ECON484 Spring 2025
HW 2

Goal: To gain experience in the analysis, design and implementation of classification algorithms.

Explanation:

Follow the steps below:

1. If we consider the letter grades of a university course as dividing students into different success classes, all the measurement and evaluation activities in the course as a whole are a classifier. Evaluate this conceptualization.

2. Interview at least 2 faculty members and at least 5 students about how the level of precision that comes with classifying a course with letters such as A, B and intermediate letters (such as BA, CB, and more precise letter ranges such as A-, B+, B-, C+ in some universities) is actually perceived. Summarize these interviews.

3. Discuss whether there are situations where it would be more appropriate to classify a course as Outstanding (A), Successful (B to D), and Failing (FD, FF, etc.).

4. 4. Use the k-NN algorithm to reclassify the ungraded student results in the attached data file as Outstanding, Successful, and Failing. While doing this process,

   1. 1. You can evaluate the end-of-term totals from 100 to 85 as Outstanding Successful, 84 to 55 as Successful, and 55 and below as Unsuccessful. For the evaluation performance of each class, manually classify the original data according to the rule (in substep number 1).

   2. Use a subset that you will select from the data that is already given as learning data. Discuss how you selected the subset that you will select (sampling method) and the effect of this method on the k-NN algorithm.

   3. Select and explain which parameters you will measure the distance between when using the k-NN method in classification. You do not have to select a distance only to the end-of-term total (one-dimensional). You can make a two- or more-dimensional classification by measuring the distance to more data that you consider important (for example, the end-of-term project and the end-of-term total). You should also choose your distance criterion (Manhattan, Euclid, etc.) together with this evaluation.

   4. Classify the remaining data by training the algorithm with the learning data.

   5. Calculate the True Positive, True Negative, False Positive, False Negative values by comparing the algorithm result and manual classification results.

Notes:

1. In order to discuss the evaluation sensitivity well, it may be more useful to meet with students who received letters that were (just barely) higher than the cut-off values or who failed the course, or with a faculty member of a course where many students received low passing grades (DC, DD).

2. In the given file, note that some students' quiz grades are between 0-10 and others' between 0-100. You will need to normalize the grades of these student ranges to the same range.

3. You do not have to write a program with Python when using the K-NN algorithm, but it should be noted that it is a much easier alternative than tools such as Excel, etc.