



Prescription

This course examines the construction of intelligent agents - software programs that can act for themselves in some part of the human world. This course focuses on agents for improving web search and includes topics such as agents for information extraction from the web, web page clustering and classification, automatic query expansion and web page ranking.

Course learning objectives

Students who pass this course should be able to:

1. achieve an understanding of the basic problems and basic principles in a variety of related research areas such as information retrieval, information extraction, clustering and classification and natural language processing. (Graduate Attributes: BE 3(a), 3(c), 3(d), 3(e); BSc COMP 1, 2, 3, 4)
2. achieve greater skill at reading, understanding, and giving presentations on papers from the research literature. (Graduate Attributes: BE 2(b); BSc COMP 2, 4)
3. achieve practical experience of building text mining systems. (Graduate Attributes: BE 3(a); BSc COMP 1)

Course content

The course focuses on text mining and Web search. We will cover a variety of topics including text representation, document classification and clustering, opinion mining, information retrieval, recommender systems, query expansion, and information extraction.

Withdrawal from Course

Withdrawal dates and process:

<https://www.wgtn.ac.nz/students/study/course-additions-withdrawals>

Lecturers

Xiaoying Gao (Coordinator)

Xiaoying.Gao@vuw.ac.nz

Teaching Format

During the trimester there will usually be two lectures per week on Tuesdays and Thursdays. Wednesday lecture time is only occasionally used for overflow and if this is the case, it will be announced in Tuesday

lectures and on our course home page.

Student feedback

Student feedback on University courses may be found at:
www.cad.vuw.ac.nz/feedback/feedback_display.php

Dates (trimester, teaching & break dates)

- Teaching: 02 March 2020 - 07 June 2020
- Break: 13 April 2020 - 27 April 2020
- Study period: 08 June 2020 - 11 June 2020
- Exam period: 12 June 2020 - 27 June 2020

Class Times and Room Numbers

02 March 2020 - 22 March 2020

- **Tuesday** 13:10 - 14:00 – 203, 83 Fairlie Tce, Kelburn
- **Wednesday** 13:10 - 14:00 – 203, 83 Fairlie Tce, Kelburn
- **Thursday** 13:10 - 14:00 – 203, 83 Fairlie Tce, Kelburn

27 April 2020 - 07 June 2020

- **Tuesday** 13:10 - 14:00 – 203, 83 Fairlie Tce, Kelburn
- **Wednesday** 13:10 - 14:00 – 203, 83 Fairlie Tce, Kelburn
- **Thursday** 13:10 - 14:00 – 203, 83 Fairlie Tce, Kelburn

Other Classes

No other classes.

Set Texts and Recommended Readings

Required

There are no required texts for this offering.

Mandatory Course Requirements

There are no mandatory course requirements for this course.

If you believe that exceptional circumstances may prevent you from meeting the mandatory course requirements, contact the Course Coordinator for advice as soon as possible.

Assessment

The assessment of the course will be based on several assignments and a final exam. All these are designed for our learning objectives outlined above and the marking is based on how well the students achieve these objectives. The assignments will consist of:

- Joining the discussions and doing the peer review.
- Giving a presentation on one paper or a self-chosen topic. The student can select a presentation time.
- Writing a review on a self chosen paper, due Week 10 Friday 5pm. (word limit: 400)
- Completing a project on text classification. This project will be explained in lectures. The baseline code is due week 4 Friday 5pm and the full code is due week 7, Monday 9am. Its assessment is based on a demonstration of the built system in lectures and a written report due week 8, Friday 5pm (Page limit:2).

Assessment Item	Due Date or Test Date	CLO(s)	Percentage
Group discussion and peer review		CLO: 1,2	3%
Presentation	Sign up time	CLO: 1,2	10%
Project baseline	week 4	CLO: 1,2,3	5%
Project full code	Week 7	CLO: 1,2,3	5%
Project report	Week 8	CLO: 1,2,3	10%
Paper review	week 10	CLO: 1,2	7%
Final examination (2 hours duration)	exam period	CLO: 1,2,3	60%

Penalties

Any assignment submitted after the deadline will be penalised by 20% per day of the full assignment marks. Individual extensions will only be granted in exceptional personal circumstances. We have a late days policy to cover minor problems.

LATE DAYS POLICY: Each student will have three "late days" which you may choose to use for any assignment or assignments during the course. There will be no penalty applied for these late days. You do not need to apply for these, instead any late days you have left will be automatically applied to assignments that you submit late.

Extensions

Individual extensions will only be granted in exceptional personal circumstances, and should be negotiated with the course coordinator before the deadline whenever possible. Documentation (eg, medical certificate) may be required.

Submission & Return

All work is submitted through the ECS submission system, accessible through the course web pages. Marks and comments will be returned using our online assessment system.

Group Work

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

Workload

In order to maintain satisfactory progress in COMP 423, you should plan to spend an average of 10 hours per week on this paper. A plausible and approximate breakdown for these hours would be:

- Lectures: 2
- Readings: 3
- Assignments: 5

Teaching Plan

See: https://ecs.wgtn.ac.nz/Courses/COMP423_2020T1/

Communication of Additional Information

All online material for this course can be accessed at https://ecs.wgtn.ac.nz/Courses/COMP423_2020T1/

Links to General Course Information

- Academic Integrity and Plagiarism: <https://www.wgtn.ac.nz/students/study/exams/integrity-plagiarism>
- Academic Progress: <https://www.wgtn.ac.nz/students/study/progress/academic-progress> (including restrictions and non-engagement)
- Dates and deadlines: <https://www.wgtn.ac.nz/students/study/dates>
- Grades: <https://www.wgtn.ac.nz/students/study/progress/grades>
- Special passes: Refer to the Assessment Handbook, at <https://www.wgtn.ac.nz/documents/policy/staff-policy/assessment-handbook.pdf>
- Statutes and policies, e.g. Student Conduct Statute: <https://www.wgtn.ac.nz/about/governance/strategy>
- Student support: <https://www.wgtn.ac.nz/students/support>
- Students with disabilities: https://www.wgtn.ac.nz/st_services/disability/
- Student Charter: <https://www.wgtn.ac.nz/learning-teaching/learning-partnerships/student-charter>
- Terms and Conditions: <https://www.wgtn.ac.nz/study/apply-enrol/terms-conditions/student-contract>
- Turnitin: <http://www.cad.vuw.ac.nz/wiki/index.php/Turnitin>
- University structure: <https://www.wgtn.ac.nz/about/governance/structure>
- VUWSA: <http://www.vuwsa.org.nz>

Offering CRN: [4962](#)

Points: 15

Prerequisites: COMP 307, one further 300-level COMP, ECEN, NWEN or SWEN course

Duration: 02 March 2020 - 28 June 2020

Starts: Trimester 1

Campus: Kelburn