KAZAKH-BRITISH TECHNICAL UNIVERSITY

Introduction to Machine Learning

Olivier JAYLET

School of Information Technology and Engineering

To contact me

- Email : o.jaylet@kbtu.kz[1]
- LinkedIn : olivier-jaylet
- In the office : Tuesday from 16h to 17h

[1] I'm almost always online during weekdays, but please let me sleep & ski in peace on Sundays

Objectives

The objective of this course is to provide students with the fundamental knowledge and skills which will enhance their competence in the field of the modern data science and machine learning.

Outcomes

At the end of the course, students must be able to :

- Describe different problems from machine learning domain.
- Solve mathematical problems that form the foundation of ML algorithm.
- Select and implement basic machine learning models and evaluate their performance.
- Be well-prepared to delve deeper into more advanced machine learning concepts, algorithms, and applications.

Syllabus

- Introduction to machine learning concepts
- Data and datasets
- Supervised learning
- Vectors and matrices
- Linear algebra recap
- Linear algebra recap (2)
- Calculus recap
- Midterm exam

Syllabus

- Calculus recap (2)
- Probability recap
- Probability recap (2)
- Statistics recap
- Probabilistic approach to machine learning
- Evaluation Metrics
- End term exam
- Final exam

What is Machine Learning?

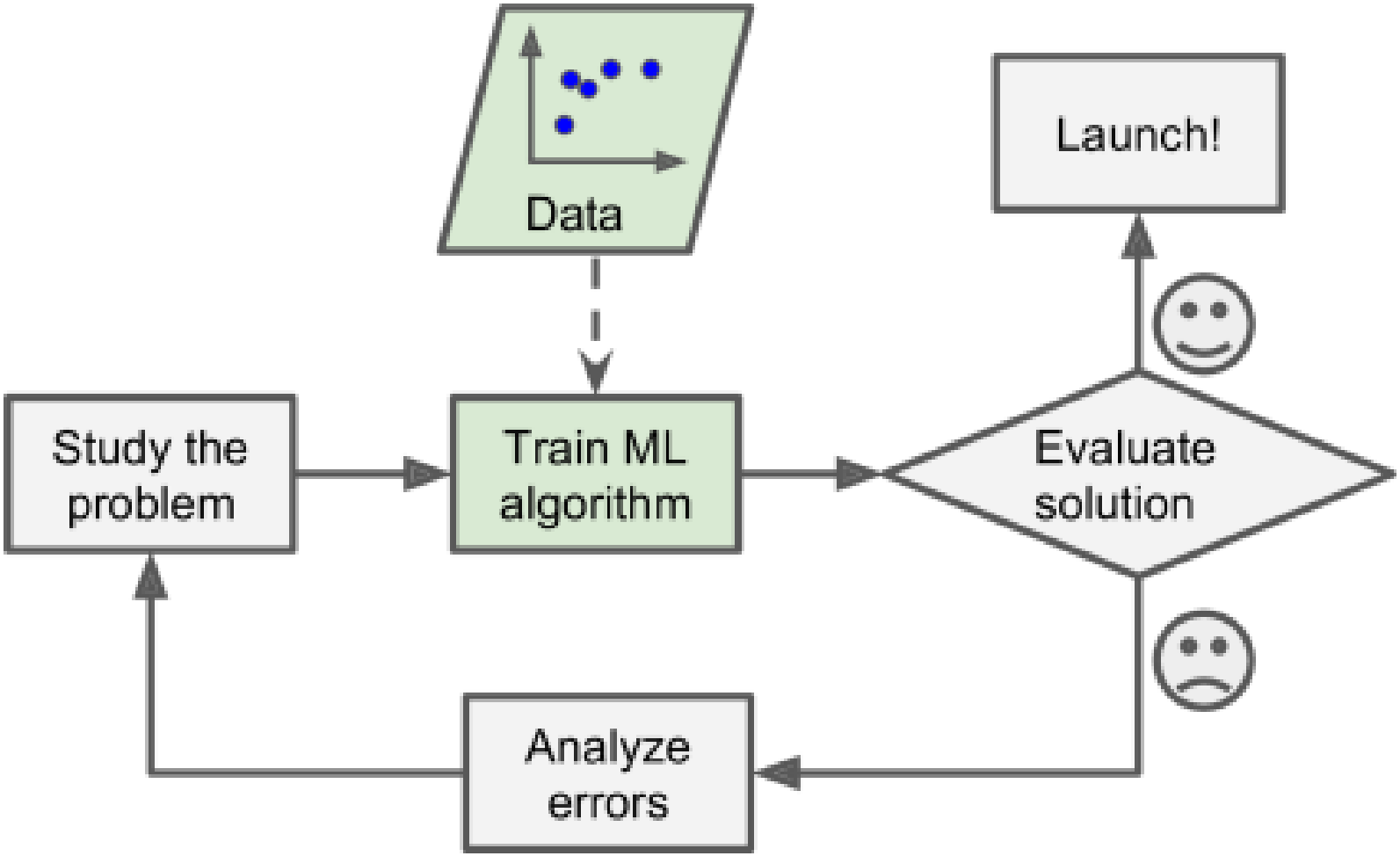
Machine learning is a branch of AI focused on building computer systems that learn from data to make predictions.
A field at the crossroads of three disciplines:

