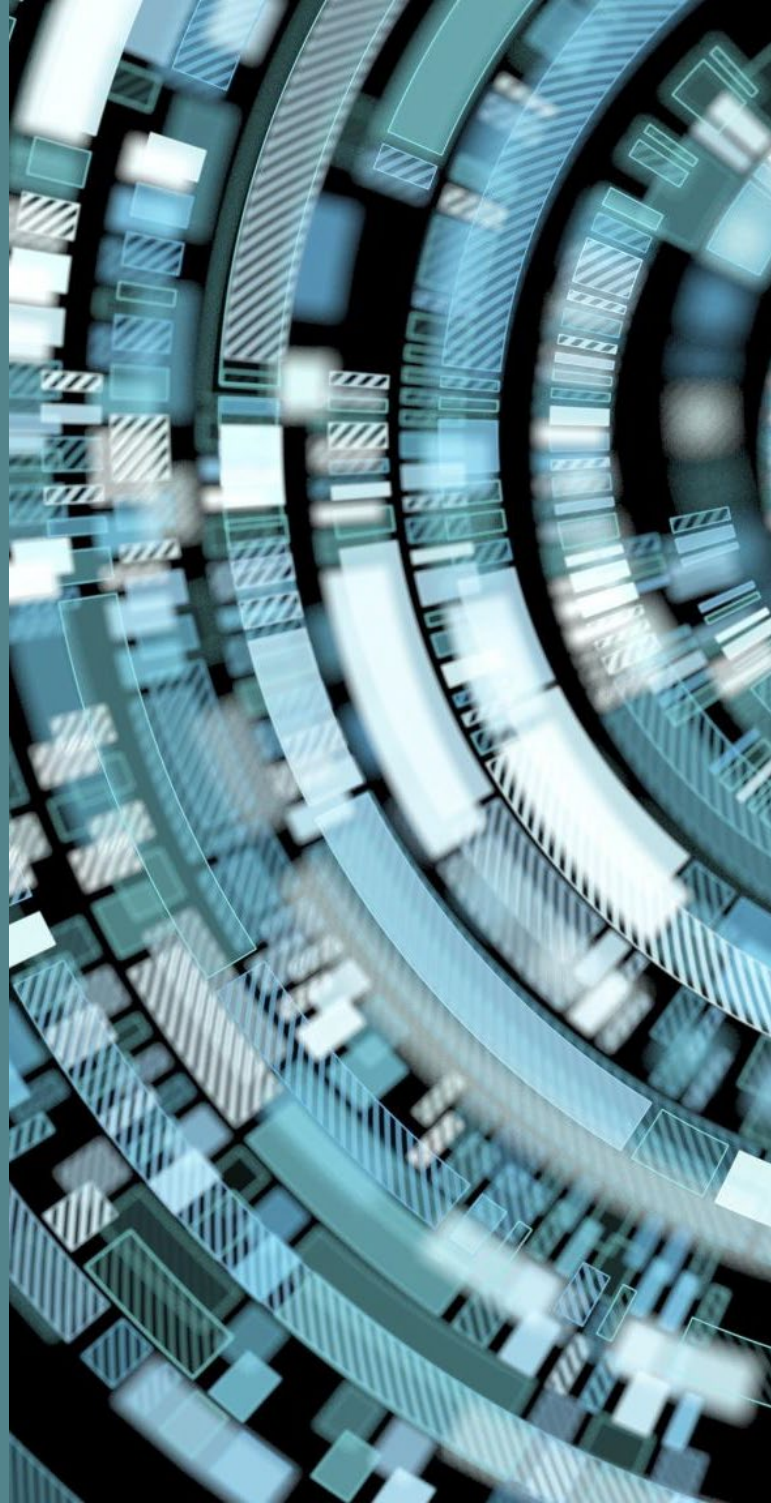
ELEC8900-57 SPECIAL TOPICS: MACHINE LEARNING

FALL - 2023

COURSE INFORMATION



INTRODUCTION

Introduction to machine learning

- Algorithms and principles for supervised learning
  - nearest neighbors, decision trees, ensembles, linear regression, logistic regression, SVMs
- Unsupervised learning: K-means
- Dimensionality Reduction: PCA
- Basics of reinforcement learning
- Deep Learning

COURSE INFORMATION

Course Website:
https://brightspace.uwindsor.ca/d2l/home

Use your Uwindsor login and password.

Office Hours

No.	Days & Times*	Room
1	Thursdays: 5:00 – 6:50 PM.	CEI Room – 3069

* Other times by appointment.

INSTRUCTOR: **DR. YASSER M. ALGINAHI**

EMAIL: ALGINAHI@UWINDSOR.CA

GRADUATE ASSISTANTS CONTACT INFORMATION

No.	Name	Office	Office Hours (& by appointment)	Email
1.	Ali Abedi	CEI 3082	Fridays: 2:00 – 4:00 P.M.	abedi3@uwindsor.ca
2.	Jiayuan Wang	CEI 3082	Wednesdays: 2:00 – 4:00 P.M.	wang621@uwindsor.ca
3.	Niloofar Naghdi Pour	CEI 3082	Tuesdays: 2:00 – 4:00 P.M.	naghdip@uwindsor.ca

COURSE DESCRIPTION

This course introduces machine learning, covering fundamental concepts, techniques, and algorithms. It explores supervised learning methods including linear regression, logistic regression, multiclass classification, neural networks (CNN, RNN, FNN, deep learning), and decision trees (bias-variance decomposition). The unsupervised learning section covers nearest neighbors, probabilistic models, principles component analysis, K-Means, EM algorithm, and provides an overview of reinforcement learning. Students will apply these techniques to real-world datasets in a programming project, gaining hands-on experience and a solid understanding of their practical applications.

COURSE REQUIREMENTS

Programming knowledge in Python or MATLAB, note that students should be familiar with Machine Learning Libraries in Python or Machine Learning Toolbox in MATLAB. Learning ML tools is the responsibility of the students.

Prerequisites: GENG8010 and GENG8030. In this course students are expected to have some knowledge of linear algebra and differential calculus.

COURSE ASSESSMENT

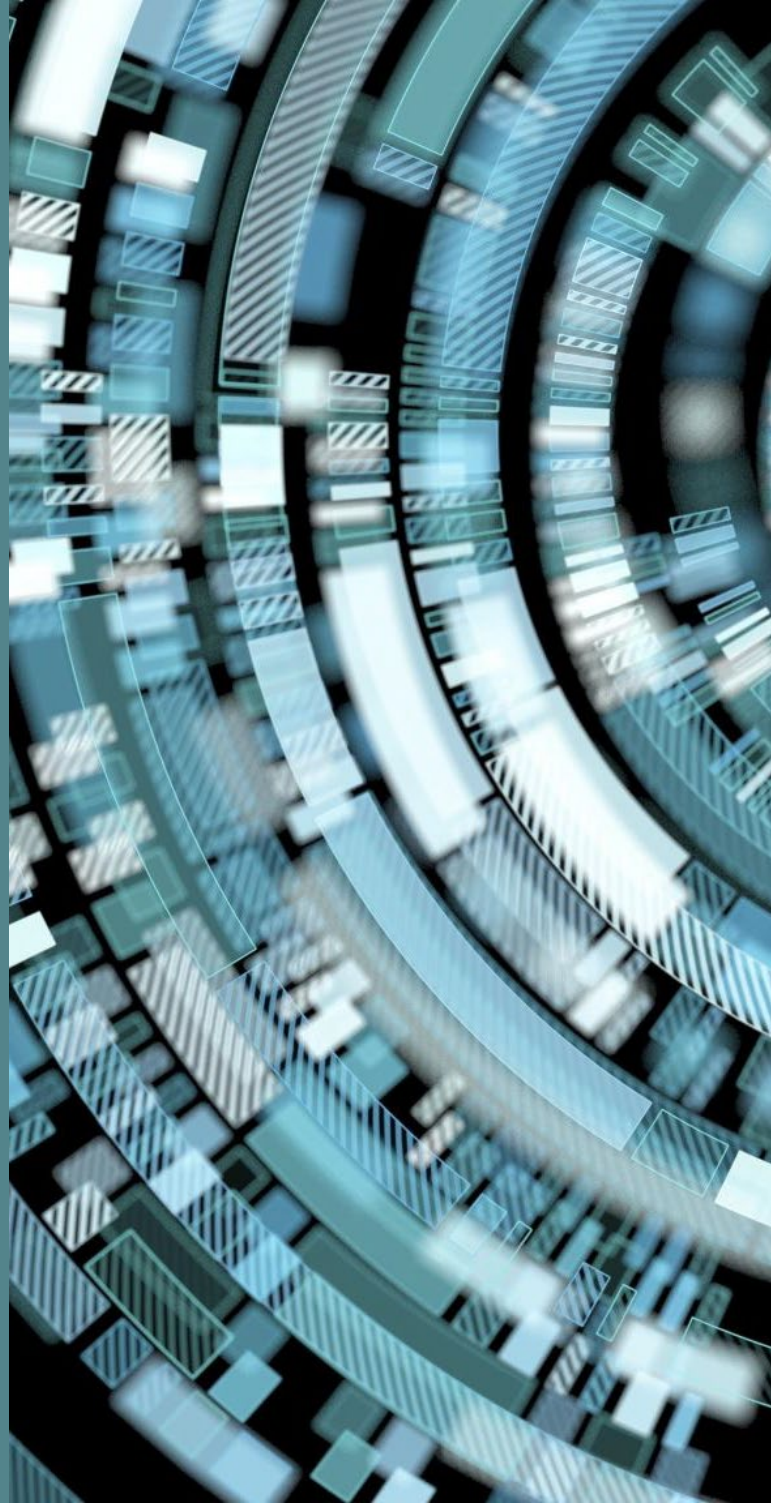
Method of Evaluation	% Of Final Grade	Due Dates*
Online assignments through Datacamp.com	20%	As posted in the datacamp course assignments
In-class Presentation	10%	*Weeks 4 – 11
Course Project	30%	Week 12
Final Exam	40%	(TBA)

*Students are divided into groups of 3 and each week 7 – 8 groups present their work in-class.

COURSE PROJECT

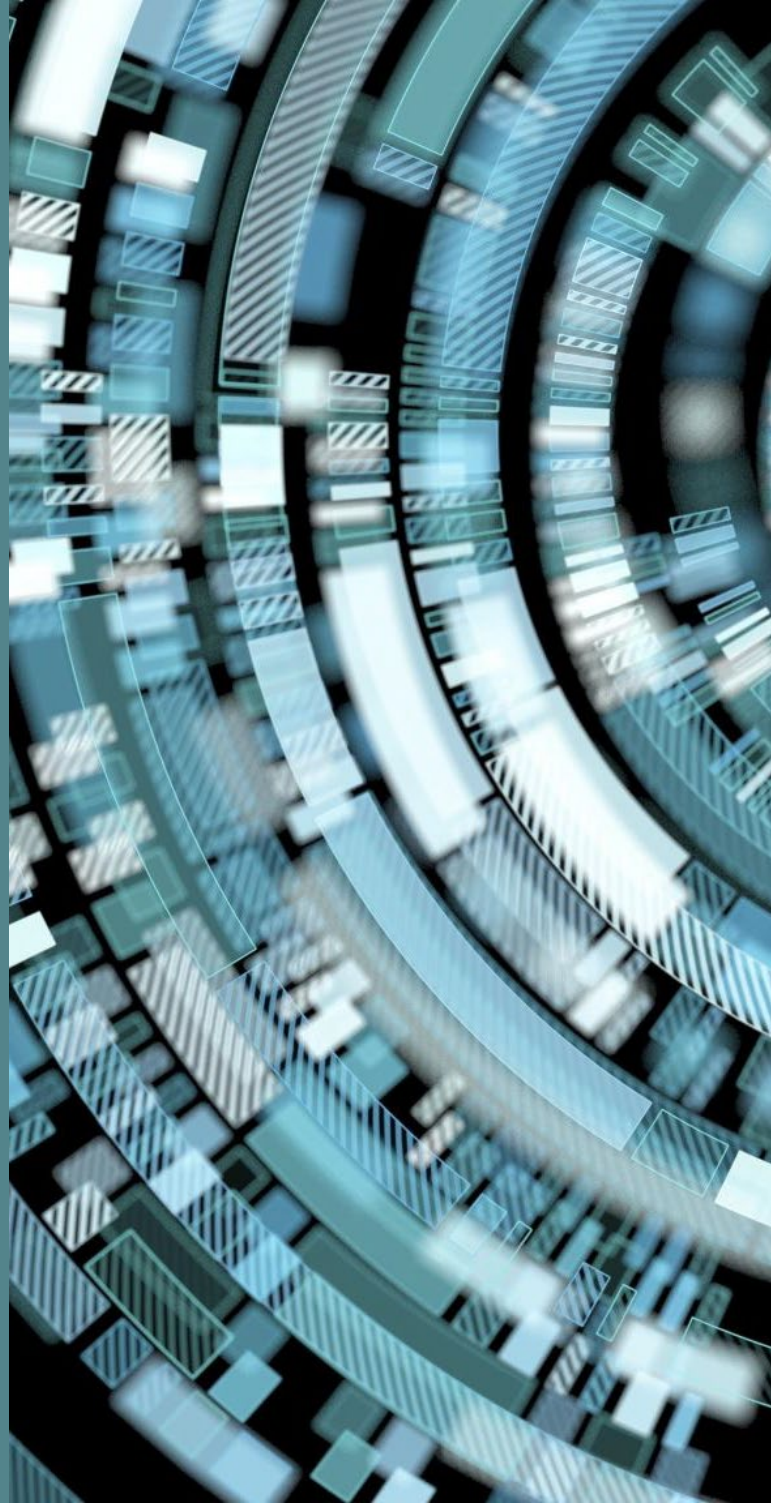
- In a group setting, explore a machine learning problem of your choice using real-world data.
- Apply multiple machine learning techniques or variations of a single approach to a dataset using programming languages (i.e., Python and/or MATLAB) then compare the results and summarize your findings in a well-structured report.
- Present your project findings in class.
- In this project, you may not use results from previous research or course projects.
- Use the following Google Sheet to record your project title and group member names: http://tinyurl.com/mtjzw59e

