Lesson 2 Naive Bayes

Supervised Classification

- learning from labeled data. After understanding the data, the algorithm determines which label should be given to new data by associating patterns to the unlabeled new data.
- Examples
    - Identifying someone from a set of pictures
    - Song recommendation based on previous liked songs

Features and Labels

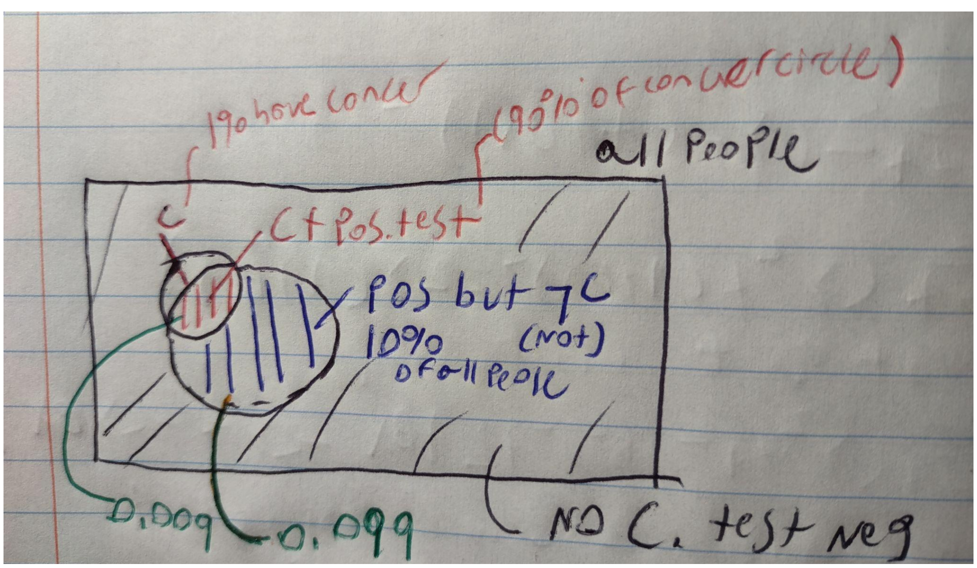
- Song example
- Features
    - Intensity, Temp, Genre, Voice gender
- Labels
    - Like, Dislike

Naive Bayes

p(c) = 0.01, c = Cancer

Test: 90% it is positive if you have C (Sensitivity)
90% it is negative if you don't have C (Specitivity)

Question: Test = Positive



What is the Probability of Having cancer?
about 8%, see diagram in written notes

- **Sensitivity** (True Positive Rate) refers to the proportion of those who received a positive result on this test out of those who actually have the condition (when judged by the 'Gold Standard').
- **Specificity** (True Negative Rate) refers to the proportion of those who received a negative result on this test out of those who do not actually have the condition (when judged by the 'Gold Standard').

Bayes Rule

Prior probability * test evidence → posterior probability

Prior: p(c) = 0.01 = 1% p(not c) = 0.99 = 99%
p(pos | c) = .9 = 90%
p(neg | not c) = 0.9 p(pos | not c) = 0.1

Posterior (joint probability):
p(c , pos.) = p(c) * p(pos | c) = 0.01 * .9 = 0.009
p(not c , pos) = p(not c) * p(pos | not c) = .99 * 0.1 = 0.099

Normalizer
p(c , pos.) + p(not c , pos) = 0.009 + 0.099 = 0.108

Posterior (actual):

p(c | pos) = $\frac{0.009}{0.108}$ = 0.0833

p(not c | pos) = $\frac{0.099}{0.108}$ = 0.9167

p(c | pos) + p(not c | pos) = 0.0833 + 0.9167 = 1

Bayes Rule Diagram