

Notes In Hindi

SUBJECT

Data warehouse and Data science

Notes In Hindi

PREPARED BY

TEAM NOTESINHINDI

Notesinhindi.com

2025 Edition

PRICE- ₹120

Topic	Page No
Data Mining in Hindi	3
Kinds of Data in Hindi	7
Data Mining Functionalities in Hindi	11
Data Reduction in Hindi	23
Data Cleaning Techniques in Hindi	29
Data Warehouse in Hindi	32
Data Cube in Hindi	37
OLAP in Hindi.....	41
Overview of Data Cube in Data Mining in Hindi	46
Correlations in Hindi	52
Frequent Patterns in Hindi.....	59
Association Mining in Hindi	63
Bayesian Classification Methods in Hindi.....	68
Tools and Libraries for Bayesian Classification in Hindi	71
Rule-Based Classification in Hindi.....	71
Ensemble Methods in Hindi	75
Neural Networks in Hindi	81
Support Vector Machines in Hindi.....	84
Cluster Analysis in Hindi	90
Clustering Structures in Data Mining in Hindi.....	94
Hierarchical Methods in Hindi	101
Density-Based Methods in Hindi	103
Outlier Analysis in Hindi	107
Web Mining in Hindi	113

Data Mining in Hindi

Data Mining एक ऐसा आसान सा process है जिसमें बड़ी मात्रा में मौजूद data के अंदर से useful information और hidden patterns निकाले जाते हैं। आज के समय में हर जगह data बहुत तेज़ी से बढ़ रहा है—जैसे shopping apps, banks, social media, hospitals, schools आदि। ऐसे में Data Mining की मदद से इस बड़े data को समझना और उससे meaningful जानकारी निकालना बहुत आसान हो जाता है।

Data Mining in Hindi का simple मतलब होता है—data me se important knowledge निकालना। Data Mining का main goal है data को useful बनाना ताकि कंपनियाँ और users बेहतर decision ले सकें। इस process में computer algorithms और AI techniques use की जाती हैं जो data को समझकर उसमें मौजूद trends, patterns और relations बताती हैं।

Data Mining क्यों जरूरी है?

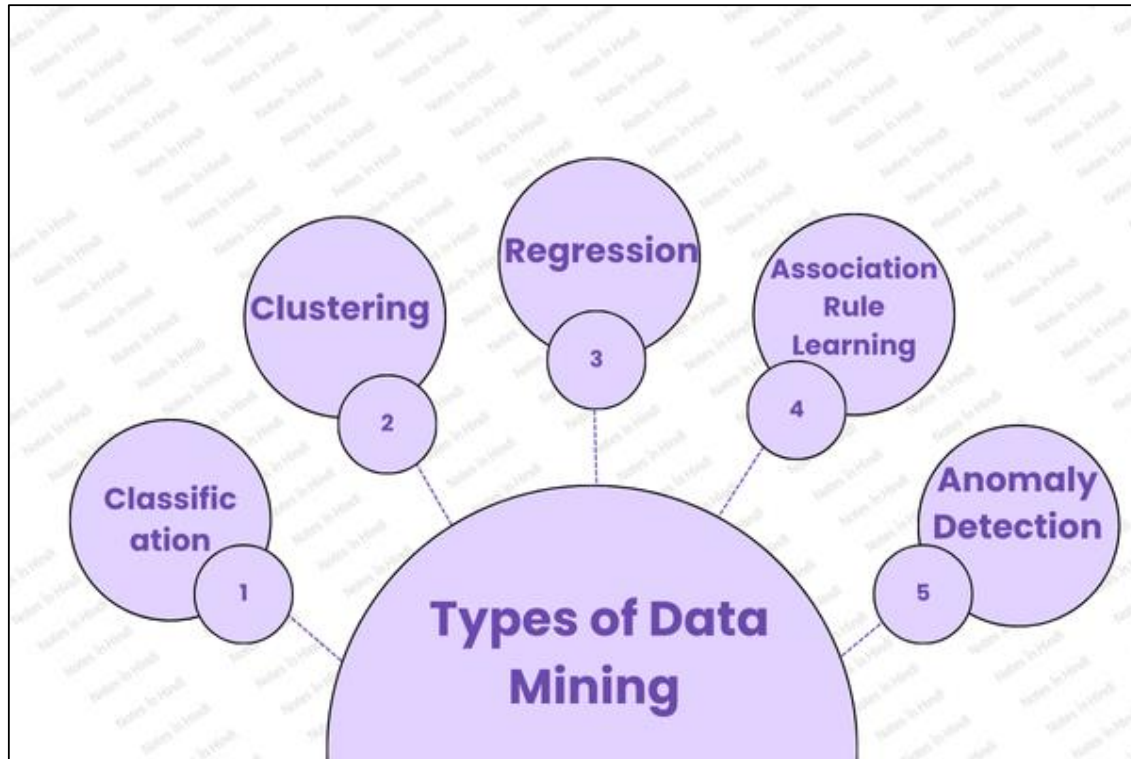
- Data बहुत तेज़ी से बढ़ रहा है, इसलिए उसे समझने के लिए Data Mining की जरूरत होती है।
- Business को future planning करने में help मिलती है।
- Customers की जरूरतों को समझने का मौका मिलता है।
- Hidden patterns मिलते हैं जो normal तरीके से नहीं दिखते।

