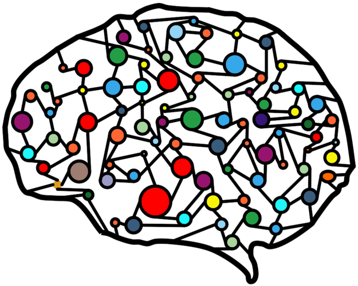
Lecture 23 – Conditional Independence, Naive Bayes



DSC 40A, Fall 2022 @ UC San Diego

Mahdi Soleymani, with help from **many others**

Agenda

- Classification and conditional independence.
- Naive Bayes.

Classification and conditional independence

Example: avocados, but with more features

color	softness	variety	ripeness
bright green	firm	Zutano	unripe
green-black	medium	Hass	ripe
purple-black	firm	Hass	ripe
green-black	medium	Hass	unripe
purple-black	soft	Hass	ripe
bright green	firm	Zutano	unripe
green-black	soft	Zutano	ripe
purple-black	soft	Hass	ripe
green-black	soft	Zutano	ripe
green-black	firm	Hass	unripe
purple-black	medium	Hass	ripe

You have a firm green-black Zutano avocado. Based on this data, would you predict that your avocado is ripe or unripe?

Example: avocados, but with more features

color	softness	variety	ripeness
bright green	firm	Zutano	unripe
green-black	medium	Hass	ripe
purple-black	firm	Hass	ripe
green-black	medium	Hass	unripe
purple-black	soft	Hass	ripe
bright green	firm	Zutano	unripe
green-black	soft	Zutano	ripe
purple-black	soft	Hass	ripe
green-black	soft	Zutano	ripe
green-black	firm	Hass	unripe
purple-black	medium	Hass	ripe

You have a firm green-black Zutano avocado. Based on this data, would you predict that your avocado is ripe or unripe?

Strategy: Calculate $P(\text{ripe}|\text{features})$ and $P(\text{unripe}|\text{features})$ and choose the class with the **larger** probability.

$$P(\text{ripe}|\text{firm, green-black, Zutano})$$

$$P(\text{unripe}|\text{firm, green-black, Zutano})$$

Example: avocados, but with more features

color	softness	variety	ripeness
bright green	firm	Zutano	unripe
green-black	medium	Hass	ripe
purple-black	firm	Hass	ripe
green-black	medium	Hass	unripe
purple-black	soft	Hass	ripe
bright green	firm	Zutano	unripe
green-black	soft	Zutano	ripe
purple-black	soft	Hass	ripe
green-black	soft	Zutano	ripe
green-black	firm	Hass	unripe
purple-black	medium	Hass	ripe

You have a firm green-black Zutano avocado. Based on this data, would you predict that your avocado is ripe or unripe?

Issue: We have not seen a firm green-black Zutano avocado before.

This means that $P(\text{ripe}|\text{firm, green-black, Zutano})$ and $P(\text{unripe}|\text{firm, green-black, Zutano})$ are undefined.

A simplifying assumption

- We want to find $P(\text{ripe}|\text{firm, green-black, Zutano})$, but there are no firm green-black Zutano avocados in our dataset.
- Bayes' theorem tells us this probability is equal to

$$P(\text{ripe}|\text{firm, green-black, Zutano}) = \frac{P(\text{ripe}) \cdot P(\text{firm, green-black, Zutano}|\text{ripe})}{P(\text{firm, green-black, Zutano})}$$

- **Key idea:** **Assume** that features are **conditionally independent** given a class (e.g. ripe).

$$P(\text{firm, green-black, Zutano}|\text{ripe}) = P(\text{firm}|\text{ripe}) \cdot P(\text{green-black}|\text{ripe}) \cdot P(\text{Zutano}|\text{ripe})$$

Example: avocados, but with more features

color	softness	variety	ripeness
bright green	firm	Zutano	unripe
green-black	medium	Hass	ripe
purple-black	firm	Hass	ripe
green-black	medium	Hass	unripe
purple-black	soft	Hass	ripe
bright green	firm	Zutano	unripe
green-black	soft	Zutano	ripe
purple-black	soft	Hass	ripe
green-black	soft	Zutano	ripe
green-black	firm	Hass	unripe
purple-black	medium	Hass	ripe

