

DATA MINING VS DATA SCIENCE

DATA SCIENCE FOUNDATION COURSE BY ENAMUL HAQUE

CHAPTER 4, LECTURE 14

ডেটা মাইনিং বনাম ডেটা সায়েন্স

ডেটা মাইনিং এবং ডেটা সায়েন্সে, তাদের মধ্যে পার্থক্য খুঁজে পাওয়া কঠিন হতে পারে।

Choosing *between two very similar fields* makes it difficult to determine which one is best for you. In data mining and data science, it can be difficult even to find differences between them. Therefore, in such areas, it becomes even more challenging to decide before.

What is data mining?
ডেটা মাইনিং কি?

Key differences between data mining and data science
মূল পার্থক্য ডেটা মাইনিং এবং ডেটা সায়েন্স

What they are

Focus

Professionals in
this field

Data type

Goal

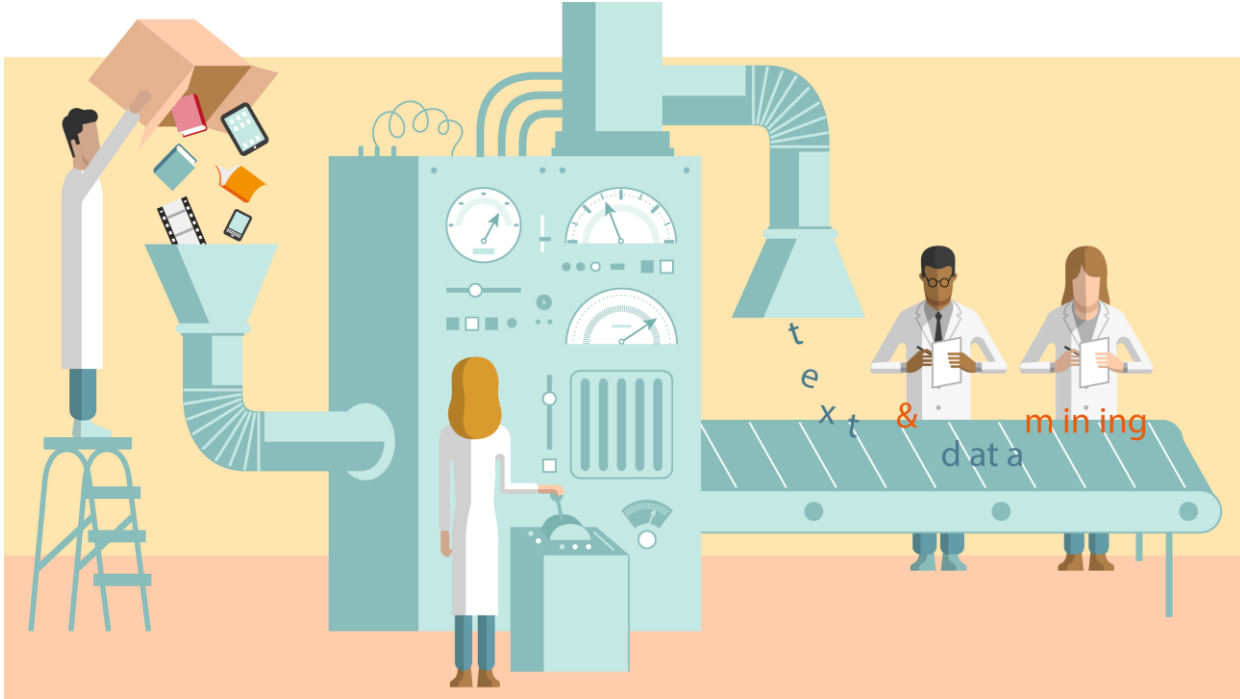
ডেটা মাইনিং কি?

ডেটা মাইনিং একটি পক্রিয়া যা কাঁচা ডেটাকে দরকারী এবং আরও বোধগম্য তথ্যে রূপান্তর করার জন্য ব্যবহার করা হয়।

Data mining is a term coined to transform raw data into useful and more understandable information.

Many companies use data mining to find and analyse patterns in their marketing, revenue, expense, and sales data. This type of information is then used to make important decisions about marketing strategies and financial management.

অনেক কোম্পানি তাদের মার্কেটিং, রাজস্ব, ব্যয় এবং বিক্রয় ডেটার নিদর্শন খুঁজে বের করতে এবং বিশ্লেষণ করতে ডেটা মাইনিং ব্যবহার করে। এই ধরনের তথ্য তারপর বিপণন কৌশল এবং আর্থিক ব্যবস্থাপনা সম্পর্কে গুরুত্বপূর্ণ সিদ্ধান্ত নিতে ব্যবহৃত হয়।



ডেটা মাইনিং কি?