INTRODUCTION TO MACHINE LEARNING



Lecture Content

- What is Machine Learning?
- Relations to AI
- History of Machine Learning
- Types of Machine Learning
- Machine Learning Workflow
- Limitations of Machine Learning
- Applications of Machine Learning



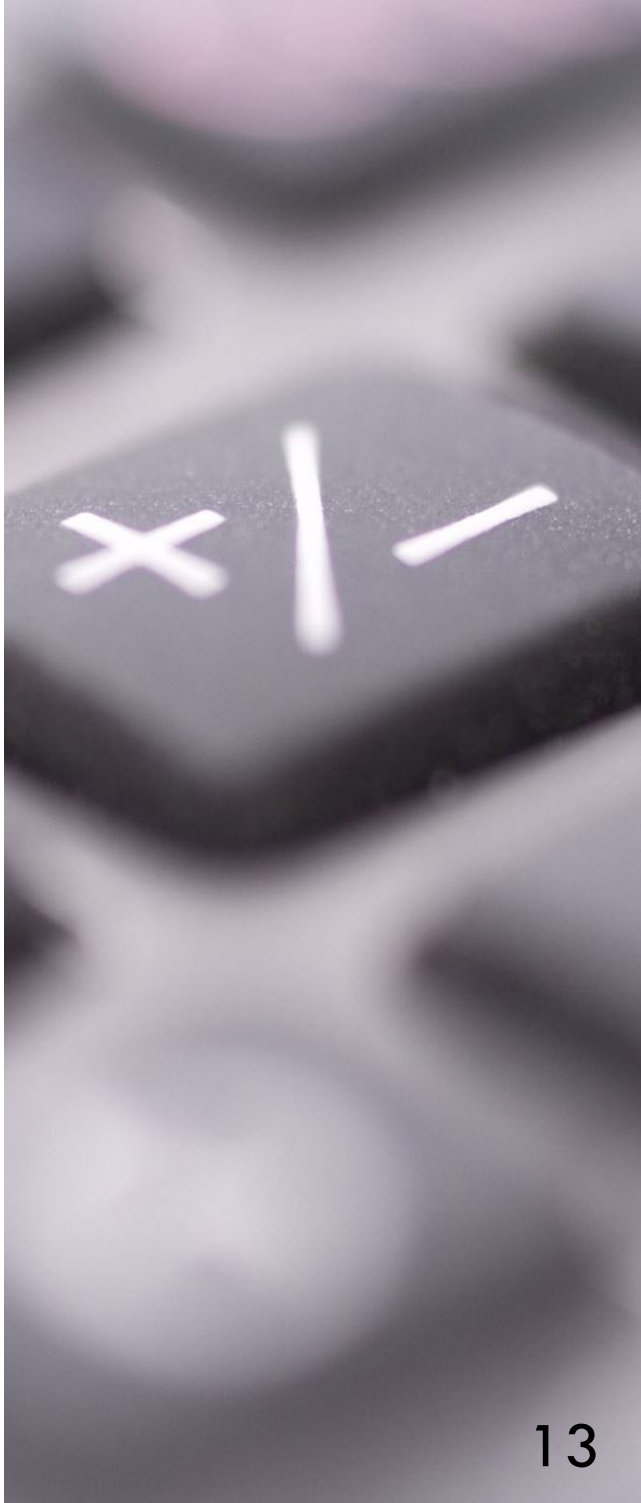
WHAT IS LEARNING?

"The activity or process of gaining knowledge or skill by studying, practicing, being taught, or experiencing something."

Merriam Webster dictionary

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

Tom Mitchell



WHAT IS MACHINE LEARNING?

- Arthur Samuel, an early American leader in the field of computer gaming and artificial intelligence, coined the term “Machine Learning ” in 1959 while at IBM.
- Samuel defined machine learning as “the field of study that gives computers the ability to learn without being explicitly programmed“.

WHAT IS MACHINE LEARNING?

- For many problems, it's difficult to manually program the correct behavior
  - recognizing people and objects
  - understanding human speech
  - Etc.
- Machine learning approach: program an algorithm to automatically learn from data, or from experience
- Why might you want to use a learning algorithm?

WHAT IS MACHINE LEARNING?

Examples why might you want to use a learning algorithm?

- hard to code up a solution by hand (e.g. vision, speech)
- system needs to adapt to a changing environment, need customized solution (e.g. spam detection, routing on a computer network, stock market analysis)
- want the system to perform *better* than the human programmers
- Human expertise does not exist (navigating on Mars)
- privacy/fairness (e.g. ranking search results)

WHAT IS MACHINE LEARNING?

It's similar to statistics...

- Both fields try to uncover patterns in data
- Both fields draw heavily on calculus, probability, and linear algebra, and share many of the same core algorithms

But it's not statistics!

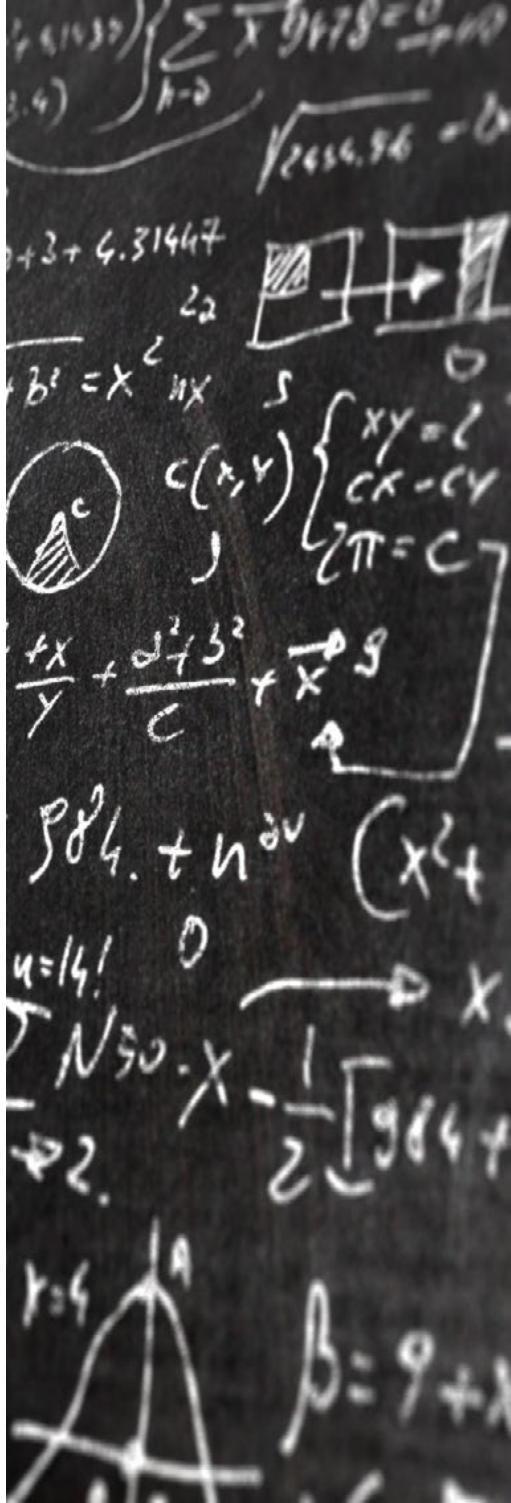
- Stats is more concerned with helping scientists and policymakers draw good conclusions; ML is more concerned with building autonomous agents
- Stats puts more emphasis on interpretability and mathematical rigor; ML puts more emphasis on predictive performance, scalability, and autonomy. ML puts more emphasis on autonomy by enabling systems to operate with minimal human intervention, and on scalability by ensuring that systems can maintain performance and efficiency as data and computational requirements grow

WHAT WE TALK ABOUT WHEN WE TALK ABOUT "LEARNİNG"

- Learning general models from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:
  - *People who bought "Da Vinci Code" also bought "The Five People You Meet in Heaven" (www.amazon.com)*
- Build a model that is *a good and useful approximation* to the data.

WHAT İS MACHİNE LEARNİNG?

- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
  - Solve the optimization problem
  - Representing and evaluating the model for inference



WHERE DO WE USE MACHINE LEARNING?

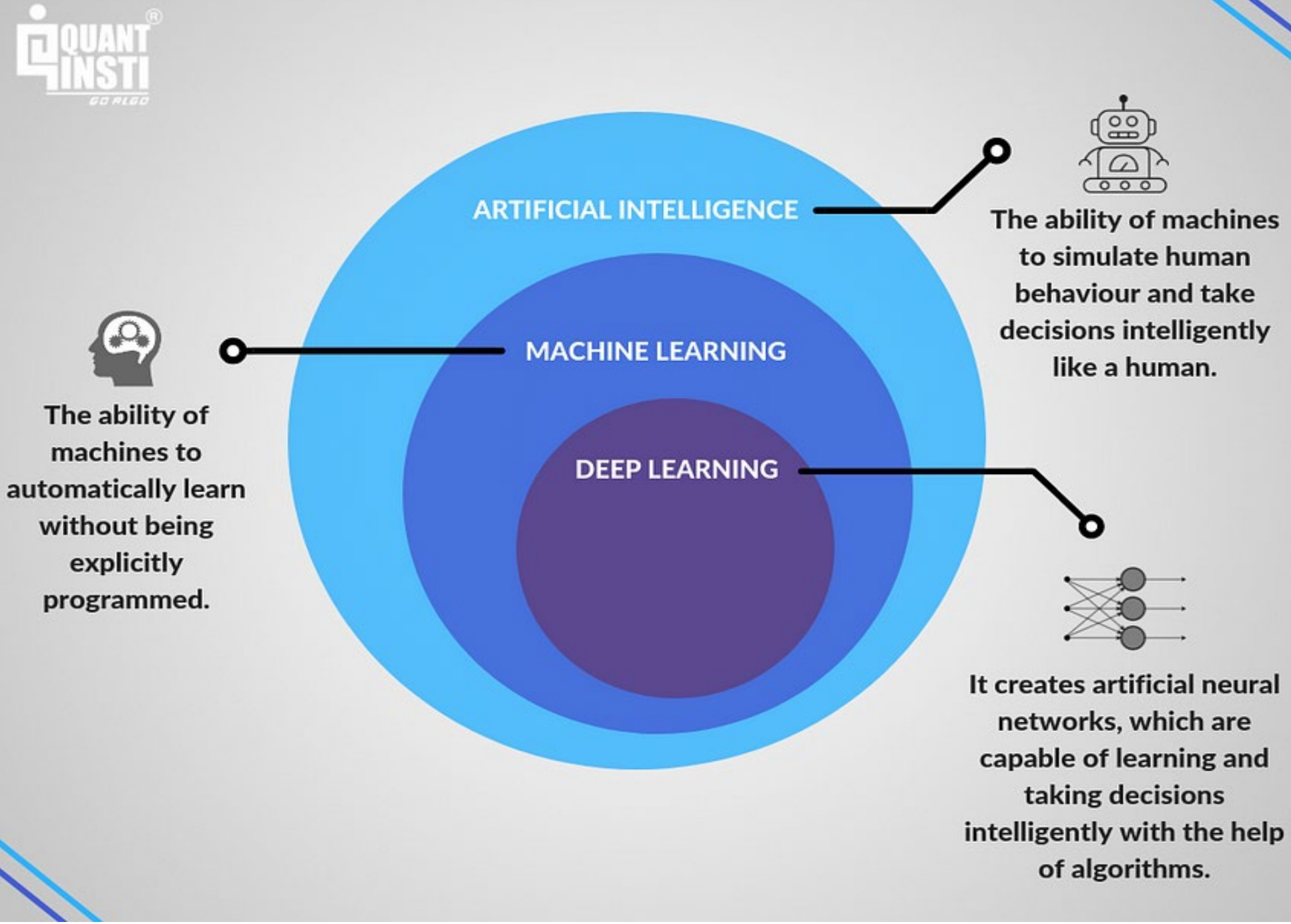
- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Optimization, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Quality of service optimization
- Bioinformatics: Motifs, alignment
- Web mining: Search engines
- … Etc.

