DATA SCIENCE IN MANUFACTURING

WEEK 7

ANDREW SHERLOCK, JONATHAN CORNEY, DANAI KORRE

BY THE END OF THIS LECTURE YOU SHOULD:

Be introduced to Machine Learning (ML) and Artificial Intelligence (AI)

Get familiar with the uses of ML and AI in manufacturing

Understand the basic principles behind ML and AI

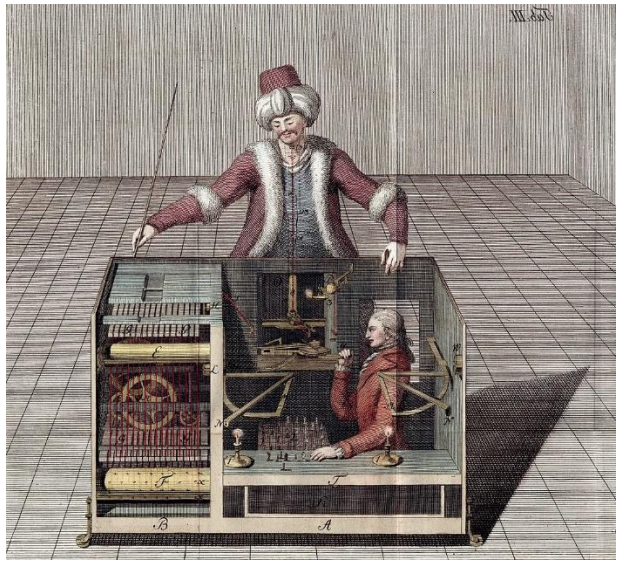
LECTURE: WEEK 7

Introduction to
Artificial Intelligence and Machine Learning
and supervised learning

John McCarthy in his paper “What is artificial intelligence” [6] defines AI as “ the science and engineering of **making intelligent machines**, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.”

ARTIFICIAL INTELLIGENCE

Making machine “appear” intelligent



AI TIMELINE

Source: The University of Queensland

1642 First mechanical calculating machine built by French mathematician and inventor Blaise Pascal.

1837 First design for a programmable machine, by Charles Babbage and Ada Lovelace.

1943 Foundations of neural networks established by Warren McCulloch and Walter Pitts, drawing parallels between the brain and computing machines.

1950 Alan Turing introduces a test—the Turing test—as a way of testing a machine's intelligence.

1955 'Artificial intelligence' is coined during a conference devoted to the topic.

1965 ELIZA, a natural language program, is created. ELIZA handles dialogue on any topic; similar in concept to today's chatbots.

1980s Edward Feigenbaum creates expert systems which emulate decisions of human experts.

1997 Computer program Deep Blue beats world chess champion Garry Kasparov.

2002 iRobot launches Roomba, an autonomous vacuum cleaner that avoids obstacles.

2009 Google builds the first self-driving car to handle urban conditions.

2011 IBM's Watson defeats champions of US game show Jeopardy!

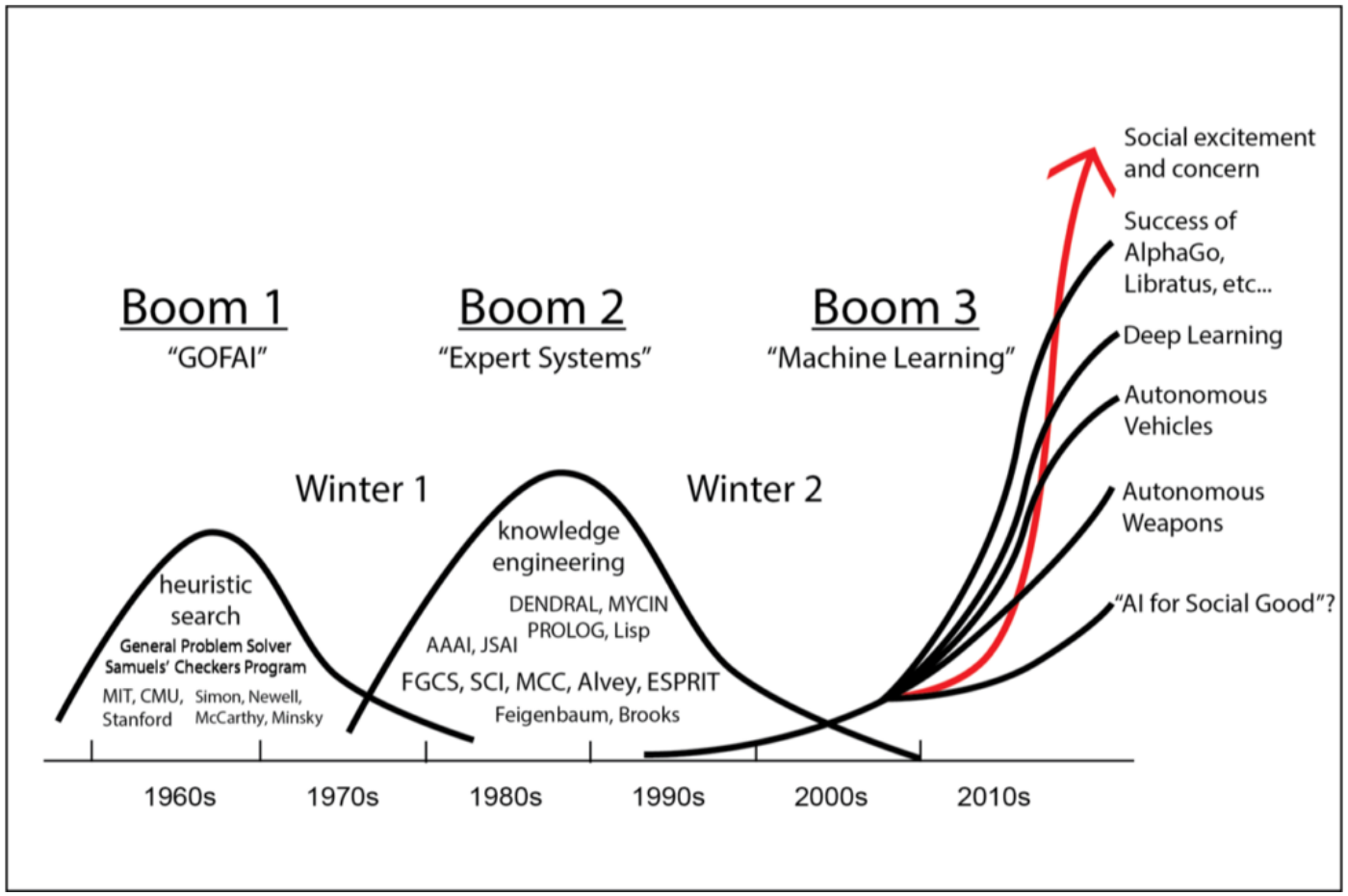
2011-2014 Personal assistants like Siri, Google Now, Cortana use speech recognition to answer questions and perform simple tasks.

2014 Ian Goodfellow comes up with Generative Adversarial Networks (GAN).

2016 AlphaGo beats professional Go player Lee Sedol 4-1.

2018 Most universities have courses in Artificial Intelligence.

A History Of Boom And Bust



SINGULARITY

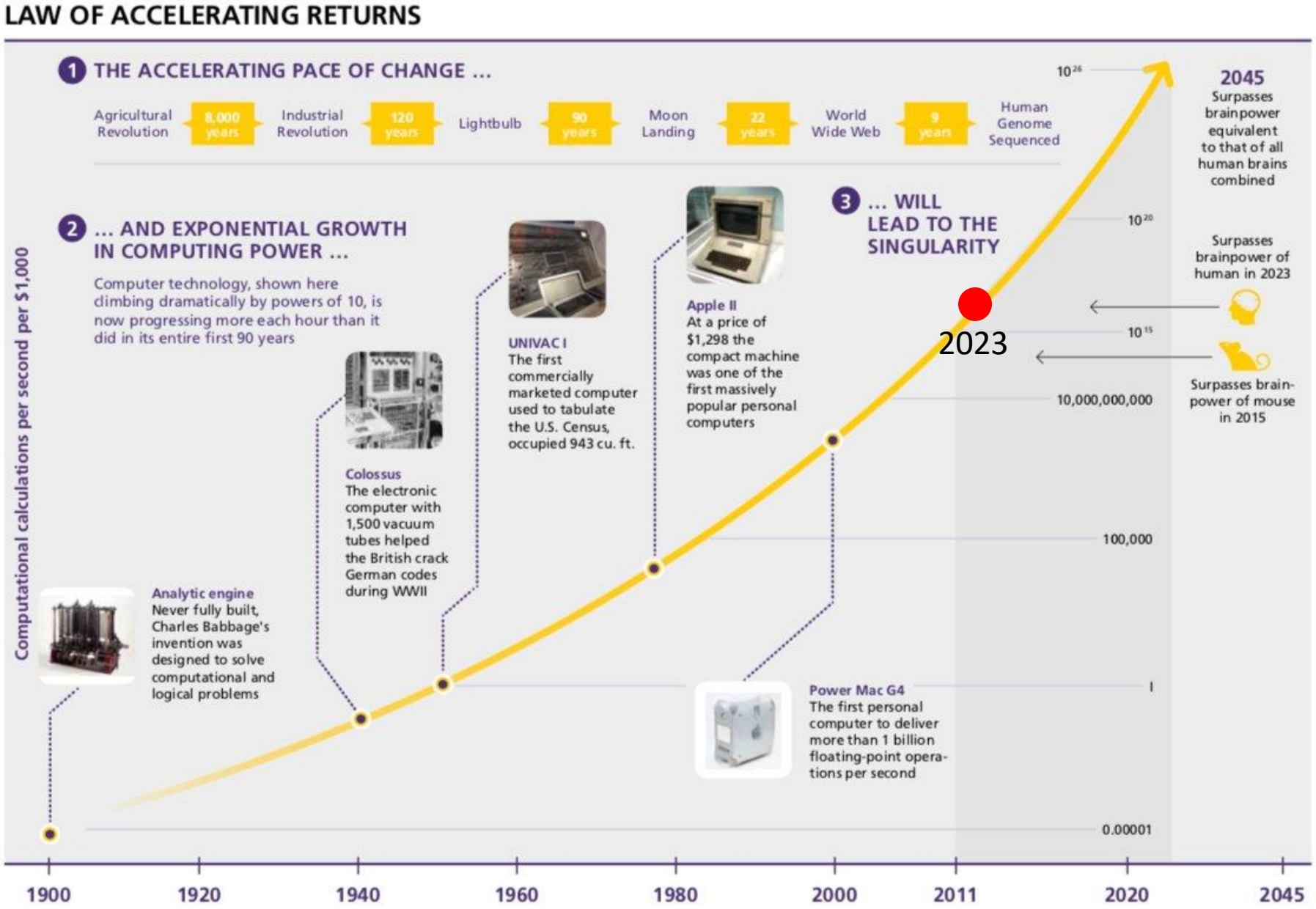


Figure 12: Ray Kurzweil's Law of Accelerating Returns depicts the exponential growth of computer processing power and technology innovations throughout history, and anticipates computers will exceed human intelligence in the future; **Source:** TIME / Wikipedia

source dhl via @mikequindazzi

SINGULARITY

“With artificial intelligence we’re summoning the demon”
– Elon Musk

“Full artificial intelligence could spell the end of the human race”
– Steven Hawking

