

Goal

Create educator modules for teachers in any discipline to include relevant Machine Learning concepts.

- Identify Relevant Topics
- Produce Ready Modules
- Disseminate Online

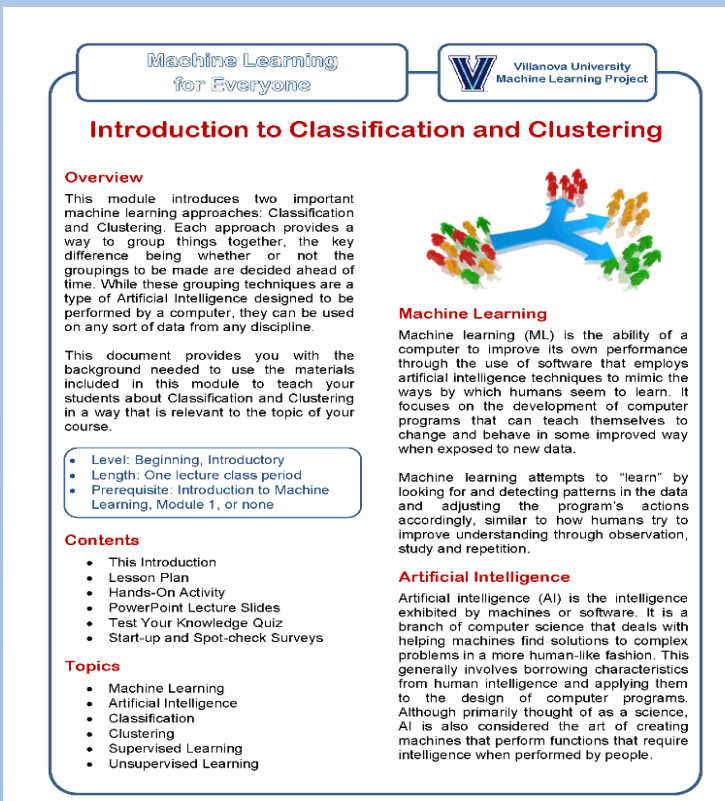
Module Example

Complete material to teach machine learning in context, understandable by non-computer scientists. Typical module contains:

- **Instructor Overview** - background, concepts
- **Handouts** - Activities
- **Data** - general and discipline specific
- **Evaluation** - pre/post tests, quizzes

Module Example: Classification & Clustering

Instructor Overview



Introduction to Classification and Clustering

Overview

This module introduces two important machine learning approaches: Classification and Clustering. Both approaches involve a way to group things together: the key difference being whether or not the groupings to be made are decided ahead of time. While these grouping techniques are a type of artificial intelligence, they can be used on any sort of data from any discipline.

Machine Learning

Machine learning (ML) is the ability of a computer to improve its own performance through the use of software that employs artificial intelligence techniques to mimic the ways by which humans learn to learn. It is based on the development of computer programs that can learn themselves to group and behave in some improved way when exposed to new data.

Artificial Intelligence

Artificial intelligence (AI) is the intelligence exhibited by machines or software. It is a branch of computer science that deals with helping machines find solutions to problems. Generally involves solving characteristics from human intelligence and applying them to the design of computer programs. Although primary thought of as a science, AI is also considered the art of making machines that perform functions that require intelligence when performed by people.

Contents

- This Introduction
- Lesson Plan
- Hands-On Activity
- Pre/post Test
- Handouts
- Final Knowledge Quiz
- Start-up and Post-test Surveys

Topics

- Machine Learning
- Artificial Intelligence
- Classification
- Clustering
- Supervised Learning
- Unsupervised Learning

