# Naive Bayes

# Do Italian guys love football?

I don’t like football!

How many italian guys you know?

What happens if you meet other 20 Italian guys who don’t love football?

Prior belief

Evidence

Posterior belief

Prior belief = “italian guys love football” (very strong)

Evidence = “Edo doesn’t love football”

Posterior belief = “italian guys love football” (less strong)

Posterior = Update(Prior, Evidence)

What's this?!



Posterior = Update(Prior, Evidence)

# Solution: Bayes Theorem

P(play tennis | weather=overcast)

=

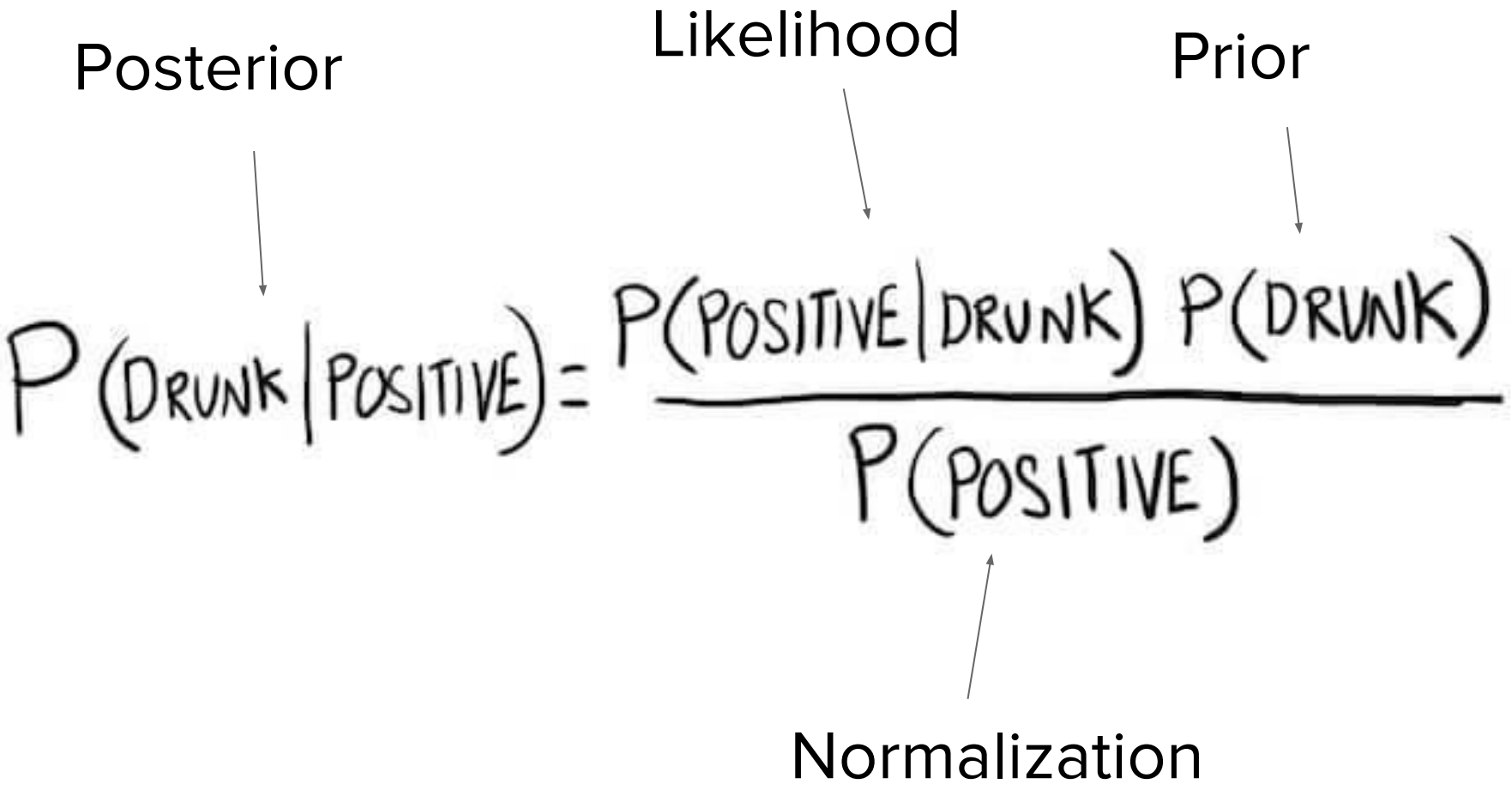
Conditional probability

Hypothesis

$$p(\mathrm{H} \mid \mathrm{E}) = \frac{p(\mathrm{E} \mid \mathrm{H})\, p(\mathrm{H})}{p(\mathrm{E})}$$

Evidence

$$P(\text{DRUNK} \mid \text{POSITIVE}) = \frac{P(\text{POSITIVE} \mid \text{DRUNK})\, P(\text{DRUNK})}{P(\text{POSITIVE})}$$


Posterior
Likelihood
Prior
P(DRUNK|POSITIVE) = P(POSITIVE|DRUNK) P(DRUNK) / P(POSITIVE)
Normalization

# QUIZ

# Example: Rare Disease

Predicted

| Real | negative | positive |
| --- | --- | --- |
| **No disease** | 90 | 10 |
| **disease** | 5 | 95 |

Is the model good?

Incidence on the population

=

1/100,000 = 0.001%

Taken a random guy from the street, if he's positive to the test, what is the probability that he has the disease?

80%?  5%?

$$P(\text{disease} \mid \text{positive}) = ?$$

$$P(\text{disease} \mid \text{positive}) = \frac{P(\text{disease}) * P(\text{positive} \mid \text{disease})}{P(\text{positive})}$$

$$P(\text{disease} \mid \text{positive}) = \frac{P(\text{disease}) * P(\text{positive} \mid \text{disease})}{P(\text{positive})}$$

0.00001

95 / 100 = 0.95

(95+10) / (200) = 0.52

$$P(\text{disease} \mid \text{positive})$$

$$=$$

$$0.0018\%$$

# Bayes Classifier

$$P(\text{class=apple} \mid \text{color=red}) \ ?$$

$$P(\text{class=pear} \mid \text{color=red}) \ ?$$

| Color (feature) | Fruit (class) |
| --- | --- |
| red | apple |
| red | apple |
| yellow | apple |
| red | pear |
| yellow | pear |
| yellow | pear |
| yellow | pear |

$$P(\text{apple} \mid \text{red}) = p(\text{apple}) * p(\text{red} \mid \text{apple}) / p(\text{red})$$

$$P(\text{pear} \mid \text{red}) = p(\text{pear}) * p(\text{red} \mid \text{pear}) / p(\text{red})$$

$$P(\text{apple} \mid \text{red}) = 3/4$$

$$P(\text{pear} \mid \text{red}) = 1/4$$

# *Naive* Bayes Classifier

P(class=apple | color=red, width=5cm)

Naive

=

It assumes features are independent

Strong assumption!

# In practice

The model works even when the assumption is violated

# When is Naive Bayes good?

1 - Quick

2 - Good for text

In general not very performing

(Like Knn, Logistic Regression, etc.)

# Naive Bayes In Sklearn

1. Multi-variate Bernoulli
2. Multinomial
3. Gaussian

## Multi-variate Bernoulli

Binary features (i.e., 0s and 1s)

Binary bag of words

**Multinomial Naive Bayes**

Discrete counts.

Traditional Bag of Words

## Gaussian

Normally Distributed features

E.g. Iris dataset - sepal width, petal width, sepal length, petal length

# SKIP
# Generative vs. Discriminative