IN SUMMARY: WHAT IS MACHINE LEARNING?

- Machine learning is a type of Artificial Intelligence that allows computers to learn from experience without being explicitly programmed.
- Machine learning algorithms are trained on large amounts of data to identify patterns and make predictions.
- Examples of machine learning in action include Amazon suggesting products based on previous purchases and Netflix recommending movies based on viewing history.

RELATIONS TO AI

- Nowadays, "Machine Learning" is often brought up with "Artificial Intelligence" (AI)
- AI does not always imply a learning-based system
  - Symbolic reasoning (Symbolic AI algorithms are designed to deal with the kind of problems that require human-like reasoning, such as planning, natural language processing, and knowledge representation.)
  - Rule based system
  - Tree search
  - etc.
- Learning based system → learned based on the data → more flexibility, good at solving pattern recognition problems.



Source: https://medium.com/analytics-vidhya/fundamental-omachine-learning-ada28afa1bd3

RELATIONS TO HUMAN LEARNING

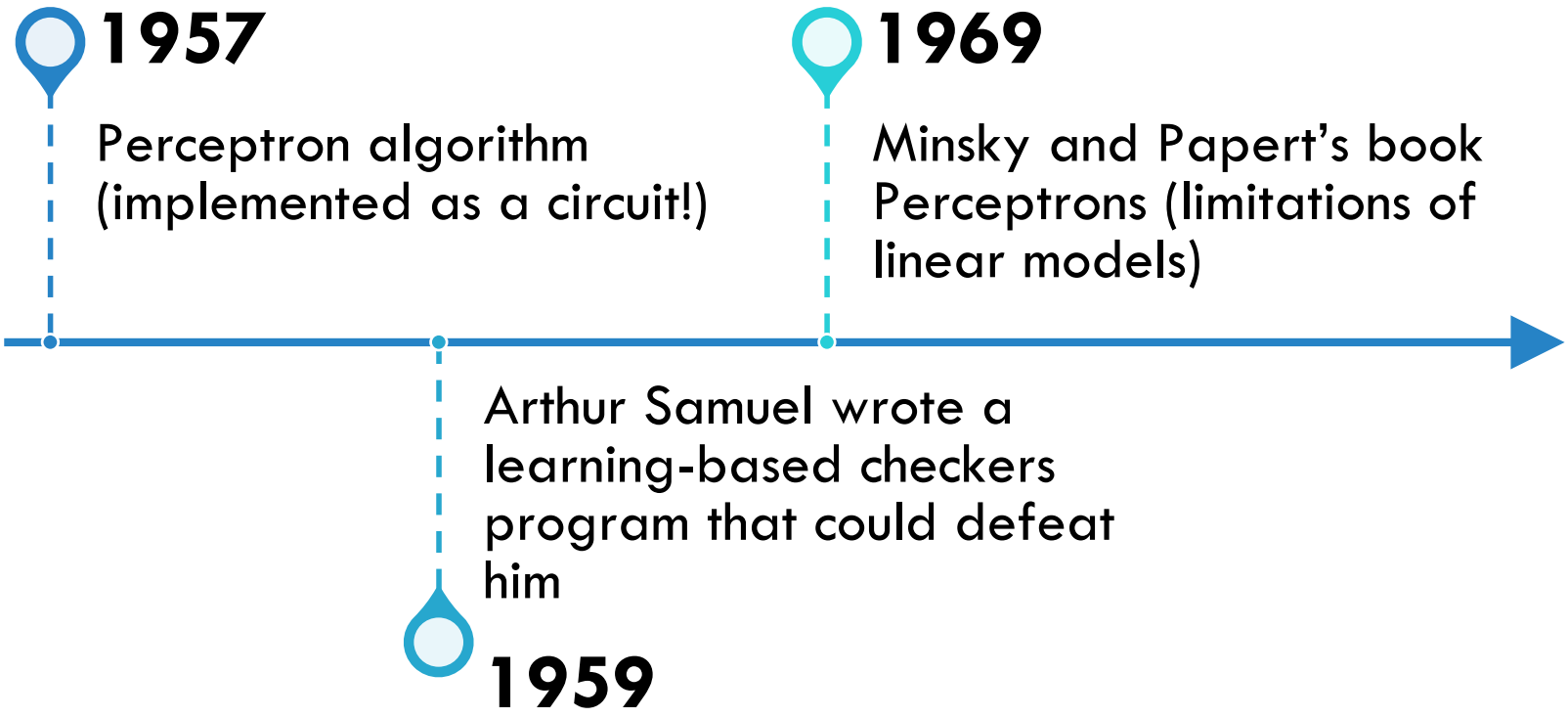
Human learning is:

- Very data efficient
- An entire multitasking system (vision, language, motor control, etc.)
- Takes lots of time

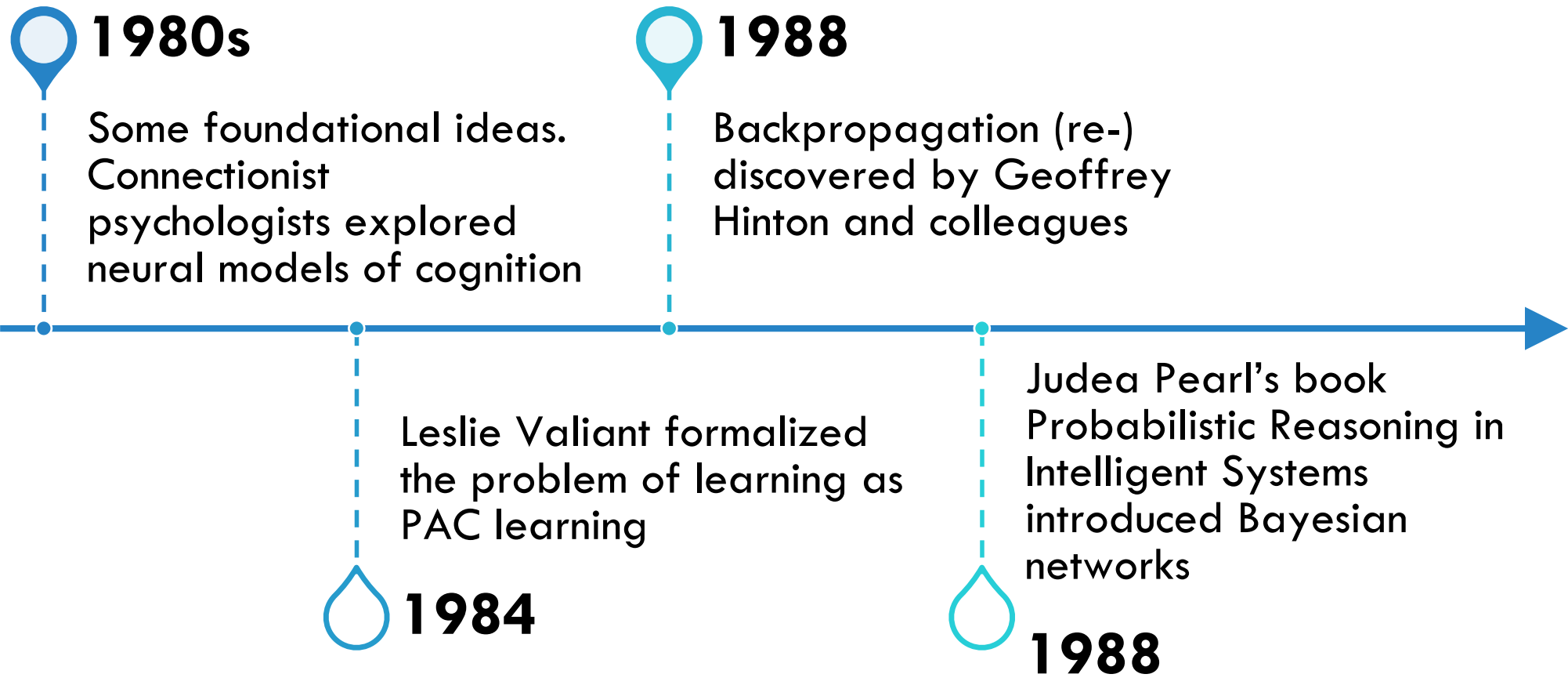
For serving specific purposes, Machine Learning doesn't have to look like human learning in the end.

It may borrow ideas from biological systems, e.g., neural networks. It may perform better or worse than humans.

