

Information Technology Department
College Of Computer
Qassim University



Course Code: IT 647

Course Title: Data Mining and Business Intelligence

Assignment – 01

Submitted to:

DR. MONA ALSALEEM

Submitted by:

Name: Md. Shakil Hossain

محمد شاكيل حسين

University ID: 472118102

Submission Date: 10/02/2026

Introduction to Data Mining

Data mining is defined as an activity that involves the ability to identify significant patterns, trends, as well as knowledge within massive amounts of data. It can be explained as knowledge discovery in data and is used to solve the contemporary problem of being immersed in data and hungry of knowledge. The field has developed with the development of the database technology, starting with the rudimentary data collection systems of the 1960s to the current data environments of the petabyte level.

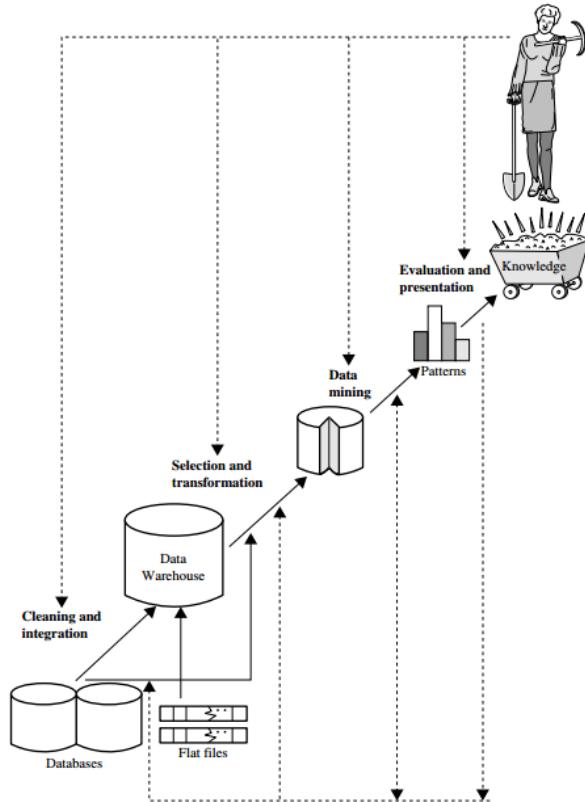


Figure: Data mining as a step in the process of knowledge discovery.

Knowledge Discovery Process (KDD)

Data mining is a fundamental phase in the bigger Knowledge Discovery from Data (KDD) process which consists of:

- Data Cleaning - Elimination of noise and inconsistencies.
- Data Integration - Integration of several data sources.
- Data Selection - Obtaining data within analysis.
- Data Transformation - Mining-friendly data conversion.
- Data Mining - Intelligent ways of extracting patterns.
- Pattern Evaluation - Finding useful and interesting patterns.
- Knowledge Presentation - Presenting results visually.

Types of Data Sources

- Relational Databases- Attributed and topped structured tables.
- Data Warehouses - Multi dimensional data cubes in the form of integrated, subject-oriented repositories.
- Transactional Databases - Accounts of the transactions.
- Complex Database - This type of data storage supports streams of data, graphs, social networks, spatial-temporal data, multimedia, and web data.

Functionalities of Data Mining

- Characterization & Discrimination- Summarizing and comparing data classes.
- Frequent Pattern Mining - Identifying repeat associations.
- Classification & Regression - Predictive models that are categorical or numeric.
- Clustering - This refers to grouping similar objects, but without the use of labels.
- Outlier Analysis - Determining anomalies that are not in the same pattern.

Technologies

- Statistics
- Machine Learning
- Database Systems
- Visualization

Applications

- Business Intelligence Sales forecast, fraud detection, customer segmentation.
- Web Search engines: crawling, indexing, ranking and personalization.
- Medical - Genomic forecasting, predictive research, health.
- Engineering - Predictive maintenance, optimization of the process.
- Recommender Systems & Social Network Analysis - Pattern recognition in the behavior of users.

Conclusion

Data mining is a crucial science in the data-driven age as it allows organizations to use raw data as an actionable data. It is the foundation of the business intelligence and it aids in decision-making in industries. The importance of data mining will only increase with the rising data volumes to find out the hidden knowledge.

References

Text Book: “Data Mining: Concepts and Techniques”, Third Edition