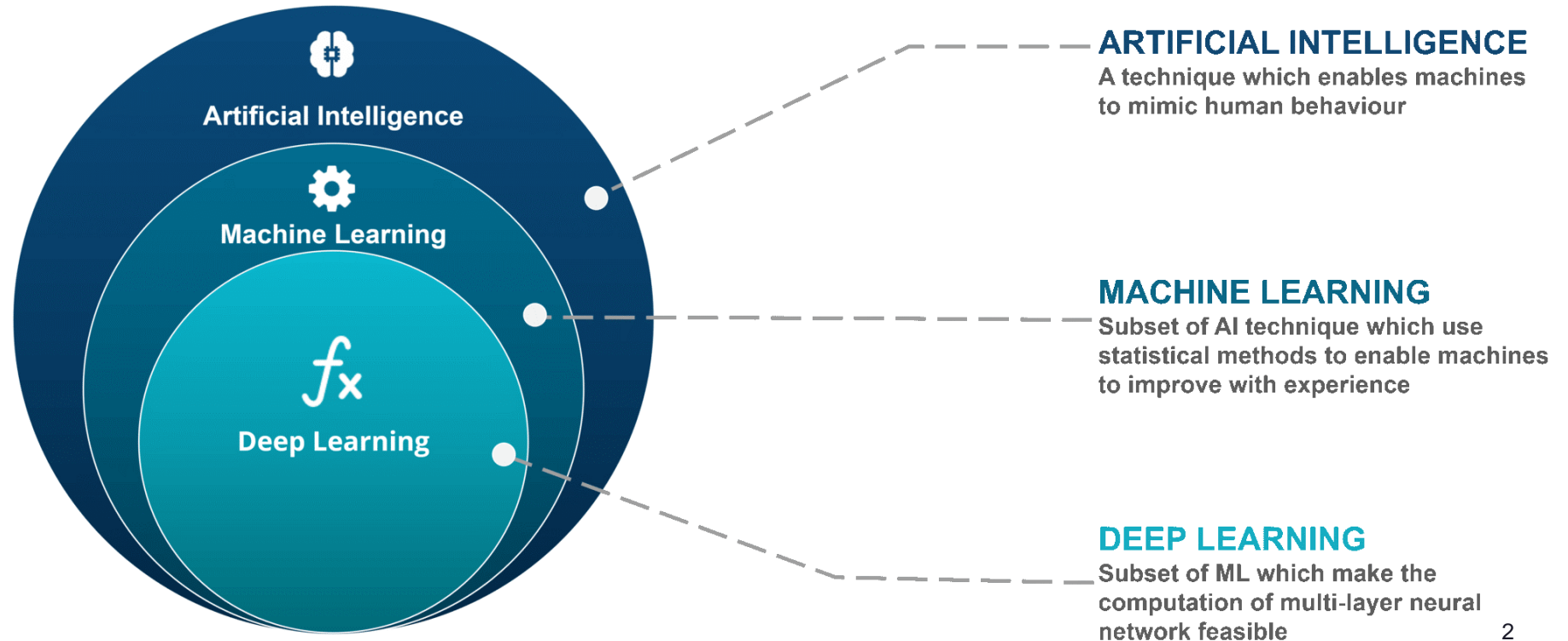# MACHINE LEARNING

Machine Learning

Machine learning is a branch of artificial intelligence that **automates** the building of systems that learn **iteratively** from data, identify patterns, and predict future results – with **minimal human intervention.** It shares many approaches with other related fields, but it focuses on predictive accuracy rather than interpretability of the model.

ARTIFICIAL INTELLIGENCE
A technique which enables machines to mimic human behaviour

MACHINE LEARNING
Subset of AI technique which use statistical methods to enable machines to improve with experience

DEEP LEARNING
Subset of ML which make the computation of multi-layer neural network feasible

2

In the new world, it is not the big fish which eats the small fish, it's the **fast** fish which eats the slow fish

Klaus Schwab
Founder and Executive Chairman
World Economic Forum

How will machine learning improve productivity?

- *automation* is the use of software to perform a repetitive task. In the business world, repetitive tasks are everywhere. A typical retail business, for example, places orders with suppliers, sends marketing material to customers, manages products in inventory, creates entries in their accounting system, makes payments to their staff, and hundreds of other things.

How do machines make decisions?

- think of machine learning as a way to arrive at a decision, based on patterns in a dataset.

# Target/outcome/dependent/predicted variables

- Target variables come in two main flavors:

- Categorical

- Continuous

- Categorical variables include things like yes or no; and north, south, east, or west. An important distinction in our machine learning work in this book is whether the categorical variable has only two categories or has more than two categories. If it has only two categories, it is called a binary target variable. If it has more than two categories, it is called a multiclass target variable.

- Continuous variables are numbers. For example, if your machine learning application predicts house prices based on features such as neighborhood, number of rooms, distance from schools, and so on, your target variable (the predicted price of the house) is a continuous variable. The price of a house could be any value from tens of thousands of dollars to tens of millions of dollars.

- Features are variables used to predict the outcome/target variable. For example, no. of rooms, distance from school in the above example.

