WEEK 1

MACHINE LEARNING

UNLIMITED

Presented by **Asst. Prof. Dr. Tuchsanai Ploysuwan**

Machine Learning

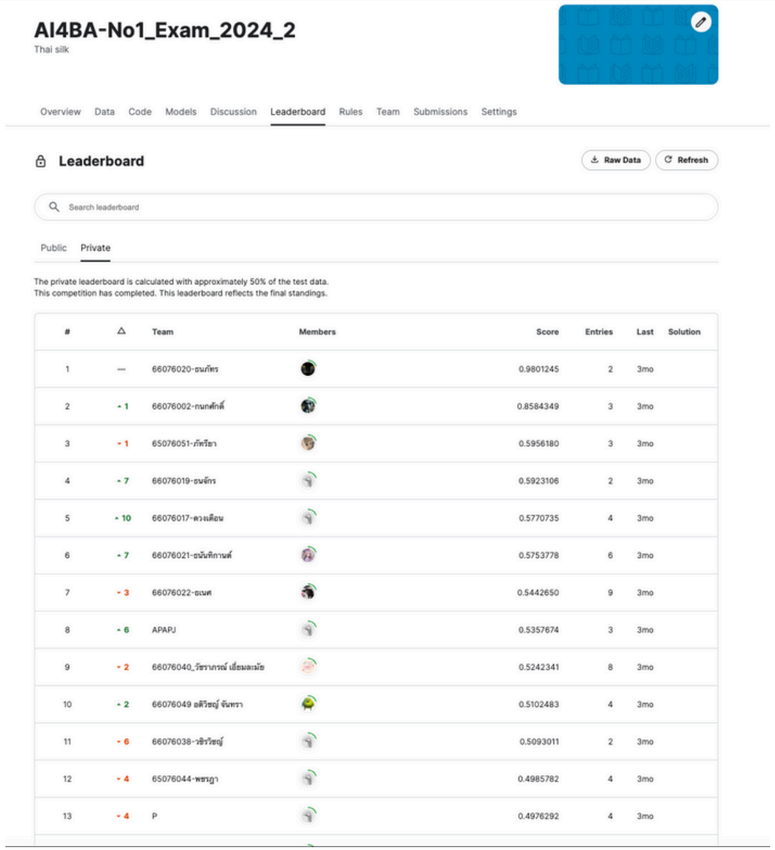
06048203

Document

Week	Topic
1	Introduction to Machine Learning
2	Fundamental and Prompt Engineering
3	Advanced Prompting Techniques
4	AI Agent Programming for Machine Learning
5	AI Agent for Machine Learning
6	Regression & Advance Regression
7	Logistic Regression
8	K-nearest neighbors
9	Support vector machine
10	Naïve Bayes
11	Decision tree
12	Random forest
13-15	Application in Real world

Criteria	Description	Weight
Kaggle Assignments	Kaggle is an online Competition (15x3)	45%
Final Exam	Performance in exams, assessing theoretical knowledge and problem-solving skills	55%

ลำดับเปอร์เซ็นต์	จำนวนคน (ประมาณ)	คะแนน	ระดับผลสัมฤทธิ์
Top 10%	4 คน	100	ยอดเยี่ยม
11–20%	4 คน	95	ดีเยี่ยม
21–40%	8 คน	90	ดีมาก
41–60%	8 คน	85	ดี
61–80%	8 คน	75	พอใช้
Bottom 20%	5 คน	65	ควรปรับปรุง



AI4BA-No1_Exam_2024_2

Thai silk

Overview Data Code Models Discussion Leaderboard Rules Team Submissions Settings

Leaderboard

Raw Data Refresh

Search leaderboard

Public Private

The private leaderboard is calculated with approximately 50% of the test data.
This competition has completed. This leaderboard reflects the final standings.

#	△	Team	Members	Score	Entries	Last	Solution
1	—	66076020-ธนภัทร		0.9801245	2	3mo	
2	▲ 1	66076002-กนกศักดิ์		0.8584349	3	3mo	
3	▼ 1	65076051-ภัทรีรา		0.5956180	3	3mo	
4	▲ 7	66076019-ธนพัทร		0.5923106	2	3mo	
5	▲ 10	66076017-ดวงเดือน		0.5770735	4	3mo	
6	▲ 7	66076021-ธนันทิกานต์		0.5753778	6	3mo	
7	▼ 3	66076022-ธนพล		0.5442650	9	3mo	
8	▲ 6	APAPJ		0.5357674	3	3mo	
9	▼ 2	66076040_วิรวากรณ์ เอื้อมดอมัย		0.5242341	8	3mo	
10	▲ 2	66076049 อดิวิชญ์ จันพรา		0.5102483	4	3mo	
11	▼ 6	66076038-วชิรวิชญ์		0.5093011	2	3mo	
12	▼ 4	65076044-พชรฐา		0.4985782	4	3mo	
13	▼ 4	P		0.4976292	4	3mo	

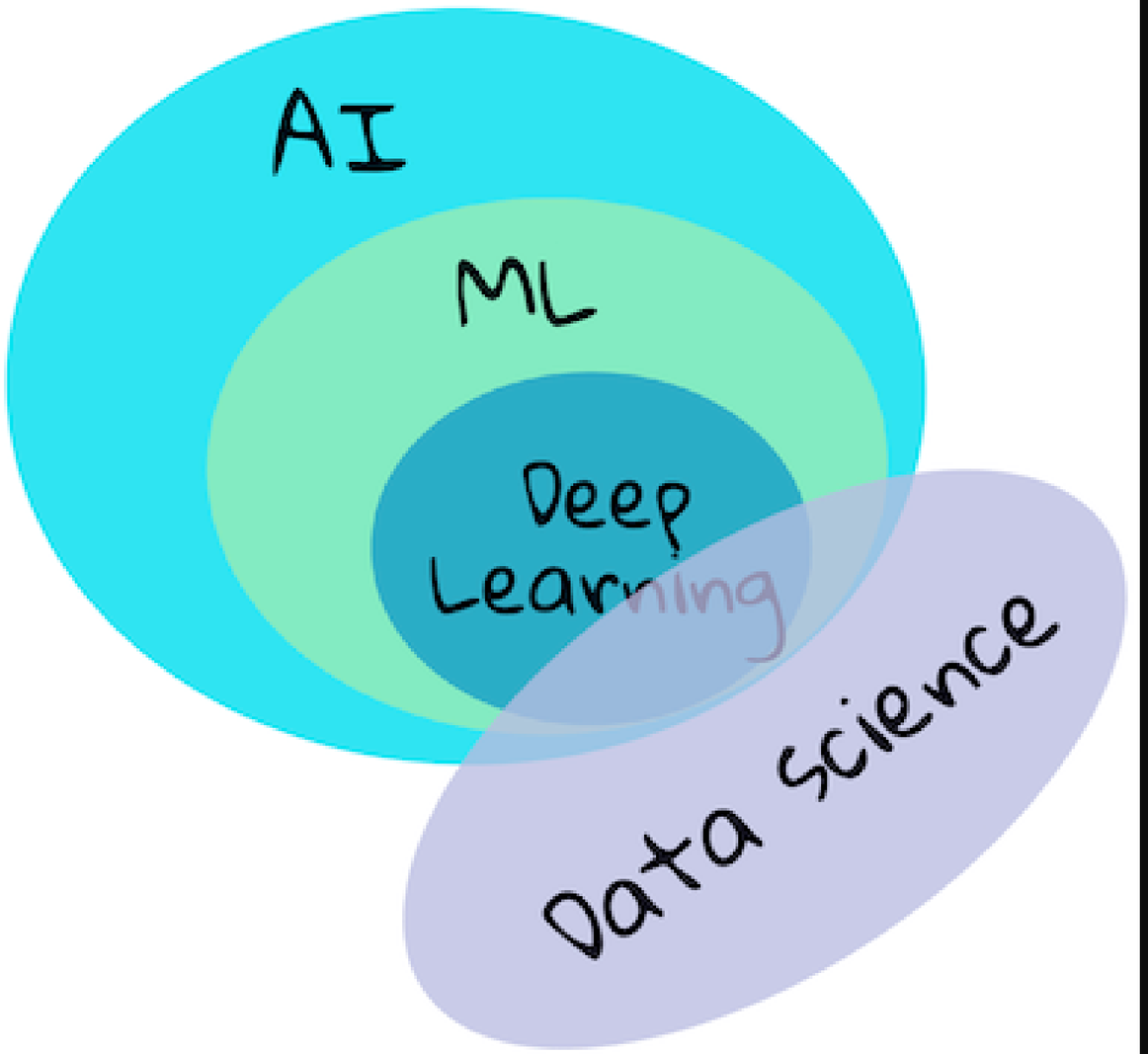
ทุจริต – ปรับตก

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรายชื่อนักศึกษา แยกตามชื่อวิชา
ภาคการศึกษา 1 ปีการศึกษา 2568
คณะทุกคณะ

