

Workshop

Data Mining and Data Validation



What is Data Mining?

- | Progress in digital data acquisition and storage technology has resulted in the growth of **huge databases**
- | Interest has grown in the possibility of **extracting valuable information** from these databases

Data mining is the automated analysis of large, observational data sets to find unsuspected relationships and to summarize the data in ways that are both understandable and useful to the data owner

- | The relationships and summaries derived through data mining are often referred to as **models** or **patterns**

Data Mining Tasks

| Description

- Find **patterns** and **relations** that meaningfully describe the data (e.g. causal relations)

| Prediction

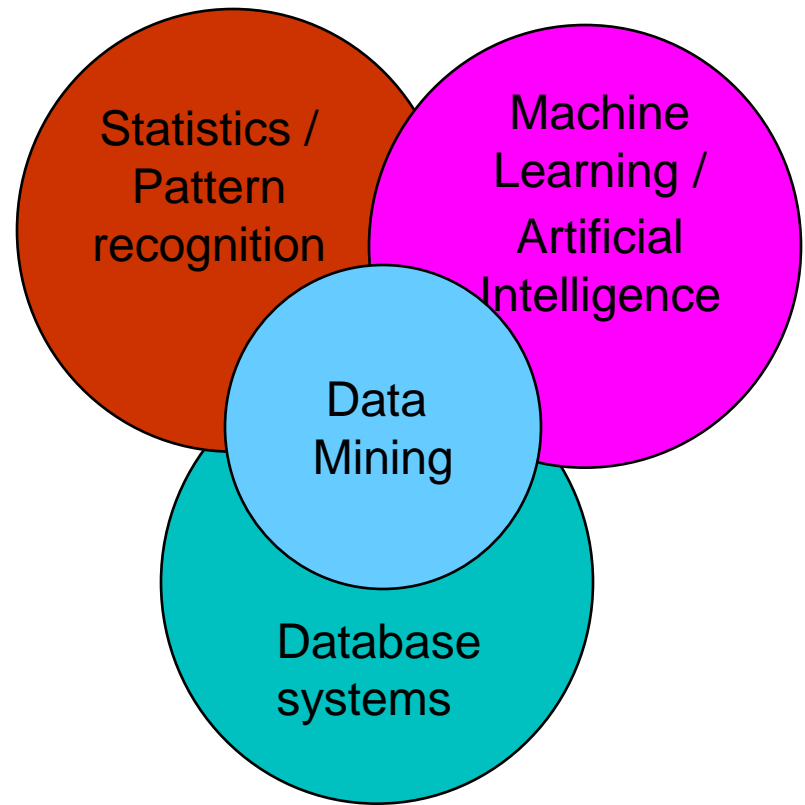
- Construct a model to foretell **values of one variable** based on the values of other variables

⇒ Not always a clear-cut distinction

⇒ Generalizability is always essential

Origins of Data Mining

- | Draws ideas from machine learning, artificial intelligence, pattern recognition, statistics, and database systems
- | **Eclectic approach:** use whatever method is useful



- **Clinical Data Mining**

Dr. Dirk Hüske-Kraus

Clinical Director, Clinical Transformation and Education, Philips Patient Care and Clinical Informatics, Böblingen, Germany

- **Can intracranial hypertension after traumatic brain injury be predicted?**

Prof. Dr.med. Geert Meyfroidt

Intensivist-Anesthesiologist, Associate Professor of Medicine, KU Leuven, Belgium

- **Discussion**