Machine Learning

- Improve decision making and solve business challenges (i.e. customer churn)
- Improve business processes
- ML features as a business
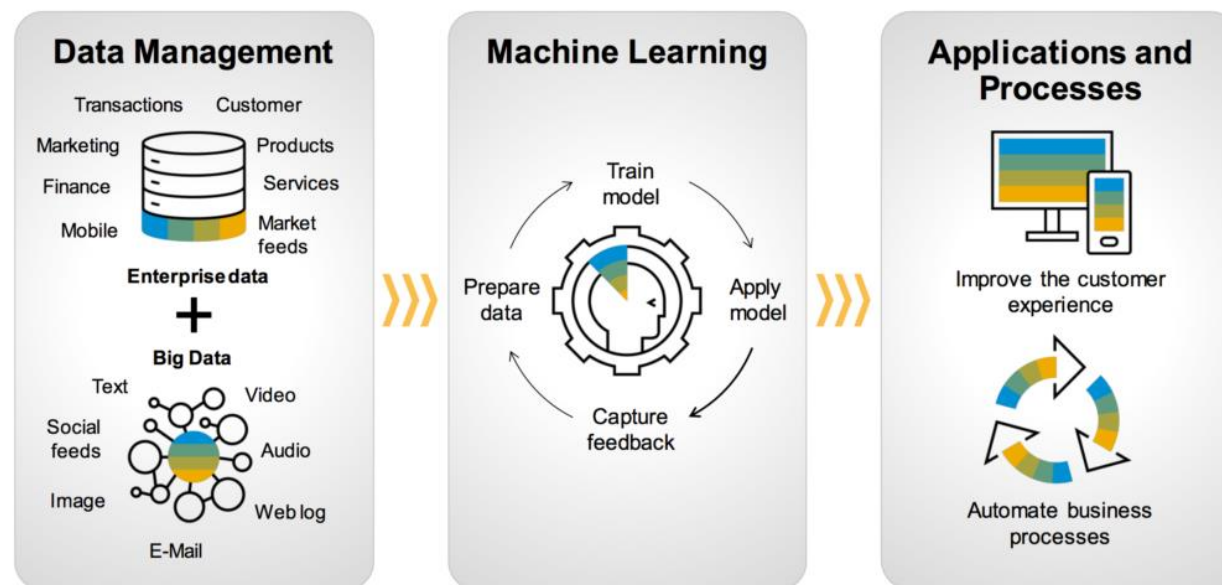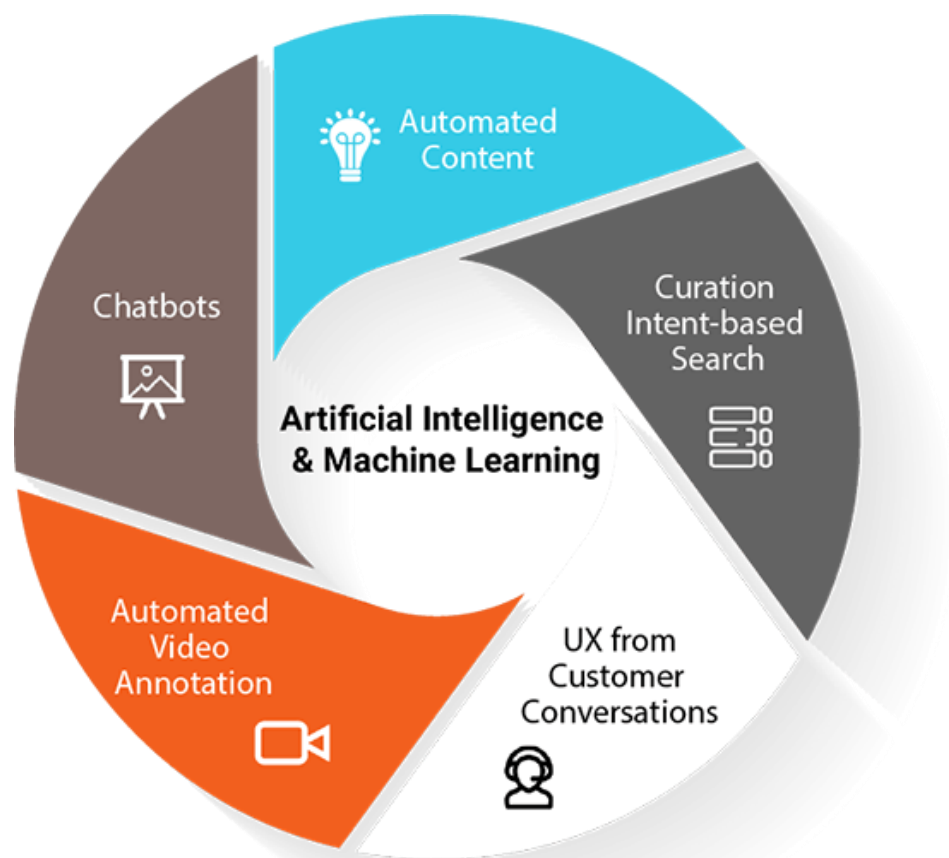
Machine Learning
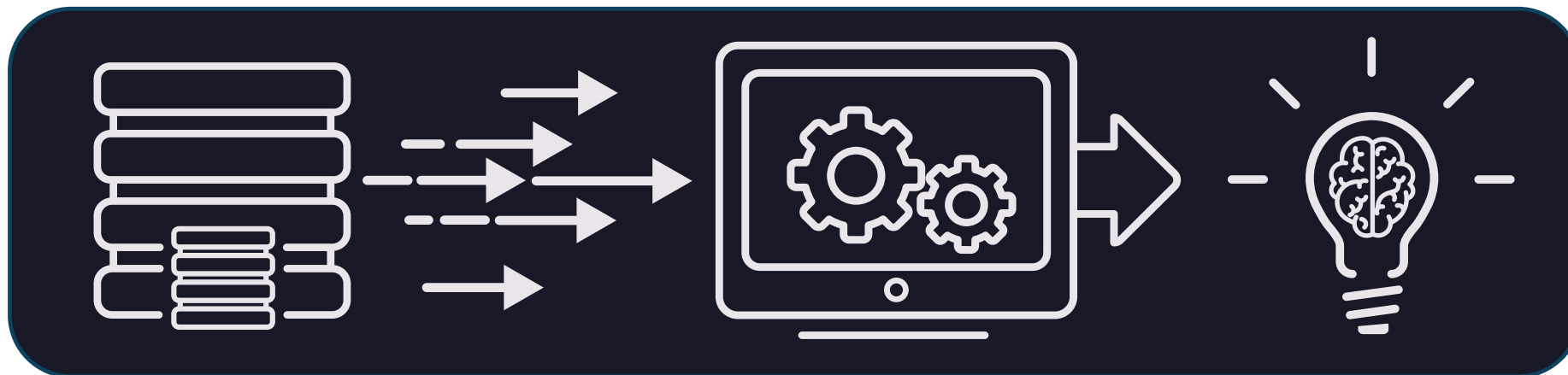
Automate

