Google Cloud

Introduction to Data Analytics

Introduction to Data Science

your lecturer: Nhan Tran

✉ nhan.th@aptechlearning.edu.vn

I'm Data & AI Expert at One Mount

Former Head of Customer Insights and Data Management at Lazada;

Lecturer at Aptech Computer Education, and FTU Hanoi;

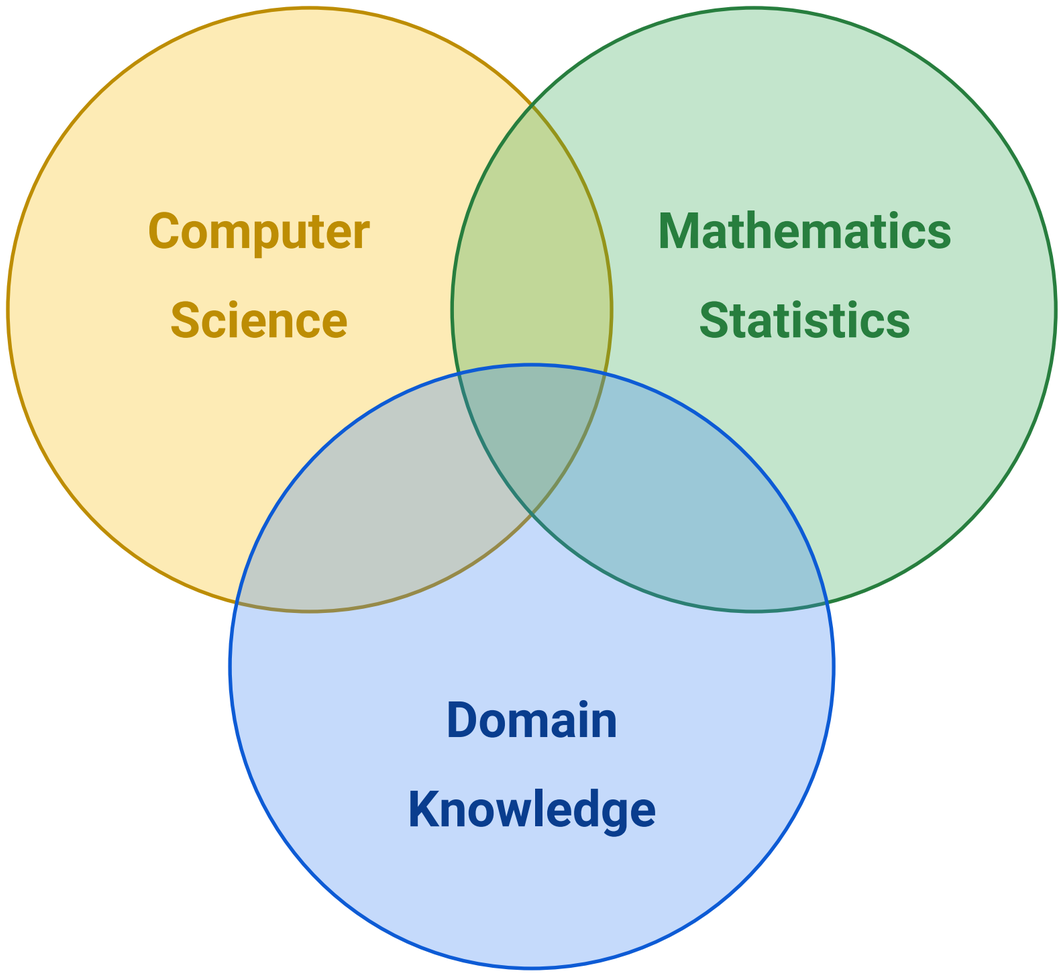
Committee Member of The International Society of Data Scientists;

Official Editor at Towards Data Science.

