MACHINE LEARNING

MEI/1

University of Beira Interior,
Department of Informatics

Hugo Pedro Proença,
hugomcp@di.ubi.pt, 2023/2024

Machine Learning

[01]

- **Assiduity (A)** To get approved at this course, students should attend to - at least - 80% of the theoretical and practical classes.
- Practical Projects (**P**). The practical projects of this course **weight 50% (10/20)** of the final mark.
  - (P1) **Project 1**: Supervised Learning (Linear Regression) (5/20). Due: Monday, October 10th, 2023, 23:59:59.
  - (P2) **Project 2**: Supervised Learning (Classification) (5/20). Due: Monday, October 31st, 2023, 23:59:59.
  - (P3) **Project 3**: Unsupervised Learning (5/20). Due: Monday, November 28th, 2023, 23:59:59.
  - (P4) **Project 4**: Reinforcement Learning (5/20). Due: Monday, December 19th, 2023, 23:59:59.
- To get approved at the course, a minimal mark of 8/20 should be obtained in the practical part.
- The practical mark is considered in all examination epochs.
- Written Test (**F**) Tuesday, January 9th, 2024, 14:00, Room 6.18.
  - Mark (M) M = [A >= 0.8] * (P * 10/20 + **F** * 10/20).
- Admission to Exams
  - Students with M >= 6 are admitted to final exams.

Machine Learning? What is It?

- Arthur Samuel
  - *"Field of study that gives computers the ability to **learn without being explicitly programmed**"*.
- Yoshua Bengio