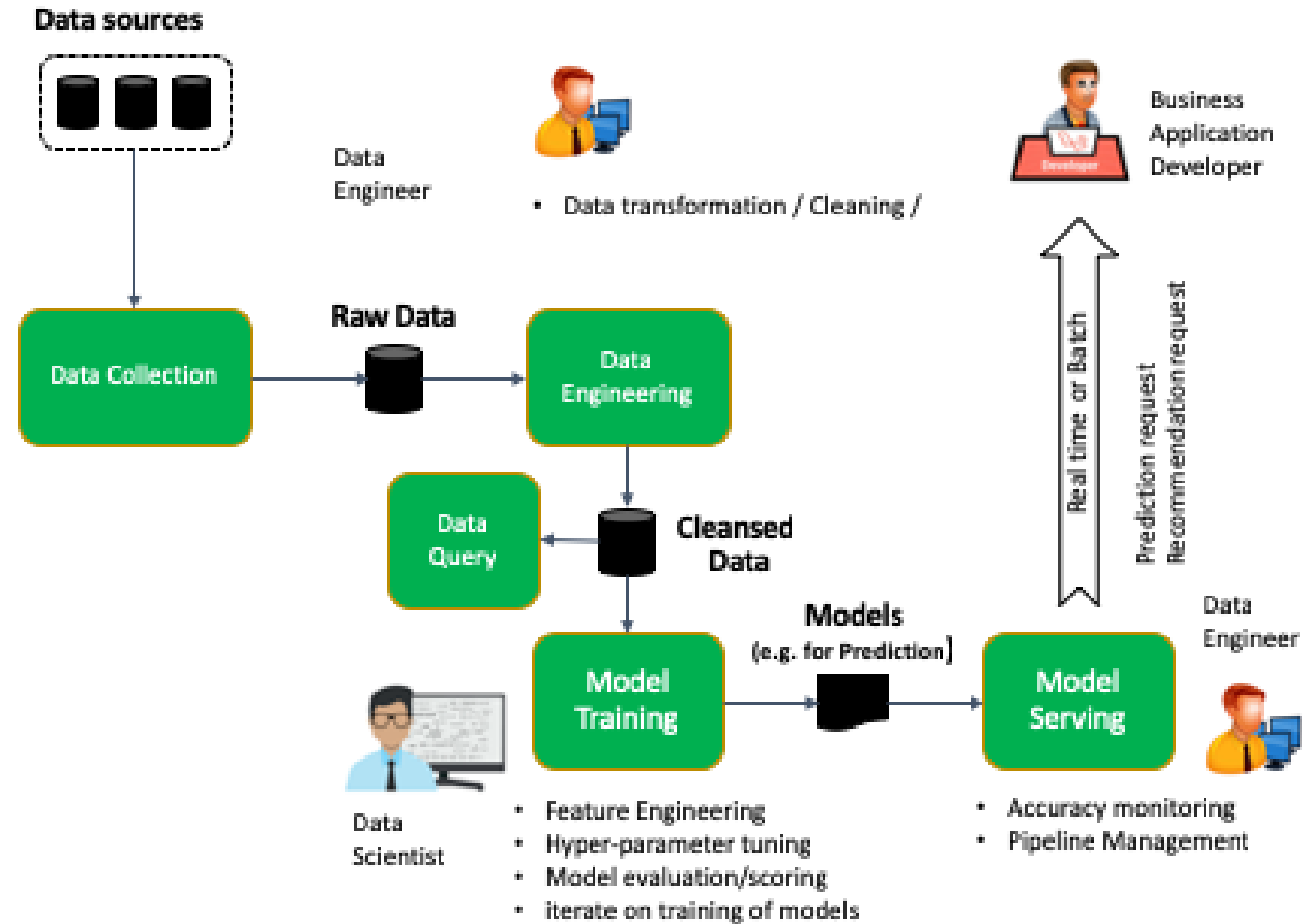Provide automation to the model building process by minimizing human intervention.

