

<https://attendance.lincoln.ac.uk>

Access Code: **280562**

CMP9794M

Advanced Artificial Intelligence

[Heriberto Cuayahuitl](#)



UNIVERSITY OF
LINCOLN

School of Computer Science

AAI Delivery Team

- **Dr. Heriberto Cuayahuitl**
hcuayahuitl@lincoln.ac.uk
office: INB3108



- **Dr. Riccardo Polvara**
rpolvara@lincoln.ac.uk
office: INB3121



- **Mr. Omar Ali (demonstrator)**
28587497@students.lincoln.ac.uk



About Me



- British computer scientist (born in Mexico),
- PhD from the University of Edinburgh in 2009,
- Areas of interest: conversational AI & robot learning.

Recommendations

- Record your attendance in every session:
 - online <https://attendance.lincoln.ac.uk>
 - paper-based register (when online system is off)
- Check Blackboard and your email frequently for updates and notifications
- Mute your mobile phone during lectures
- If you want to ask/say something, raise your hand or just interrupt your lecturer (that's okay)

Lectures and Workshops

- Lectures: Monday 9-11am in MB0302
- Workshops: Monday 11am-1pm in INB1102

<div><div><</div><div>Week 1</div><div>></div></div>		<div><div>24-25</div><div></div></div>		<div><div>Single Week</div><div></div></div>		<div><div>Full</div><div></div></div>		<div><div>Table (Horizontal)</div><div></div></div>		<div><div>Export/Print</div><div></div></div>						
<div>Monday</div> <div>(23rd Sep 24)</div>	9am				10am				11am				12pm			
	CMP9794								CMP9794							
	<div>MB0302</div> <div>9am - 11am</div> <div>Weeks 1-5, 7-12, 16</div> <div>Heriberto Cuayahuitl Portilla, Riccardo Polvara</div> <div>LECTURE</div>								<div>INB1102</div> <div>INB1301</div> <div>11am - 1pm</div> <div>Weeks 1-5, 7-12, 16</div> <div>Heriberto Cuayahuitl Portilla, Riccardo Polvara</div> <div>WORKS</div>							
<div>Tuesday</div> <div>(24th Sep 24)</div>																
<div>Wednesday</div> <div>(25th Sep 24)</div>																
<div>Thursday</div> <div>(26th Sep 24)</div>																
<div>Friday</div> <div>(27th Sep 24)</div>																

Main Topics in this Module

- **Quantifying uncertainty**
 - Introduction to probability theory
- **Probabilistic reasoning**
 - Bayesian Nets (BNs), Gaussian Processes (GPs)
- **Reasoning over time**
 - Dynamic BNs, Hidden Markov Models (HMMs)
- **Making complex decisions**
 - Markov Decision Processes (MDPs)
 - Reinforcement learning (without neural nets)

Agenda

Week	Commencing on	Topic	Delivered by
1	23/09/2024	Introduction to module	Heriberto Cuayahuitl
2	30/09/2024	Bayesian nets w/exact inference	Heriberto Cuayahuitl
3	07/10/2024	Structure learning	Heriberto Cuayahuitl
4	14/10/2024	Bayesian nets w/approx. inference	Heriberto Cuayahuitl
5	21/10/2024	Gaussian Bayesian networks	Heriberto Cuayahuitl
6	28/10/2024	Reading week (no lectures/workshops)	
7	04/11/2024	Gaussian processes	Heriberto Cuayahuitl
8	11/11/2024	Probabilistic reasoning over time I	Ricardo Polvara
9	18/11/2024	Probabilistic reasoning over time II	Ricardo Polvara
10	25/11/2024	Introduction to complex decision making	Ricardo Polvara
11	02/12/2024	Markov decision processes	Ricardo Polvara
12	09/12/2024	Reinforcement learning	Ricardo Polvara
13-15		Christmas break	
16	06/01/2025	Present and Future of AI: Discussion	Riccardo and guests

Learning Objectives

LO1. Critically appraise a range of AI techniques for knowledge representation, reasoning and decision-making **under uncertainty**, identifying their strengths and weaknesses, and selecting appropriate methods to serve particular roles

LO2. Design and develop software algorithms for solving complex (non-trivial) AI problems in an application domain of interest.

Assessments*

- Assignment (50%):
 - Bayesian Networks (BNs)
 - Gaussian BNs
 - Gaussian Processes
- In-class test (50%):
 - Mock-test (last workshop)
 - Test (see Hand-in deadlines and your timetable)



*You should **READ** the Assessment docs in Blackboard

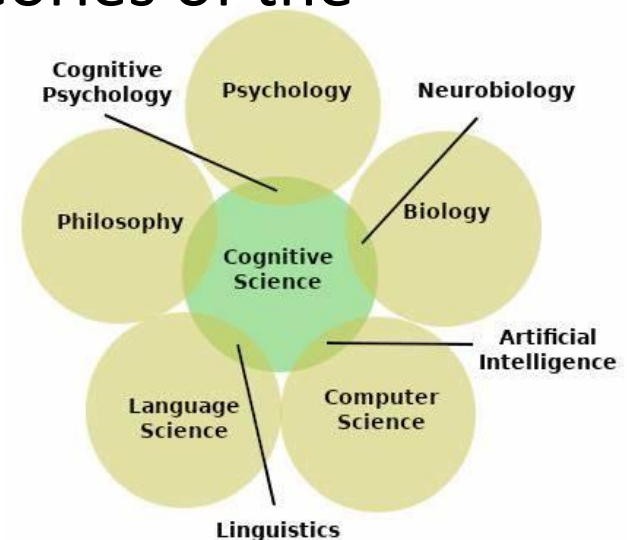
What is AI?

Many definitions. Example “the science and engineering for equipping machines/robots to acquire their own behaviour”.

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

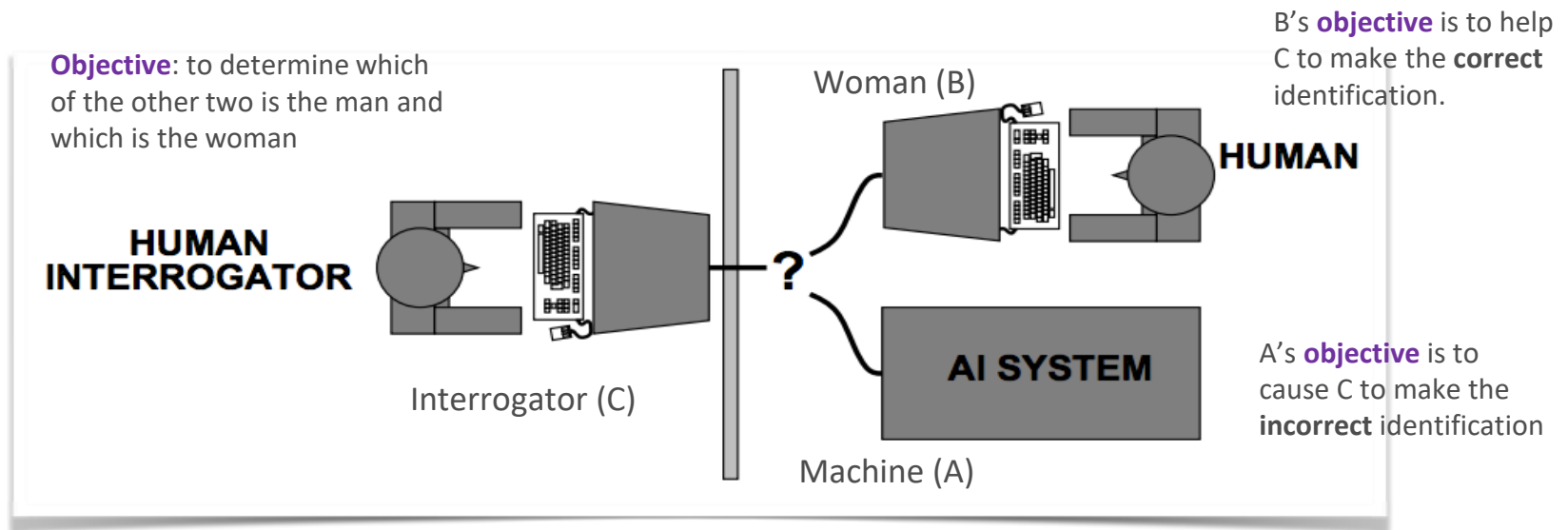
Thinking Humanly: Cognitive Modelling

- We must have a way of determining how humans think:
 - Introspection (catching our own thoughts as they happen)
 - Psychological experiments (observing a person in action)
 - Brain imaging (observing brain in action)
- Cognitive science brings multiple fields together to try to construct precise and testable theories of the workings of the human mind.



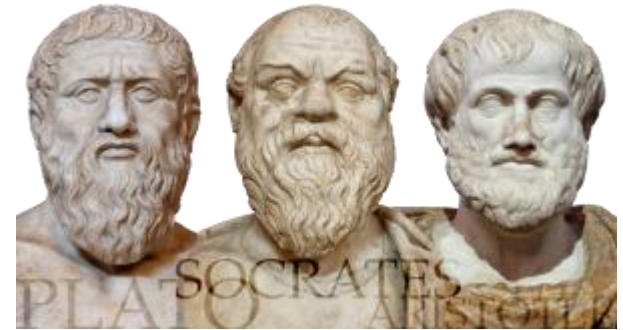
Acting Humanly: The Turing Test (1950)

- Turing (1950) “Computing machinery and intelligence”
- Can machines behave intelligently?
- [Summary of Alan Turing’s paper \(1950\)](#)



Thinking Rationally: Laws of Thought

- Several Greek scholars developed various forms of logic: **notation** and **rules** for thoughts
- AI hopes to create intelligent systems using logic programming (e.g., Prolog).
- However, it is not easy to represent informal knowledge using logical notation, particularly when knowledge is **not 100% certain**.



Acting Rationally

- **Rational behaviour:** doing the right thing
- **The right thing:** which is expected to maximise goal achievement – given the available information to the AI.
- **Rational Agent** is one that achieve the best outcome or, when there is **uncertainty**, the best expected outcome.

Birth of AI

- The Dartmouth Conference (1956) brought together researchers in a variety of topics:
 - complexity theory, language simulation, neuron nets, abstraction of content from sensory inputs, relationship of randomness to creative thinking, learning machines



John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



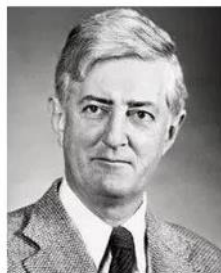
Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