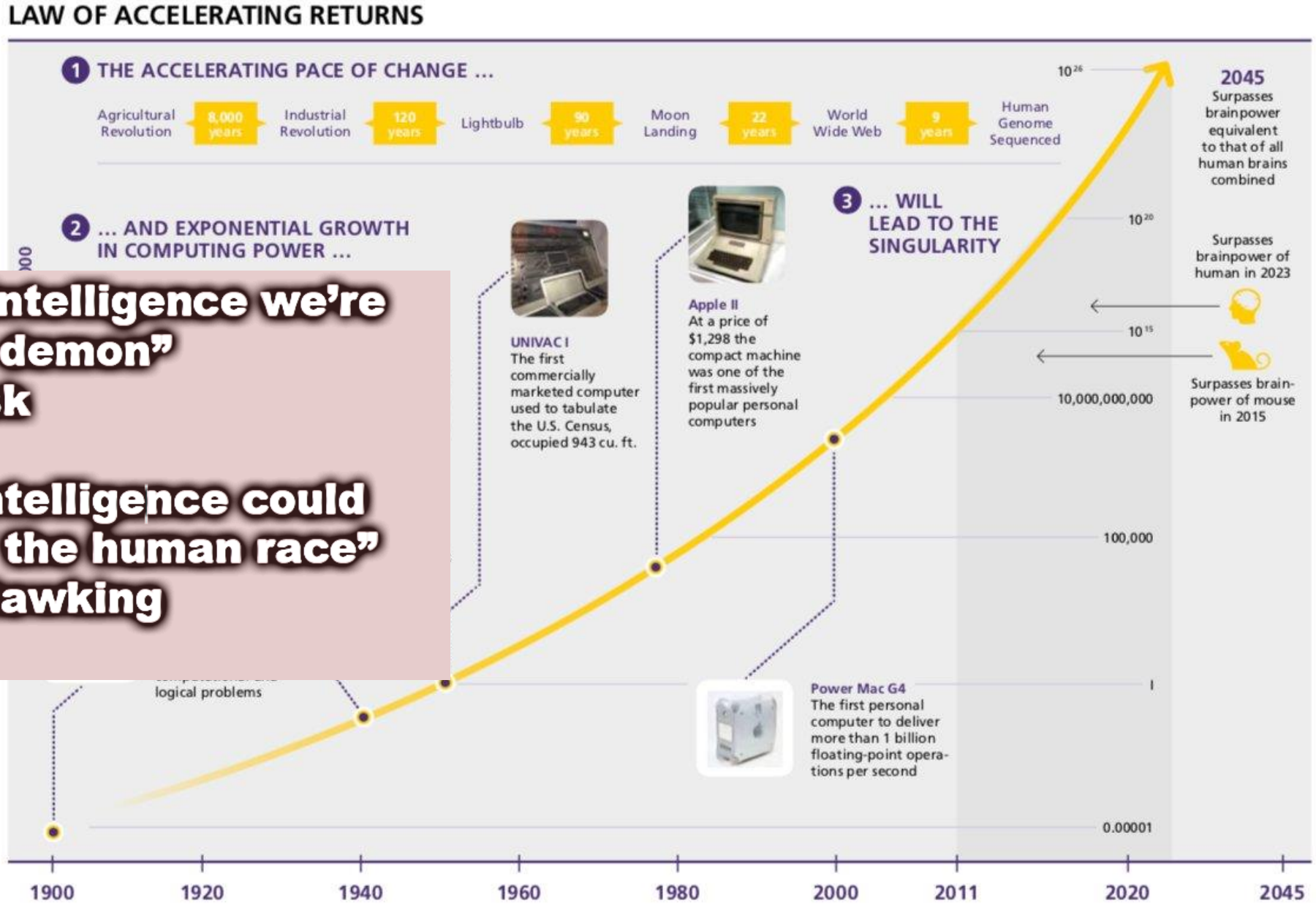
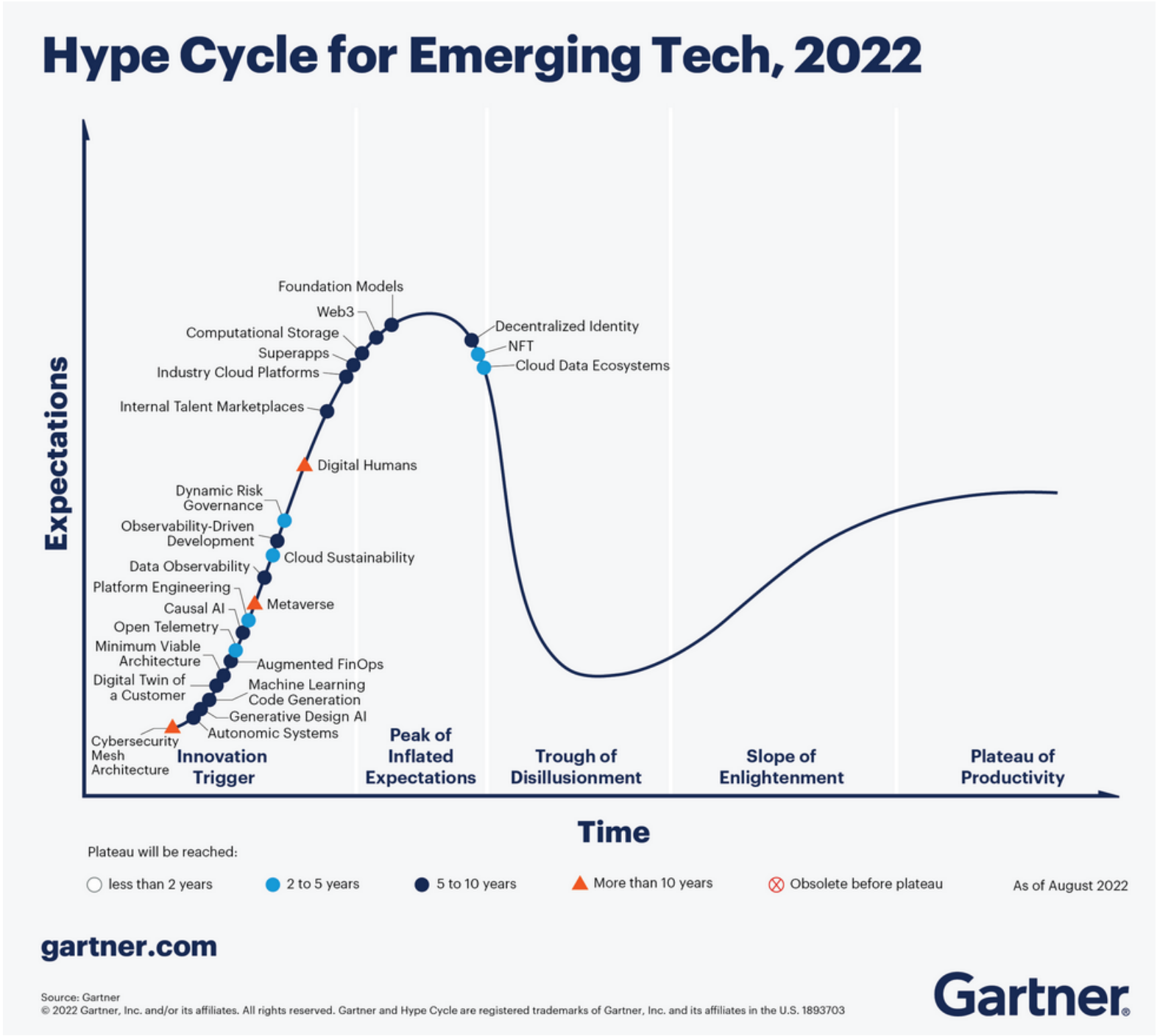


Figure 12: Ray Kurzweil's Law of Accelerating Returns depicts the exponential growth of computer processing power and technology innovations throughout history, and anticipates computers will exceed human intelligence in the future; **Source:** TIME / Wikipedia

source dhl via @mikequindazzi

The 2022 Gartner Hype Cycle identifies 25 must-know emerging technologies designed to help enterprise architecture and technology innovation leaders

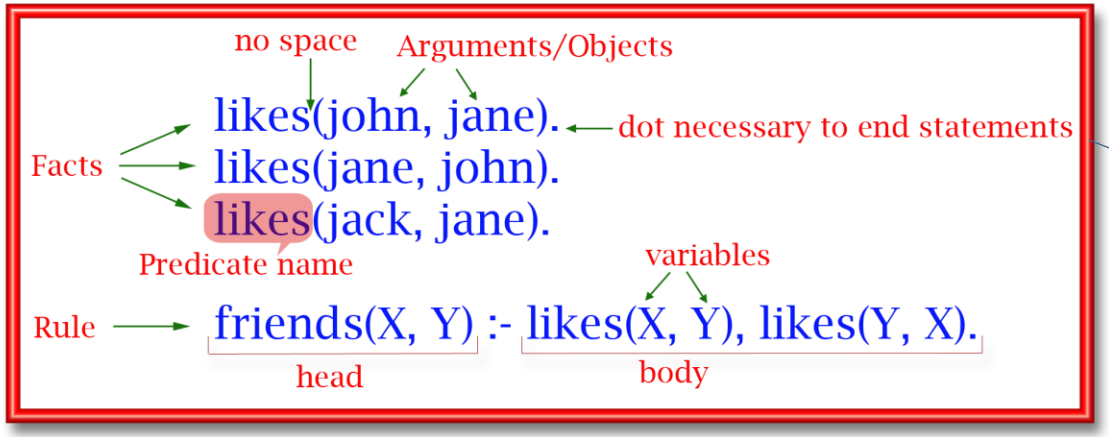


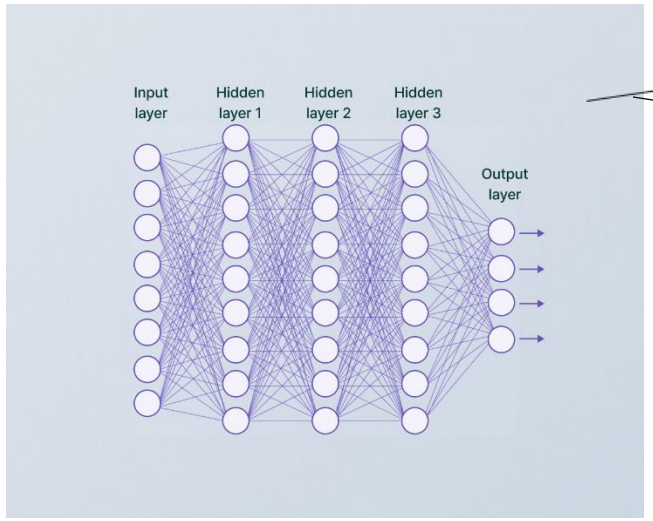
THE UNIVERSITY of EDINBURGH
School of Engineering

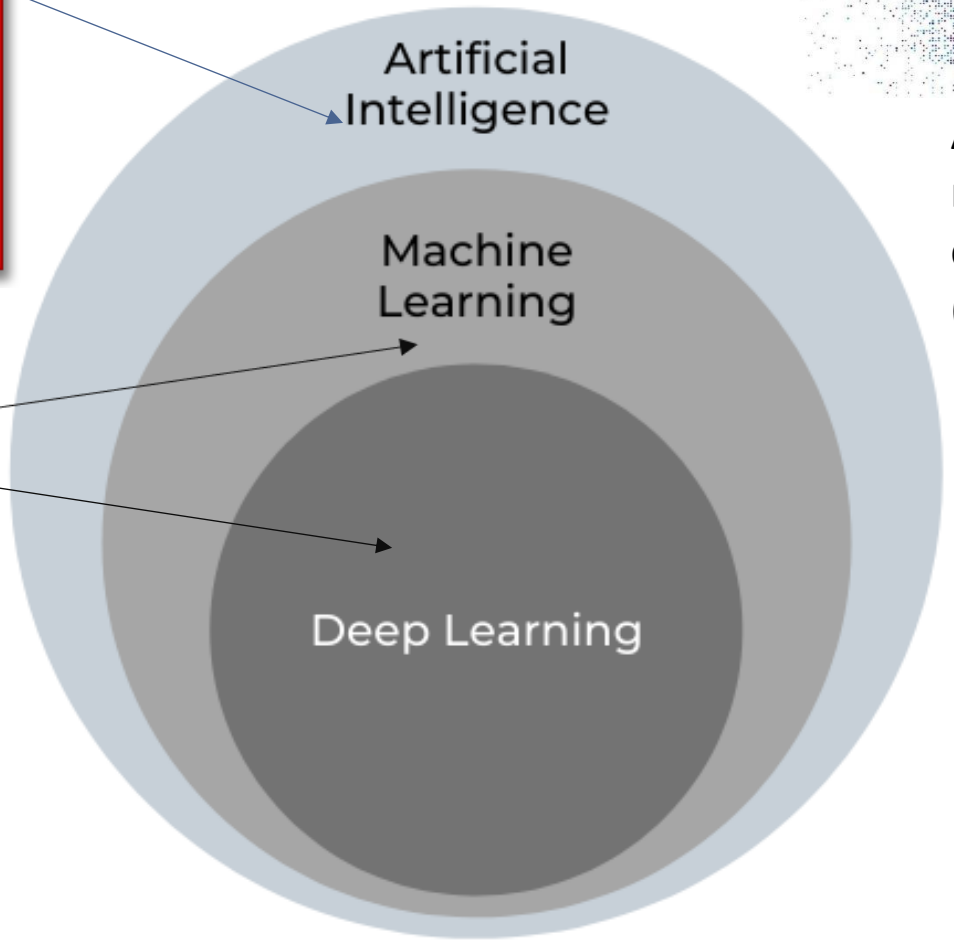
DEEP LEARNING VS. MACHINE LEARNING

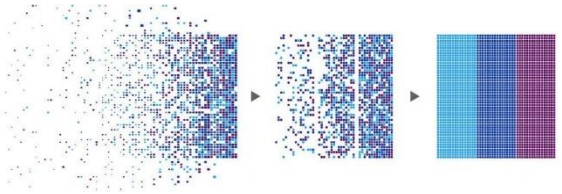
A broad range of technologies (e.g.) Prolog

Program Window

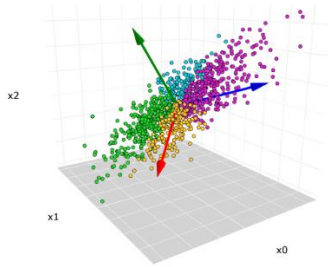








A range of pattern recognition and clustering technologies (lots of maths)!