Customize

Build powerful models using state-of-the-art algorithms

Accelerate

Fast response time for sophisticated analytics applied to data of any size or complexity.
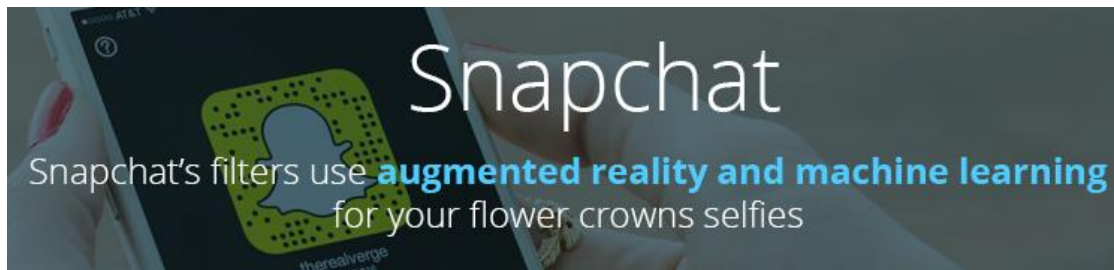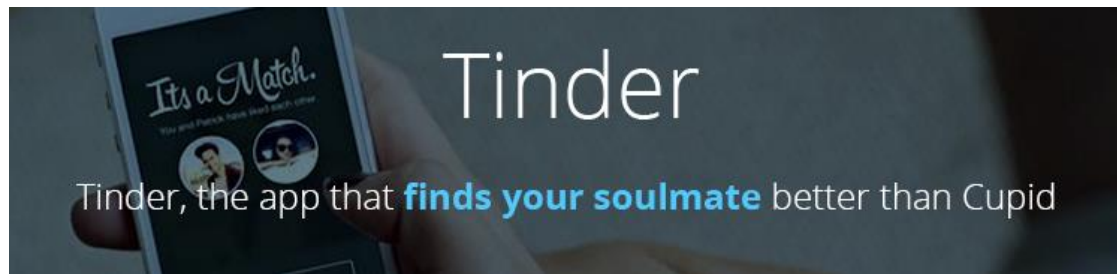
Technological advances

- **The advent of the Internet economy and the explosion in mobile apps** has caused a deluge of data waiting to be turned into value.

- industries that have made the transition to the digital world — think newspapers, TV, hospitality, food delivery, insurance, etc.

- **The sharp decrease in costs associated with data storage and processing.**

Global Mobile Data Traffic Growth / Regions
MEA has the Highest Growth Rate (72%) from 2014–2019
APAC will Generate 39% of all Mobile Data Traffic by 2019

57% CAGR 2014–2019

Exabytes per Month

- Latin America (LATAM)
- Middle East and Africa (MEA)
- Central and Eastern Europe (CEE)
- Western Europe (WE)
- North America (NA)
- Asia Pacific (APAC)

2019 values: 8.4%, 12.5%, 14.4%, 9.9%, 15.7%, 39.1%

Years: 2014, 2015, 2016, 2017, 2018, 2019

Source: Cisco VNI Global Mobile Data Traffic Forecast, 2014–2019

CISCO

© 2014 Cisco and/or its affiliates. All rights reserved. Cisco Confidential 8

Examples of ML applications

Netflix, the app that **knows what you want**

Tinder, the app that **finds your soulmate** better than Cupid

Oval Money, a finance mobile app that knows how to use **machine learning the right way**

Snapchat's filters use **augmented reality and machine learning** for your flower crowns selfies

Google Maps knows where are **the best parking spots** thanks to machine learning

Machine learning for business

- Approving purchase orders

- Evaluating which customers are at risk of churning

- Identifying support requests that should be escalated immediately

- Auditing invoices from suppliers

- Forecasting operational trends, such as power consumption, using time-series data

- Incorporating additional features into time-series forecasting
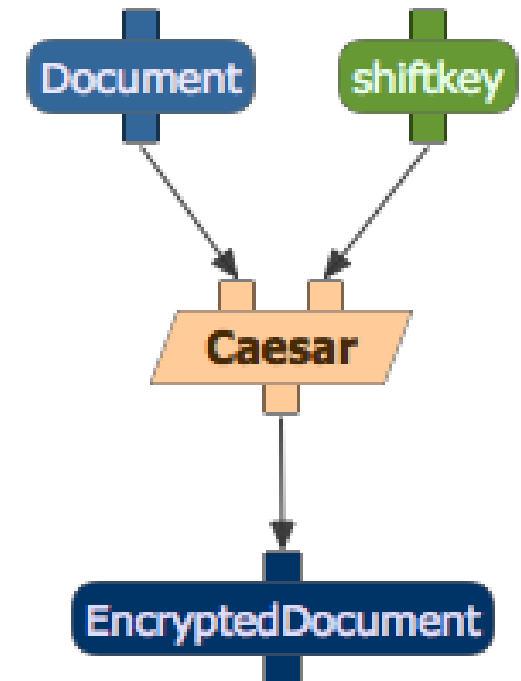
Today's Business Challenges

Fraud

Targeted Marketing

Financial Risk

Churn

Treat Programs as "Black Boxes"

- You don't have to understand complex mathematics and programming in order to use software

- This is why we often refer to software as a "black box"

- You only need to understand inputs and outputs and the program's function in order to use it correctly

Learning Approaches

Supervised Learning

- The training data is annotated with information to help the learning system

  - **Estimation** of a continuous outcome (ex: profit, cost, delivery time, etc.). For example, regression.
  - **Classification** of a categorical outcome (win/lose/tie, fraud/not fraud, etc). For example, logistic regression, KNN, decision trees.