  - *"Part of research on artificial intelligence, seeking to provide **knowledge** to computers **through data**, observations and interacting with the world. That acquired knowledge allows computers to correctly generalize to new settings".*
- Wikipedia
  - *"Is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead."*
- *NVIDIA*
  - "*Machine Learning at its most basic is the practice of using algorithms to **parse data**, **learn from it**, and then make a determination or prediction about something in the world*."
- Google Glossary
  - "*A program or system that **trains a model** from **input data**. The trained model can make useful **predictions** from **new** (never-before-seen) data drawn **from the same distribution** as the one used to train the model*."

Machine Learning? Basics

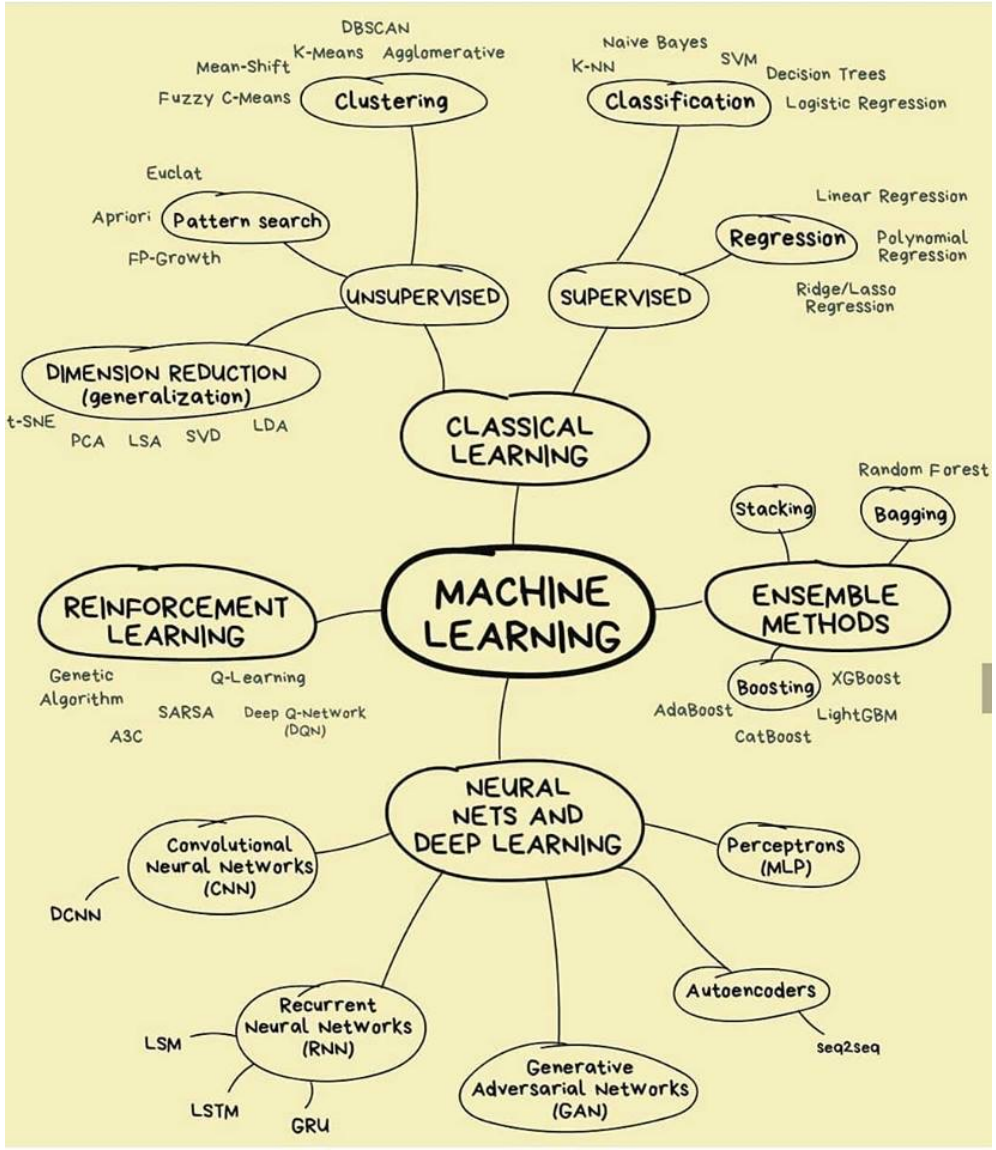
- There are nowadays many different machine learning algorithms, which are typically grouped by the kind of learning (supervised/unsupervised), or their similarity strategy/function (classification, regression, grouping,…)
- At the bottom line, every machine learning algorithm consists of three components: 1) **Representation** (the language, or instances that the algorithm understands); 2) **Evaluation** (how to distinguish between two instances); and 3) **Optimization** (how to search for the optimal solution).

