

DS2000

4/7- Fri

Admin

- HW8 due 9pm

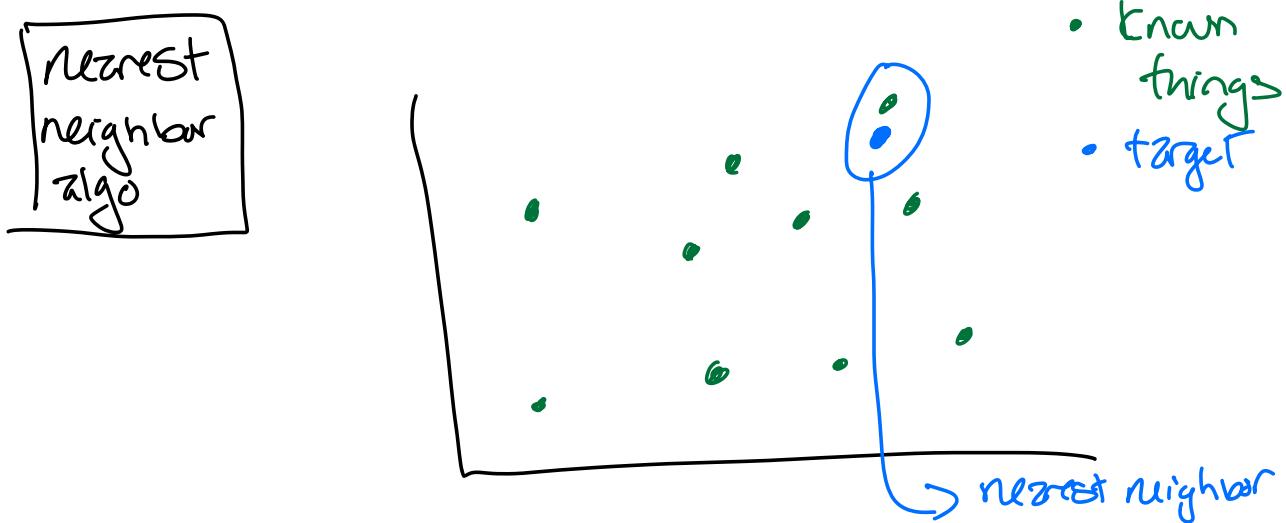
Agenda

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 \rightarrow kNN classifier

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

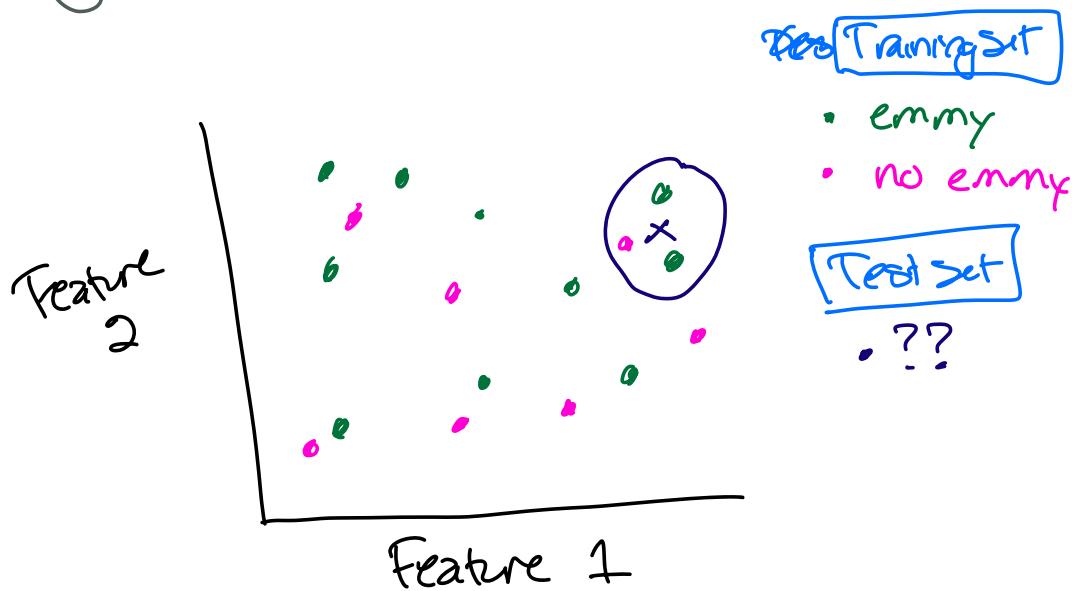
Machine learning



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



$K = \# \text{ of nearest neighbors}$
(here, $K = 3$)

Today's Data

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

Category: emmy (no emmy)