Unsupervised Learning

- The training data is not annotated with any extra information to help the learning system

Example of Approaches:

- Association Rules

- Cluster Analysis

Reinforcement Learning

Example:
https://aws.amazon.com/deepracer/

Examples of approaches (to be discussed in this class)

- Association rules:

  "80% of customers who buy *cheese* and *milk* also buy *bread*, and 5% of customers buy all of them together"

- Clustering

  identifying a set of similarity groups in the data

- KNN

  to classify a categorical outcome or predict a continuous outcome

  - Classification:

  mining patterns that can classify future data into known classes.

Why unsupervised Learning?

• The main task of unsupervised learning is to **find patterns in the data**.
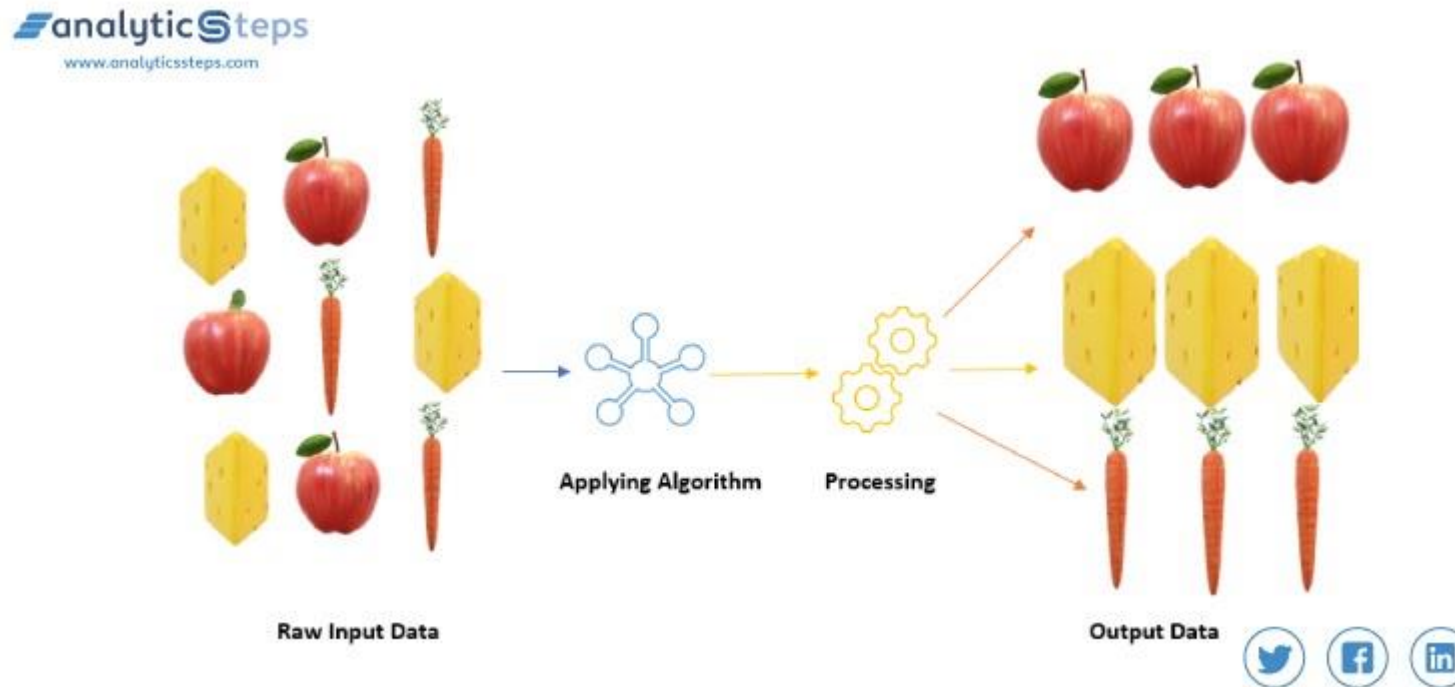
Why Supervised Learning?

- Supervised learning allows you to collect data or produce a data output from the previous experience.

- Supervised machine learning helps you to solve various types of real-world computation problems.

- For example, you want to train a machine to help you predict <u>how long it will take you to drive home from your workplace</u>.
  - Weather conditions, Time of the day, day of the week, Holidays are your inputs. The output is the amount of **time** it took to drive back home on that specific day.