HISTORY OF MACHINE LEARNING



HISTORY OF MACHINE LEARNING

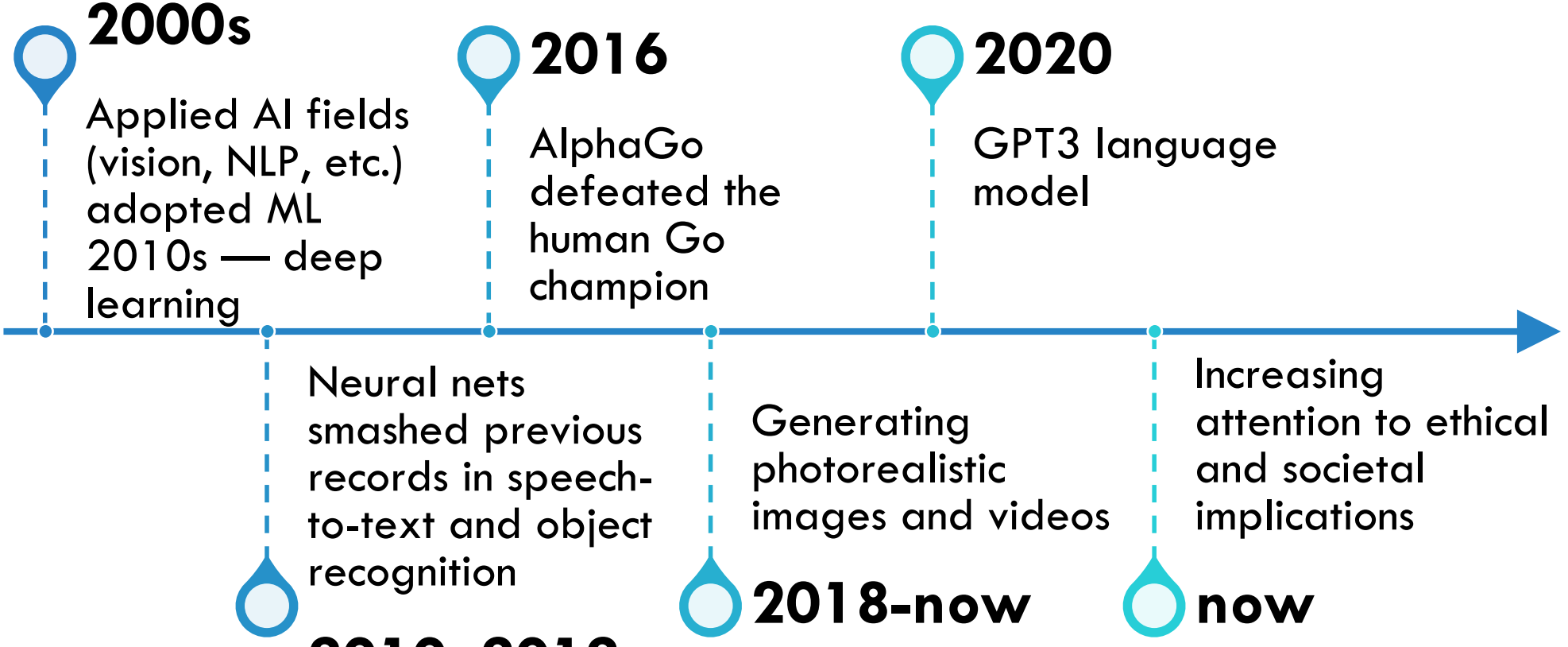


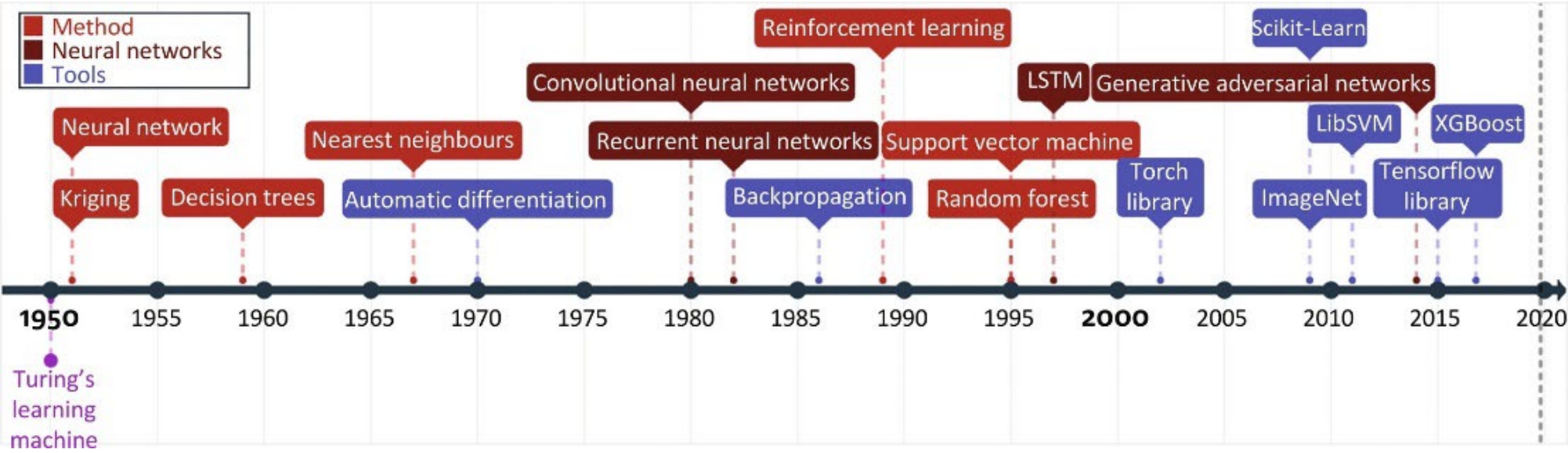
HISTORY OF MACHINE LEARNING

1990s — the "AI Winter", a time of pessimism and low funding But looking back, the '90s were also sort of a golden age for ML research

- Markov chain Monte Carlo
- Variational inference
- Kernels and support vector machines
- Boosting
- Convolutional networks
- Reinforcement learning

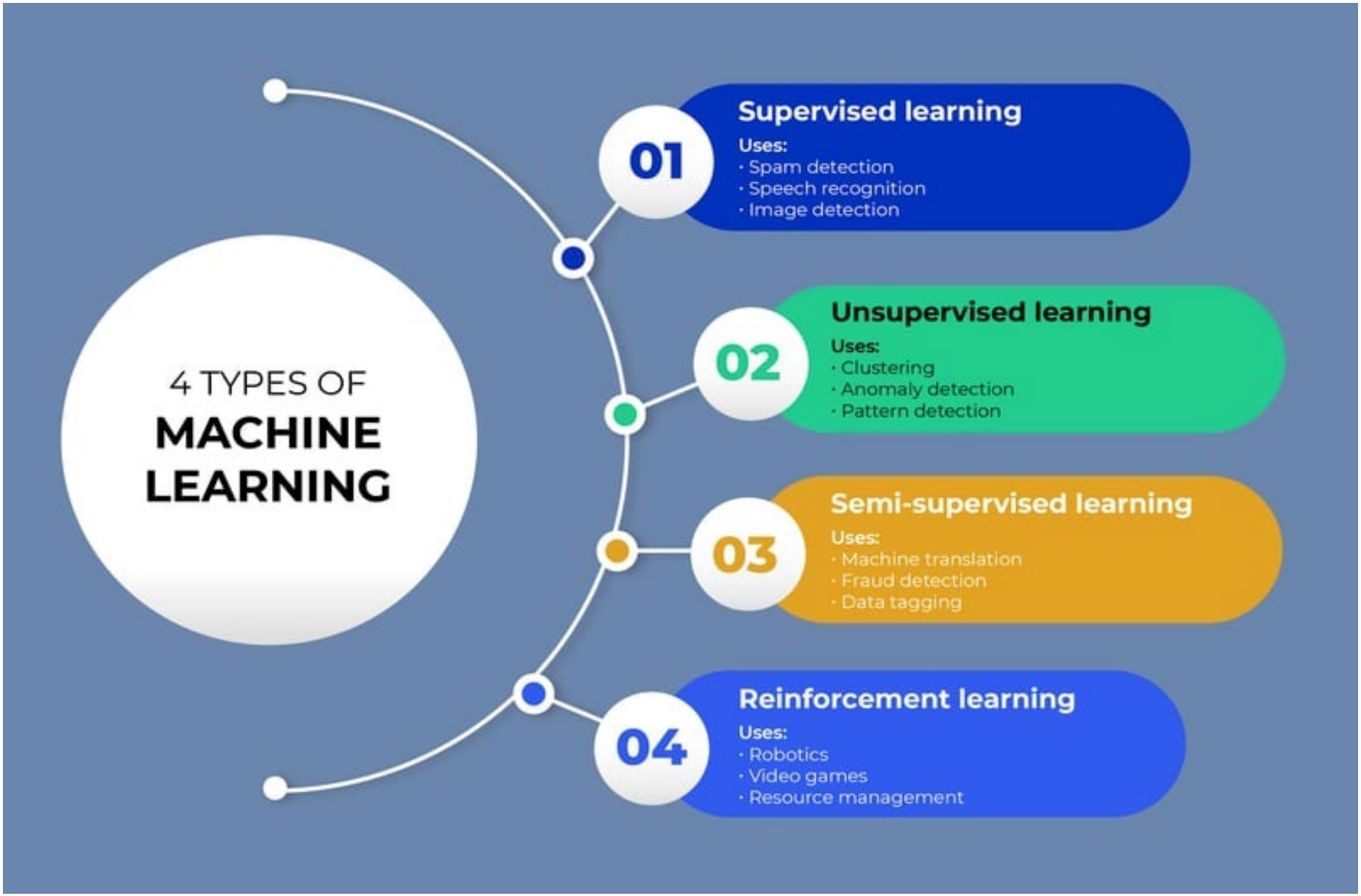
HISTORY OF MACHINE LEARNING





[Source: https://ars.els-cdn.com/content/image/1-s2.0-S0065268720300054-f01-01-9780128216699_lrg.jpg]

TYPES OF MACHINE LEARNING?



[Source: https://www.algotive.ai/blog/machine-learning-what-is-ml-and-how-does-it-work

TYPES OF MACHINE LEARNING?

- **Supervised learning:** have labeled examples of the correct behavior
- **Reinforcement learning:** learning system (agent) interacts with the world and learns to maximize a scalar reward signal
- **Unsupervised learning:** no labeled examples – instead, looking for "interesting" patterns in the data

SUPERVISED LEARNING

- Supervised Learning: In supervised learning, the algorithm learns from labeled data, where the input data is paired with corresponding output labels or target variables.
- The algorithm learns to map inputs to outputs based on the provided examples.
- Common algorithms include linear regression, logistic regression, decision trees, random forests, support vector machines (SVM), and neural networks.

UNSUPERVISED LEARNING

- Unsupervised Learning: Unsupervised learning deals with unlabeled data, where the algorithm learns patterns, structures, or relationships within the data without any predefined labels or target variables.
- It is used to discover inherent structures, clusters, or anomalies in the data.
- Common algorithms include clustering algorithms like K-means, hierarchical clustering, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), as well as dimensionality reduction techniques such as principal component analysis (PCA) and others

REINFORCEMENT LEARNING

- Reinforcement learning involves an agent that interacts with an environment and learns to make optimal decisions or actions to maximize a cumulative reward.
- The agent receives feedback in the form of rewards or penalties based on its actions.
- The algorithm learns through trial and error, optimizing its actions to achieve long-term goals.
- Reinforcement learning is often used in robotics, game playing, and autonomous systems.
- Common algorithms include Q-learning, Deep Q-Networks (DQN), and Policy Gradient methods.

SEMI-SUPERVISED LEARNING

- Semi-supervised Learning: Semi-supervised learning combines elements of supervised and unsupervised learning.
- It uses a small amount of labeled data along with a larger amount of unlabeled data to improve the learning performance.
- It can be useful when labeled data is scarce or expensive to obtain.
- Algorithms in this category often leverage unsupervised learning techniques in combination with supervised learning algorithms

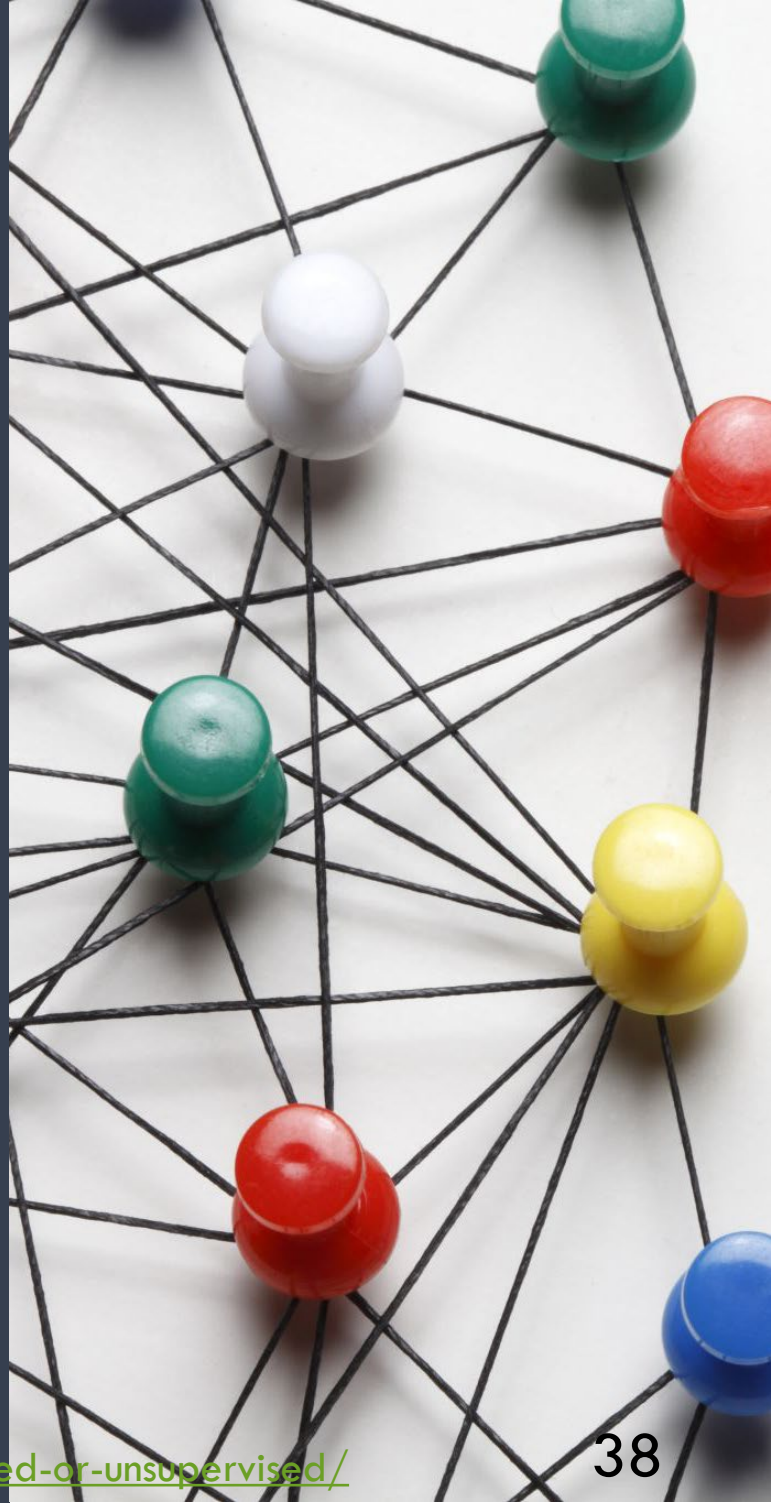
DEEP LEARNING

- Deep Learning: Deep learning is a subset of machine learning that focuses on using artificial neural networks with multiple layers (deep neural networks) to learn and extract hierarchical representations from data.
- Deep learning has achieved remarkable success in various domains such as computer vision, natural language processing, and speech recognition.
- Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are commonly used in deep learning