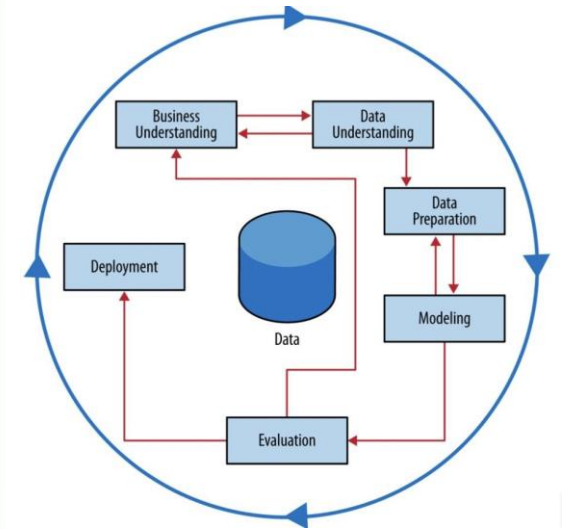
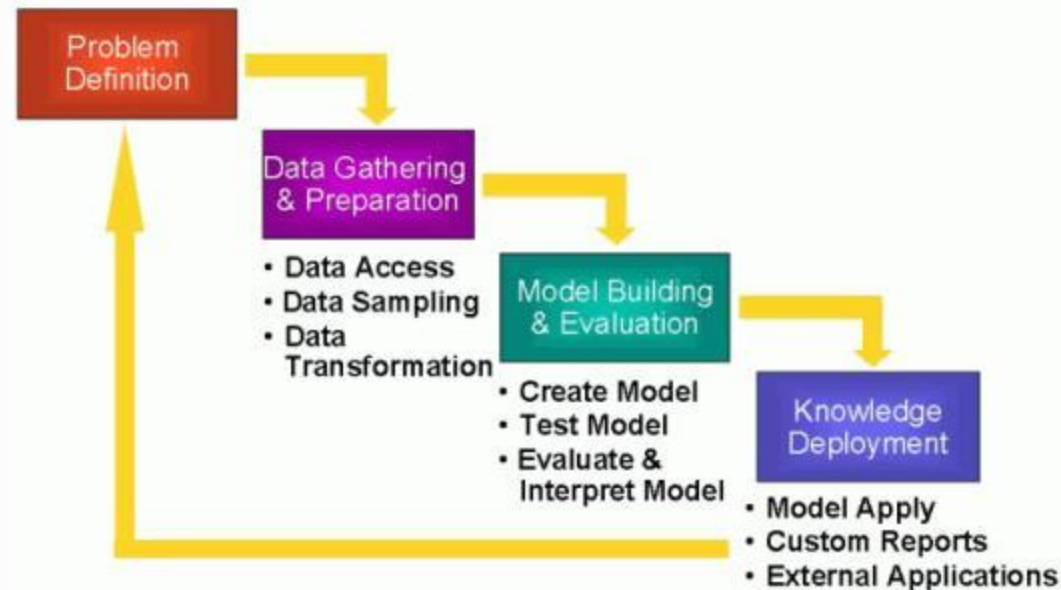
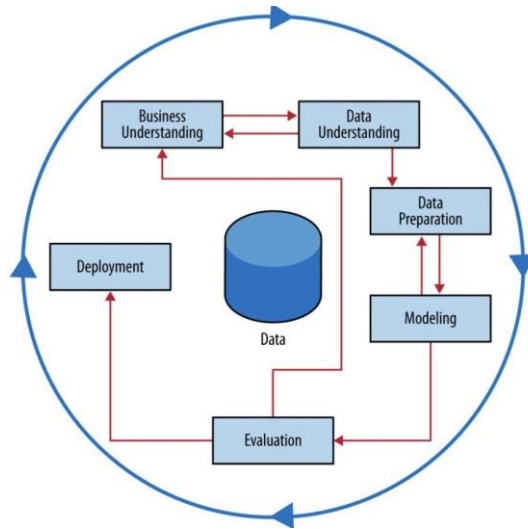
ডেটা মাইনিং একটি পক্রিয়া যা কাঁচা ডেটাকে দরকারী এবং আরও বোধগম্য তথ্যে রূপান্তর করার জন্য ব্যবহার করা হয়।

Data mining is a term coined to transform raw data into useful and more understandable information.

The data mining process begins with companies collecting data and uploading it to the data warehouses where it is stored and managed. Often it also loads into the cloud for storage. Business analysts are then involved in studying the data and deciding how best to organise and display it. This information is transferred to software designed to sort the data and then sorted and displayed as a diagram, graph, or table.

ডেটা মাইনিং প্রক্রিয়া শুরু হয় কোম্পানিগুলো ডেটা সংগ্রহ করে এবং ডেটা ওয়ারহাউসে আপলোড করে যেখানে এটি সংরক্ষণ এবং পরিচালিত হয়। প্রায়শই এটি সঞ্চয়ের জন্য ক্লাউডে লোড হয়। ব্যবসায় বিশ্লেষকরা তখন ডেটা অধ্যয়ন করে এবং কীভাবে এটি সংগঠিত এবং প্রদর্শন করা যায় তা নির্ধারণ করেন। এই তথ্য ডেটা সাজানোর জন্য ডিজাইন করা সফটওয়্যারে স্থানান্তর করা হয় এবং তারপর সাজানো এবং একটি ডায়াগ্রাম, গ্রাফ বা টেবিল হিসাবে প্রদর্শিত হয়।