รหัสวิชา 06048203 รายวิชา MACHINE LEARNING หน่วยกิต 3 กลุ่ม รวมทุกกลุ่ม
อาจารย์ผู้สอน ผศ. ดร.ทัศน์ย พลอยสุวรรณ

ลำดับที่	รหัส น.ศ.	ชื่อ-นามสกุล	หมายเหตุ
1	68076005	นายกฤตนัย กัปตพล	
2	68076006	นายก้องภพ อภิศักดิ์กุล	
3	68076009	นางสาวจิรา วารีรักษ์	
4	68076012	นายชนาธิป อมรชีวศิลปกุล	
5	68076013	นายชยานันต์ นันทกิจอำรง	
6	68076015	นางสาวชาลิสา น้อยบุญเติม	
7	68076022	นายณัฐพล ทันสมัย	
8	68076023	นายณัฐพล แพรศรี	
9	68076028	นายเตชินท์ เจริญพร้อม	
10	68076029	นายทองแท้ ศิริมาก	
11	68076033	นายธีระศักดิ์ เจิมทอง	
12	68076035	นางสาวนลินี อัสแสง	
13	68076037	นางสาวนันทัชพร เปลี่ยนปรีชา	
14	68076038	นายนิติพัฒน์ บุญเกตุ	
15	68076040	นางสาวเบญจมาภรณ์ เจียนเกาะ	
16	68076043	นายปิยพนธ์ เปี่นวงษ์	
17	68076044	นายปิยวัฒน์ ยศธนายน	
18	68076045	นายไปรย์ อัครปาลิน	
19	68076047	นายพัทธดนย์ จันทร์เกตุ	
20	68076050	นางสาวแพรว พลอยกระจ่างศรี	
21	68076052	นายภาณุพงศ์ บุตรดี	
22	68076053	นายภาณุวิชญ์ บุญหนัก	
23	68076054	นายวรวุฒิ จรุงคงเดช	
24	68076056	นายวิหวัส แสนนรินทร์	
25	68076057	นายวิริยะ พลซำนิ	
26	68076059	นายศิภิศักย์ สุขโต	
27	68076060	นายศิริพล ศรีเฮงไพบูลย์	
28	68076061	นางสาวศิริลักษณ์ ทองพัด	
29	68076063	นางสาวศุภิสรา ศิริเสรี	
30	68076064	นายสมศักดิ์ เนื่องมัจฉา	
31	68076065	นายสรรเสริญ มากเจริญ	
32	68076066	นางสาวสรุตา น้อยหมอ	
33	68076069	นายสุทธวัฒน์ หวังดีธนาโชค	
34	68076071	นายอดิธัช ติรสุวรรณวาสี	
35	68076072	นายอนุชา เอื้อสุขกุล	
36	68076073	นายอาทิตย์ อัสโม	
37	68076078	นายศุภณัฐ ธีรรักษ์ตระกูล	

รวมจำนวนนักศึกษาทั้งสิ้น 37 คน

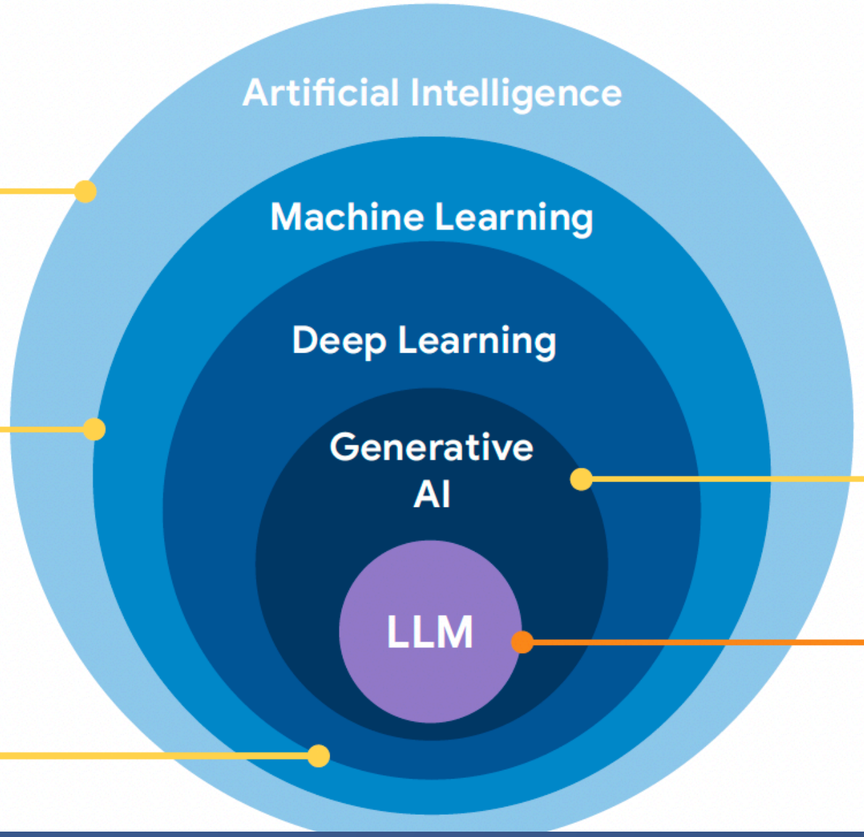


- Artificial Intelligence (AI) is a field focused on creating intelligent systems that can perform tasks requiring human-like intelligence. It encompasses both narrow AI, which is designed for specific tasks, and general AI, which aims to achieve human-level intelligence.
- Machine Learning (ML) is a subset of AI that involves developing algorithms and models that allow computers to learn from data and improve their performance without being explicitly programmed. ML focuses on pattern recognition, prediction, and decision-making based on learned patterns.
- Deep Learning is a specialized branch of ML that utilizes deep neural networks with multiple layers to automatically learn hierarchical representations of data. It excels at tasks involving complex patterns and feature extraction, and often requires substantial computational resources.
- Data Science is an interdisciplinary field that involves extracting knowledge and insights from data using scientific methods, processes, algorithms, and tools. It encompasses various techniques, including AI, ML, and Deep Learning, to analyze and interpret large and complex datasets to solve problems and make informed decisions.

Artificial Intelligence (AI)
A field, which combines computer science and robust datasets, to enable problem-solving

Machine Learning (ML)
A branch of AI and computer science which focuses on the use of data and algorithms to imitate the way that humans learn

Deep Learning (DL)
A neural network with three or more layers which attempt to simulate the behavior of the human brain, allowing it to "learn" from large amounts of data



Generative AI
A branch of AI that involves using machine learning algorithms to create new content using previously created content

Large Language Model (LLM)
A generative AI model that is designed to understand human-like text and can perform a variety of natural language processing tasks.

“Why use machine learning (or deep learning)?”

(maybe not very simple…)

“If you can build a simple rule-based system that doesn’t require machine learning, do that.”

Good reason:

Why not?

Better reason: For a complex problem, can you think of all the rules?

(probably not)

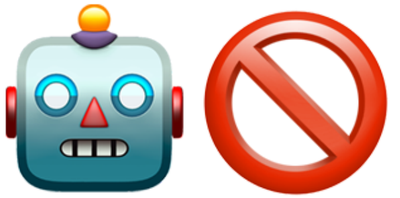
What deep learning is good for

Problems with long lists of rules—when the traditional approach fails, machine learning/deep learning may help.

Continually changing environments—deep learning can adapt (‘learn’) to new scenarios.

Discovering insights within large collections of data—can you imagine trying to hand-craft rules for what 101 diSerent kinds of food look like?

What deep learning is not good for



When you need explainability—the patterns learned by a deep learning model are typically uninterpretable by a human.

When the traditional approach is a better option — if you can accomplish what you need with a simple rule-based system.

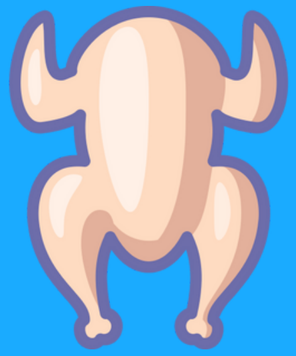
When errors are unacceptable — since the outputs of deep learning model aren’t always predictable.

When you don’t have much data — deep learning models usually require a fairly large amount of data to produce great results.

