Fall 2010 Course

Instructor: Dr. Michael M. Richter:

CPSC 601.44
CPSC 601.45

Title: Applied Machine Learning

Description

The objective of both courses is to

- Introduce into essential aspects of machine learning
- Learn to perform its applications
-

Machine learning is a hot topic because it is applicable in many interesting and different areas
The course 6x is addressed to computer science students while the audience for the course 6y is coming from other departments.

Machine learning is a wide range computer science techniques of particular interest when other disciplines than are involved.
The topic is an umbrella term that covers a variety of techniques. There is supervised and unsupervised learning, symbolic and subsymbolic learning, clustering and data mining techniques.
The course will introduce into the general foundations, the algorithms and the practical applications where the latter provide the motivations.
Supervised symbolic learning is presented for concept learning, decision trees and inductive logic programming.. The formalisms are those of logic. Special attention is paid to evolutionary learning and genetic algorithms.
Unsupervised learning is concerned with clustering and knowledge discovery. Intermediate techniques are reinforcement learning, analogy learning and PAC-learning, learning of informal concepts and learning organizations. Subsymbolic learning is concerned with neural nets and other methods.
Besides the techniques the students shall learn which method is suitable for which task.

In order to get a feeling for the intended application areas there will be practical work in the assignment as part of the whole lecture. The students should

- Select an application of their interest. That means, they should have some experience in the area and there should be a particular problem they know something about. This problem could e.g. occur in a thesis or in an external project..
- Select learning method and a tool. For both support will be given.
- Implement learning techniques with visualization (if applicable).

There will be suggestions given for the application areas. Examples are given in www.cpsc.ucalgary.ca/~mrichter (Applied Machine Learning, Technical Report)

The chapters are:

- Introduction (5%)
- Concept Learning (5% of the course)
- Decision Trees (ID3 *C4.5 C5*) (10%)
- Unsupervised Learning of Concepts and Clustering (10%)
- Evolutionary Algorithms (10%)
- Reinforcement Learning (5%)
- PAC-Learning (5%)
- Inductive Logic Programming (ILP) (5%)

- Experience Based Methods (10%)
- Learning Informal Concepts (5%)
- Data Mining and Knowledge Discovery (5%)
- Support Vector Machines (5%)
- Neural Nets (10%)
- Data Preprocessing and Visualization (5%)

This is flexible, depending on the chosen topics for the projects. Some topics may be extended and others shortened. Not all of the material will be presented in the course. The selection is done according to the types of participating students and their chosen projects. Also, the ordering of the chapters depend on the topics of the projects.
Detailed slides for the course are available under
www.cpsc.ucalgary.ca/~mrichter
These slides contain all necessary material but additional references will be given. This will be updated at the end of August.
A recommended text is:
T. Mitchell: Machine Learning, McGraw Hill.
Also, the ordering of the material will be somehow flexible because the students get those contents related to the assignments early so that they can start to work on it after the first day of class.

The task of the assignment is:

The assignment will be in the form of a project. The project can be performed by one or two students.
The first step is to choosing and describing an application area together with a learning goal and an outline of the method types. This should be chosen in such a way that an understanding of the area is present.
The projects have to be formed in the first week of class and have to presented of the first week of class (e.g. as a ppt.presentation). This is, of course, very flexible.
This learning method has to be implemented (suggestions for tools are given; it is not recommended programming from scratch)

In the middle of the term the groups shall present a presentation in class with a progress report, describing the state of the project and the expected risk.
The project has to be presented at the end of the course in the form of a demo and an oral presentation; the implementation and the description have to be delivered on a CD.
Those students who take this as a graduate course should in addition give an evaluation of the used method.

I will be regularly available to discuss the progress of the projects.
There is no mid-term exam and final exam and no extra homework.