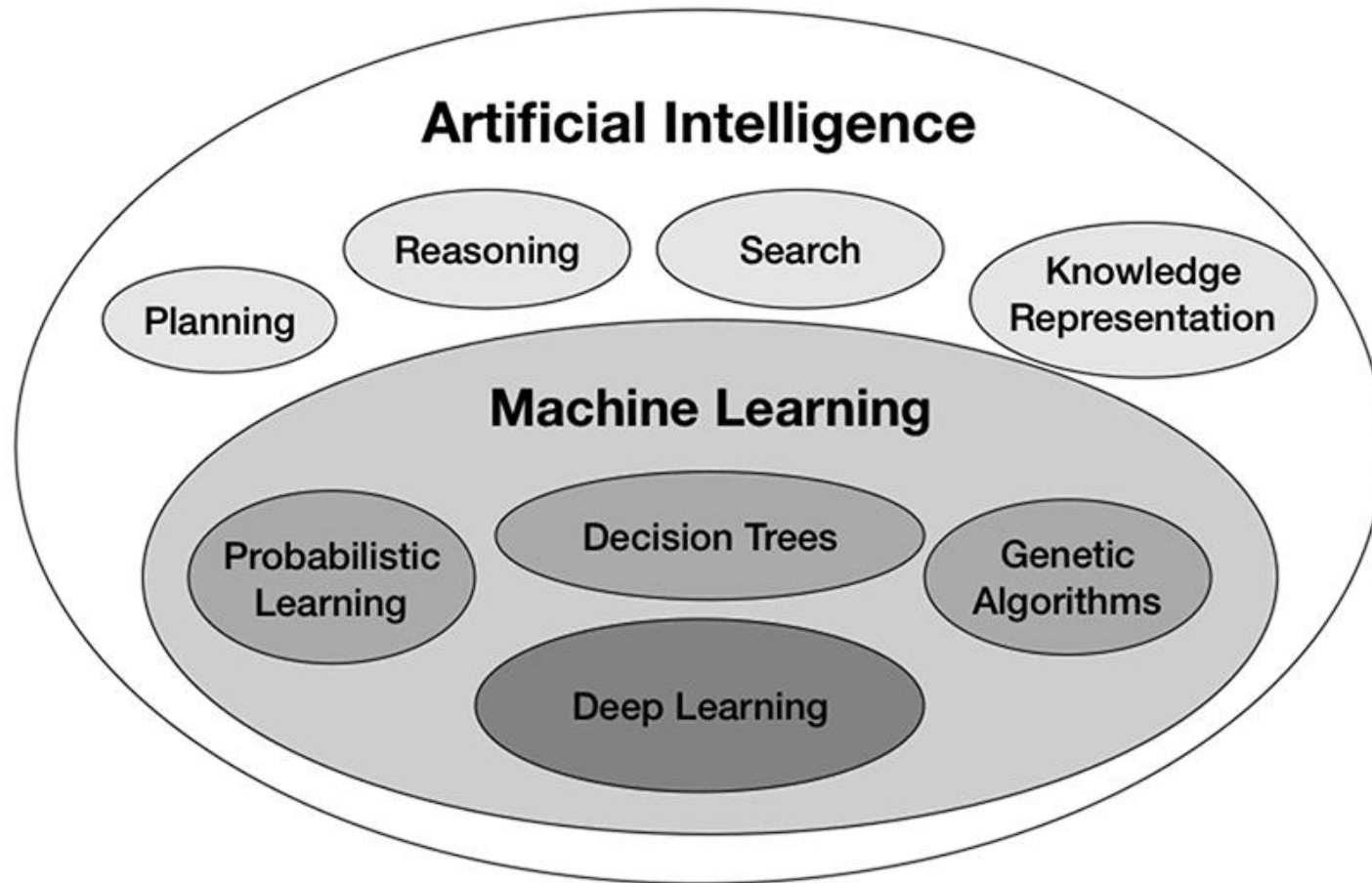
Nathaniel Rochester



Trenchard More

Picture from
<https://www.scienceabc.com>

Pillars of AI



Major Areas in AI

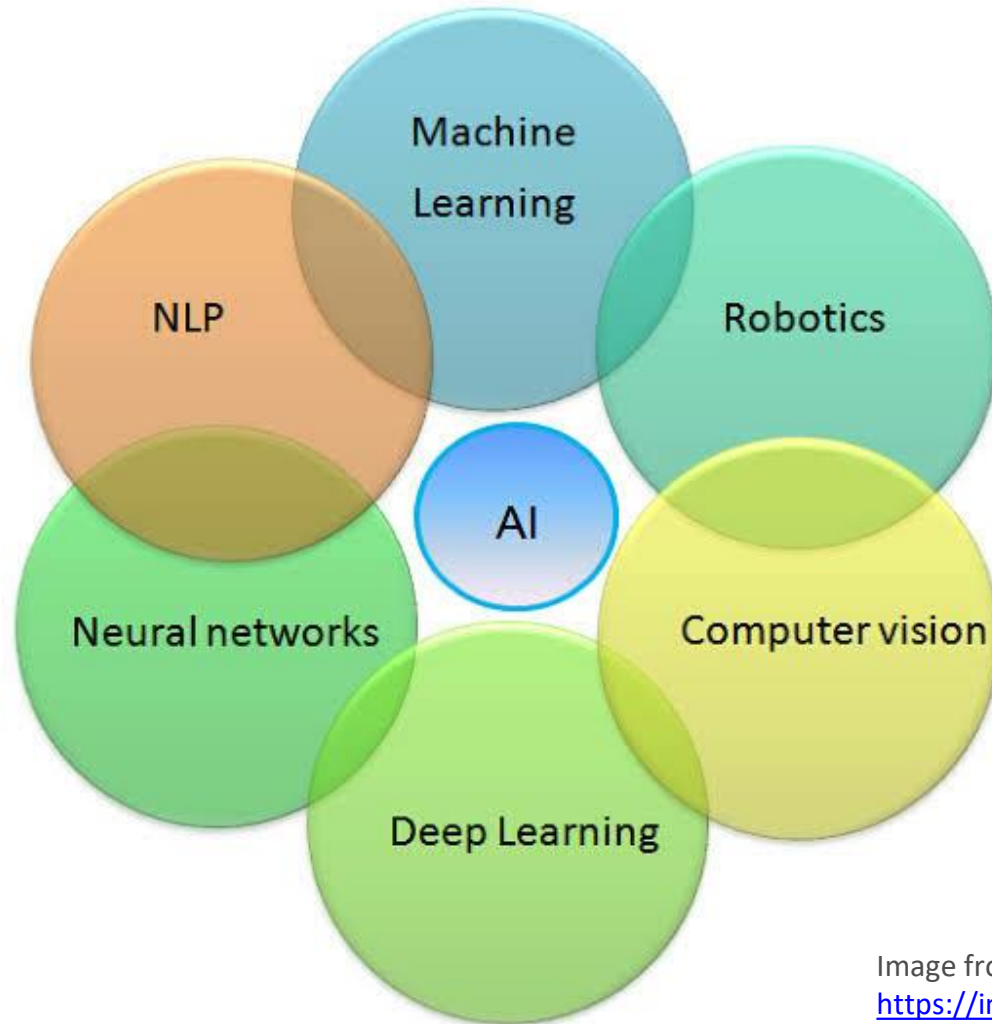
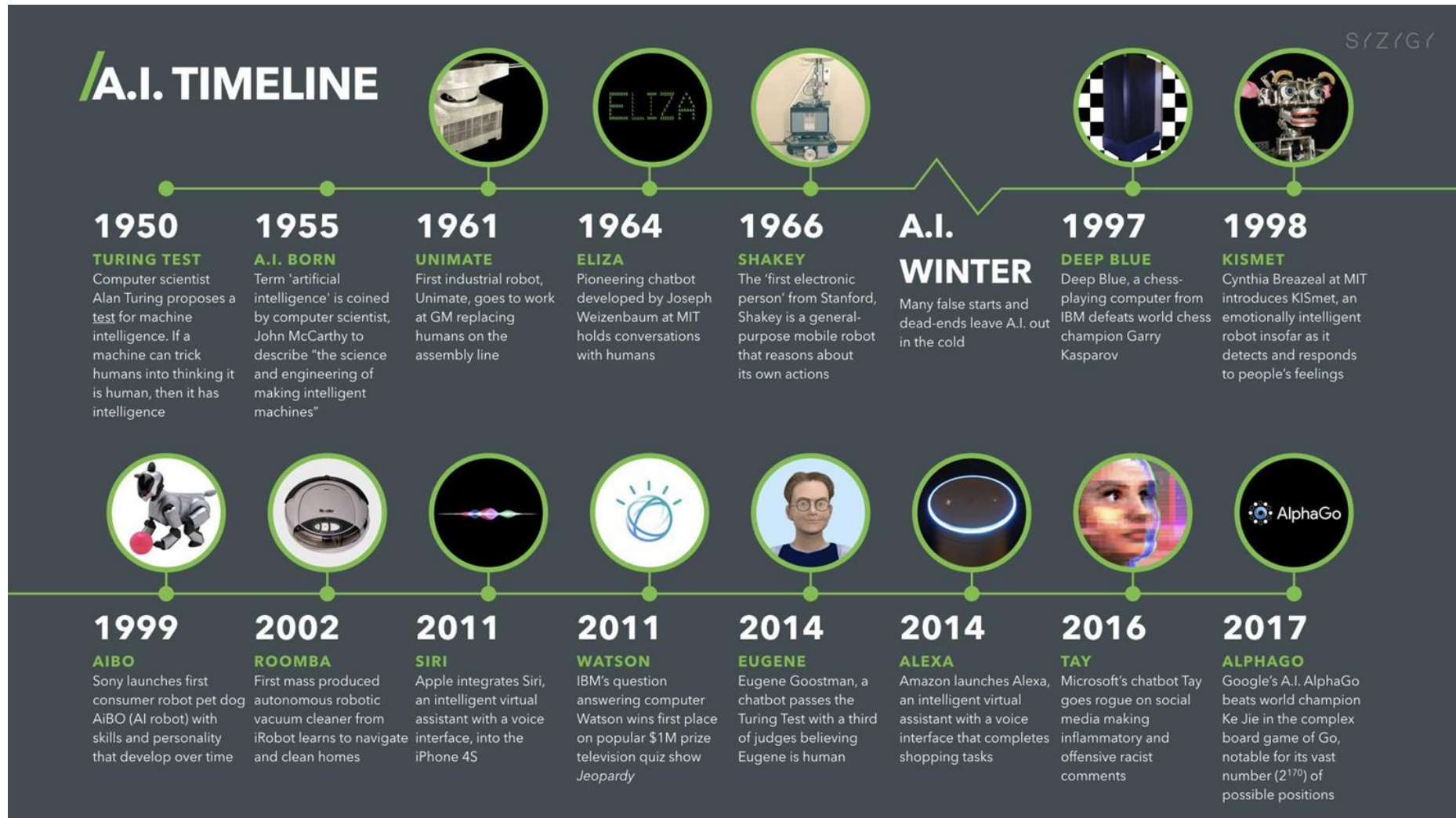
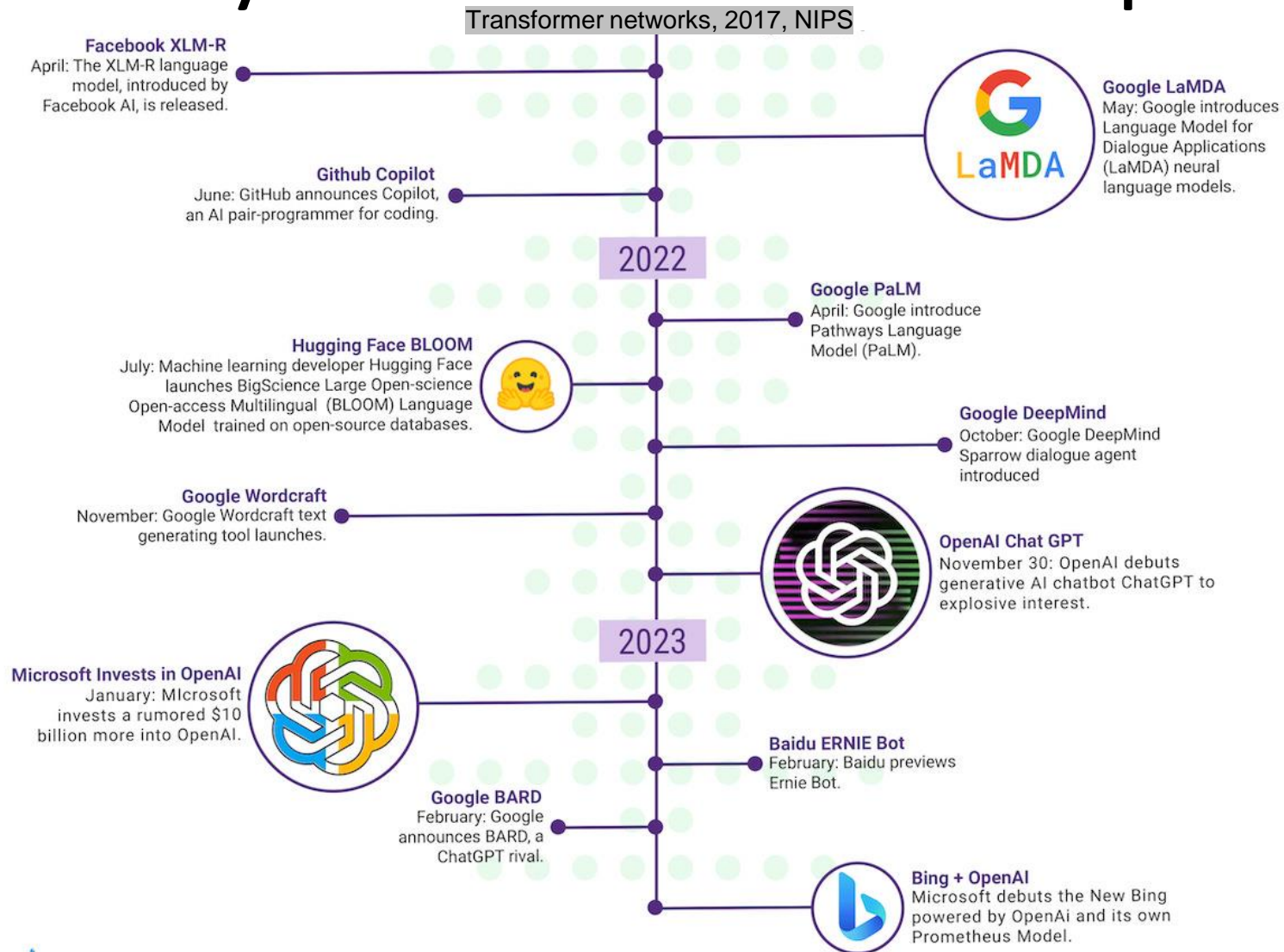