(though we’ll see how to get great results without huge amounts of data)

Traditional programming

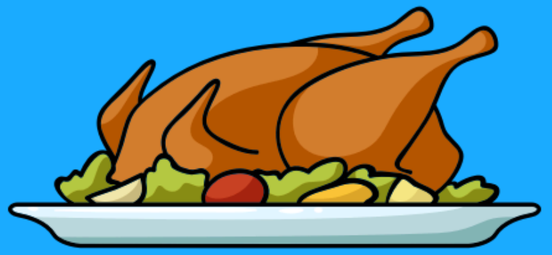
Inputs



Rules

1. Cut vegetables
2. Season chicken
3. Preheat oven
4. Cook chicken for 30-minutes
5. Add vegetables

Output

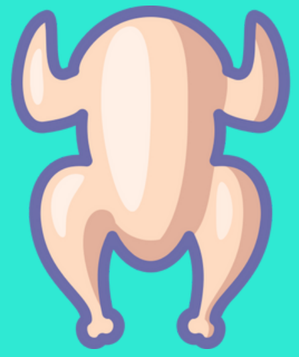


Starts with

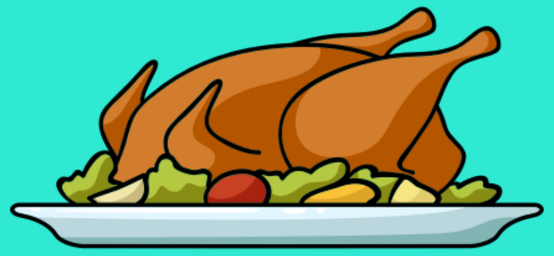
Makes

Machine learning algorithm

Inputs



Output



Rules

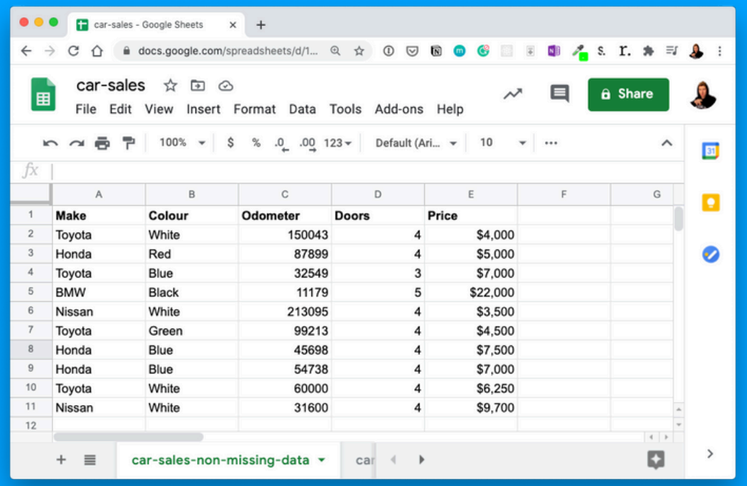
1. Cut vegetables
2. Season chicken
3. Preheat oven
4. Cook chicken for 30-minutes
5. Add vegetables

Starts with

Figures out

Machine Learning vs. Deep Learning

Machine Learning



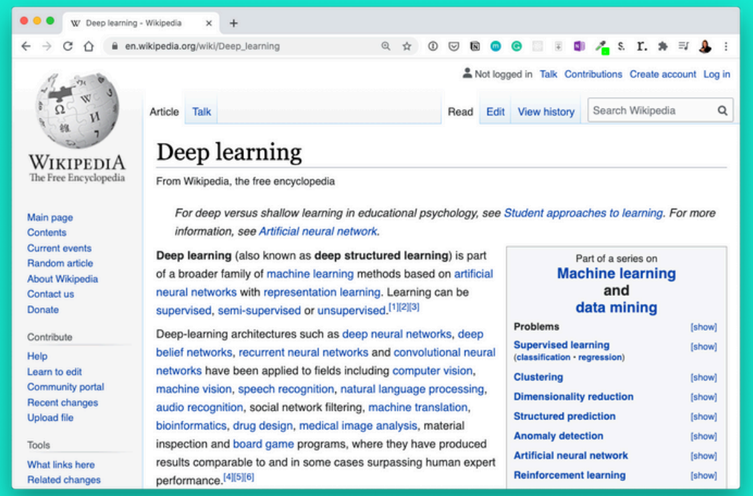
	Make	Colour	Odometer	Doors	Price
1	Toyota	White	150043	4	$4,000
2	Honda	Red	87899	4	$5,000
3	Toyota	Blue	32549	3	$7,000
4	BMW	Black	11179	5	$22,000
5	Nissan	White	213095	4	$3,500
6	Toyota	Green	99213	4	$4,500
7	Honda	Blue	45698	4	$7,500
8	Honda	Blue	54738	4	$7,000
9	Toyota	White	60000	4	$6,250
10	Nissan	White	31600	4	$9,700

Algorithm: gradient boosted machine

dmlc XGBoost

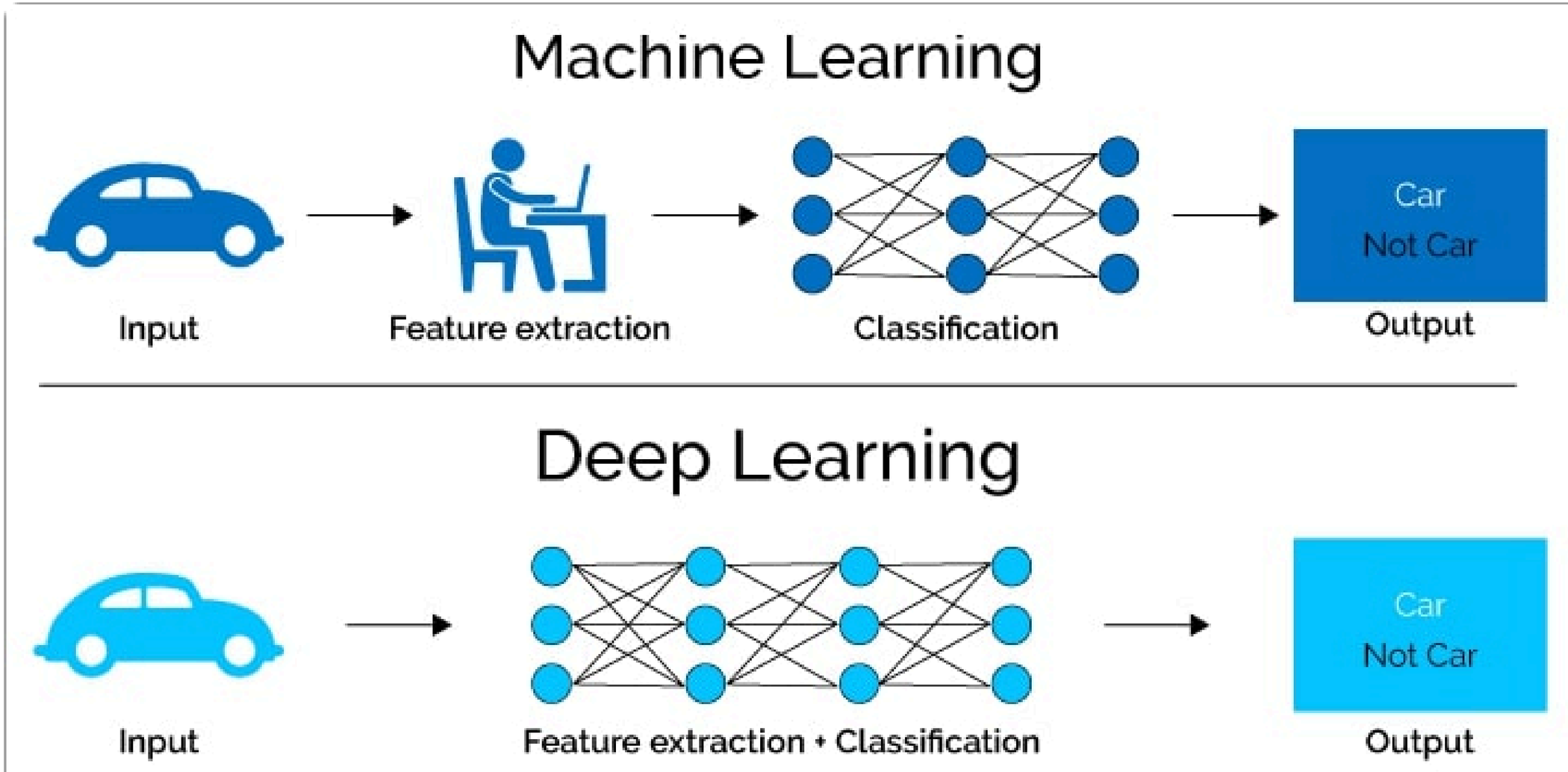
Structured data

Deep Learning



Algorithm: neural network

Unstructured data


Machine Learning
Input
Feature extraction
Classification
Car
Not Car
Output
Deep Learning
Input
Feature extraction + Classification
Car
Not Car
Output

