AI & Robotics

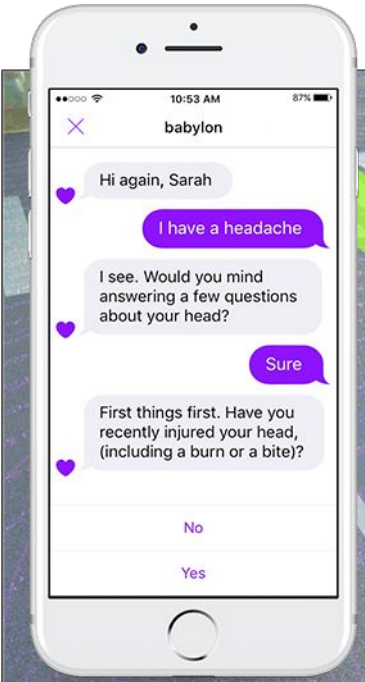
Supervised Learning

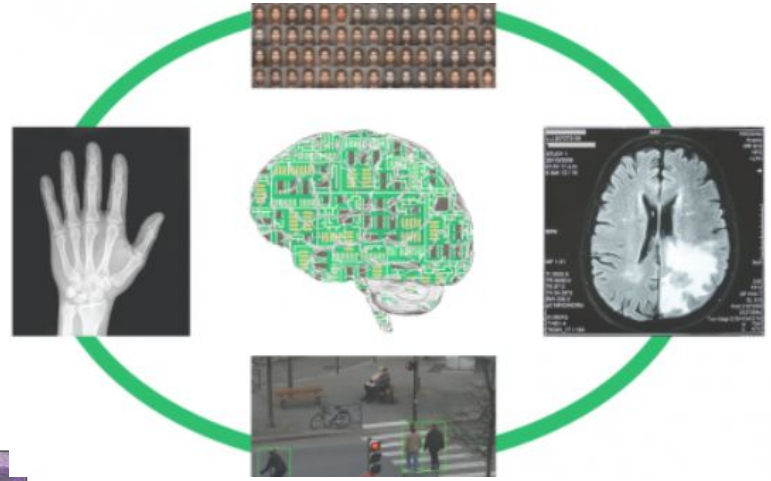
PXL AI & ROBOTICS LAB

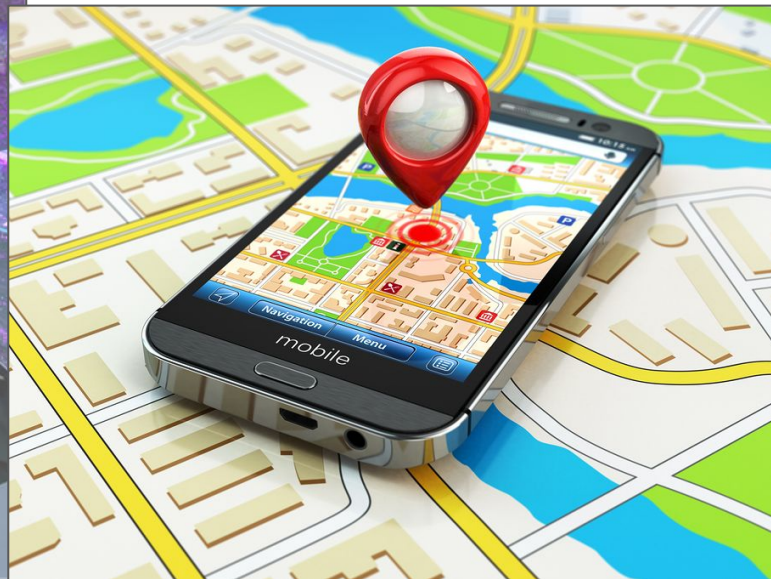
Goals

The **junior-colleague**

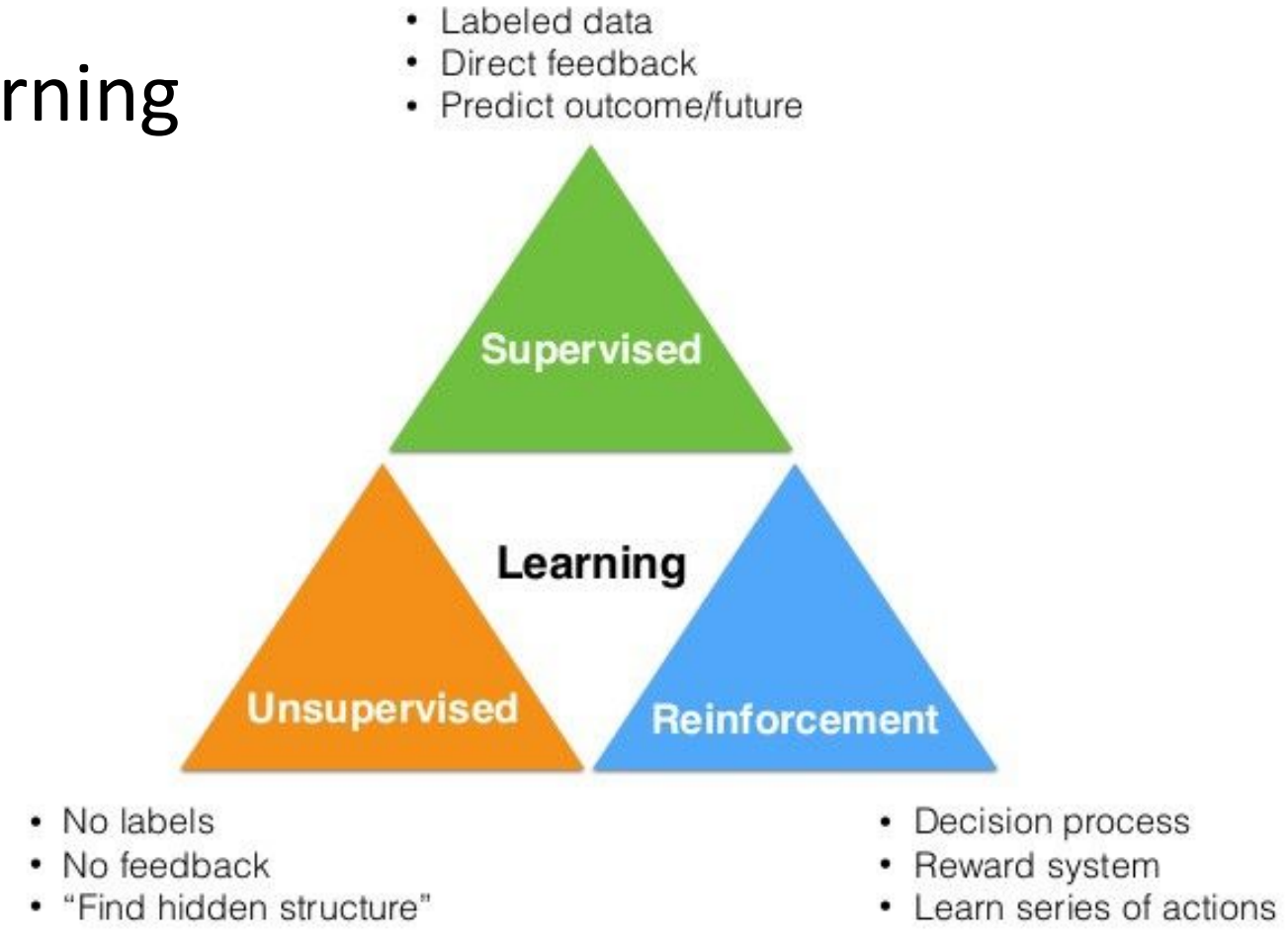
- can explain Supervised Learning in their own words
- can describe the general flow of a supervised learning pipeline
- can explain the difference between classification and regression
- can explain the difference between linear and nonlinear spread of classification examples
- can explain the difference between structured and unstructured data


