

Workshop on Belief Functions Classification – Computer Project

Thierry Denœux

July-August 2017

The data file `Mode` in the R package `Ecdat` contains cross-sectional data about the choice of transportation mode by individuals. We want to predict the variable `choice` from the other cost and time of transportation modes, using the evidential K -NN rule and the evidential neural network classifier in the package `evclass`.

1. Remove the mode `'carpool'` as well as the variables `cost.carpool` and `time.carpool`.
2. Randomly split the data into a training set with approximately 2/3 of the observations, and a test set.
3. Classify the test data using the evidential K -NN rule with $K = 5$. Give the test error rate and the confusion matrix.
4. Plot the test error rate as a function of K . What is the best choice for K ?
5. Apply the evidential neural network to these data. Experiment with different numbers of prototypes.
6. Optional: optimize hyperparameter μ by cross-validation.