Lecture 1:
General
introduction

Sophie Robert

General
definitions

Classification
problems

Regression
problems

Clustering
problems

Exercises

Lecture 1: General introduction
Introduction to Machine Learning

Sophie Robert

L3 MIASHS | Semestre 2

2023-2024

Lecture 1:
General
introduction

Sophie Robert

General
definitions

Classification
problems

Regression
problems

Clustering
problems

Exercises

General definitions

What is Machine Learning ?

Machine Learning algorithms

Algorithms able to **learn** and **adapt** without following explicit instructions by **drawing inferences from patterns in data**.

What is Machine Learning ?

Machine Learning algorithms

Algorithms able to **learn** and **adapt** without following explicit instructions by **drawing inferences from patterns in data**.

Given a **training** dataset, Machine Learning* algorithms are able to **find patterns in data** to **predict** or **infer** information on new data.

What is Machine Learning ?

Machine Learning algorithms

Algorithms able to **learn** and **adapt** without following explicit instructions by **drawing inferences from patterns in data**.

Given a **training** dataset, Machine Learning* algorithms are able to **find patterns in data** to **predict** or **infer** information on new data.

Question

Can you give me some examples of Machine Learning models you have previously studied ?

What is Machine Learning ?

Lecture 1:
General
introduction

Sophie Robert

General
definitions

Classification
problems

Regression
problems

Clustering
problems

Exercises

Problems can usually be divided into two main types:

What is Machine Learning ?

Lecture 1:
General
introduction

Sophie Robert

General
definitions

Classification
problems

Regression
problems

Clustering
problems

Exercises

Problems can usually be divided into two main types:

Problems can usually be divided into two main types:

- **Supervised learning***: the algorithm **should learn** from example data to predict the value for **unseen data**.

What is Machine Learning ?

Problems can usually be divided into two main types:

- **Supervised learning***: the algorithm **should learn** from example data to predict the value for **unseen data**.
    - **Classification* problems**: a **class** (**categorical variable***) is predicted

What is Machine Learning ?

Lecture 1:
General
introduction

Sophie Robert

General
definitions

Classification
problems

Regression
problems

Clustering
problems

Exercises

Problems can usually be divided into two main types:

- **Supervised learning***: the algorithm **should learn** from example data to predict the value for **unseen data**.
    - **Classification* problems**: a **class** (**categorical variable***) is predicted
    - **Regression* problems**: a **metric** (**numerical variable***) is predicted

What is Machine Learning ?

Lecture 1:
General
introduction

Sophie Robert

General
definitions

Classification
problems

Regression
problems

Clustering
problems

Exercises

Problems can usually be divided into two main types:

- **Supervised learning***: the algorithm **should learn** from example data to predict the value for **unseen data**.
  - **Classification* problems**: a **class** (**categorical variable***) is predicted
  - **Regression* problems**: a **metric** (**numerical variable***) is predicted
- **Unsupervised learning***: the algorithm **should find some patterns in the data** to provide a better understanding of the data.

What is Machine Learning ?

Problems can usually be divided into two main types:

- **Supervised learning***: the algorithm **should learn** from example data to predict the value for **unseen data**.
    - **Classification* problems**: a **class** (**categorical variable***) is predicted
    - **Regression* problems**: a **metric** (**numerical variable***) is predicted
- **Unsupervised learning***: the algorithm **should find some patterns in the data** to provide a better understanding of the data.

We will study both in this course !

Classification problems

Example of classification problems include:
Handwritten number recognition (OCR)
Given pixel repartition, learn to match handwritten numbers
with their true value.

Flower species

Given flowers characteristics, predict their species.

iris setosa   iris versicolor   iris virginica

petal   sepal   petal   sepal   petal   sepal

Regression problems

Price predictions
Given house characteristics (size, location, number of bedrooms), predict the selling prices of houses.

Clustering problems

Infer Netflix watching behavior
Given a set of user characteristics (time spent per day, number of different items watched, number of series and movies), group users into subgroups.

Exercises

Given their size, their weight, and their color, predict the fish specie.
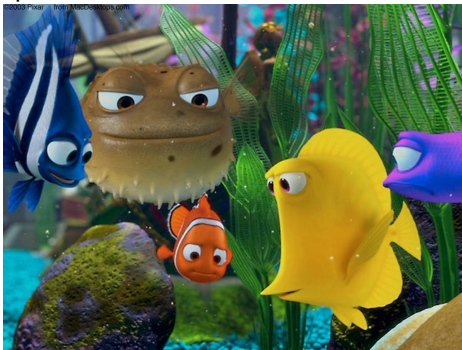
Can you infer the type of problem ?

Given their previous results (number of race won, previous times, total training times …), predict athletes time in a race.

Given their medical records (weight, age, height, smoking status), predict patients with diabetes.

Can you infer the type of problem ?

Given their spending habit (number of bought items, total money spent …), identify different shopper behaviors.

Given tumors characteristics (size, height, width, color ...),
predict if they are cancerous tumors.
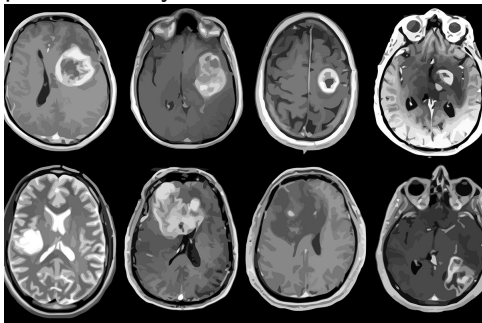
Given baby whales characteristics (specie, location, parents ...),
predict their total adult weights.

Questions ?

Questions ?