DATA SCIENCE
MACHINE LEARNING
with PYTHON

Instructor:

Rashedul Alam Shakil

Founder of aiQuest Intelligence

M. Sc. in Data Science (Major ML/AI)

FAU Erlangen Nuremburg, Germany

Email: rashedul.alam@aiquest.org

What, Why, How?

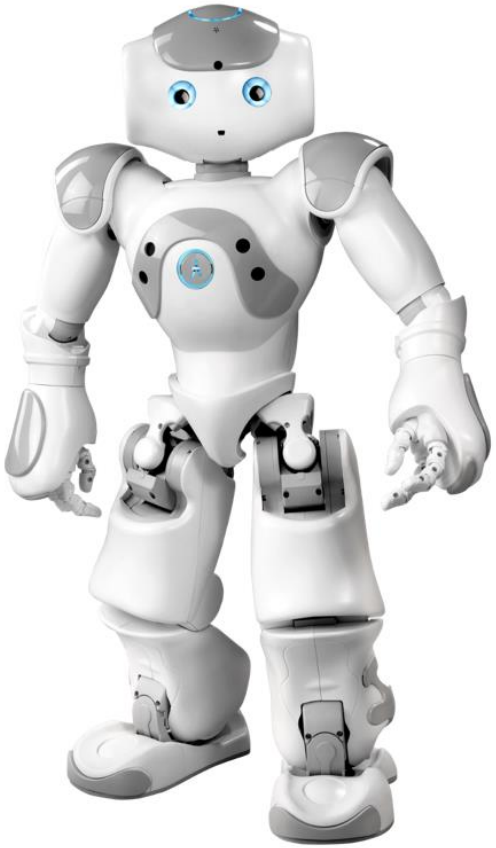
Data Science

What is Data Science?

Why it is So Popular?

Does Data Science Have Future?

How Much Money Do Data Scientists Make?

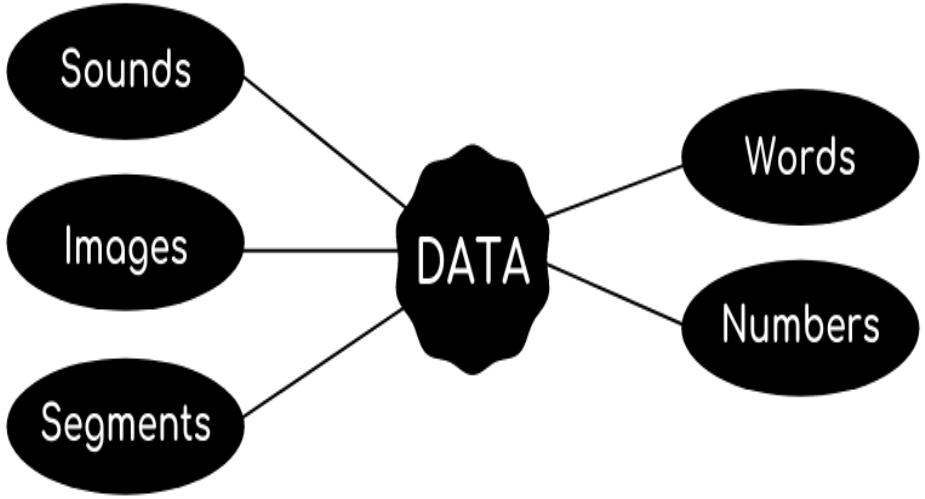


What is Data?

Definition

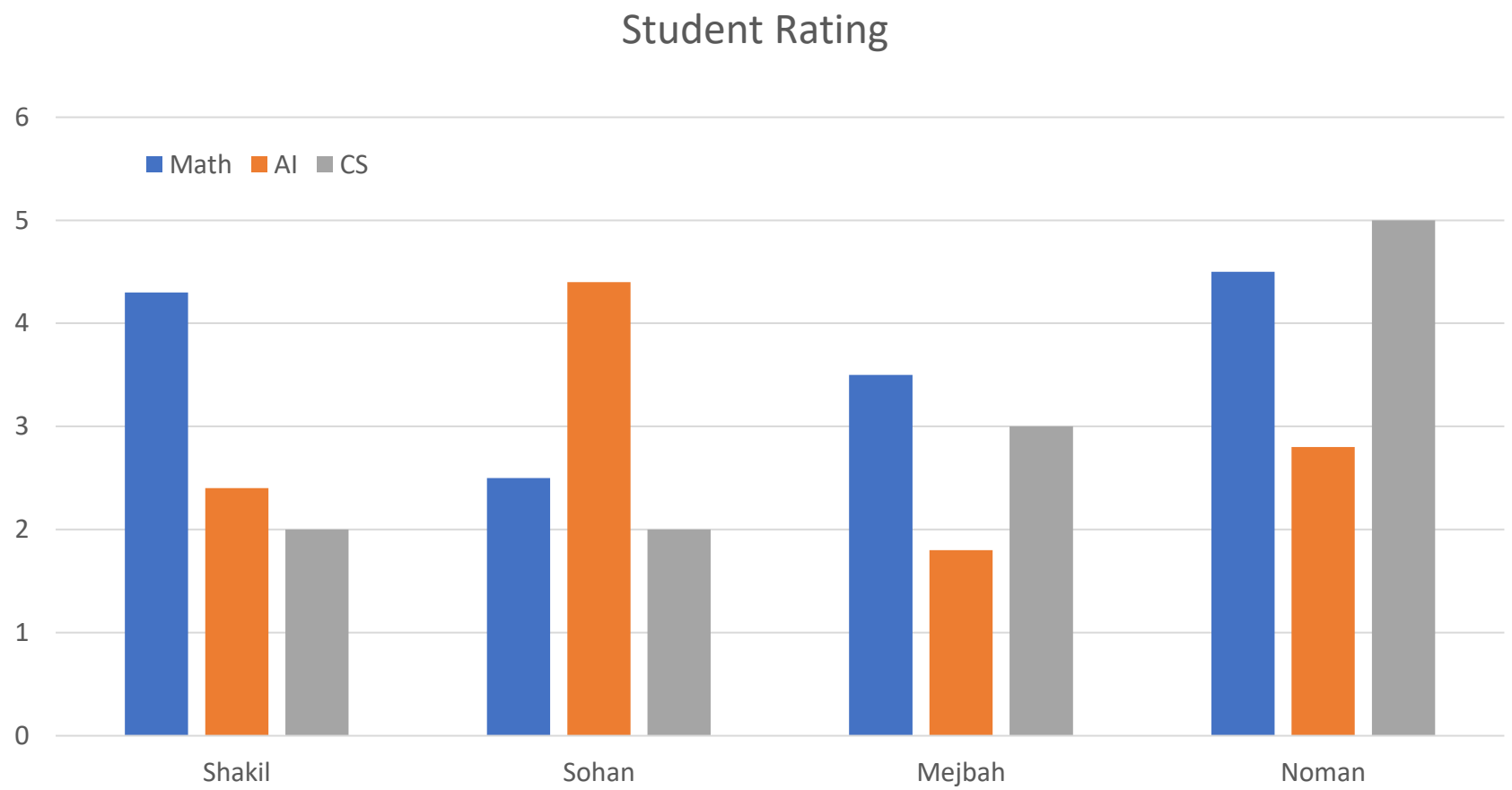
Data is defined as a collection of organized or unorganized facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, processing by humans, or some automatic means such as computers, ATMs.