You have a firm green-black Zutano avocado. Based on this data, would you predict that your avocado is ripe or unripe?

$$P(\text{ripe}|\text{firm, green-black, Zutano}) = \frac{P(\text{ripe}) \cdot P(\text{firm, green-black, Zutano}|\text{ripe})}{P(\text{firm, green-black, Zutano})}$$

Example: avocados, but with more features

color	softness	variety	ripeness
bright green	firm	Zutano	unripe
green-black	medium	Hass	ripe
purple-black	firm	Hass	ripe
green-black	medium	Hass	unripe
purple-black	soft	Hass	ripe
bright green	firm	Zutano	unripe
green-black	soft	Zutano	ripe
purple-black	soft	Hass	ripe
green-black	soft	Zutano	ripe
green-black	firm	Hass	unripe
purple-black	medium	Hass	ripe

You have a firm green-black Zutano avocado. Based on this data, would you predict that your avocado is ripe or unripe?

$$P(\text{unripe}|\text{firm, green-black, Zutano}) = \frac{P(\text{unripe}) \cdot P(\text{firm, green-black, Zutano}|\text{unripe})}{P(\text{firm, green-black, Zutano})}$$

Conclusion

- The numerator of $P(\text{ripe}|\text{firm, green-black, Zutano})$ is $\frac{6}{539}$.
- The numerator of $P(\text{unripe}|\text{firm, green-black, Zutano})$ is $\frac{6}{88}$.
  - Both probabilities have the same denominator, $P(\text{firm, green-black, Zutano})$.
  - Since we're just interested in seeing which one is larger, we can ignore the denominator and compare numerators.
- Since the numerator for unripe is **larger** than the numerator for ripe, we **predict that our avocado is unripe**.

Naive Bayes

Naive Bayes classifier

- We want to predict a class, given certain features.

- Using Bayes' theorem, we write

$$P(\text{class}|\text{features}) = \frac{P(\text{class}) \cdot P(\text{features}|\text{class})}{P(\text{features})}$$

- For each class, we compute the numerator using the **naive assumption of conditional independence of features given the class**.

- We estimate each term in the numerator based on the training data.

- We predict the class with the largest numerator.
  - Works if we have multiple classes, too!

na·ive

/nīˈēv/

adjective

(of a person or action) showing a lack of experience, wisdom, or judgment.
"the rather naive young man had been totally misled"

- (of a person) natural and unaffected; innocent.
  "Andy had a sweet, naive look when he smiled"

  Similar: innocent unsophisticated artless ingenuous inexperienced

- of or denoting art produced in a straightforward style that deliberately rejects sophisticated artistic techniques and has a bold directness resembling a child's work, typically in bright colors with little or no perspective.

Example: comic characters

ALIGN	SEX	COMPANY
Bad	Male	Marvel
Neutral	Male	Marvel
Good	Male	Marvel
Bad	Male	DC
Good	Female	Marvel
Bad	Male	DC
Good	Male	DC
Bad	Male	Marvel
Good	Female	Marvel
Bad	Female	Marvel

My favorite character is a male Marvel character. Using Naive Bayes, would we predict that my favorite character is bad, good, or neutral?

ALIGN	SEX	COMPANY
Bad	Male	Marvel
Neutral	Male	Marvel
Good	Male	Marvel
Bad	Male	DC
Good	Female	Marvel
Bad	Male	DC
Good	Male	DC
Bad	Male	Marvel
Good	Female	Marvel
Bad	Female	Marvel

Example: comic characters

ALIGN	SEX	COMPANY
Bad	Male	Marvel
Neutral	Male	Marvel
Good	Male	Marvel
Bad	Male	DC
Good	Female	Marvel
Bad	Male	DC
Good	Male	DC
Bad	Male	Marvel
Good	Female	Marvel
Bad	Female	Marvel

My other favorite character is a female Marvel character. Using Naive Bayes, would we predict that my favorite character is bad, good, or neutral?