ENSEMBLE LEARNING

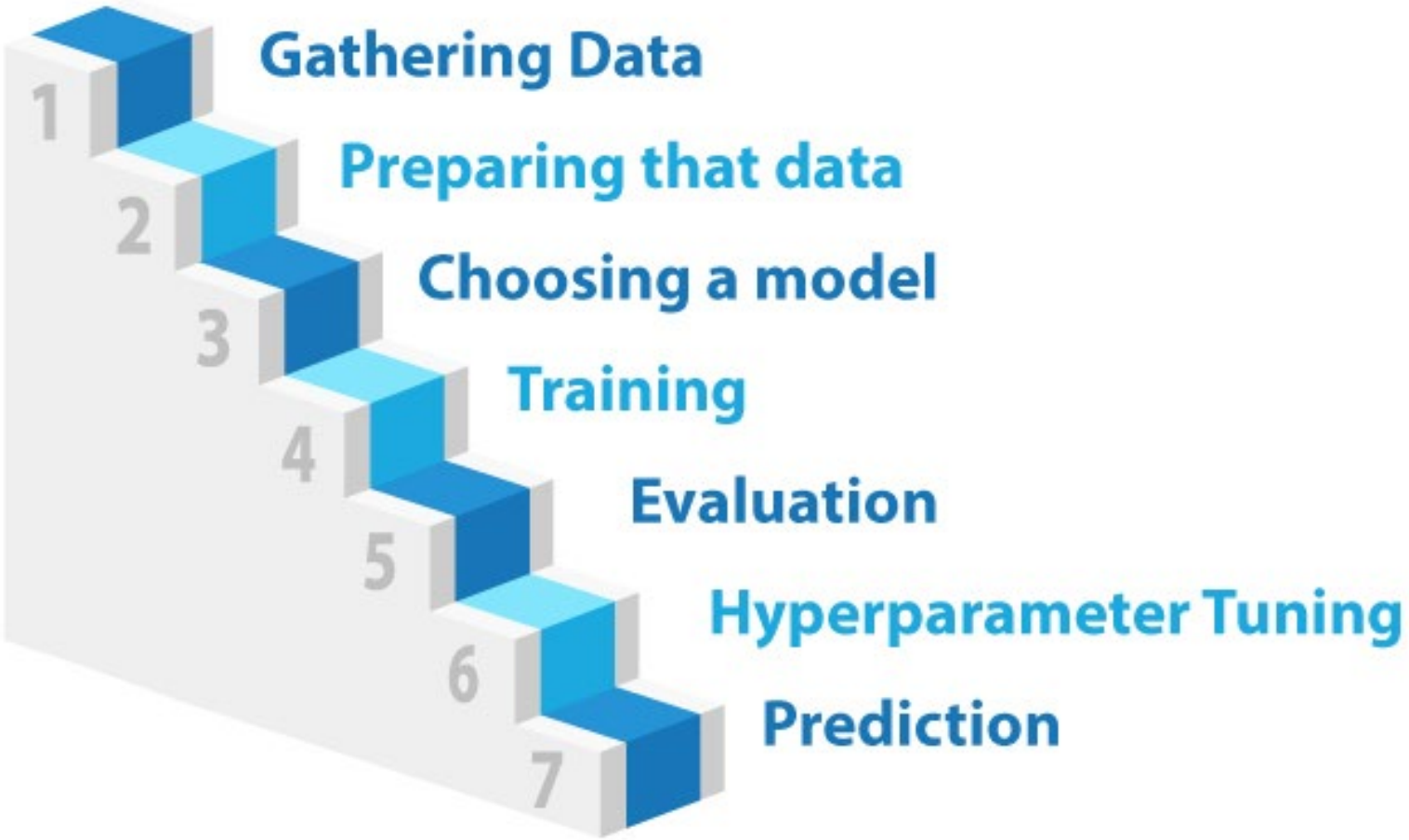
- Ensemble Learning: Ensemble learning combines multiple individual models to create a more robust and accurate predictive model.
- It can improve generalization and reduce the risk of overfitting.
- Examples of ensemble learning methods include bagging (e.g., random forests), boosting (e.g., AdaBoost, Gradient Boosting), and stacking.

The type of learning used depends on the problem context, the type of data available, and the specific outcomes desired.



Source: https://www.aiplusinfo.com/blog/is-deep-learning-supervised-or-unsupervised/

7 steps of Machine Learning

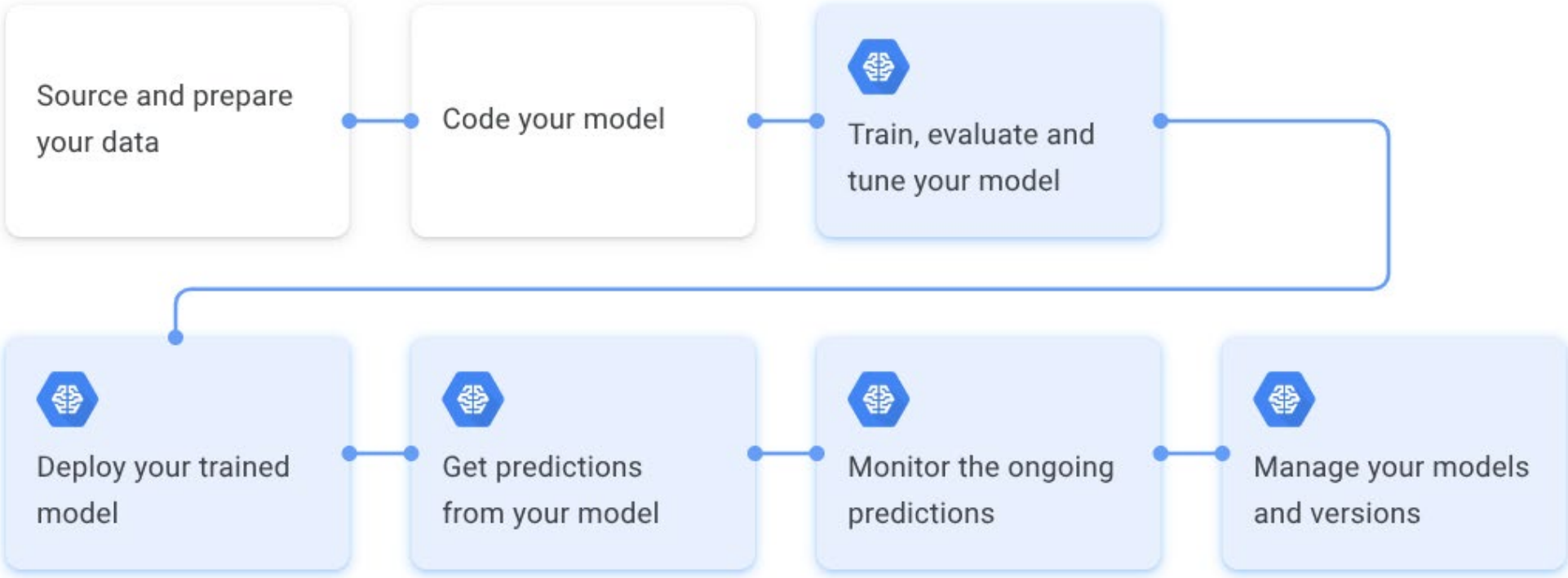


Source: https://tinyurl.com/3hy78es9

MACHINE LEARNING WORKFLOW

1. Should I use ML on this problem?
   - Is there a pattern to detect?
   - Can I solve it analytically?
   - Do I have data?
2. Gather and organize data.
   - Preprocessing, cleaning, visualizing.
3. Establishing a baseline.
4. Choosing a model, loss, regularization, ...
5. Optimization (could be simple, could be a Phd...).
6. Hyperparameter search.
7. Analyze performance & mistakes, and iterate back to step 4 (or 2).

MACHINE LEARNING WORKFLOW

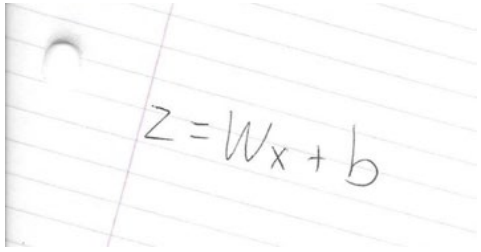


ML workflow

Source: https://cloud.google.com/ai-platform/docs/ml-solutions-overview

IMPLEMENTING ML SYSTEMS

- You will often need to derive an algorithm (with pencil and paper), and then translate the math into code.
- Array processing (NumPy)
  - ✓ **vectorize** computations (express them in terms of matrix/vector operations) to exploit hardware efficiency
  - ✓ This also makes your code cleaner and more readable!



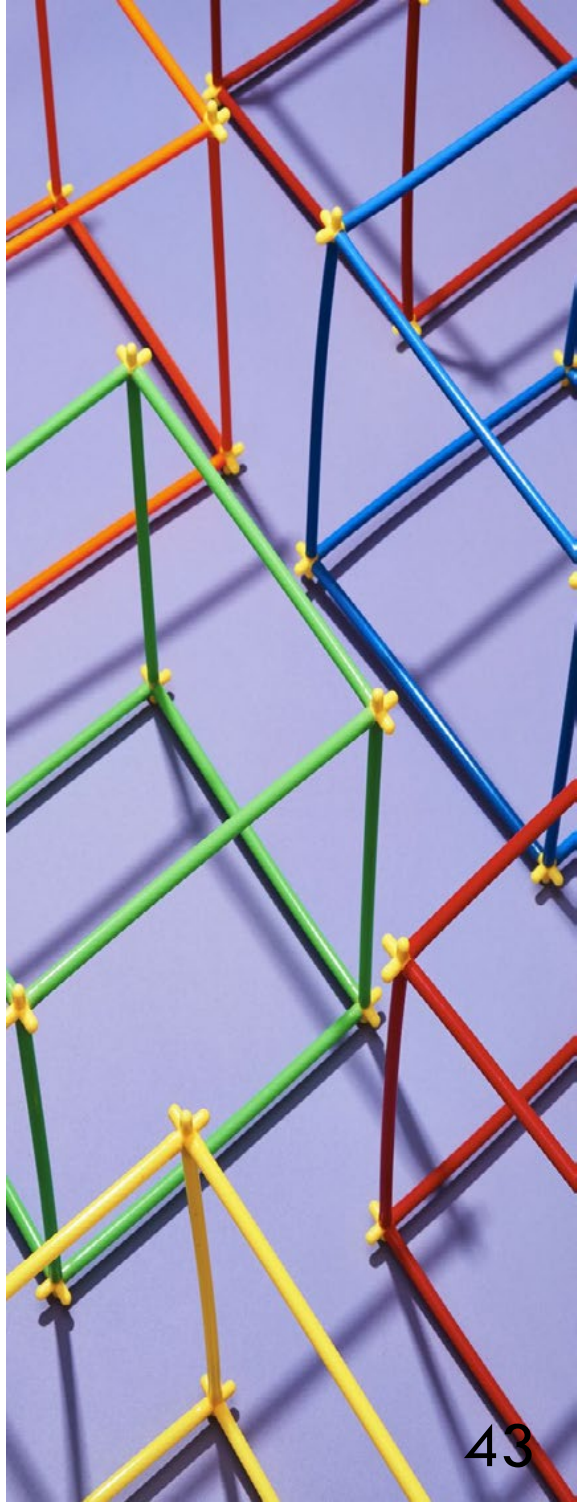

```
z = np.zeros(m)
for i in range(m):
    for j in range(n):
        z[i] += W[i, j] * x[j]
    z[i] += b[i]
```