The Data Mining Process



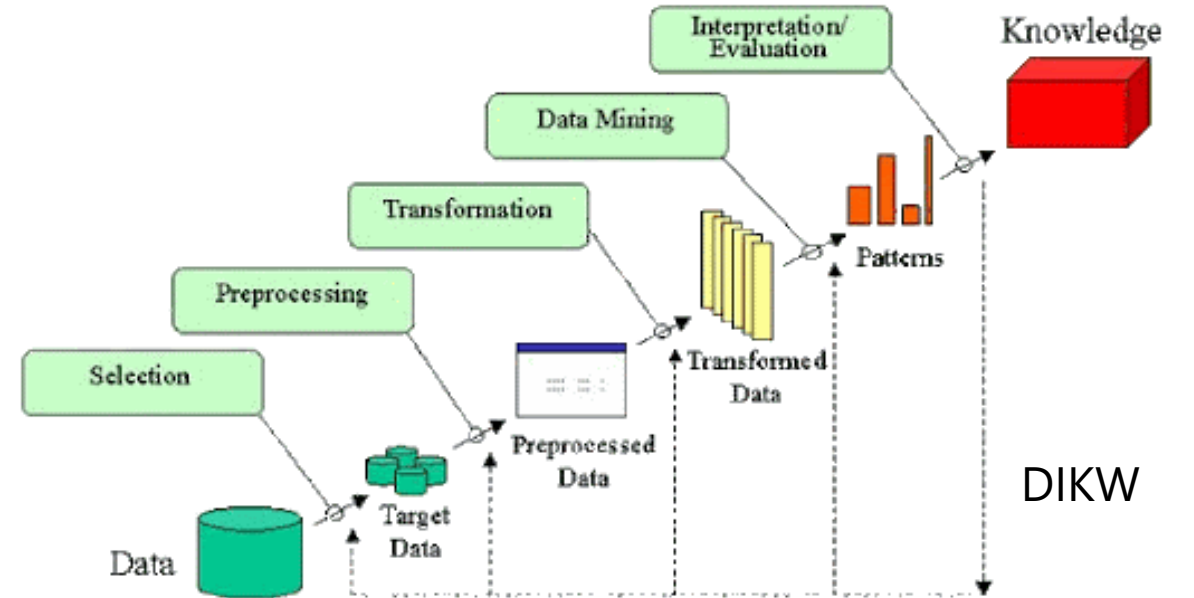
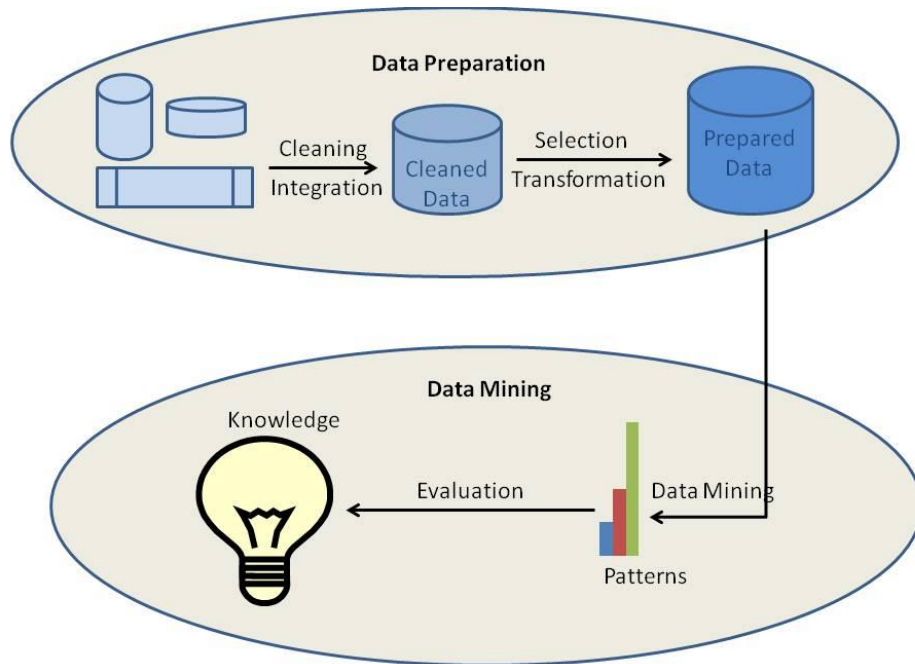
ডেটা মাইনিং কি?