Data Mining के Important Elements

- **Data Collection** – data को कई sources से collect किया जाता है।
- **Data Cleaning** – गलत, incomplete या duplicate data को हटाया जाता है।
- **Pattern Identification** – useful pattern और relations खोजे जाते हैं।
- **Knowledge Presentation** – result को understandable तरीके से दिखाया जाता है।

Types of Data Mining in Hindi

Data Mining के कई types होते हैं और हर type data को अलग-अलग तरीके से analyze करता है। Types of Data Mining in Hindi समझने से हमें यह पता चलता है कि कौन सा method किस जगह काम आता है।



1. Classification

- Data को pre-defined categories में divide करता है।
- जैसे: email को “spam” और “not spam” में classify करना।

2. Clustering

- Data को groups में divide करता है जहाँ हर group में similar items होते हैं।
- जैसे customers को buying behavior के आधार पर group करना।

3. Regression

- Values predict करने में use होता है, जैसे price prediction।
- Continuous output देता है।

4. Association Rule Learning

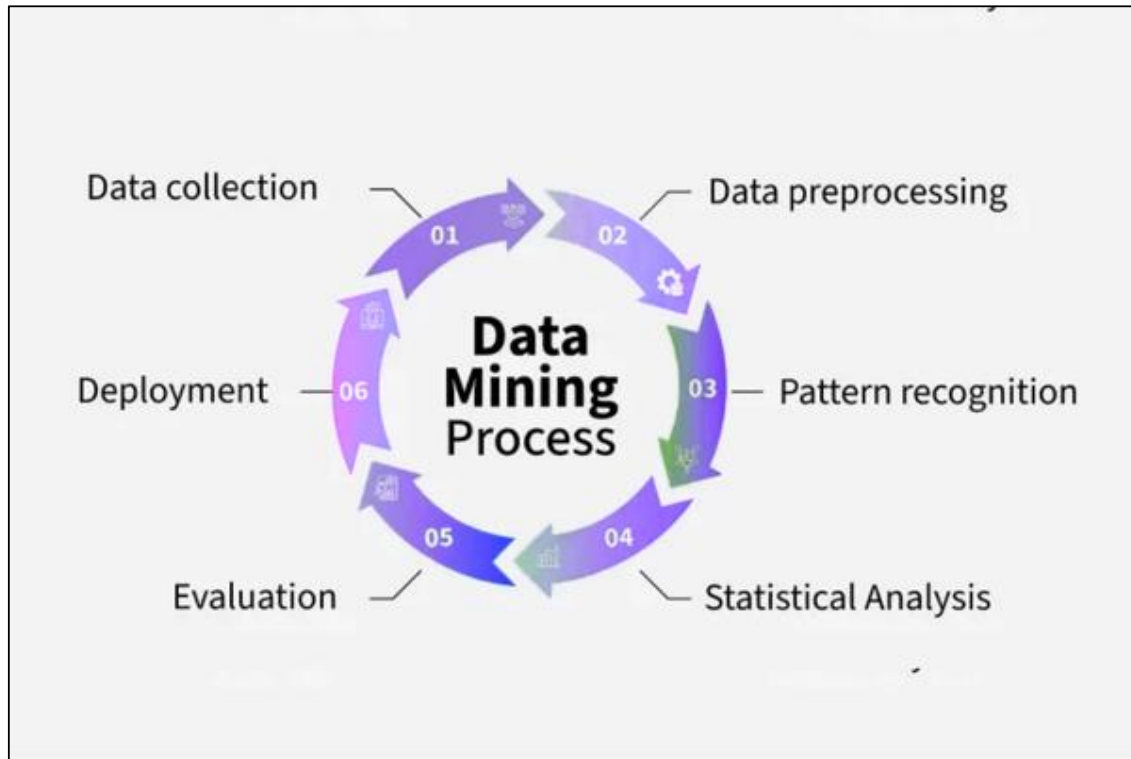
- Market Basket Analysis में उपयोग होता है।
- जैसे: जो लोग bread खरीदते हैं वे butter भी खरीदते हैं।