Hands-on Activity

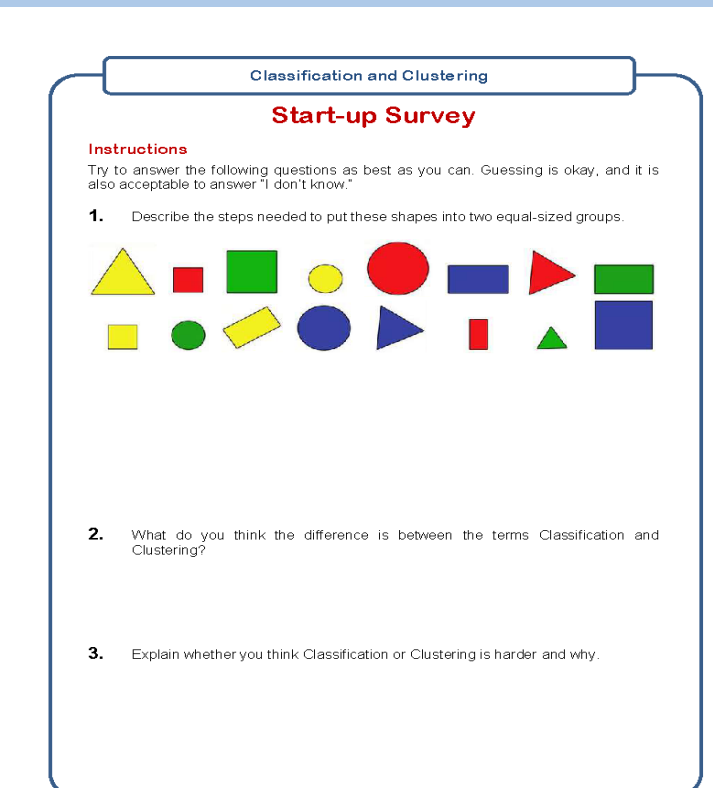


Example: Classification

Activity: Classification and Clustering

Create groups using example groups or by finding similarities

Pre & Post Test



Classification and Clustering Start-up Survey

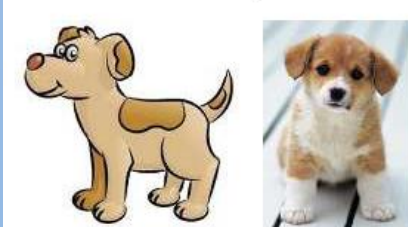
Instructions

Try to answer the following questions as best as you can. Guessing is okay, and it is also acceptable to choose "I don't know."

1. Describe the steps needed to put these shapes into two equal-sized groups.
2. What do you think the difference is between the terms Classification and Clustering?
3. Explain whether you think Classification or Clustering is harder and why.

Classification

Group A



Group B



Clustering

Group 1



Group 2



Clustering Again!