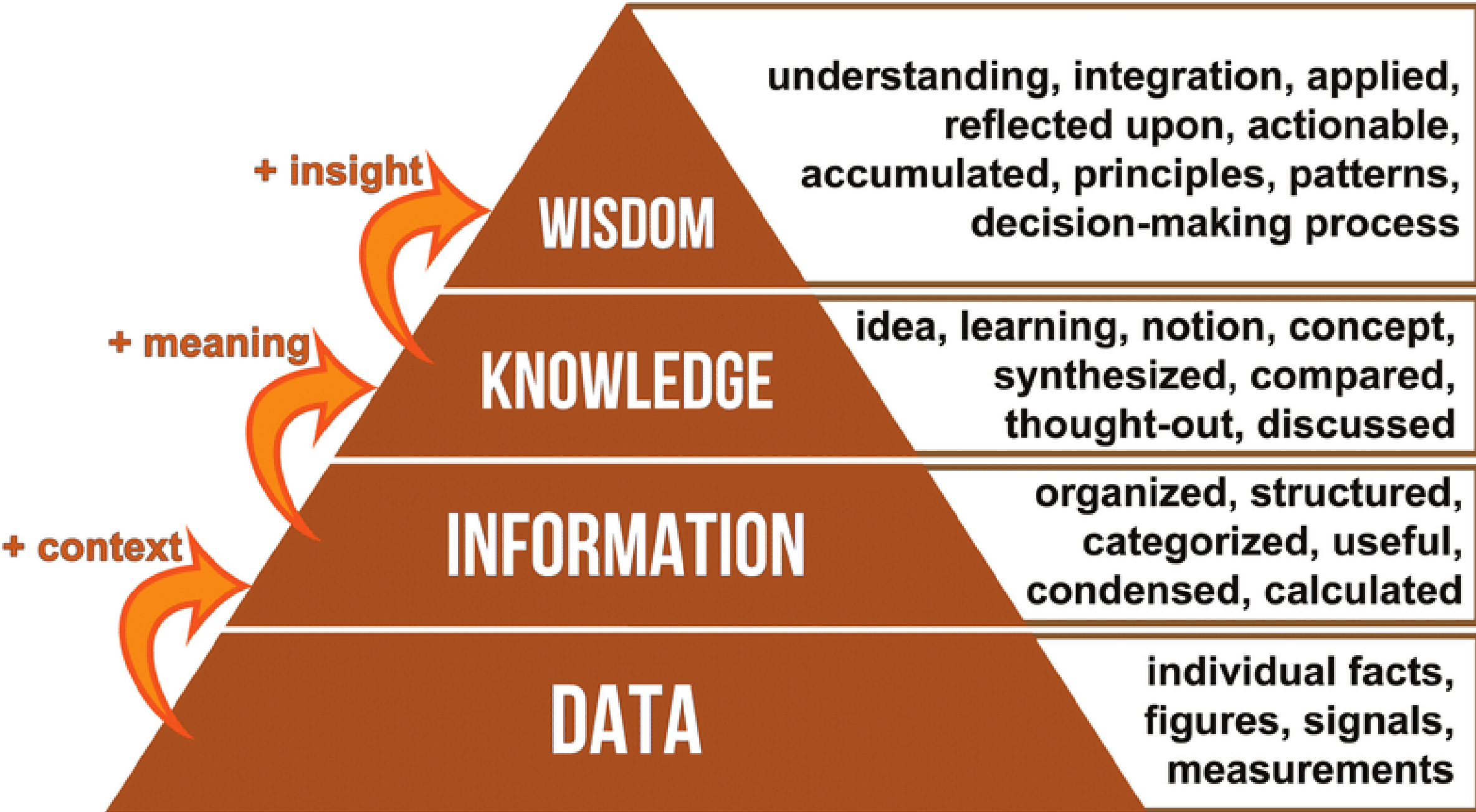
ডেটা মাইনিং একটি পক্রিয়া যা কাঁচা ডেটাকে দরকারী এবং আরও বোধগম্য তথ্যে রূপান্তর করার জন্য ব্যবহার করা হয়।

Data mining is a term coined to transform raw data into useful and more understandable information.

Data analysts do quite a lot of work to find these useful ideas. Typically, they are instructed to search for suitable data sets and variables to study, collect both structured and unstructured data, analyse and interpret data, and explain their findings to stakeholders in an understandable way

ডেটা বিশ্লেষকরা এই দরকারী ধারণাগুলি খুঁজে পেতে বেশ কাজ করেন। সাধারণত, তাদের অধ্যয়ন করার জন্য উপযুক্ত ডেটা সেট এবং ভেরিয়েবল অনুসন্ধান করতে, কাঠামোগত এবং অসংগঠিত উভয় ডেটা সংগ্রহ করতে, ডেটা বিশ্লেষণ এবং ব্যাখ্যা করতে এবং তাদের ফলাফলগুলি স্টেকহোল্ডারদের বোধগম্য উপায়ে ব্যাখ্যা করার নির্দেশ দেওয়া হয়।