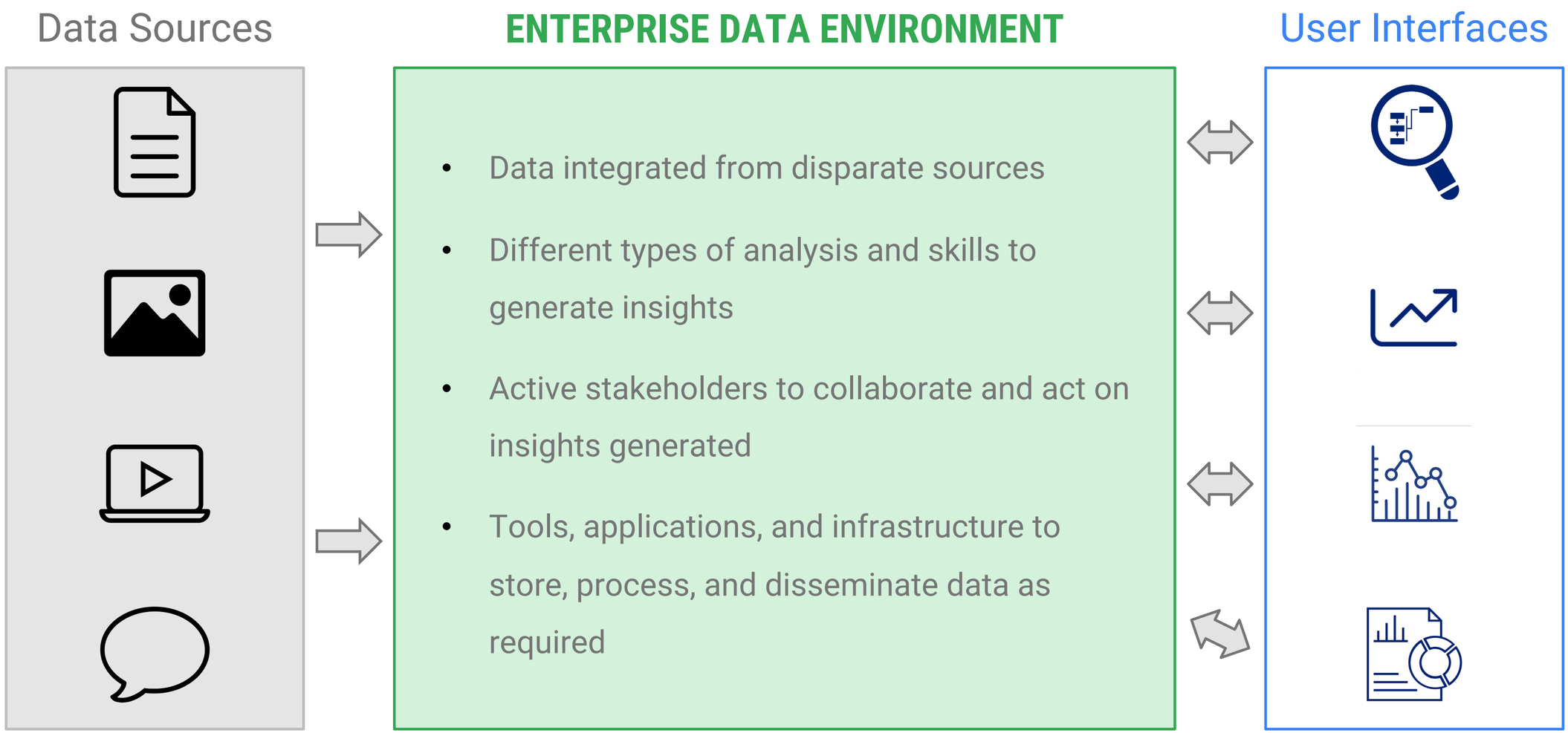
Objectives

1. Introduction to Data Science
2. Introduction to Data Analysis
3. Data Science processes
4. AI vs ML vs DL
5. Fallacies

1a. Data Science



1b. Data Ecosystem



1c. Enterprise Data Environment technologies

Cloud Technologies

Machine Learning

Big Data

1d. The Purpose of Data Ecosystem

Fraud Detection

Recommendation Engines

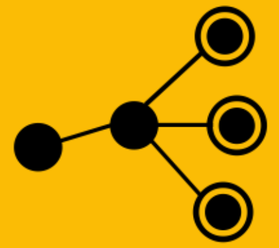
Data Mining

Customer Insights

2a. Data Analysis Introduction

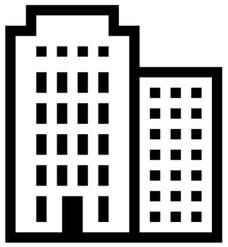
Data Analysis helps business:

- Understand the historical data (their performance);
- Validate course of action: save times, resources, ensure success;
- Take informed decisions.