10:53 AM
87%
babylon
Hi again, Sarah
I have a headache
I see. Would you mind answering a few questions about your head?
Sure
First things first. Have you recently injured your head, (including a burn or a bite)?
No
Yes

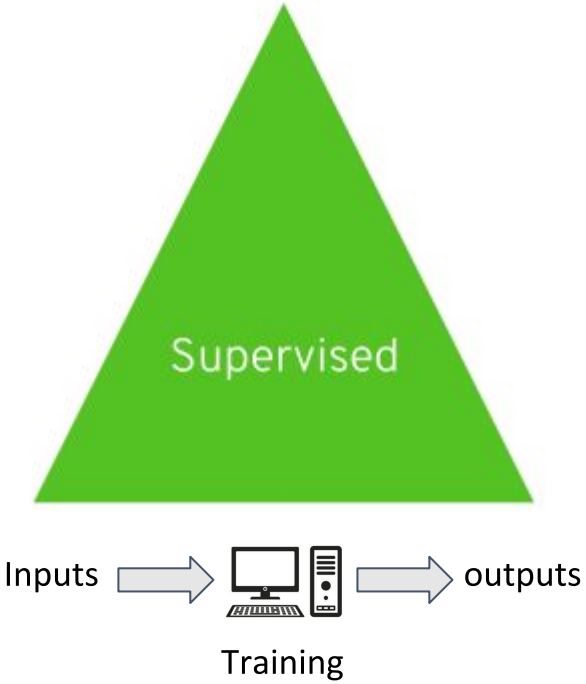



Navigation
Menu
mobile

Machine Learning

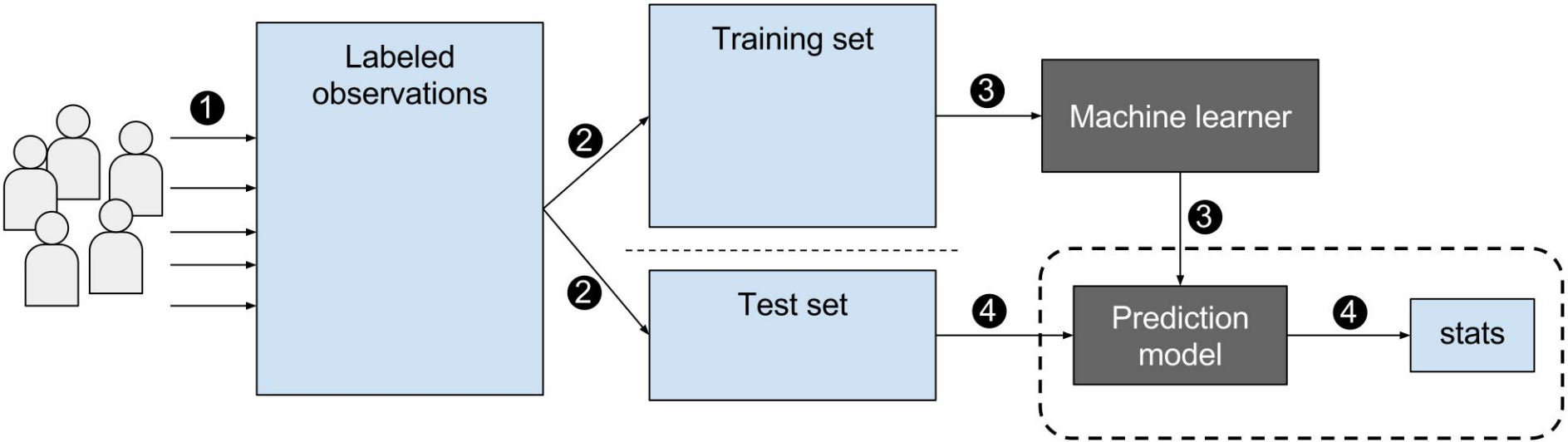


Supervised Learning



- Labeled Data
- Direct feedback
- Predict outcome
- Meaningful patterns in labeled data
- Most commonly used
- Most ready for real-world