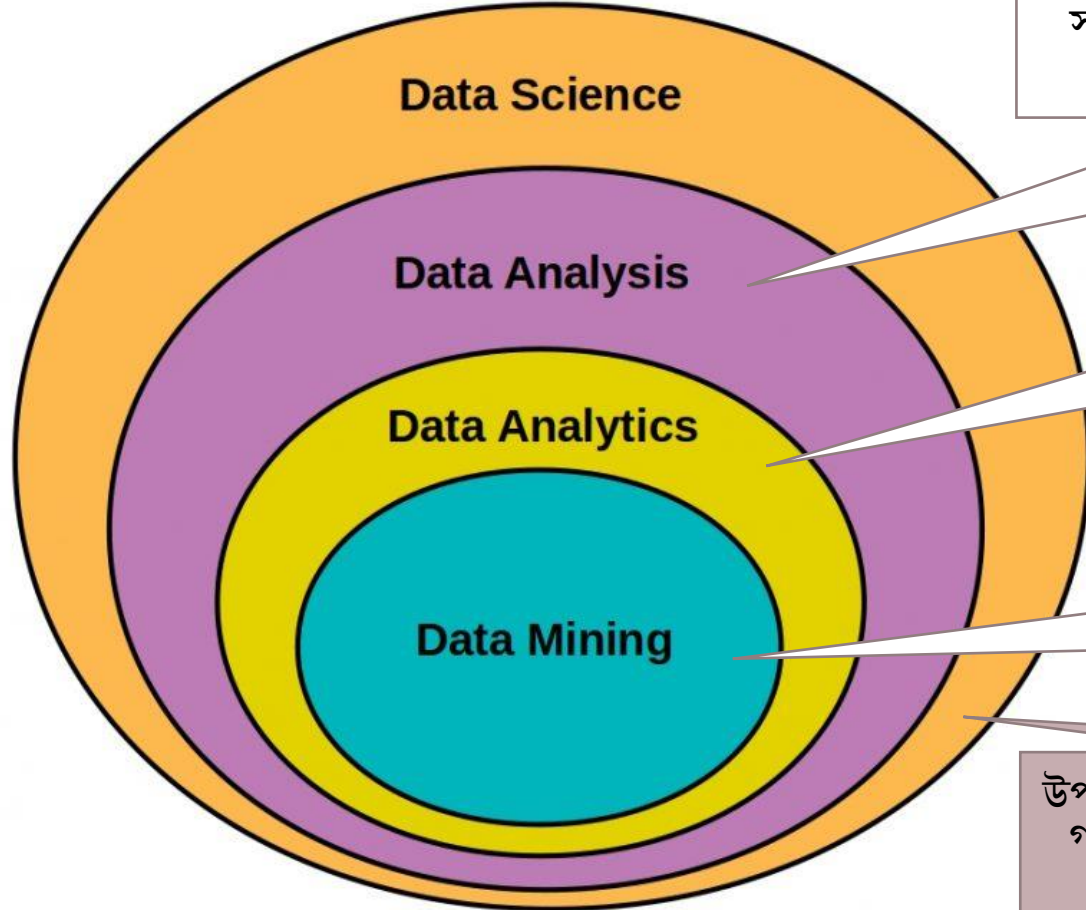


DECISION RISK

ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।



ডেটা বিশ্লেষণ হল উপযোগী তথ্য আবিষ্কার, উপসংহার জানানো এবং সিদ্ধান্ত গ্রহণকে সমর্থন করার লক্ষ্যে ডেটা পরিদর্শন, পরিষ্কার, রূপান্তর এবং মডেলিং করার একটি প্রক্রিয়া।

ডেটা অ্যানালিটিক্স হল তথ্য সম্পর্কে সিদ্ধান্ত নেওয়ার জন্য কাঁচা ডেটা বিশ্লেষণ করার বিজ্ঞান। ডেটা বিশ্লেষণের জন্য অনেক কৌশল এবং যান্ত্রিক প্রক্রিয়া এবং অ্যালগরিদম ব্যবহার করা হয়।

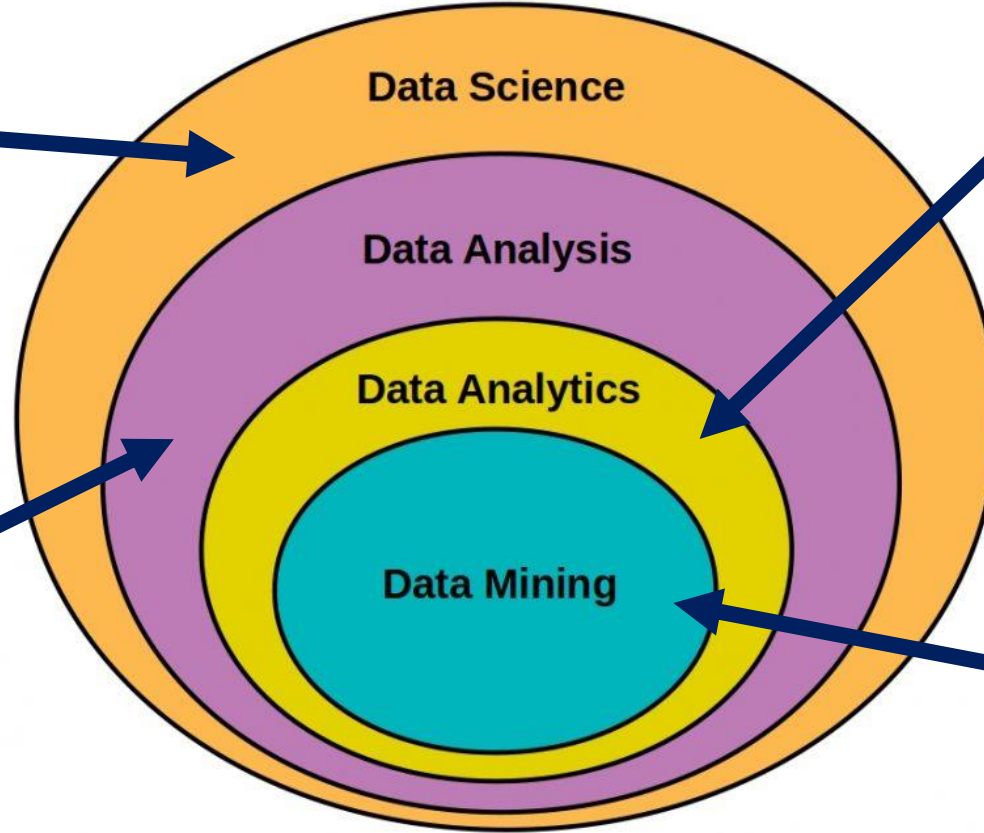
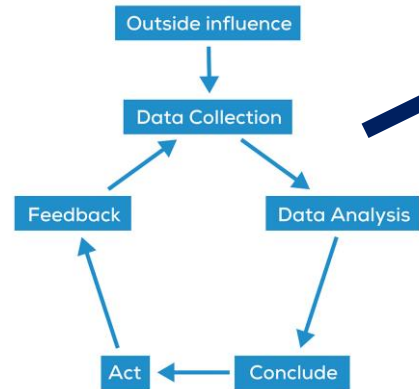
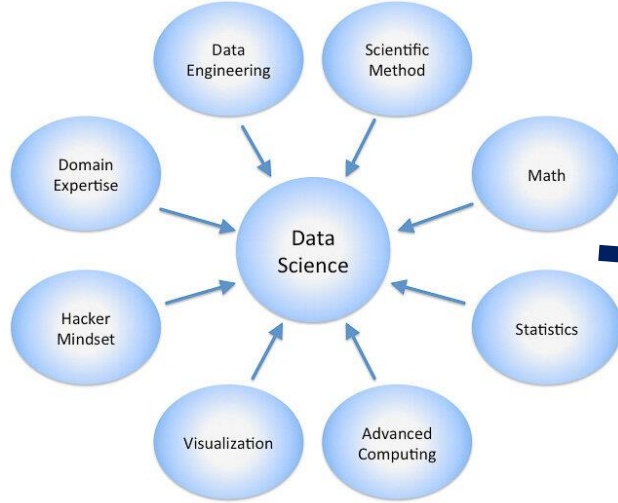
ডেটা মাইনিং একটি প্রক্রিয়া যা কাঁচা ডেটাকে দরকারী এবং আরও বোধগম্য তথ্যে রূপান্তর করার জন্য ব্যবহার করা হয়।