Descriptive Analytics	Diagnostic Analytics	Predictive Analytics	Prescriptive Analytics
			
"What happened?"	**"Why did it happen?"**	**"What will happen?"**	**"What should be done?"**

2b. Introduction to Data Analyst

The role of Data Analyst depends on type of your company, and the extent to which it has adapted data-driven practices

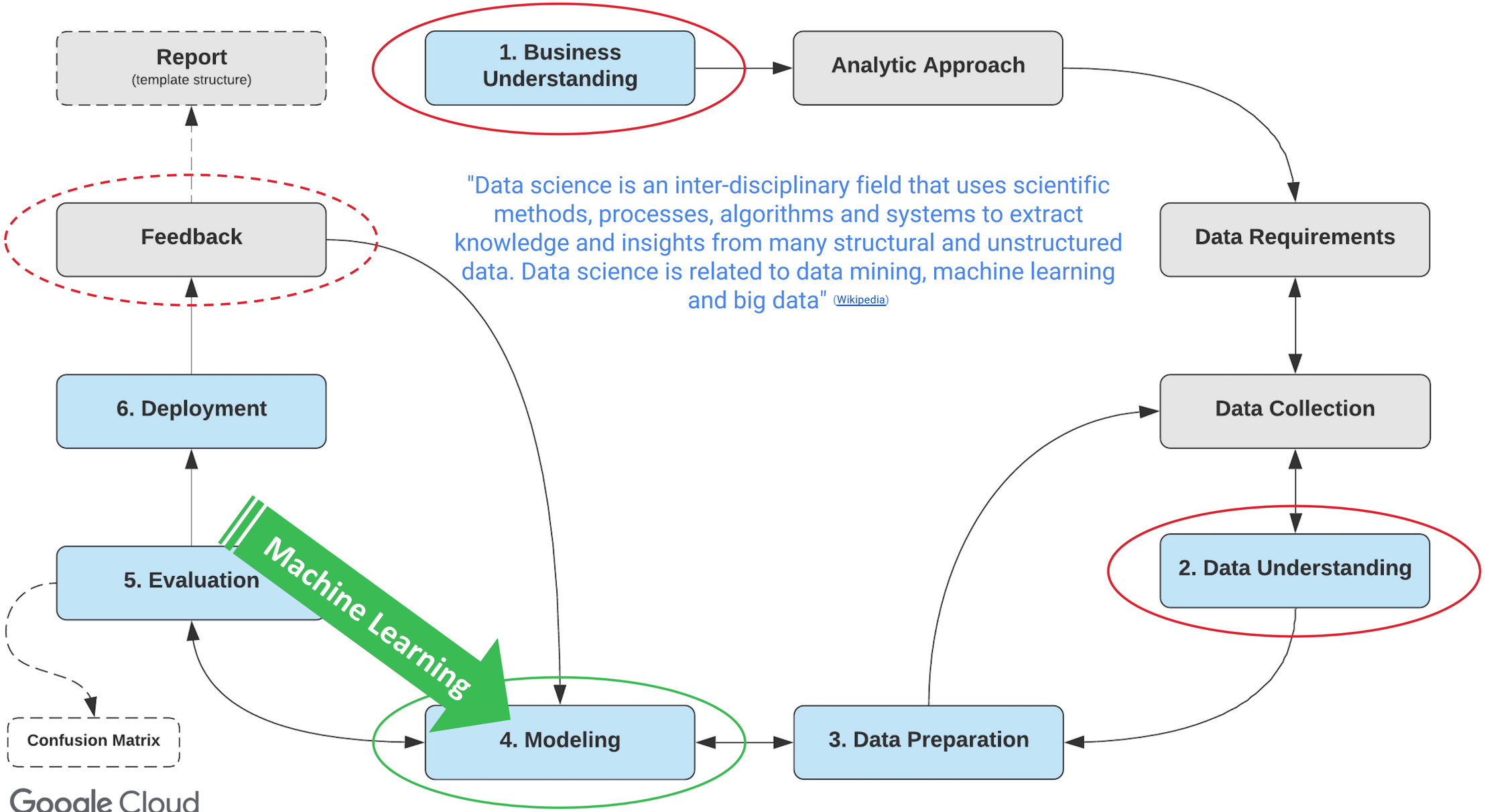


Your company

Your team's practices

- Acquiring data;
- Creating queries to extract required data;
- Filtering, cleaning, standardizing, and reorganizing data;
- Using statistical tools, and techniques;
- Analyzing patterns;
- Making charts, and reports;
- Creating appropriate documentation