Supervised Learning



Supervised Learning idea

- Learning a function that maps an input to an output based on example input-output pairs
- Learn f: X -> y

  With:

  X: a feature vector of independent variables (x_0, …, x_k)

  y: the dependent variable (the variable we are trying to predict

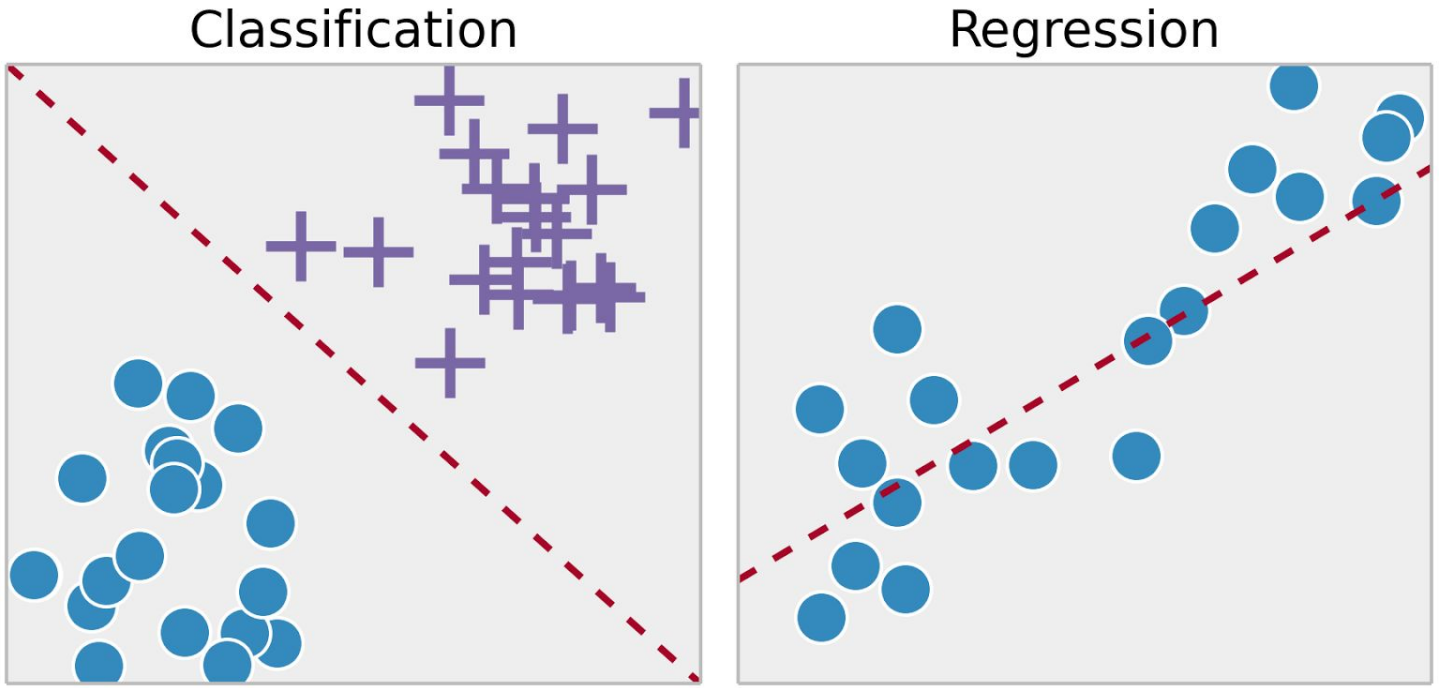
From a bunch of training samples of the form {(X_0, y_0), …, (X_n, y_n)}

Variables are also often called attributes or features. We will use these terms interchangeably throughout the course.