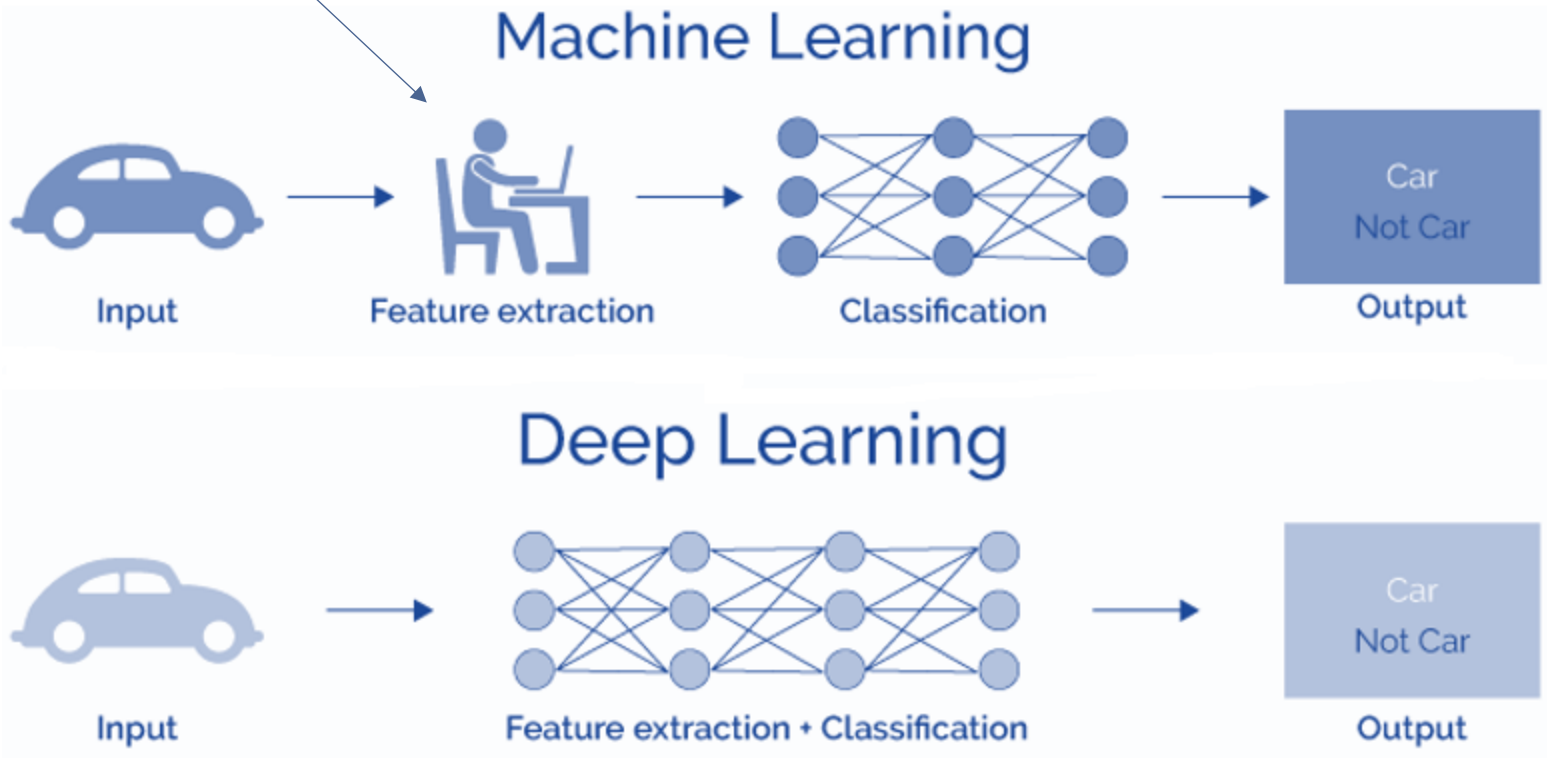


A.I.

In layman's terms, AI is the endeavour of automating tasks of intellectual nature usually performed by humans. A program doesn’t have to be complex to be considered AI. Even a simple program like rock, paper, scissors can be considered AI. Both deep learning and machine learning are sub-fields of artificial intelligence, and deep learning is actually a sub-field of machine learning.

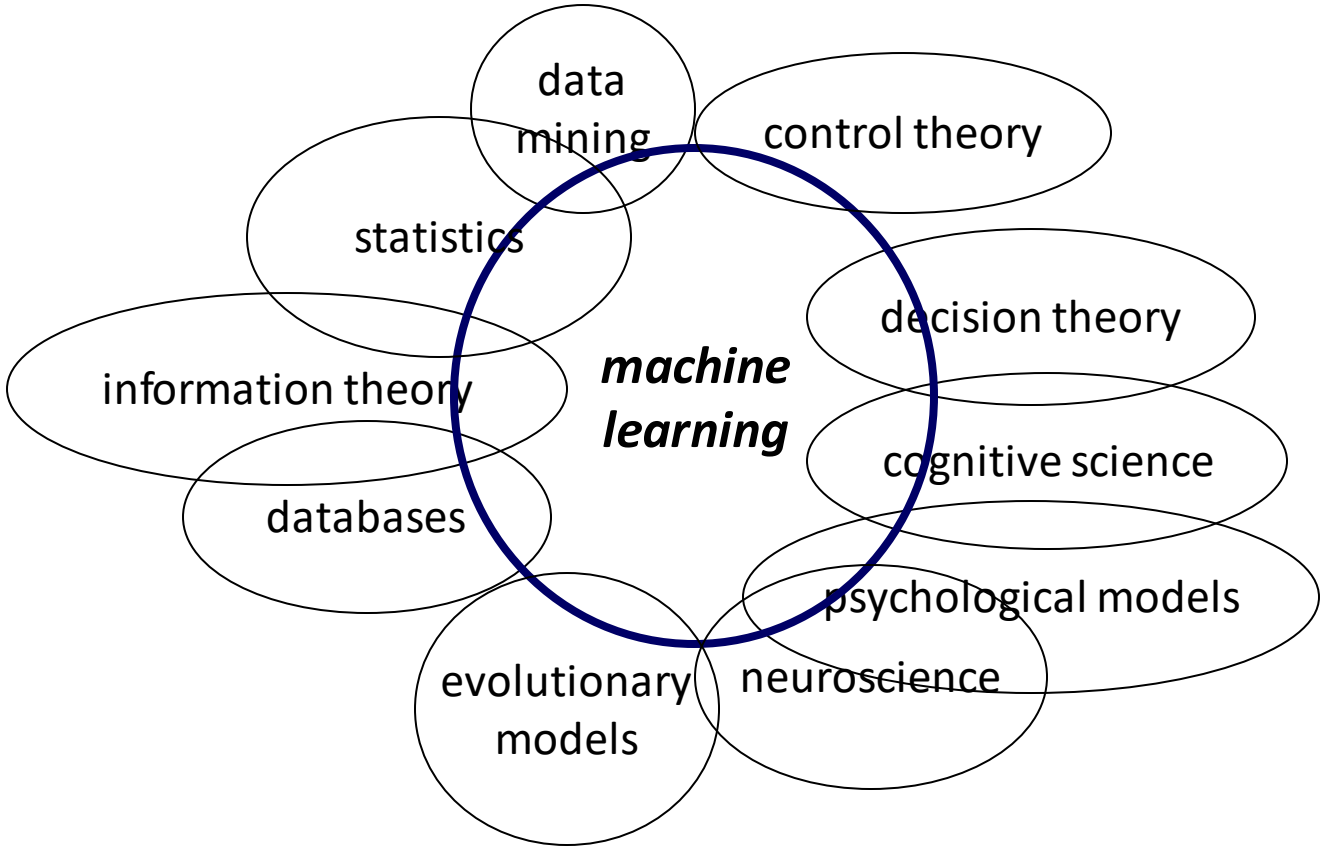
DEEP LEARNING

In deep learning or neural networks, we have many layers of representation of data.



Machine Learning (ML)

OVERVIEW: RELATED FIELDS



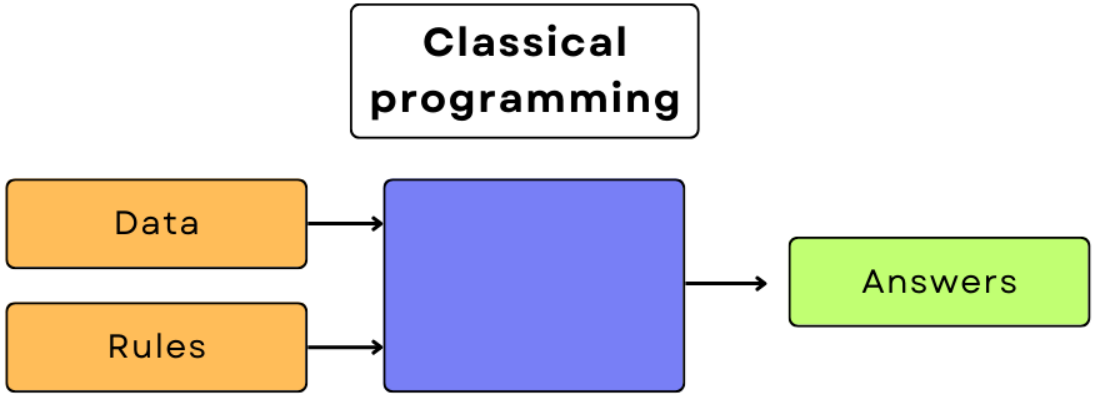
Source: SRI International

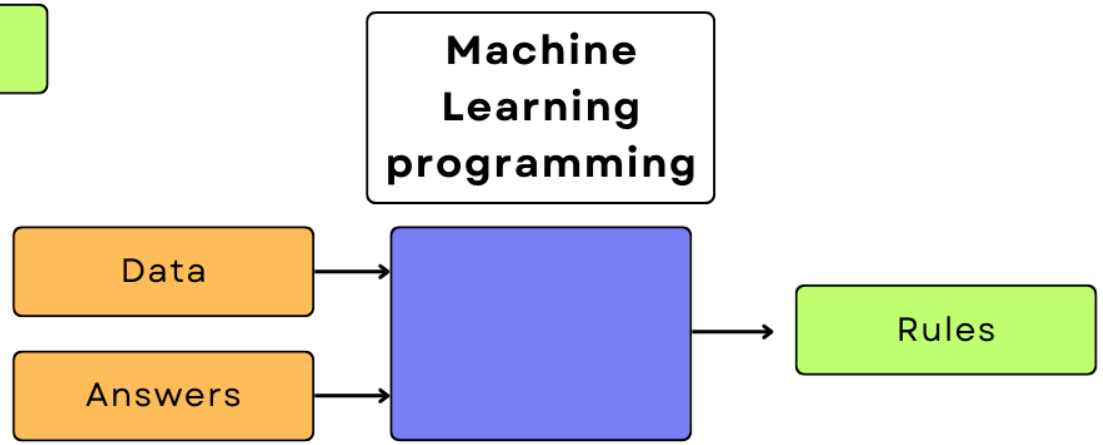
MACHINE LEARNING (ML)

- What is machine learning?
- What is supervised, unsupervised and reinforcement learning?

Regardless of the Hype it's a new way of programming!