The main examples of data are phone numbers, weights, prices, costs, number of items sold, product names, addresses, registration marks, etc.



What is Information?

Definition



What is Database?

Definition

A database is an organized collection of related data or information stored and accessed electronically within a computer system.

For example, SQL, MongoDB, Oracle Database, etc. are all examples of different databases. These modern databases are managed by DBMS. Structured Query Language, or SQL as it is more widely known, is used to operate on the data in a database.

What is Data Warehouse?

Definition

A data warehouse is a centralized storage system that allows for the storing, analyzing, and interpreting of data in order to facilitate better decision-making.

- Snowflake
- Databricks
- Amazon Redshift
- Azure Synapse Analytics
- Google Big Query

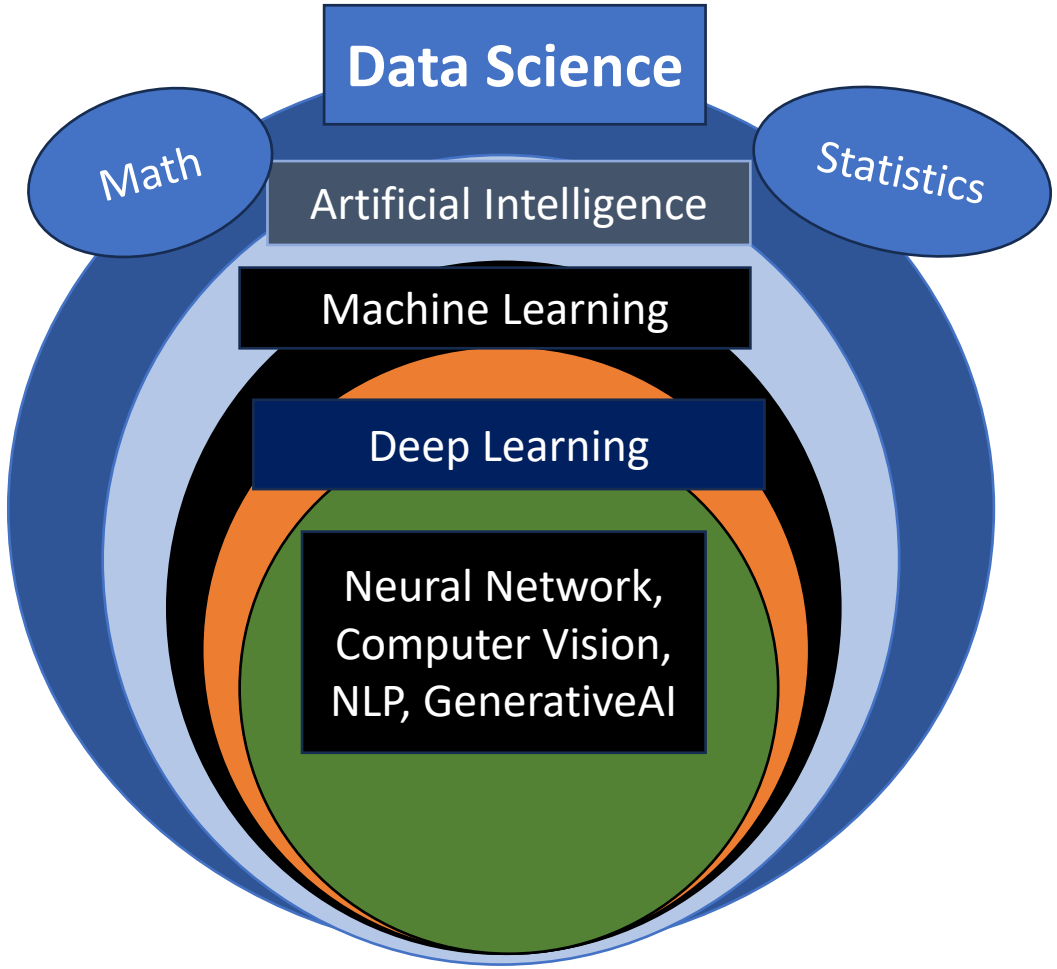
How Much Data is in the World Today?

Year	Data Generated	Change Over Previous Year
2021*	79 zettabytes	↑ 14.8 zettabytes
2022*	97 zettabytes	↑ 18 zettabytes
2023*	**120 zettabytes**	↑ 23 zettabytes
2024*	147 zettabytes	↑ 27 zettabytes

1 Zetabyte (ZB) = 1,000 Exabytes (EB), 1 Exabyte (EB) = 1,000 Petabytes (PB)
1 Petabyte (PB) = 1,000 Terabytes (TB)

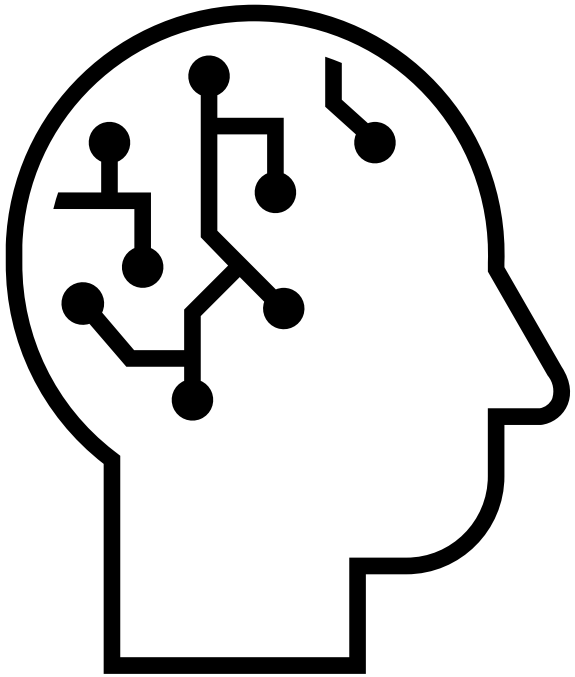
How to Use Data for a Smart Career?

