Lab: Decision Trees

Erick Ibarra
Carlos Gómez

Explain the advantages and disadvantages of writing a program on your own vs using a pre-created suite such as WEKA.

The main advantage of writing the program on your own is personalization. Making the program from the scratch ensure that all the functionality that you want in the system will be covered (if done right). Another advantage is learning. Since you must understand completely the concept before starting, when developing you ensure the understanding of the concepts. The main disadvantage of doing it your own is the time it will take to implement vs. the pre-created suite. The recommendation could be: Implement it yourself the first time to learn, and after that, use pre-created suites when they offer you the characteristics needed in the project.

Explain what criteria you followed to choose the datasets for your tree and the WEKA tests.

We selected the following dataset:
https://archive.ics.uci.edu/ml/datasets/Audiology+(Original)

The reasons to select it are the following: The data set characteristics is multivariable. The attribute characteristics are categorical, and the associated tasks is classification.

Apart from the characteristics of the data set, the main reason to use it was the fact that the data set is about a real problem, audiology, so using it let us see a real application of the trees, which make it more interesting since we can see our work having results.

Include the graphics of the trees or part of the trees you generated in WEKA and your own program. Are they different, and if so, why?

Based in what you have learned so far where would you use decision trees?

We take decisions every day, a lot of times a day, and many of those decisions are almost automatic. However, when we face a difficult decision that can have many outcomes and where are several things to consider as cost and utility. In this type of scenarios, decision trees are useful to analyze all the different possible outcomes, so the use basically would be in decision analysis. There are two key advantage in decision trees: data required little preparation, and they are easy to implement and explain.

Sometimes people use technology forgetting that it should serve people, not the other way, so as engineers our mission is to develop tech that can be use by common people and have a high contribution to their lives. This happen with decision trees since they are both effective, and easy to use.