উপাত্তবিজ্ঞান/তথ্য বিজ্ঞান (Data Science) হচ্ছে একটি বহুবিষয়ী ক্ষেত্র যেখানে বিভিন্ন গঠিত ও অগঠিত উপাত্ত বিশ্লেষণে ও সেখান থেকে জ্ঞান আহরণে বৈজ্ঞানিক পদ্ধতি, প্রক্রিয়া, অ্যালগরিদম ও বিভিন্ন সিস্টেম ব্যবহার হয়।

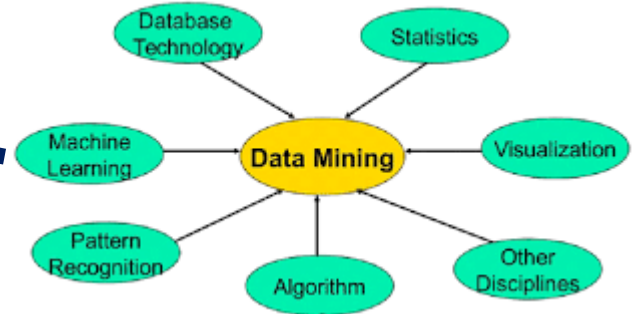
ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।



Healthcare	Science	Security	Business
 ✓ It allow us to find new cures and better understand and predict disease patterns. ✓ This leads to saving more lives.	 ✓ It creates new possibilities and ways to conduct research which would otherwise be impossible, helping us to make new discoveries.	 ✓ Police forces use big data tools to predict criminal activities, conduct investigations and ultimately to catch criminals faster.	 ✓ It helps us to improve and optimize the ways we do business by making data-driven decisions.



ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।

What they are

The biggest difference between data mining and data science is simply what it is. Although data science is an extensive field of science, data mining is just a method used in this area. This means that data science encompasses a wider range of research and methods, while data mining focuses solely on data collection and transformation within a single process.

যদিও ডেটা সায়েন্স বিজ্ঞানের একটি বিস্তৃত ক্ষেত্র, ডেটা মাইনিং এই ক্ষেত্রে ব্যবহৃত একটি পদ্ধতি মাত্র। এর মানে হল যে ডেটা সায়েন্স বিস্তৃত গবেষণা এবং পদ্ধতিগুলি অন্তর্ভুক্ত করে, যখন ডেটা মাইনিং শুধুমাত্র একক প্রক্রিয়ার মধ্যে ডেটা সংগ্রহ এবং রূপান্তরকে কেন্দ্র করে।

Data Mining vs Data Science

Comparison Chart

Data Mining	Data Science
It deals with data analysis and modeling to understand the complex world of data.	It is the science of extracting useful information from large data sets or databases.
Data mining is just a part of a broader 'knowledge discovery from data' process.	It is a multidisciplinary field that covers a wide range of techniques, applications, and disciplines.
It is about uncovering valuable information from the tremendous amounts of data and to transform such data into organized knowledge.	It uses a blend of tools, algorithms, and machine principles to extract usable information from both structured and unstructured data.
Applications of data mining include predicting customer behavior, credit analysis and fraud detection, discovering relationships between diseases, business intelligence, and more.	Common applications include fraud and risk detection, targeted advertising, speech recognition, airline route planning, image recognition, delivery logistics, etc.

ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।

Focus

Data mining is usually used as **part of the business analysis process**. Typically, data mining is not used outside the business environment because it is explicitly designed to help companies collect and understand their data. On the other hand, data science is a scientific study. Data scientists use this study, among other things, to create predictive models, conduct experiments and social analysis.