  - Other examples:
    - Output: Profit. Input: advertising expenses.
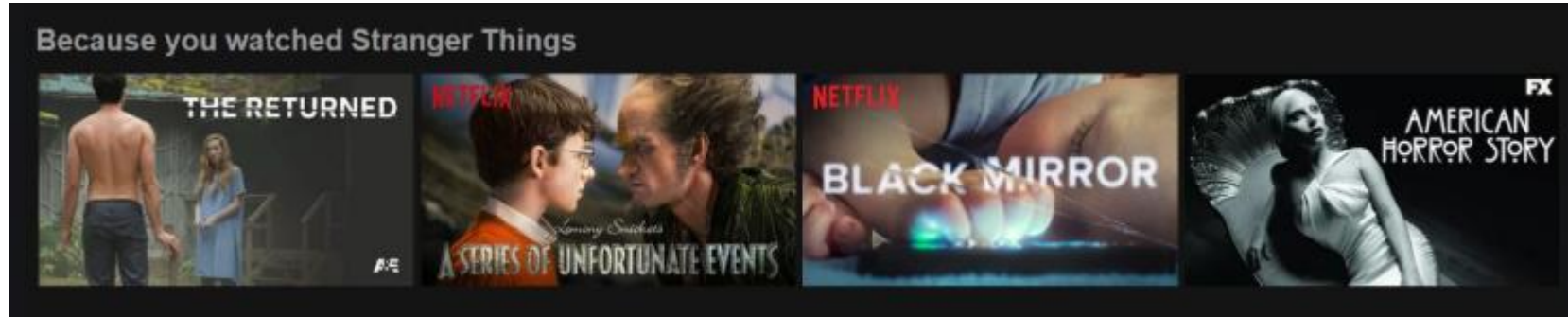   - Output: Health index. Input: Age, gender, income, etc.

Example – Recommender systems
Would You Like to Find a New Movie?

Content-based filtering

- Is it because the movies are like a movie you recently enjoyed?
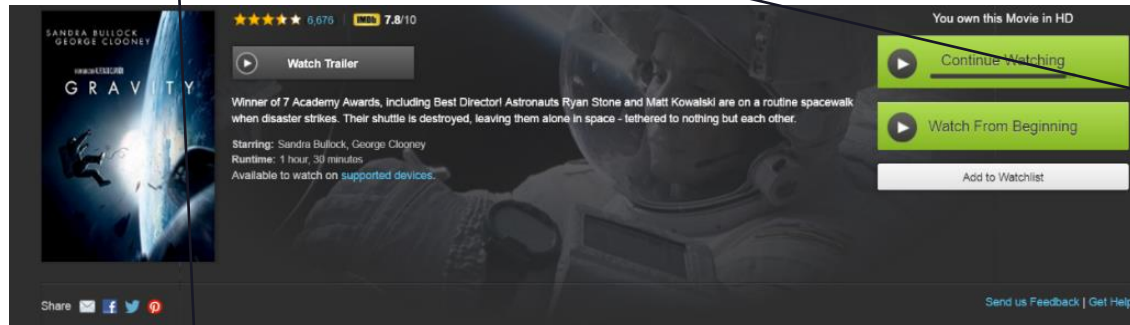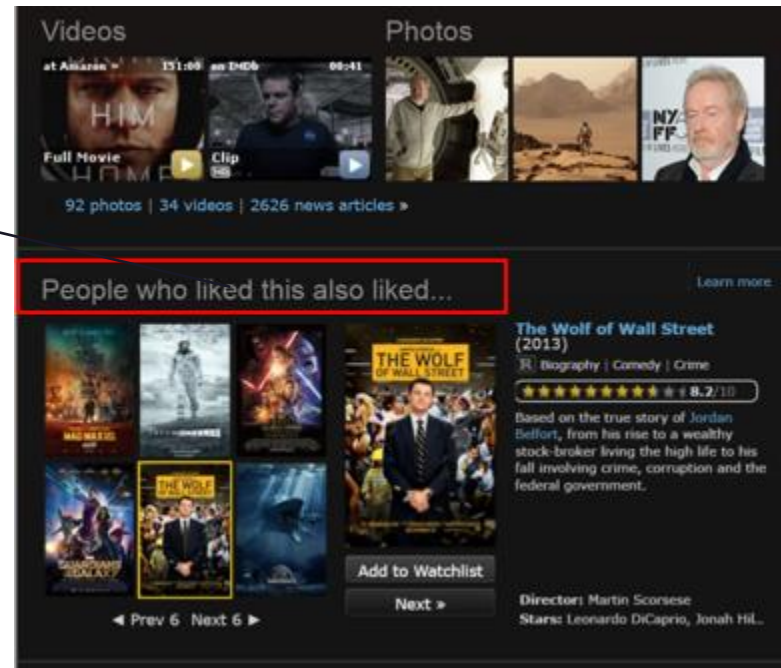 - You enjoyed movie E. How about A, B, C, D?

Because you watched Stranger Things

Collaborative filtering

Is it because viewers like you enjoyed these movies?

- People like you also enjoyed movie A and movie B.