Machine Learning vs. Deep Learning

(common algorithms)

- Random forest
- Gradient boosted models
- Naive Bayes
- Nearest neighbour
- Support vector machine
- …many more

(since the advent of deep learning these are often referred to as "shallow algorithms")

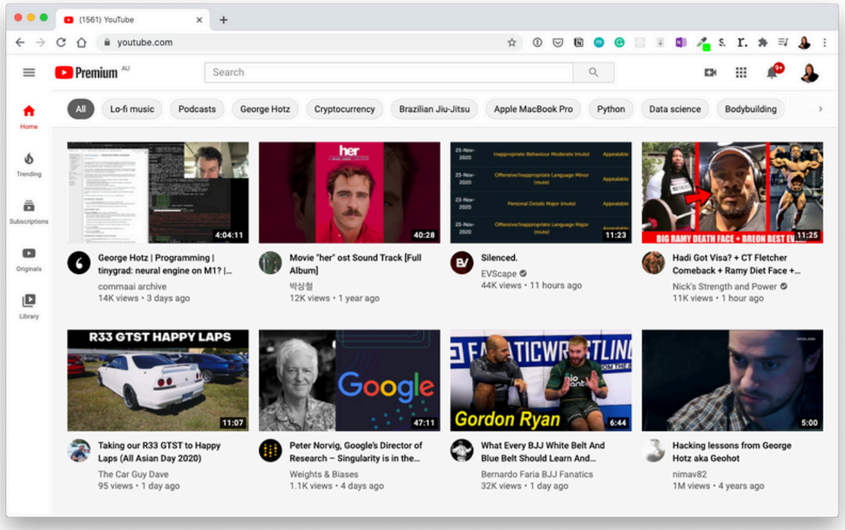
- Neural networks
- Fully connected neural network
- Convolutional neural network
- Recurrent neural network
- Transformer
- …many more

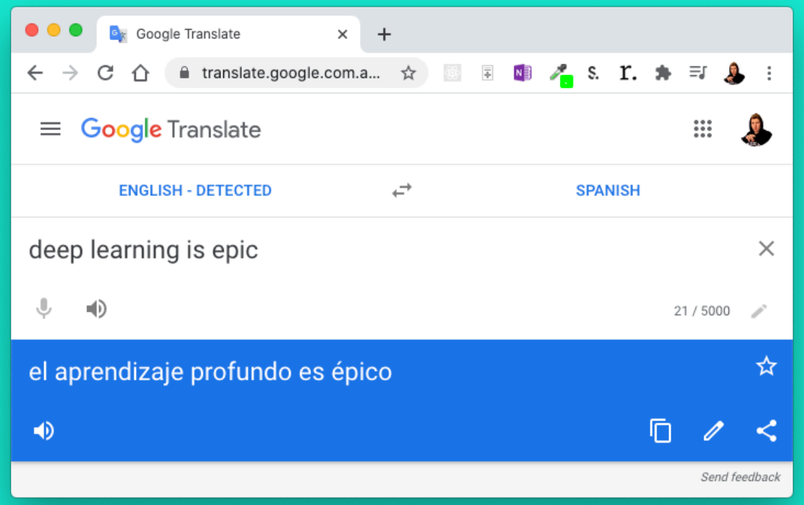
What we're focused on building (with PyTorch)

(depending how you represent your problem, many algorithms can be used for both)

Structured data ⟷ **Unstructured data**

Deep Learning Use Cases





Translation

“Hey Siri, who’s the bigge
big dog of them all?”

Speech recognition
Sequence to sequence
(seq2seq)

Computer Vision

To: daniel@mrdbourke.com
Hey Daniel,

This deep learning course is incredible! I can’t wait to use what I’ve learned!

Not spam

To: daniel@mrdbourke.com
Hay daniel…

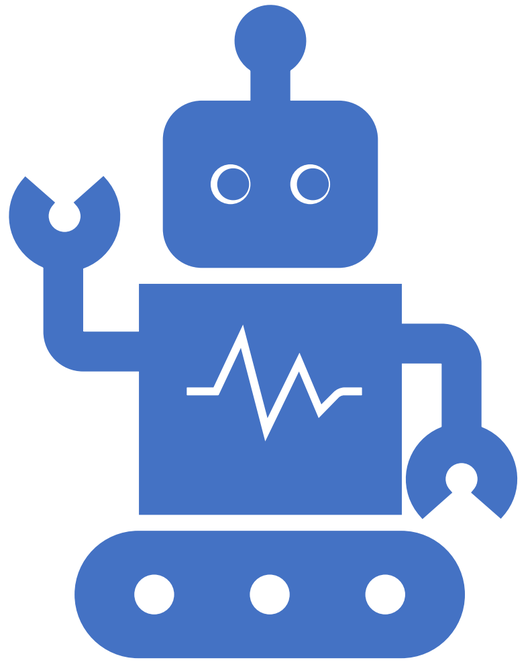
C0ongratu1ations! U win $1139239230

Spam

Natural Language Processing (NLP)

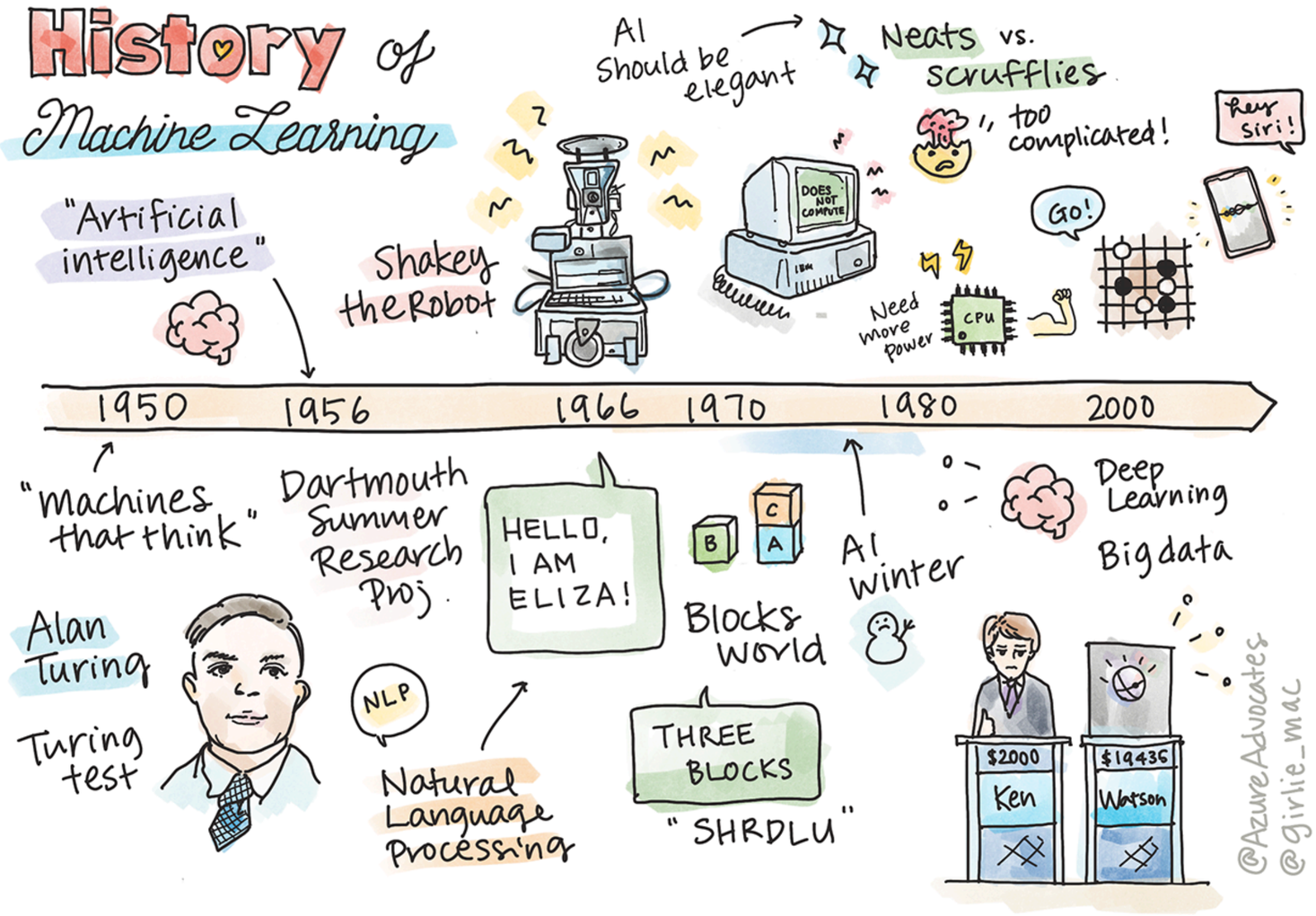
Example of Application of machine learning

- To predict the likelihood of disease from a patient's medical history or reports.
- To leverage weather data to predict weather events.
- To understand the sentiment of a text.
- To detect fake news to stop the spread of propaganda.



