



CS 412 Intro. to Data Mining

Chapter 6. Mining Frequent Patterns, Association and Correlations: Basic Concepts and Methods

Jiawei Han, Computer Science, Univ. Illinois at Urbana-Champaign, 2017



Chapter 6: Mining Frequent Patterns, Association and Correlations: Basic Concepts and Methods

การหา pattern ที่ซ่อนใน Data

- ☐ Basic Concepts 
- ☐ Efficient Pattern Mining Methods
- ☐ Pattern Evaluation
- ☐ Summary

Pattern Discovery: Basic Concepts

- ❑ What Is Pattern Discovery? Why Is It Important?
- ❑ Basic Concepts: Frequent Patterns and Association Rules
- ❑ Compressed Representation: Closed Patterns and Max-Patterns

What Is Pattern Discovery?

❑ What are patterns?

Item หลายๆอย่างที่มีมักจะซื้อพร้อมกัน , set of item จะเกิดขึ้นซ้ำๆบ่อยๆ

- ❑ **Patterns:** A set of items, subsequences, or substructures that occur frequently together (or strongly correlated) in a data set
- ❑ Patterns represent **intrinsic** and **important properties** of datasets

❑ **Pattern discovery:** Uncovering patterns from massive data sets

สามารถค้นหาแล้วเจอได้

❑ Motivation examples:

- ❑ What products were often purchased together?
สินค้าอะไรที่คนจะซื้อพร้อมกันเสมอ
- ❑ What are the subsequent purchases after buying an iPad?
หลังจากซื้อiPadไปแล้ว ครั้งหน้าจะกลับมาซื้ออะไร
- ❑ What code segments likely contain copy-and-paste bugs?
มี code ของคนที่ทำก่อนหน้ามาใช้แก้ปัญหาได้ แต่จะมี bug ให้ดูตัวแปร
- ❑ What word sequences likely form phrases in this corpus?
มันจะสามารถคาดเดาต่อไปให้ได้