Program: M. Sc. Computer Science at University of Camerino
Course Name: Knowledge Engineering and Business Intelligence
Resource Person: Nadeem Qaisar Mehmood

Weka 3.6 Introduction

What is WEKA?

Waikato Environment for Knowledge Analysis. It's a data mining/machine learning tool developed by Department of Computer Science, University of Waikato, New Zealand. Weka is also a bird found only on the islands of New Zealand.

Download and Install

- Website: http://www.cs.waikato.ac.nz/~ml/weka/
- Support multiple platforms (written in java):
  o   Windows, Mac OS X and Linux

Main Features

  o 49 data preprocessing tools
    o 76 classification/regression algorithms
    o 8 clustering algorithms
    o 3 algorithms for finding association rules
    o 15 attribute/subset evaluators + 10 search algorithms for feature selection

LAB Live Demo!

Those who missed live demonstration of the software in the class can follow documentation at this page: www.cs.waikato.ac.nz/ml/weka/documentation.html. It contains presentation slides about WEKA interfaces and other features. You have to cover the following topics mainly.

Outline

Data

ARFF files, CVS (relation, attributes)

The Explorer:

Graphical User Interface

Preprocess data:

Using filters for cleaning, transformation and removing noise from the data. Performing preprocessing steps such as discretization, binary transformation etc.

Classification:

Learning Decision Trees such as ID3(does not handle continuous data), J48 which handles real values also), generation of the tree and transformation of it into rules.