Machine Learning

Lecture 1

10/18/22

Class Rules

- Respect to be respected.
- 75% attendance mandatory
- In case of any questions or setting up a meeting outside office hours please **EMAIL** at **Saniya.Ashraf@buitms.edu.pk**
- **Late assignment policy:** 2 late days allowed with10% penalty per day. Not acceptable after that.

- Sessionals are divided as follows
  - 5 marks for Behavior and Class ethics
  - 20 marks for assignments/presentations and **surprise** quizzes
  - Sessionals gained are **ABSOLUTE**. If by the end of the semester you are missing a grade by even a single score, it will **NOT** be entertained so plan accordingly. Those students who do not miss **any assignments** might be given this exception.

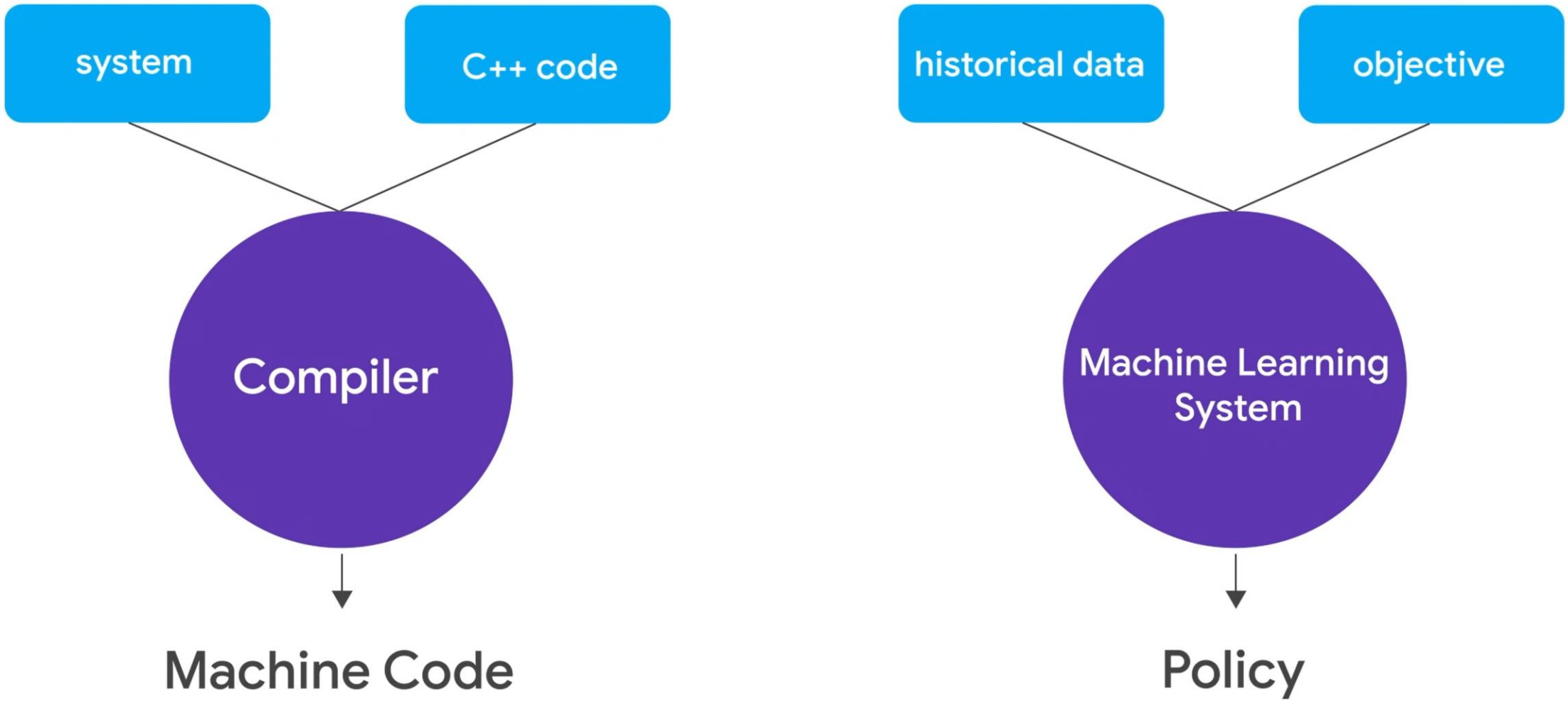
What is machine learning?

