Data Mining, Neural Networks and Genetic Programming - Course Outline

COMP 422: 2013 Trimester 2

This document sets out the workload and assessment requirements for COMP 422. It also provides contact information for staff involved in the course. If the contents of this document are altered during the course, you will be advised of the change by an announcement in lectures and/or on the course web site. A printed copy of this document is held in the School Office.

Objectives

This course will address a variety of data mining concepts, algorithms, methods and techniques, particularly learning/adaptive methods. The practical projects in the course will enable you to gain experience with a number of these algorithms and methods and explore how they work and their limitations. The exam will assess your understanding of the relevant concepts, theories and algorithms.

This course will also provide the background for doing a research project or a thesis in data mining and machine learning related areas, particularly neural networks and genetic programming techniques on pattern recognition tasks, such as classification and clustering problems.

By the end of the course, students should be able to achieve the following objectives:

- Understand the key concepts, theories, and tasks of data mining and knowledge discovery in databases (KDD). (BE 3(a)); (BSc COMP 3)
- Understand the main strengths and limitations of commonly used data mining algorithms and how to apply them to (Engineering) applications. (BE 3(a), 3(d)); (BSc COMP 1, 3)
- Understand key concepts and tasks in computer vision and image processing. (BE 3(a)); (BSc COMP 2, 3)
- Select/develop good features and algorithms for object recognition, particularly classification. (BE 3(a), 3(b)); (BSc COMP 1, 2, 3, 4)
- Use neural networks and genetic programming techniques for data mining tasks such as regression and classification. (BE 3(a), 3(b), 3(d)); (BSc COMP 1, 2, 3, 4)
- Select an appropriate criterion to evaluate a data mining/learning system such as a classifier. (BE 3(a), 3(d)); (BSc COMP 2, 4)
- Learn/practise oral communication skills via presentations or demonstrations and writing skills via project reports. (BE 2(b)); (BSc COMP 4)

Textbook

Due to the nature of this course (crossing different areas), we do not have a particular text book. However, there are some books in the library for Data Mining, Neural Networks and Genetic Programming.

- Wolfgang Banzhaf, Peter Nordin, Robert, E. Keller, and Frank, D. Francone. Genetic programming: an introduction on the automatic evolution of computer programs and its applications. Morgan Kaufmann Publishers. 1998. ISBN: 1-55860-510-X.
- Margaret H. Dunham. Data Mining: Introductory and Advanced Topics. Prentice Hall, 2003.
- A. Forsyth and Jean Ponce. Computer Vision: A Modern Approach. Prentice Hall, 2003.
- Earl Gose, Richard Johnsonbaugh, and Steve Jost. Pattern Recognition and Image Analysis. Prentice Hall PTR, Upper Saddle River, NJ 07458, 1996. ISBN 0-13-236415-8.
- John R. Koza. Genetic programming : on the programming of computers by means of natural selection. MIT Press, London, England, 1992.
- John R. Koza. Genetic programming II: automatic discovery of reusable programs. Cambridge, Mass. : MIT Press, London, England, 1994.

- T. Mitchell. Machine Learning. McGraw-Hill, New York, 1997.
- Stuart J. Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall, Pearson Education Inc., 2nd edition, 2003.
- Sergios Theodoridis, Konstantinos Koutroumbas. Pattern recognition (4th edition). Academic Press. 2009.
- Richard Jensen, Qiang Shen. Computational intelligence and feature selection : rough and fuzzy approaches.
- Michael Affenzeller. Genetic algorithms and genetic programming : modern concepts and practical applications. CRC Press, 2009.

In addition, many other materials will be posted on the course web site (Useful Links).

Lectures, Tutorials, Laboratories, and Practical work

A schedule of lecture topics, readings, and assignment due dates is available online

Lectures for COMP 422 are:

- Mondays 4:10 - 5:00pm, MY 303
- Tuesdays 4:10 - 5:00pm, MY 303
- Wednesdays 4:10 - 5:00pm, MY 303

Note that some of the lecture slots may be used as tutorials, demonstrations, helpdesks or project work. As a 400 level course, there will be no scheduled labs.

Projects

There will be two projects. The first is about basic data mining algorithms and vision applications, which is related to the first four objectives and the last objective of this course. The second project is to use neural network and genetic programming (evolutionary computation) techniques for data mining application tasks, particularly regression and classification tasks. This project is related to the last three objectives of the course. The two projects will involve both programming and writing, and they will need to be submitted via the School online *web submission system* and also to the School *handin boxes*. Details of the submission methods will be described in the project descriptions, which will be posted onto our course web site.