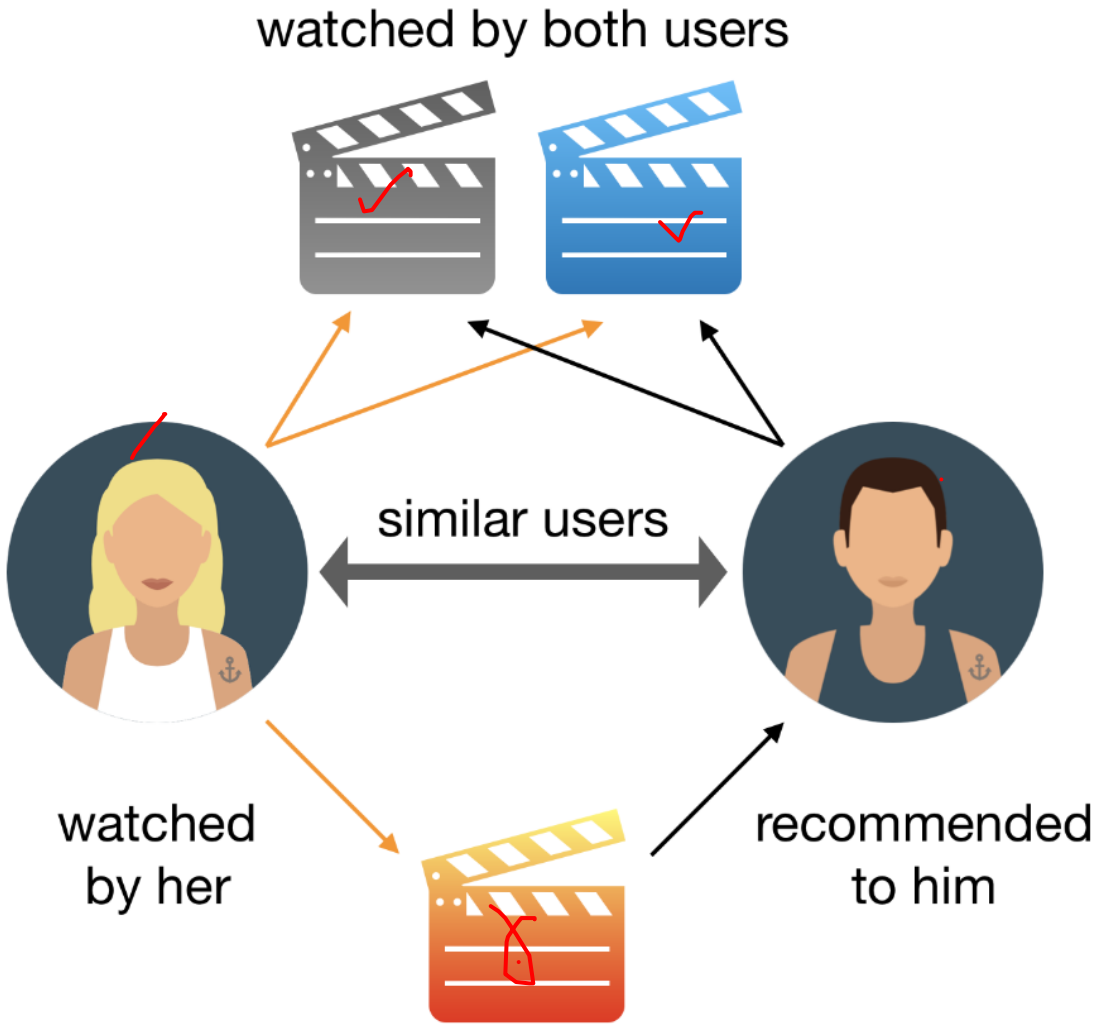
- Computer Science
- Mathematics
- Statistics



What is AI?

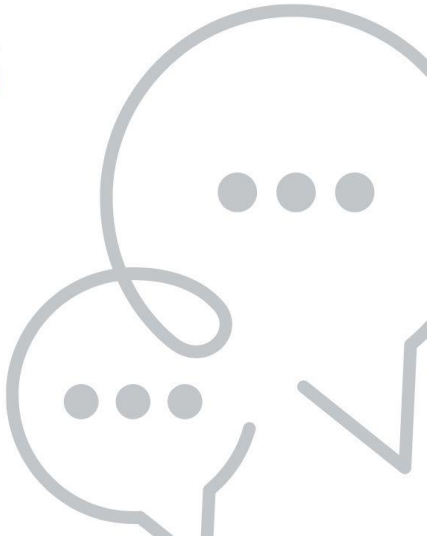
Definition





AI Chat Bot for your Business

IMPROVE YOUR CUSTOMER SERVICE WITH AI CHAT BOT.