Finance, economics, earth science, space exploration, biomedical engineering, cognitive science, and even fields in the humanities have adapted machine learning to solve the arduous, data-processing heavy problems of their domain.

History of learn



```
Welcome to
                  EEEEEE  LL      IIII   ZZZZZZ   AAAAA
                  EE      LL       II       ZZ   AA   AA
                  EEEEE   LL       II     ZZZ    AAAAAAA
                  EE      LL       II    ZZ      AA   AA
                  EEEEEE  LLLLLL  IIII ZZZZZZ    AA   AA

  Eliza is a mock Rogerian psychotherapist.
  The original program was described by Joseph Weizenbaum in 1966.
  This implementation by Norbert Landsteiner 2005.


ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

Eliza, an early 'chatterbot', could converse with people and act as a primitive 'therapist'. You'll learn more about Eliza in the NLP lessons.

Notable discoveries

1950s:

- - 1950: Alan Turing proposes the "Turing Test" as a measure of machine intelligence.
- - 1956: John McCarthy organizes the Dartmouth Workshop, marking the birth of artificial intelligence (AI) as a field of study.

1960s:

- - 1967: The concept of nearest neighbor algorithms is introduced by Evelyn Fix and Joseph Hodges.
- - 1969: The "Perceptron" algorithm, a type of artificial neural network, is developed by Frank Rosenblatt.

1970s:

- - 1970: The "ID3" algorithm is introduced by J.R. Quinlan, becoming one of the first decision tree learning algorithms.
- - 1974: Paul Werbos introduces the concept of backpropagation, a crucial technique for training artificial neural networks.

1980s:

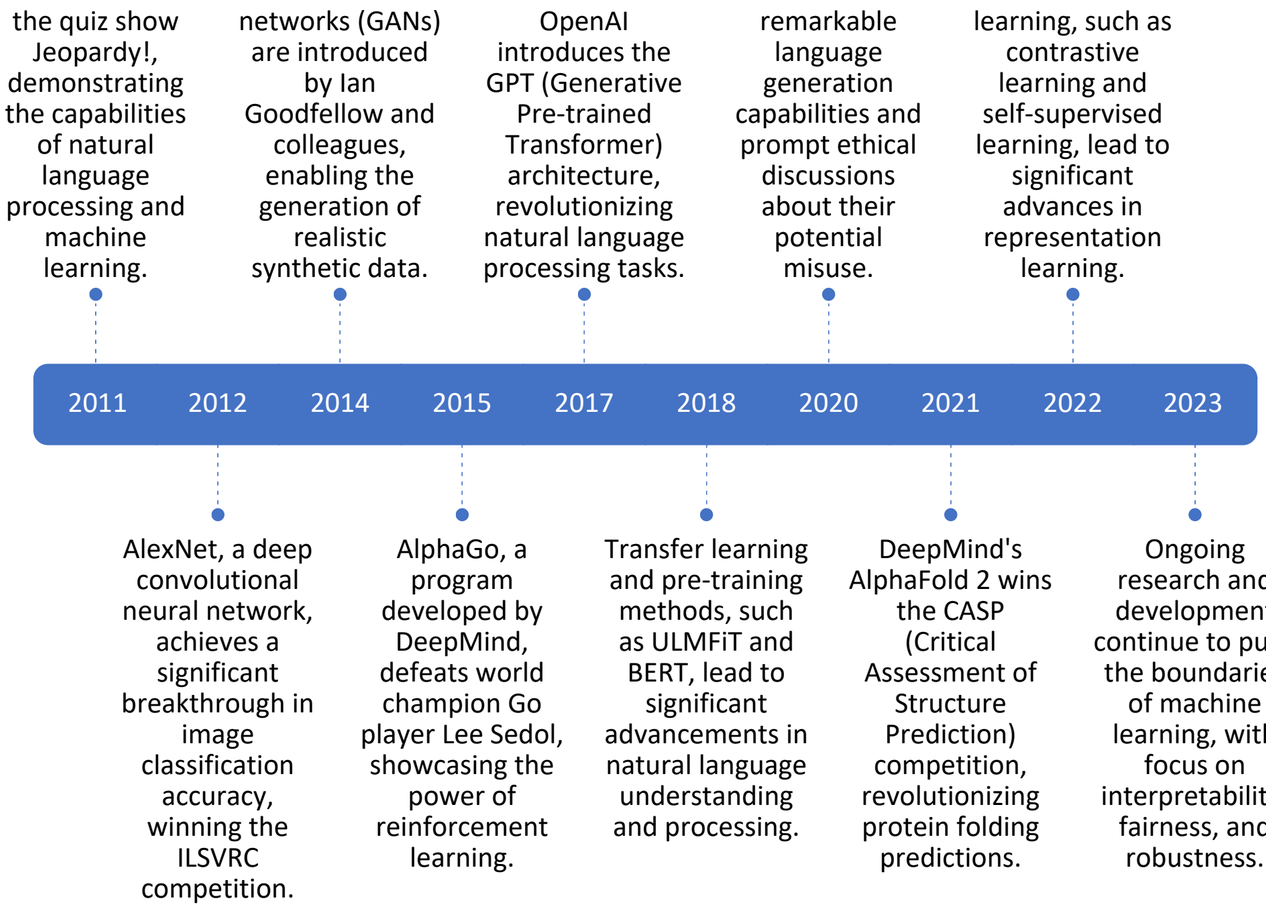
- - 1981: Gerald DeJong develops the "Explanation-Based Learning" approach, enabling systems to learn from explicit explanations.
- - 1986: Geoffrey Hinton, David Rumelhart, and Ronald Williams demonstrate the practical effectiveness of backpropagation, leading to a resurgence of interest in neural networks.

1990s:

- - 1995: Vladimir Vapnik and Alexey Chervonenkis propose the Support Vector Machines (SVM) algorithm for classification tasks.
- - 1997: IBM's Deep Blue defeats world chess champion Garry Kasparov, showcasing the power of machine learning in complex games.

2000s:

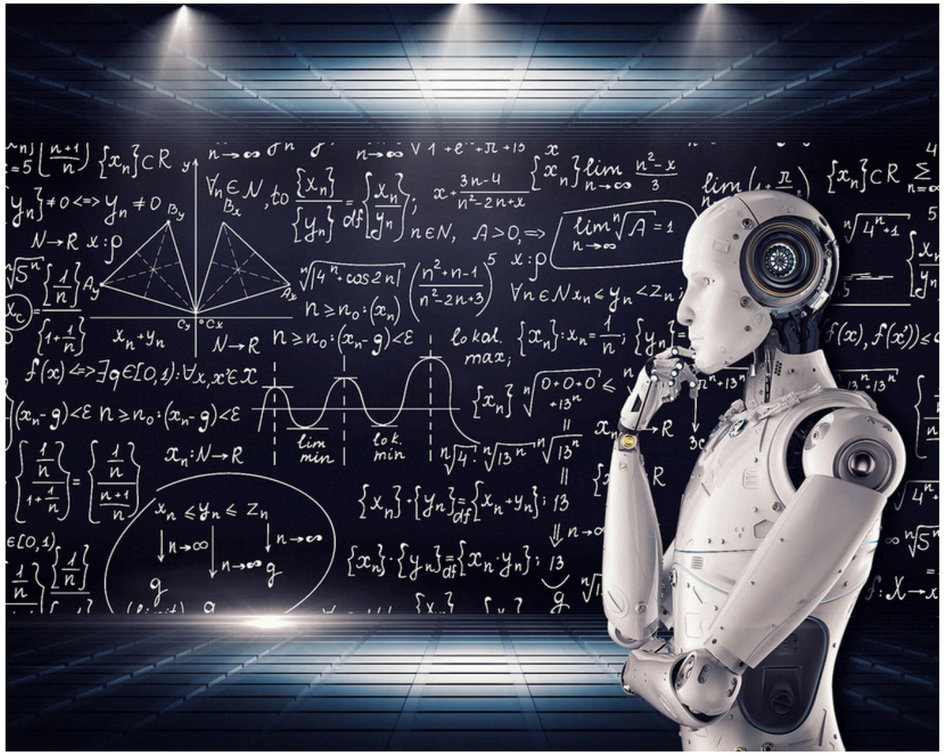