Workload

In order to maintain satisfactory progress in COMP 422, you should plan to spend an average of at least *10* hours per week on this course. A plausible and approximate breakdown for these hours would be:

- Lectures and tutorials: 3 hours
- Readings and projects: 7 hours

School of Engineering and Computer Science

The School office is located on level three of the Cotton Building (Cotton 358).

Staff

The course coordinator and lecturer for COMP 422 is Mengjie Zhang. His contact details are:

- *Mengjie Zhang*
- Cotton 355
- +64 4 463 5654
- Mengjie.Zhang@ecs.vuw.ac.nz

Tutor details: Su Nguyen and Bing Xue will serve as a tutor/casual lecturer of this course. Their email addresses are su.nguyen@ecs.vuw.ac.nz, bing.xue@ecs.vuw.ac.nz.

You can discuss course related issues with Mengjie Zhang, Su Nguyen, and/or Bing Xue by email or in person.

Announcements and Communication

The main means of communication outside of lecture will be the COMP 422 web area at http://ecs.victoria.ac.nz/Courses/COMP422_2013T2/. There you will find, among other things, this document, the lecture notes and project handout, and the COMP 422 Forum. The forum is a web-based bulletin board system providing a discussion platform between students. Questions and comments can be posted to the forum and other students can reply to the questions. Staff will also read these posts and respond to them from time to time if necessary. If you want to get a answer from the Staff, you need to email them directly.

Assessment

Your grade for COMP 422 will be determined based on the following assessment weightings:

Item	Weight
Project 1	20%
Project 2	20%
Final Examination	60%

Exams

The timetable for final examinations will be available from the University web site and will be posted on a notice board outside the faculty office. The final examination will be three hours long. No computers, electronic calculators or similar devices will be allowed in the final examination. Paper non-English to English dictionaries will be permitted. Non-programmable calculators will be allowed. The examination period for trimester 2 is 25 October - 16 November.

Practical Work

Description of projects, including rough dates and submission processes

There are two projects in COMP 422, each of which is worth 20% of the final grade. The first project will be tentatively due on Monday 19 August 2012 (Week 6). The second project is tentatively due on Monday 14 October 2012 (week 12).

Policies and penalties for late submission

Requests for extension of the deadline will only be granted in rare circumstances. Unless an arrangement has been approved, projects handed in late will be penalised 10% per day, and will not be accepted beyond a week after the deadline.

Plagiarism

We encourage you to discuss the principles of the course and assignments/projects with other students, to help and seek help with programming details and problems involving the lab machines. However, any work you hand in must be your own work.

The School policy on Plagiarism (claiming other people's work as your own) is available from the course home page. Please read it. We will penalise anyone we find plagiarising, whether from students currently doing the course, or from other sources. Students who knowingly allow other students to copy their work will also be penalised. If you have had help from someone else (other than a tutor), it is always safe to state the help that you got. For example, if you had help from someone else in writing a component of your code, it is not plagiarism as long as you state (eg, as a comment in the code) who helped you in writing the method.

Mandatory Requirements

- Submit both projects;
- Obtain a D grade on the final exam.

Passing COMP 422

To pass COMP 422, a student must satisfy mandatory requirements and gain at least a **C** grade overall.

Withdrawal

The last date for withdrawal from COMP 422 with entitlement to a refund of tuition fees is Friday 26 July 2013. The last date for withdrawal without being regarded as having failed the course is Friday 27 September 2013 -- though later withdrawals may be approved by the Dean in special circumstances.

Rules & Policies

Find key dates, explanations of grades and other useful information at http://www.victoria.ac.nz/home/study.

Find out about academic progress and restricted enrolment at http://www.victoria.ac.nz/home/study/academic-progress.

The University's statutes and policies are available at http://www.victoria.ac.nz/home/about/policy, except qualification statutes, which are available via the Calendar webpage at http://www.victoria.ac.nz/home/study/calendar (See Section C).

Further information about the University's academic processes can be found on the website of the Assistant Vice-Chancellor (Academic) at http://www.victoria.ac.nz/home/about/avcacademic

All students are expected to be familiar with the following regulations and policies, which are available from the school web site:

Grievances
Student and Staff Conduct
Meeting the Needs of Students with Disabilities
Student Support
Academic Integrity and Plagiarism
Dates and Deadlines including Withdrawal dates
School Laboratory Hours and Rules
Printing Allocations
Expectations of Students in ECS courses

The School of Engineering and Computer Science strives to anticipate all problems associated with its courses, laboratories and equipment. We hope you will find that your courses meet your expectations of a quality learning experience.

If you think we have overlooked something or would like to make a suggestion feel free to talk to your course organiser or lecturer.

Course Outline as PDF