Representation	Evaluation	Optimization
Instances	Accuracy/Error rate	Combinatorial optimization
K-nearest neighbor	Precision and recall	Greedy search
Support vector machines	Squared error	Beam search
Hyperplanes	Likelihood	Branch-and-bound
Naive Bayes	Posterior probability	Continuous optimization
Logistic regression	Information gain	Unconstrained
Decision trees	K-L divergence	Gradient descent
Sets of rules	Cost/Utility	Conjugate gradient
Propositional rules	Margin	Quasi-Newton methods
Logic programs		Constrained
Neural networks		Linear programming
Graphical models		Quadratic programming
Bayesian networks		
Conditional random fields		

credits: Pedro Domingos, U. Washington

Machine Learning? A Taxonomy

- The number of available Machine Learning algorithms is astonishing.
- However, they fall mainly into “supervised”, “unsupervised”, “statistical”, “neural” and “reinforcement learning” categories.



Machine Learning? Applications

- The use of sorting robots improves efficiency, accuracy and security during the sorting process and also saves 70% of manual work, dramatically reducing labor costs

credits: youtube.com

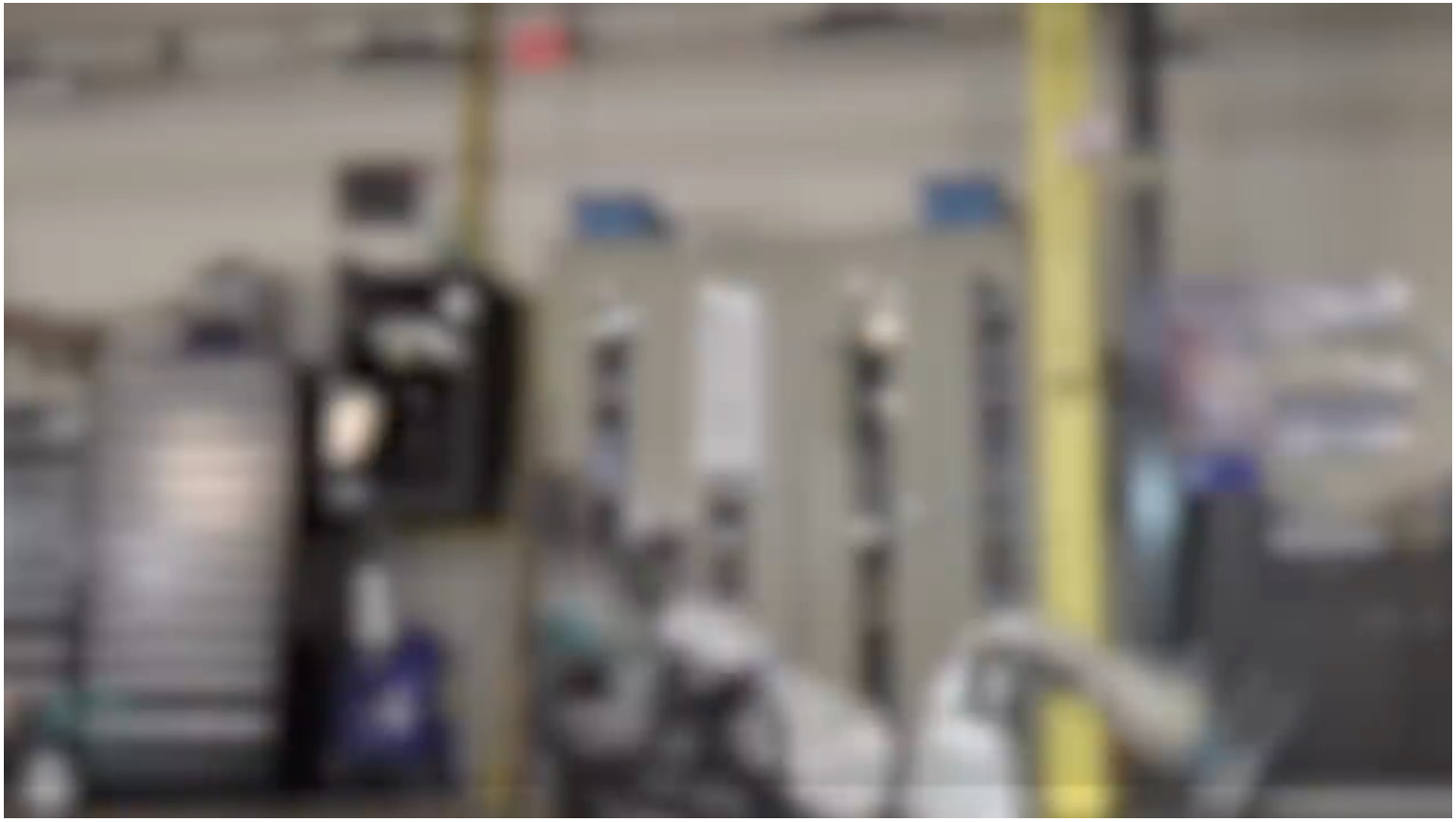
Machine Learning? Applications

- MOTOBOT 2 is capable of circulating a racetrack on its own. I September 2017, one of the project's milestones was reached as MOTOBOT successfully hit speeds exceeding 200 km/h. MOTOBOT then attempted to better the lap time set by MotoGP star, Valentino Rossi.

credits: youtube.com

Machine Learning? Applications

- Atlas is capable of performing human movements, even a tad bit better than actual humans. This incredible robot can understand obstacles easily and can even jump on or over them conveniently.



credits: youtube.com

Machine Learning? Applications

- Deep Blue
  - Twenty years ago IBM's Deep Blue defeated previously unbeaten chess grandmaster Gary Kasparov. Deep Blue won its first game against a world champion on 10 February 1996.
    - Deep Blue was the first computer system to defeat a reigning world champion in a match under standard chess tournament time controls. Kasparov accused IBM of cheating and demanded a rematch. IBM refused and dismantled Deep Blue.

credits: youtube.com

Machine Learning? Applications

- Self Driving Cars
  - Waymo cars (formerly known as “Google self-driving car”) might be the state-of-the-art in this extremely demanding application.