Watched and liked
Vs
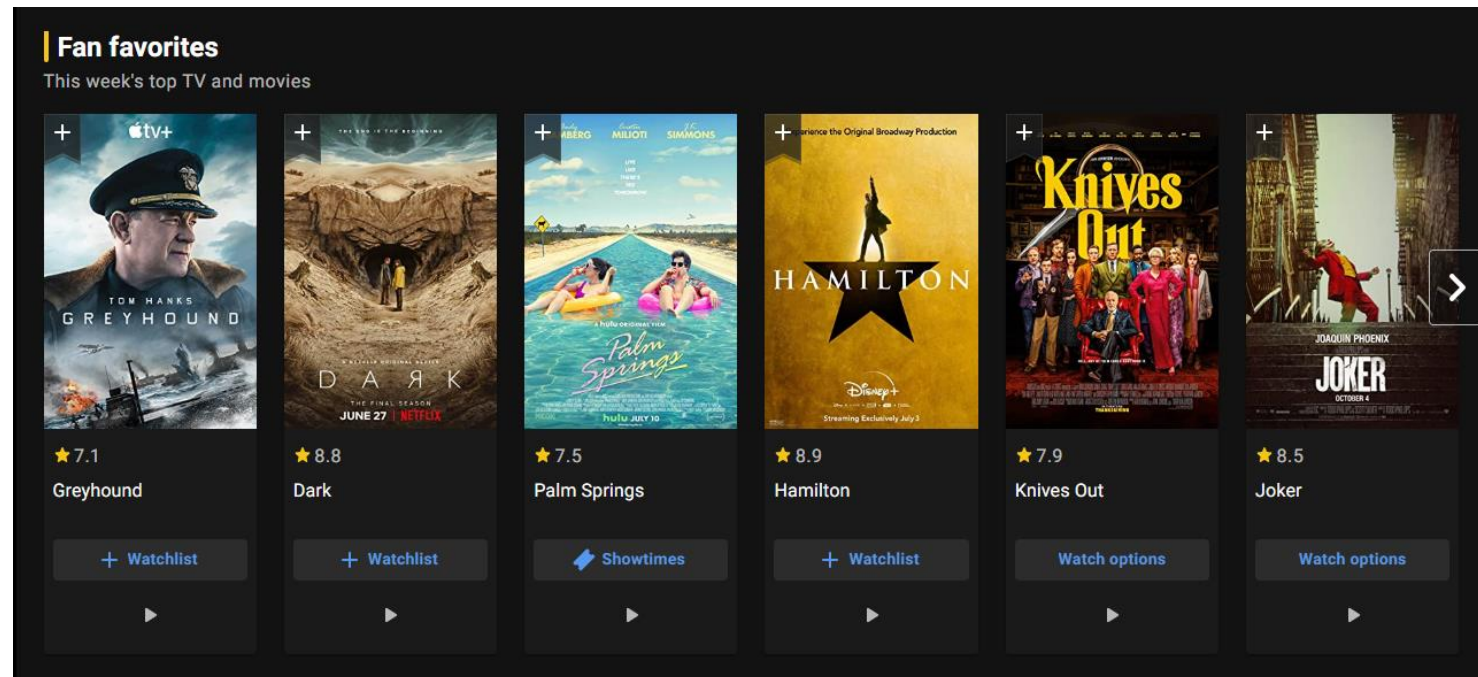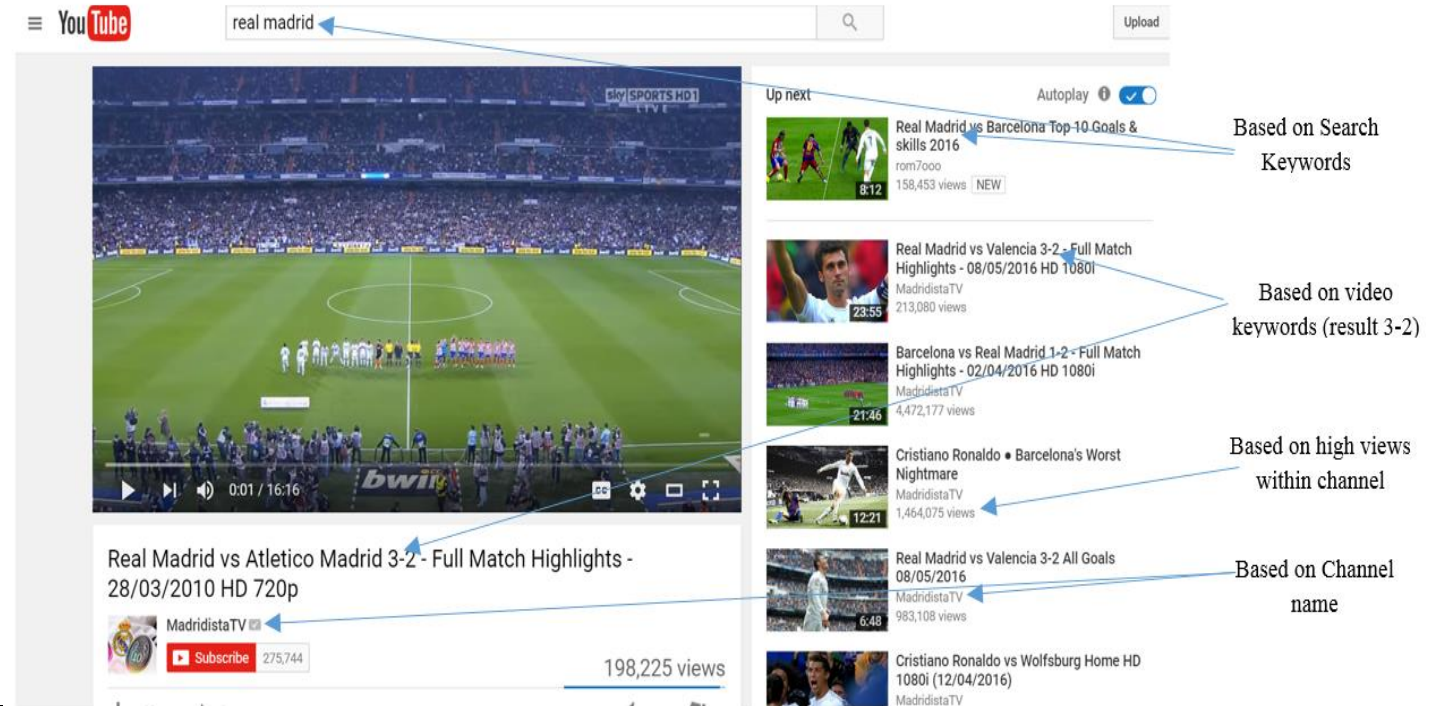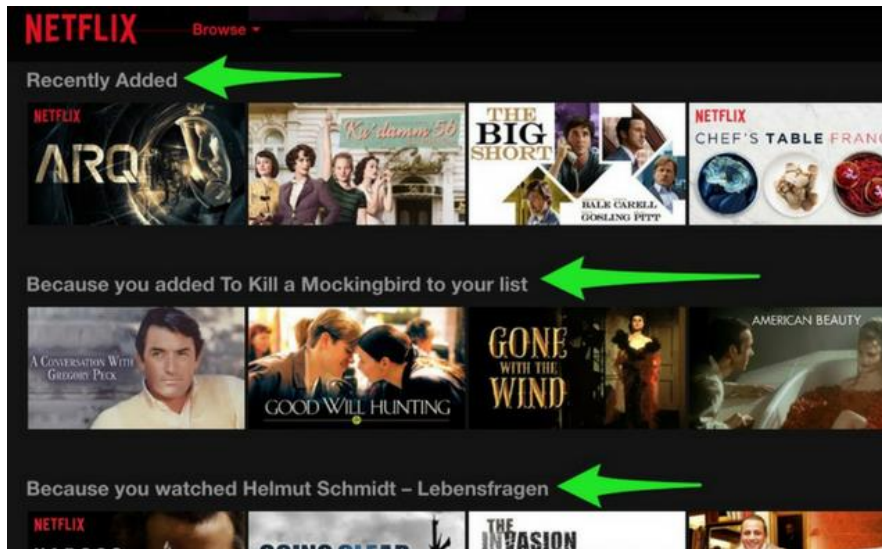Watched

