Diagnostic Testing

STA 198: Introduction to Health Data Science

Yue Jiang

May 29, 2020

The following material was used by Yue Jiang during a live lecture.

Without the accompanying oral comments, the text is incomplete as a record of the presentation.

Conditional probabilities

Suppose we care about the probability someone has HIV, denoted by $P(HIV+)$

- What if they have a positive HIV test? $P(HIV+|Test+)$
- What if they have a negative HIV test? $P(HIV+|Test-)$

Knowing the result of the HIV test changes our estimate of their HIV probability

Which of the three probabilities would we expect to be the highest? The lowest?

Medical diagnostics

Suppose we're interested in the performance of a diagnostic test. Let A be the event that someone has a condition of interest, and let B be the event that a test for that condition is positive

- Prevalence: $P(A)$
- Sensitivity: $P(B|A)$, or the true positive rate
- Specificity: $P(B^c|A^c)$, or 1 minus the false positive rate
- Positive Predictive Value: $P(A|B)$
- Negative Predictive Value: $P(A^c|B^c)$

What do these probabilities mean in plain English?

Rapid self-administered HIV tests

From the FDA package insert for the OraQuick ADVANCE Rapid HIV-1/2 Antibody Test,

- Sensitivity, $P(Test+|HIV+)$: 99.3%
- Specificity, $P(Test-|HIV-)$: 99.8%

From CDC statistics for 2016, 14.3/100,000 Americans aged $\geq$13 are HIV+.

Suppose a randomly selected American aged $\geq$13 has a positive test on this test. What do you think is the probability they are HIV+?

Let me know here!

Bayes' rule

Using the law of total probability, we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$
$$= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Let A be the event of being HIV+ and B be testing positive,

$$P(HIV+|Test+) = \frac{P(Test+|HIV+)P(HIV+)}{P(Test+|HIV+)P(HIV+) + P(Test+|HIV-)P(HIV-)}$$

Back to OraQuick...

Suppose a randomly selected American aged $\geq$13 has a positive test on this test. What do you think is the probability they are HIV+?

- Is this calculation surprising?
- What is the explanation?
- Is this calculation reasonable to perform?
- What if a randomly selected adult in Botswana tested positive (HIV prevalence $\approx$25%)?

Discrimination thresholds

Oral HIV tests give positive or negative results depending on levels of HIV antibodies detected in saliva

- If antibody levels are above a certain threshold, it is classified as a positive test
- Varying the threshold for a positive vs. negative test will result in a test in different characteristics
- But what is this threshold of "high" antibody level that classifies them as positive?
- We are looking for a binary classifier that classifies a patient as being positive or negative based on a threshold value of a continuous variable.
- At each threshold value, there is a trade-off between sensitivity and specificity
- How might you suggest an HIV diagnostic rule that has 100% sensitivity? Specificity?

ROC curves

Receiver Operating Characteristic curves show how specificity and specificity change as the discrimination threshold changes

Plotting true vs. false positive rates

ROC curves show, for each false positive rate, what the true positive rate is corresponding to that threshold value (how do these relate to sensitivity and specificity?).

Back to the first day of class...

ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

Annals of Oncology 29: 1836–1842, 2018
doi:10.1093/annonc/mdy166
Published online 28 May 2018

ORIGINAL ARTICLE

Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists

H. A. Haenssle[1*,†], C. Fink[1†], R. Schneiderbauer[1], F. Toberer[1], T. Buhl[2], A. Blum[3], A. Kalloo[4], A. Ben Hadj Hassen[5], L. Thomas[6], A. Enk[1] & L. Uhlmann[7]

From Haenssle et al. (2018)

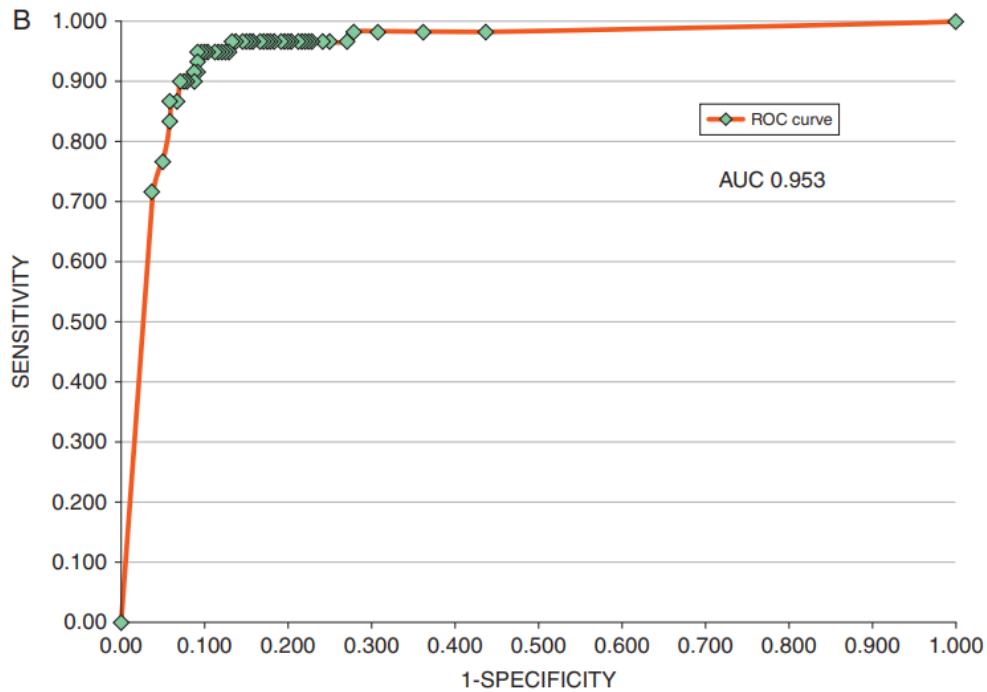

Figure 2. (A) ROC curve of the CNN in relation to the average (±SD) sensitivity and specificity of all dermatologists [mean: green (online) circle; ±SD: green (online) error bars] in set-100 (dichotomous classification, study level-I) and the dermatologists' mean sensitivity and specificity in relation to their level of experience. (B) ROC curve of the CNN in set-300.

In which area of the plot would a "good" classifier lie?