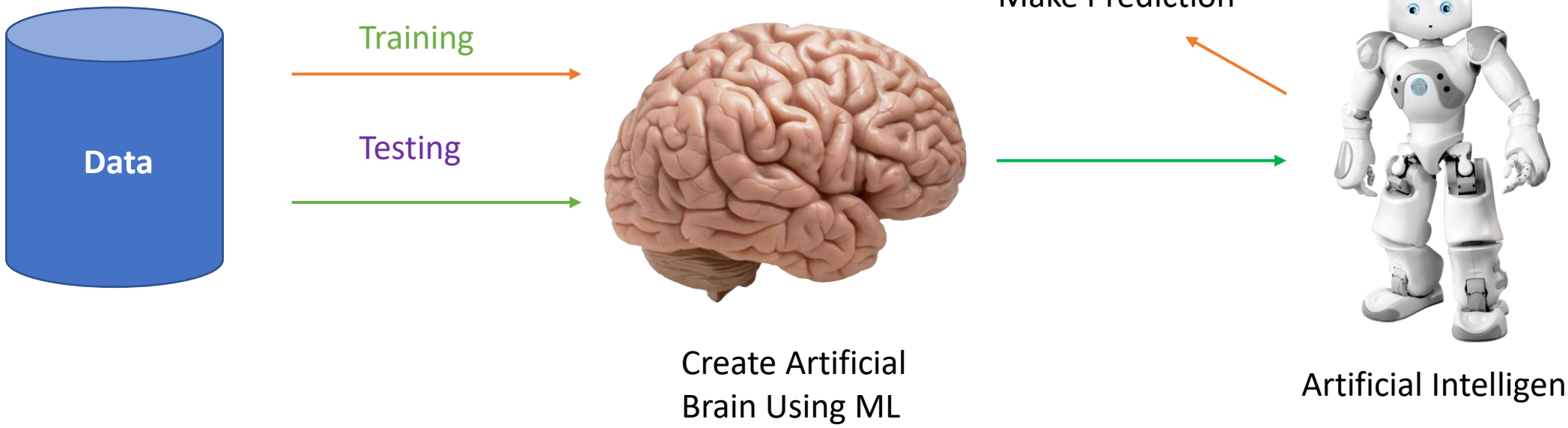


Machine Learning

Definition

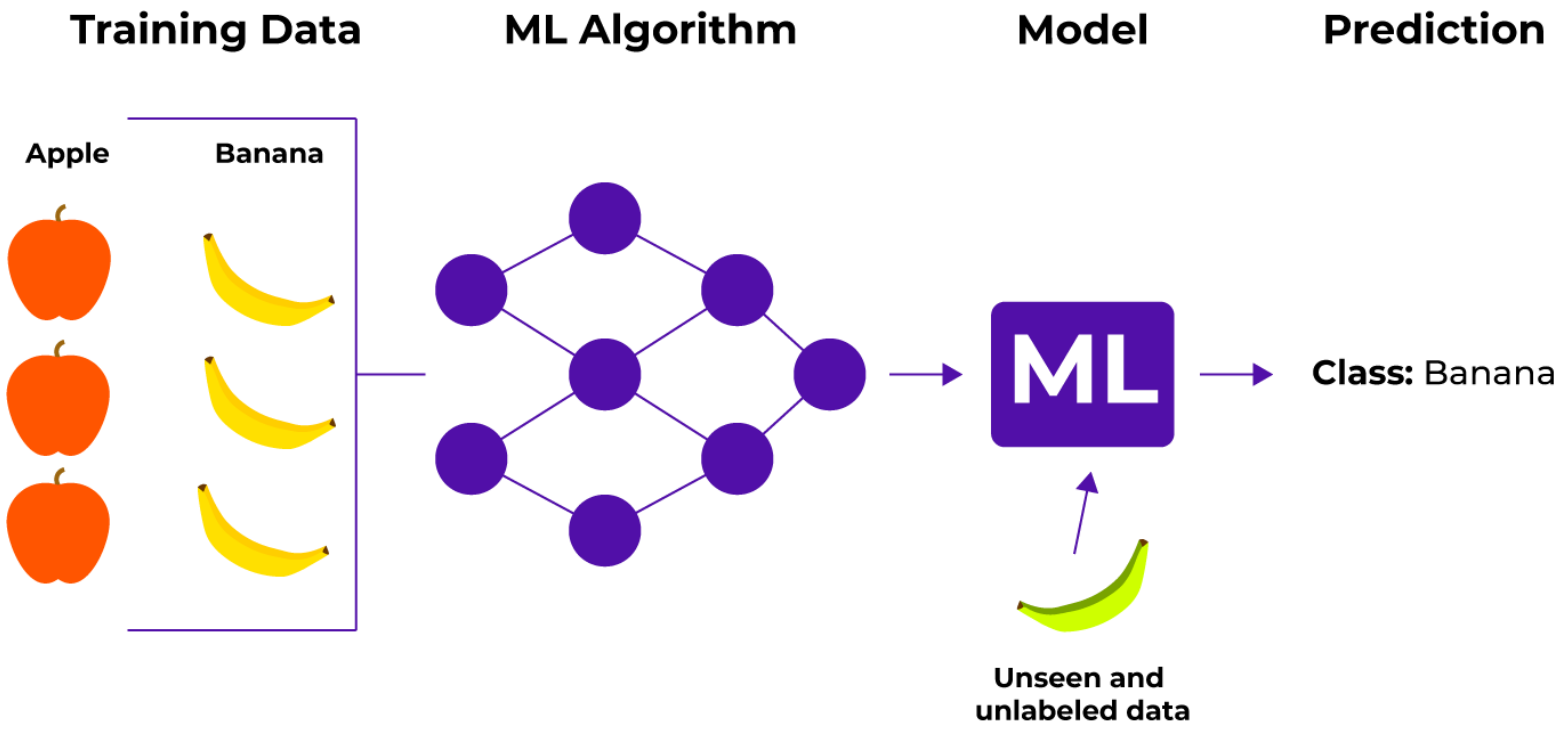
What is Machine Learning?