- Mathematics (Linear algebra, Calculus & optimization)
- Statistics
- Computer science

Two more definitions

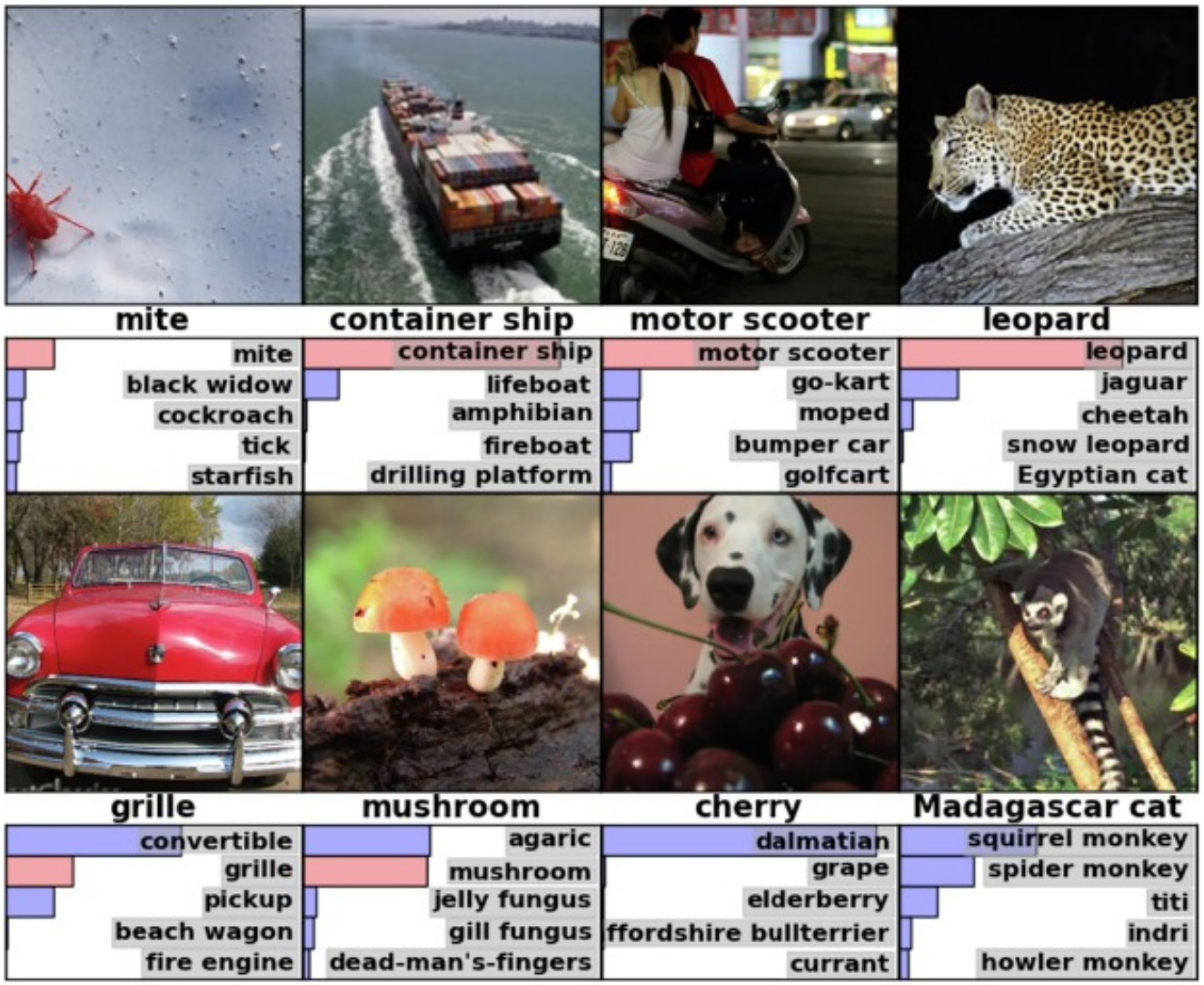
- According to Murphy (2012):
  - *machine learning* is defined as a set of methods that can **automatically detect patterns in data**, and then use the uncovered patterns to **predict future data**, or to perform other kinds of **decision making under uncertainty**.