credits: youtube.com

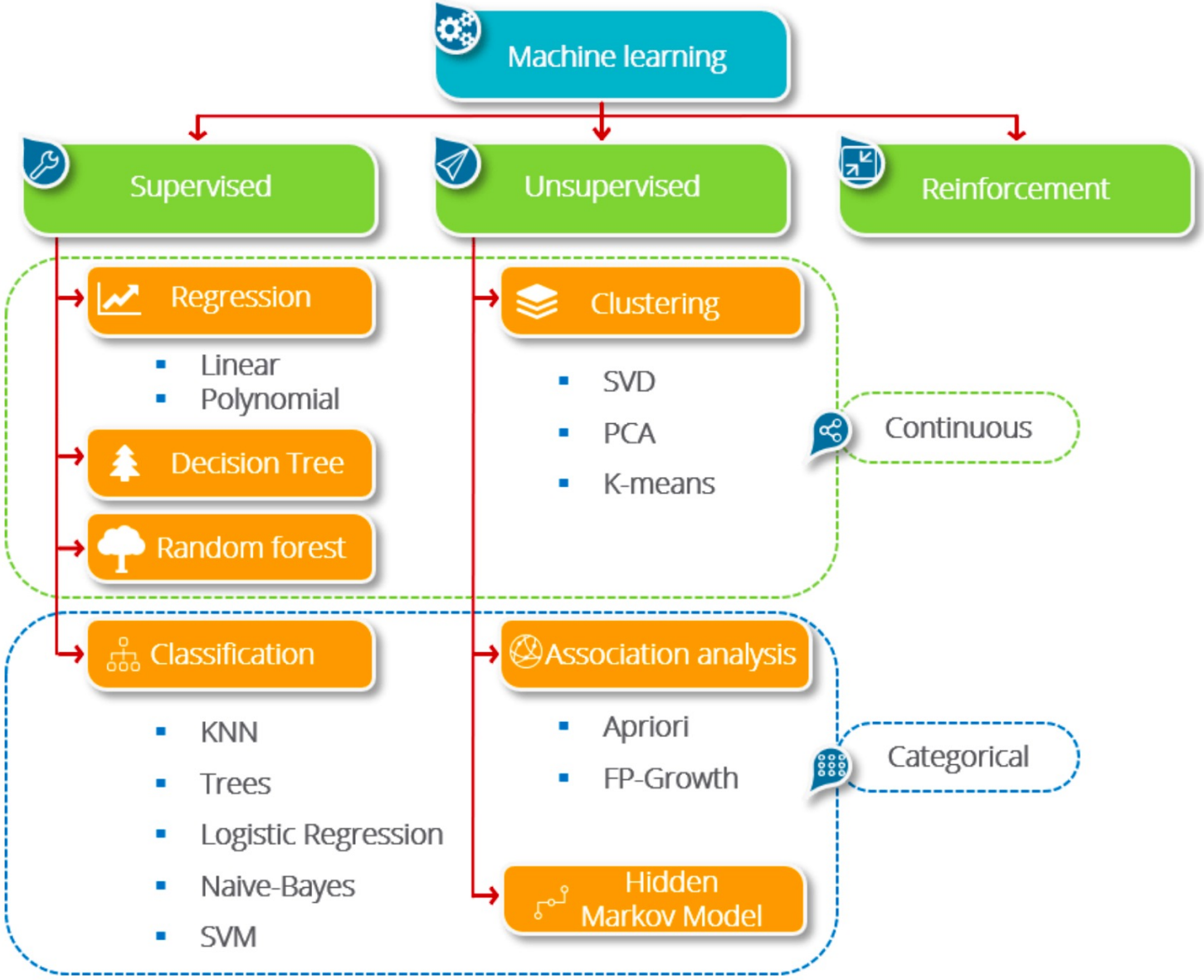
Machine Learning? Applications

- Watson is a question-answering computer system capable of answering questions posed in natural language, developed in IBM.
- The computer system was initially developed to answer questions on the quiz show Jeopardy! and, in 2011, the Watson computer system competed on Jeopardy! against human champions Brad Rutter and Ken Jennings, winning the first place prize of $1 million.
  - "Jeopardy!" is a classic game show, with a twist. The answers are given first, and the contestants supply the questions.

