CS 260: Foundations of Data Science

Prof. Sara Mathieson

Spring 2025

HAVERFORD COLLEGE

Admin

- **Lab 4** due Wednesday

Outline

- Go over Lab 2
- Finish Evaluation Metrics
- Intro to probability
- Intro to Bayesian models

Outline

- Go over Lab 2
- Finish Evaluation Metrics
- Intro to probability
- Intro to Bayesian models

Lab 2: not posted online

Outline

- Go over Lab 2
- Finish Evaluation Metrics
- Intro to probability
- Intro to Bayesian models

Handout 8 pg 1

① (A)

true \ pred	−	+	
−	77	3	= 80 = N
+	13	7	= 20 = P

$p^* = 10$

② precision $= \frac{7}{10}$

recall (TPR) $= \frac{7}{20}$

FPR $= \frac{3}{80}$

④ (B)

true \ pred	−	+
−	68	12
+	2	18

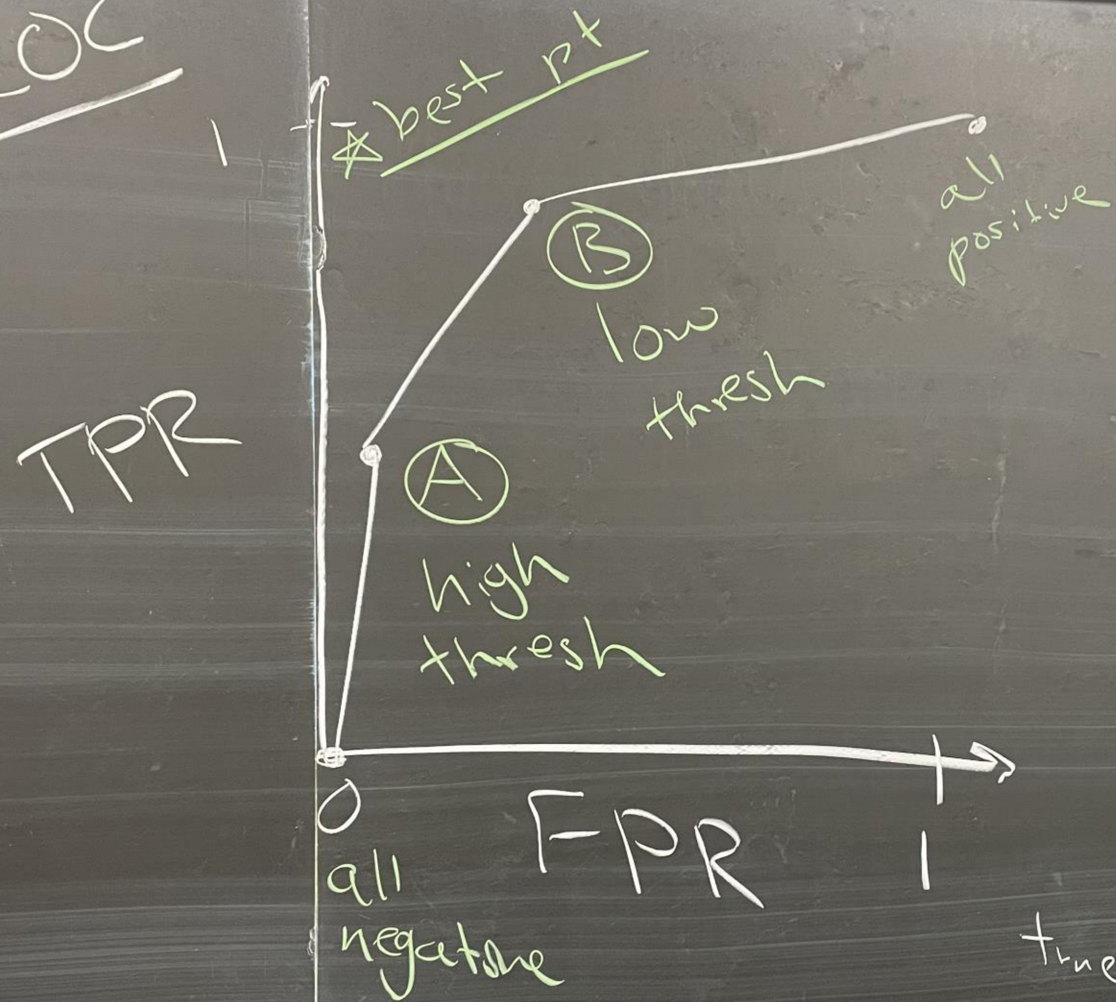
"miss" false negative

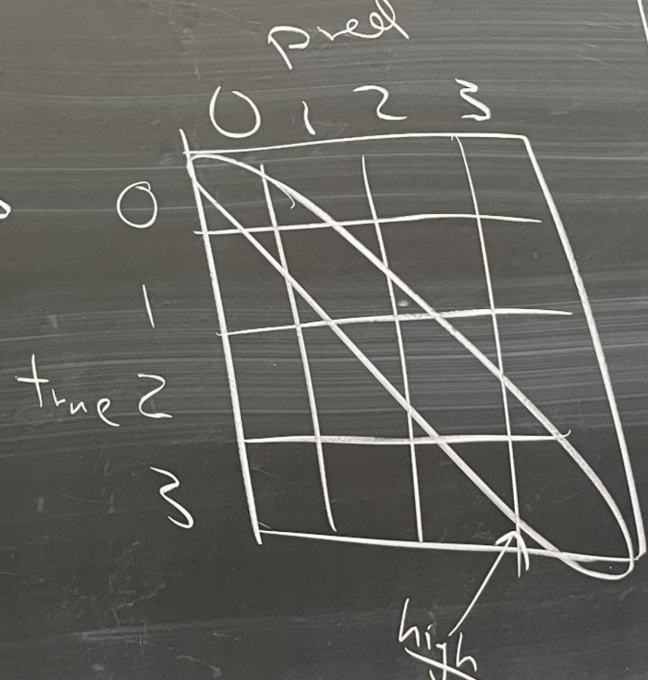
FPR $= \frac{12}{80}$

TPR $= \frac{18}{20}$

"alarm" false positive