- What would you do if you have to predict rainfall?
- Traditional programming approach: create a physics-based model of the Earth's atmosphere and surface, computing massive amounts of fluid dynamics equations
- Problem? Extremely difficult, often tedious, and prone to errors.
- ML approach: feed the ML model large amounts of weather data until the ML model eventually ***learns*** the mathematical relationship between weather patterns that produce differing amounts of rain.
- Then feed it the current data and it will predict the amount of rain.

What is machine learning?

- ML is the process of training a piece of software, called a **model**, to make useful predictions from data.
- An ML model represents the **mathematical** relationship between the elements of data that an ML system uses to make predictions.

Compilers & Machine Learning



What is machine learning?

- ML systems learn

  How to combine input

  to produce useful predictions

  on never-seen-before data

Machine learning advantages

- Reduce time spent on programming.
  - Example: a program to correct spelling errors
- Customize finished software products to make them better for specific groups of people.
  - Example: expanding the program for each language.
- ML lets you solve programs that you, as a programmer, have no idea how to do by hand.
  - The program to recognize faces.
- ML allows things to be possible that at a point no one has ever thought of.
  - www.thispersondoesnotexist.com

Types of ML Systems

- ML systems fall into three distinct categories based on how they learn to make predictions:
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning

Supervised Machine learning

- Supervised learning models can make predictions after seeing lots of data with the correct answers and then discovering the connections between the elements in the data that produce the correct answers. This is like a student learning new material by studying old exams that contain both questions and answers. Once the student has trained on enough old exams, the student is well prepared to take a new exam. These ML systems are “supervised” in the sense that a human gives the ML system data with the known correct results.

- Two of the most common use cases for supervised learning are regression and classification.

Regression vs. Classification

- A **regression** model predicts continuous values. For example, regression models make predictions that answer questions like the following:
- What is the value of a house in California?
- What is the probability that a user will click on this ad?
- A **classification** model predicts discrete values. For example, classification models make predictions that answer questions like the following:
- Is a given email message spam or not spam?
- Is this an image of a dog, a cat, or a hamster?

Regression vs. Classification

enjoy algorithms

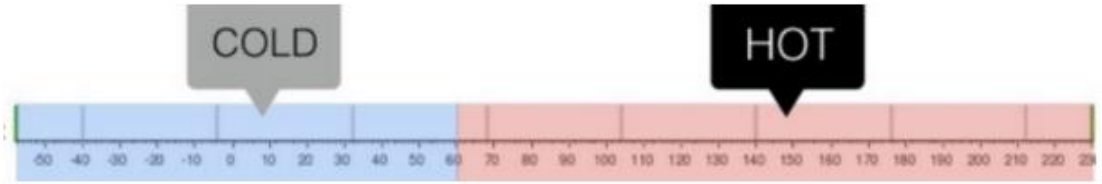
Regression

What will be the temperature tomorrow?

Fahrenheit

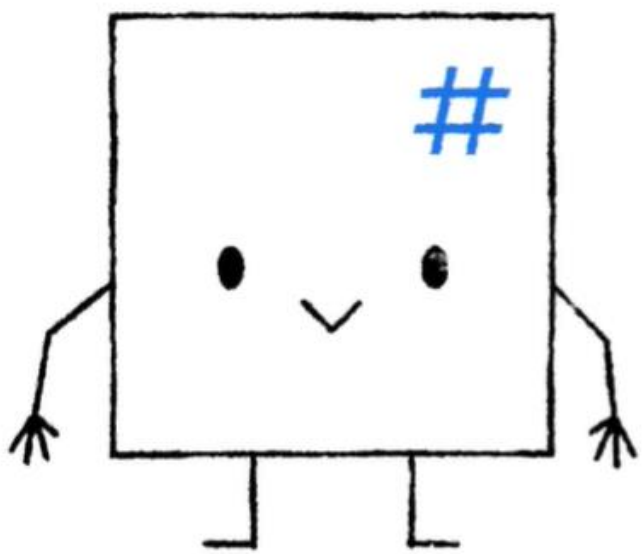
Classification

Will it be hot or cold tomorrow?