ML = Machine + Learning



Machine Learning

Example



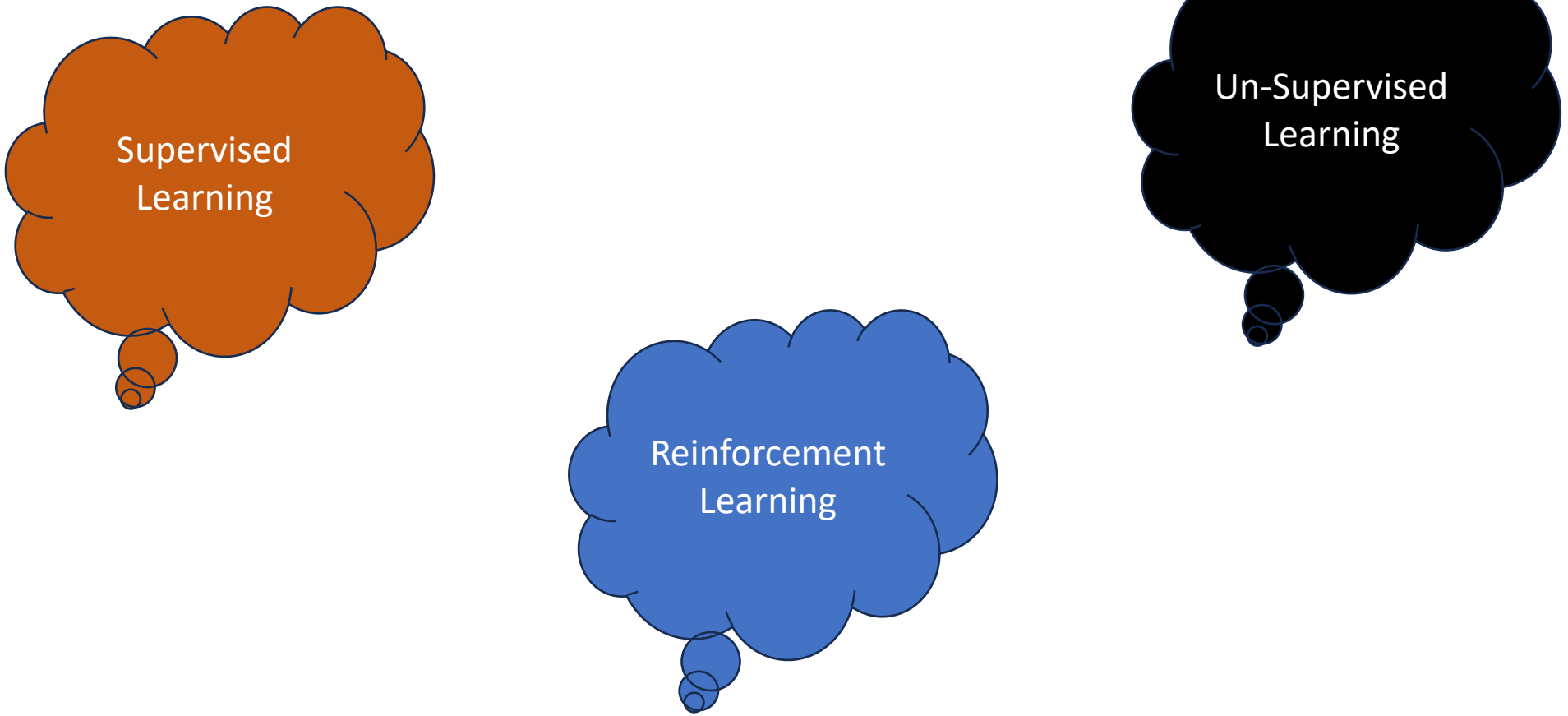
Types Machine Learning Model

Generalized classification

Discriminative Models

Generative Models

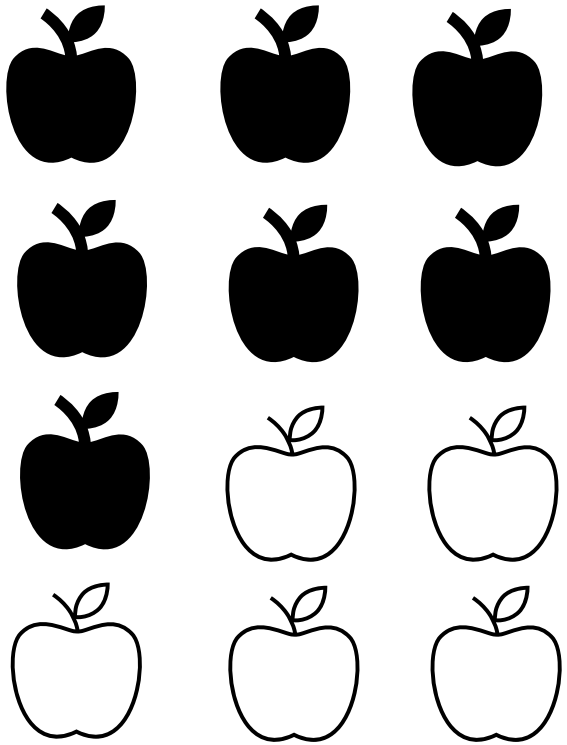
Types of Machine Learning



Statistical Machine Learning

Supervised Learning

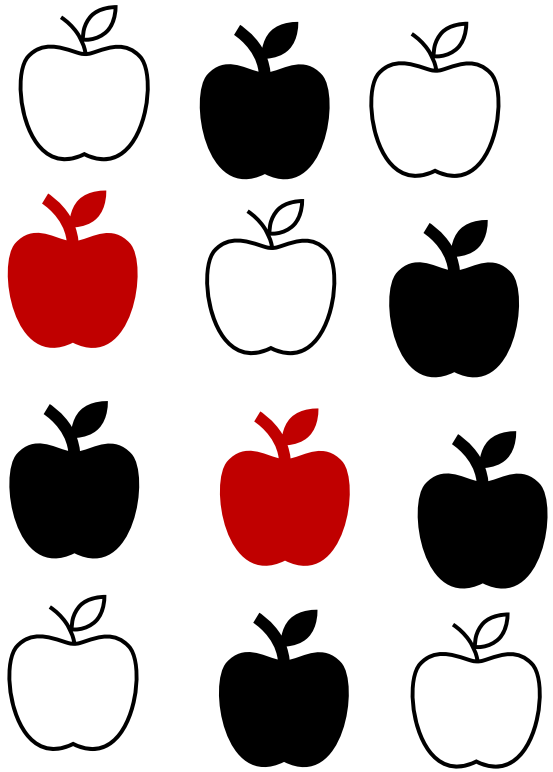
1. Classification Tasks



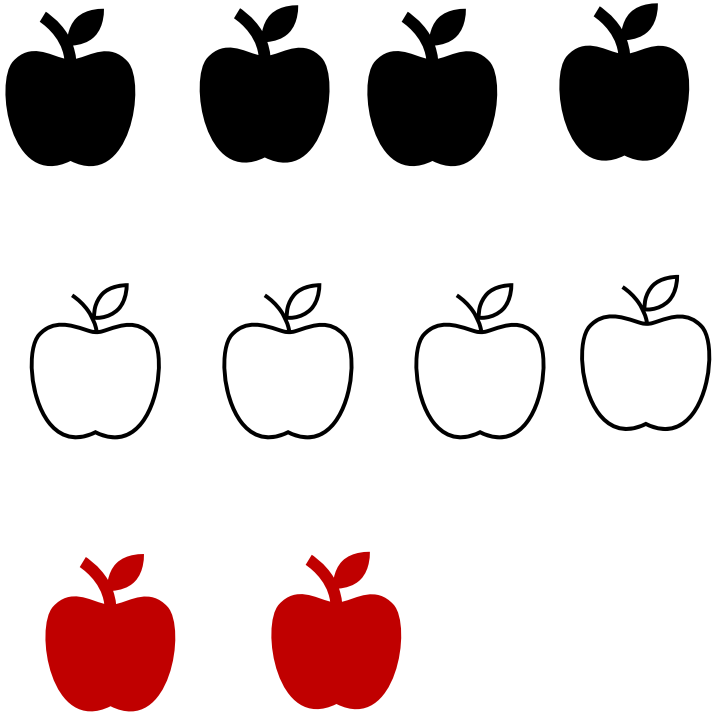
2. Regression Tasks

