Data Mining
13016366
Semester 1 Year 2018

Course Information

Instructor : Dr.Pimprapai Thainiam
Office : IC07 Logistics and Supply Chain Management on 8th floor
Email : pimprapai.th@kmitl.ac.th
Office Hours : Monday 1.00 pm – 4.00 pm or by appointment
Class time : Tuesday 9.00 am – 12.00 pm
Class room : IC03 on 8th floor

Learning Outcomes

1. Students will be able to explain and use the mining process for descriptive and predictive analytics.
2. Students will be able to explore data using various statistical and visualization techniques.
3. Students will be able to understand and apply the core data mining methods of Classification, Association Rules Mining, Cluster Analysis, and Regression.
4. Students will be able to conduct a complete data mining project including data preparation and documentation of the results.

Textbooks

Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, University of Minnesota, University of Minnesota, Addison Wesley, 2006. Note: You can use any edition of this book

Data Mining: Concepts and Techniques, Jiawei Han, Jian Pei, and Micheline Kamber. Elsevier, 2011. Note: You can use any edition of this book

Prerequisites

13006009 Probability and Statistics
13016237 Information Systems and Databases

Course Topics and Class Outlines

Week	Date	Topic	Note
1	August 7th	Introduction to Data Mining and R	
2	August 14th	Data and Data Quality	
3	August 21st	Data Preprocessing	
4	August 28th	No class (IC Quality Assurance)	

Week	Date	Topic	Note
5	September 4th	Data Exploration and Data Visualization	
6	September 11th	Basic Classification I	Assignment 1 due
7	September 18th	Basic Classification II	
8	September 25th	Midterm Exam	
9	October 2nd	Alternative Classification I	Open books and notes
10	October 9th	No class (attend the training course)	
11	October 16th	Alternative Classification II	
12	October 23rd	No class (Chulalongkorn Day)	Assignment 2 due
13	October 30th	Basic Association Rules Mining	
14	November 6th	Alternative Association Rules Mining	
15	November 13th	Basic Clustering	Assignment 3 due
16	November 20th	Alternative Clustering, Regression	Last day of the class
18	December 4th	Final Exam	Open books and notes Assignment 4 due

Course Work and Grading

Topic	Percentage
Assignment 1	15%
Assignment 2	15%
Assignment 3	15%
Assignment 4	15%
Midterm Exam	20%
Final Exam	20%

Assignments

You will work on four assignments throughout the semester. The result of each assignment is a detailed report (following the categories of the CRISP-DM framework). The reports need to be submitted to **pimprapai.th@kmitl.ac.th** not later than 9.00 am on their assigned due dates. **Late submission will be penalized 1 point per day** (24-hour period) (Each assignment has 10 points). You can expect up to 90% if you address all project questions adequately. For a higher grade, you need to present exceptional work (e.g., in-depth explanation why one method works better, developing special preprocessing, using and explaining a method not covered in class).

Attendance Policy

Students are expected to attend class regularly. If a student is absent from class, it is that student's responsibility to make arrangements with the instructor to make up any work missed or to ensure that reports are submitted on time or early. Late submission will be accepted in case of documented emergency (e.g., medical emergency). Any assignments that will be missed (including those due to university-sanctioned events) must be completed before the due date.