DATA 1202

Spring 2024

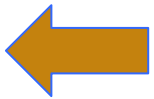
Lecture 25

Classification

Announcements

- **Homework 13** due Wednesday 4/17
- **Project 2** due Friday 4/19

Course Content

1. Computation: Python and tables
2. Describing data
    a. By visualizing
    b. By quantifying
3. Probability
4. Inference
5. Prediction 

Prediction

Guessing the Value of an Attribute

- Based on incomplete information
- One way of making predictions:
  - To predict an outcome for an individual,
  - find others who are like that individual
  - and whose outcomes you know.
  - Use those outcomes as the basis of your prediction.

- Two Types of Prediction
  - Regression = Numeric; Classification = Categorical

Prediction Example: Spam or Not?

You made a Wells Fargo payment - wellsfargo.com You recently submitted a payment The ...

BUSINESS TRUST - -- I have a legal business proposal for you worth $23,000,000. If you kn...

Hi - Today???!!!! What a wonderful day! Congrats again! I am definitely not doing s...

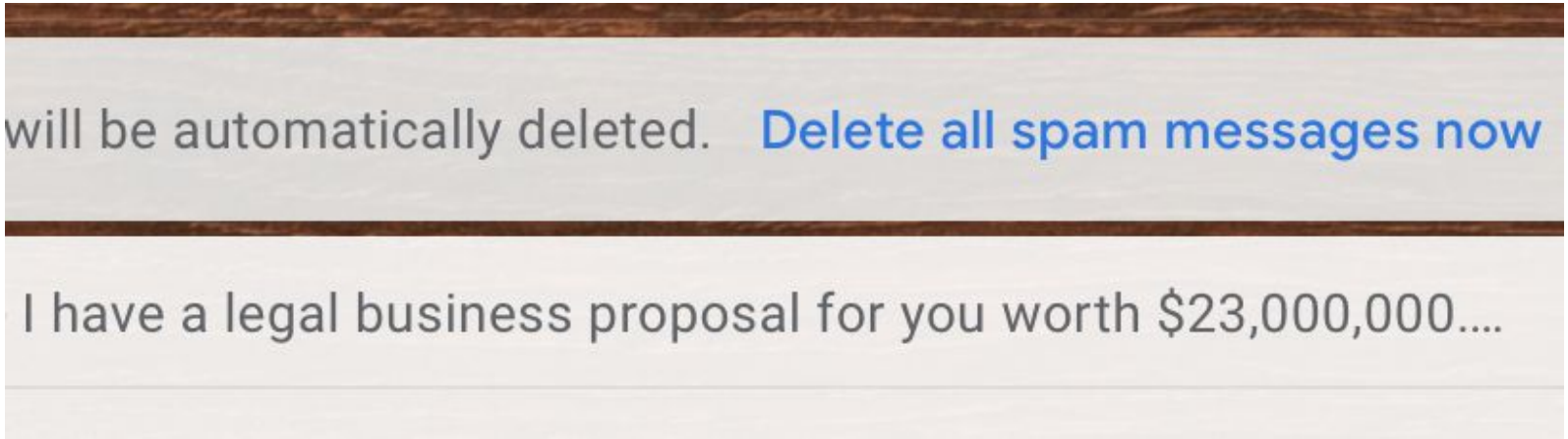
Michael Kors Handbags Up To 84% Plus Free Shipping! - Shop Handbags Online & In Store...

Machine Learning Algorithm

- A mathematical model
- calculated based on sample data
  - "training data"
- that makes predictions or decisions without being explicitly programmed to perform the task