Image from
<https://infohubns.blogspot.com/2020/07/research-areas-of-ai.html>

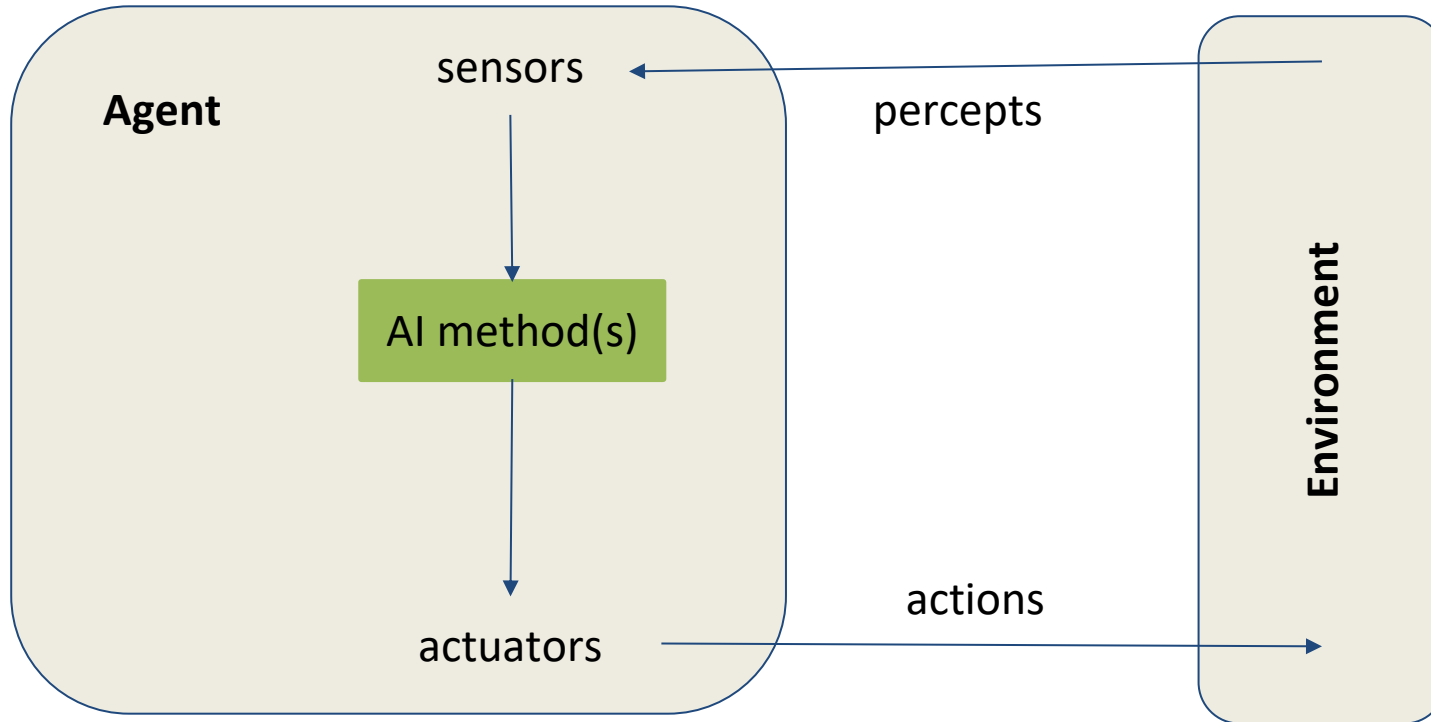
History of AI in a Nutshell – part 1



History of AI in a Nutshell – part 2



AI Agents



AI Agent { perceives its environment via its sensors,
makes decisions using AI techniques, &
executes decisions using its actuators.

Properties of Task Environments

- Fully-observable vs. partially observable
- Single-agent vs. multi-agent
- Deterministic vs. stochastic
- Episodic vs. sequential
- Static vs. dynamic
- Discrete vs. continuous
- Known vs. unknown

More details of those properties in Russell and Norvig Chapter 2.

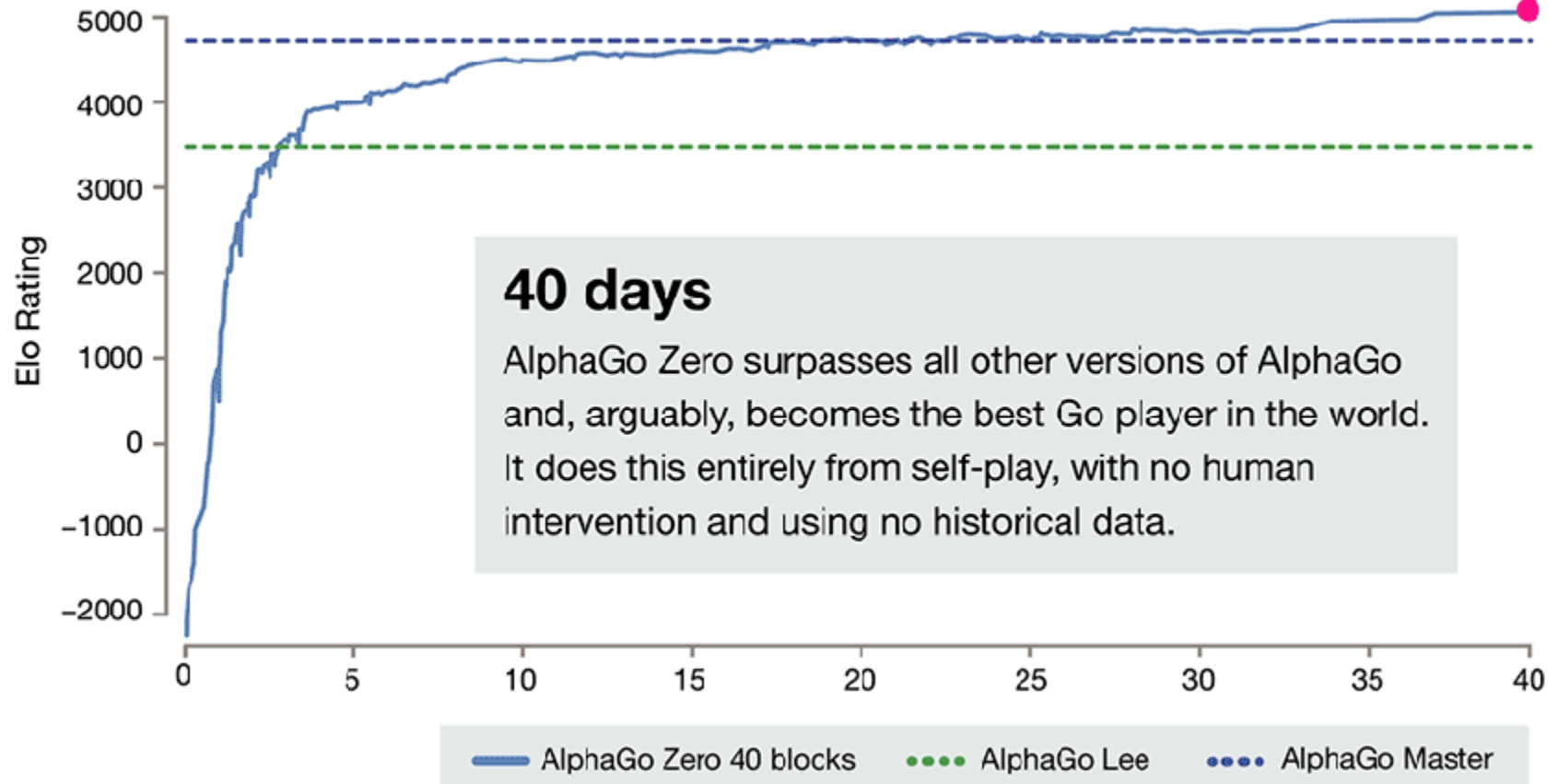
Hardest case (in red)



Example1: Autonomous cars driving on unfamiliar roads.

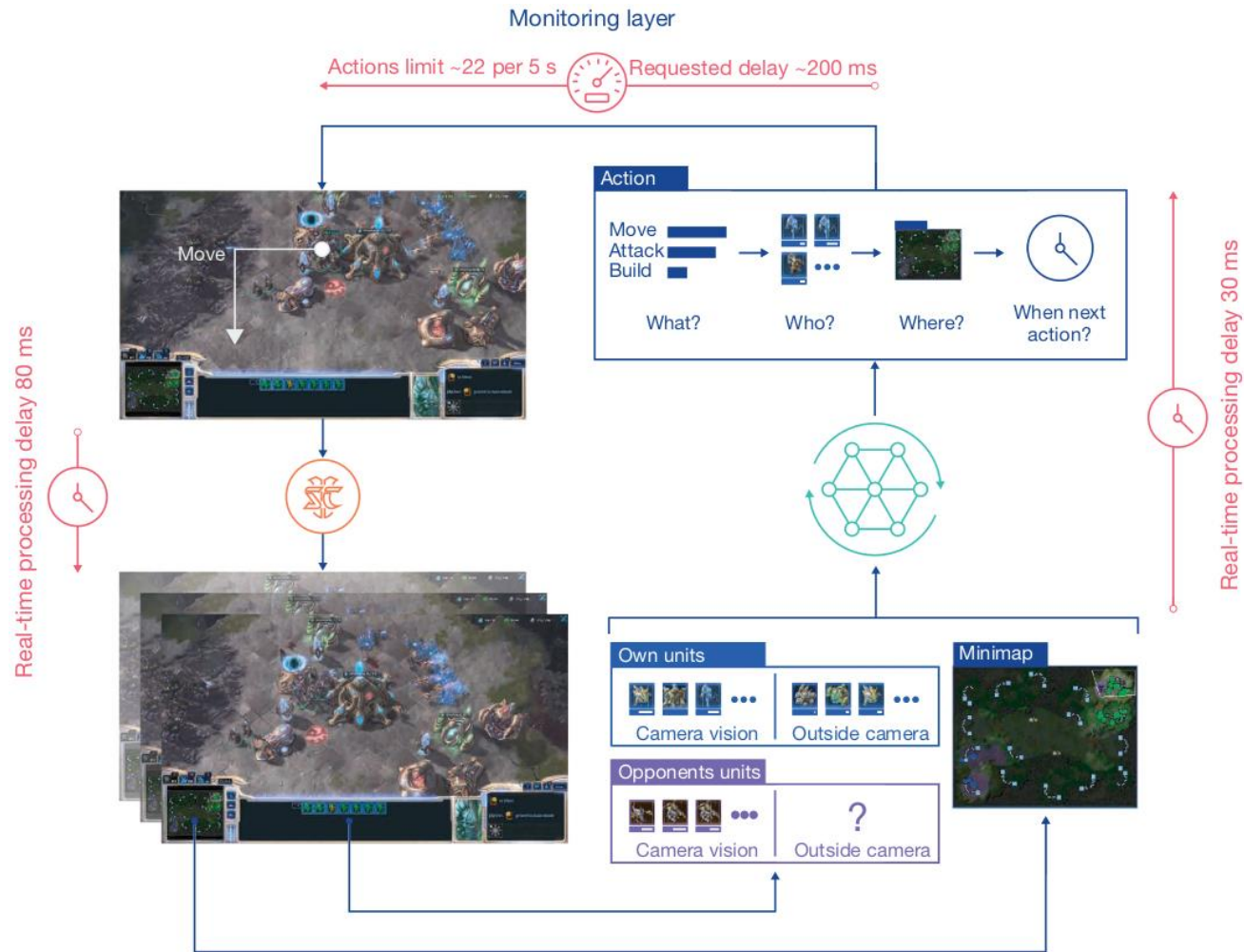
Example2: Robots cooking food in a changing/dynamic kitchen.