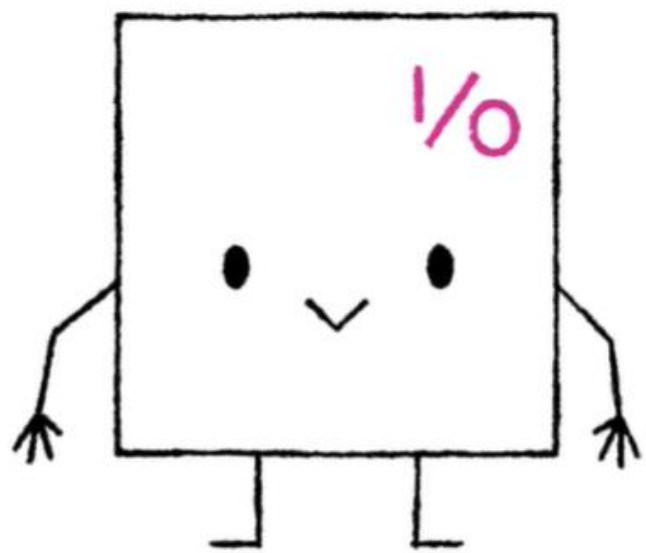


Fahrenheit

Regression



Binary Classification

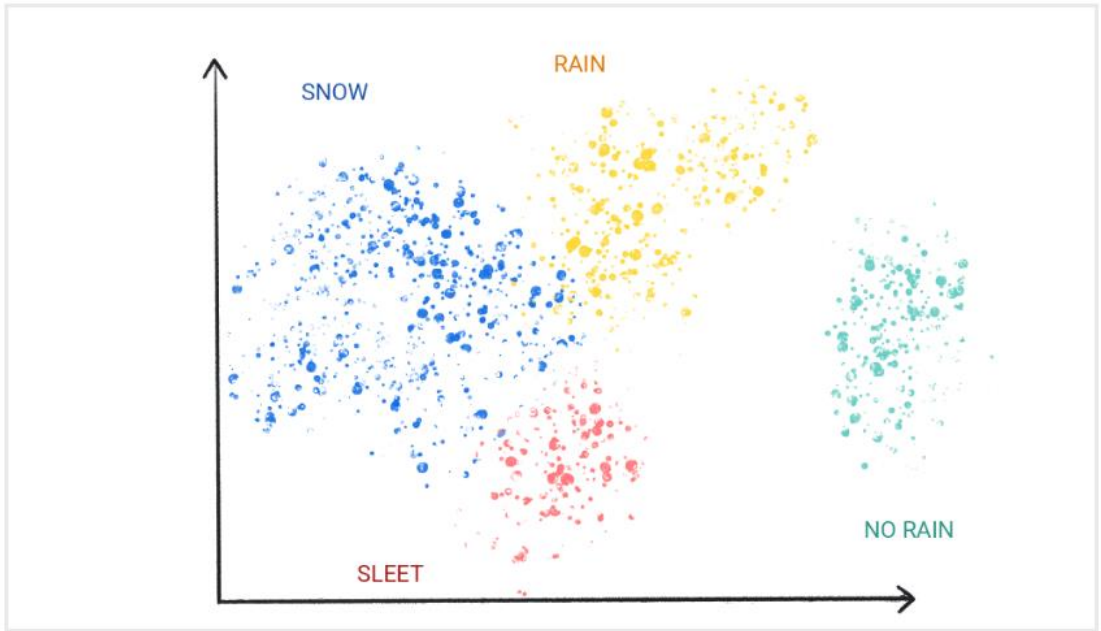


Multiclass Classification

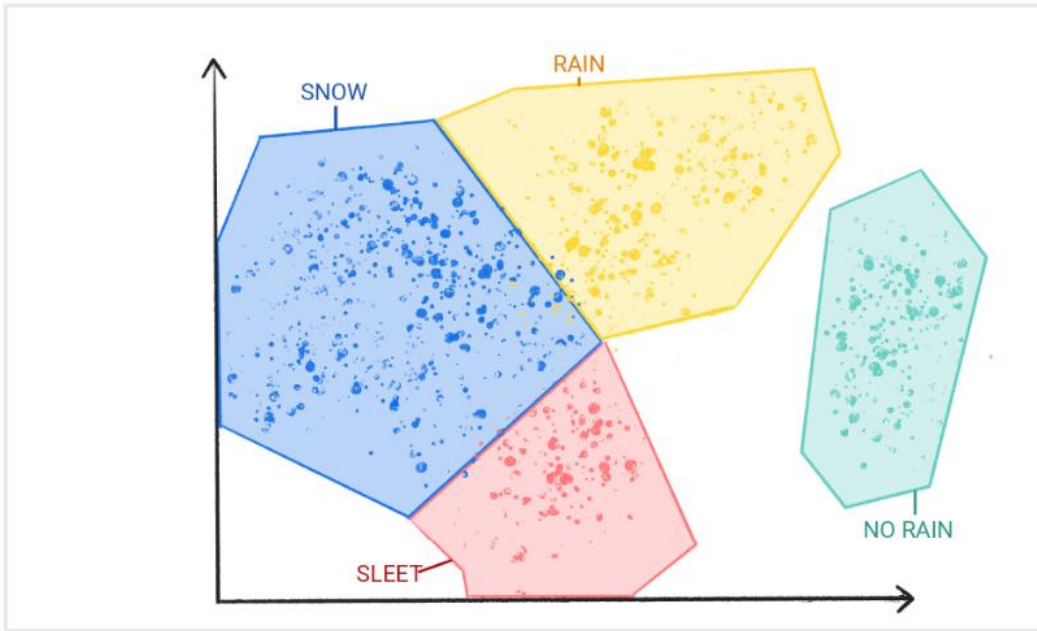
Unsupervised learning

- Unsupervised learning models make predictions by given data that does not contain any correct answers. An unsupervised learning model's goal is to identify meaningful patterns among the data. In other words, the model has no hints on how to categorize each piece of data, but instead, it must infer its own rules.

- A commonly used unsupervised learning model employs a technique called clustering. The model finds data points that makes boundaries around natural groupings.



An ML model clustering similar weather patterns.

