

Cervical Cancer Behaviour Risk

The goal of this assignment is to classify a sample into two categories as described in the dataset (given in this link: <https://archive.ics.uci.edu/ml/datasets/Cervical+Cancer+Behavior+Risk>). You need to apply condensed k-nearest neighbour (CNN in short) to classify the points. Use the last 15 examples in the dataset as the test set and the remaining examples as the training set. For CNN, while condensing, use only the closest neighbour (1-NN).

You must write the CNN code from scratch without using any CNN library. After condensation, set the value of k to 5 for performing classification on the test data. You should also compare the performance of CNN with standard k-NN on the same test set.

You need to submit the following files in your moodle account. Put all the files in a folder, zip it and then upload in moodle.

1. The code file.
2. Result file (.txt format). The result file must have the accuracy of standard K-NN, and your CNN on the test set.

Deadline: 30th January (11.55 pm IST).