Statistical Machine Learning

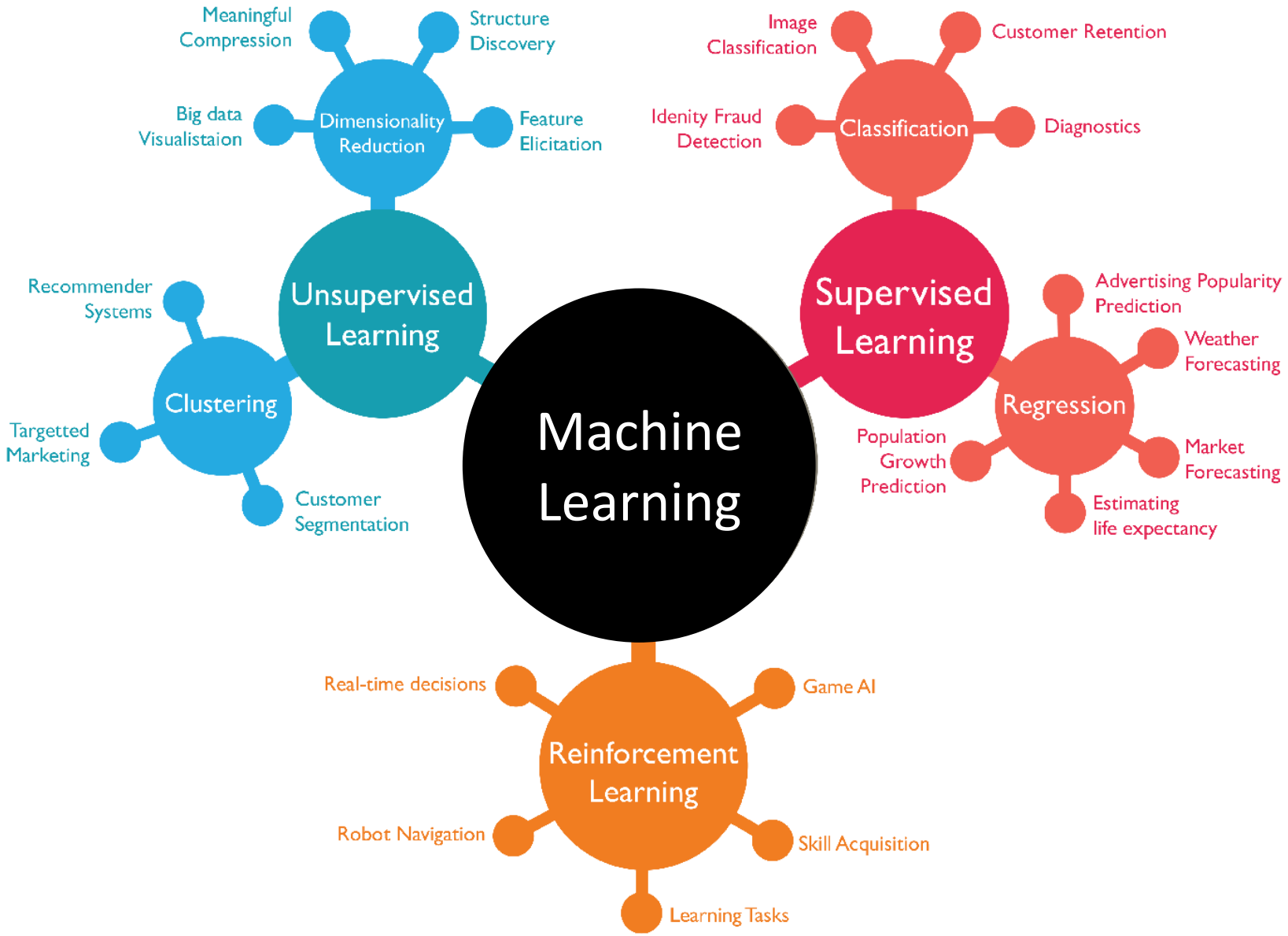
Before Cluster



After Cluster



Statistical Machine Learning



Neural Machine Learning

1. Convolutional Neural Networks (CNNs)
2. Recurrent Neural Networks (RNNs)
3. Long Short-Term Memory (LSTM) Networks
4. Generative Adversarial Networks (GANs)
5. Variational Autoencoders (VAEs)
6. Transformer Networks
7. N-Gram Model
8. Bayesian Neural Networks (BNNs)

