

# Data Mining

Presented by: NPower Group 5

**Yusuf Hekimoglu - Laher Pereira - Komal Patel - Kownain Shahzad  
- Zubair Mohammed Iftekhar - Kenil Bhadani - Leon Wainaina**

# What is Data Mining?

- Data mining is the process of sorting through large data sets to identify patterns and relationships that can help solve business problems through data analysis.
- Data mining is a key part of data analytics and one of the core disciplines in data science.
- Data mining is more useful today due to the growth of big data and data warehousing.



# History and Origins

- Data analysis technologies emerged in the late 1980s and early 1990s.
- The term "Data Mining" was coined by Michael Lovell in 1983, gaining popularity by 1995.
- The inaugural International Conference on Knowledge Discovery and Data Mining was held in 1995.
- ACM SIGKDD conference was established in 1999, becoming pivotal.
- Academic journals like Data Mining and Knowledge Discovery (1997) and the American Journal of Data Mining and Knowledge Discovery (2016) contribute significantly.
- Data mining and process mining enhance organizational performance with distinct methodologies.

# Why is Data Mining important?

- Data mining is a crucial component of successful analytics initiatives in organizations.
- Effective data mining aids in various aspects of planning business strategies and managing operations.
- Data mining supports fraud detection, risk management, cybersecurity planning and many other critical business use cases.
- It plays an important role in other areas, including healthcare, government, scientific research, mathematics and sports.

# Data Mining Process

The core elements of data mining include machine learning and statistical analysis, along with data management tasks done to prepare data for analysis.