Recent Major Developments in AI



[DeepMind's AlphaGo learned to play the game of Go from scratch, 2017](#)

Recent Major Developments in AI



DeepMind's AlphaStar learns to play StarCraft II at GrandMaster level, 2019

Some Recent Developments in AI

Human View

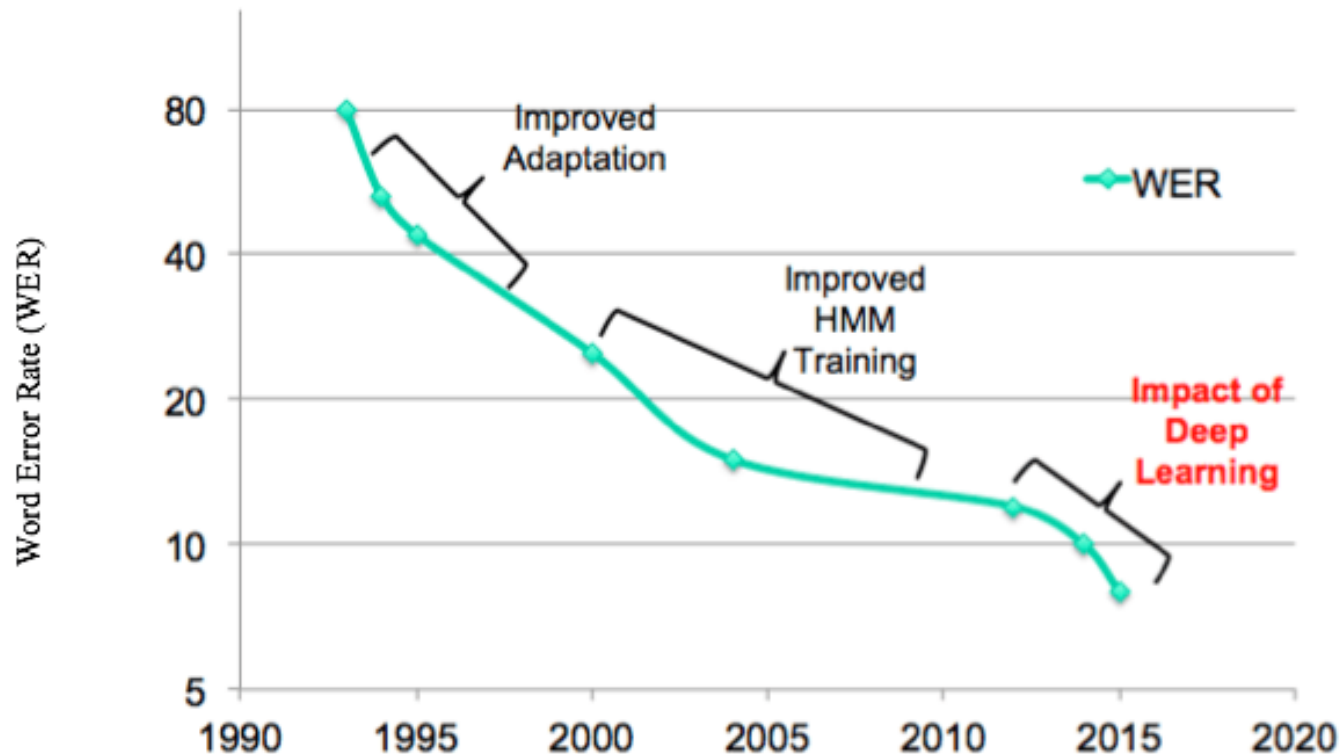


AI View

3.006	-1.386	-0.4695	0.883	1	0.84
-0.3154	-0.5425	-0.5	0.866	0	0.82
3.11	-1.36	-0.9336	0.3584	1	0.78
-2.324	2.863	0.9746	0.225	0	0.86
3.037	-1.361	-0.7773	0.6294	1	0.82
-1.387	2.951	0.988	0.1565	0	0.74
3.023	-0.9395	0.05234	-0.9985	0	0.66
2.951	-0.5747	0.01746	1	0	0.72
2.963	-1.303	0.3906	0.9204	0	0.68
2.834	-3.164	0.01746	-1	0	0.68
3.127	-1.368	0.6562	0.755	1	0.55
3.088	-1.366	0.4695	0.883	0	0.55
2.984	-1.398	-0.225	0.9746	1	0.55
3.037	-1.391	0.788	0.6157	0	0.55
3.076	-1.438	0.883	0.4695	0	0.55
-2.412	2.846	0.996	0.08716	1	0.3

[OpenAI Five is the 1st AI system to defeat world champs at e-sports, 2019](#)

Some Recent Developments in AI

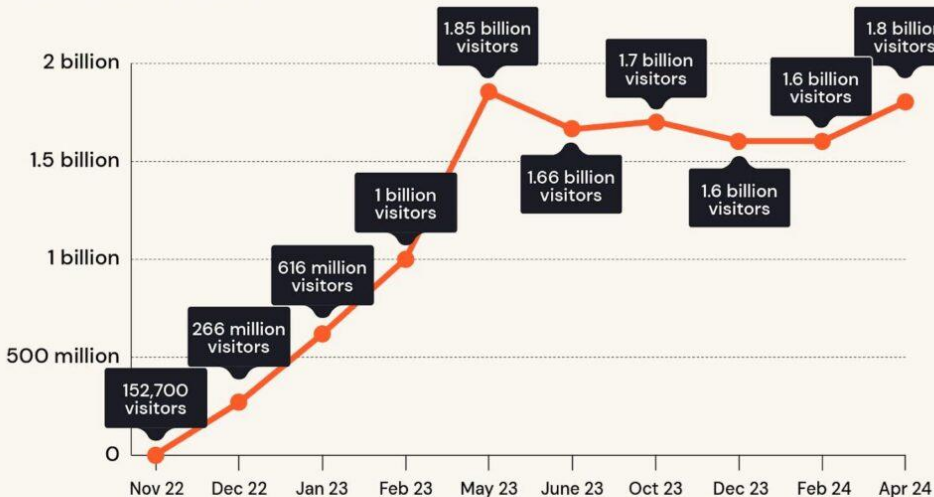


[Deep learning brought substantial improvements in ASR performance](#)

Some Recent Developments in AI

CHATGPT STATISTICS

Change in ChatGPT website visitors since launch



Read the full report at tooltester.com/en/blog/chatgpt-statistics

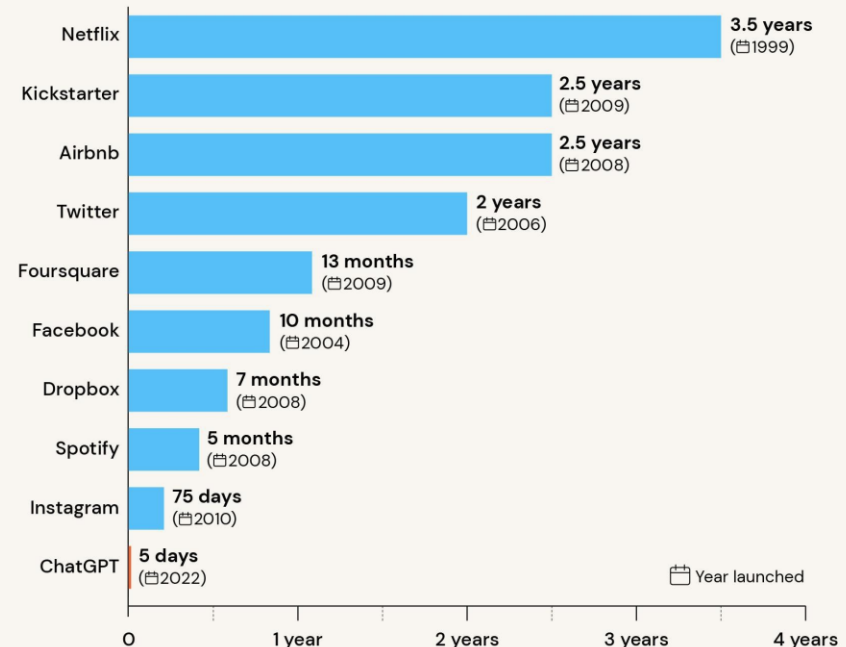
tooltester

ChatGPT received around 1.8 billion visitors in April 2024, with ~100 million active users.

OpenAI's ChatGPT, a 175B parameter neural net, learns to chat on 570GB of text in 34 days, 2022

