BIOM/SYSC5405 – Pattern Classification and Experiment Design

Assignment 3— Peer Evaluation Guide

Overall Guidance

- Only evaluate "presentation" once (in the first rubric criterion). Please don't penalize poor presentation quality repeatedly throughout the rubric.
- Code can be either inline or in an appendix.
- Given word limits for answers are just a guide. Please allow ~150% of that limit before considering a response to be unduly long.
- The peer review has five phases:
    - 4 Nov: submit your own solution
    - 7 Nov: evaluate three of your peers' solutions
    - 9 Nov: deadline to review the evaluations <u>that you have received</u> and provide feedback by replying to your reviewers' comments
    - 11 Nov: deadline to double-check all three of the evaluations <u>that you gave</u> to see if you received any replies to your original grading.
    - **For each reply, you must respond with a new comment justifying your grade (whether you agree to change it or not)**. Respond by adding a 'reply' to each of the 'replies' that you received.
    - If you fail to respond to any replies, your own grade may be impacted.
    - 13 Nov: deadline to let me know if there are unresolved issues with the grading that you have received.
- Please keep in mind that changing a single grade by 15% (e.g., "QA goes from 77% to 90% for one of your graders") will have a negligible impact on your final grade:

  +13% / (3 graders) * (15% per question) / (4 assignments) * (30% of final grade) = +0.05

Question A

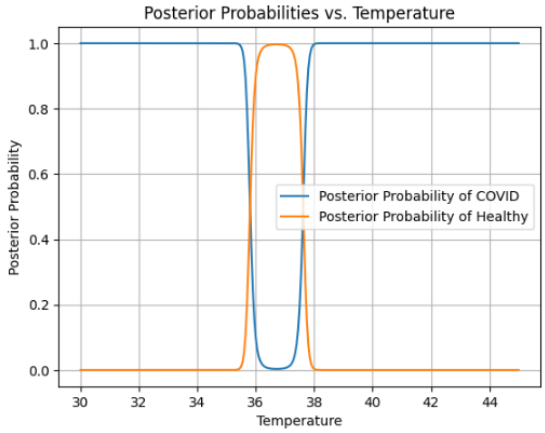
For Question A, please grade **Question 2**:

COVID is happily now less prevalent than it was when the data were first collected. We can now assume that, for every one person with COVID, there are 9 people without COVID in the population. i) Use Bayes' theorem to compute the posterior probability that a patient with a temperature of 37.5 degrees and a respiration rate of 23 is healthy. ii) Determine (*analytically or through trial-and-error*), what is the minimum temperature at which this patient (RR=23) will be classified as having `covid`.

The student should provide:

1) The **posterior probability** $\boldsymbol{P(HEALTY|x = [37.5, 23]) = 0.808,}$ which can be computed from:
    a. the prior (0.1 for COVID; 0.9 for HEALTHY)
    b. the class-conditional distributions estimated in Q1
    c. p(x), which is the prior-weighted sum of class conditionals $p(x = [37.5,23]) = p(x|H)P(H) + p(x|C)P(C)$.
2) The minimum temperature at which a patient with RR=23 will be classified as COVID. We will accept **either** **negative infinity** $(-\infty)$ **OR** 0 Kelvin **OR** -273 degrees Celsius **OR** `37.6`

*For bonus marks, the student may provide a diagram of the posterior probability of COVID (or both classes) and discuss how there are actually two separate ranges of temperatures that lead to a COVID diagnosis (T<x **OR** T>`37.63`). A sample plot is shown below (thanks to M.Abourobea):*



Common errors:

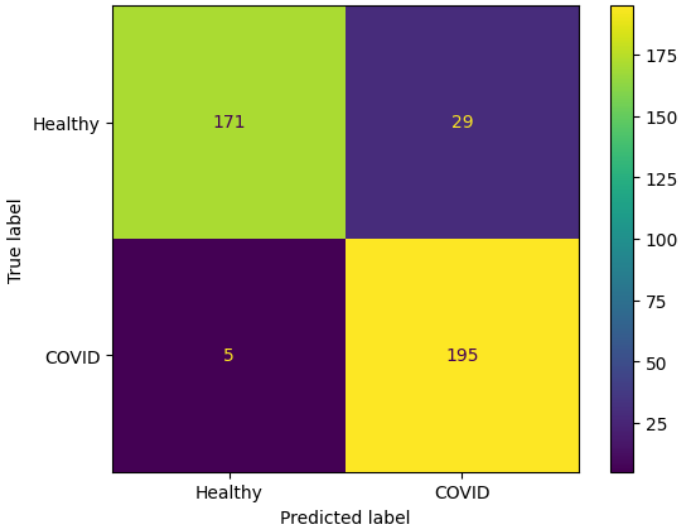
- The posterior probability is incorrectly by more than 5% (roughly [0.768,0.845])
- The min temperature for COVID classification is incorrect (let's accept range [37.55,38.0] **or** $-\infty$ or -273 degrees Celsius or 0 degrees Kelvin)
- Note: If one error causes other errors, just count it as one error. If everything else looks correct (peak at code), but one error (e.g., incorrect posterior) causes downstream errors, that should be treated as a single error.

Question B

For Question B, please grade **Question 4.**

Given the high cost of false positives, you decide that your false positive rate must be below 15%. i) What is the maximum sensitivity we can achieve? ii) What is the maximum precision that we can achieve? iii) Report a confusion matrix for this decision threshold.

```
i)   The max Sn @ FPR<=15% = 0.975 (accept [0.970,0.980])
ii)  The max Pr @ FPR<=15% = 0.871 or 1.0 (accept [0.865,0.880] OR 1.0)
iii) Confusion matrix for FPR=15% with threshold = 0.0162:
       Allow differences of up to 3 in each cell.
```



Note that students may supply two confusion tables (one for max Sn, one for max Pr). One of the confusion tables should match the one above; ignore the second table.

Common Errors

- Any of the answers is wrong (see allowable ranges above; one minor error per erroneous answer)
- Confusion matrix is missing labels on rows/columns *(heatmap is not required)*

Question C

For Question C, please grade **Question 7.**

We will now use a K-nearest-neighbour classifier to classify all passengers in the original 400-patient data set (ignore prior information). Report the apparent error rate for K-NN classifiers with K={1,5,15,25}. Which value of the hyperparameter, K, performs best and why? *(50 words)*

Answer:

```
K=1, apparent error rate = 0.0
K=5, apparent error rate = 0.06
K=15, apparent error rate = 0.0775
K=25, apparent error rate = 0.085

K=1 performs best because we have used the same data for training and
testing, so the nearest sample to each test sample is itself, with a
distance of 0! (This classifier is almost certainly overfit..)
```

Common Errors:

- The student failed to report the apparent error rate for each K
- The apparent error rates are incorrect (likely because the student used a different train/test protocol, rather than reporting the **apparent error** (i.e., resubstitution error)
- The optimal K does not agree with their stated results (note that if their apparent error rates differ from the solution, they may arrive at a different optimal K value)
- The student did not provide a reasonable explanation for why the observed best K is 'best'.
    - An explanation of "because it has the lowest error" is insufficient.
    - If they found that a different value of K was optimal (because they made a previous error), then we will accept any reasonable explanation.
    - If they found that K=1 is optimal, then I'm looking for an explanation very similar to mine.