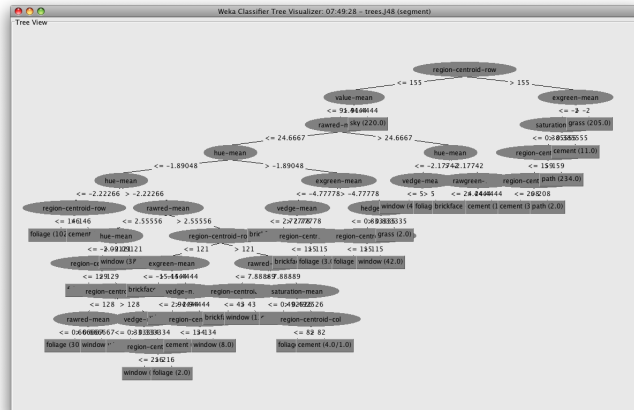
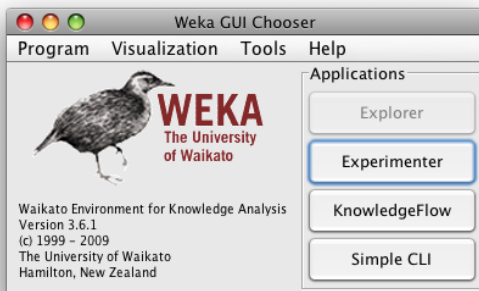


CMSC 360: Intelligent Robotics and Perception

Lab 4: Camels – Machine Learning

Due December 2, 2009



Overview

In the third unit of the course we are studying robots that rely on more complex perceptual and learning algorithms. In this lab, we will explore different machine learning algorithms. Machine learning has found applications in many different areas, but it's potential in robotics is particularly exciting. A robot's sensors provide rich streams of noisy data over extended periods of time. Algorithms that learn from experience, and that can adapt to changing environments, unburden us from pre-programming every little detail.

In this lab we will use the **weka**¹ open source machine learning environment. Weka is a Java framework for machine learning that implements many standard learning algorithms. Not only does it have a Java API, but it also provides an interactive environment for data visualization, clustering, and classification.

Learning Objectives

- Evaluate a supervised machine learning algorithm
- Evaluate an unsupervised machine learning algorithm
- Use an open-source machine learning package
- Access and process online datasets

¹<http://www.cs.waikato.ac.nz/ml/weka/>

Part One - Pick a Dataset

First, you will have to locate an interesting dataset. The more attributes and instances, the better. In your lab report, be sure to justify why you thought this dataset was interesting. University California Irvine (UCI) has a large number of machine learning datasets online². Also, a large number of robotics datasets have been made available under the Radish³ project. You are encouraged to use a dataset you collect and create (e.g. based on your semester project).

Part Two - Evaluate an Unsupervised Learner

Choose one the clustering or pre-processing algorithms (unsupervised learners) implemented in weka (e.g. principal components analysis, k-means) that looks interesting and seems reasonable for your dataset. Write a page or so describing how this particular algorithm works. Report the results of running this algorithm on the dataset you chose in part one. You should explore the algorithm by varying at least one of its critical parameters. For example, how the k-means clustering changes as you add more cluster centers.

Part Three - Evaluate a Supervised Learner

Choose one the classifiers (supervised learners) implemented in weka (e.g. decision tree (J48), neural network) that looks interesting and seems reasonable for your dataset. Write a page or so describing how this particular algorithm works. Report the results of running this algorithm on the dataset you chose in part one. You should explore the algorithm by varying at least one of its critical parameters. For example, how the classification performance of the neural network changes with the number of hidden nodes.

²<http://archive.ics.uci.edu/ml/>

³<http://radish.sourceforge.net/>