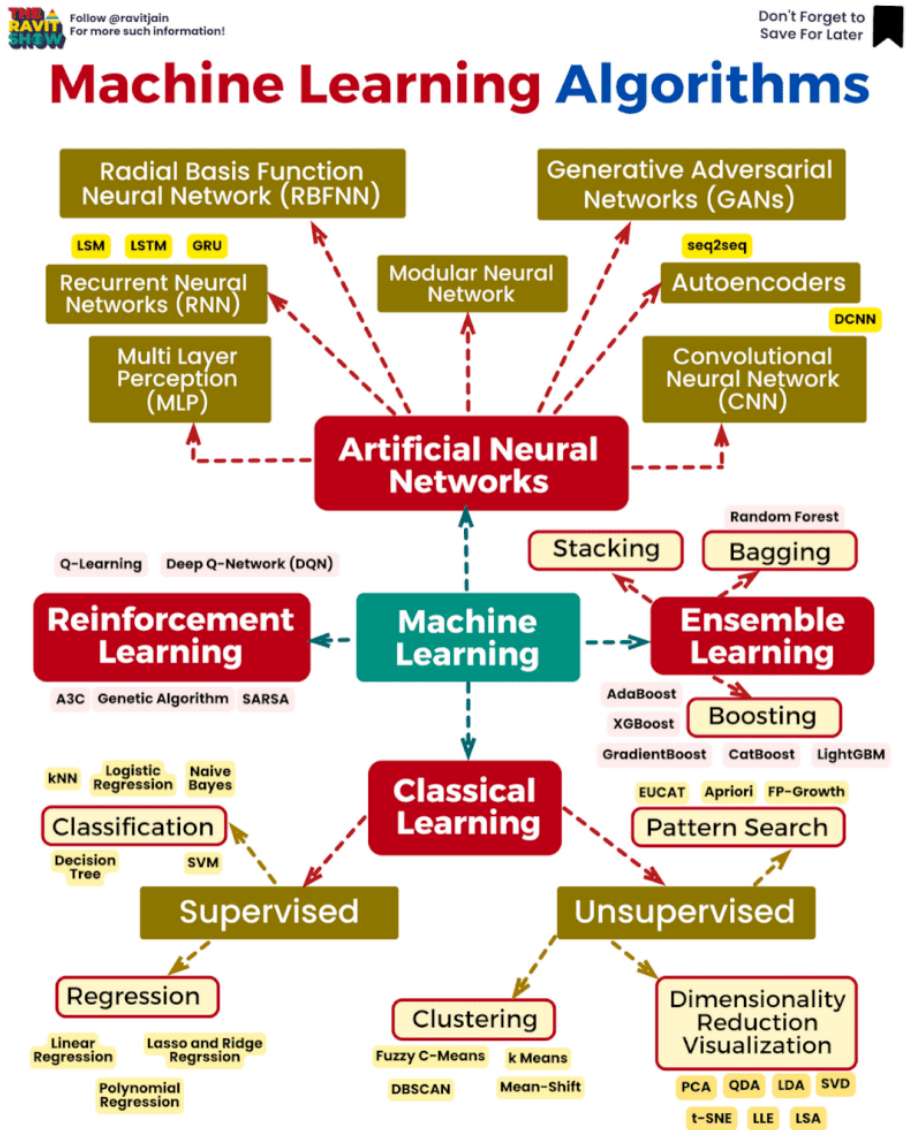
```
z = W @ x + b
```

IMPLEMENTING ML SYSTEMS

Neural net frameworks: PyTorch, TensorFlow, JAX, etc.

- automatic differentiation
- compiling computation graphs
- libraries of algorithms and network primitives
- support for graphics processing units (GPUs)





[Source: https://www.linkedin.com/feed/update/urn:li:activity:7095438905629421568?utm_source=share&utm_medium=member_android]

LIMITATIONS OF MACHINE LEARNING

- Data Quality
- Data Quantity
- Bias and Fairness
- Overfitting and Underfitting
- Privacy and Security
- Interpretability

DATA QUALITY

- Machine learning algorithms rely on data to learn and make predictions.
- If the data used to train the algorithm is of poor quality, inaccurate, or biased, the algorithm will produce inaccurate results.
- It is important to ensure that the data used to train machine learning algorithms is accurate, relevant, and unbiased.

DATA QUANTITY

- Machine learning algorithms require large amounts of data to learn effectively.
- If there is not enough data available, the algorithm may not be able to learn effectively and may produce inaccurate results.
- It is important to ensure that there is enough data available to train the algorithm effectively.

BIAS & FAIRNESS

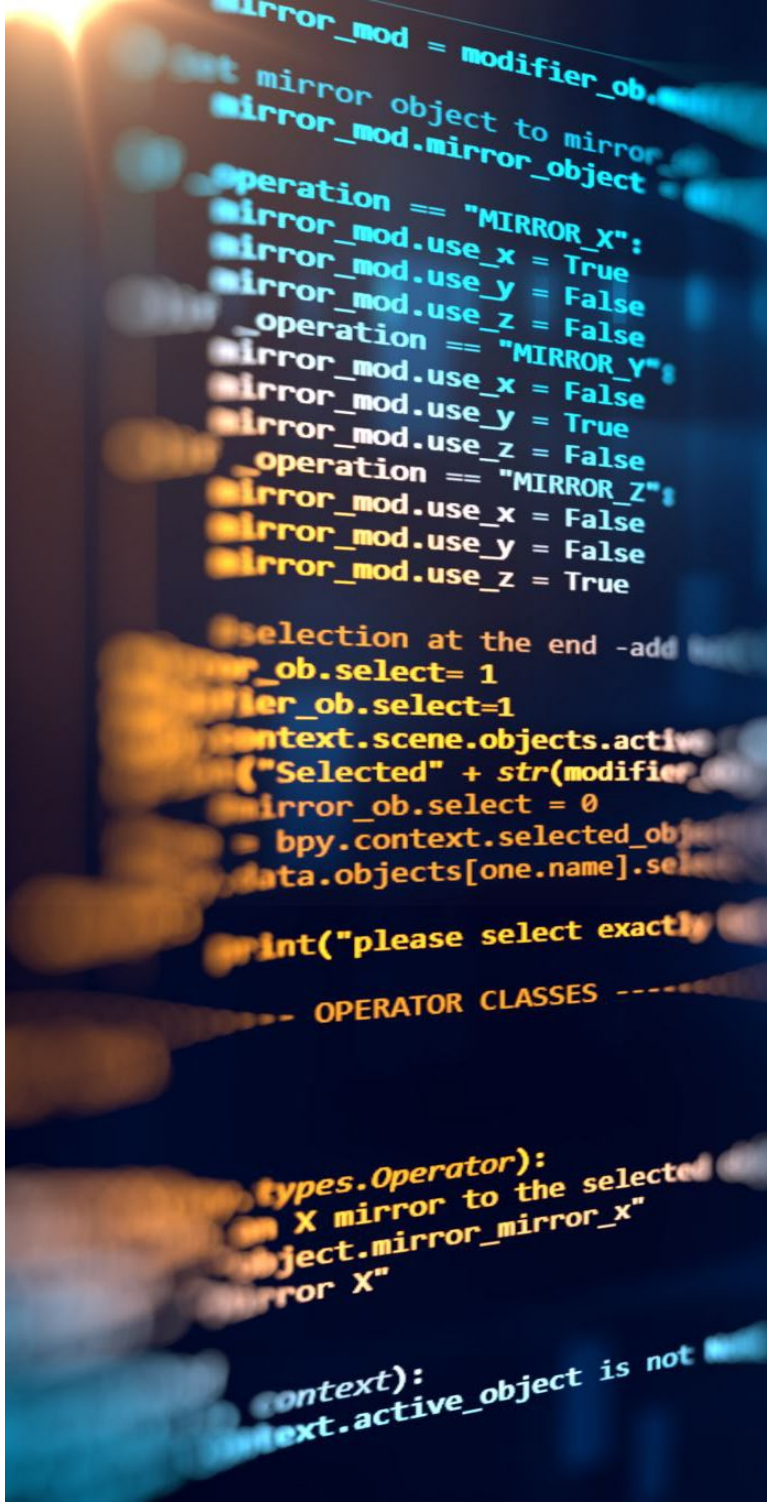
- Machine learning algorithms can be susceptible to bias if the data used to train them is biased. This can result in unfair or discriminatory decisions.
- It is important to ensure that the data used to train machine learning algorithms is unbiased and that measures are taken to prevent bias from being introduced into the algorithm.

OVERFITTING & UNDERFITTING

- Overfitting occurs when a machine learning algorithm is too complex and fits the training data too closely, resulting in poor performance on new data.
- Underfitting occurs when a machine learning algorithm is too simple and does not fit the training data well enough, resulting in poor performance on both the training and new data.
- It is important to find the right balance between complexity and simplicity when designing machine learning algorithms.

PRIVACY & SECURITY

- Machine learning algorithms often require access to sensitive data, such as personal information, to learn effectively.
- It is important to ensure that this data is handled securely and that measures are taken to protect the privacy of individuals.



INTERPRETABILITY

Machine learning algorithms can be opaque, making it difficult to understand how they make decisions.

This can make it difficult to debug the algorithm or to provide explanations for its decisions.

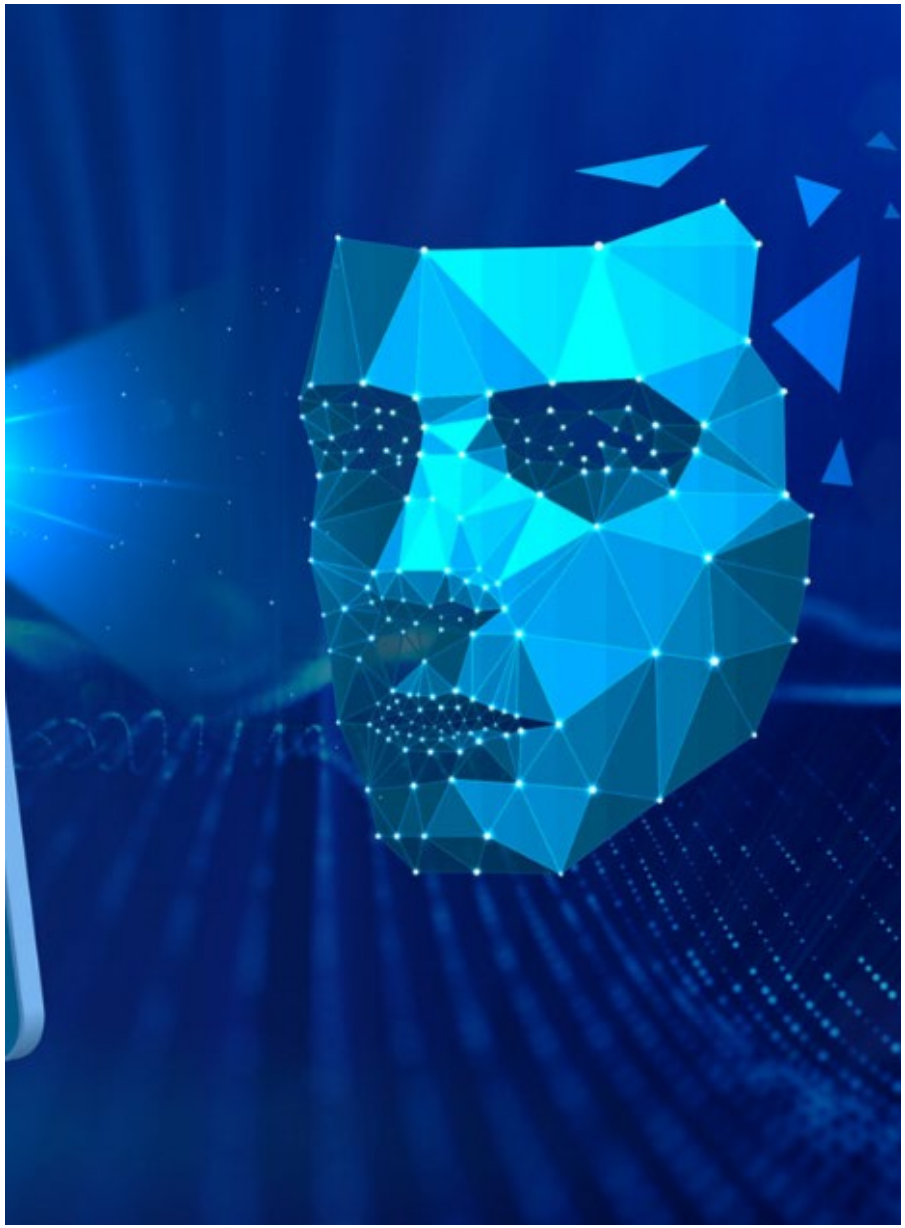
It is important to develop methods for making machine learning algorithms more interpretable so that their decisions can be better understood.