CHATGPT STATISTICS

Time to reach 1 million users



Read the full report at tooltester.com/en/blog/chatgpt-statistics

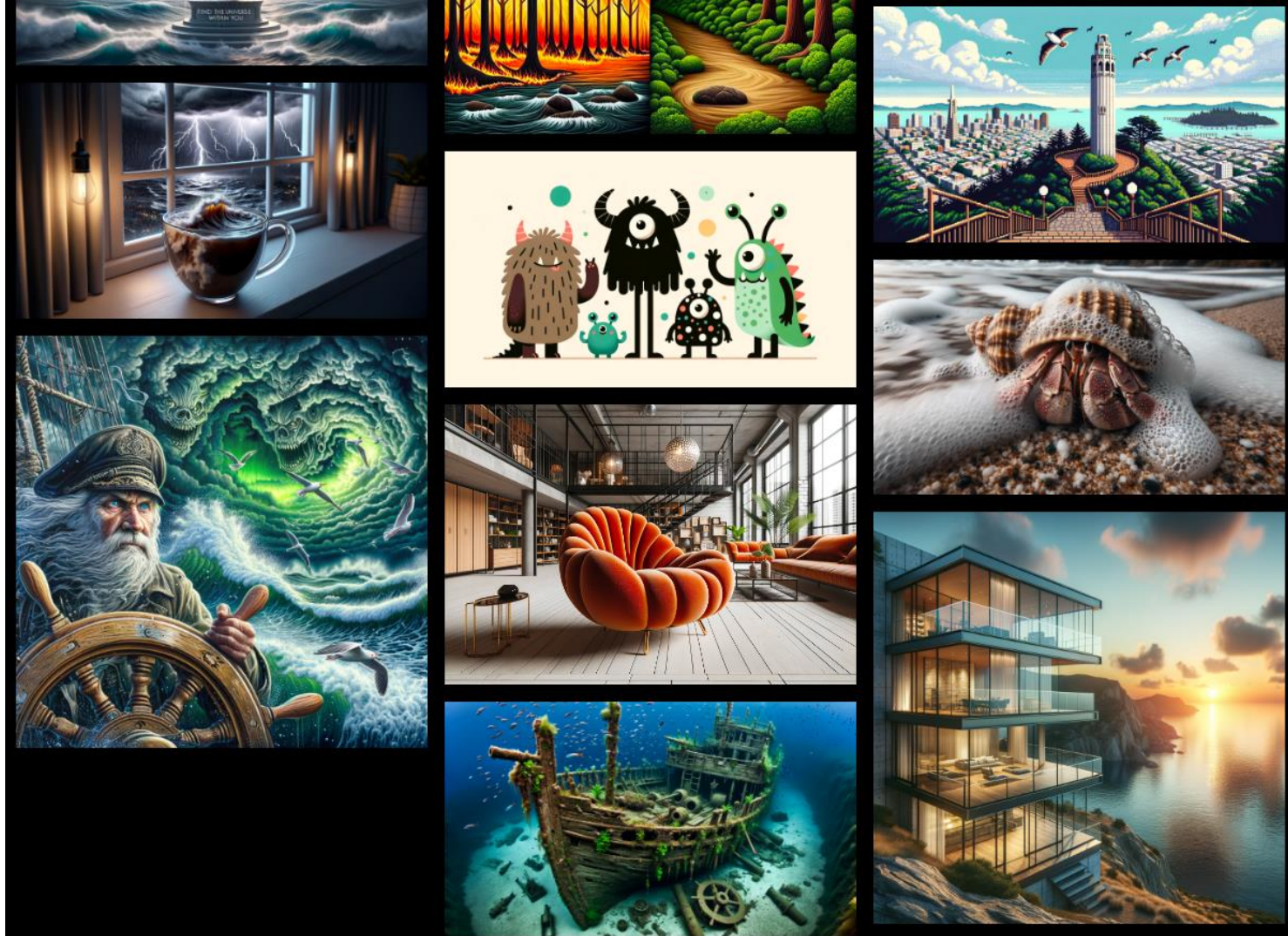
tooltester

Some Recent Developments in AI



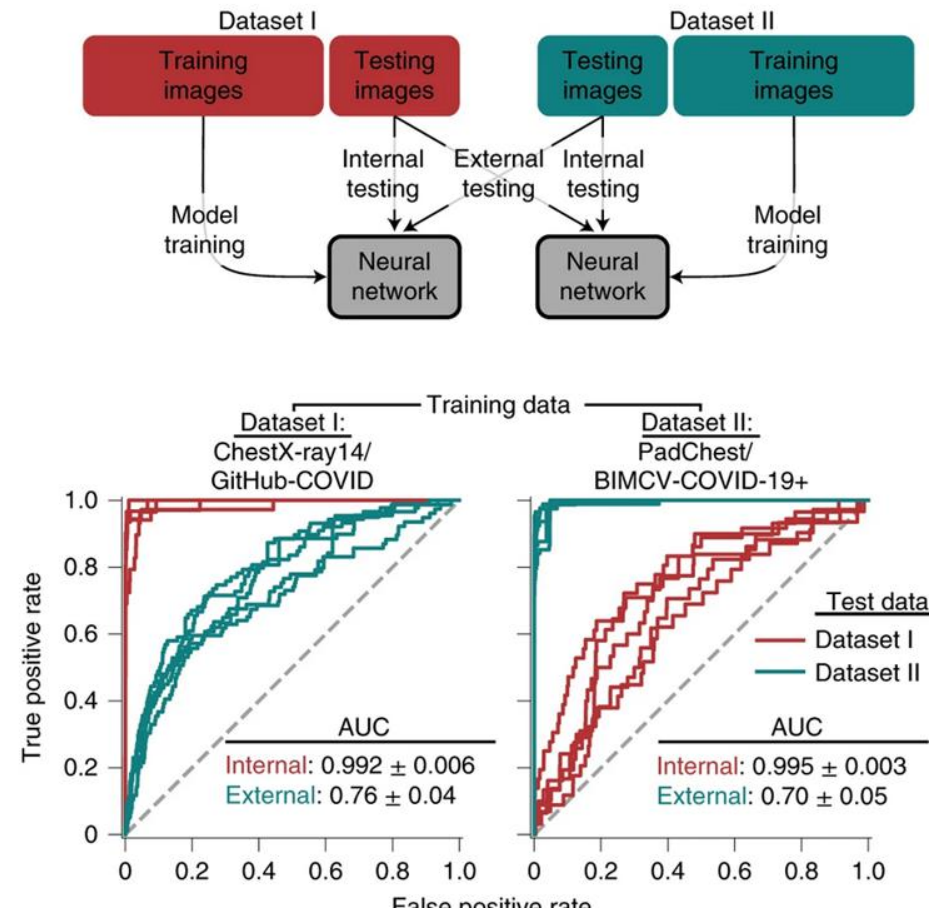
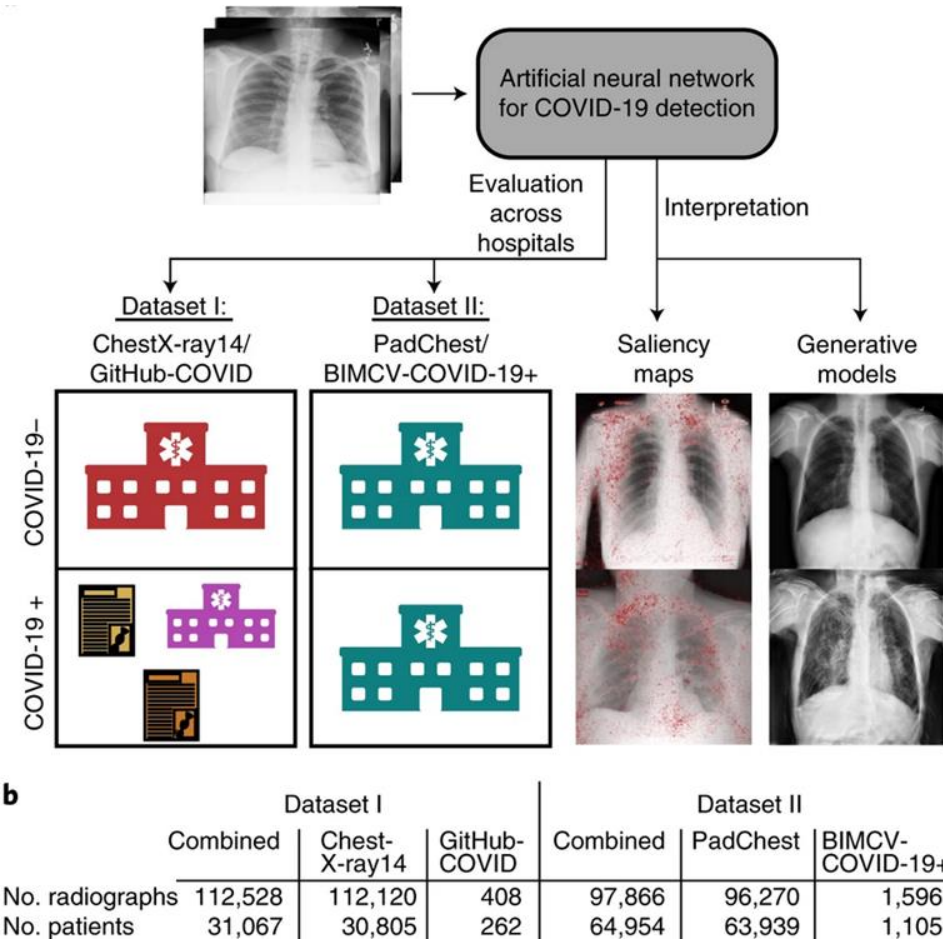
[Aachen+MPI+Tuebingen train neural nets to track multiple objects, 2019](#)

Some Recent Developments in AI



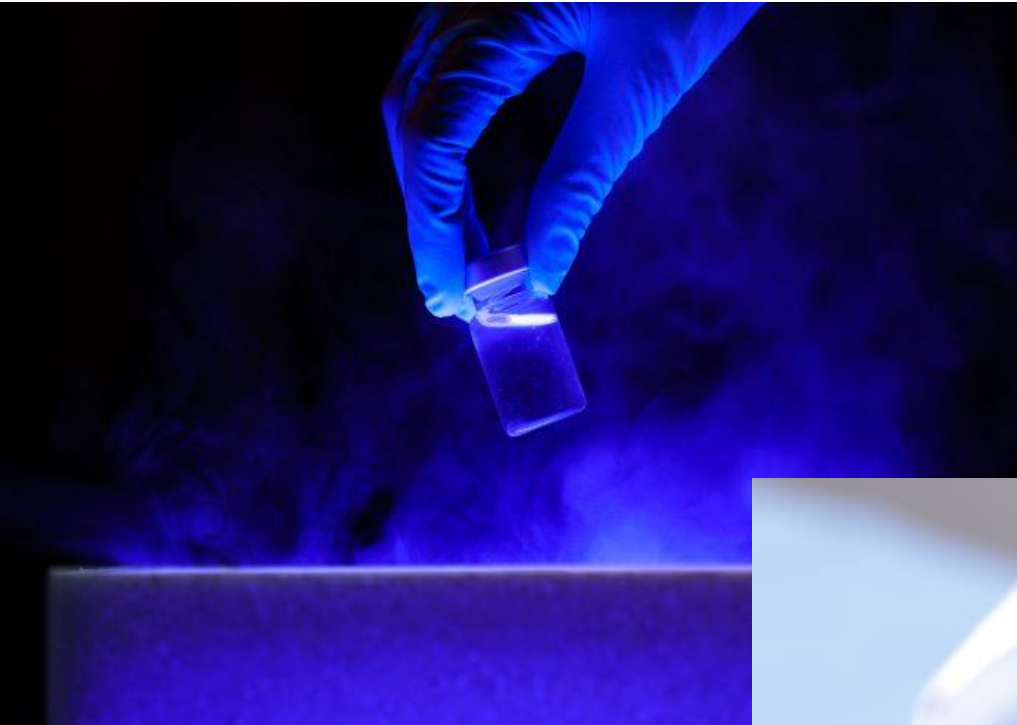
[OpenAI's DALL.E 3 is able to generate images that exactly adheres to text provided, 2023](#)

Some Recent Developments in AI



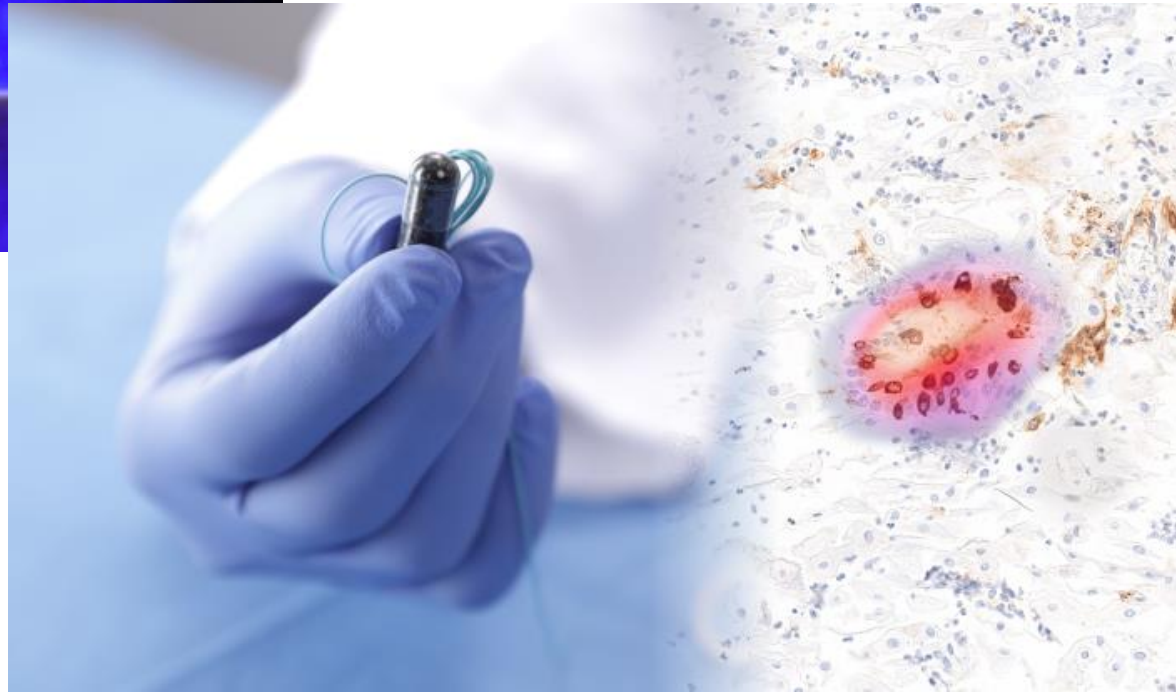
[Medical-imaging AI system learns to detect Covid-19 from chest radiographs, 2021](#)