5. Anomaly Detection

- Data में unusual या abnormal patterns ढूँढता है।
- जैसे credit card fraud detection।

Process of Data Mining in Hindi

Process of Data Mining in Hindi एक step-by-step तरीका है जिससे data को useful knowledge में बदला जाता है। यह process simple लेकिन logical steps में divide होता है जिससे पूरे data mining work को आसानी से समझा जा सके।



Data Mining Process स्टेप्स

1. Data Cleaning

- गलत data, missing values और duplicate entries को हटाया जाता है।
- Clean data accurate results देता है।

2. Data Integration

- Data को अलग-अलग sources से combine करके एक जगह लाया जाता है।
- जैसे SQL database + Excel sheet data।

3. Data Selection

- Analysis के लिए useful data को select किया जाता है।
- Irrelevant data को remove किया जाता है।

4. Data Transformation

- Data को format किया जाता है ताकि algorithm आसानी से समझ सके।

- Normalization, aggregation आदि techniques use होती हैं।

5. Data Mining

- Main step जिसमें machine learning algorithms apply किए जाते हैं।
- Patterns, relations और predictions generate होते हैं।

6. Pattern Evaluation

- जो patterns मिले हैं उनमें से useful patterns चुने जाते हैं।
- Useless patterns remove किए जाते हैं।

7. Knowledge Presentation

- Results को charts, graphs, tables में present किया जाता है।
- Decision makers के लिए findings easy बनाई जाती हैं।

Classification in Data Mining in Hindi

Classification in Data Mining in Hindi का मतलब है data को predefined categories में classify करना। यह Data Mining का सबसे common और important part है।

Classification future prediction और decision making के लिए बहुत useful है। Machine learning models जैसे Decision Tree, Naive Bayes, Neural Networks आदि classification में use होते हैं।

Classification कैसे काम करता है?

- पहले training dataset लिया जाता है जिसमें हर data already labelled होता है।
- Model dataset को पढ़कर rules सीखता है।
- फिर new data पर apply करके उसका class predict करता है।

Classification के Examples

- Email को Spam / Not Spam classify करना।
- Loan application को Approve / Reject करना।
- Disease detection – Patient “healthy” है या “infected”।

Classification के Popular Algorithms

ALGORITHM	USE
Decision Tree	Simple classification models बनाने में
Naive Bayes	Email spam detection में popular