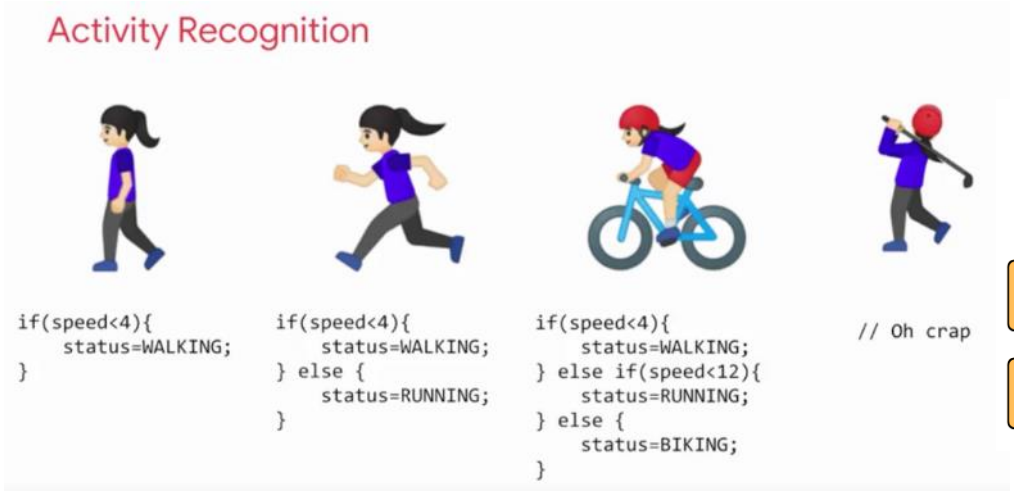
Conventional Programming is writing a program in a traditional procedural language, such as assembly language or a high-level compiler language (C, C++, Java, JavaScript, Python, etc).



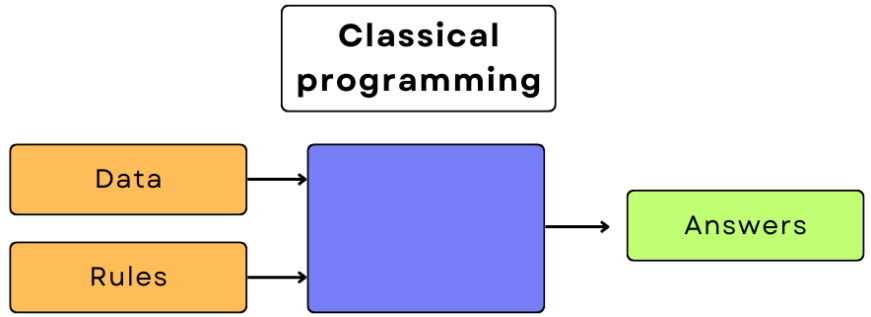


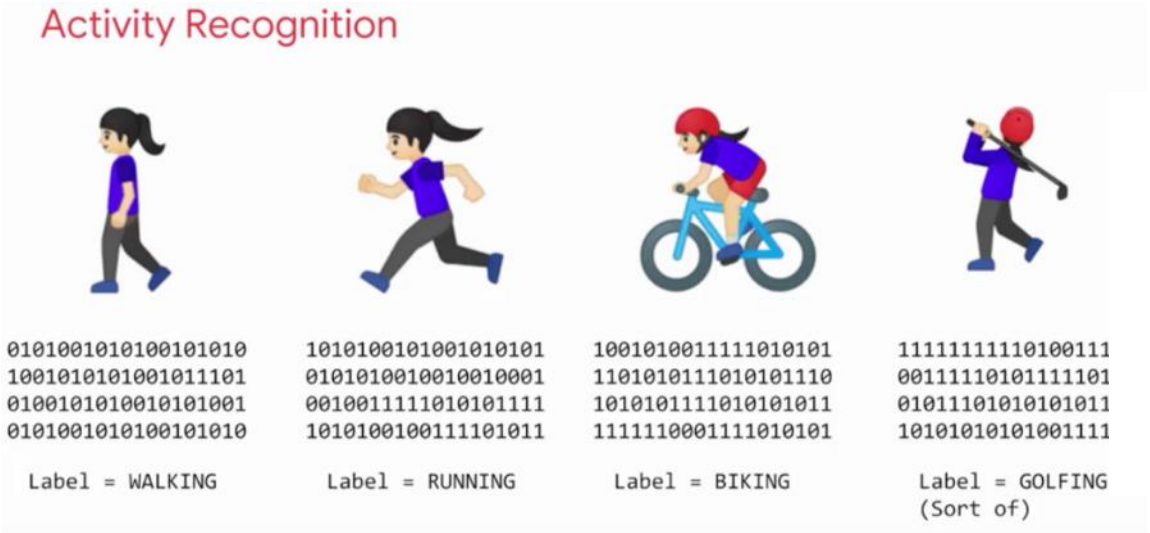
Machine Learning solves this problem by modeling this data with **train data** and **test data** and then ***predict*** the result.

TEACHABLE MACHINES

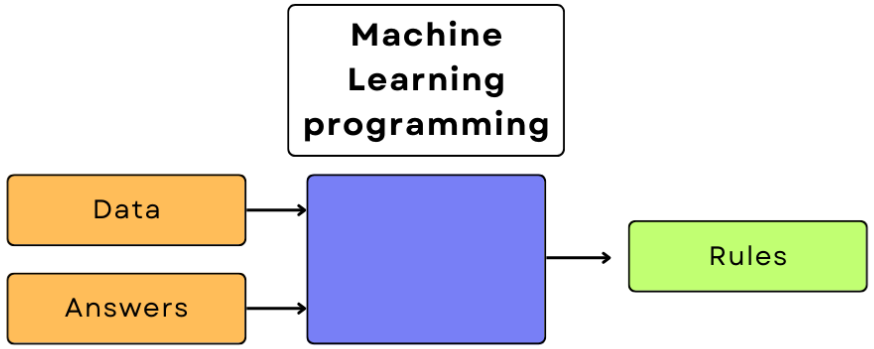


Conventional Programming

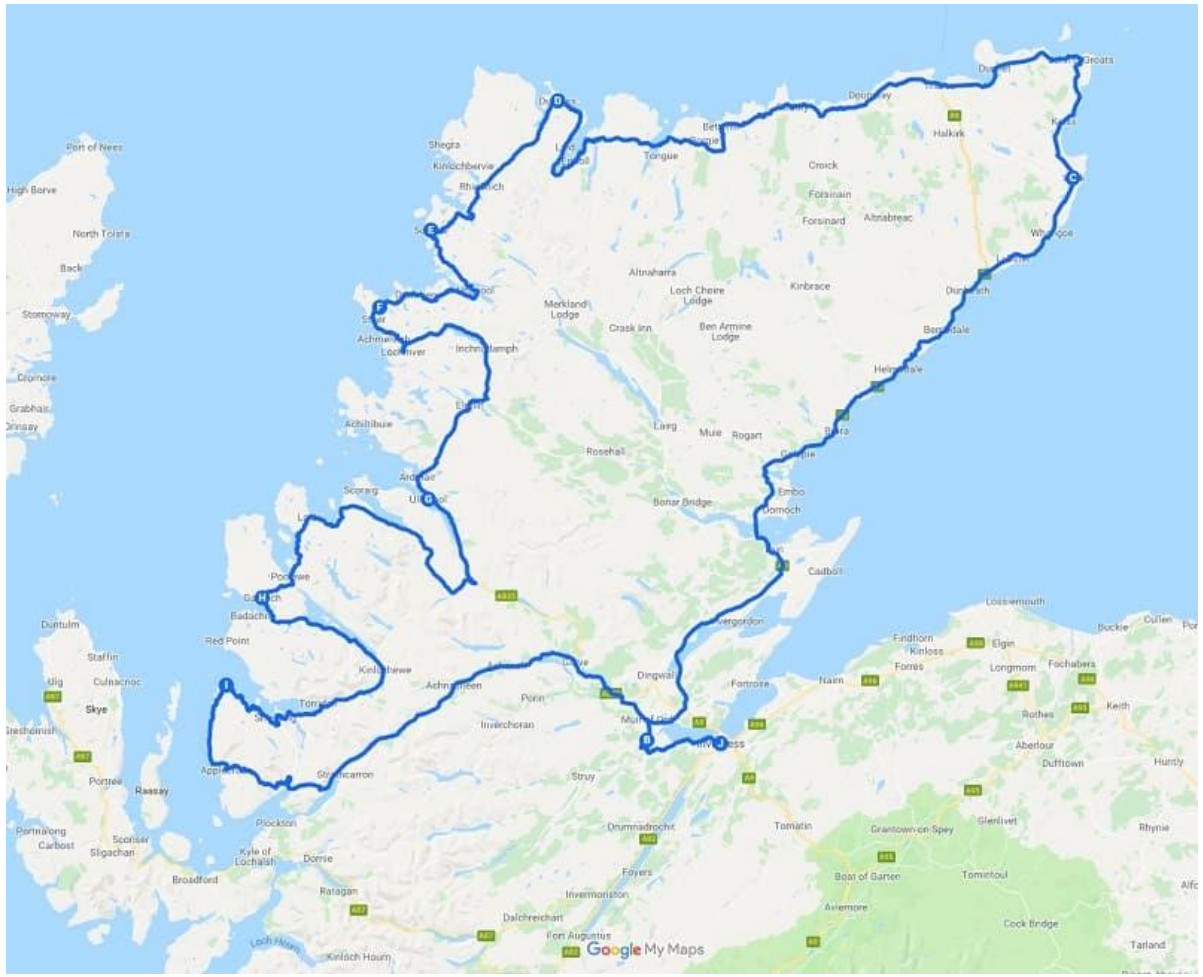




Machine Learning



A TOUR OF ML TECHNOLOGIES



WHEN DO WE USE MACHINE LEARNING

- When we don’t have human expertise on the matter (e.g. navigating on Mars)
- When human expertise cannot be explained (e.g. speech recognition)
- When the models are based on vast amounts of data (genetics and genomics)
- When we need tailor-made models (personalized medicine)

Based on slide by E. Alpaydin

TYPES OF MACHINE LEARNING

SUPERVISED LEARNING

UNSUPERVISED LEARNING

REINFORCEMENT LEARNING

SUPERVISED VS. UNSUPERVISED LEARNING

Supervised learning: classification is supervised learning from examples.

- Supervision: The data (observations, measurements, etc.) are labeled with pre-defined classes, which is
- like a “teacher” gives us the classes (supervision).

Unsupervised learning (clustering)

- Class labels of the data are not given or unknown
- **Goal**: Given a set of data, the task is to establish the existence of classes or clusters in the data

Source: Bing Liu, UIC

REINFORCEMENT LEARNING

Reinforcement learning is a type of machine learning that involves an agent learning to make decisions in an environment by interacting with it and receiving feedback in the form of rewards or punishments. The goal of reinforcement learning is for the agent to learn the optimal actions to take in order to maximize the cumulative reward over time.

Example: Training a robot to navigate a maze by rewarding it for reaching the end of the maze and punishing it for hitting obstacles along the way.

Source: openai.com

TOP PREDICTION ALGORITHMS

	TYPE	NAME	DESCRIPTION	ADVANTAGES	DISADVANTAGES
Linear		**Linear Regression**	**The "best fit" line** through all data points. Predictions are numerical.	Easy to understand — you clearly see what the biggest drivers of the model are.	Sometimes too simple to capture complex relationships between variables. Does poorly with correlated features.
		**Logistic Regression**	The adaptation of **linear regression** to problems of classification (e.g., yes/no questions, groups, etc.	Also easy to understand.	Sometimes too simple to capture complex relationships between variables. Does poorly with correlated features.