credits: youtube.com

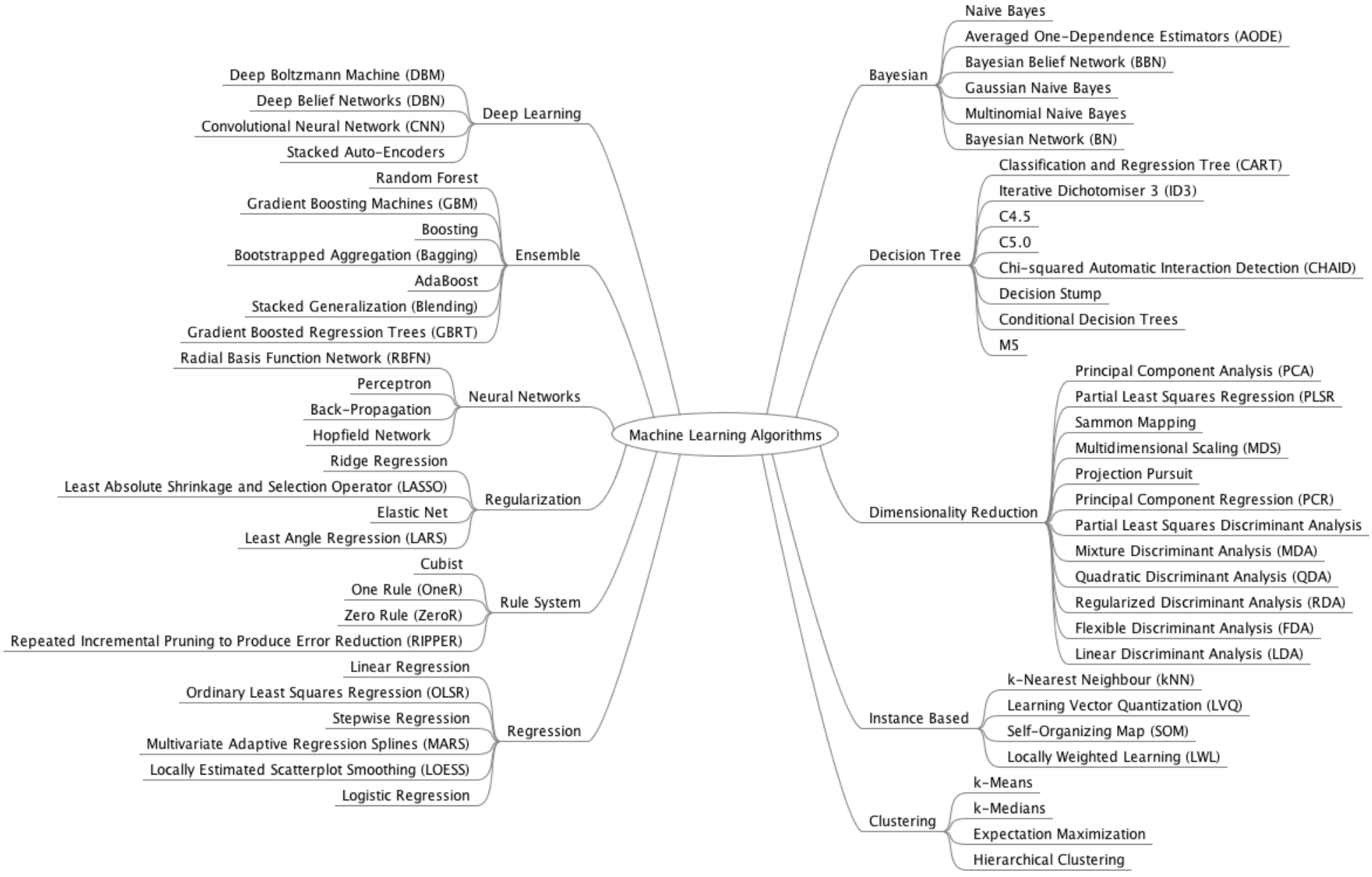
Machine Learning Taxonomy

- Machine Learning algorithms can be divided into three main types: *1) Supervised Learning; Unsupervised Learning; and 3) Reinforcement Learning*