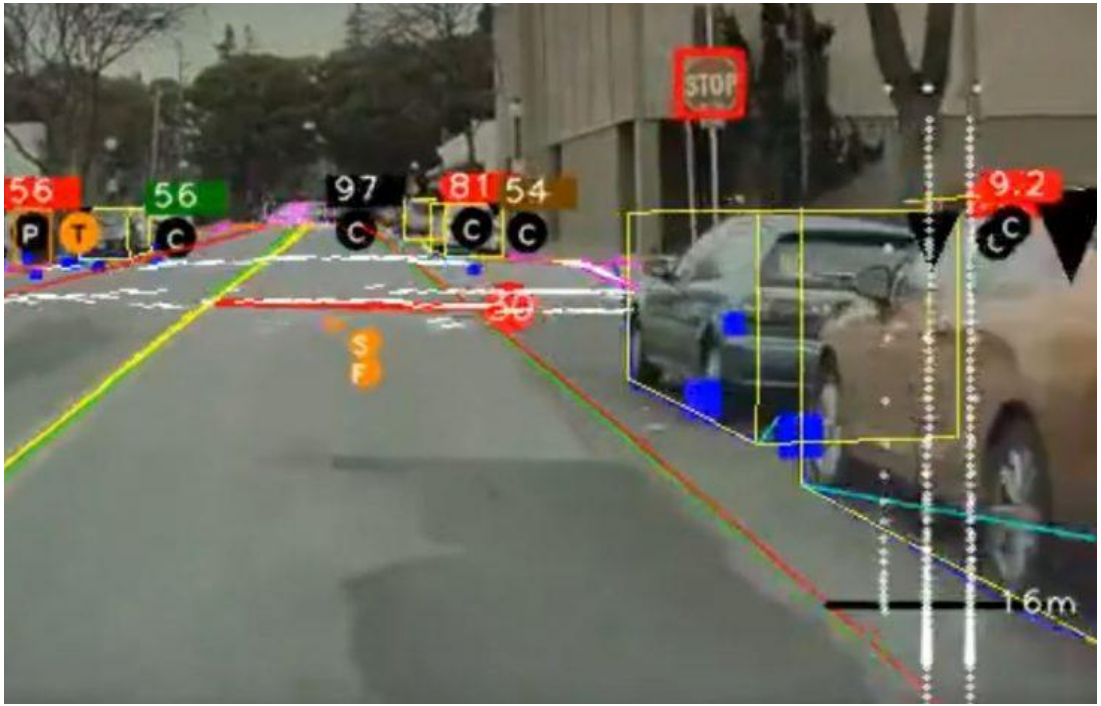
Classification

Classification Examples: Text



Output: (Spam, Not Spam)

Classification Examples: Image



Output:

(Car,
Road,
Tree,
Sky,
Traffic Sign)

Classification Examples: Video

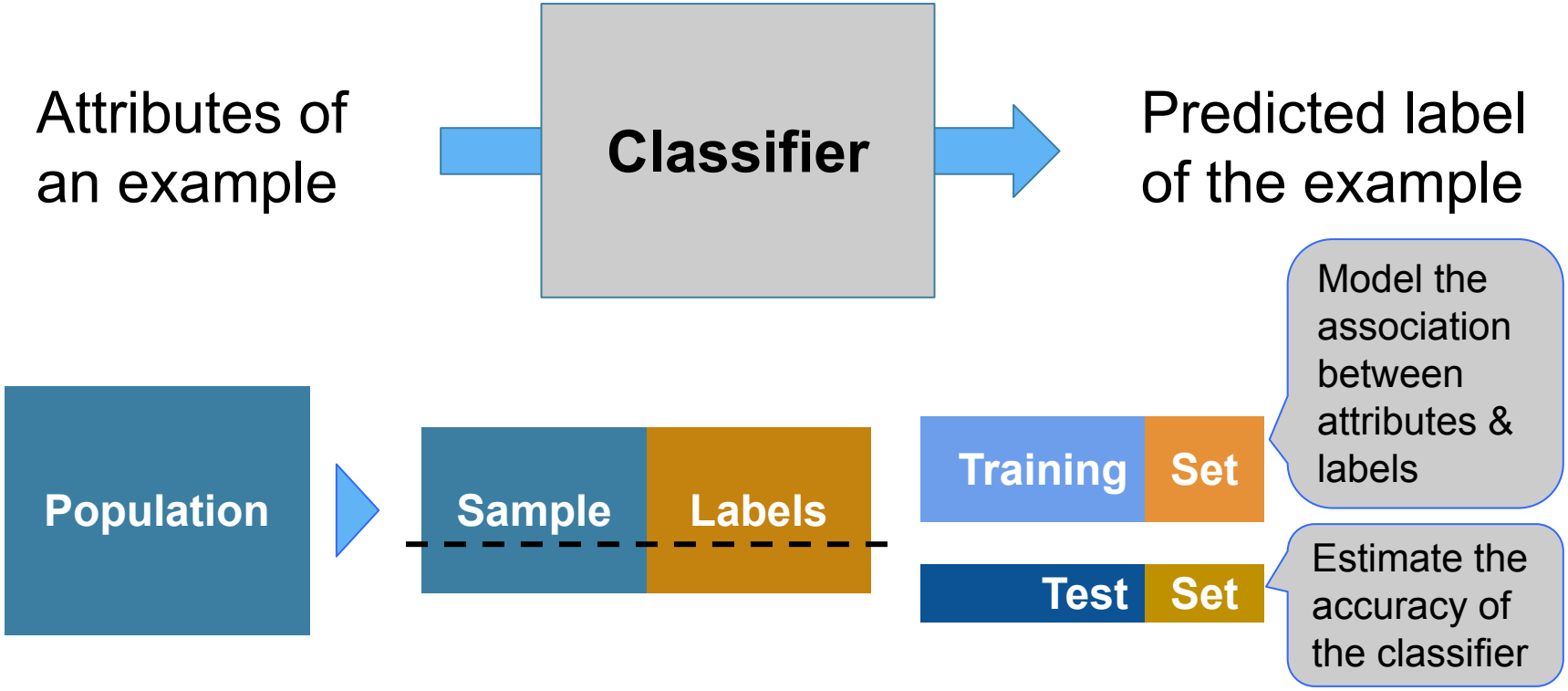
Output:

(In, Out)

(Demo)

Classifiers

Training a Classifier



Nearest Neighbor Classifier