Source: data iku

TOP PREDICTION ALGORITHMS

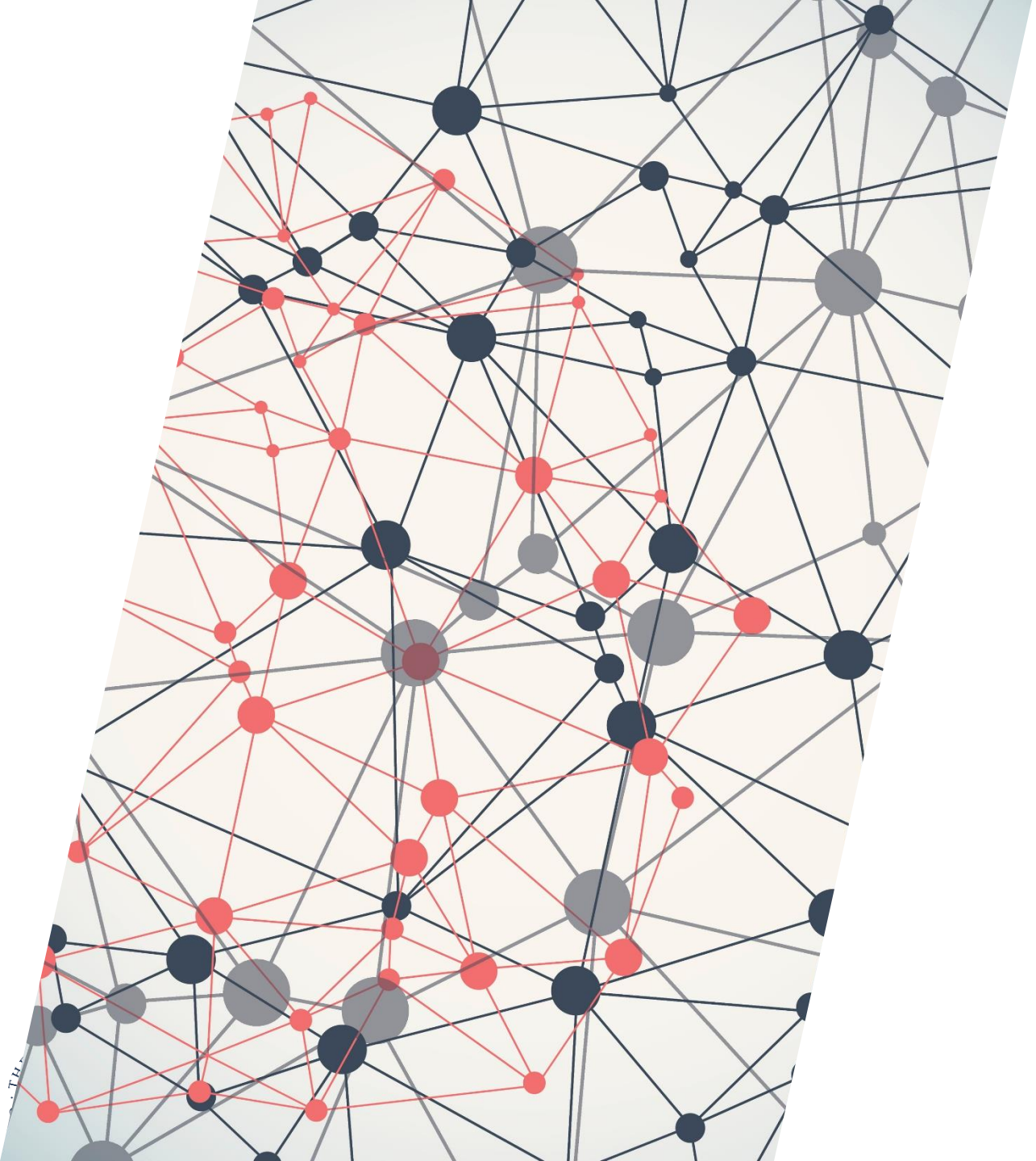
Tree-Based		**Decision Tree**	**A series of yes/no rules based on the features,** forming a tree, to match all possible outcomes of a decision.	Easy to understand.	Not often used on its own for prediction because it's also often too simple and not powerful enough for complex data.
		**Random Forest**	Takes advantage of many decision trees, with rules created from subsamples of features. Each tree is weaker than a full decision tree, but **by combining them we get better overall performance.**	A sort of "wisdom of the crowd". Tends to result in very high quality models. Fast to train.	Models can get very large. Not easy to understand predictions.
		**Gradient Boosting**	Uses even weaker decision trees, that are increasingly **focused on "hard" examples.**	High-performing.	A small change in the feature set or training set can create radical changes in the model. Not easy to understand predictions.

Source: data iku

TOP PREDICTION ALGORITHMS

Neural Networks		**Neural Networks**	Interconnected "neurons" that pass messages to each other. Deep learning uses several **layers of neural networks stacked on top of one another.**	Can handle extremely complex tasks — no other algorithm comes close in image recognition.	Very slow to train, because they often have a very complex architecture. Almost impossible to understand predictions.

Source: data iku



SUPERVISED MACHINE LEARNING (SML)

What is SML?

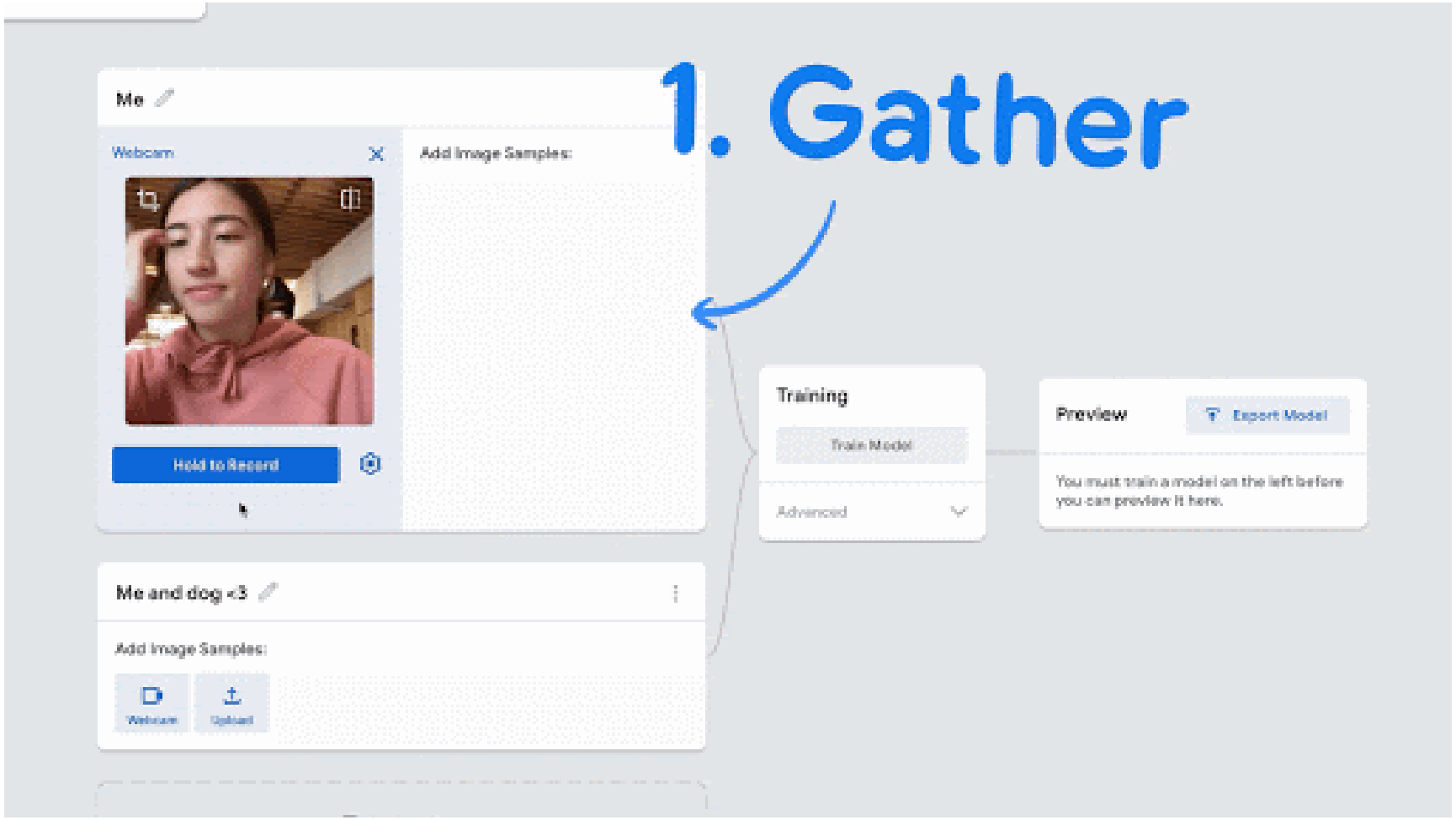
How it works

Advantages of SML

Disadvantages of SML

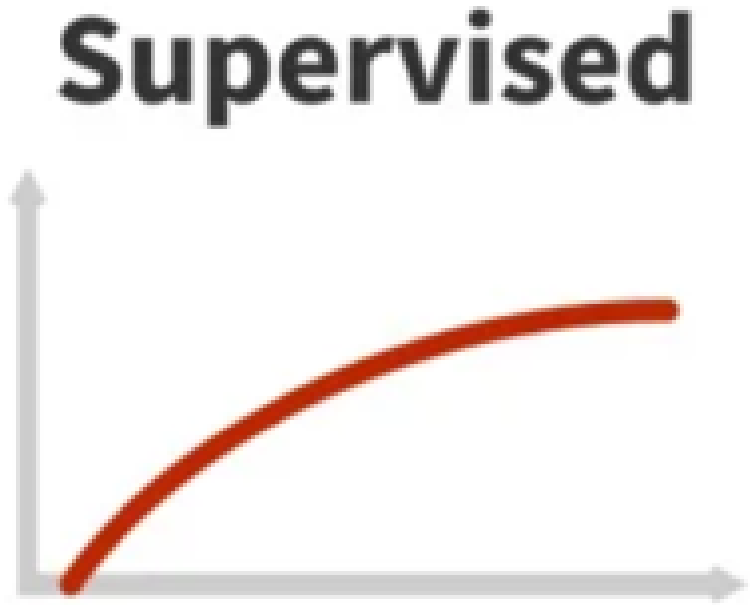
Types of SML algorithms

https://teachablemachine.withgoogle.com/



Supervised Learning

WHAT IS SUPERVISED LEARNING?



Source: Ben Freundorfer

Doug Rose defines supervised learning as "When a data scientist acts like a tutor for the machine, training it by showing it basic rules and giving it an overall strategy." [5]

- **Regression model**
- **Classification model**

REGRESSION VS CLASSIFICATION

The main difference between regression and classification is the type of output they produce. Regression algorithms predict continuous numeric values, while classification algorithms predict categorical values.

Examples:

- Regression: Predicting the price of a house based on its features such as the number of bedrooms, square footage, etc.
- Classification: Predicting whether a customer will churn (cancel their subscription) based on their demographics, transaction history, etc.

Source: openai.com

SUPERVISED MACHINE LEARNING

- We humans learn from past experiences.
- A computer does not "experience."
  - **A computer system learns from data, which** represents "past experiences" in an application domain.
- Our focus: learn **a target function** that can be used to predict the values (**labels**) of a discrete class attribute, e.g.,
  - **high-risk** or **low risk** and **approved** or **not-approved**.
- The task is commonly called: **supervised learning, classification,** or **inductive learning**.

Source: Bing Liu, UIC

EXAMPLE APPLICATION

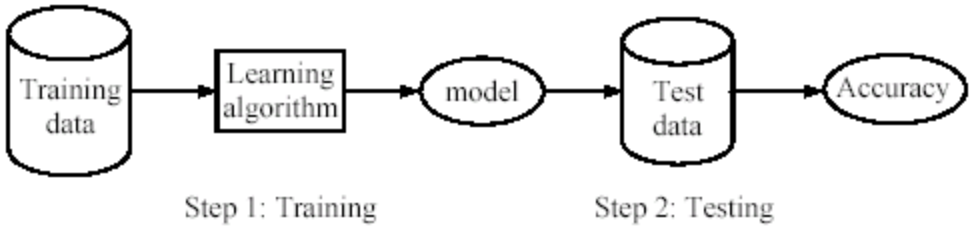
- A credit card company receives thousands of applications for new cards. Each application contains information about an applicant,
  - age
  - annual salary
  - outstanding debts
  - credit rating
  - etc.
- Problem: Decide whether an application should approved, i.e., **classify** applications into two categories, **approved** and **not approved**.