Some Recent Developments in AI

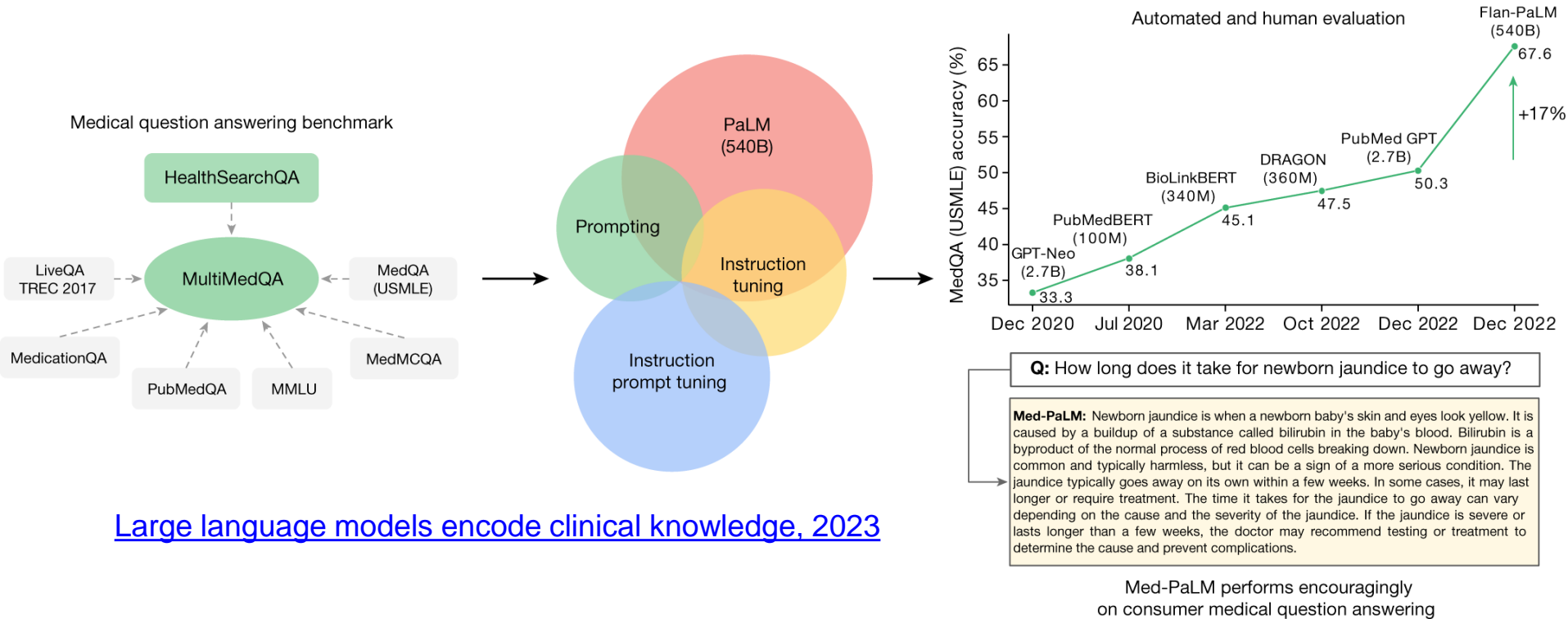


AI system can find COVID vaccine candidates within seconds, 2021

AI for early Cancer detection, 2021

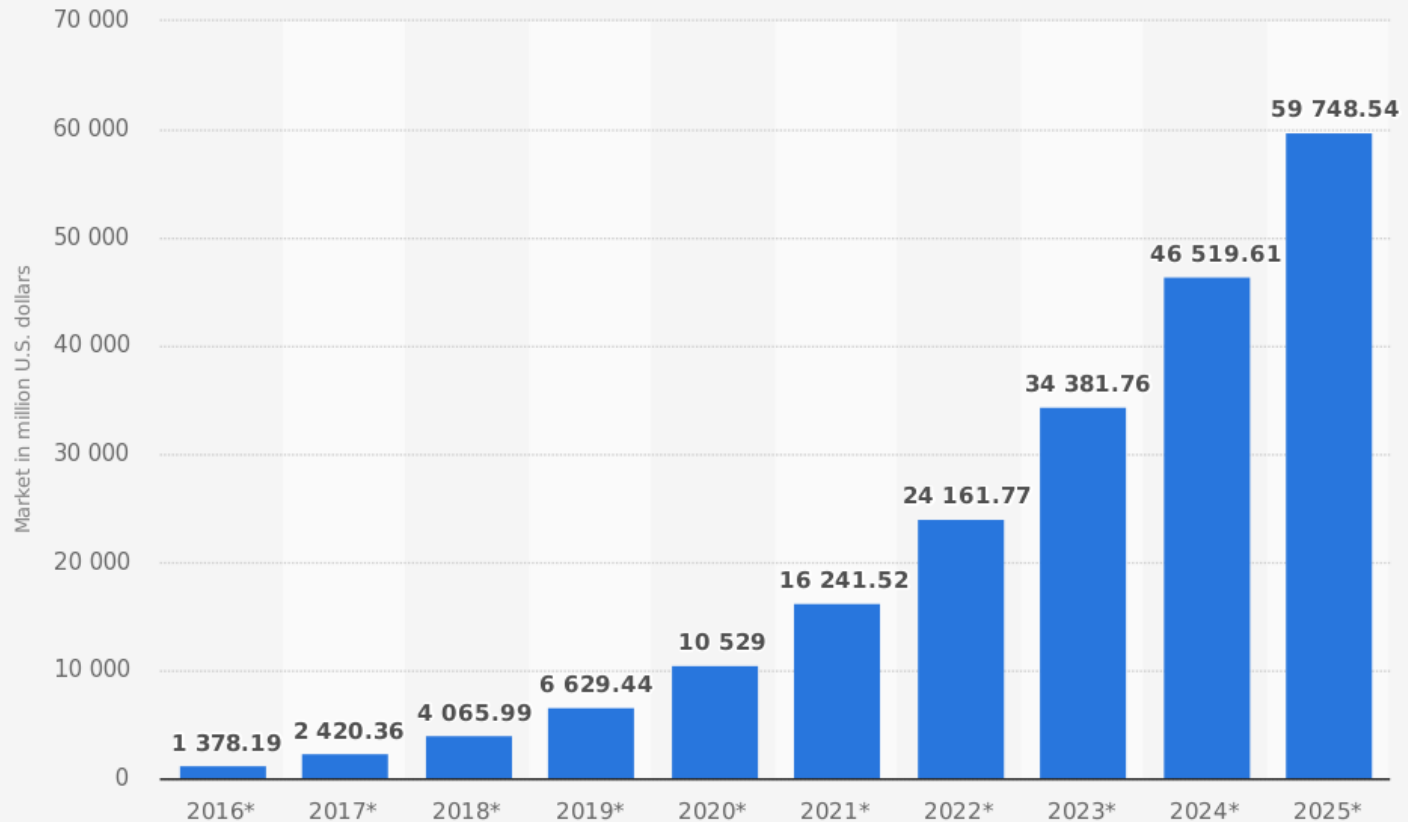


Some Recent Developments in AI



AI: Economic Impact

Revenues from the artificial intelligence (AI) market worldwide, from 2016 to 2025 (in million U.S. dollars)

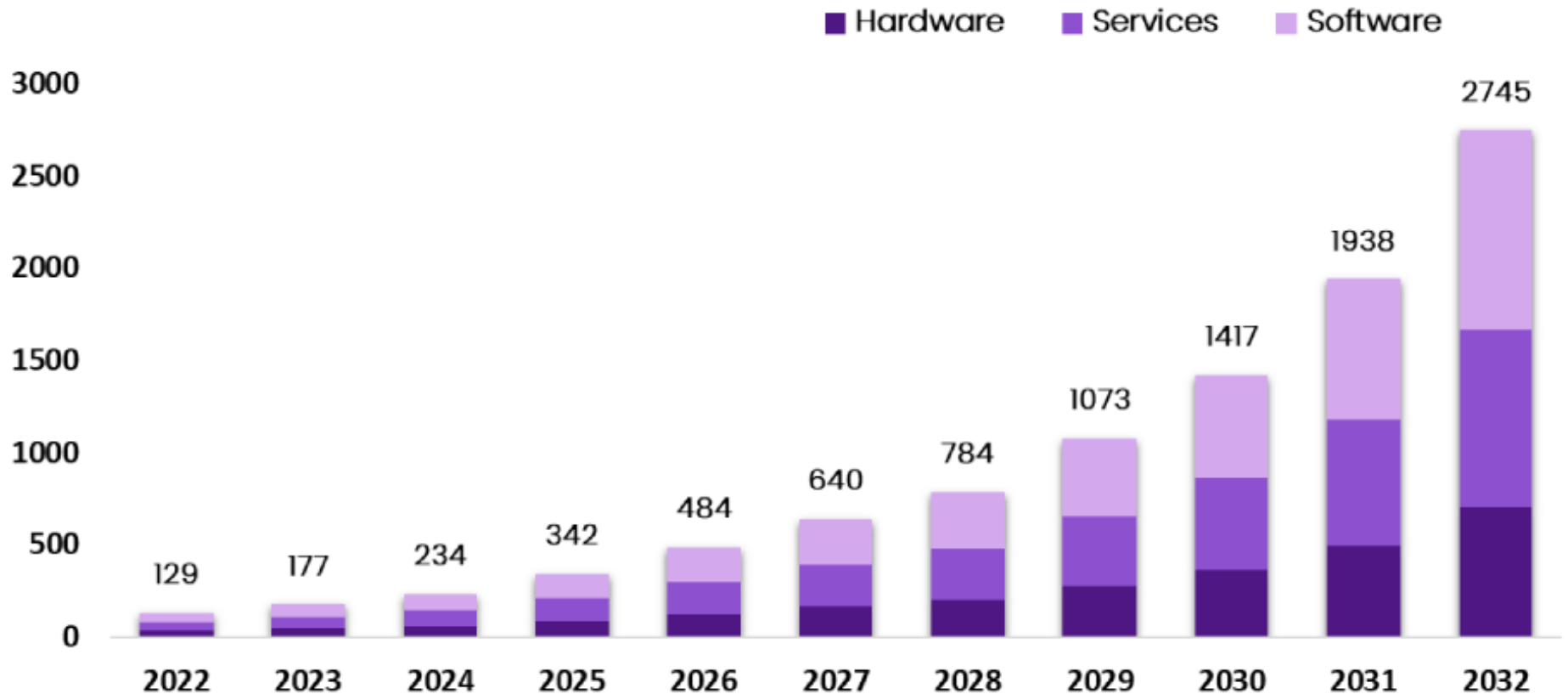


Source
Tractica
© Statista 2017

Additional Information:
Worldwide; 2017

AI: Economic Impact

Global Artificial Intelligence Market



The Market will Grow
At the CAGR of:

36.8%

The forecasted market
size for 2032 in USD:

\$2745B

 **market.us**
ONE STOP SHOP FOR THE REPORTS

[source](#)

Ethics in AI

Fairness

All AI systems that process social or demographic data pertaining to features of human subjects must be designed to meet a minimum threshold of **discriminatory non-harm**. This entails that the datasets they use be equitable; that their model architectures only include reasonable features, processes, and analytical structures; that they do not have inequitable impact; and that they are implemented in an unbiased way.

F

Accountability

Accountability By Design: All AI systems must be designed to facilitate end-to-end answerability and auditability. This requires both responsible humans-in-the-loop across the entire design and implementation chain and activity monitoring protocols that enable end-to-end oversight and review.

A

Sustainability

Designers and users of AI systems must remain aware that these technologies have transformative effects on individuals and society. They must thereby proceed with a continuous sensitivity to real-world impacts. They must also keep in mind that the technical sustainability of these systems depends on their safety: their accuracy, reliability, security, and robustness.