ডেটা মাইনিং সাধারণত ব্যবসায়িক বিশ্লেষণ প্রক্রিয়ার অংশ হিসেবে ব্যবহৃত হয়। সাধারণত, ডেটা মাইনিং ব্যবসার পরিবেশের বাইরে ব্যবহার করা হয় না কারণ এটি স্পষ্টভাবে কোম্পানিগুলিকে তাদের ডেটা সংগ্রহ এবং বুঝতে সাহায্য করার জন্য ডিজাইন করা হয়েছে। অন্যদিকে, তথ্য বিজ্ঞান একটি বৈজ্ঞানিক গবেষণা। তথ্য বিজ্ঞানীরা এই গবেষণাকে অন্যান্য বিষয়ের মধ্যে ব্যবহার করে ভবিষ্যদ্বাণীমূলক মডেল তৈরি করতে, পরীক্ষা-নিরীক্ষা এবং সামাজিক বিশ্লেষণ করতে।



ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।

Professionals in this field

In data mining, professionals are only expected to understand how to collect, organise, understand, and accurately display data. On the other hand, data scientists should have at least some qualifications in many fields, such as AI research, data engineering, data analysis, programming, and domain knowledge. To use data mining, you need to have some of the knowledge and skills that data scientists have, but not as great.

ডেটা মাইনিংয়ের ক্ষেত্রে, পেশাদাররা কেবল ডেটা সংগ্রহ, সংগঠিত, বোঝা এবং সঠিকভাবে প্রদর্শন করতে বোঝার প্রত্যাশা করে। অন্যদিকে, তথ্য বিজ্ঞানীদের অনেক ক্ষেত্রে কমপক্ষে কিছু যোগ্যতা থাকা উচিত, যেমন এআই গবেষণা, ডেটা ইঞ্জিনিয়ারিং, ডেটা বিশ্লেষণ, প্রোগ্রামিং এবং ডোমেন জ্ঞান। ডেটা মাইনিং ব্যবহার করার জন্য, আপনার কিছু জ্ঞান এবং দক্ষতা থাকা দরকার যা ডেটা বিজ্ঞানীদের আছে, কিন্তু ততটা মহান নয়।



ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

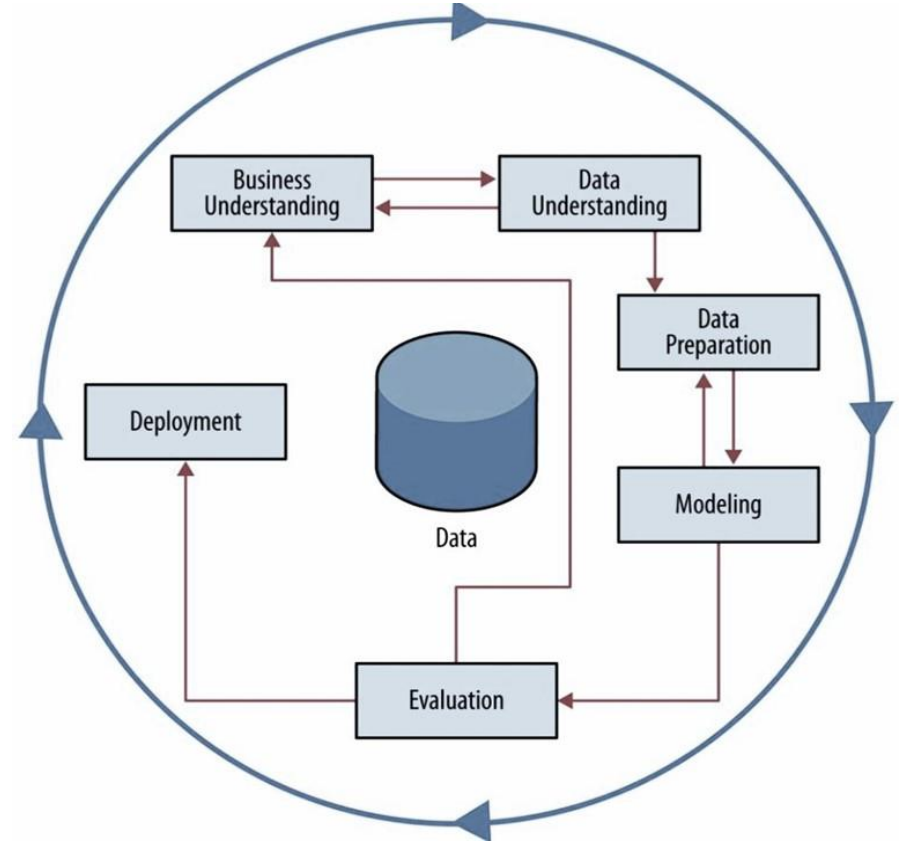
Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।

Data type

As a rule, data mining focuses only on structured data, although unstructured data can be used. For data scientists, using structured, unstructured and semi-structured data is common. Data mining is a little easier in this aspect because professionals cannot know how to work with all types of data, while data experts will probably need to know all types.