[Source: https://data-flair.training/blogs/machine-learning-applications/]

APPLICATIONS OF MACHINE LEARNING

IMAGE RECOGNITION

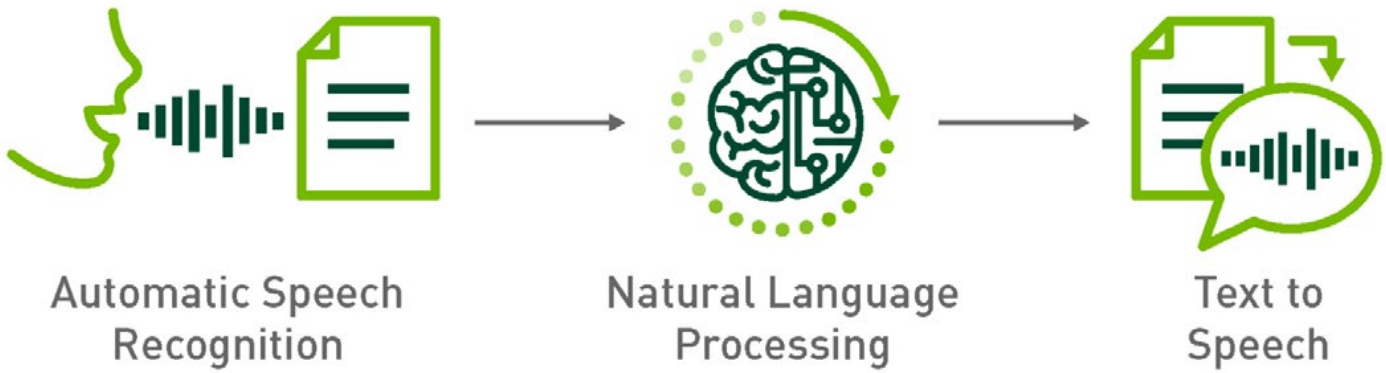
Machine learning algorithms can accurately identify objects in images, making it possible to automatically tag and organize photos. For example, Facebook uses machine learning to automatically recognize faces in photos and suggest tags for them.



https://www.flickr.com/photos/mikemacmarketing/30188201497/in/photostream/

SPEECH RECOGNITION

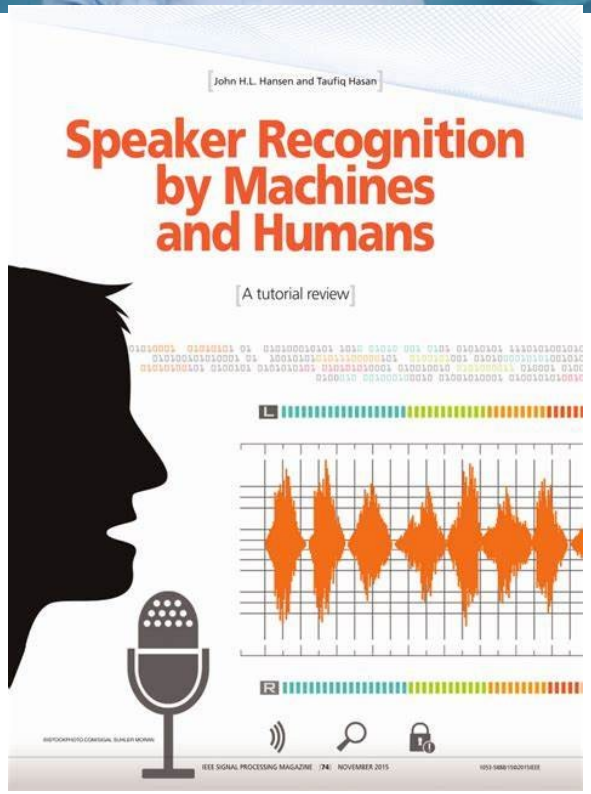
Companies like Apple and Google use machine learning to power their virtual assistants, Siri and Google Assistant, respectively. Machine learning algorithms analyze audio data to understand spoken language and respond with appropriate actions.



[Source: https://developer.nvidia.com/blog/wp-content/uploads/2019/12/Conversation_AI_Workflow_v02.png]

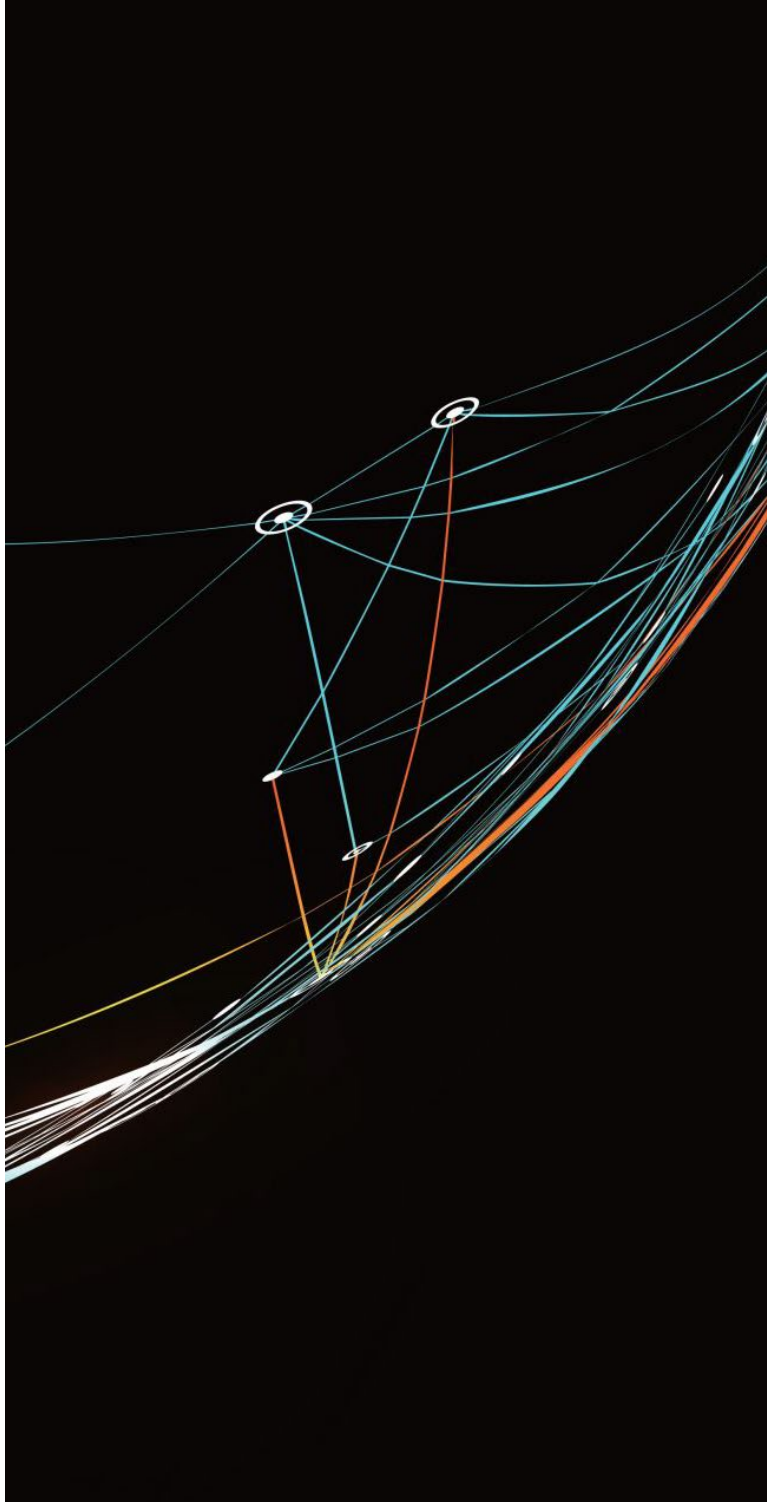
EXAMPLES - ML APPLICATIONS

Speech: Speech to text, personal assistants, speaker identification...



MEDICAL DIAGNOSTICS

- ML helps solve diagnostic and prognostic problems in medical domains.
- ML is used for prognosis, prediction of disease progression, outcomes research, therapy planning, patient management, data analysis, and intelligent alarming.
- Successful implementation of ML can improve the efficiency and quality of medical care.
- In medical diagnosis, ML improves accuracy by analyzing patient data from medical tests, diagnostics, symptoms, and physical information.
- Doctors can use ML results to narrow down on the disease inflicting the patient.

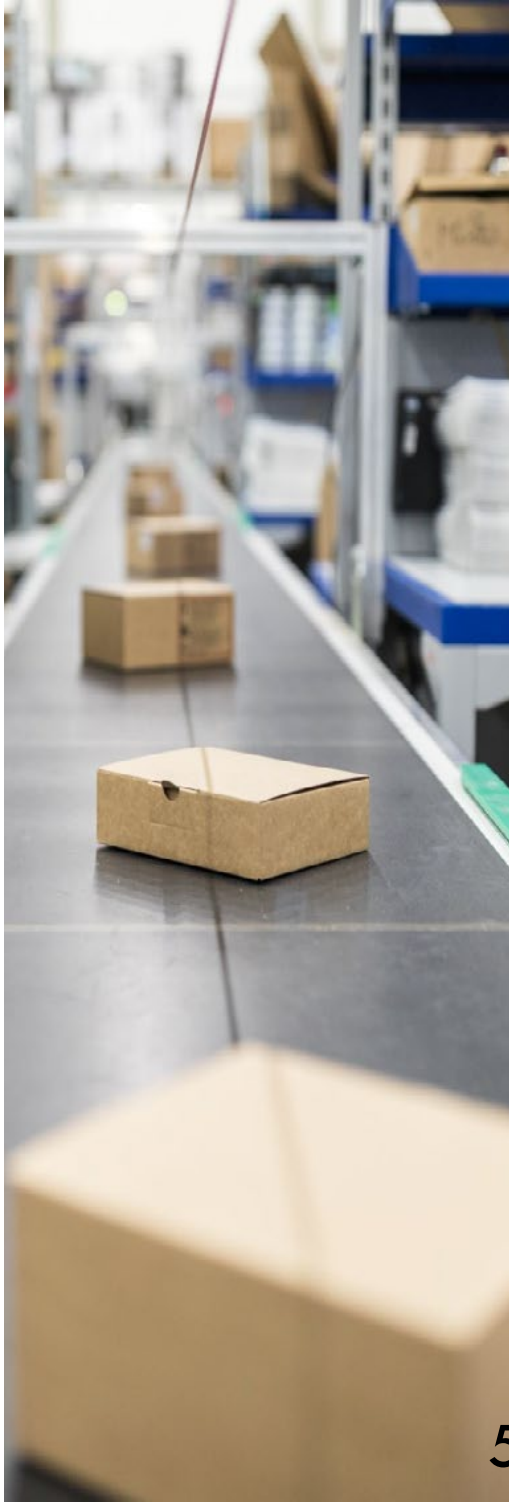


STATISTICAL ARBITRAGE

- Statistical arbitrage involves automated, short-term trading of many securities.
- Algorithms use historical correlations and economic variables.
- Machine learning methods, including linear regression and Support Vector Regression (SVR), are applied for index arbitrage strategy.
- PCA reduces feature space dimension.
- Trading signals are generated by modeling regression residuals as mean-reverting.
- Classification categories include sell, buy, or do nothing.
- Expected security return can be predicted and used for trading decisions.

LEARNING ASSOCIATION

- Learning association is the process of developing insights into associations between products.
- Machine learning can study the association between products people buy, known as basket analysis.
- Knowing these relationships can help suggest associated products to customers and bundle products for a better package.
- Associations are found by examining large amounts of sales data and developing rules to derive probability tests in learning conditional probability.



CLASSIFICATION

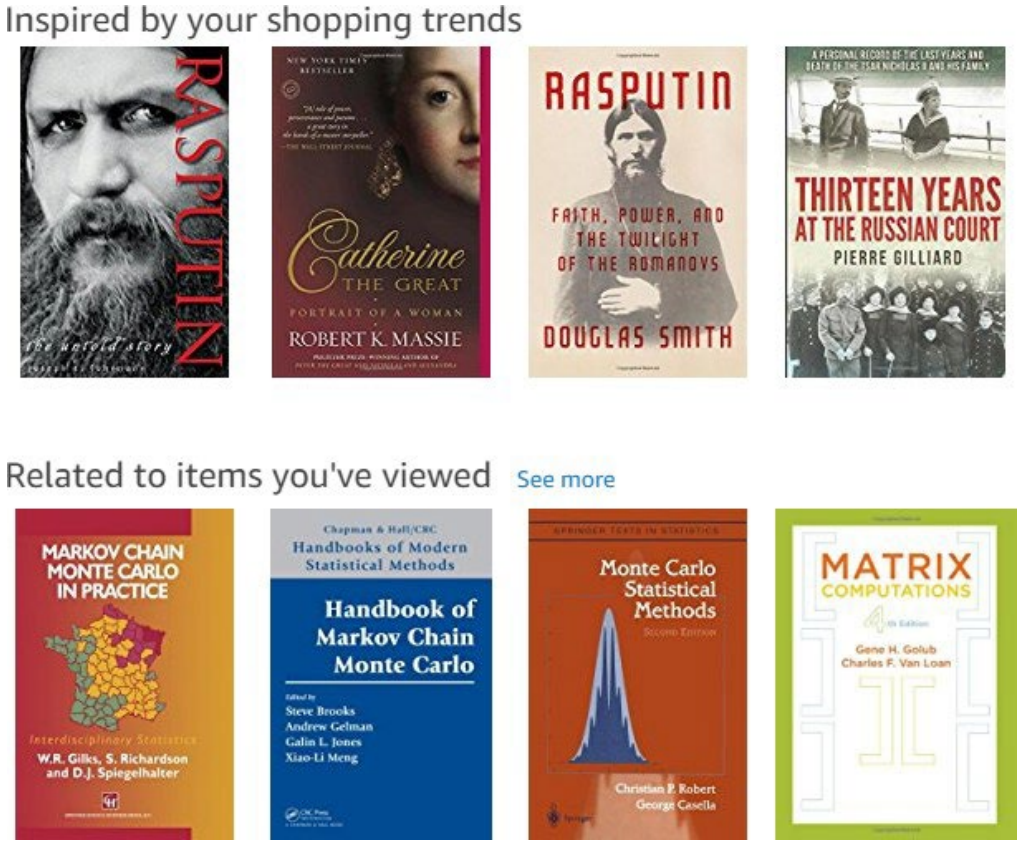
- Classification is the process of placing individuals from a population into classes based on independent variables.
- Classification helps analysts use measurements of an object to identify its category.
- Data consisting of examples of objects with their correct classification is used to establish an efficient rule.
- For example, banks assess customers' ability to repay loans by considering factors such as earnings, age, savings, and financial history.
- This information is taken from past loan data to create a relationship between customer attributes and related risks.