S

Transparency

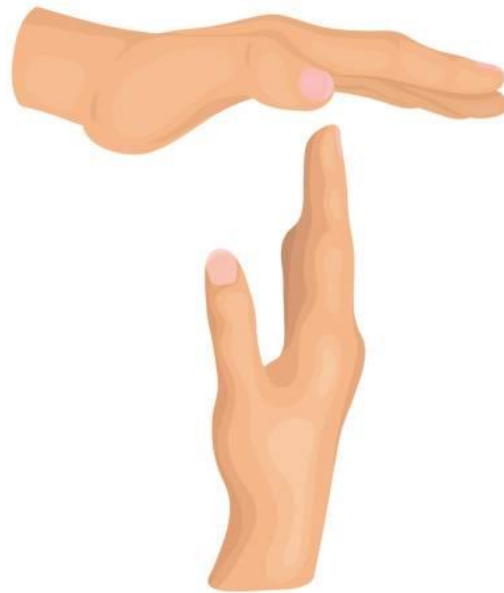
Designers and implementers of AI systems must be able (1) to explain to affected stakeholders in everyday language how and why a model performed the way it did in a specific context and (2) to justify the ethical permissibility, the discriminatory non-harm, and the public trustworthiness both of its outcome and of the processes behind its design and use.

T

[Understanding Artificial Intelligence ethics and safety](#)

[Data ethics and AI guidance landscape](#)

Break



(Discrete) Random Variables

- The events we are interested in have a set of possible values. Examples of random variables:

Coin toss (C): {heads, tails}

Roll a die (D): {1, 2, 3, 4, 5, 6}

Weather (W): {snowy, sunny, rainy, foggy}

Measles (M): {true, false}

- For each event, a random variable takes a value from the associated set. Then we have:

$P(C = \text{tails})$ or $P(\text{tails})$

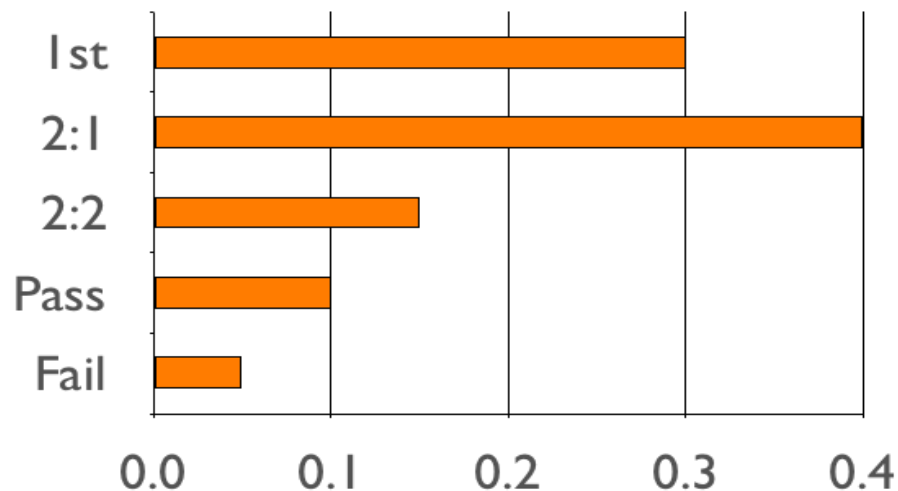
$P(D = 1)$ or $P(1)$

$P(W = \text{sunny})$ or $P(\text{sunny})$

$P(M = \text{true})$ or $P(\text{measles})$

Discrete Probability Distribution

- A probability distribution is a listing of probabilities for *every possible* value of a single random variable



$$P(1st)=0.30$$

$$P(2:1)=0.40$$

$$P(2:2)=0.15$$

$$P(\text{Pass})=0.10$$

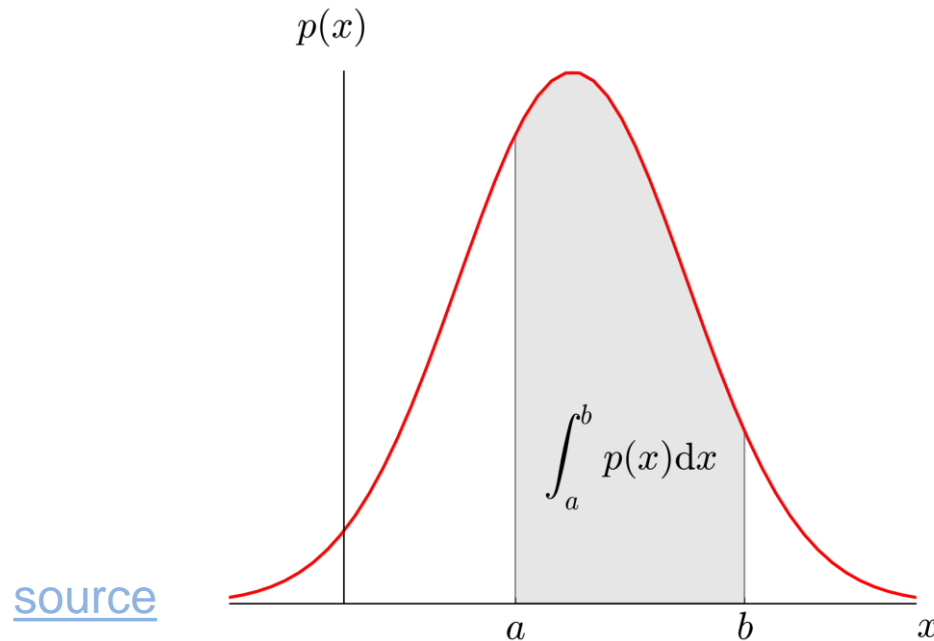
$$P(\text{Fail})=0.05$$

$$\sum_{e \in S} P(e) = 1$$

- Probability distributions can be estimated from data

Continuous Probability Distribution

- The distribution of a *continuous random variable* is represented by a **probability density function (PDF)**



- Due to the infinite number of possible values of a continuous variable (e.g., speed, temperature, time, ...)

Joint Probabilities

Joint probabilities represent the whole joint probability distribution

S1=1st	S2=1st	Probability
TRUE	TRUE	0.2
TRUE	FALSE	0.1
FALSE	TRUE	0.1
FALSE	FALSE	0.6

$$\sum_{e \in S} P(e) = 1$$

Joint Probability Distribution

Sometimes a joint probability distribution looks like the one below, which has the same information as the table on the previous slide.

	S2 =1st	$\neg(\text{S2 =1st})$
S1=1st	0.2	0.1
$\neg(\text{S1=1st})$	0.1	0.6

Marginal Probabilities

	S2 =1st	¬(S2 =1st)
S1=1st	0.2	0.1
¬(S1=1st)	0.1	0.6

Marginalisation => summing up the probabilities of the other variables—i.e. $P(X) = \sum_y P(X, Y = y)$

Example: What is the probability of S1 getting a 1st?

$$\begin{aligned} P(S1=1st) &= P(S1=1st \wedge S2=1st) + P(S1=1st \wedge \neg(S2=1st)) \\ &= 0.2 + 0.1 = 0.3 \end{aligned}$$

Marginal Probabilities

	S2 =1st	¬(S2 =1st)
S1=1st	0.2	0.1
¬(S1=1st)	0.1	0.6

$$\begin{aligned}P(S1=1st) &= P(S1=1st \wedge S2=1st) + P(S1=1st \wedge \neg(S2=1st)) \\&= 0.2 + 0.1 = 0.3\end{aligned}$$

$$\begin{aligned}P(\neg(S1=1st)) &= P(\neg(S1=1st) \wedge S2=1st) + P(\neg(S1=1st) \wedge \neg(S2=1st)) \\&= 0.1 + 0.6 = 0.7\end{aligned}$$

Note that $P(S1=1st) + P(\neg(S1=1st)) = 0.3 + 0.7 = 1$

Marginal Probabilities

	S2 =1st	¬(S2 =1st)
S1=1st	0.2	0.1
¬(S1=1st)	0.1	0.6

$$\begin{aligned} P(S2=1st) &= P(S1=1st \wedge S2=1st) + P(\neg(S1=1st) \wedge S2=1st) \\ &= 0.2 + 0.1 = 0.3 \end{aligned}$$

$$\begin{aligned} P(\neg(S2=1st)) &= P(S1=1st \wedge \neg(S2=1st)) + P(\neg(S1=1st) \wedge \neg(S2=1st)) \\ &= 0.1 + 0.6 = 0.7 \end{aligned}$$

Note that $P(S2=1st) + P(\neg(S2=1st)) = 0.3 + 0.7 = 1$

Conditional Probability

- Conditional probabilities answer the question:
*Given that some event **B happened**, what is the probability of **A happening** too?*

- $P(A|B)=?$

- A conditional probability is defined as:

$$P(A | B) = P(A \wedge B) / P(B), \text{ where } P(B) \neq 0$$

Product Rule

$$P(A \wedge B) = P(A | B) * P(B)$$

Chain Rule for Probabilities

- $P(A \wedge B) = P(A \mid B) * P(B)$
- $$\begin{aligned} P(A \wedge B \wedge C) &= P(A \mid B \wedge C) * P(B \mid C) * P(C) \\ &= P(C) * P(B \mid C) * P(A \mid B \wedge C) \\ &= P(A) * P(B \mid A) * P(C \mid B \wedge A) \\ &= P(C \mid B \wedge A) * P(B \mid A) * P(A) \end{aligned}$$
- $$P(A_1 \wedge A_2 \wedge \dots \wedge A_n) = P(A_n \mid A_{n-1} \wedge \dots \wedge A_1) * P(A_{n-1} \mid A_{n-2} \wedge \dots \wedge A_1) * \dots * P(A_2 \mid A_1) * P(A_1)$$

Conditional Probability

	S2=1st	$\neg(S2=1^{st})$
S1=1st	0.2	0.1
$\neg(S1=1^{st})$	0.1	0.6

- $P(S2=1^{st} \mid S1=1^{st})?$
- $= P(S2=1^{st}, S1=1^{st}) / P(S1=1^{st})$
- $= 0.2 / 0.3 = 0.66666666$

$$P(A \mid B) = P(A \wedge B) / P(B)$$

Substituting A with S2=1st

Substituting B with S1=1st

If the first student has a 1st, the second has a 66% chance of having a 1st too!

Conditional Probability

	S2=1st	$\neg(\text{S2=1st})$
S1=1st	0.2	0.1
$\neg(\text{S1=1st})$	0.1	0.6

- $P(\neg(\text{S2=1st}) \mid \text{S1=1st})?$
- $= P(\neg(\text{S2=1st}), \text{S1=1st}) / P(\text{S1=1st})$
- $= 0.1 / 0.3 = 0.33333333$

$$P(A \mid B) = P(A \wedge B) / P(B)$$

Substituting A with $\neg(\text{S2=1st})$

Substituting B with S1=1st

If the first student (S1) has a 1st, the second has a 33% chance of not getting a 1st.