Group 1



Group 2



Group 3



Modules

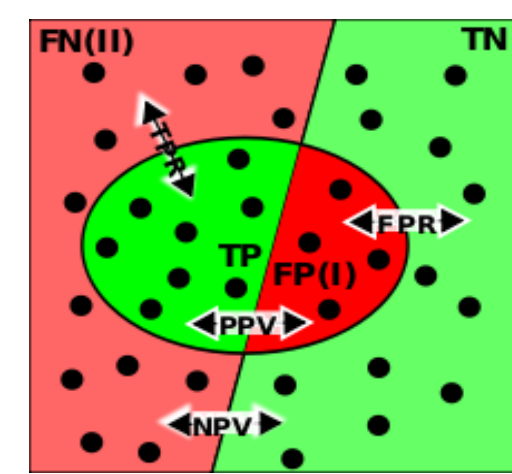
Intro

Fundamentals, background, Animal Game



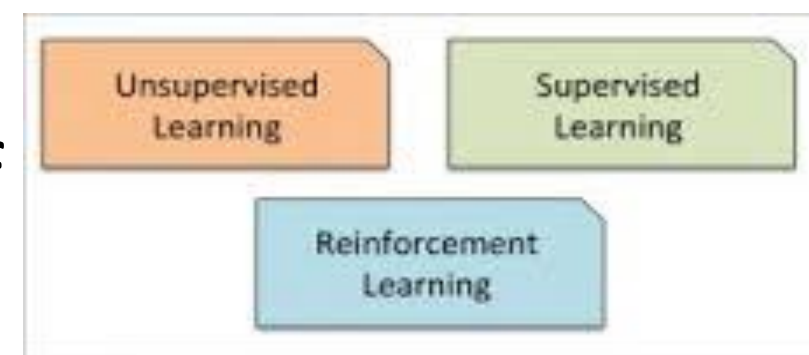
Evaluating Classifiers

How to compare results



Kinds of ML

Overview of ML areas



Decision Trees

Uses WEKA and data sets



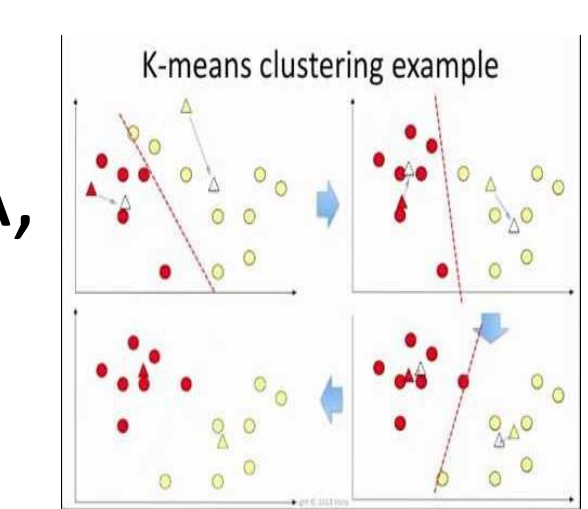
Text

Classification introduces WEKA, classify tweets, authors



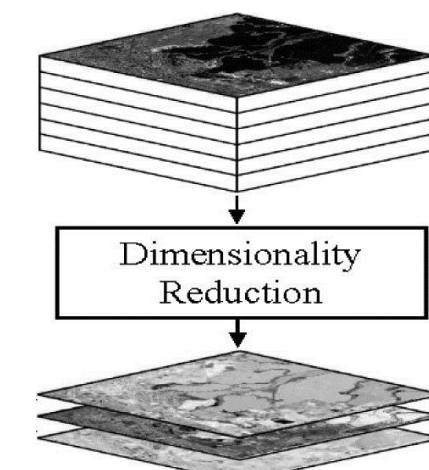
K - Means

Uses WEKA, advanced topic



Dimensionality Reduction

Uses WEKA, advanced topic



Choosing Inputs

Approaches to data sets



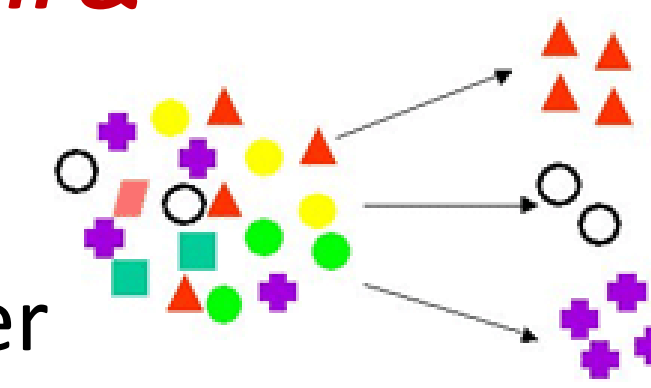
Neural Networks

Uses SimBrain software



Classification & Clustering

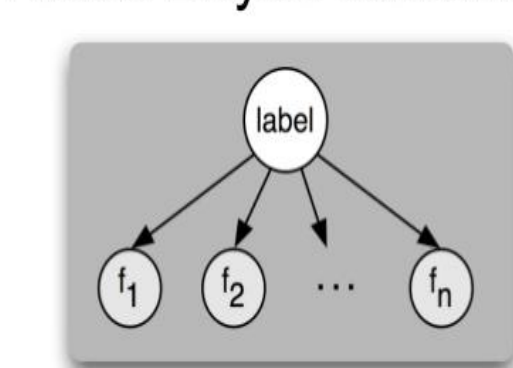
Hands-on, no computer needed



Naïve Bayes

Uses Python or WEKA

Naïve Bayes Classifier



Future Plans

- Complete model design
- Gather more domain-specific data sets
- Disseminate via: ComputingPortal.org/MachineLearning