- We can also read in Athey (2018):
  - *machine learning* is a field that **develops algorithms** designed to be applied to **datasets**, with the main areas of focus being **prediction** (regression), **classification**, and **clustering or grouping tasks**.


Data
Launch!
Study the problem
Train ML algorithm
Evaluate solution
Analyze errors

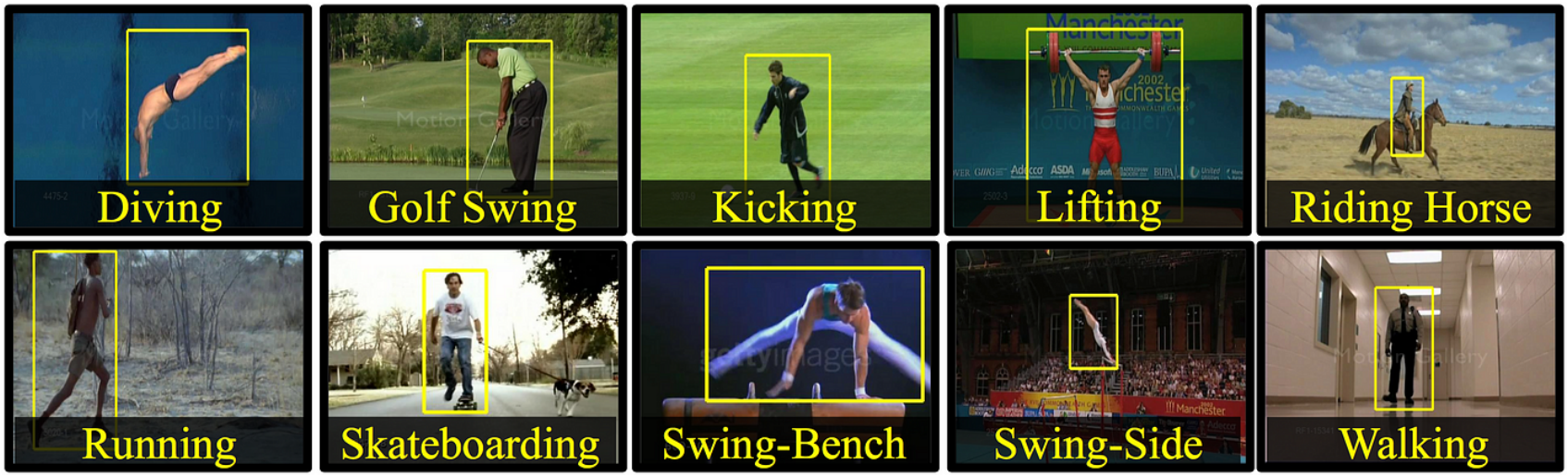
Applications of ML

Image classification



Object detection