01. Data Gathering

02. Data Preparation

03. Data Mining

04. Data Analysis & Interpretation

# DATA MINING TECHNIQUES

Association Rule Mining

Neural Networks

Regression

Sequence and Path Analysis

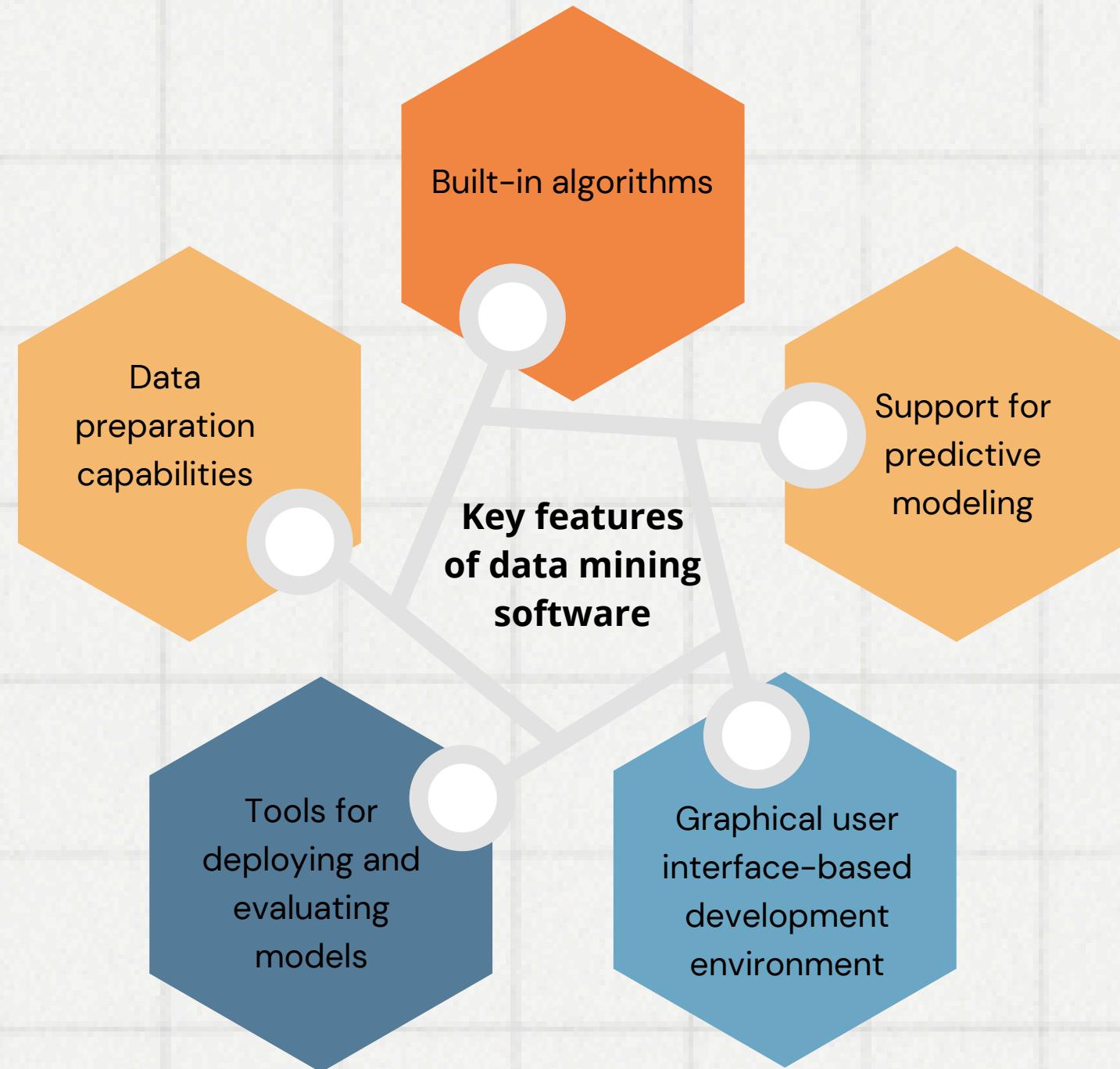
Classification

Decision Trees

Clustering

KNN (K-nearest neighbors)

# Data Mining Software and Tools

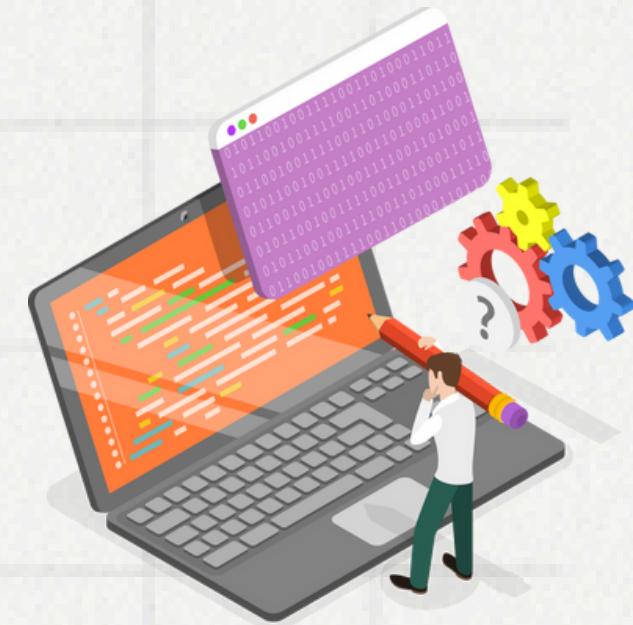


## Data mining tools

ALTERYX, DATAIKU, H2O.AI, IBM, KNIME, MICROSOFT, ORACLE, RAPIDMINER, SAP, SAS INSTITUTE, AND TIBCO SOFTWARE.