Cold start solution

- Is it because most people like these movies?
    - Movie E is an award winner, and it is a safe recommendation if I know nothing about you.
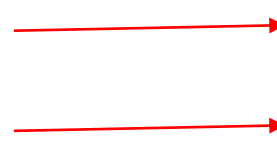
Hybrid system

- Is it a little bit of each of these?

Based on Search Keywords

Based on video keywords (result 3-2)

Based on high views within channel

Based on Channel name

Illustration

	M1	M2	M3	M4	M5
U1	5	-	4	-	3
U2	2	-	-	4	5
U3	-	4	3	3	-
U4	1	2	1	-	-
U5	1	3	4	-	5
U6	4	4	-	5	4
U7	4	4	5	-	4
U8	3	-	-	4	5
U9	-	3	2	1	1
U10	5	4	-	3	-

The 6 core AI Intents

Accelerate research and discovery

Spend less time looking for the information you need and more time acting on it. Easily process millions of data points to focus on the work that matters most.

For example, 2.5 million scientific papers are published in English every year. When this unstructured text becomes structured, researchers spend their time looking at the most important insights.

The 6 core AI Intents

Recommend with confidence

Teach systems the nuances of your business to ensure every factor is considered so that you can make better informed decisions, give tailored advice, and deepen customer relationships.

For example, customers enjoy tools or experts that help them purchase the right product. This could be anything from shopping for a dress to choosing the right savings plan.

The 6 core AI Intents

Enrich your interactions

Train your AI on common requests to reduce response times, increase the number of transactions, and make interactions more productive.

For example, AI can enable call center agents to better address caller needs. This helps provide a consistently positive experience and saves businesses time and money.

The 6 core AI Intents

Anticipate and preempt disruptions

Continuously monitor the condition of your systems to mitigate problems before they disrupt your work. Use AI to catch potential issues in the systems and processes that are essential to your business.

One application of this is public transit maintenance. $398M is lost every year due to subway delays in New York City. Predictive maintenance can help avoid downtime for users and save money for the city.

The 6 core AI Intents

Scale expertise and learning

By combining employee expertise with your industry's latest learnings, each of you knows as much as all of you. Make deep instructional knowledge available to everyone in your organization.

This use of AI offers solutions to problems faced in legal practice, for example. 38% of law partners will retire in the next decade. AI can comb through legal documents to help summarize successful tactics for legal cases.

The 6 core AI Intents

Detect liabilities and mitigate risk

Train systems to understand and keep up with constantly-changing regulations and privacy obligations. Identify compliance issues quickly and easily to protect your business and your people.

For example, about €20M in fines could be imposed on companies failing to comply with GDPR. With AI, analysts can help uncover compliance breaches and follow the right course of action.

Types of Data needed for AI

Public data

This can be found in the world (for free or for purchase). For example, census data is free and highly accessible. Anyone can go download those spreadsheets. Even social media hashtags, likes, and comments are public data. Sometimes, public data isn't enough, though. For example, businesses regularly purchase additional demographic data for user segmentation.

Private data

This type of data is owned and held by your business, which often provides a competitive advantage. This sometimes takes the form of employee records, hardware assets, inventory logs, payroll, and more.

User data

User data can be linked to individual users. Think contact information, medical history, or geographic location. All of these variables belong to the user. Businesses must ask for access to store or use the data.