

OHLONE COLLEGE
Ohlone Community College District
OFFICIAL COURSE OUTLINE

I. Description of Course:

1. Department/Course: CS - 139

2. Title: Data Mining

3. Cross Reference:

4. Units: 3.00

Lec Hrs: 3.00

Lab Hrs: 0.00

5. Repeatability: Yes Times:2

6. Grade Options:

Letter Grade, May Petition Credit/No
Credit (GC)

7. Degree/ Applicability:

Credit, Degree Applicable, Not
Transferable (D)

8. General Education:

9. CAN Numbers:

10. Field Trips: Not Required

11. Requisites:

None

12. Catalog Description:

Data mining has attracted a great deal of attention in the information industry in recent years due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge. This course offers a sound understanding of data mining principles and teaches proven methods for knowledge discovery in large corporate databases. Topics covered include Data Description for Data Mining, Predictive Data Mining, Data Mining Models and Algorithms, and The Data Mining Process.

13. Class Schedule Description:

Data mining concepts and techniques for uncovering data patterns hidden in data sets are discussed.

14. Counselor Information:

This is an introductory course in Data mining. Student will learn the principles and techniques for extracting information and knowledge discovery in large corporate databases.

II. Student Learning Outcomes

The student will:

1. Explain the fundamental concepts of data mining Develop and practice techniques used in data mining
2. Describe, explain, and apply the data mining strategies
3. Display the ability to apply data mining algorithms
4. Describe and explain the most commonly used data mining models
5. Describe and explain the data mining process
6. Evaluate data warehouse design
7. Describe and explain the methods for knowledge discovery

III. Course Outline:

- A. Each week has 3 hours lecture
- B. Week 1 Introduction
 - 1. - Data mining: a definition
 - 2. - History and application
- C. Week 2 Data Mining Fundamentals
 - 1. - What can computers learn?
 - 2. - Is data mining appropriate for my problem?
- D. Week 3 A Simple Data Mining Process Model
 - 1. - Assembling the data
 - 2. - The Data Warehouse
 - 3. - Mining the data
 - 4. - Interpreting the results
- E. Week 4 Data Mining Applications
 - 1. - Example applications
- F. Week 5 Data Mining Strategies
 - 1. - Classification, Estimation, Prediction, Unsupervised clustering
 - 2. - Association rules
 - 3. - Evaluating performance
- G. Week 6 Data Mining Techniques
 - 1. - Decision Trees
 - 2. - General association rules
 - 3. - The K-means algorithm
- H. Week 7 A five-step approach for Unsupervised clustering
- I. -Applying the five-step approach for Unsupervised clustering
- J.
- K. A six-step approach for Supervised learning
- L. -Applying the six-step approach for Supervised learning
- M.
- N. Week 8 Techniques for Generating Rules
 - 1. - Instance typicality
 - 2. - Special considerations and features
- O. Week 9 Midterm
- P. Week 10 Knowledge discovery in Databases
 - 1. - A KDD process model
 - 2. - The CRISP-DM process model
- Q. Week 11 The Data Warehouse
 - 1. - Operational Databases
 - 2. - Data Warehouse design
 - 3. - On-line Analytical processing
- R.
- S. Week 12 Formal Evaluation Techniques
 - 1. - Tools for evaluation
 - 2. - Computing test set confidence interval
 - 3. - Comparing Supervised learner models
- T. Week 13 Formal Evaluation Techniques & continue
 - 1. - Unsupervised evaluation techniques
 - 2. - Attribute evaluation
 - 3. - Evaluating Supervised models with numeric output
- U.
- V. Week 14 Neural Networks

1. - Feed-Forward neural networks
 2. - Neural network training: a conceptual view
 3. - Neural network explanation
- W.
- X. Week 15 Neural Networks ζ continue
1. - General considerations
 2. - Neural network training: a detailed view
- Y. Week 16 Statistical Techniques
1. - Linear Regression Analysis
 2. - Logistic Regression
 3. - Bays Classifier
 4. - Clustering algorithms
- Z. Week 17 Specialized Techniques
1. - Time-Series analysis
 2. - Mining the web
 3. - Mining Textural Data
- AA. Week 18 Final Exam

IV. Course Assignments:

Reading Assignments

Writing Assignments

Homework assignments.

- (1) Students need to search on Web for advanced study on data warehouse and data mining subject
- (2) Written assignment should cover some research field such as Biotech, data communications

Projects, Activities, and other Assignments

V. Methods of Evaluation:

1. A. Written Assignments: (approximate 30 % of grade)
- B. Describe:
- C. Homework assignments.
- D. (1) Students need to search on Web for advanced study on data warehouse and data mining subject
- E. (2) Written assignment should cover some research field such as Biotech, data communications
 1. B. Computational or Non-Computational Problem Solving (approximate % of grade)
- F. Describe
 1. C. Skill Demonstrations (approximate 20 % of grade)
- G. Describe
- H. Project PowerPoint presentation: (1) Main topic student is assigned and contents (2) diagram of data warehouse and data mining student designed (3) programming demo
 1. D. Objective Exams (approximate 50 % of grade)
- I. Describe
- J. Concept tests in multiple choice, true/false questions and skill tests in short-answer problems.

Methods of Instruction:

Discussion

Lecture

VI. Textbooks:

Required

1. Richard Roiger, Michael Geatz *Data Mining* first Edition, Addison Wesley, 2003, Boston, MA 02116, 2004 ISBN: 0201741288

Optional

VII. Supplies:

1. none