KNN

Similar data points के आधार पर classify

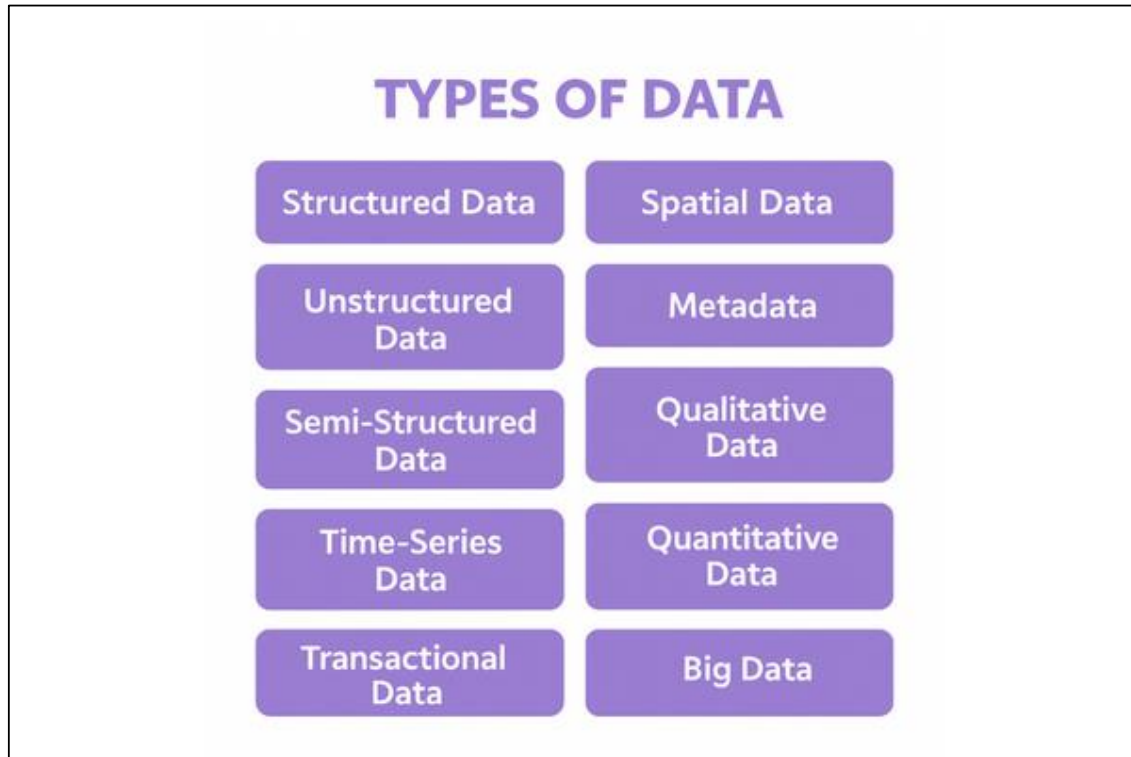
Kinds of Data in Hindi

Data आज की digital दुनिया का सबसे important हिस्सा है। किसी भी सिस्टम, mobile app, website या business में Data ही वो base होता है जिस पर फैसले लिए जाते हैं। Kinds of Data in Hindi को समझना इसलिए जरूरी है ताकि आपको साफ पता चल सके कि कौन-सा Data किस काम में आता है और उसकी खासियत क्या होती है। नीचे हम सभी kinds of data को बहुत आसान भाषा में समझेंगे।

Data क्या होता है?

Data किसी भी प्रकार की जानकारी होती है जिसे हम store, manage और use कर सकते हैं। जैसे- numbers, text, images, videos, records आदि। Business, education, research और technology में अलग-अलग types का data इस्तेमाल होता है। इसी वजह से Data को अलग-अलग categories में divide किया जाता है।

Types of Data



- Structured Data
- Unstructured Data
- Semi-Structured Data
- Time-Series Data
- Transactional Data
- Spatial Data
- Metadata

- Qualitative Data
- Quantitative Data
- Big Data

Structured Data in Hindi

Structured Data वो Data होता है जो एक fixed format में store किया जाता है। यह पूरी तरह organized होता है और इसे search करना, update करना या analyze करना बहुत आसान होता है। Structured Data in Hindi mostly rows और columns के रूप में database में रखा जाता है जैसे कि MySQL, SQL Server, Oracle आदि।