ROC


best pt
1
TPR
A
high thresh
B
low thresh
all positive
0
all negative
FPR
1


pred
0 1 2 3
true
0
1
2
3
high

Outline

- Go over Lab 2
- Finish Evaluation Metrics
- Intro to probability
- Intro to Bayesian models

Example

R = rain

U = umbrella

Q: If $P(R) = 20\%$

and $P(R, U) = 15\%$

$\uparrow$ "and" / joint

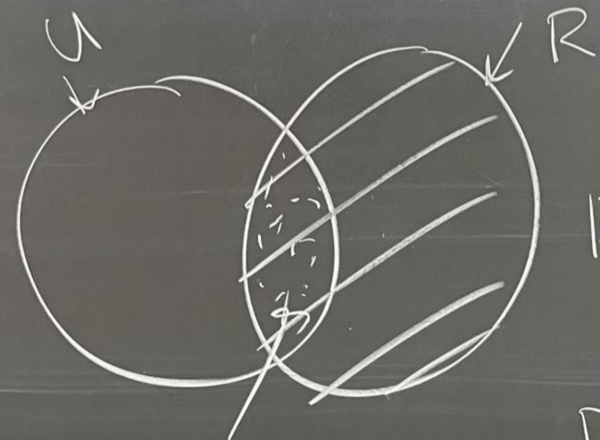
what is $P(U|R)$?

$\uparrow$ "given"

Bayes Rule   true always!

$$\left.\begin{aligned} P(A,B) &= P(A)\,P(B|A) \\ P(A,B) &= P(B)\,P(A|B) \end{aligned}\right\} =$$

$$\Rightarrow \boxed{P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}}$$



$\blacksquare = 0.2$

$\boxdot = 0.15$

$$P(U \mid R) = \frac{P(R, U)}{P(R)} = \frac{15}{20} = \boxed{.75}$$

Sum over all events should be 1

$$\sum_{a \in vals(A)} P(a) = 1 \quad , \quad \underline{ex}: \; P(R) + P(\bar{R}) = 1$$