Source: Bing Liu, UIC

SUPERVISED LEARNING PROCESS: TWO STEPS

- Learning or training: Learn a model using the training data (with labels)
- Testing: Test the model using unseen test data (without labels) to assess the model accuracy

$$Accuracy = \frac{\text{Number of correct classifications}}{\text{Total number of test cases}},$$



Source: Bing Liu, UIC

SUPERVISED LEARNING: USES

- Prediction of future cases: Use the rule to predict the output for future inputs
- Knowledge extraction: The rule is easy to understand
- Compression: The rule is simpler than the data it explains
- Outlier detection: Exceptions that are not covered by the rule, e.g., fraud

ADVANTAGES AND DISADVANTAGES OF SUPERVISED MACHINE LEARNING

Advantages

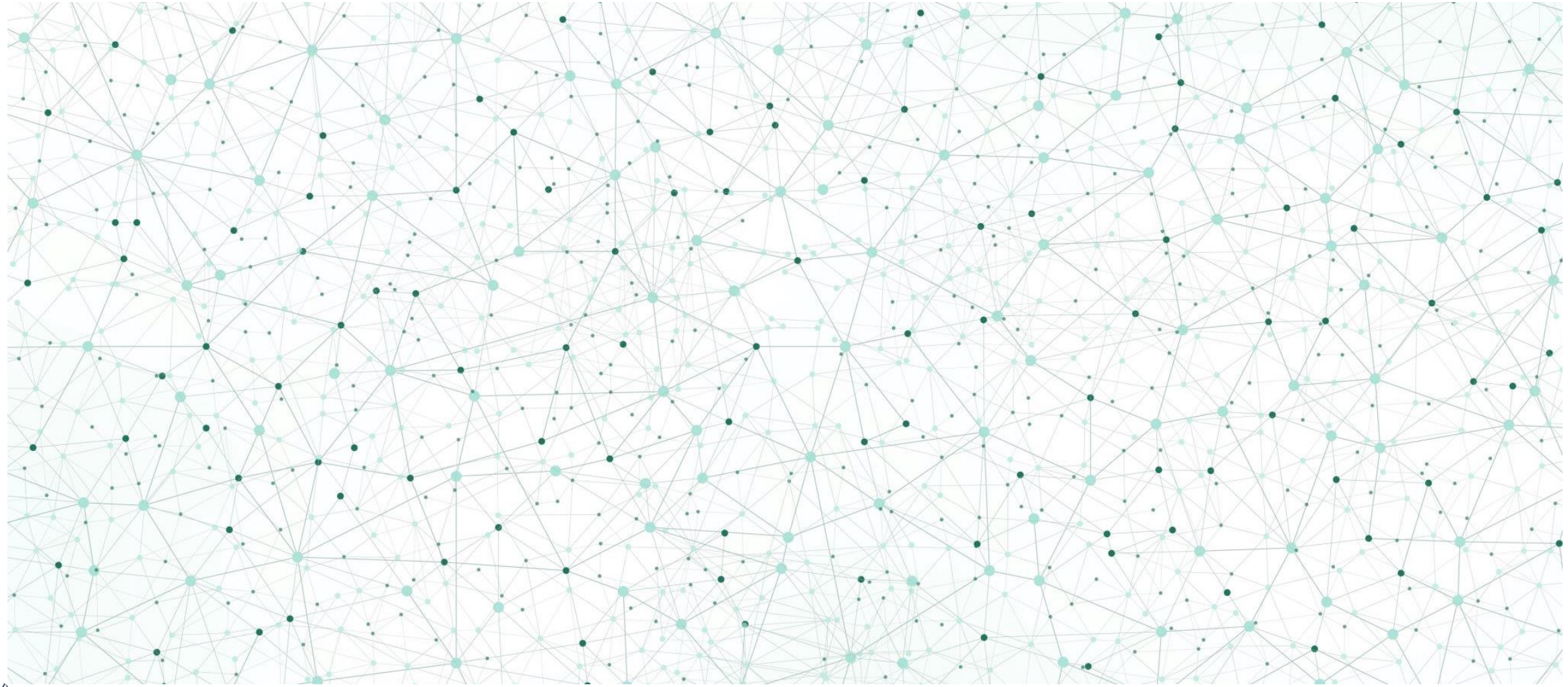
- Accuracy
- Efficiency
- Flexibility
- Transparency

Disadvantages

- Dependence on labelled data
- Overfitting
- Bias
- Lack of adaptability

Source: openai.com

TYPES OF SUPERVISED LEARNING ALGORITHMS



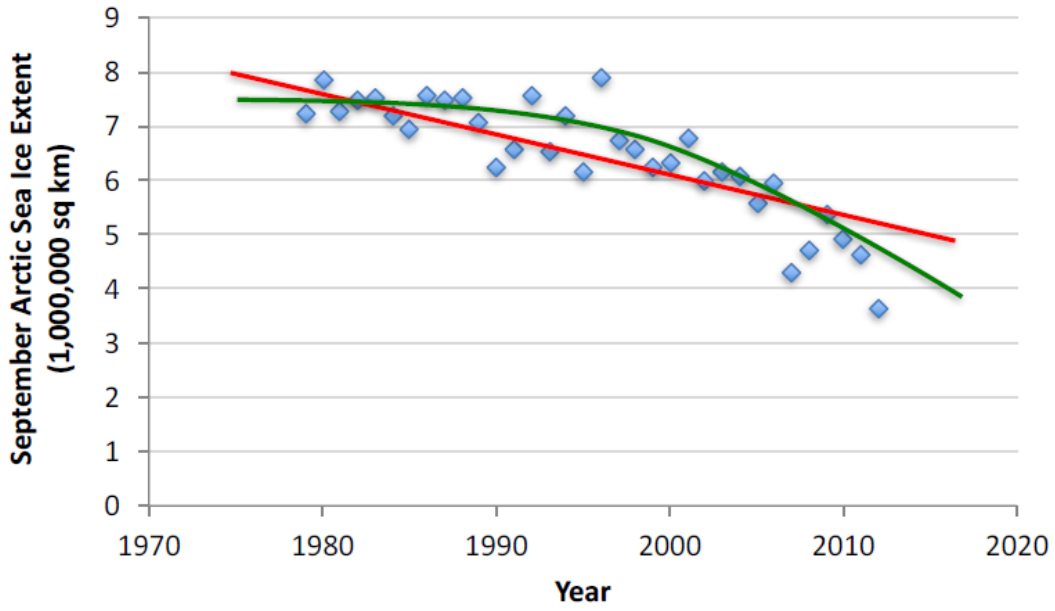
REGRESSION

Regression algorithms are used to predict continuous values, such as predicting the price of a house or the temperature at a specific time. Linear regression, polynomial regression, and support vector regression (SVR) are examples of regression algorithms.

Source: openai.com

REGRESSION MODEL

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression



Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013)

Source: Eric Eaton

REGRESSION
TYPES OF REGRESSION ALGORITHMS:

- Simple Linear Regression
- Multiple Linear Regression
- Polynomial Regression
- Support Vector Regression
- Decision Tree Regression
- Random Forest Regression

LINEAR REGRESSION

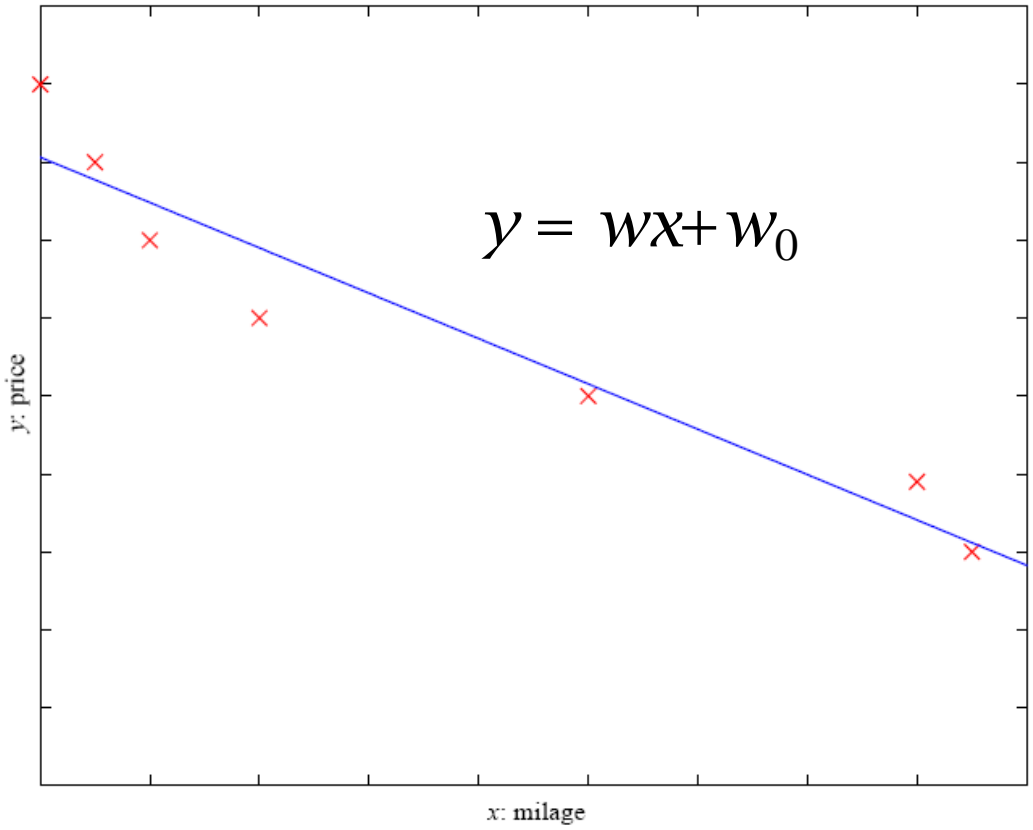
- Example: Price of a used car
- x : car attributes

  y : price

  $$y = g(x \mid \vartheta)$$

  $g(\ )$ model,

  ϑ parameters



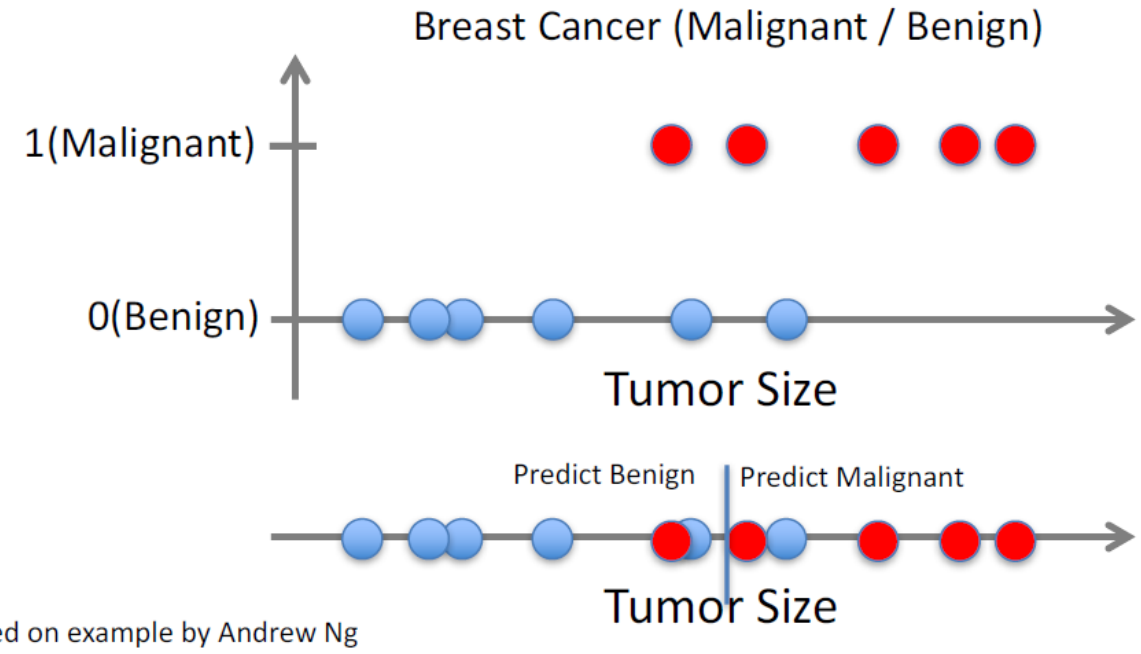
CLASSIFICATION

Classification algorithms are used to classify data into discrete categories, such as classifying an email as spam or not spam. Logistic regression, k-nearest neighbors (KNN), decision trees, random forests, and support vector machines (SVM) are examples of classification algorithms.

Source: openai.com

CLASSIFICATION MODEL

- Given (x_1, y_1), (x_2, y_2), ..., (x_n, y_n)
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Based on example by Andrew Ng

Source: Eric Eaton

CLASSIFICATION

Types of ML Classification Algorithms:

Classification Algorithms can be further divided into the following types:

- Logistic Regression
- K-Nearest Neighbours
- Support Vector Machines
- Kernel SVM
- Naïve Bayes
- Decision Tree Classification
- Random Forest Classification

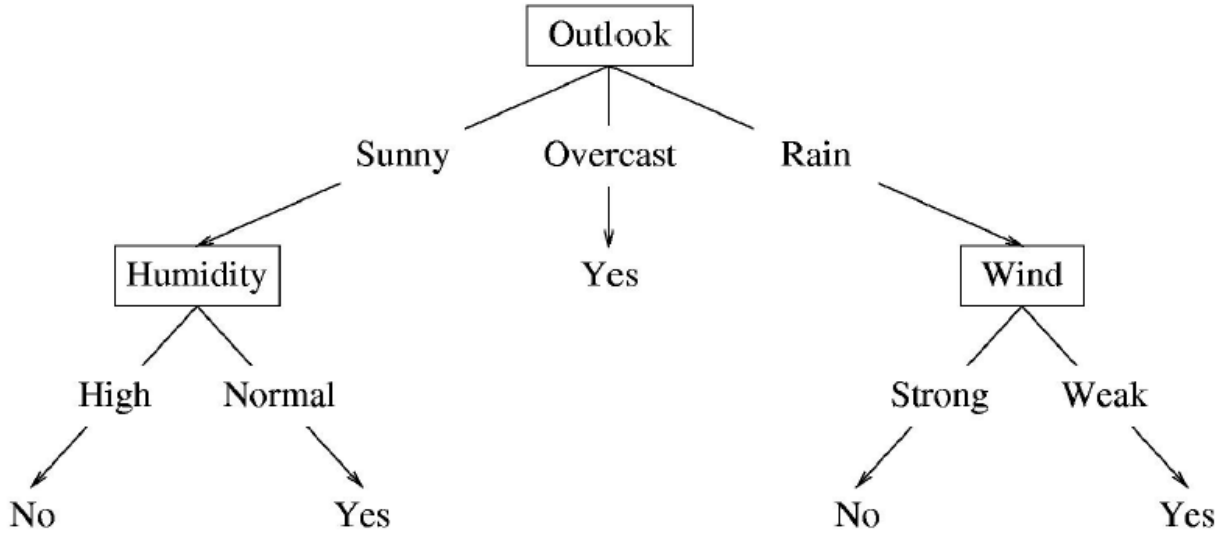
DECISION TREES

Decision trees are a type of algorithm that is used for both regression and classification tasks. They represent decisions and their possible consequences in a tree-like structure and can be used for tasks such as credit risk analysis or medical diagnosis.

