

Data Mining in Applications

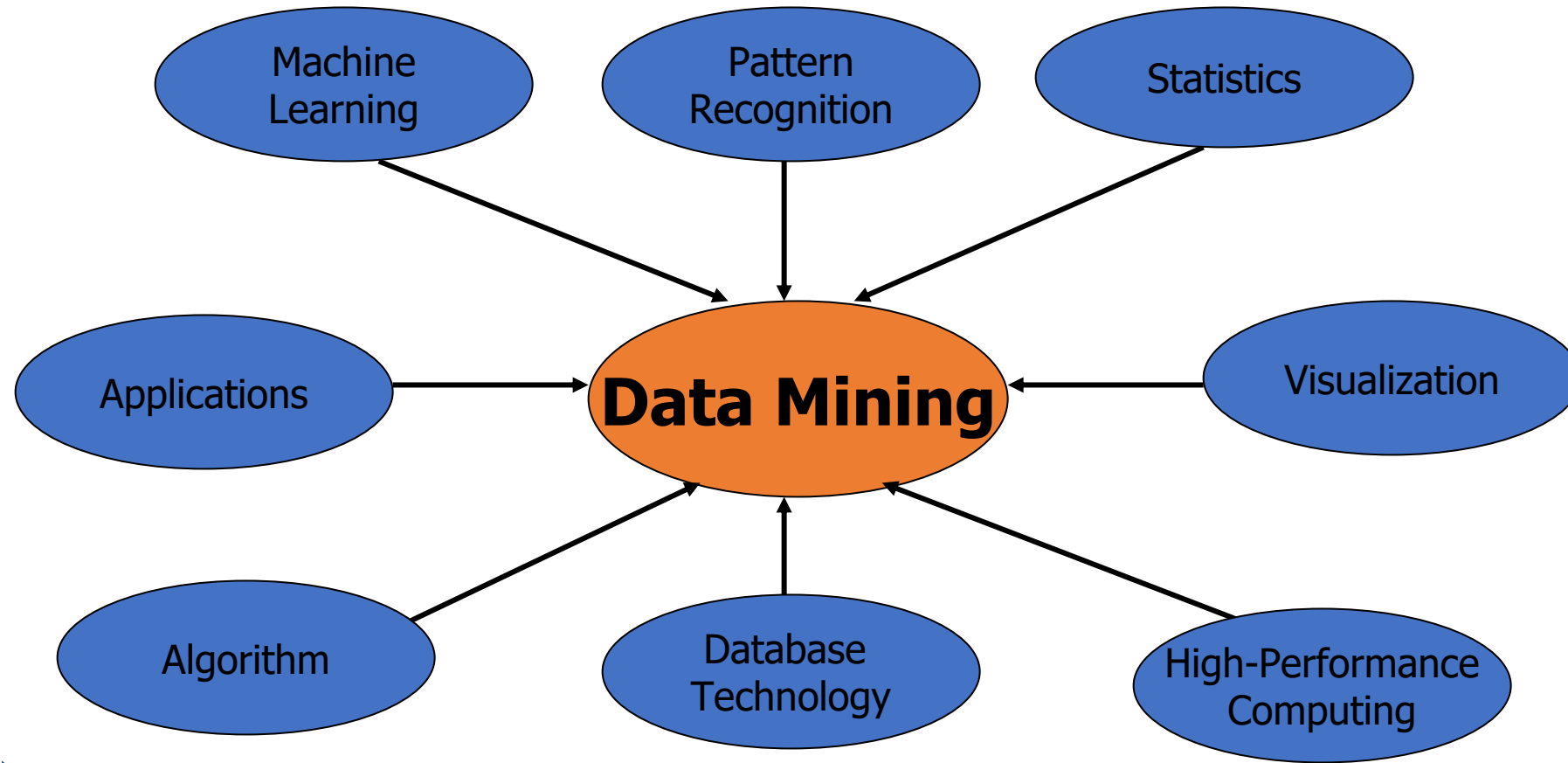
Data Mining 101 Introduction

Tim Pengajar:

- Ratih Ardiati Ningrum, S.Si., M.S., M.Stat.
- Indah Fahmiyah, S.Si., M.Stat.



