## Theory of belief functions: Application to machine learning and statistical inference Exercises on evidential classification

Thierry Denœux

Summer 2023

## 1 Transportation mode prediction

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 (EKNN) and the evidential neural network (ENN) classifiers in the package evclass.

- 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 EKNN rule with K = 5. Give the test error rate and the confusion matrix.
- 4. Plot the leave-one-out error rate as a function of K. What is the best choice for K?
- 5. Apply the ENN classifier to these data. Experiment with different numbers of prototypes.
- 6. Optimize hyperparameter  $\mu$  by cross-validation.
- 7. Compare the previous results with those of voting and multilayer perceptron classifiers (use packages FNN and nnet).

## 2 Credit scoring

1. Load the credit\_approval dataset. Remove the observations with missing values.

- 2. Split the data into training and test sets.
- 3. Plot the ROC curves of the EKNN and ENN classifiers with hyperparameters tuned by cross-validation.
- 4. Change the decision losses and observe the impact on the confusion matrix.
- 5. Decision with a reject option: for each of the two classifiers, plot the error rate vs. the rejection rate for different decision rules.