Action recognition



Fraud detection

Car prices predictions

$24435.00

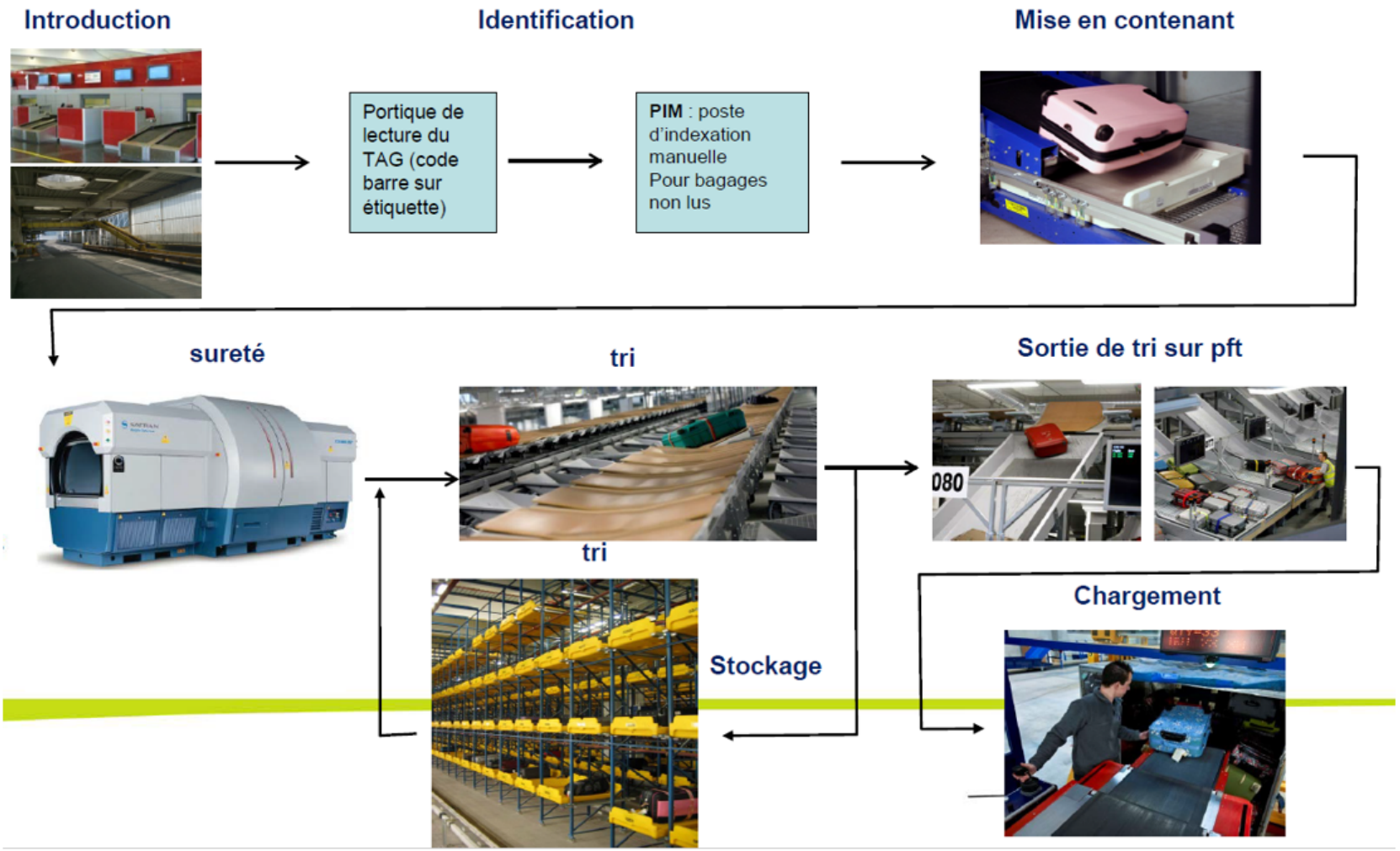
$36527.00

$18810.00

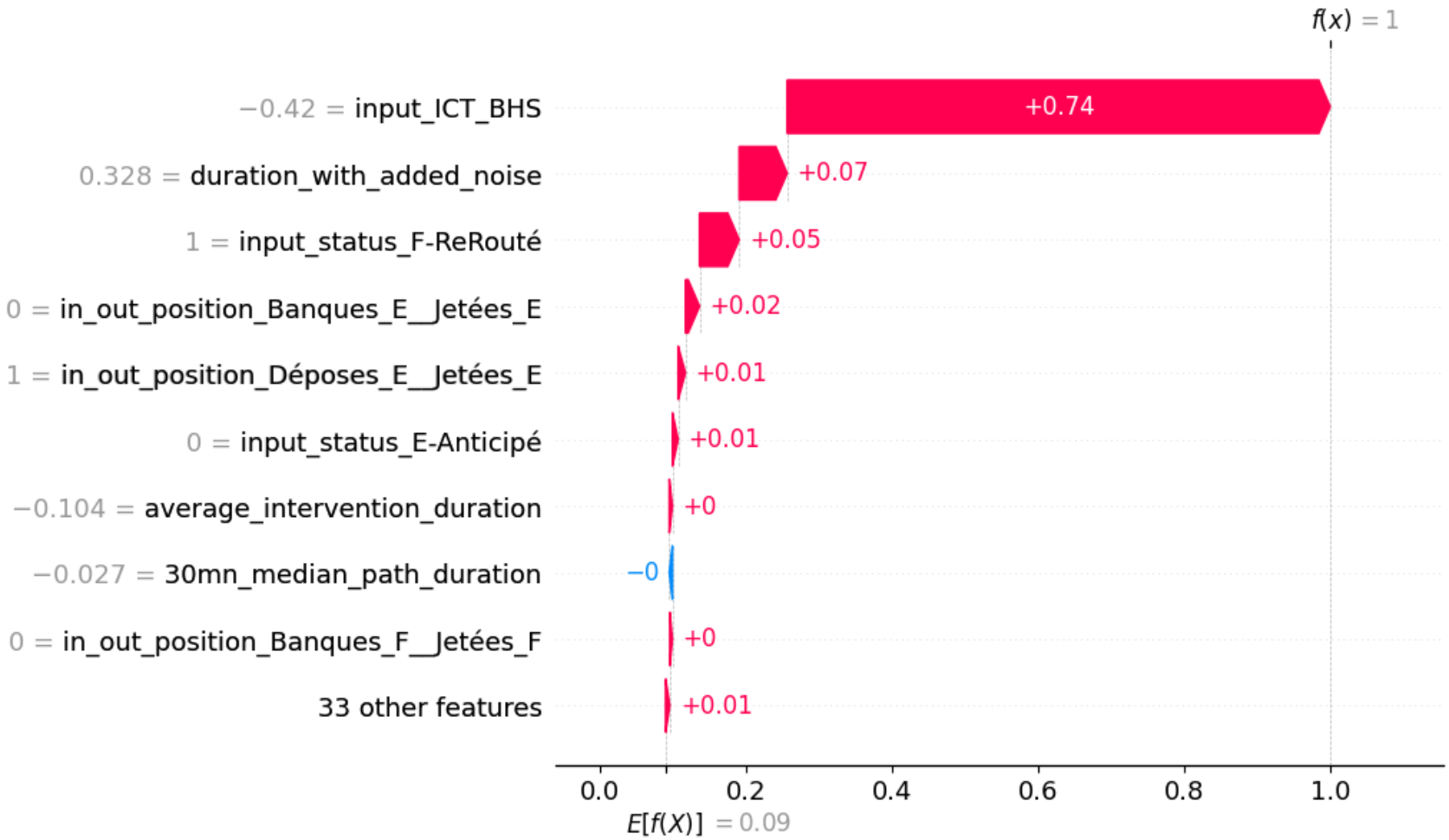
$26095.00

$5810.00

Mishandled baggage prediction