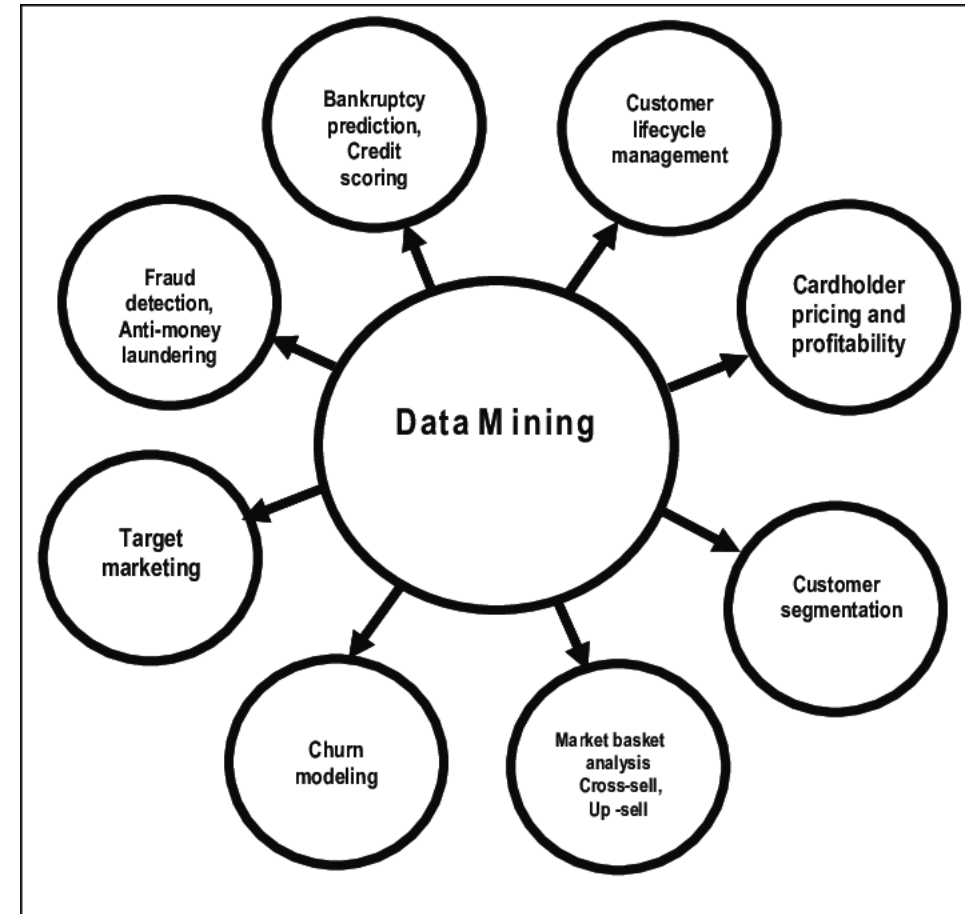
Data Mining: Confluence of Multiple Disciplines



Data mining in applications



Source: Kanchan, B. & Mahale, Kishor. (2016). Study of machine learning algorithms for special disease prediction using principal of component analysis. 5-10. 10.1109/ICGTSPICC.2016.7955260.



Source: Vadlamani, Ravi. (2011). Chapter I Introduction to Banking Technology and Management. Advances in Banking Technology and Management: Impacts of ICT and CRM. 10.4018/978-1-59904-675-4.ch001.



Classification: Application 1

Direct Marketing

- Goal: Reduce cost of mailing by *targeting* a set of consumers likely to buy a new cell-phone product.
- Approach:
 - Use the data for a similar product introduced before.
 - We know which customers decided to buy and which decided otherwise. This *{buy, don't buy}* decision forms the *class attribute*.
 - Collect various demographic, lifestyle, and company-interaction related information about all such customers.
 - Type of business, where they stay, how much they earn, etc.
 - Use this information as input attributes to learn a classifier model.

From [Berry & Linoff] Data Mining Techniques, 1997



Classification: Application 2

Fraud Detection

- Goal: Predict fraudulent cases in credit card transactions.
- Approach:
 - Use credit card transactions and the information on its account-holder as attributes.
 - When does a customer buy, what does he buy, how often he pays on time, etc
 - Label past transactions as fraud or fair transactions. This forms the class attribute.
 - Learn a model for the class of the transactions.
 - Use this model to detect fraud by observing credit card transactions on an account.



Clustering: Application 1

Market Segmentation:

- Goal: subdivide a market into distinct subsets of customers where any subset may conceivably be selected as a market target to be reached with a distinct marketing mix.
- Approach:
 - Collect different attributes of customers based on their geographical and lifestyle related information.
 - Find clusters of similar customers.
 - Measure the clustering quality by observing buying patterns of customers in same cluster vs. those from different clusters.



Clustering: Application 2

Document Clustering:

- Goal: To find groups of documents that are similar to each other based on the important terms appearing in them.
- Approach: To identify frequently occurring terms in each document. Form a similarity measure based on the frequencies of different terms. Use it to cluster.
- Gain: Information Retrieval can utilize the clusters to relate a new document or search term to clustered documents.



Association Rule Discovery: Application 1

Marketing and Sales Promotion:

- Let the rule discovered be
 $\{Bagels, \dots\} \rightarrow \{Potato\ Chips\}$
- Potato Chips as consequent => Can be used to determine what should be done to boost its sales.
- Bagels in the antecedent => Can be used to see which products would be affected if the store discontinues selling bagels.
- Bagels in antecedent and Potato chips in consequent => Can be used to see what products should be sold with Bagels to promote sale of Potato chips!



Terima
Kasih!