- - 2001: The "AdaBoost" algorithm, which combines multiple weak classifiers to form a strong classifier, is introduced by Yoav Freund and Robert Schapire.
- - 2006: Geoffrey Hinton and colleagues publish a seminal paper on deep belief networks, paving the way for modern deep learning architectures.
- - 2009: The "ImageNet" dataset and the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) are introduced, spurring advancements in computer vision.

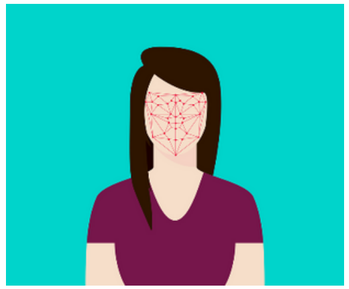

2011
the quiz show Jeopardy!, demonstrating the capabilities of natural language processing and machine learning.
2012
AlexNet, a deep convolutional neural network, achieves a significant breakthrough in image classification accuracy, winning the ILSVRC competition.
2014
networks (GANs) are introduced by Ian Goodfellow and colleagues, enabling the generation of realistic synthetic data.
2015
AlphaGo, a program developed by DeepMind, defeats world champion Go player Lee Sedol, showcasing the power of reinforcement learning.
2017
OpenAI introduces the GPT (Generative Pre-trained Transformer) architecture, revolutionizing natural language processing tasks.
2018
Transfer learning and pre-training methods, such as ULMFiT and BERT, lead to significant advancements in natural language understanding and processing.
2020
remarkable language generation capabilities and prompt ethical discussions about their potential misuse.
2021
DeepMind's AlphaFold 2 wins the CASP (Critical Assessment of Structure Prediction) competition, revolutionizing protein folding predictions.
2022
learning, such as contrastive learning and self-supervised learning, lead to significant advances in representation learning.
2023
Ongoing research and development continue to push the boundaries of machine learning, with focus on interpretability, fairness, and robustness.

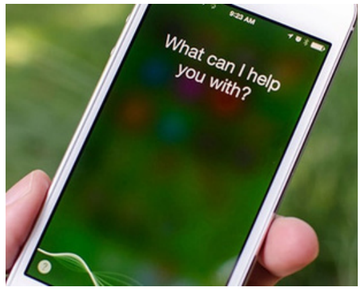
MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE: INTRODUCTION

STEMPLICITY
SCIENCE TECHNOLOGY ENGINEERING & MATH MADE SIMPLE

- Artificial Intelligence/Machine learning does not only mean robots or Sci-Fi movies!
- Machine learning applications are everywhere!
- Google search engine, amazon recommender systems, Facebook facial recognition (tagging), Siri







