Module Outline

Introduction to Machine Learning

Mairéad Stephens | 17 July 2013 | 1.0

Introduction

This document is the course outline for the "Introduction to Machine Learning" course. This workshop runs over a half-day and comprises of a series of hands-on exercises and slide-based background material.

Attendees will leave with two key skills – how to cast their research problem as a machine learning problem, and how to determine an appropriate evaluation technique that is relevant to their research goals. They will also be exposed to and have a chance to experiment with a collection of machine learning algorithms.

Summary

This module is targeted at participants who:

- would like to take a first step into machine learning, either to apply it to their research or to add it to their arsenal of eResearch skills

This module assumes knowledge of:

- Basic mathematical ability

Description

Machine-learning is the automated, data driven discovery of programs or hypotheses that can perform tasks that are otherwise difficult to program. It includes sub-fields for classification, planning and parameter estimation, and has close ties to robotics, optimisation, medical diagnosis and expert systems. The field is one that is still rapidly evolving, but is well past the point where some of the key techniques are now available in a conveniently packaged form, both in terms of APIs as well as interactive workbenches.

This workshop will introduce some of the fundamental concepts of classification machine learning, data representation and how to design experiments that evaluate machine learning algorithms.

External Resources

In these modules, you will make use of some external materials. These materials were not created by Intersect, but are either standard tools used in the area or are particularly useful for teaching purposes.

Resources Used in this Module

For this module, the following resources are used

Location	Description
http://www.cs.waikato.ac.nz/ml/weka	ML Workbenches (WEKA)

Units Outline

Name	Concepts Introduced
Classification and Classifiers	This talk and exercises introduces the fundamentals of classification machine learning. By the end of this topic, attendees will understand • what a machine learning problem looks like, • some of the available arsenal of algorithms (nearest neighbour, SVMs, decision trees, information gain) and • the rudiments of using WEKA to run an algorithm on a problem
Evaluation	This talk and exercises introduces the main concepts in evaluation, including • Fundamental measurements of performance (accuracy, sensitivity, specificity, roc curves) • Choosing the right measure • Working out a test regime (n-fold, leave-one-out, folding)
Advanced Topics	This talk presents some of the next steps that interested attendees could follow up on. They include • Multiclass extensions to 2-class learning algorithms (e.g. 1-vs-1 and 1-vs-rest) • The challenges of data representation and management • Parallel and distributed machine learning

Revision History

Version No.	Revision Date	Summary of Changes	Revised by
1.0	17/07/2013	Initial draft	Mairéad Stephens