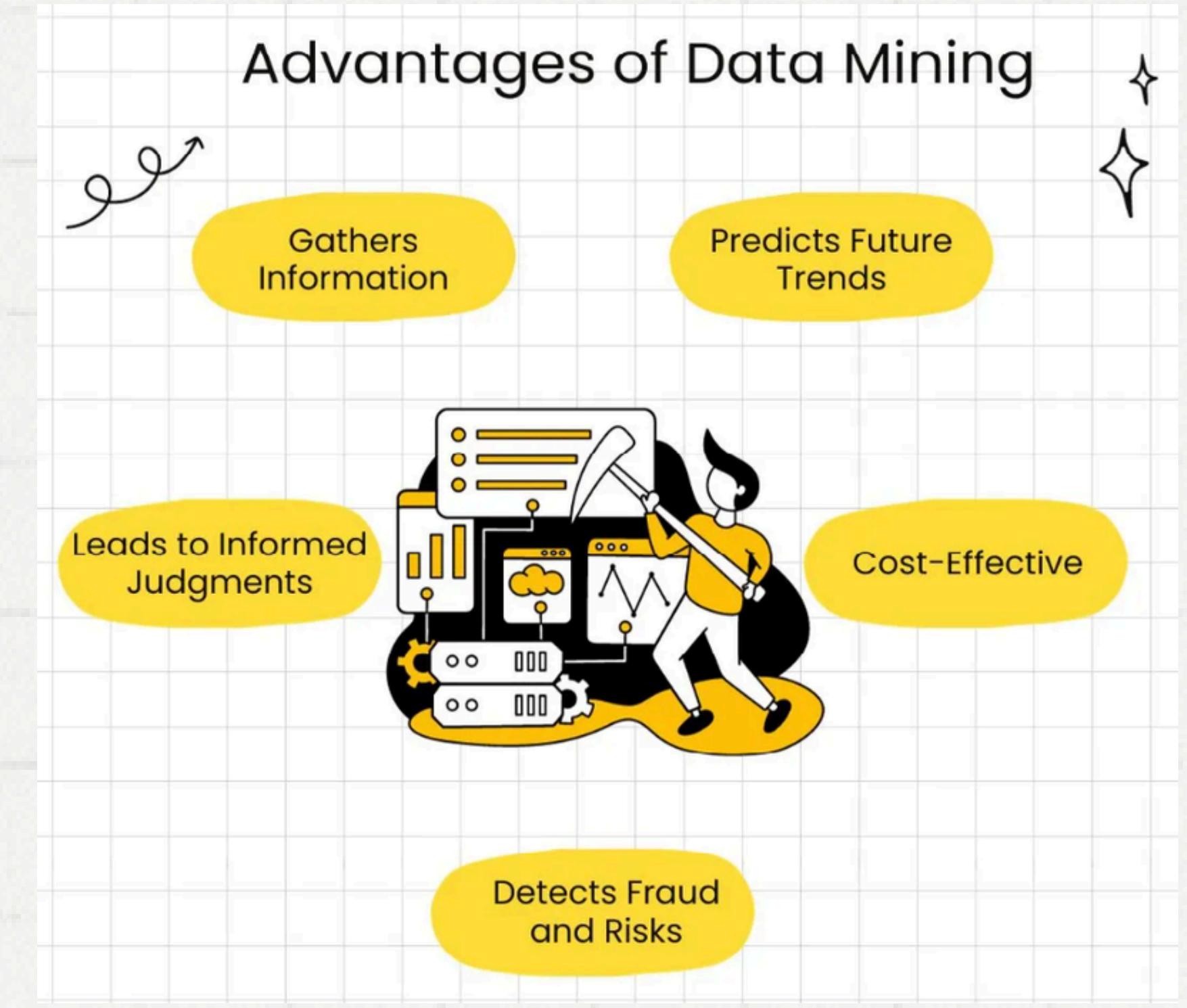
## Free open source technologies for data mining

DATAMELT, ELKI, ORANGE, RATTLE, SCIKIT-LEARN, AND WEKA.