not rain

$$\sum_{a \in vals(A)} P(a \mid B) = 1$$

Independence

$$P(A, B) = P(A)P(B)$$

not always true

$$\Downarrow$$

$$\cancel{P(B)}P(A|B) = P(A)\cancel{P(B)}$$

Conditional Independence

$$P(A|B, C) = P(A|C)$$

"A is independent of B given C"

thunder rain lightning

Marginalization

$$P(A) = \sum_{b \in vals(B)} P(A, B=b)$$

only weather

ex: $P(U) = P(U, sun) + P(U, rain)$

only if independent

$$P(U)p(sun) + P(U)p(rain)$$

$$P(U)\left(\underbrace{P(sun) + P(rain)}_{1}\right)$$

Intro to Probability

- The **probability** of an **event** e has a number of epistemological interpretations
- Assuming we have **data**, we can count the number of times e occurs in the dataset to estimate the probability of e, $P(e)$.

$$P(e) = \frac{\text{count}(e)}{\text{count}(\text{all events})}.$$

- If we put all events in a bag, shake it up, and choose one at random (called **sampling**), how likely are we to get e?

Intro to Probability

- Suppose we flip a fair coin
- What is the probability of heads, $P(e = H)$?

Intro to Probability

- Suppose we have a fair 6-sided die.

$$\frac{count(s)}{count(1) + count(2) + count(3) + \cdots + count(6)} = \frac{1}{1+1+1+1+1+1} = \frac{1}{6}$$

Intro to Probability

- What about a die with on ly three numbers $\{1, 2, 3\}$, each of which appears twice?
- What's the probability of getting "1"?

Intro to Probability

- The set of all probabilities for an event e is called a **probability distribution**
- Each die roll is an independent event (Bernoulli trial).

Intro to Probability

- Which is greater, $P(HHHHH)$ or $P(HHTHH)$?

Intro to Probability

Probability Axioms

1. Probabilities of events must be no less than 0. $P(e) \geq 0$ for all e.
2. The sum of all probabilities in a distribution must sum to 1. That is, $P(e_1) + P(e_2) + \ldots + P(e_n) = 1$. Or, more succinctly,

$$\sum_{e \in E} P(e) = 1.$$

Intro to Probability

Joint Probability

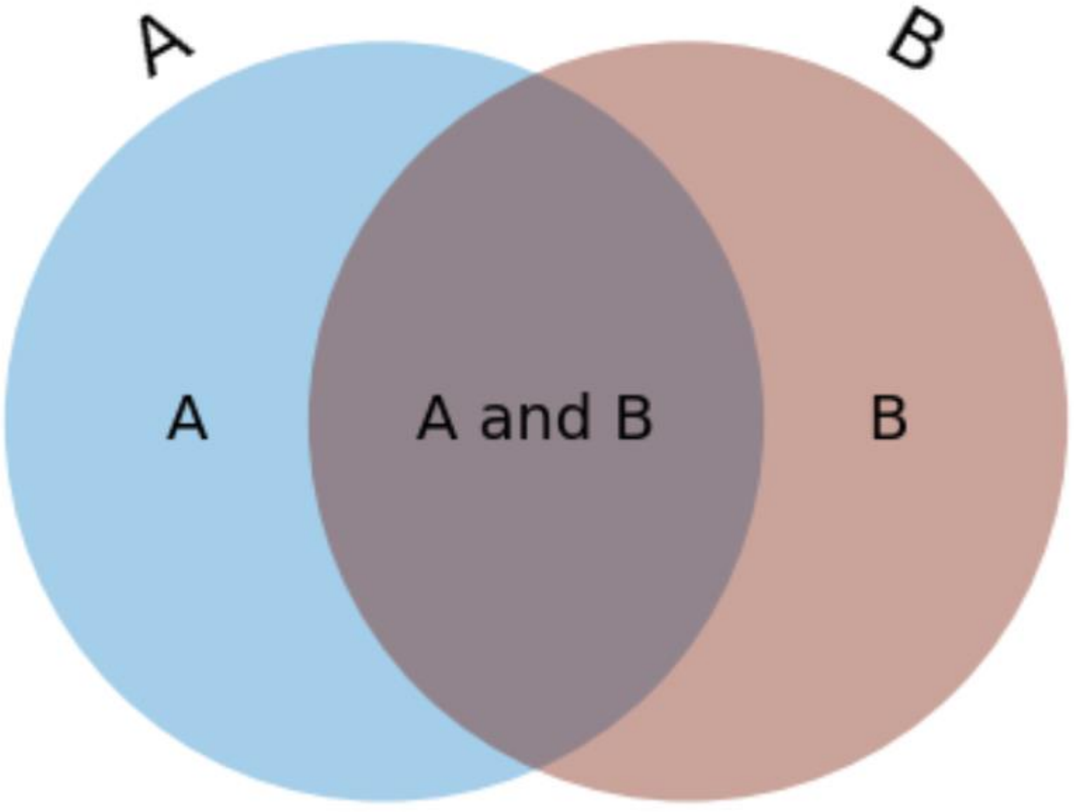
The probability that two independent events e_1 and e_2 *both* occur is given by their product.