3. The Data Science processes

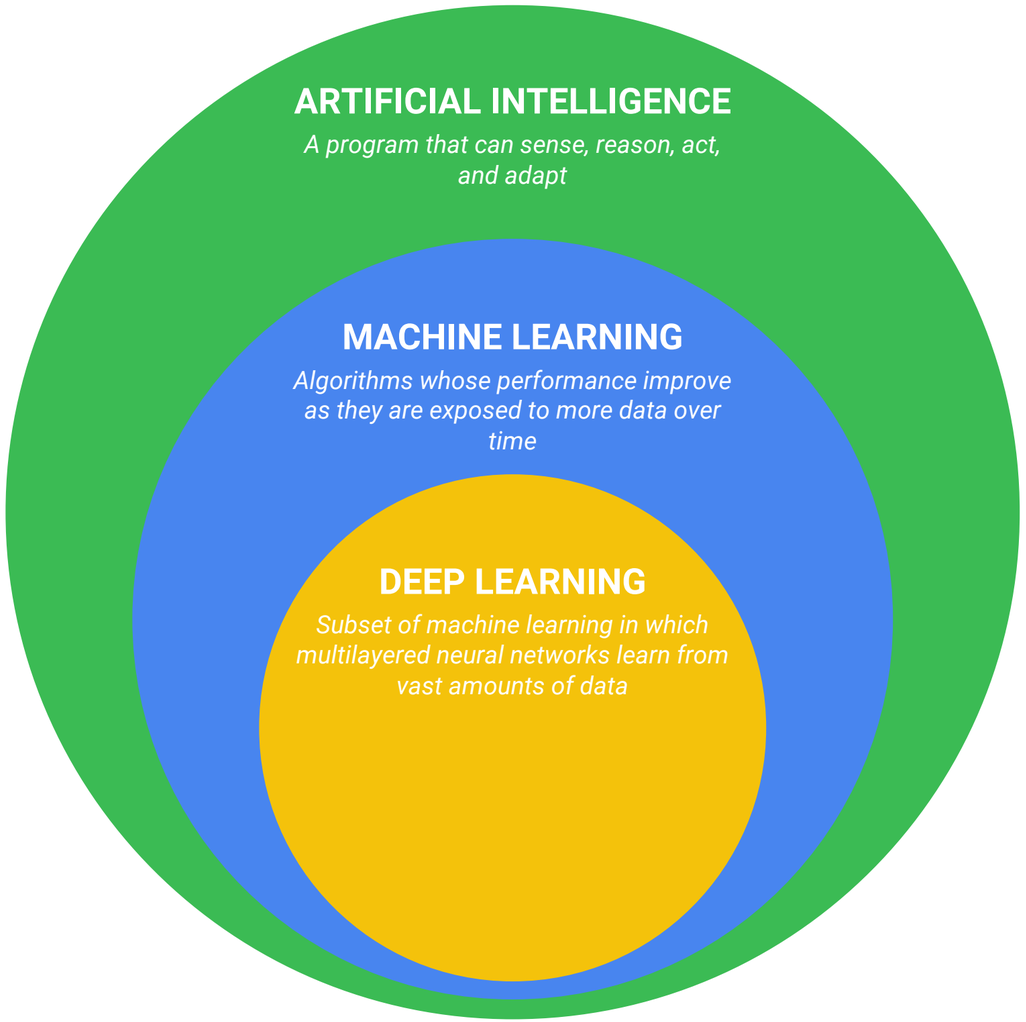


"Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data. Data science is related to data mining, machine learning and big data" (Wikipedia)

Google Cloud

Footnote: source on Lucid Chart

4. AI vs ML vs DL



Artificial Intelligence (AI) is a **discipline**

Machine Learning (ML) is a **specific way of solving** AI problems

Deep Learning (DL) is used when you **can't explain the labeling** rules

Google Cloud

Pitfall: do not misunderstanding Artificial Intelligence and Software Engineering

5a. Machine Learning fallacy

dead mindset #1: “in machine learning, we only use data and algorithms”