Conditional Probability

	S2=1st	$\neg(\text{S2=1}^{\text{st}})$
S1=1st	0.2	0.1
$\neg(\text{S1=1}^{\text{st}})$	0.1	0.6

Since $P(\text{S2=1st} \mid \text{S1=1st}) = P(\text{S2=1st}, \text{S1=1st}) / P(\text{S1=1st})$
 $= 0.2 / 0.3 = 0.6666666$

and $P(\neg(\text{S2=1st}) \mid \text{S1=1st}) = P(\neg(\text{S2=1st}), \text{S1=1st}) / P(\text{S1=1st})$
 $= 0.1 / 0.3 = 0.3333333$

Then $P(\text{S2=1st} \mid \text{S1=1st}) + P(\neg(\text{S2=1st}) \mid \text{S1=1st}) = 1$

More on Joint Probabilities

Given a joint probability table, we have all the information we need about the domain. We can calculate the probability of any logical formula.

$$P(S1=first \vee S2=first) = 0.2 + 0.1 + 0.1 = 0.4$$

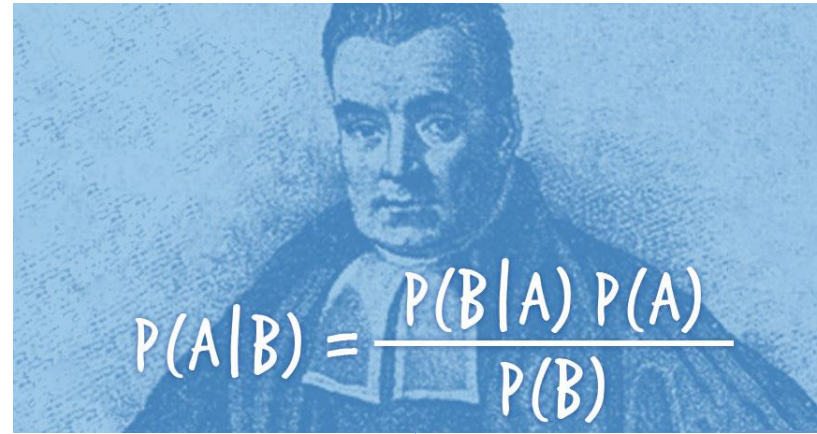
	S2 =1st	$\neg(S2 =1st)$
S1=1st	0.2	0.1
$\neg(S1=1st)$	0.1	0.6

Bayes Rule

- $P(A \wedge B) = P(B \wedge A)$
- $P(A \mid B) * P(B) = P(B \mid A) * P(A)$
- $P(A \mid B) = (P(B \mid A) * P(A)) / P(B)$
- $P(B \mid A) = (P(A \mid B) * P(B)) / P(A)$

Bayes Rule

- $P(A \wedge B) = P(B \wedge A)$
- $P(A | B) * P(B) = P(B | A) * P(A)$
- $P(A | B) = (P(B | A) * P(A)) / P(B)$
- $P(B | A) = (P(A | B) * P(B)) / P(A)$



Thomas Bayes
(1701-1761)

Bayes Rule for Classification

- Given inputs X and outputs Y , the Bayes rule can be written as

$$P(Y = y_k | X = x_i) = \frac{P(X = x_i | Y = y_k)P(Y = y_k)}{\sum_j P(X = x_i | Y = y_j)P(Y = y_j)}$$

where

y_k is the possible value for Y

x_i is the possible vector value for X

- Use training data to estimate $P(X|Y)$ and $P(Y)$.

Difficulty in Unbiased Bayesian Classifiers

- Accurately estimating $P(X|Y)$ requires a set of parameters such as $\theta_{ij} = P(X = x_i | Y = y_j)$, where
index j refers to 2 possible values
index i refers to 2^n possible values
- This requires 2^{n+1} parameters or probabilities.
- A vector X with 30 Boolean inputs requires 2.15B parameters

Naïve Bayes Classifier

- Conditional independence: $P(X_1 \dots X_n | Y) = \prod_{i=1}^n P(X_i | Y)$

- Thus

$$P(Y = y_k | X_1 \dots X_n) = \frac{P(Y = y_k) P(X_1 \dots X_n | Y = y_k)}{\sum_j P(Y = y_j) P(X_1 \dots X_n | Y = y_j)}$$

can be rewritten as

$$P(Y = y_k | X_1 \dots X_n) = \frac{P(Y = y_k) \prod_i P(X_i | Y = y_k)}{\sum_j P(Y = y_j) \prod_i P(X_i | Y = y_j)}$$

Naïve Bayes Classifier

- Given a new instance/example $X^{new} = \langle X_1 \dots X_n \rangle$, the most probable value of Y can be obtained with

$$Y = \arg \max_{y_k} \frac{P(Y = y_k) \prod_i P(X_i | Y = y_k)}{\sum_j P(Y = y_j) \prod_i P(X_i | Y = y_j)}$$

- Since the denominator does not depend on y_k , the equation can be simplified as

$$Y = \arg \max_{y_k} P(Y = y_k) \prod_i P(X_i | Y = y_k)$$

Naïve Bayes for Discrete Inputs

- Estimate two sets of parameters:

$$\theta_{ijk} = P(X = x_{ij} | Y = y_k)$$
$$\pi_k = P(Y = y_k)$$

- According to

$$\theta_{ijk} = P(X = x_{ij} | Y = y_k) = \frac{\#D\{X_i = x_{ij} \wedge Y = y_k\} + l}{\#D\{Y = y_k\} + lJ}$$
$$\pi_k = P(Y = y_k) = \frac{\#D\{Y = y_k\}}{|D|}$$

where J is the number of unique values in X_i , l (e.g., $l = 1$) avoids zero estimates, and $|D|$ is the num. of elements in the training set.

Example: Training Data

Day	<i>Outlook</i>	<i>Temperature</i>	<i>Humidity</i>	<i>Wind</i>	<i>PlayTennis</i>
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Data from: [Mitchell, T. "Machine Learning", McGraw Hill, 1997.](#)

Example: Estimated Parameters

$$P(\text{PlayTennis} = \text{yes}) = \frac{9}{14} = 0.643$$

$$P(\text{PlayTennis} = \text{no}) = \frac{5}{14} = 0.357$$

$$P(\text{Outlook} = \text{sunny} | \text{PlayTennis} = \text{yes}) = 2/9 = 0.222$$

$$P(\text{Outlook} = \text{sunny} | \text{PlayTennis} = \text{no}) = 3/5 = 0.60$$

$$P(\text{Outlook} = \text{overcast} | \text{PlayTennis} = \text{yes}) = 4/9 = 0.444$$

$$P(\text{Outlook} = \text{overcast} | \text{PlayTennis} = \text{no}) = 0/5 = 0.0$$

$$P(\text{Outlook} = \text{rain} | \text{PlayTennis} = \text{yes}) = 3/9 = 0.333$$

$$P(\text{Outlook} = \text{rain} | \text{PlayTennis} = \text{no}) = 2/5 = 0.4$$

...

We want to avoid
zero probabilities

Example: Estimated Parameters

$$P(\text{Temperature} = \text{hot} | \text{PlayTennis} = \text{yes}) = 2/9 = 0.222$$

$$P(\text{Temperature} = \text{hot} | \text{PlayTennis} = \text{no}) = 2/5 = 0.4$$

$$P(\text{Temperature} = \text{mild} | \text{PlayTennis} = \text{yes}) = 4/9 = 0.444$$

$$P(\text{Temperature} = \text{mild} | \text{PlayTennis} = \text{no}) = 2/5 = 0.4$$

$$P(\text{Temperature} = \text{cool} | \text{PlayTennis} = \text{yes}) = 3/9 = 0.333$$

$$P(\text{Temperature} = \text{cool} | \text{PlayTennis} = \text{no}) = 1/5 = 0.2$$

$$P(\text{Humidity} = \text{high} | \text{PlayTennis} = \text{yes}) = 3/9 = 0.333$$

$$P(\text{Humidity} = \text{high} | \text{PlayTennis} = \text{no}) = 4/5 = 0.8$$

$$P(\text{Humidity} = \text{normal} | \text{PlayTennis} = \text{yes}) = 6/9 = 0.666$$

$$P(\text{Humidity} = \text{normal} | \text{PlayTennis} = \text{no}) = 1/5 = 0.2$$

$$P(\text{Wind} = \text{strong} | \text{PlayTennis} = \text{yes}) = 3/9 = 0.333$$

$$P(\text{Wind} = \text{strong} | \text{PlayTennis} = \text{no}) = 3/5 = 0.6$$

$$P(\text{Wind} = \text{weak} | \text{PlayTennis} = \text{yes}) = 6/9 = 0.666$$

$$P(\text{Wind} = \text{weak} | \text{PlayTennis} = \text{no}) = 2/5 = 0.4$$

Example: Classifying a New Instance

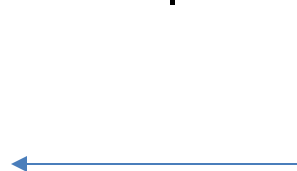
$$P(\text{yes})P(\text{sunny}|\text{yes})P(\text{cool}|\text{yes})P(\text{high}|\text{yes})P(\text{strong}|\text{yes}) = \\ = 0.643 * 0.222 * 0.333 * 0.333 * 0.333 = 0.0053$$

$$P(\text{no})P(\text{sunny}|\text{no})P(\text{cool}|\text{no})P(\text{high}|\text{no})P(\text{strong}|\text{no}) = \\ = 0.357 * 0.60 * 0.2 * 0.8 * 0.6 = 0.0206$$



$$P(\text{PlayTennis} = \text{yes}|\text{evidence}) = \frac{0.0053}{0.0053 + 0.0206} = 0.205$$
$$P(\text{PlayTennis} = \text{no}|\text{evidence}) = \frac{0.0206}{0.0053 + 0.0206} = 0.795$$

Outlook=sunny
Temperature=cool
Humidity=high
Wind=strong



Same example with Log Probabilities

$$\begin{aligned} &P(\text{yes})P(\text{sunny}|\text{yes})P(\text{cool}|\text{yes})P(\text{high}|\text{yes})P(\text{strong}|\text{yes}) \\ &= (-0.442) + (-1.504) + (-1.098) + (-1.098) + (-1.098) \\ &= -5,242 \end{aligned}$$

$$\begin{aligned} &P(\text{no})P(\text{sunny}|\text{no})P(\text{cool}|\text{no})P(\text{high}|\text{no})P(\text{strong}|\text{no}) \\ &= (-1.029) + (-0.51) + (-1.609) + (-0.223) + (-0.51) \\ &= -3.883 \end{aligned}$$



$$P(\text{PlayTennis} = \text{yes}|\text{evidence}) = \frac{e^{-5,242}}{e^{-5,242} + e^{-3.883}} = 0.205$$

$$P(\text{PlayTennis} = \text{no}|\text{evidence}) = \frac{e^{-3.883}}{e^{-5,242} + e^{-3.883}} = 0.795$$

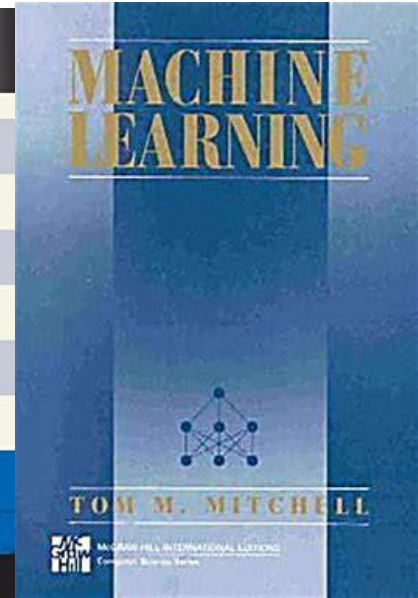
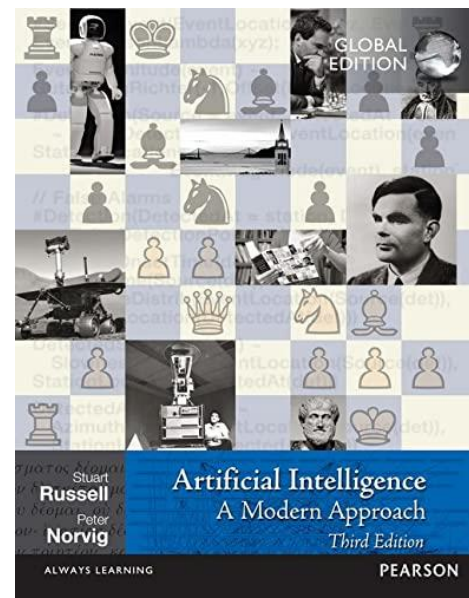
Today

- Elaborated on the term '**Artificial Intelligence**'
- Some recent major developments in AI
- Introduction to probability theory
- Probabilistic reasoning with Naïve Bayes

Readings:

Russell & Norvig 2016. [Chapters 1, 2, 13](#)

Mitchell, T. 2017; 2nd Ed. [Chapter 3](#)



This and Next Week

Workshop (today):

Exercises on probability theory
Python programs for Bayesian classifiers

Lecture (next week):

Bayesian Networks with Exact Inference
Readings: Russell & Norvig 2016. Chapters 13-14.4

Questions?