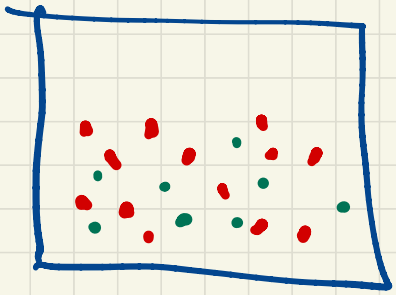
Basic Probability



100 balls total
75 red balls
25 green balls

$P(\text{red ball}) = \frac{75}{100} = 0.75$

$P(\text{green ball}) = \frac{25}{100} = 0.25$

$P(\text{red ball}, \text{green ball}) = ??$

Conditional Probability

$P(x,y) = P(x|y) \cdot P(y)$ →

① Keep the red ball out (w/o replacement)

$\frac{75}{100} \cdot \frac{25}{99} \approx \boxed{0.19}$

② Put red ball back

Independence

$P(x,y) = P(x) \cdot P(y)$ →

$\frac{75}{100} \cdot \frac{25}{100} = \frac{3}{16} = \boxed{0.1875}$

Bayes' Rule

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)}$$

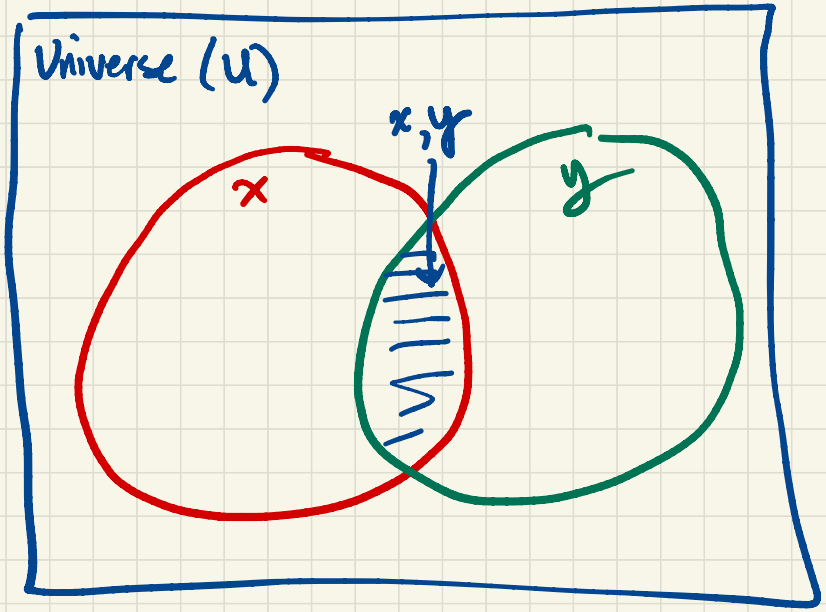
$$P(x) = \frac{|x|}{|u|}$$

$$P(y) = \frac{|y|}{|u|}$$

$$P(x,y) = \frac{|xy|}{|u|}$$

$$P(y|x) = \frac{|xy|}{|x|} = \frac{\frac{|xy|}{|u|}}{\frac{|x|}{|u|}} = \frac{P(xy)}{P(x)}$$

$$P(x|y) = \frac{P(xy)}{P(y)}$$



$$\cancel{P(x,y)} = P(x|y) \cdot P(y) = P(y|x) \cdot P(x)$$

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)}$$

label / prediction

observations / features

Naive Bayes Classifier

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)}$$

label / prediction — observations / features

Predict $\hat{y}$, having observed $x = \{x_1 \dots x_n\}$.

$$\hat{y} = \underset{y}{\operatorname{argmax}} \{ P(y = C \mid x = x_1, x_2 \dots x_n) \}$$

need to compute! ← Bayes rule

observe / estimate → $P(x = x_1 \dots x_n \mid y)$ · $P(y)$ ← easy

$$= \underset{y}{\operatorname{argmax}} \left\{ \frac{P(x = x_1 \dots x_n \mid y) \cdot P(y)}{\cancel{P(x)}} \right\}$$

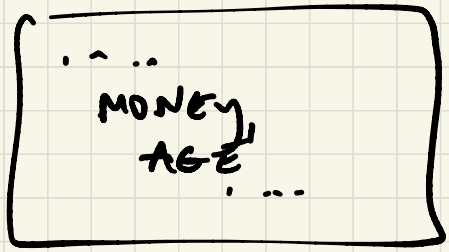
independence

$$\approx \underset{y}{\operatorname{argmax}} \{ P(x_1|y) \cdot P(x_2|y) \cdots P(x_n|y) \cdot P(y) \}$$

Example: Spam Filtering

$$\underset{y}{\operatorname{argmax}} \{ P(x_1|y) \cdot P(x_2|y) \cdots P(x_n|y) \cdot P(y) \}$$

TRAINING ($P(x_i|y_i)$)



TESTING

SPAM (y = SPAM)

- MONEY
- ROLEX
- PHARMACY, AGE

$P(\text{MONEY}|\text{SPAM}) = 1/3$

$P(\text{AGE}|\text{SPAM}) = 1/3$

① $P(\text{MONEY}|\text{SPAM}) \cdot P(\text{AGE}|\text{SPAM}) \cdot P(\text{SPAM})$

$= \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{9}{10} = \frac{9}{90} = 0.1$

HAM (y = HAM)

- CLASS
- MACHINE LEARNING
- AGE

$P(\text{MONEY}|\text{HAM}) = 0/3 \rightarrow \alpha$ (small)

$P(\text{AGE}|\text{HAM}) = 1/3$

② $P(\text{MONEY}|\text{HAM}) \cdot P(\text{AGE}|\text{HAM}) \cdot P(\text{HAM})$

$= 0 \cdot \frac{1}{3} \cdot \frac{1}{10} = 0$

$= \alpha \cdot \frac{1}{3} \cdot \frac{1}{10} = \frac{\alpha}{30}$

(e.g., if $\alpha = \frac{1}{3}$, then $1/90$)