Types Machine Learning Model

Generalized classification

Discriminative Models

Generative Models

Fig: Training Model

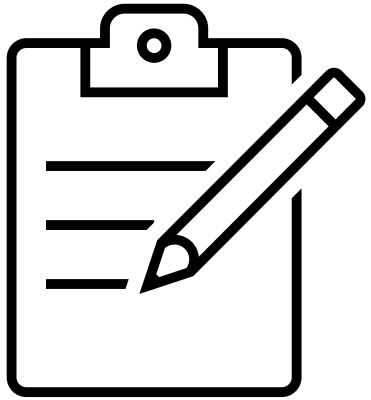
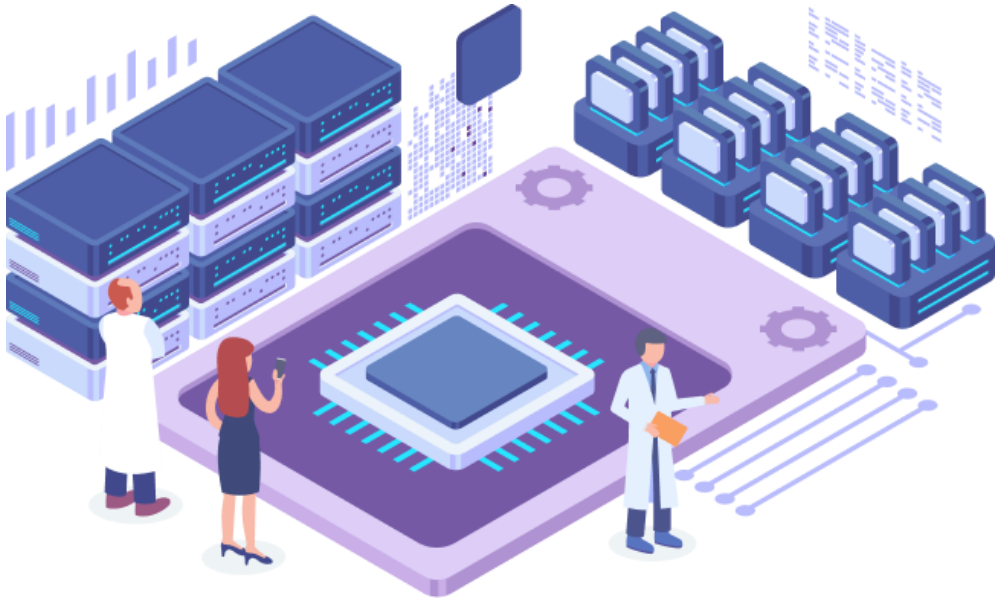


Fig: Generating Answer

Large Language Models: LLMs

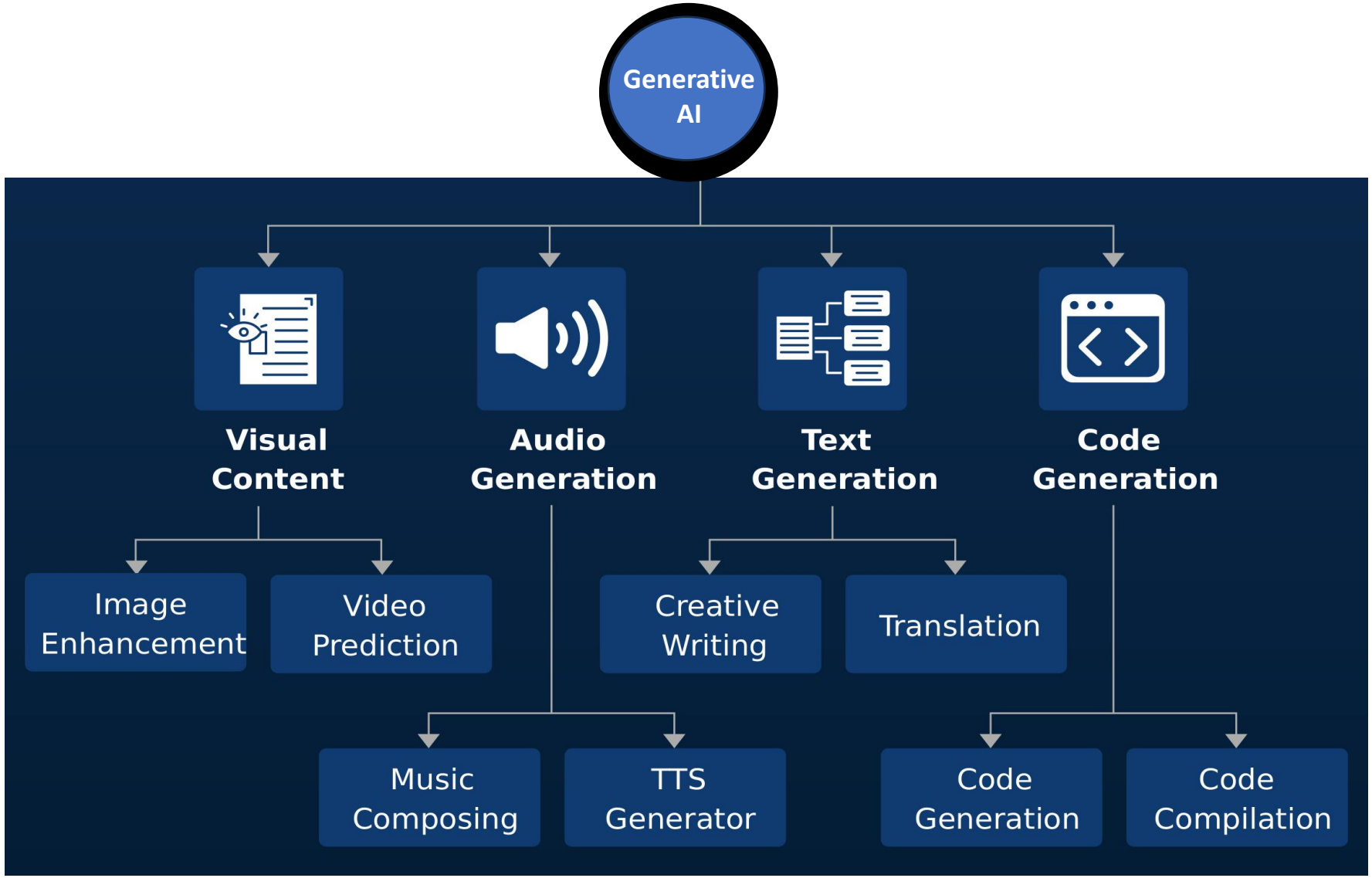
Brief Discussion

Large language models (LLMs) are natural language processing computer programs that use artificial neural networks to generate text. Some notable ones are GPT-3, GPT-4, LaMDA (Bard), BLOOM, and LLaMA. LLMs power many applications, such as AI chatbots and AI search engines.