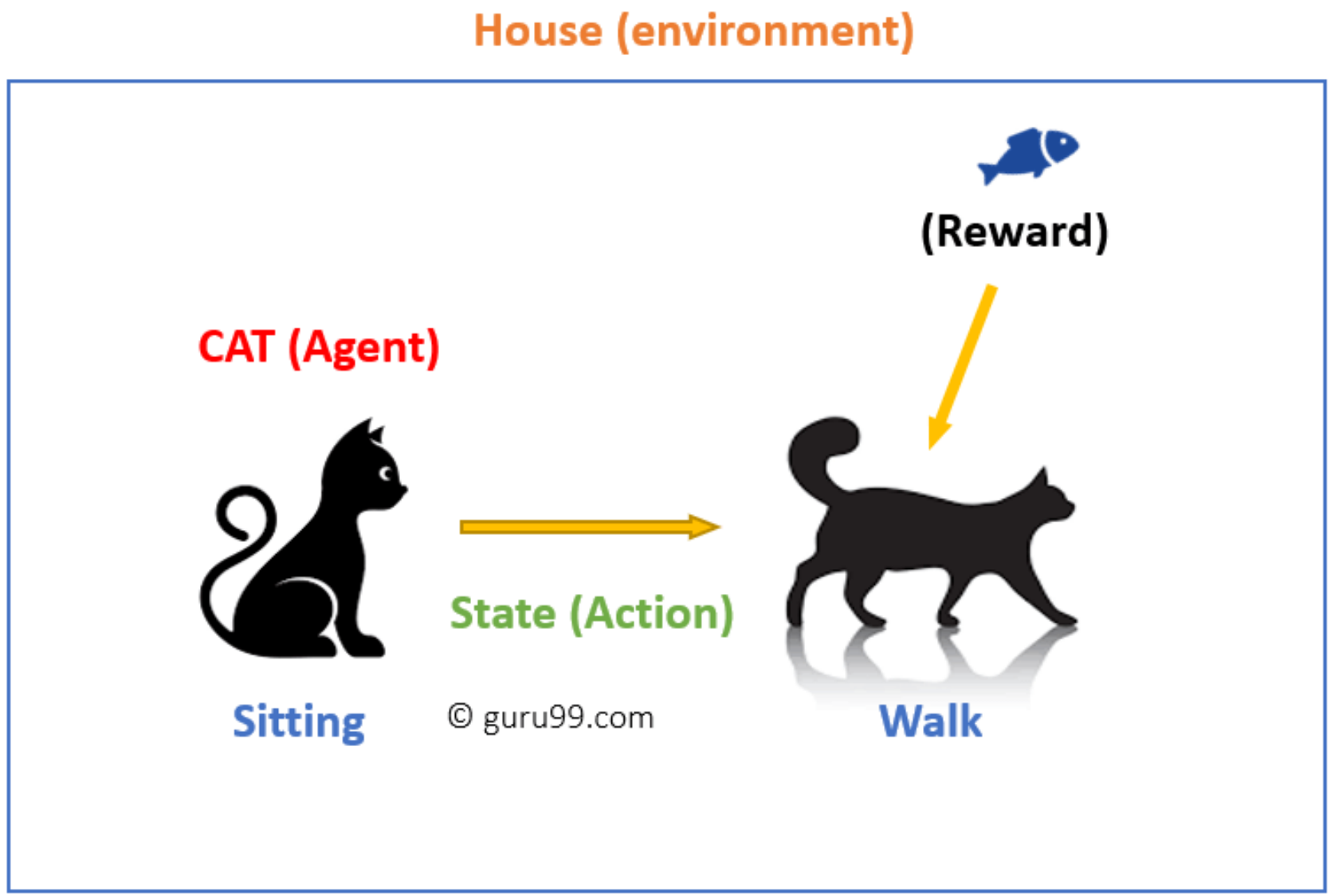


Clusters of weather patterns labeled as snow, sleet, rain, and no rain.

Reinforcement learning

- Reinforcement learning models make predictions by getting rewards or penalties based on actions performed within an environment. A reinforcement learning system generates a policy that defines the best strategy for getting the most rewards.

- Reinforcement learning is used to train robots to perform tasks, like walking around a room, and software programs like AlphaGo to play the game of Go.


House (environment)
(Reward)
CAT (Agent)
State (Action)
Sitting
© guru99.com
Walk

Fundamental machine learning terminology.

- **Labels**
  - A label is the thing we're predicting—the y variable. The label could be the future price of wheat, the kind of animal shown in a picture, the meaning of an audio clip, or just about anything.
- **Features**
  - A feature is an input variable—the x variable. A simple machine learning project might use a single feature, while a more sophisticated/complex machine learning project could use millions of features i-e., x_1,x_2,x_3…x_N

Binary classification and Multiclass classification

- **Binary classification** models output a value from a class that contains only two values, for example, a model that outputs either rain or no rain. 0/1
- **Multiclass classification** models output a value from a class that contains more than two values, for example, a model that can output either rain, hail, snow, or sleet