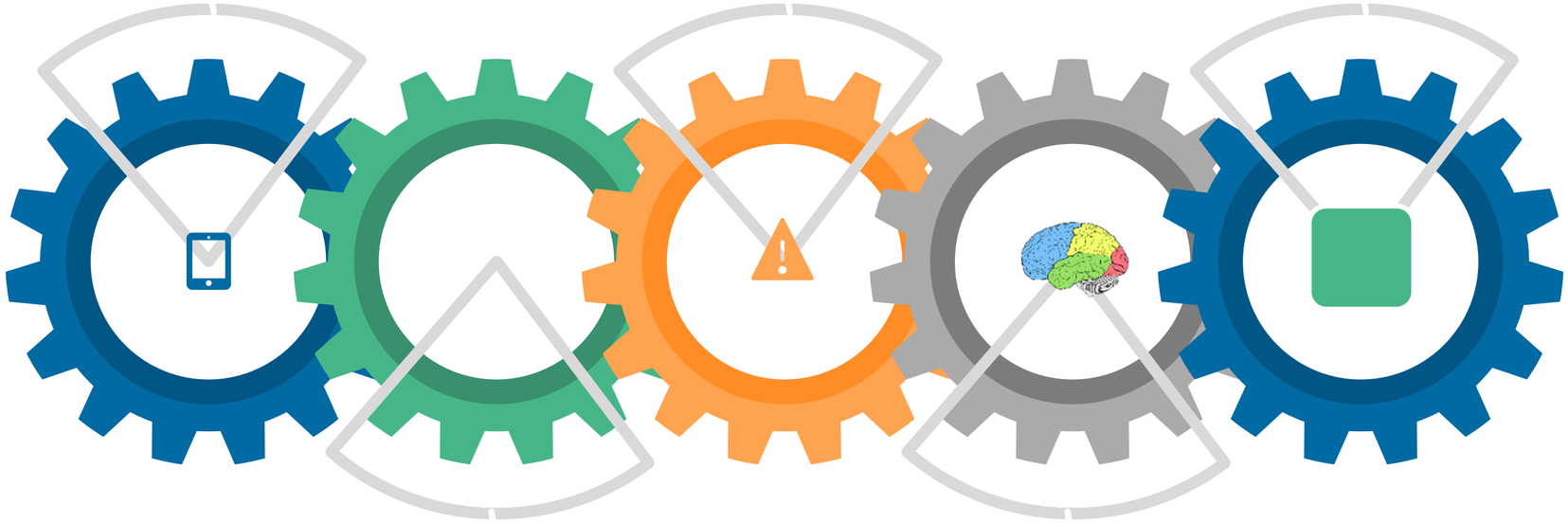
MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE: PROCESS

- Machine learning is the study of algorithms that teach computers to learn from experience.
- Through experience (more training data), computers can continuously improve their performance.

STEP 1
Feed training data

STEP 3
Find patterns

STEP 5
Learn from experience



STEP 2
Analyze Data

STEP 4
Perform prediction/Make a decision

MACHINE LEARNING: BIG PICTURE

STEMPLICITY
SCIENCE TECHNOLOGY ENGINEERING & MATH MADE SIMPLE

ARTIFICIAL INTELLIGENCE
Science that enables computers to mimic human intelligence. Subfields: Machine Learning, robotics, and computer vision

MACHINE LEARNING
Subset of AI that enable machines to improve at tasks with experience

SUPERVISED LEARNING
Training algorithms using labeled input/output data.

- CLASSIFICATION
- REGRESSION

UNSUPERVISED LEARNING
Training algorithms with no labeled data. It attempts at discovering hidden patterns on its own.

- CLUSTERING

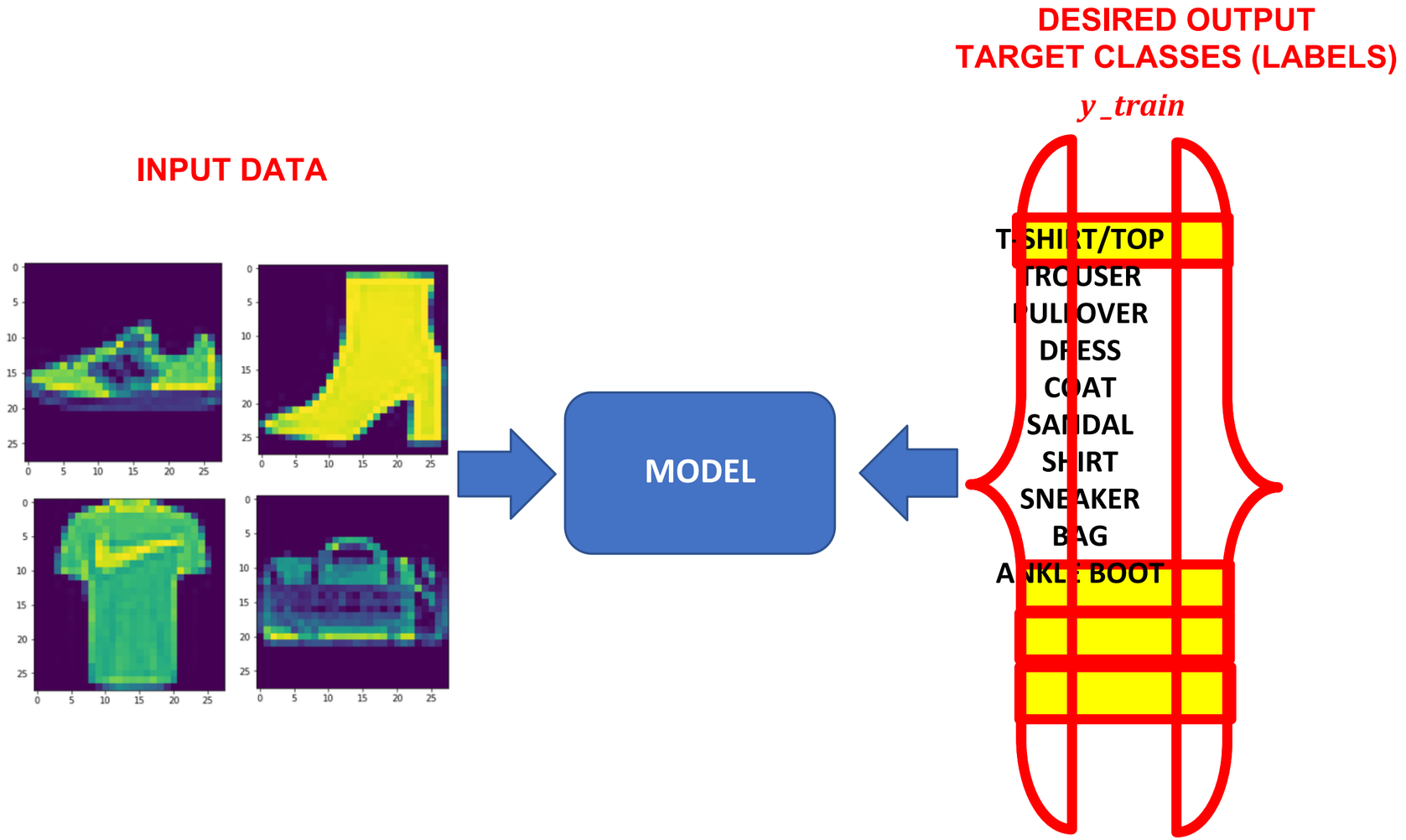
REINFORCEMENT LEARNING
Algorithm take actions to maximize cumulative reward.

MACHINE LEARNING: WHICH TECHNIQUE SHOULD

Classification	Regression	Clustering
○ **Support Vector Machines** ○ **Naïve Bayes** ○ **Random forest** ○ **K Nearest Neighbours** ○ **Logistic regression**	○ Predicting continuous values such as temperature ○ **Simple Linear Regression** ○ **Multiple Linear Regression** ○ **Polynomial Regression**	○ No labelled data ○ Finding Patterns in data is required ○ Market segmentation ○ **K-Means Clustering**

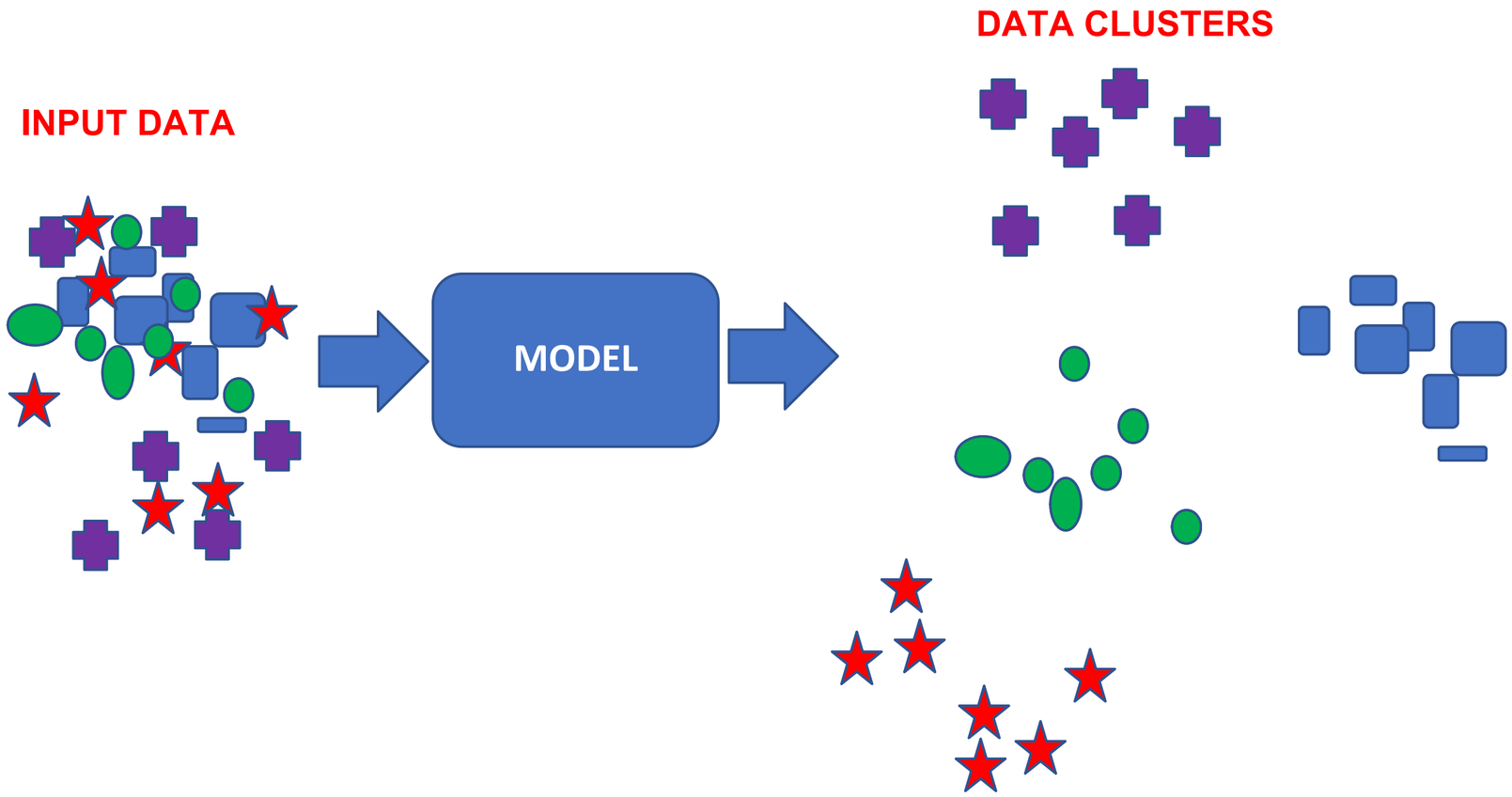
MACHINE LEARNING: SUPERVISED LEARNING

- **Supervised:** used to train algorithms using labeled input and output data.
- Performance is assessed by comparing trained model prediction vs. real output.



MACHINE LEARNING: UNSUPERVISED LEARNING

- **Unsupervised learning:** provides the algorithm with no labeled data.
- The algorithm attempts at discovering hidden patterns within the training data.
- Unsupervised learning methods can analyze complex data that humans might find difficult to interpret.
- No feedback!

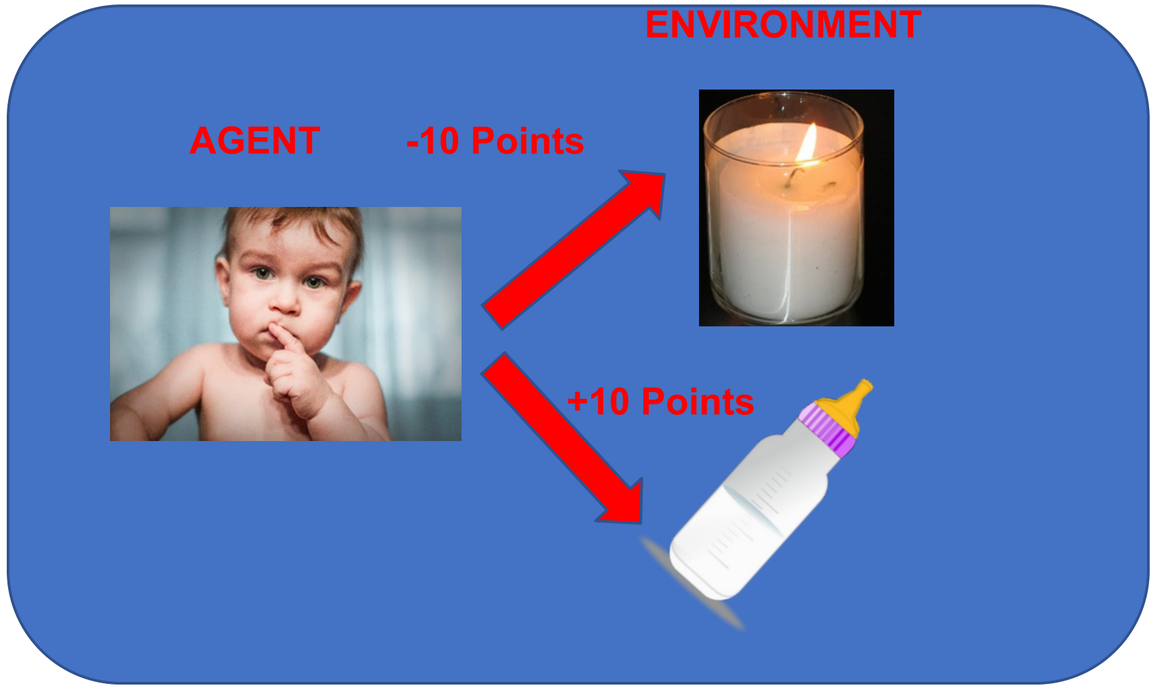


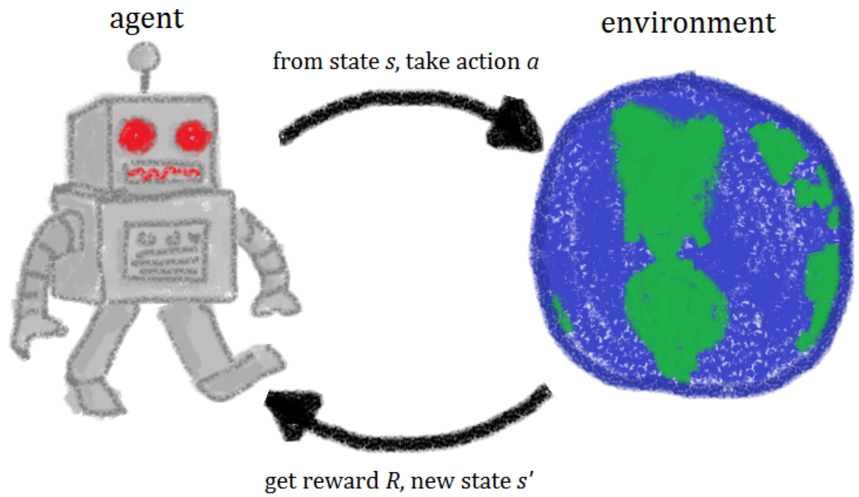
Machine learning