Mishandled baggage prediction

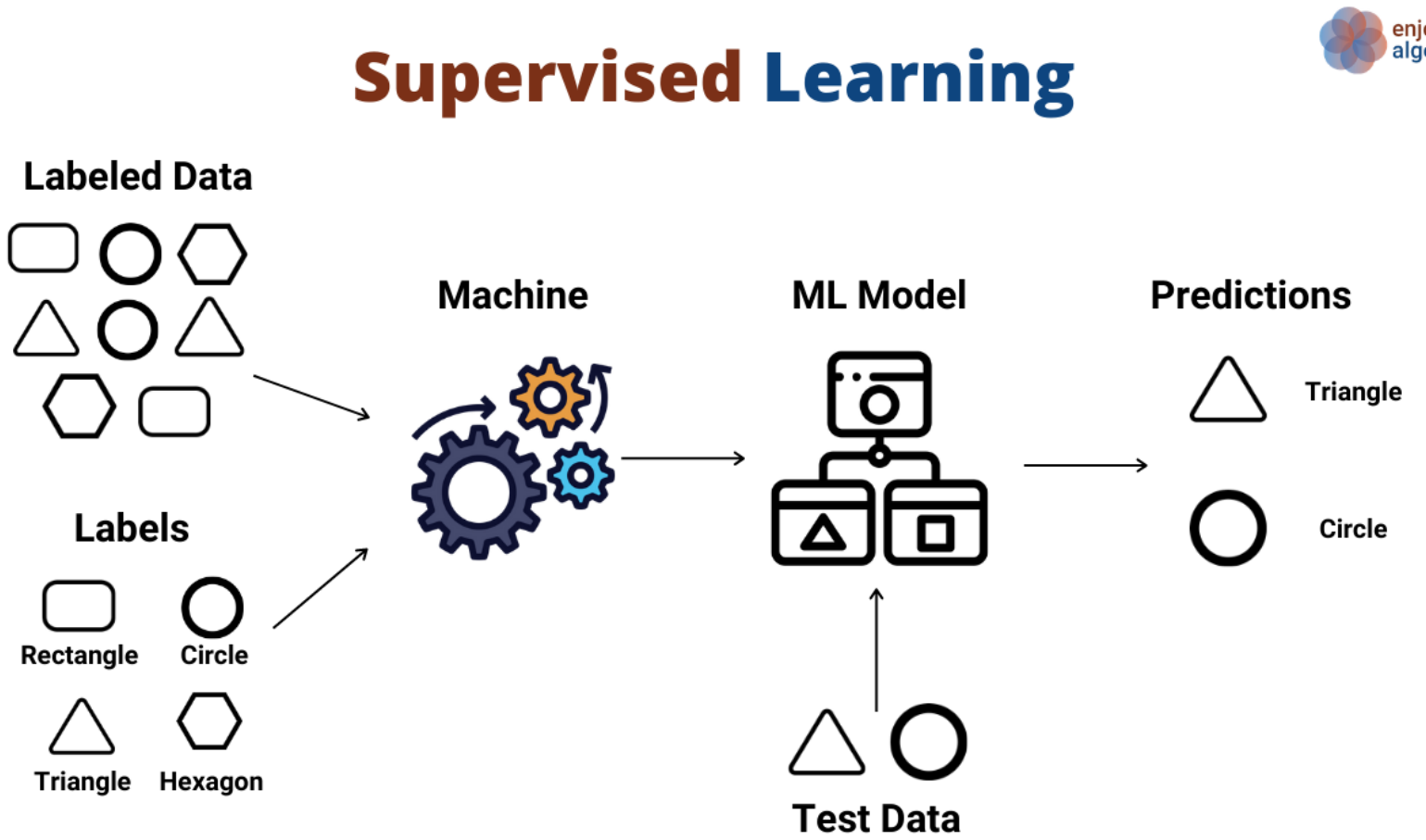


Mishandled baggage prediction

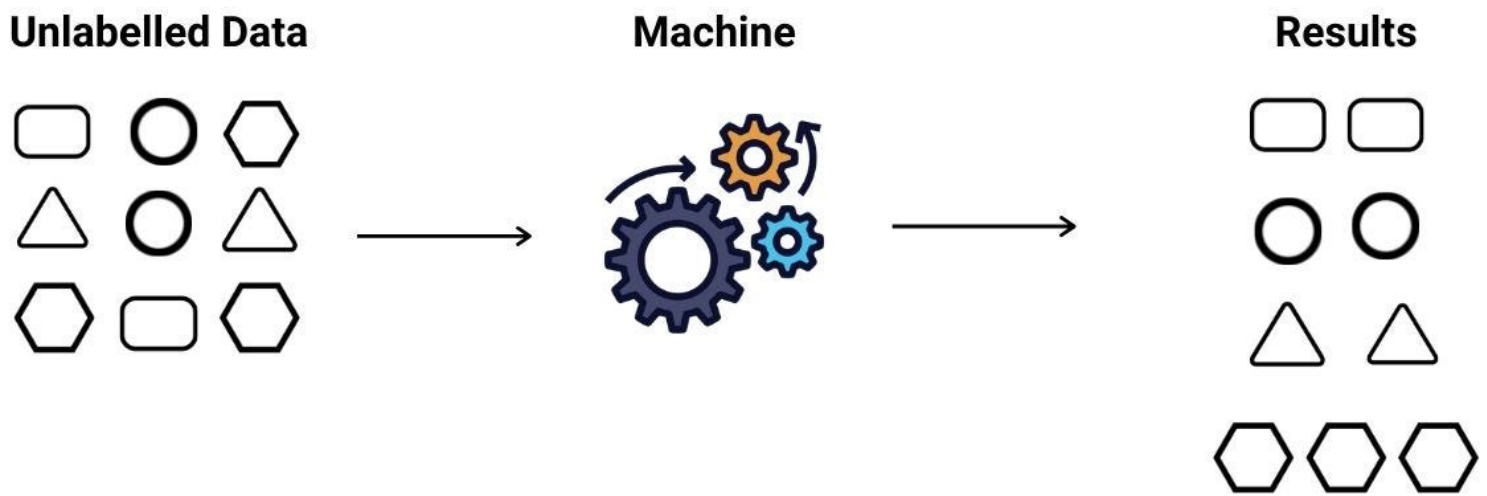
Two ML families

Usually, machine learning is divided in two categories :