Fundamental machine learning terminology.

- **Examples** An **example** is a particular instance of data, **x**. (We put **x** in boldface to indicate that it is a vector.) We break examples into two categories:
  - labeled examples
  - unlabeled examples
- A **labeled example** includes both feature(s) and the label (x, y). We use labeled examples to **train** the model.

Labeled examples

housingMedianAge (feature)	totalRooms (feature)	totalBedrooms (feature)	medianHouseValue (label)
15	5612	1283	66900
19	7650	1901	80100
17	720	174	85700
14	1501	337	73400
20	1454	326	65500

Fundamental machine learning terminology.

- **unlabeled examples:** contains features but not labels.
- **Models:** A model defines the relationship between features and label. For example, a spam detection model might associate certain features strongly with "spam". Let's highlight two phases of a model's life:

- **Training** means creating or learning the model. That is, you show the model labeled examples and enable the model to gradually learn the relationships between features and label.

- **Inference** means applying the trained model to unlabeled examples. That is, you use the trained model to make useful predictions (y'). For example, during inference, you can predict medianHouseValue for new unlabeled examples.

Assignment 1: Due on Thursday 29/9/2022

- Download the Assignment Instructions here.
- You must submit a PRINTED REPORT IN CLASS it will not be accepted otherwise.
- Use LATEX if you can. Acceptable in word as well.