PREDICATION

- Prediction is the output of a machine learning algorithm that forecasts outcomes based on historical data.
- Predict functions are used to make these forecasts.
- Predictions can be applied across sectors for varied purposes, such as predicting loan defaults or future sales.
- These predictions can be made during or after model creation, or even after a failure.
- For example, in retail, businesses use predictions to make procurement and stock decisions.

EXAMPLES - ML APPLICATIONS

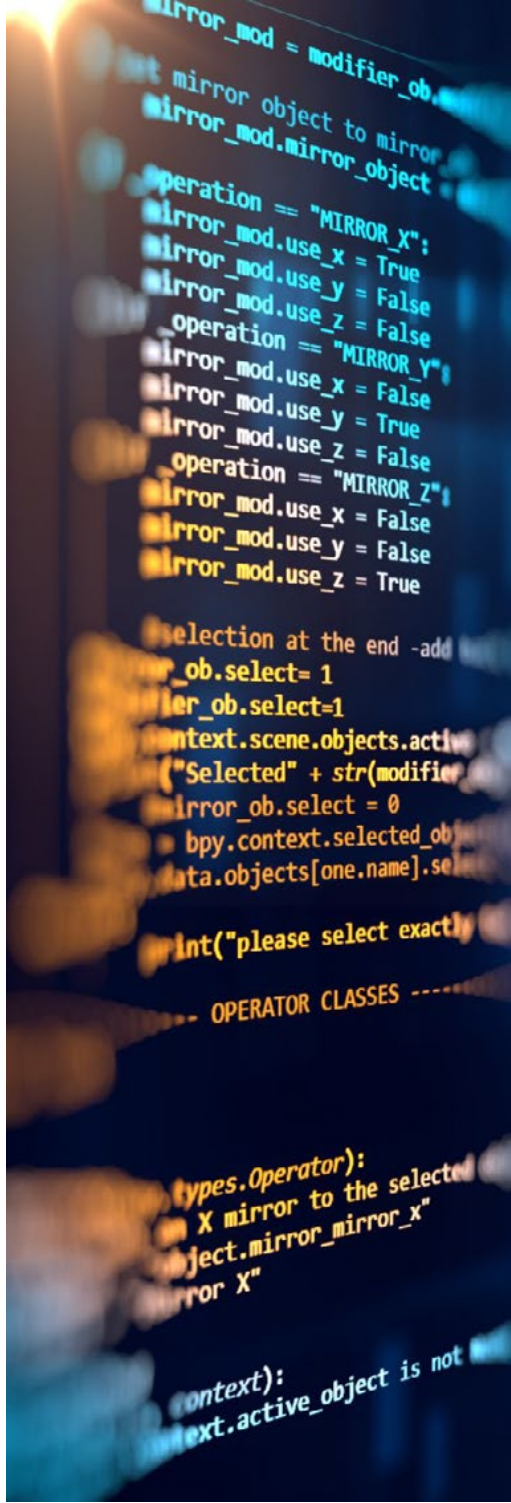
E-commerce & Recommender Systems : Amazon, netflix, ...



Source: https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-applications

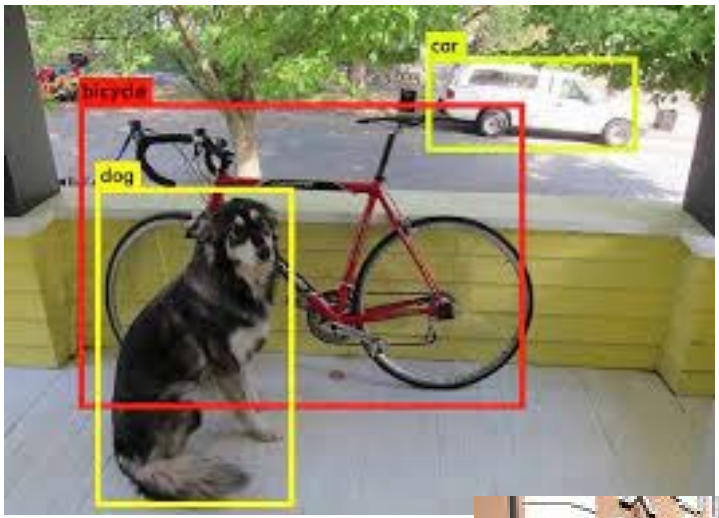
EXTRACTION

- Information Extraction (IE) is the process of extracting structured information from unstructured data, such as web pages, articles, and e-mails.
- The output produced by IE is maintained in a relational database and is in a summarized form, such as an excel sheet or table.
- IE is becoming a key in the big data industry due to the huge volume of unstructured data being generated.
- The first key challenge is handling unstructured data and converting it to structured form based on some pattern so that it can be stored in RDBMS.
- Nowadays, businesses want data as soon as it is generated, i.e. in real time, rather than in batches like End-of-Day (EOD).

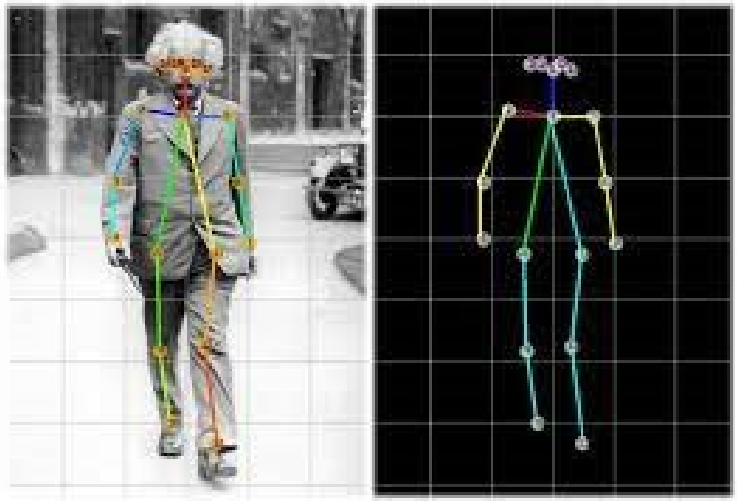


EXAMPLES - ML APPLICATIONS

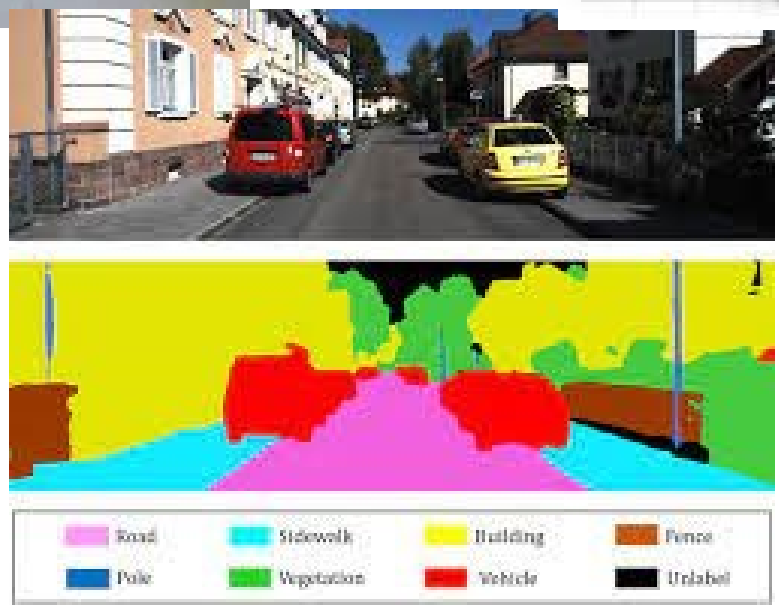
Machine Vision:



Object detection



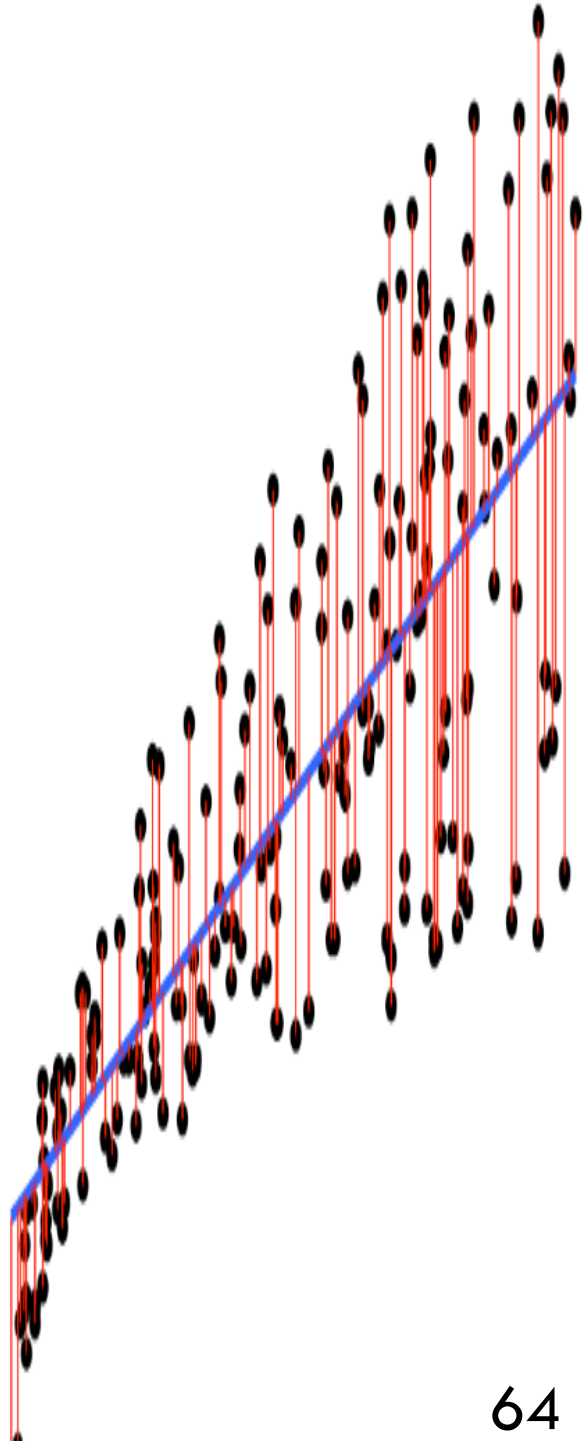
pose estimation



Semantic Segmentation

REGRESSION

- Machine learning can be applied to regression to produce an outcome variable (y) based on input variables (x).
- A model expresses the relationship between parameters as Y=g(x), where g is a function of the model's characteristics.
- Machine learning optimizes parameters to minimize approximation error and calculate the closest outcome.
- Inputs can be altered for a better model, known as response surface design.



OTHER APPLICATIONS

Social Media Features	Product Recommendations	Sentiment Analysis
Automating Employee Access Control	Marine Wildlife Preservation	Regulating Healthcare Efficiency and Medical Services
Predict Potential Heart Failure	Banking Domain	Language Translation

Source: https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-applications

REFERENCES

https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-applications

MACHINE LEARNING WITH PYTHON By SAI LIKHITH PANUGANTI


HISTORIA DE LA FILOSOFÍA
BERTRAND RUSSELL
HISTORIA UNIVERSAL
J. M. ROBERTS
HISTORIA DE LA LITERATURA UNIVERSAL
SUMMA LITERARIA
6
HESÍODO
TEOGONÍA
HOMERO
DIDEA
TOM SHARPE
Vicios ancestrales
Antonio Gala
El manuscrito carmesí
Historia de los Egipcios

QUESTIONS