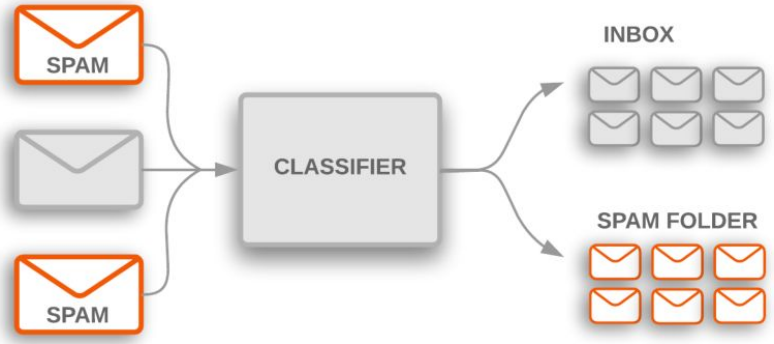
Supervised Learning examples

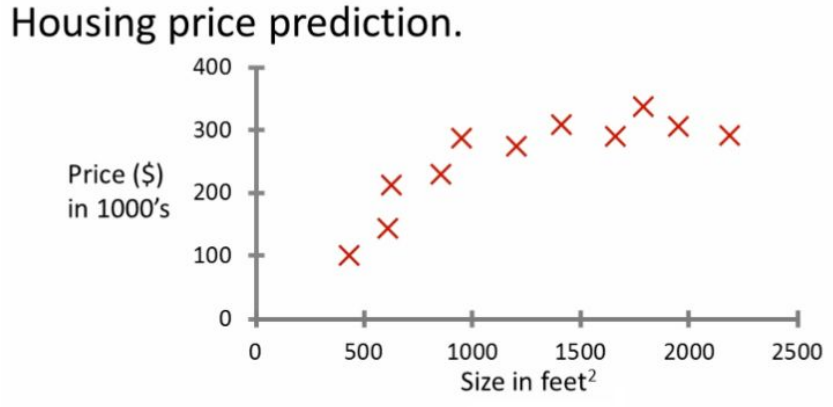
- Predicting sale price based on historical data containing attributes like dates, quantities, resources, etc.
- Spam filtering based on email addresses, words in the subject or contents of the mail, etc.
- Demand forecasting
- Facial recognition
- Self-driving cars
- ...

