BIOM/SYSC5405 – Pattern Classification and Experiment Design

Assignment 4— 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:
    - 20 Nov: submit your own solution
    - 23 Nov: evaluate three of your peers' solutions
    - 26 Nov: deadline to review the evaluations <u>that you have received</u> and provide feedback by replying to your reviewers' comments
    - 28 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.
    - 1 Dec: 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) * (16% per question) / (4 assignments) * (30% of final grade) = +0.05%

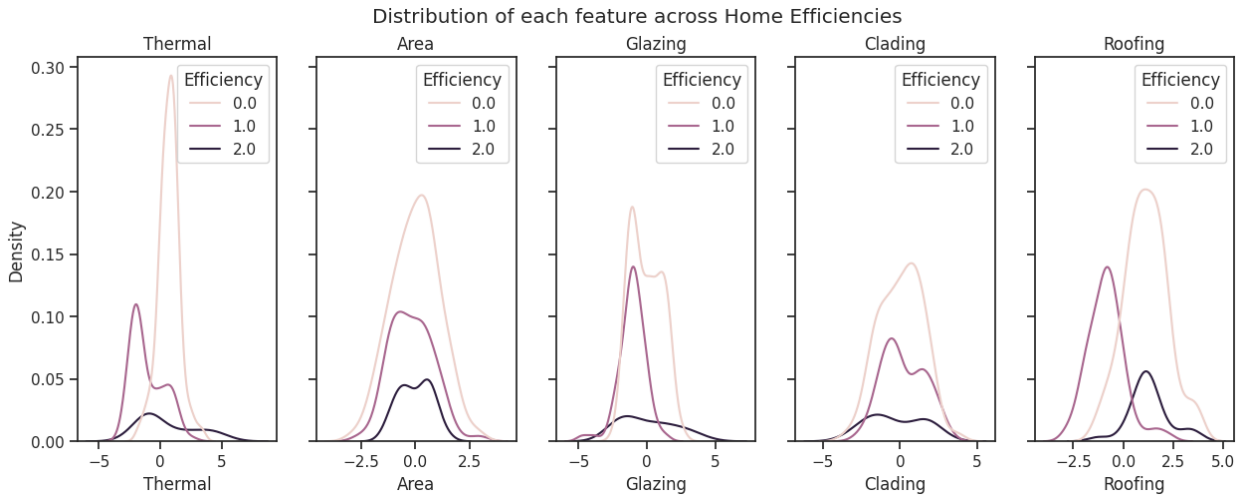
Question A

For Question A, please grade **Question 3**:

Using the training set, for each feature, plot the feature distribution for each class. You can either use five histograms or five 1D kernel density plots. Label each sub-plot by the feature name. The distribution of feature values should be visible for all three (potentially overlapping) classes on each of the five plots. Which feature looks most useful and why? Which home efficiency class do you think will have the lowest accuracy and why? (60 words max)

The student should provide:

1) Five histograms or five kernel density plots, one for each feature. Each sub-plot should be labeled with the name of the feature. A sample figure is shown below:



2) A statement of which feature is most useful and some explanation for why (e.g., "best separates the three classes"). We will accept **either Thermal OR** Roofing. Valid reasons include (*full marks for this or something similar*):
    a. At least two of the three classes are most clearly separated using this feature
3) A statement of which efficiency (class) will have the lowest accuracy and why. Based on the figure above, an efficiency of 2.0 will have the lowest accuracy. Valid reasons include (*full marks for either one of these*):
    a) since it has the lowest number of exemplars
    b) since that class overlaps with the other classes in all five features.

Common errors:

- The figures are missing y-axis labels or titles, or the three distributions cannot be clearly seen (i.e., the three data series are occluding each other), or no legend is provided
- The wrong feature is selected of no reason why is provided
- The wrong class/efficiency is selected or no valid reason why is provided

Question B

For Question B, please grade **Question 4 and Question 5.**

4. Complete 5-fold-cross-validation over the train subset using an SVM classifier with a polynomial kernel with degree=3 and C=0.8. **Report** the accuracy over each fold, the average accuracy across all five folds, and the standard deviation across the five accuracy measurements.

5. Train another SVM model (same kernel & C) on all of your training samples. Test on the test subset. **Report** the accuracy on the test subset. Does it fall within 1 standard deviation of the average accuracy observed in Question 4?

The student should provide:

1) Five accuracies for the five folds of cross-validation across the training set. Accept values in the range [0.650,0.920]; if outside of that range, double-check code for correct implementation.
2) The average accuracy across all five folds and the standard deviation. Accept an average in the range of [0.750,0.850] and SD in range (0.04-0.085); if outside of that range, double-check code for correct implementation. Clearly, the average should be the arithmetic average of the student's observed performance on all five folds.
3) Test classification accuracy across the test subset. Accept values in range [0.650,0.920]; if outside of that range, double-check code for correct implementation.
4) A statement of whether the test accuracy falls within 1 SD of the training accuracy (it should; if it does not, then double-check the code for errors)

Note that we expect some variation in classifier performance due to the stochastic nature of splitting the data into train/test and then into folds. There is also randomness in training an SVM.

Common Errors

- Any of the answers is wrong (see allowable ranges above; one minor error per erroneous answer)

Question C

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

Using 5-CV across only the training subset, perform a hyperparameter sweep of the number of hidden nodes in a 3-layer feedforward neural network. Report your accuracy for numH=[1, 10, 100] hidden nodes. Use the 'adam' solver, a hyperbolic tangent activation function for the hidden layers.

The student should provide the accuracy, averaged over their five folds, for each value of `numH`. We are expecting answers in the range of: `[0.750,0.900]`

Common Errors:

- The student failed to report the accuracy for each value of numH
- The network structure (number of layers, activation function, solver, etc.) does not match the requirements in the question text.