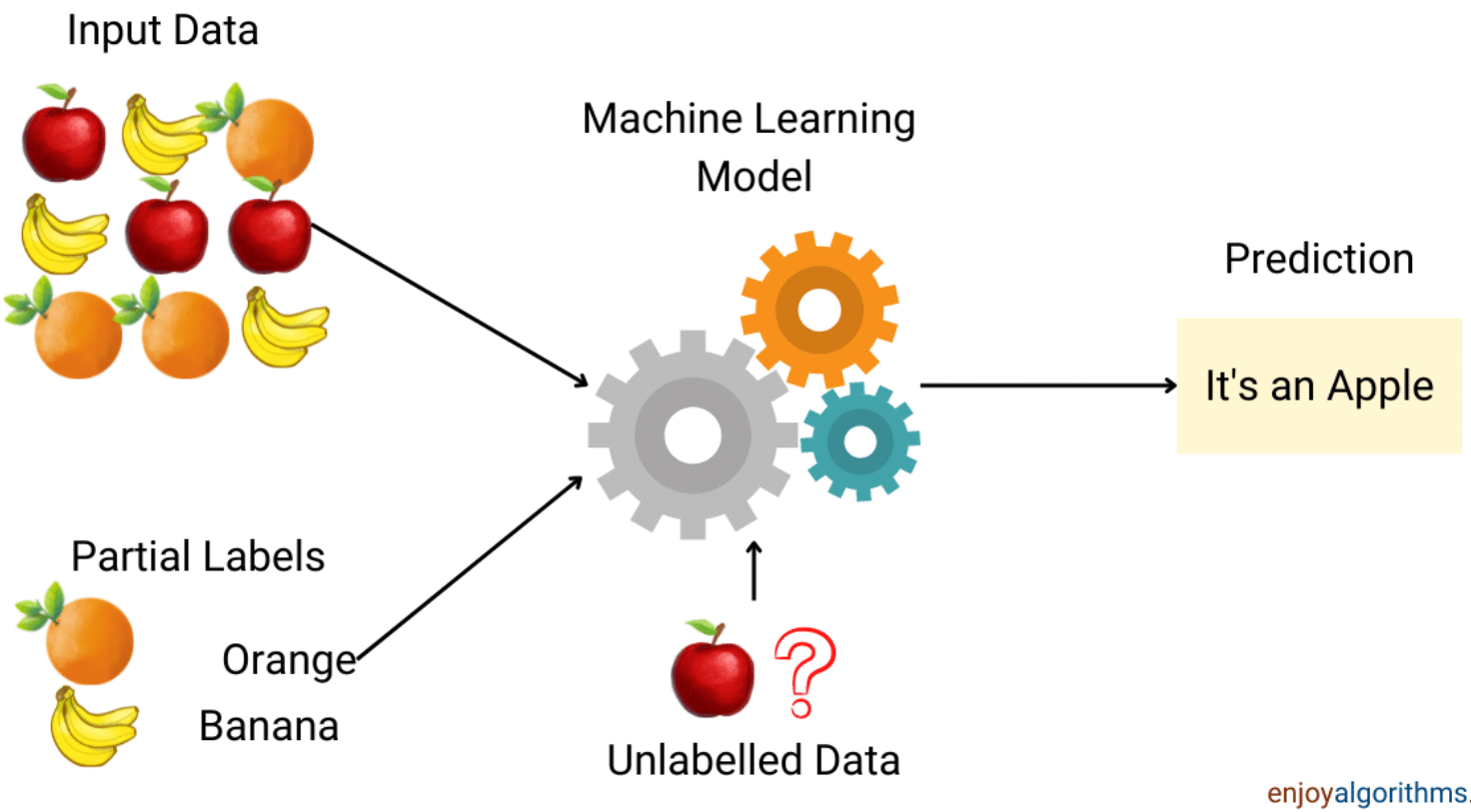
- the predictive or supervised learning approach.
- the descriptive or unsupervised learning approach.