credits: edureka.co

Machine Learning Algorithms Taxonomy



credits: MLMastery.com

Machine Learning Taxonomy

•Supervised Learning

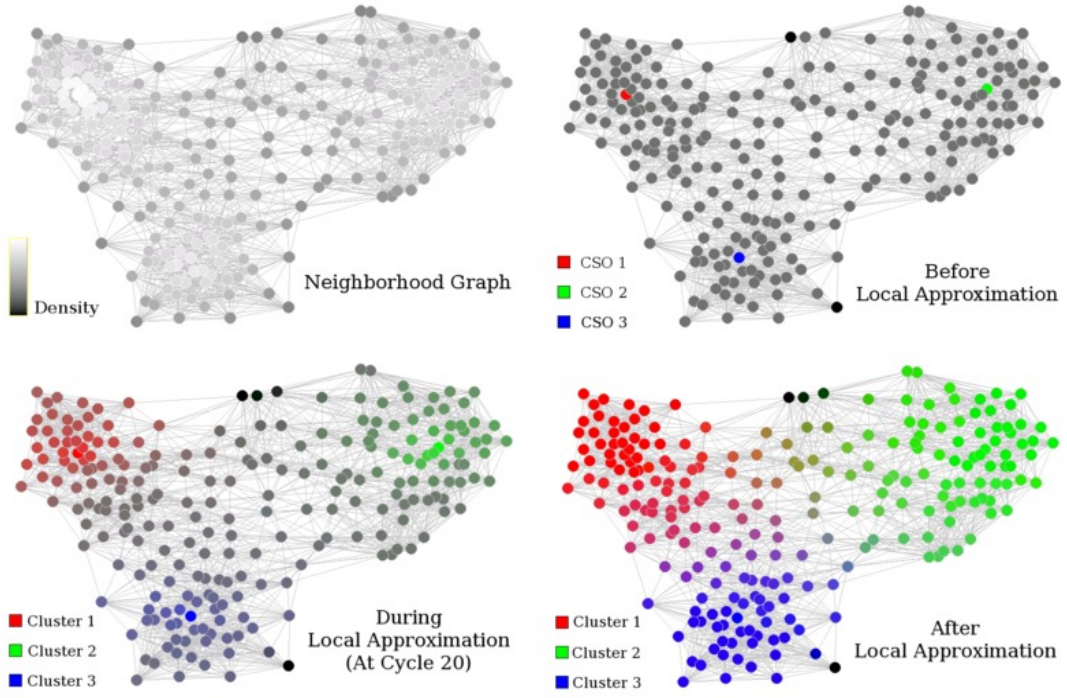
- The input data is called training (learning)data.
- Every instance has a known label or result, such as spam/not-spam or a stock price at a time.
- A model is inferred through a training/optimization process in which it is required to make predictions and is corrected when those predictions are wrong.
- The training process continues until the model achieves a desired level of accuracy on the training data.
- The typical problems in this family can be divided into classification and regression.
  - Classification problems typically evolve a “label” or “class” that should be automatically;
  - Regression problems evolve to predict one real value, by modelling the relationship between variables

credits: edureka.co

Machine Learning Taxonomy

- **Unsupervised Learning**

  - The model learns through observation to find structures in the data, which are modelled as “clusters”
  - It finds patterns and relationships in the dataset by creating such clusters.

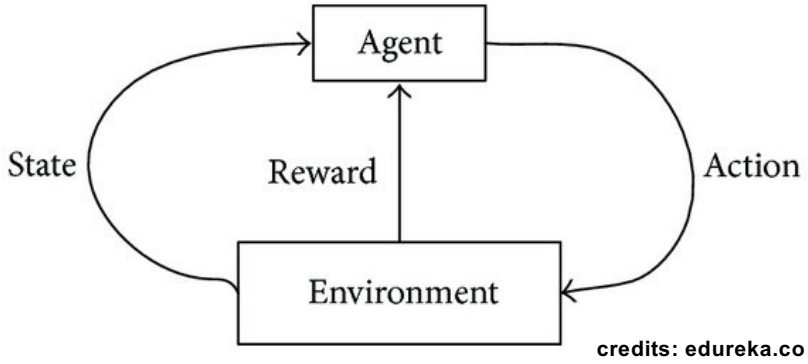


credits: Wikipedia.com

Machine Learning Taxonomy

- **Reinforcement Learning**

  - It is the ability of an automated agent to interact with the environment and find out what is the best outcome.
  - It follows the concept of hit and trial method. The agent is rewarded or penalized with a point for a correct or a wrong answer, and on the basis of the positive reward points gained the model trains itself.



  - Some of the most common algorithms in this family are Q-Learning, Temporal Difference and Deep Adversarial Networks.