weakness.*

***** Exam on the basic concepts and equations with simple problem solving in the closed-book format as a review, whereas the complicated/integrated problem solving will be worked in the open-book format.*

Program: ECE
Degree Level: Master

Faculty/College: TGGS

Section 4: Lesson Plan and Evaluation

1. **Lesson Plan**

Week	Topics/Details	CLOs	Hours	Learning and teaching activities; teaching media (if any)	Lecturer
1	**Introduction** - data mining overview - recent research **Review** - data objects - statistics and similarity measurement	CLO 1	3.0	• Lecture presentation slides • Q&A	Yodsawalai
2-3	**Data Preprocessing** - data quality - data cleaning - data integration - data reduction - data transformation - data discretiazation	CLO 1 CLO 3	6.0	• Lecture presentation slides • Q&A	Yodsawalai
4-6	**Mining Frequent Patterns** - basic concepts - efficient pattern mining method - mining closed/max patterns - mining colossal patterns - mining sequential patterns - contrainted-based mining - association rules mining	CLO 2 CLO 3	9.0	• Lecture presentation slides • Q&A	Yodsawalai
7	*Literature survey presentation*	CLO 4	3.0	In-class presentation	
8	**Classification** - basic concepts and methods - model evaluation	CLO 2 CLO 3	3.0	• Lecture presentation slides • Q&A • Assignment #2	Yodsawalai
9	**Clustering Analysis** - basic concepts - partitioning methods - hierarchical methods - density-based methods - grid-based methods - clustering evaluation	CLO 2 CLO 3	3.0	• Lecture presentation slides	Yodsawalai

Program: ECE
Degree Level: Master

Faculty/College: TGGS

Week	Topics/Details	CLOs	Hours	Learning and teaching activities; teaching media (if any)	Lecturer
10-12	**Network Mining** - motif in network - link analysis - heterogeneous information network - network clustering - network cascading	CLO 2 CLO 4	6.0	• Lecture presentation slides • Q&A	Yodsawalai
13	**Outlier Analysis**	CLO 3	3.0	• Lecture presentation slides • Q&A	Yodsawalai
14	**Data Exchange** - schema mapping - basic algorithms in data integrations	CLO 3	3.0	• Lecture presentation slides • Q&A	Yodsawalai
15	*Final Exam*	CLO 1 CLO 2 CLO 3 CLO 4	3.0		Yodsawalai
16	*Project presentation*	CLO 5	3.0	• In-class presentation	
		Total	48.0		

2. **Evaluation Plan** (**in accordance with OBE** 2 **mapping framework**)

Course Learning Outcomes (CLOs)	Evaluation Methods	Week of Evaluation	Percentage of Evaluation
CLO 4, CLO 5	1 Literature Survey	7, 16	40%
CLO 5	1 Project	16	20%
CLO 1, 2, 3	3 Exams	15	40%

Section 5 Teaching/Learning Resources

Textbooks and materials

- J. Han, J. Pei, and M. Kamber. “Data mining: concepts and techniques.” Elsevier, 2011.
- Y. Sun, and J. Han. “Mining heterogeneous information networks: principles and methodologies.” Morgan & Claypool Publishers, 2012.

- C. M. Bishop, “Pattern Recognition and Machine Learning,” Springer, 2007..
- S. Chakrabarti, “Mining the web: statistical analysis of hypertext and semi-structured data,” Morgan Kaufmann, 2002.
- T. Hastie, R. Tibshirani, and J. Friedman, “The elements of statistical learning: data mining, inference and prediction,” Springer-Verlag, 2001.
- C. Shi, Y. Li, J. Zhang, Y. Sun, and P. S. Yu. “A survey of heterogeneous information network analysis.” IEEE TKDE, 2017.
- A. Doan, A. Halevy, and Z. Ives. “Principles of data integration.” Elsevier, 2012.
- C. Zhai, and S. Massung. “Text data management and analysis: a practical introduction to information retrieval and text mining”. Morgan & Claypool, 2016.
- A. Barabási. "Network science." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 371.1987, 2013.
- X. Dong, and T. Rekatsinas. "Data integration and machine learning: A natural synergy." Proceedings of the 2018 international conference on management of data. 2018.
- J. Leskovec, A. Rajaraman, and J. D. Ullman. "Mining of massive data sets. Cambridge university press”, 2020.

Section 6 Course Evaluation and Improvement

1. Course evaluation by students

The students will have an opportunity to evaluate the effectiveness of the course in a form of paper survey and group interview at the end of each semester. The results of survey and interview including the grading will be reviewed by the curriculum meeting to evaluate the course’s effectiveness.

2. Strategies for assessing learning management

The students will have an opportunity to evaluate the teaching of the course in a form of paper survey and group interview at the end of each semester. The results of survey and interview including the grading will be reviewed by the curriculum meeting to evaluate the teaching as well as returning to the lecturer for further improvement.

3. Improvement schemes of course implementation

The evaluation from the students including the grading will be submitted to the curriculum meeting for reviewing and brainstorming to improve teaching of each course. Comments and suggestions given by the curriculum meeting will be informed to the responsible lecturer of each course.

4. Verification of students’ learning outcomes, referred to OBE 2 and 3

The grading of this course will be evaluated and reviewed by the Department meeting and the TGGS executive board meeting in order to verify its appropriateness before the final approval.

5. Course review and improvement plans

The results of the grading evaluation and student evaluation will be submitted to the curriculum meeting for reviewing and brainstorming to improve the effectiveness of the offered courses. Comments and suggestions will be informed to the responsible lecturer of each course.

.