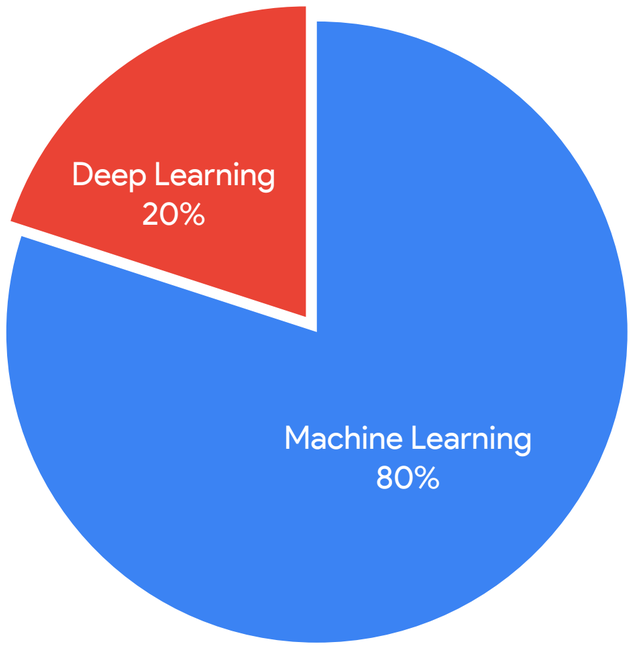
in reality, machine learning is the combination of:

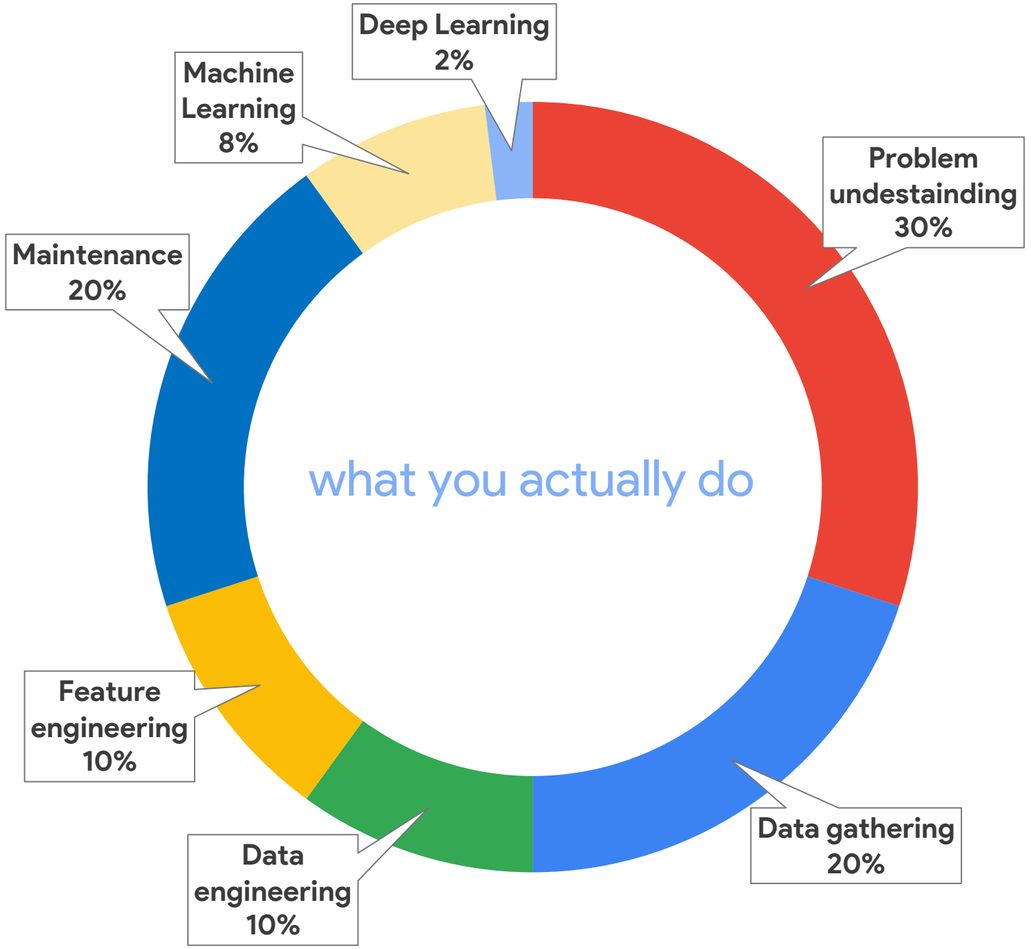
- programming skills (mostly Python, then R, MATLAB, C/C++, Java)
- analytics skills, including mathematics and statistics, hypotheses testing, and visualization
- engineering skills to retrieve data, and put your models to production
- business understanding

5b. Data Scientist fallacy

dead mindset #2: “Data Scientist job is to build machine learning models”

what you think





Google Cloud


Google Cloud