

## • Integration

→ Both methods can be used.

For instance, unsupervised learning can be used to uncover patterns or group data, which can then be labeled for supervised learning.

## Practice cases:

- A conservation organization has deployed cameras in a wildlife reserve. " - "

- supervised learning (because the data is already labeled)

- classification (because the goal is to identify different animal species)

- A music streaming service. " - "

- unsupervised learning (the data isn't labeled)

- clustering (categorizing songs into different genres using audio features of the tracks)

- A marketing firm. " - "

- ~~unsupervised~~ learning (unlabeled data)

- association (discovering the "rules" that



govern these emerging trends in order to label them)

- A manufacturing firm. " - "
- ~~structured data~~ <sup>supervised learning</sup> (the data is known and a lot of additional information already exists, therefore it could be considered labeled)
- ~~regression~~ classification (categorize data into pass or fail)
- A tech company is developing a language translation tool. " - "
- supervised learning (labeled data → <sup>bilingual</sup> corpus)
- classification (categorizing words ~~as nouns~~ as verbs, nouns, etc., grammar rules, ...)

## Basics of model building

- 1) Define the problem
  - understand the problem
  - set objectives
- 2) Collect and Prepare data
  - Data collection → scraping websites, querying databases, using APIs, etc.
  - Data cleaning → handling missing values, removing duplicates, fixing errors.
  - Data preprocessing → convert data into usable format