

University of Ottawa

School of Electrical Engineering and Computer Science

CSI5155 Fall 2019

Homework Exercises 1: Detecting Seismic Bumps

For this homework, please use the (quite imbalanced) seismic bumps dataset from the UCI Machine Learning Repository. (The direct link is <https://archive.ics.uci.edu/ml/datasets/seismic-bumps> or alternatively you may need to go to <https://archive.ics.uci.edu/ml/index.php> and enter *seismic bumps* in the keyword search function.)

The aim of this machine learning task is to predict seismic activity, given some historic measurements.

Instruction:

1. Submit your assignment using the uOttawa Virtual Campus, before the due date.
2. No late assignments will be accepted.
3. This is an individual assignment. Recall that assignments are awarded a participation mark and that you are required to answer all the questions, in order to obtain full marks.
4. Use *either* WEKA or Scikit-Learn to complete the assignments.

Question 1: Explorative Machine Learning

[5]

Import the data into your machine learning environment. Next, construct models using the following four (4) types of algorithms: a decision tree, a rule-based learning, a Naïve Bayesian classifiers and a k-nearest neighbor classifier. You should use 10 fold cross validation (the default setting in WEKA).

Submit the following:

- The four (4) models constructed by the algorithms
- The four (4) confusion matrix corresponding to the models.
- The *minority class* recall and precision for the four (4) models.

Question 2: ROC Curve

[5]

Draw the ROC Curve in order to contrast the four (4) models, using the software package or library of your own choice.

Question 3: Lessons Learned

[5]

Submit a 200 word summary discussing the results you obtained and the lessons you learned when analysing this data. Your summary should include a **decision** as to which one of the four (4) machine learning algorithms you would use as well as a **motivation** for your choice.