Source: openai.com

DECISION TREE

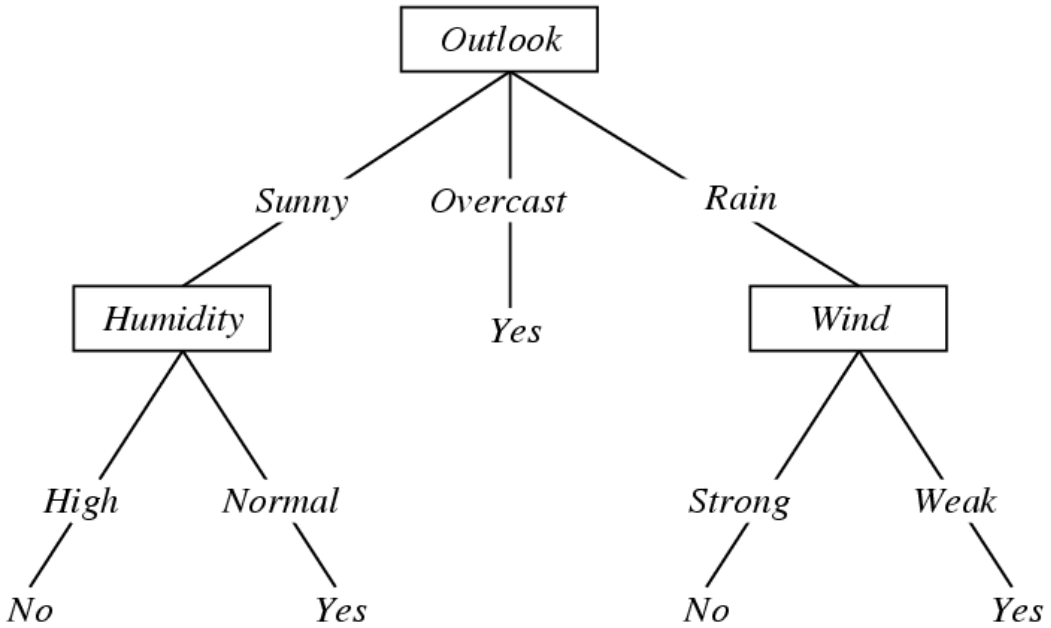
- A possible decision tree for the data:



- Each internal node: test one attribute X_i
- Each branch from a node: selects one value for X_i
- Each leaf node: predict Y (or $p(Y \mid \boldsymbol{x} \in \text{leaf})$)

DECISION TREE LEARNING

Decision Tree for *PlayTennis*

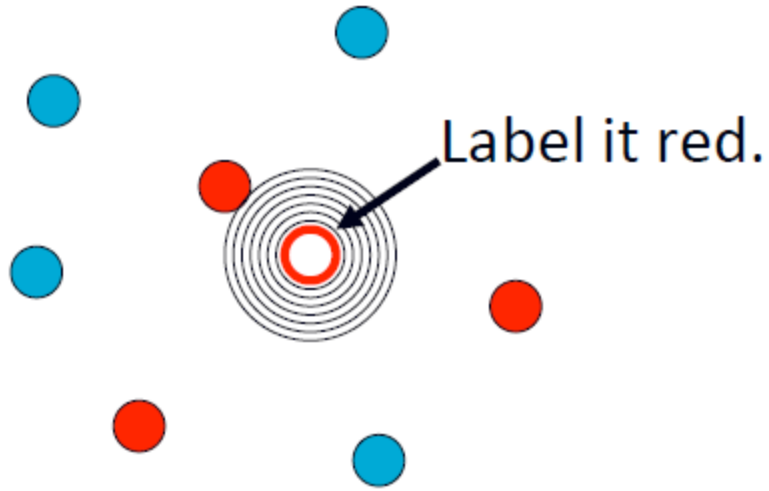


Problem Setting:

- Set of possible instances *X*
  - each instance *x* in *X* is a feature vector
  - e.g., <*Humidity=low, Wind=weak, Outlook=rain, Temp=hot*>
- Unknown target function $f : X \rightarrow Y$
  - *Y* is discrete valued
- Set of function hypotheses $H=\{ h \mid h : X \rightarrow Y \}$
- each hypothesis *h* is a decision tree
- trees sorts *x* to leaf, which assigns *y*

K-NEAREST NEIGHBOUR

- **1-Nearest Neighbour**
- One of the simplest of all machine learning classifiers
- Simple idea: label a new point the same as the closest known point



Source: Eric Eaton

K-NEAREST NEIGHBOUR

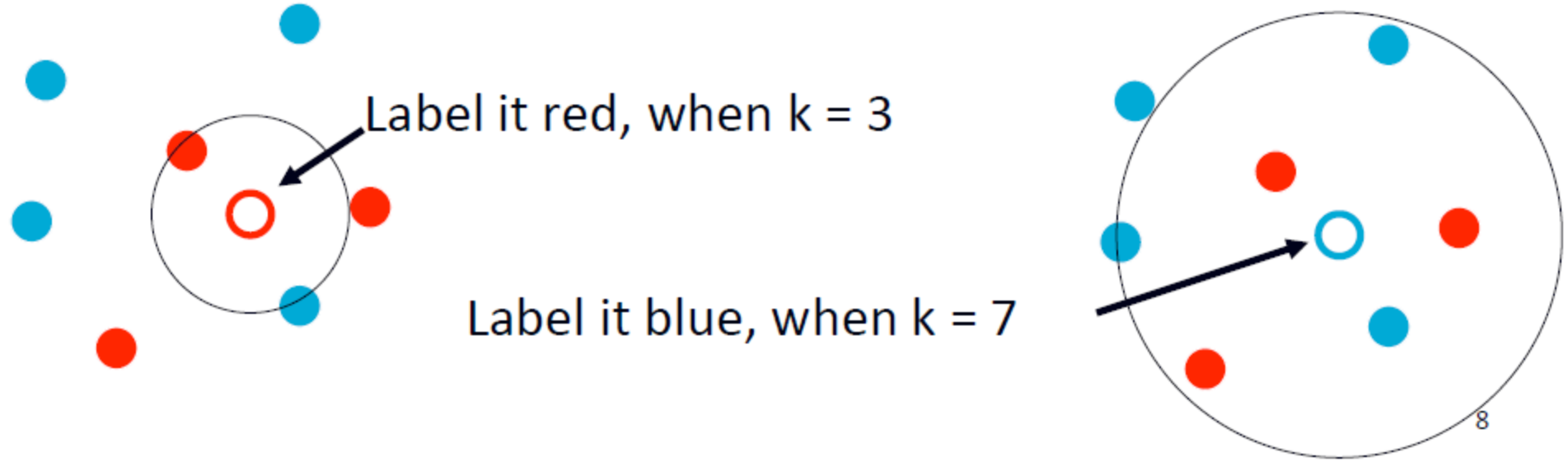
- **Four Aspects of an Instance---Based Learner:**
- 1. A distance metric
- 2. How many nearby neighbours to look at?
- 3. A weighting function (optional)
- 4. How to fit with the local points?

Adapted from "Instance---Based Learning" lecture slides by Andrew Moore, CMU.

Source: Eric Eaton

K-NEAREST NEIGHBOUR

- Generalizes 1---NN to smooth away noise in the labels
- A new point is now assigned the most frequent label of its *k* nearest neighbours



Source: Eric Eaton

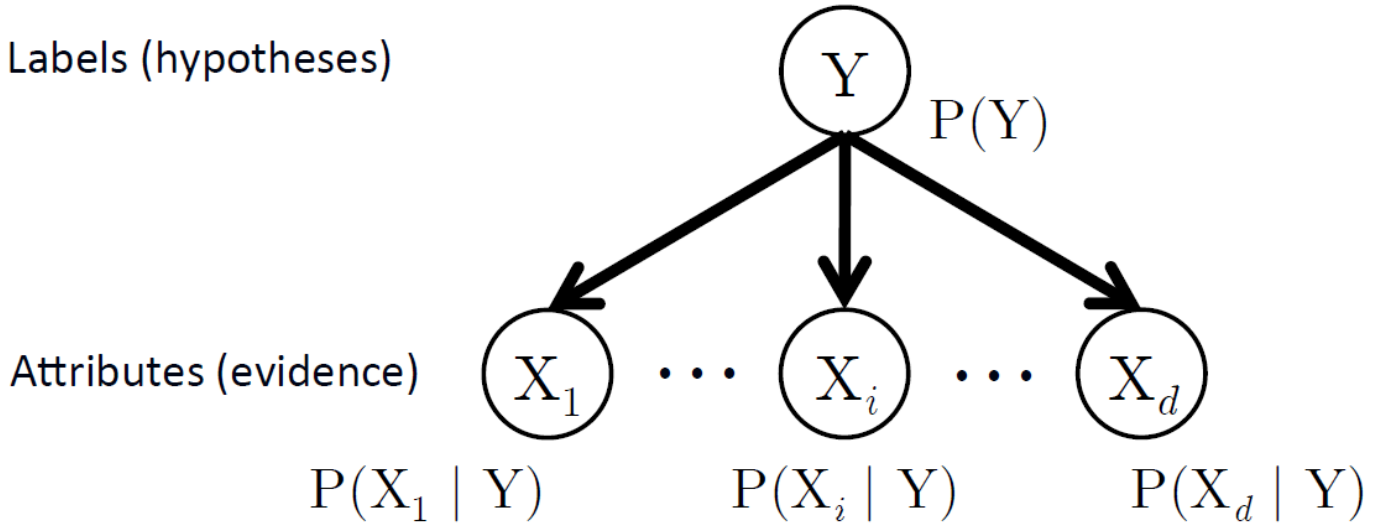
NAIVE BAYES

Naive Bayes is a probabilistic algorithm that is commonly used for text classification tasks, such as sentiment analysis or spam detection.

