Programming Project 1 – Data Editing and Reduction

Choose your preferred programming language (Python, R, Java, C, Perl, SQL, etc) and implement three programs with the following requirements:

1. Assume that the training datasets are csv files comprised of numerical attributes only with the exception of the class attribute. You can also assume that the class attribute is the last attribute of the file.
2. Implement program **NormalizeValues** that takes as input a csv file and normalizes the values of all its attributes (i.e., transforms them in the [0, 1] range). Obviously, the class attribute should be excluded. The normalized file should be written to disk.
3. Implement program **ENN** that takes as input a normalized csv file and applies the editing algorithm ENN on it. The edited file should be written to disk.
4. Implement program **IB2** that takes as input a normalized csv file and applies the instance reduction algorithm IB2 on it. The reduced file should be written to disk.
5. Test your implementations with the provided **iris.csv** and **letter-recognition.csv** datasets.

Submit in a zip file with
(a) your code (for compiled languages like Java, C, etc., also submit the executable files),
(b) the edited datasets (irisENN.csv and letter-recognitionENN.csv), and,
(c) the reduced datasets (irisIB2.csv and letter-recognitionIB2.csv).

If you decide to use SQL, assume that the initial files were loaded as tables (iris and letter-recognition) and submit your SQL code that generates new tables with the edited or reduced content.