Structured Data की खासियतें

- Rows और Columns में store होता है, इसलिए समझना आसान होता है।
- Data का format हमेशा fixed रहता है।
- SQL जैसी भाषा से आसानी से manage किया जा सकता है।

Structured Data के examples

- Student records (Name, Roll Number, Class)
- Bank account records
- Employee records

Unstructured Data in Hindi

Unstructured Data वो Data होता है जिसका कोई fixed format नहीं होता है। इसे समझना और manage करना थोड़ा मुश्किल होता है क्योंकि यह किसी defined structure में नहीं रहता। Unstructured Data in Hindi में images, audio, videos, emails, social media posts जैसी चीजें आती हैं।

Unstructured Data की खासियतें

- कोई specific structure नहीं होता।
- Data बहुत बड़ा और complex होता है।
- AI और Machine Learning की मदद से analyze किया जाता है।

Unstructured Data के examples

- YouTube videos
- Photos
- WhatsApp messages
- Social media comments

Semi-Structured Data in Hindi

Semi-Structured Data ऐसा Data होता है जिसमें थोड़ा structure होता है और थोड़ा unstructured रहता है। इस Data में tags, markers या की-values होती हैं जिससे इसे समझना आसान होता है। Semi-Structured Data in Hindi जैसे JSON, XML आदि।

Semi-Structured Data की खासियतें

- पूरी तरह structured नहीं लेकिन identifiable pattern होता है।
- NoSQL databases इसे स्टोर करने में बेहतर होते हैं।
- Fast processing और flexible format।

Semi-Structured Data examples

- JSON Files
- XML Files
- Emails (subject, body, sender mixture)

Time-Series Data in Hindi

Time-Series Data वो Data होता है जिसे time के sequence में store किया जाता है। यह Data समय के हिसाब से बदलता रहता है। Time-Series Data in Hindi mostly finance, stock market, weather forecasting में इस्तेमाल होता है।

Time-Series Data की खासियतें

- Time के order में store होता है।
- Trend analysis में useful होता है।
- Prediction (Forecasting) में काम आता है।

Time-Series Data examples

- Stock market price
- Temperature records
- Daily sales data

Transactional Data in Hindi

Transactional Data वो Data होता है जो किसी transaction के दौरान generate होता है। यह Data real समय में change होता है। Transactional Data in Hindi जैसे online shopping records, payment records आदि।

Transactional Data की खासियतें

- हर छोटे-बड़े transaction का detail contain करता है।
- Fast processing के लिए optimized होता है।
- Financial और E-commerce systems में सबसे ज्यादा use होता है।

Transactional Data examples

- UPI payment records
- Order history
- ATM transactions

Spatial Data in Hindi

Spatial Data उस Data को कहते हैं जो किसी location या जगह से related होता है। Spatial Data in Hindi maps, GPS, Google Maps tracking में बहुत इस्तेमाल होता है।

Spatial Data की खासियतें

- Location, area, distance जैसी जानकारी store करता है।
- GIS systems में use होता है।
- Navigation और tracking में helpful।

Spatial Data examples

- Google Maps routes
- City boundary maps
- GPS coordinates

Metadata in Hindi

Metadata वो Data होता है जो किसी दूसरे Data की जानकारी देता है। Metadata in Hindi को simple शब्दों में "Data about Data" कहा जाता है।

Metadata की खासियतें

- किसी file का size, date, creator आदि बताता है।
- Data को organize और search करना आसान बनाता है।
- Digital libraries में बहुत important।

Metadata examples

- Photo details (resolution, camera name)
- Document properties
- File creation date

Qualitative Data in Hindi

Qualitative Data वो Data होता है जो किसी चीज़ की quality या nature को describe करता है। यह non-numerical होता है। Qualitative Data in Hindi mostly feelings, behaviours और descriptions होता है।

Qualitative Data features

- Numbers में नहीं होता, words में होता है।
- Human behaviour समझने में useful।
- Subjective nature का होता है।

Qualitative Data examples

- Feedback
- Opinions
- Emotions

Quantitative Data in Hindi