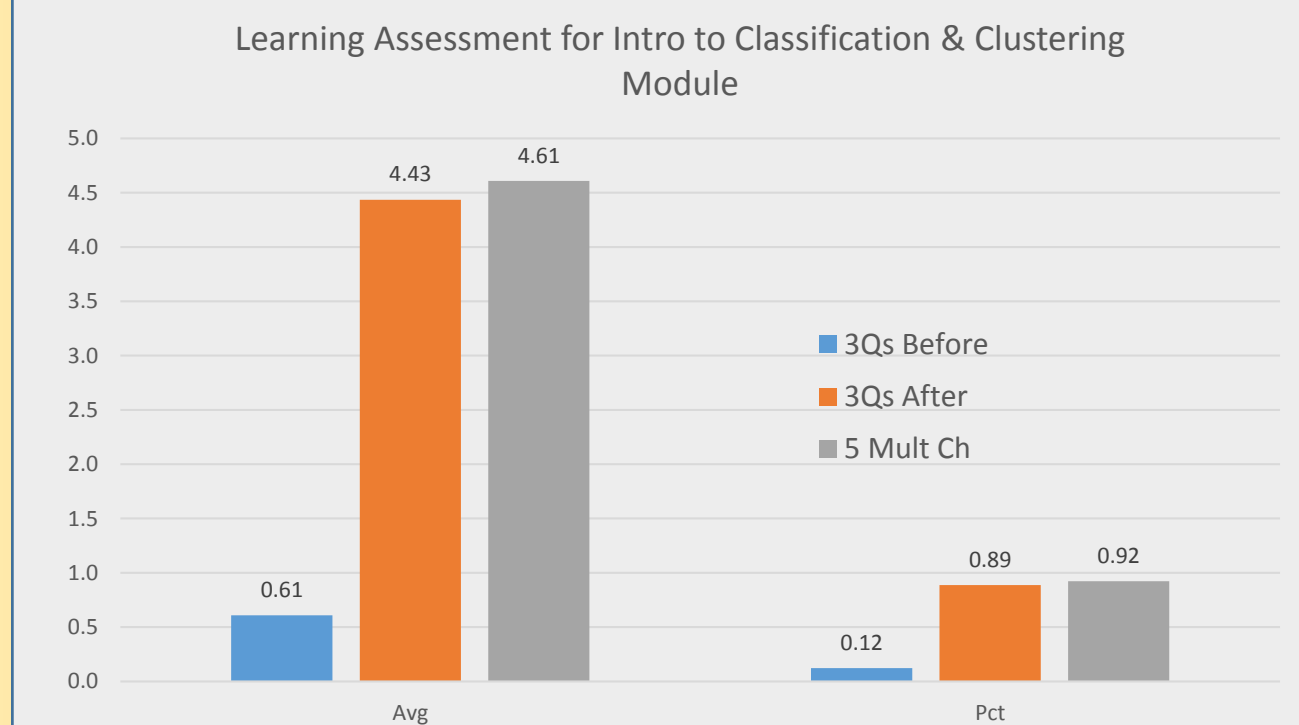
Machine Learning

Computer programs that find patterns in data, enabling them to "learn" and make decisions based on that learning.

Results

Results of pre/post tests for Classification & Clustering module show strong learning and good retention.

Pre test given **before** any material.
Post test given **2 days later**.



41 students pre/post test scores went from low (12%) to solid (89%) understanding on identical questions, with strong ability to apply knowledge (92%) to similar problems.

Acknowledgements

Project funded in part by NSF DUE award 1141033. Thanks to Carol Weiss for learning measure design and evaluation.

21st Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE)
11th – 13th July 2016, Arequipa, Peru