Linear Regression: A linear regression algorithm fits a linear equation to the data by minimizing the difference between the predicted and actual values. It is used for regression tasks where the target variable is continuous.
Logistic Regression: Logistic regression is used for binary classification problems. It estimates the probability of an instance belonging to a particular class by fitting a logistic function to the data.
Decision Trees: Decision trees create a flowchart-like structure where each internal node represents a decision based on a feature, and each leaf node represents a class label or a numerical value. They are versatile and can be used for both classification and regression tasks.
Random Forests: Random forests are an ensemble learning method that combines multiple decision trees. Each tree is built on a random subset of features, and the final prediction is made based on the majority vote or average of the predictions from individual trees.
Support Vector Machines (SVM): SVM is a powerful algorithm used for classification tasks. It finds a hyperplane that maximally separates the data into different classes. SVM can handle linear and non-linear separable data by using different kernel functions.
Naive Bayes: Naive Bayes is a probabilistic algorithm based on Bayes' theorem. It assumes that features are conditionally independent given the class label and uses this assumption to calculate the probability of an instance belonging to a particular class.

MACHINE LEARNING: REINFORCEMENT LEARNING

- Reinforcement learning allows machines take actions to maximize cumulative reward.
- Reinforcement algorithms learn by trial and error through reward and penalty.
- Two elements: **environment** and **learning agent**.
- The environment rewards the agent for correct actions.
- Based on the reward or penalty, agent improves its environment knowledge to make better decision.

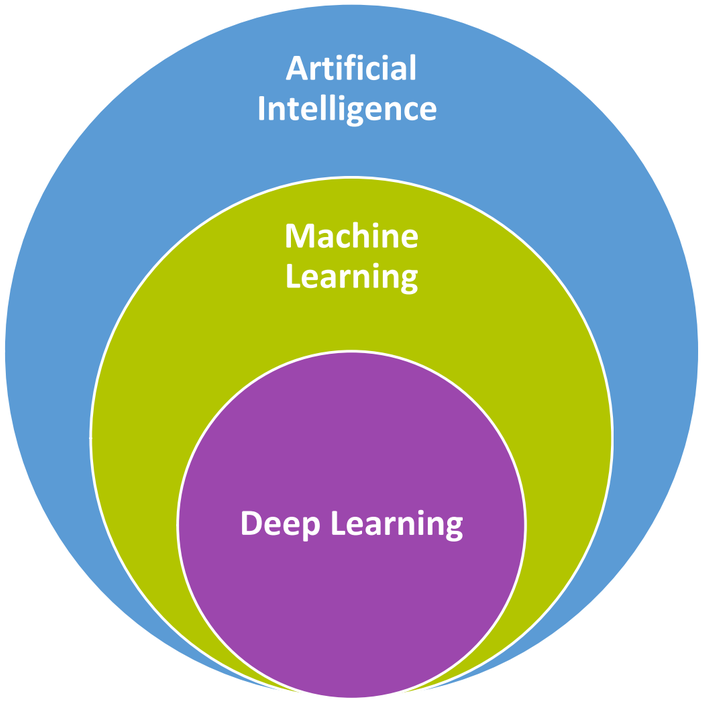




https://commons.wikimedia.org/wiki/File:Rl_agent.png

MACHINE LEARNING: DEEP LEARNING

- Deep learning is a subset of machine learning that utilizes multi-layer Artificial Neural Networks.
- Deep Neural Networks are inspired by the human brain and mimics the operation of biological neurons.



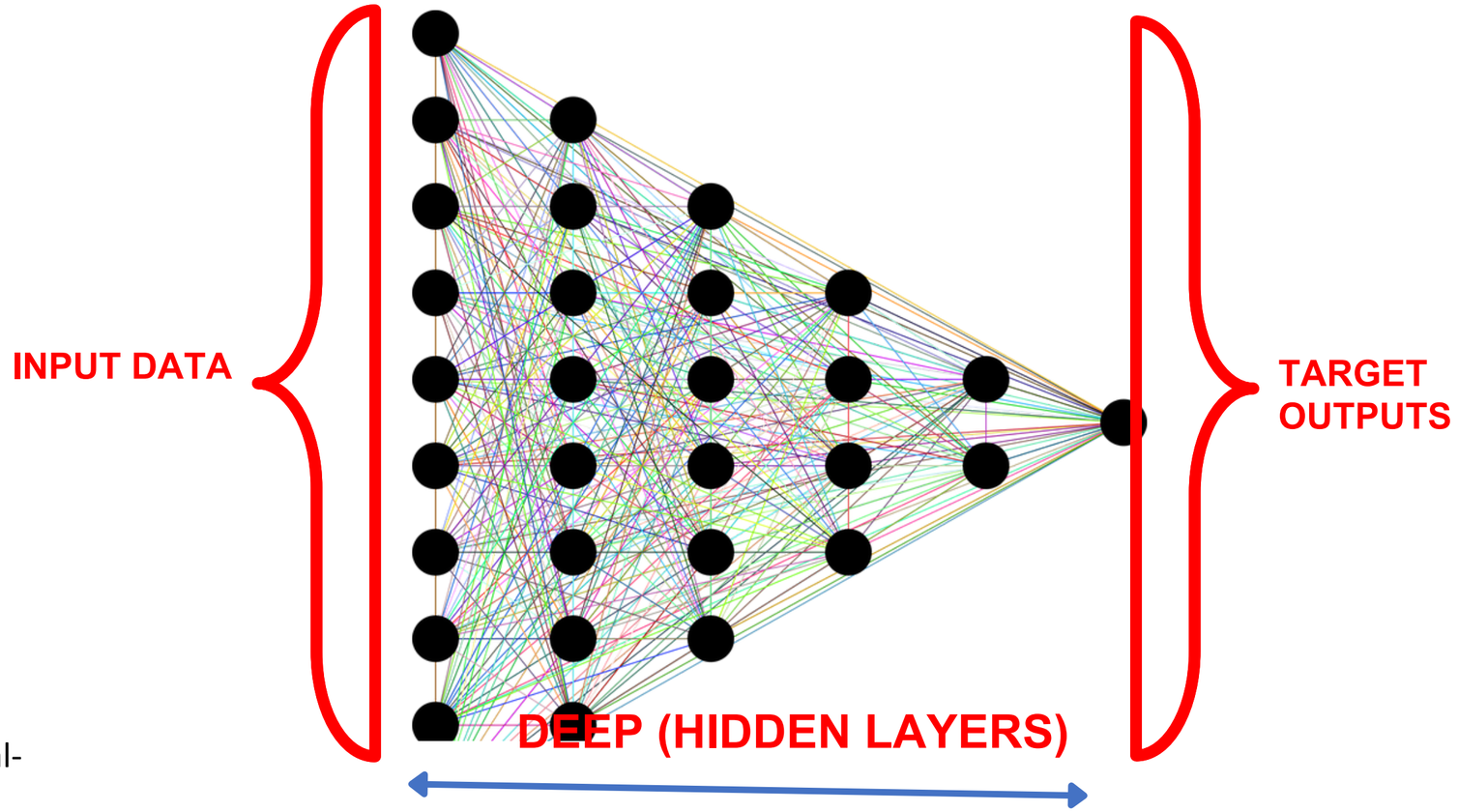


Photo Credit: https://pixabay.com/en/neural-network-thought-mind-mental-3816319/