Supervised Learning
enjoy algorithms
Labeled Data
Machine
ML Model
Predictions
Triangle
Circle
Labels
Rectangle
Circle
Triangle
Hexagon
Test Data

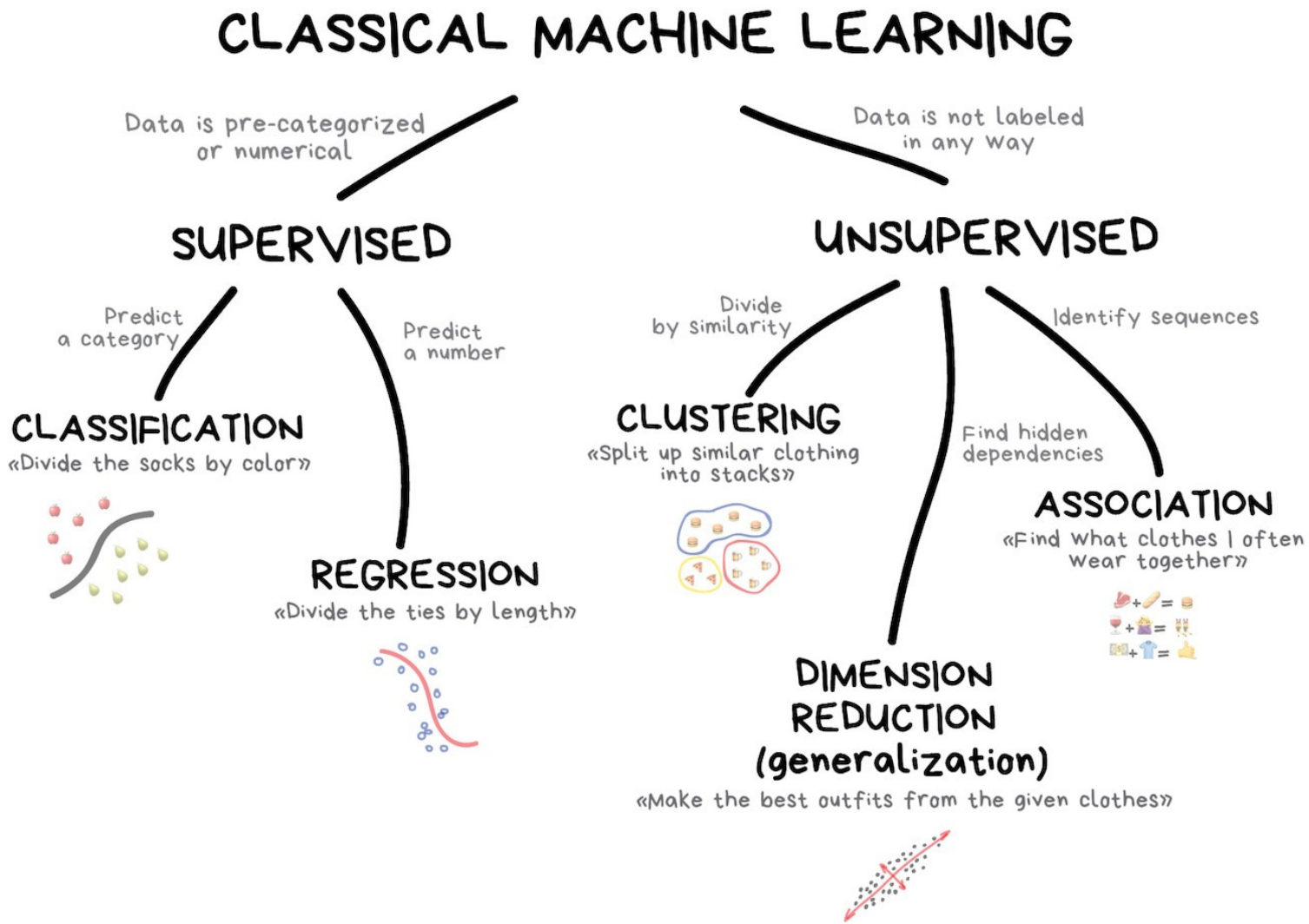
Unsupervised Learning



Semi-supervised learning



Types of ML


CLASSICAL MACHINE LEARNING
Data is pre-categorized or numerical
Data is not labeled in any way
SUPERVISED
UNSUPERVISED
Predict a category
Predict a number
Divide by similarity
Identify sequences
CLASSIFICATION
«Divide the socks by color»
REGRESSION
«Divide the ties by length»
CLUSTERING
«Split up similar clothing into stacks»
Find hidden dependencies
ASSOCIATION
«Find what clothes I often wear together»
DIMENSION REDUCTION (generalization)
«Make the best outfits from the given clothes»

Thank you for your attention