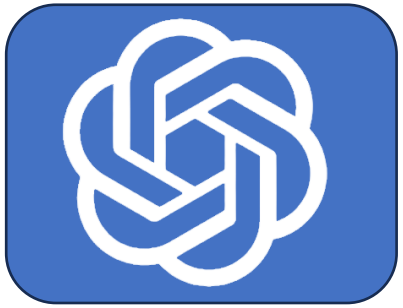
Generative AI Scopes

General Scopes

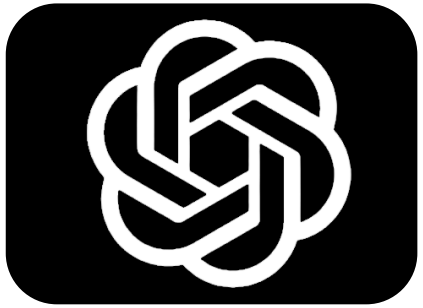


Generative AI Tools

Some Popular Tools



Chat-GPT



DALLE / DALLE 2

Google Bard AI

Midjourney AI

Recurrent Neural Networks (RNNs)

Long Short-Term Memory (LSTM) Networks

Generative Adversarial Networks (GANs)

Variational Autoencoders (VAEs)

Transformer Networks

N-Gram Model

Common Skills in Data Science

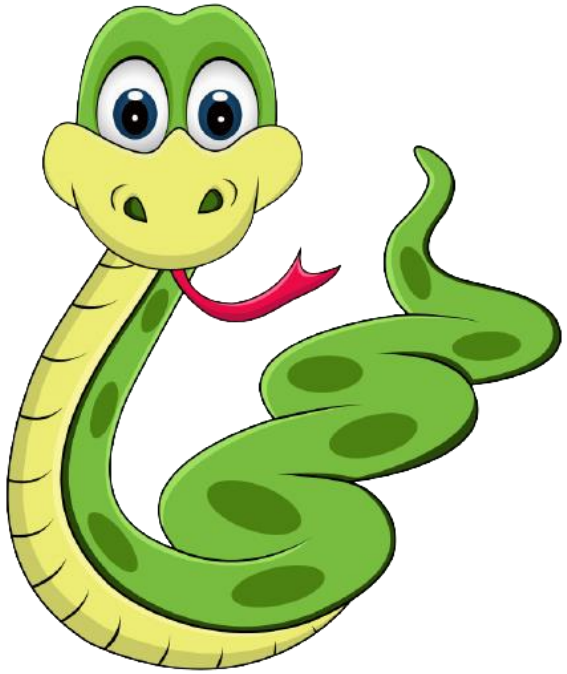
- Programming Languages: Python, R, etc.
- Statistics and Mathematics
- Data Manipulation with Pandas
- Machine Learning
- Deep Learning
- Data Visualization
- SQL
- Big Data Tools: Hadoop, Spark
- Feature Engineering
- Data Communication
- Domain Knowledge
- Experimentation and Evaluation
- Data Ethics and Privacy
- Version Control: Git
- Problem-Solving
- Data Science Libraries: pandas, NumPy, scikit-learn
- Data Storytelling
- Collaboration and Teamwork
- Cloud Computing: AWS, Azure, Google Cloud
- Time Management

What Are We Going To Learn?

Module 01: Introduction to Data Science & AI
Module 02: Basics Python
Module 03: Regression Algorithms
Module 04: Feature Engineering
Module 05: Basics of Linear Algebra
Module 06: Statistics for Data Science
Module 07: Classification Algorithms
Module 08: ML Model Evaluation
Module 09: Unsupervised Learning
Module 10: Hyperparameter Optimization
Module 11: Deep Learning & Neural Networks

Download Module!

Why Python for Data Science & AI

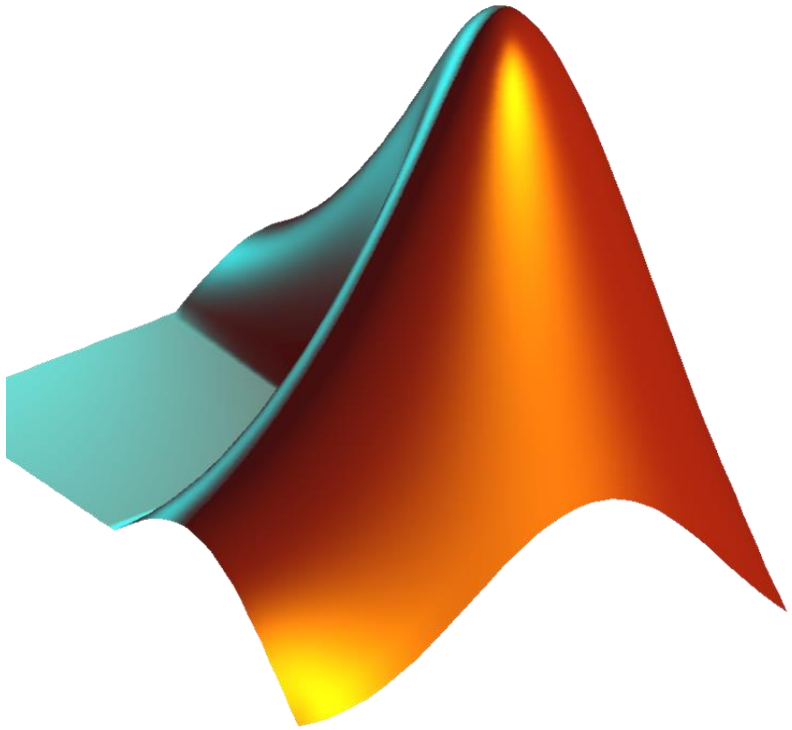


1. Easy Syntax, Flexible, Support OOP & Faster
2. Python has Machine Learning Libraries
3. Python has Data Analysis Library
4. Python has Data Frame Library
5. Python has Calculator Library
6. Python is Significant for Deep Learning
7. Keras, Tensorflow, Pytorch
8. Web (Django & Flask)
9. Open Resources

- Data Analysis
- Data Visualization
- Statistical Libraries
- Statistics and Research Methods

MATLAB for AI

- Numerical Computing
- Matrix Operations
- Machine Learning Toolbox
- Deep Learning Support
- Data Visualization
- Simplicity and User-Friendly Interface
- Community Support
- Integration with Other Languages



Data Science Jobs

- Data Scientist ($139,840/year)
- Machine Learning Engineer ($114,826/year)
- Data Architect ($108,278/year)
- Data Engineer ($102,864 /year)
- Business Intelligence (BI) Developer($81,514 /year)
- Statistician ($76,884/year)
- Database Administrator ($72,400/year)
- Data Analyst($62, 453/year)

aiQuest's all courses are available: at https://www.aiquest.org

Build Your Strong Social Network

[LinkedIn]

[GitHub]

kaggle

[Kaggle]

[Download the Books]

Thank you!