একটি নিয়ম হিসাবে, ডেটা মাইনিং শুধুমাত্র স্ট্রাকচার্ড ডেটার উপর ফোকাস করে, যদিও অসংগঠিত ডেটা ব্যবহার করা যেতে পারে। তথ্য বিজ্ঞানীদের জন্য, কাঠামোগত, অবকাঠামো এবং আধা কাঠামোগত ডেটা ব্যবহার করা সাধারণ। ডেটা মাইনিং এই দিক থেকে একটু সহজ কারণ পেশাদাররা সব ধরনের ডেটা নিয়ে কীভাবে কাজ করতে হয় তা জানতে পারে না, অন্যদিকে ডেটা বিশেষজ্ঞদের সম্ভবত সব ধরনের জানতে হবে।



ডেটা মাইনিং এবং ডেটা সায়েন্স - মূল পার্থক্য

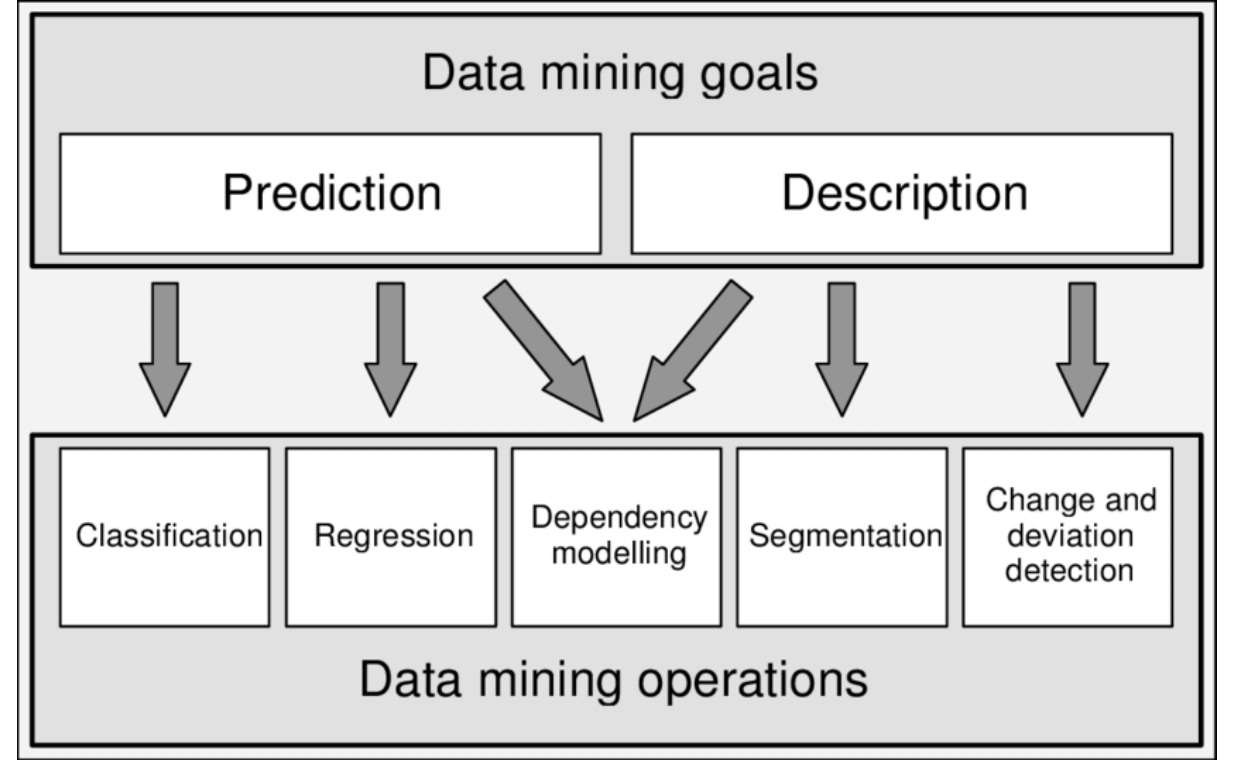
Data mining is a process, and data science is an area of research.

ডেটা মাইনিং একটি প্রক্রিয়া এবং ডেটা সায়েন্স হচ্ছে গবেষণার একটি ক্ষেত্র।

Goal

Data mining's primary goal is to make business data easy to understand and therefore available for use. Data science aims to achieve scientific advances and create data-driven products for use by different organisations. In general, data mining has a much more specific purpose than data science. The whole purpose of data mining is to study and organise company data and identify previously unknown trends.

ডেটা মাইনিং এর প্রাথমিক লক্ষ্য হল ব্যবসার ডেটা বোঝা সহজ এবং তাই ব্যবহারের জন্য উপলব্ধ। ডেটা সায়েন্সের লক্ষ্য বৈজ্ঞানিক অগ্রগতি অর্জন করা এবং বিভিন্ন সংস্থার দ্বারা ব্যবহারের জন্য ডেটা চালিত পণ্য তৈরি করা। সাধারণভাবে, ডেটা মাইনিংয়ের ডেটা সায়েন্সের চেয়ে অনেক বেশি নির্দিষ্ট উদ্দেশ্য রয়েছে। ডেটা মাইনিংয়ের পুরো উদ্দেশ্য হল কোম্পানির ডেটা অধ্যয়ন এবং সংগঠিত করা এবং পূর্বে অজানা প্রবণতা চিহ্নিত করা।



NEXT

Data Science in Self-Driving Cars