$$P(e_1 \wedge e_2) = P(e_1 \cap e_2) = P(e_1)P(e_2) \text{ when } e_1 \cap e_2 = \emptyset$$

- Intuitively, think of every probability as a *scaling factor*.
- You can think of a probability as the fraction of the probability space occupied by an event e_1.
  - $P(e_1 \wedge e_2)$ is the fraction of of e_1's probability space wherein e_2 also occurs.
  - So, if $P(e_1) = \frac{1}{2}$ and $P(e_2) = \frac{1}{3}$, then $P(e_2, e_2)$ is a third of a half of the probability space or $\frac{1}{3} \times \frac{1}{2}$.

Intro to Probability

Joint Probability



Materials by Alvin Grissom II

Intro to Probability

Conditional Probability

- A **conditional probability** is the probability that one event occurs given that we take another for granted.
- The probability of e_2 given e_1 is $P(e_2 \mid e_1)$.
- This is the probability that e_2 will occur given that we take for granted that e_1 occurs.

Intro to Probability

Marginal Probability Distributions

Given a discrete joint probability distribution function $P(X, Y)$, how would we find $P(X)$?

- "Marginalize out" the Y (sum over all all $y \in Y$).
- Fix the X.
- Discrete Case: $p(x) = \sum_{y \in Y} P(x, y)$
- Continuous Case: $p(x) = \int p(x, y) dy$

Example

$$\underbrace{p(\overbrace{\text{spam}}^{\text{Y or N}} \mid \overbrace{\text{words}}^{\text{email}})}_{\text{want to compute}} = \frac{p(\text{spam}, \text{words})}{p(\text{words})}$$

"X" "data"
very difficult!

"Posterior"

$$= \frac{p(\text{spam}, \text{words})}{p(\text{words}, \text{spam}) + p(\text{words}, \overline{\text{spam}})}$$

"prior"

"likelihood"
(generative)

$$= \frac{p(\text{spam})\, p(\text{words} \mid \text{spam})}{\underbrace{p(\text{spam})\, p(\text{words} \mid \text{spam}) + \underbrace{p(\overline{\text{spam}})}_{1-\text{prior}}\, p(\text{words} \mid \overline{\text{spam}})}_{\text{"evidence"}}}$$

Handout 8, pg 2

$$P(D\mid pos) = \frac{P(D)\,P(pos\mid D)}{P(pos)}$$

$$= \frac{P(D)\,P(pos\mid D)}{P(pos, D) + P(pos, \bar{D})}$$

$$= \frac{P(D)\,P(pos\mid D)}{P(D)\,P(pos\mid D) + P(\bar{D})\,P(pos\mid \bar{D})}$$

$$p(\text{neg} \mid H) = 0.9$$

$$p(\text{neg} \mid H) \underset{\uparrow \\ \text{or}}{+} p(\text{pos} \mid \underset{\sim \\ \bar{D}}{H}) = 1$$

$$\frac{\frac{1}{\cancel{100}} \cdot \frac{9}{\cancel{10}}}{\frac{1}{\cancel{100}} \cdot \frac{9}{\cancel{10}} + \frac{99}{\cancel{100}} \cdot \frac{1}{\cancel{10}}} = \frac{9}{108} = \frac{1}{12} \approx \boxed{8\%}$$

Outline

- Go over Lab 2
- Finish Evaluation Metrics
- Intro to probability
- Intro to Bayesian models

Next time!