# Benefits of Data Mining

- More effective Marketing and Sales
- Better Customer Service
- Improved Supply Chain Management
- Increased Production Uptime
- Stronger Risk Management
- Lower Costs



## FINANCIAL SERVICES



- FINANCIAL RISK MODELS
- DETECT FRAUDULENT TRANSACTIONS
- VET LOAN
- CREDIT APPLICATIONS

## MANUFACTURING



- UPTIME
- OPERATIONAL EFFICIENCY
- SUPPLY CHAIN PERFORMANCE
- PRODUCT SAFETY

## RETAIL



- RECOMMENDATION FOR VISITORS
- INVENTORY
- SCM (SUPPLY CHAIN MANAGEMENT)
- MARKETING
- CUSTOMER BEHAVIOUR

## INSURANCE



- RISK MODELING
- APPROVE POLICY APPLICATIONS
- PRICING THE POLICIES
- MANAGING PROSPECTIVE CUSTOMERS

# Industry Examples

## Data Mining



## HR



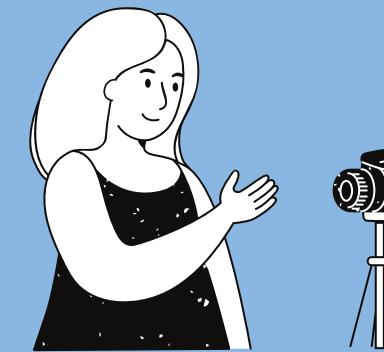
- SORTING APPLICANT DATA
- COLLECTING POSSIBLE CANDIDATES

## ENTERTAINMENT



- PERSONALIZE RECOMMENDATION

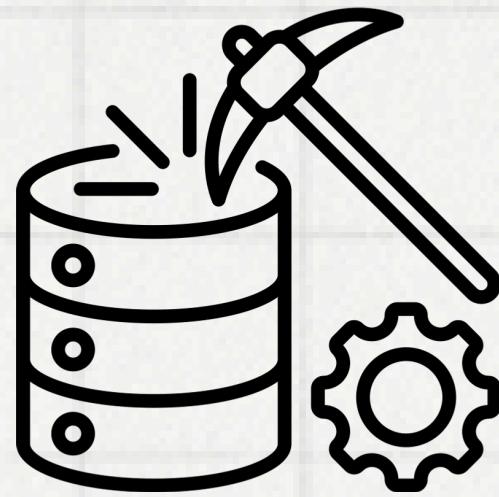
## SOCIAL MEDIA



- GATHERING DATA FROM USERS FOR 2 DIFFERENT REASONS;  
1. ADVERTISING  
2. SOLD TO THIRD PARTIES

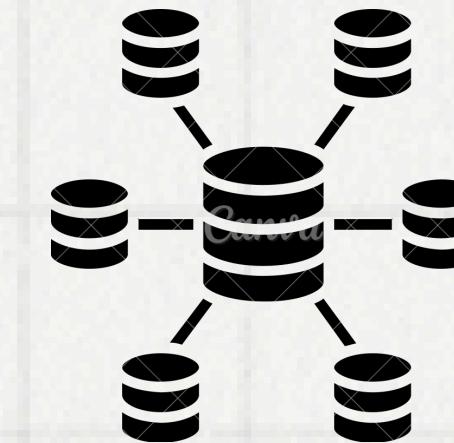
# DATA MINING VS. DATA ANALYTICS AND DATA WAREHOUSING

## DATA MINING AND ANALYTICS



- Aspect of data analytics, automating analysis of large data sets
- Helps discover undetectable and useful information in data science and BI applications
- Utilizes data lakes to store historic and current data
- Data lakes are based on platforms like hadoop, Spark, NoSQL, cloud storage

## DATA WAREHOUSING



- Supports data mining by providing storage for data sets
- Stores historic data in enterprise data warehouse as well as smaller data marts

# Thank You

for  
listing!!

our  
**Reference**

1. <https://www.techtarget.com/searchbusinessanalytics/definition/data-mining#:~:text=Data%20mining%20is%20the%20process,make%20more%2Dinform%20business%20decisions>.