Classification vs. Regression

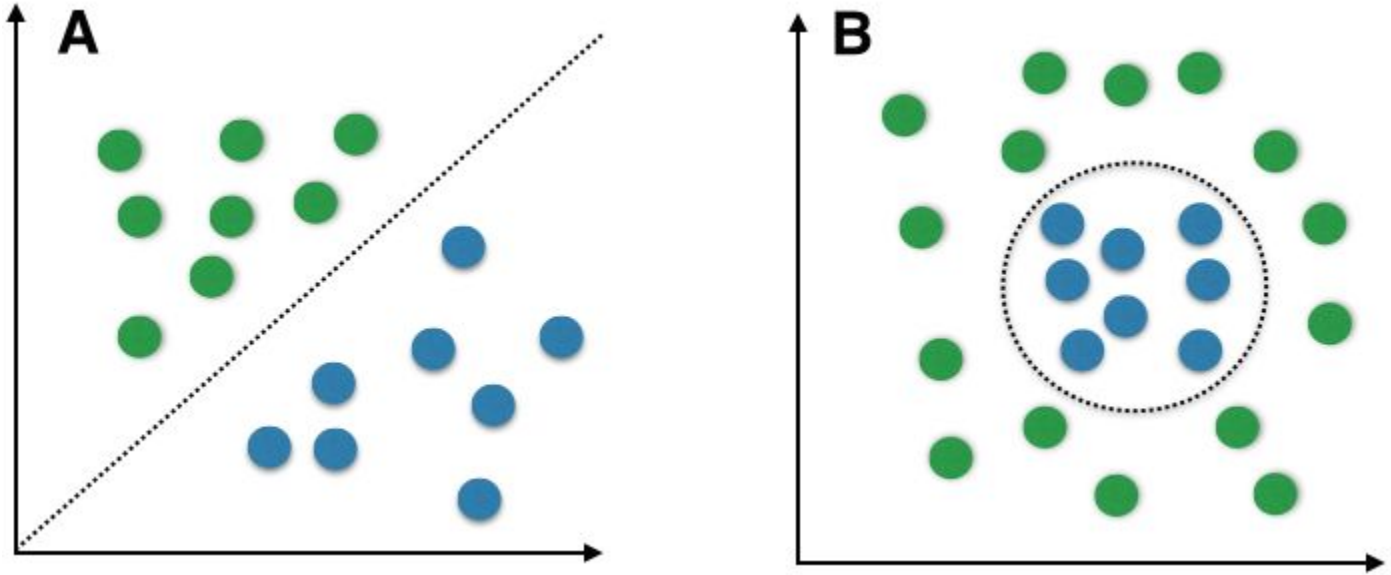


Classification vs. Regression

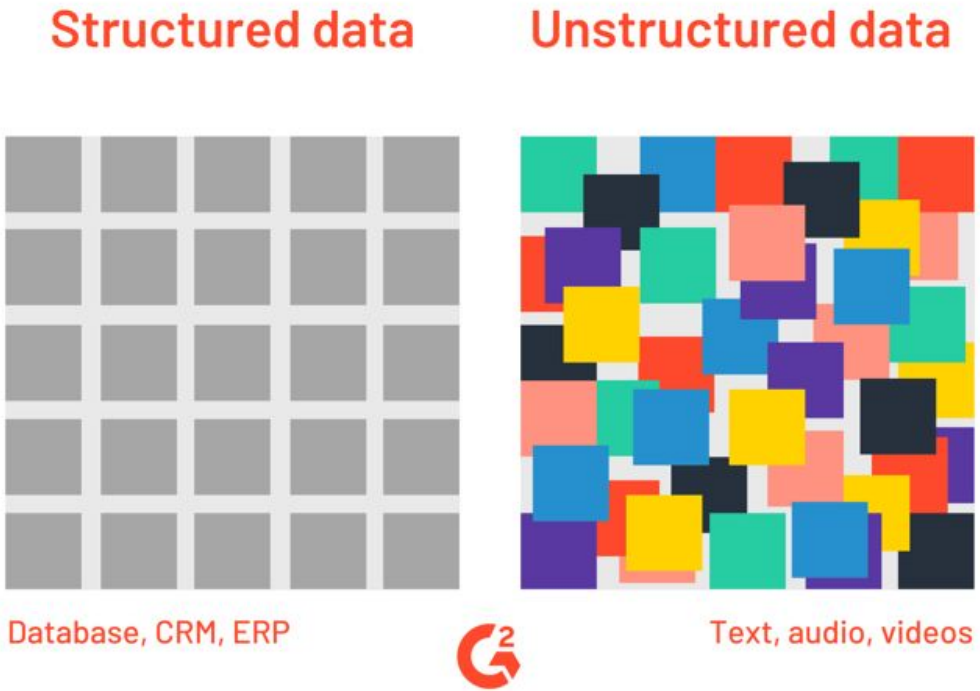




Linear vs. non-linear



Structured vs. unstructured data



Structured vs. unstructured data

	Structured Data	Unstructured Data
Characteristics	• Pre-defined data models • Usually text only • Easy to search	• No pre-defined data model • May be text, images, sound, video or other formats • Difficult to search
Resides in	• Relational databases • Data warehouses	• Applications • NoSQL databases • Data warehouses • Data lakes
Generated by	Humans or machines	Humans or machines
Typical applications	• Airline reservation systems • Inventory control • CRM systems • ERP systems	• Word processing • Presentation software • Email clients • Tools for viewing or editing media
Examples	• Dates • Phone numbers • Social security numbers • Credit card numbers • Customer names • Addresses • Product names and numbers • Transaction information	• Text files • Reports • Email messages • Audio files • Video files • Images • Surveillance imagery