

DS2000

4/7 - Fri :)

Admin

- HW8 due 9pm

Agenda

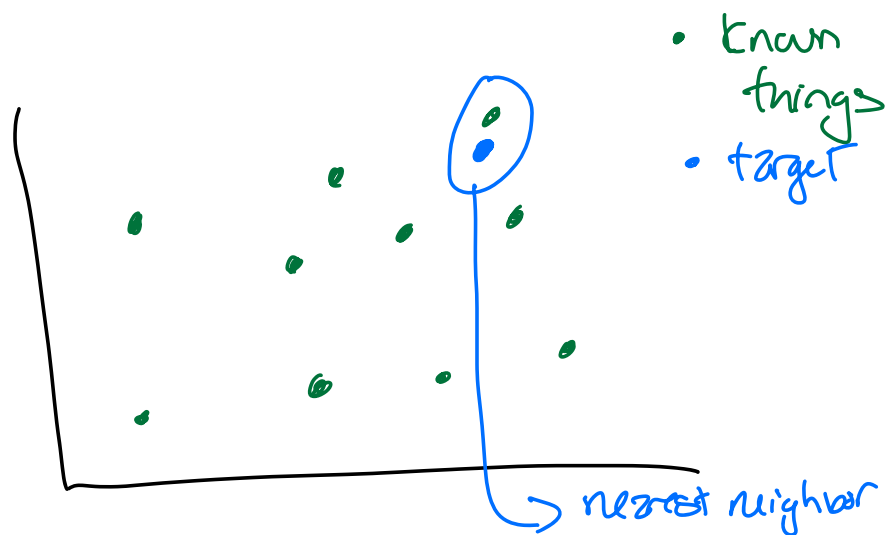
1. K-nearest neighbor classification
  2. Python
-

# 1. K nearest neighbor Classification

What we know: nearest neighbor algorithm

- target thing: what we want to know about
- known things: things we already know, want to find most similar

nearest  
neighbor  
algo



nearest neighbor → knn classifier

- things that can be categorized
- start with known things
- guess the category of unknown thing

machine learning

↓  
predictive model

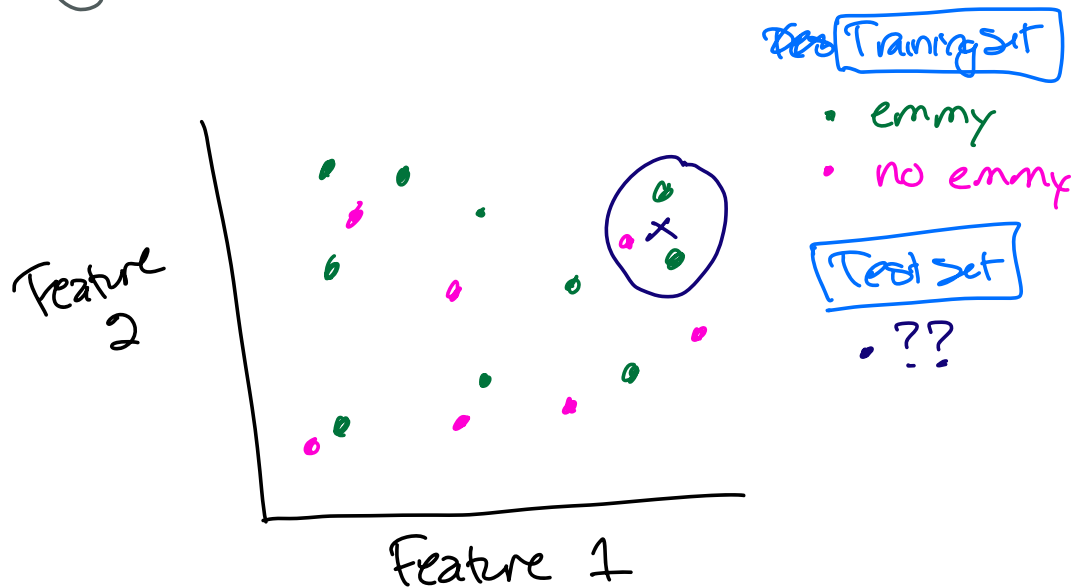
today ~ KNN classifier for TV Shows

- training set: know features and categories

data about  
each item

(label  
(emmy))

- testing set: know features, not category



$k$  = # of nearest neighbors  
(here,  $k = 3$ )

## Today's Data

Features: violence, drama, funny (1-5)

Category: enemy (no enemy)