ALIGN	SEX	COMPANY
Bad	Male	Marvel
Neutral	Male	Marvel
Good	Male	Marvel
Bad	Male	DC
Good	Female	Marvel
Bad	Male	DC
Good	Male	DC
Bad	Male	Marvel
Good	Female	Marvel
Bad	Female	Marvel

Uh oh...

- There are no neutral female characters in the data set.
- The estimate $P(\text{female}|\text{neutral}) \approx \frac{\text{\# female neutral characters}}{\text{\# neutral characters}}$ is 0.
- The estimated numerator,
  $P(\text{neutral}) \cdot P(\text{female, Marvel}|\text{neutral}) =$
  $P(\text{neutral}) \cdot P(\text{female}|\text{neutral}) \cdot P(\text{Marvel}|\text{neutral})$,
  is also 0.
- But just because there isn't a neutral female character in the data set, doesn't mean they don't exist!
- **Idea:** Adjust the numerators and denominators of our estimate so that they're never 0.

Smoothing

note for Janine in the future –– remove this slide, only smooth the conditional probs

- **Without** smoothing:

$$P(\text{bad}) \approx \frac{\#\ \text{bad}}{\#\ \text{bad} + \#\ \text{good} + \#\ \text{neutral}}$$

$$P(\text{good}) \approx \frac{\#\ \text{good}}{\#\ \text{bad} + \#\ \text{good} + \#\ \text{neutral}}$$

$$P(\text{neutral}) \approx \frac{\#\ \text{neutral}}{\#\ \text{bad} + \#\ \text{good} + \#\ \text{neutral}}$$

- **With** smoothing:

$$P(\text{bad}) \approx \frac{\#\ \text{bad} + 1}{\#\ \text{bad} + 1 + \#\ \text{good} + 1 + \#\ \text{neutral} + 1}$$

$$P(\text{good}) \approx \frac{\#\ \text{good} + 1}{\#\ \text{bad} + 1 + \#\ \text{good} + 1 + \#\ \text{neutral} + 1}$$

$$P(\text{neutral}) \approx \frac{\#\ \text{neutral} + 1}{\#\ \text{bad} + 1 + \#\ \text{good} + 1 + \#\ \text{neutral} + 1}$$

Smoothing

- **Without** smoothing:

$$P(\text{female}|\text{neutral}) \approx \frac{\#\ \text{female neutral}}{\#\ \text{female neutral} + \#\ \text{male neutral}}$$

$$P(\text{male}|\text{neutral}) \approx \frac{\#\ \text{male neutral}}{\#\ \text{female neutral} + \#\ \text{male neutral}}$$

- **With** smoothing:

$$P(\text{female}|\text{neutral}) \approx \frac{\#\ \text{female neutral} + 1}{\#\ \text{female neutral} + 1 + \#\ \text{male neutral} + 1}$$

$$P(\text{male}|\text{neutral}) \approx \frac{\#\ \text{male neutral} + 1}{\#\ \text{female neutral} + 1 + \#\ \text{male neutral} + 1}$$

- When smoothing, we add 1 to the count of every group whenever we're estimating a probability.

Example: comic characters

Using smoothing, let's determine whether Naive Bayes would predict a female Marvel character to be bad, good, or neutral.

ALIGN	SEX	COMPANY
Bad	Male	Marvel
Neutral	Male	Marvel
Good	Male	Marvel
Bad	Male	DC
Good	Female	Marvel
Bad	Male	DC
Good	Male	DC
Bad	Male	Marvel
Good	Female	Marvel
Bad	Female	Marvel

Summary

Summary

- In classification, our goal is to predict a discrete category, called a **class**, given some features.
- The Naive Bayes classifier works by estimating the numerator of $P(\text{class}|\text{features})$ for all possible classes.
- It uses Bayes' theorem:

$$P(\text{class}|\text{features}) = \frac{P(\text{class}) \cdot P(\text{features}|\text{class})}{P(\text{features})}$$

- It also uses a simplifying assumption, that features are conditionally independent given a class:

$$P(\text{features}|\text{class}) = P(\text{feature}_1|\text{class}) \cdot P(\text{feature}_2|\text{class}) \cdot \ldots$$