Source: openai.com

NAIVE BAYES

The Naïve Bayes Graphical Model

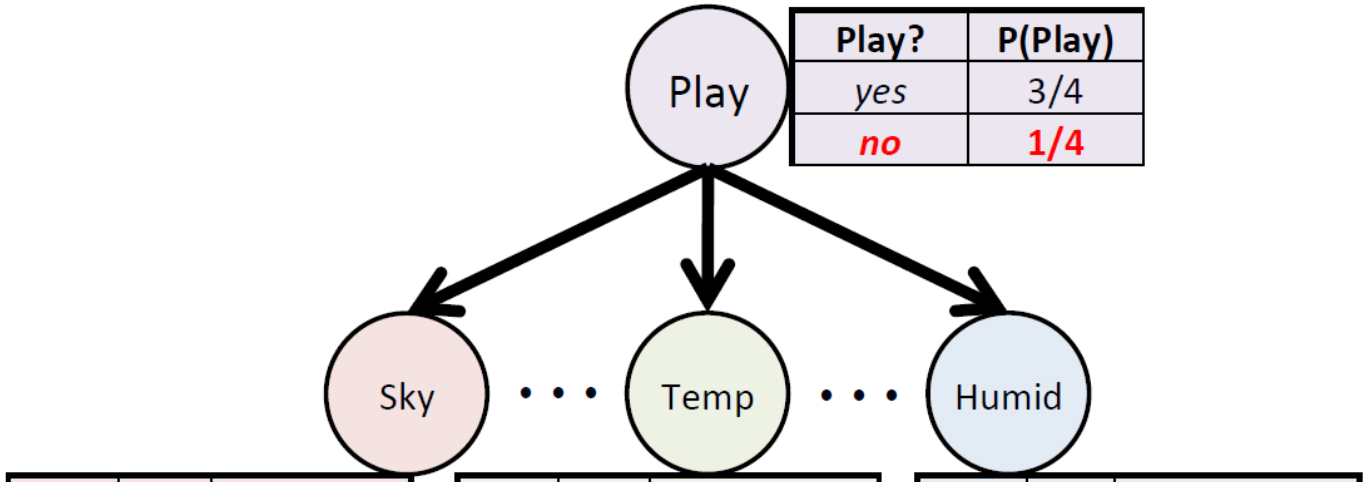


- Nodes denote random variables
- Edges denote dependency
- Each node has an associated conditional probability table (CPT), conditioned upon its parents

Source: Eric Eaton

NAIVE BAYES

Example NB Graphical Model



Play?	P(Play)
yes	3/4
no	**1/4**

Sky	Play?	P(Sky | Play)
sunny	*yes*	4/5
rainy	***yes***	**1/5**
sunny	*no*	1/3
rainy	***no***	**2/3**

Temp	Play?	P(Temp | Play)
warm	*yes*	4/5
cold	***yes***	**1/5**
warm	*no*	1/3
cold	***no***	**2/3**

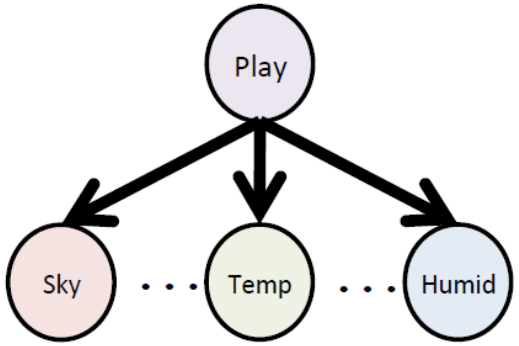
Humid	Play?	P(Humid | Play)
high	*yes*	3/5
norm	***yes***	**2/5**
high	*no*	2/3
norm	***no***	**1/3**

- Some **redundancies** in CPTs that can be eliminated

Source: Eric Eaton

NAIVE BAYES

Example Using NB for Classification



Play?	P(Play)
yes	3/4
no	1/4

Temp	Play?	P(Temp | Play)
warm	*yes*	4/5
cold	*yes*	1/5
warm	*no*	1/3
cold	*no*	2/3

Sky	Play?	P(Sky | Play)
sunny	*yes*	4/5
rainy	*yes*	1/5
sunny	*no*	1/3
rainy	*no*	2/3

Humid	Play?	P(Humid | Play)
high	*yes*	3/5
norm	*yes*	2/5
high	*no*	2/3
norm	*no*	1/3

$$h(\mathbf{x}) = \arg\max_{y_k} \log P(Y = y_k) + \sum_{j=1}^{d} \log P(X_j = x_j \mid Y = y_k)$$

Goal: Predict label for $\mathbf{x}$ = (rainy, warm, normal)

Source: Eric Eaton

NAIVE BAYES

- **Advantages:**
  - Fast to train (single scan through data)
  - Fast to classify
  - Not sensitive to irrelevant features
  - Handles real and discrete data
  - Handles streaming data well
- **Disadvantages:**
  - Assumes independence of features

NEURAL NETWORKS

Neural networks are a type of machine learning algorithm that use multiple layers of artificial neurons to learn from labeled data. They can be used for both regression and classification tasks and have shown excellent results in areas such as image and speech recognition.

Source: openai.com

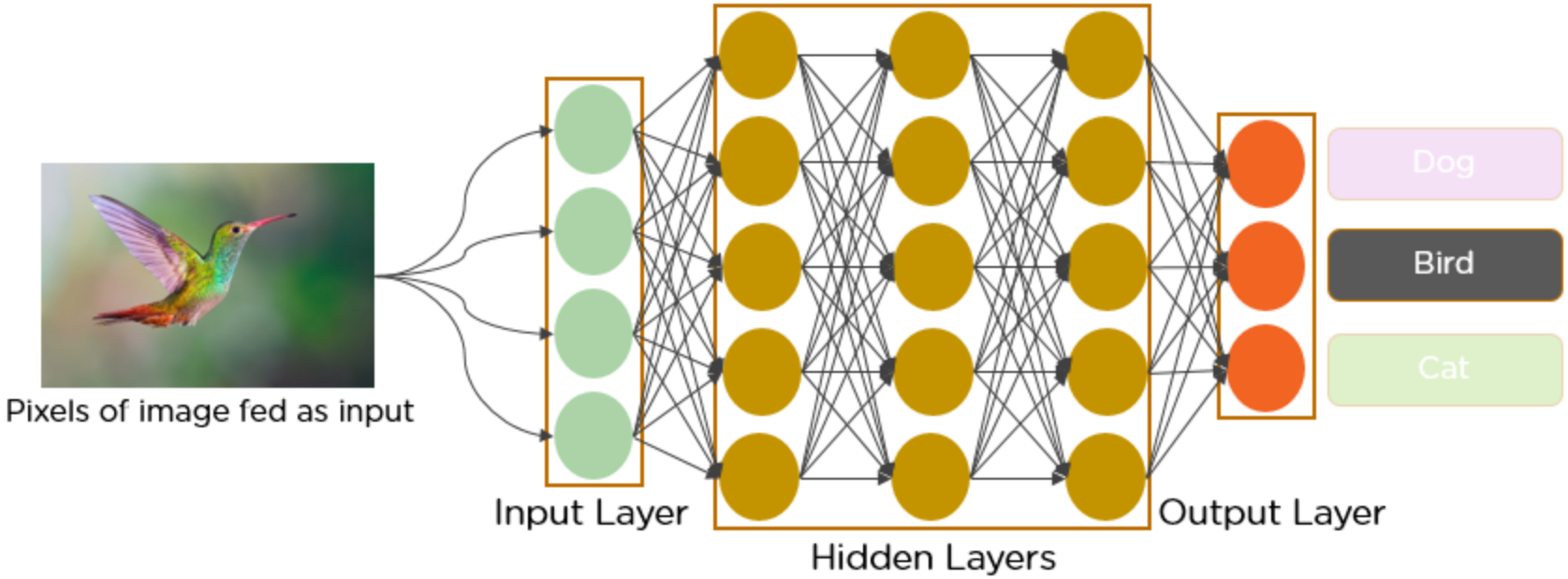
NEURAL NETWORKS

Neural networks refer to a biological phenomenon comprised of interconnected neurons that exchange messages with each other. This idea has now been adapted to the world of machine learning and is called ANN (Artificial Neural Networks).

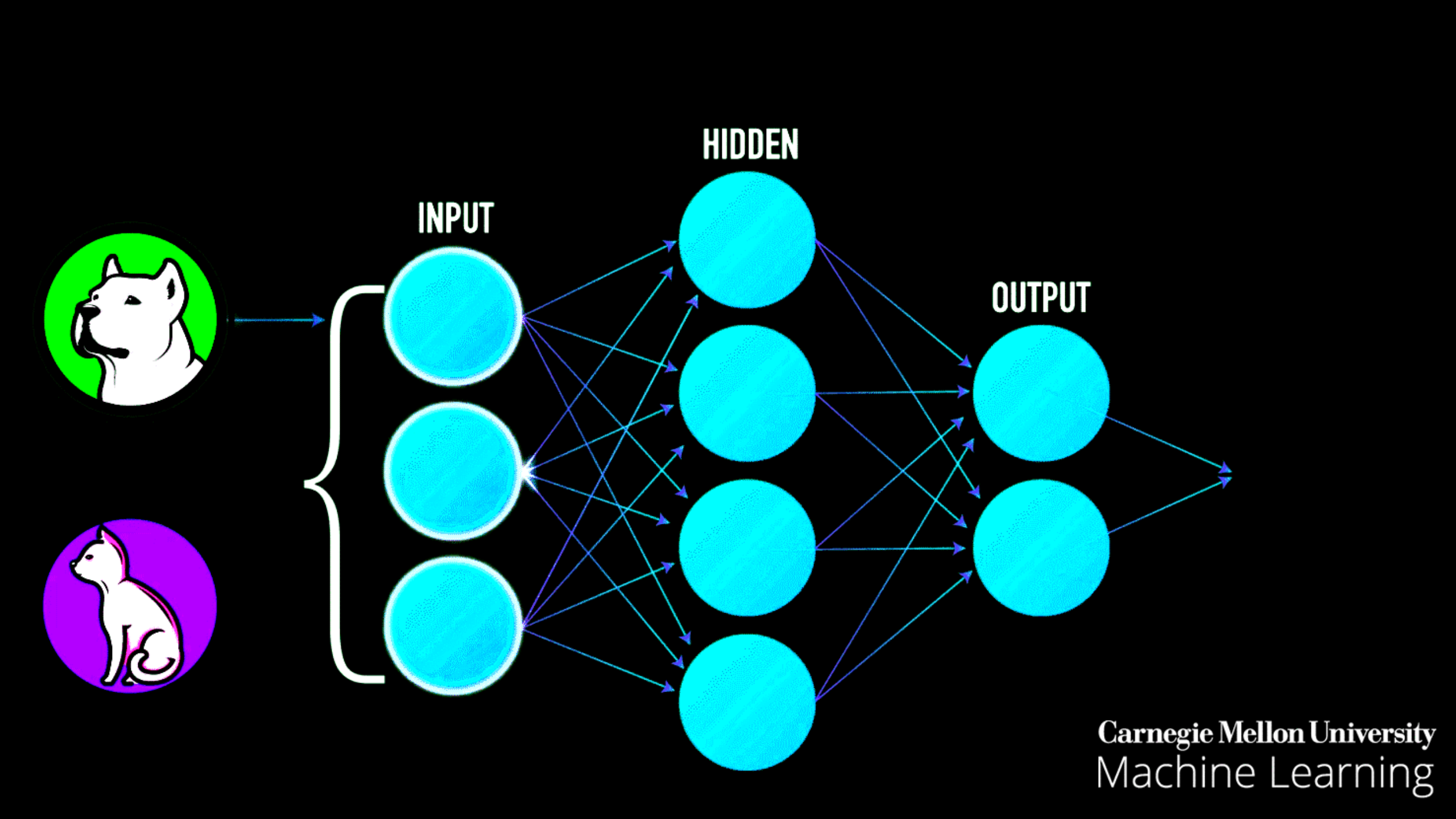
Deep learning, which you've heard a lot about, can be done with several layers of neural networks put one after the other. ANNs are a family of models that are taught to adopt cognitive skills.

Source: Dataiku

MACHINE LEARNING CLASSIFICATION NETWORK



The training process assigns weights to the arcs of the network so it output the right classifications



Carnegie Mellon University
Machine Learning

REFERENCES

1. Zhydik, O., 2021. *Six Powerful Use Cases for Machine Learning in Manufacturing*. [online] Eleks.com. Available at: https://eleks.com/blog/machine-learning-in-manufacturing/
2. Green, J., 2020. *5 ways you can use Machine Learning in manufacturing*. [online] Ancoris.com. Available at: https://www.ancoris.com/blog/5-ways-machine-learning-manufacturing
3. Doty, C., 2021. *10 Ways Machine Learning Will Reshape Manufacturing in 2021 - RapidMiner*. [online] RapidMiner. Available at: https://rapidminer.com/blog/10-ways-machine-learning-in-manufacturing-2021/
4. Rai, R., Tiwari, M., Ivanov, D. and Dolgui, A., 2021. Machine learning in manufacturing and industry 4.0 applications. *International Journal of Production Research*, 59(16), pp.4773-4778.
5. GLOSSARY, Artificial Intelligence Foundations: Machine Learning, LinkedinLearn
6. McCarthy, J., 2004. WHAT IS ARTIFICIAL INTELLIGENCE?. Computer Science Department Stanford University Stanford.

RESOURCES

AAAI. Machine Learning. *http://www.aaai.org/Pathfinder/html/machine.html*

Dietterich, T. (2003). Machine Learning. *Nature Encyclopedia of Cognitive Science.*

Doyle, P. Machine Learning. *http://www.cs.dartmouth.edu/~brd/Teaching/AI/Lectures/Summaries/learning.html*

Dyer, C. (2004). Machine Learning. *http://www.cs.wisc.edu/~dyer/cs540/notes/learning.html*

Mitchell, T. (1997). *Machine Learning*.

Nilsson, N. (2004). Introduction to Machine Learning. *http://robotics.stanford.edu/people/nilsson/mlbook.html*

Russell, S. (1997). Machine Learning. *Handbook of Perception and Cognition*, Vol. 14, Chap. 4.

Russell, S. (2002). *Artificial Intelligence: A Modern Approach*, Chap. 18-20. *http://aima.cs.berkeley.edu*