UNIVERSITY OF INFORMATION TECHNOLOGY

Faculty of Information Systems

Introduction to the course

Data Mining

Cao Thi Nhan

INTRODUCTION

1. Course name: Data Mining
2. Number of credits: 4 (3 Theory – 1 Lab)
3. Description
4. Content
5. Requirements and Expectations
6. Assessment
7. References

Description

- The course provides basic concepts of data mining and data mining process as well as main phases of data mining process.
- Students are equipped many commonly used methods such as: classification, clustering, regression, association rules…
- Practice teamwork skills

Content

- Chapter 1: Overview of Data mining
- Chapter 2: Data preprocessing
- Chapter 3: Frequent itemsets and association rule
- Chapter 4: Frequent episodes
- Chapter 5: Rough set
- Chapter 6: Data classification
- Chapter 7: Data clustering
- Chapter 8: Data mining and database technology

Requirements and Expectations

- Preparation
- Attendance (NOT absent over 3 class sessions)
- Teamwork
- Homework
- Project: Group: 2-3 students

Assessment

- Midterm exam: 20% (6.11.2018)
- Project: 30%
- Final exam: 50%
- Bonus score: Max +5% to Project
  - Discussion, homework
  - Other activities

Project

- Group: 2-3 students/group.
- Students have to register at a link in course by: 22-09-2018
- Students can choose favorite topics (not the same among groups):
  - Data set
  - Data preprocessing
  - Data mining algorithm
  - Assess result
  - Build an application based on the data mining result

References

1. **Jiawei Han, Micheline Kamber, and Jian Pei,** *Data Mining Concepts and Techniques,* 3 edition, Morgan Kaufmann Publishers, 2011.
2. **Jiawei Han, Micheline Kamber, and Jian Pei,** *Data Mining Concepts and Techniques Slides.*
3. **Ian H. Witten, Eibe Frank, Mark A. Hall,** *Data Mining* **–** *Practical Machine Learning Tools and Techniques, 3rd,* Elsevier *2011.*
4. **Do Phuc,** *Data Mining***,** Vietnam National University, 2006.
5. **Ho Tu Bao**, *Introduction to knowledge discovery and data mining*, IOIT, 2001.


QUESTIONS
ANSWERS