

# Data Mining

(2023095)

2 Credits

Professors: CHEN Xi  
E-mail: chen\_xi@zju.edu.cn  
Office hours: tbc  
Teaching Assistant: tbc

## Pre-requisites:

Management Statistics

## Course Materials (Readings/textbooks):

Lecture notes will be provided

## Course Description

With the rapid developments in computer and data storage technologies, the fundamental paradigms of classical data analysis are mature for change. Data mining techniques aim at helping people to work smarter by revealing underlying structure and relationships in large amounts of data. This course takes a practical and research oriented approach to introduce the new generation of statistical data mining techniques and show how to use them to make better decisions. Topics include data mining project models, leading data mining software, important data mining models and their business applications, such as clustering, association rules, decision trees, and neural network etc.

## Learning Objectives/Measurable Learning Outcomes

### Learning Objectives

By the conclusion of the course of study, students are expected to be able to:

The Students will learn:

- Understand the conceptual steps involved in data mining and learn how to apply it for solving business problems.
- Evaluate the advantages, disadvantages, and applicability of data mining techniques like decision trees, neural networks, regression, and clustering and learn to use these techniques appropriately for business data mining.
- Use various available packages, functionalities and analysis capabilities of state-of-the-art software like Python or IBM SPSS Modeler to analyze business data and perform hands-on

data mining.

- Understand how to using data mining techniques for business academic research.

### **Measurable Learning Outcomes**

Expected learning outcomes for students include:

- 1) To recognize and engage with key concepts, algorithms and theories serving the study of related data mining.
- 2) To correctly associate business and research problems with data mining tasks
- 3) To conduct data mining tasks following CRISP-DM process model.
- 4) To innovatively design data mining solutions to provide empirical evidence to support research hypothesis.
- 5) To identify innovative research insights from the interpretation of data mining tasks

### **Main Contents**

1	Introduction to Data Mining
2	Probability
3	Basic Concepts of learning theory
4	Support vector machine
5	Neural Network and Decision Tree
6	Unsupervised learning and Latent variable models
7	Experiments
8	Group presentation

### **Student Evaluation**

1. In-Class Exercises: 2-3 times 30%

2. Home work: 1-2 times 20%
3. Group project: 40%
4. Participation: 10%