

CIS*6650*02 – Data Mining Course Outline – Fall 2017

Instructor: Luiza Antonie

Office : MacLachlan Building, Room 225

Phone: 519-824-4120 ext. 54979

Email: lantonie@uoguelph.ca

Timetable

Lecture: Monday, Wednesday 14:30-15:50, MCKN, Room 310

Course Description

With the continuous growth of digital collections, the research field of data mining has gained a substantial interest among practitioners in a variety of fields and industries. Nowadays, almost every company collects and stores data, with the purpose of data analysis to support better decision making. Data mining algorithms and techniques are used in various domains including market basket analysis, medical applications, business management, space exploration, financial data analysis, etc. This course introduces concepts, algorithms and techniques of data mining.

Lecture Topics

- Introduction to data mining.
- Mining frequent patterns and association rules.
- Classification.
- Cluster and outlier analysis.
- Data cleaning and integration.
- Applications.

Course Materials

No textbooks. Lecture slides will be provided. Published research papers will be recommended for reading and presentation.

The following books are recommended:

1. Data Mining: Concepts and Techniques
Jiawei Han, Micheline Kamber and Jian Pei
Morgan Kaufmann, , 3rd ed. 2011.

2. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations
I. H. Witten and E. Frank
Morgan Kaufmann, 2nd ed. 2005.

3. Introduction to Data Mining
P.-N. Tan, M. Steinbach, and V. Kumar
Addison Wesley, 2005.

Course Web Page:

<http://courselink.uoguelph.ca>

Tentative Schedule

Date	Lecture	Assignments
Sep 11th	Administrative Introduction to Data Mining	
Sep 18th	Introduction to Data Mining Know Your Data	
Sep 25th	Mining Frequent Patterns and Association Rules	Assignment 1 Out
Oct 2nd	Mining Frequent Patterns and Association Rules Classification	
Oct 9th	Classification Paper Presentation and Discussion	Assignment 1 Due Assignment 2 Out
Oct 16th	Classification	Project Proposal Due
Oct 23rd	Cluster and Outlier Analysis Paper Presentation and Discussion	Assignment 2 Due Assignment 3 Out
Oct 30th	Cluster and Outlier Analysis	
Nov 6th	Data Cleaning and Integration Paper Presentation and Discussion	Assignment 3 Due
Nov 13th	Data Cleaning and Integration Advanced Concepts	Project Progress
Nov 20th	Applications	
Nov 27th	Project Presentations	
Dec 15th		Project Report Due

Course Assessment

			Learning Outcomes
Assessment 1:	Assignments (3)	30%	2,5,6
Assessment 2:	Paper Presentation	15%	1,3,4
Assessment 3:	Final Project	45%	2,5,6
Assessment 4:	Class Participation	10%	1,2,3,4
Total:		100%	

Notes:

- All assignments will be posted on the course web page and they will be submitted via DropBox on CourseLink.
- All students will benefit from a 48 hour grace period after the original due date of the assignments. In all but exceptional circumstances, this should allow enough flexibility to accommodate any unforeseen events that could otherwise impact your work. In fairness to the vast majority of students who are responsible in handing in their work on time, a deduction of 10% will be applied on the first minute that the assignment is handed in after the grace period. An additional 20% deduction will be applied every 24 hours thereafter.
- It is academic misconduct to collaborate on assignments which do not explicitly allow for group work. It is also misconduct to represent other student's work as your own. There are penalties for doing this. All parties to misconduct will be reported to the Director of the School of Computer Science as participants in academic misconduct. See Academic Misconduct in Section VIII of the Undergraduate Calendar for the causes and penalties of misconduct.
- Failure to submit assignments correctly will result in a substantial loss of marks.

Course Learning Outcomes

Knowledge and Understanding:

1. The student should be able to describe fundamental data mining tasks like frequent pattern mining, classification, and clustering.
2. The student should be able to demonstrate the ability to understand and implement key algorithms for fundamental data mining tasks.
3. The student should be able to read, understand and communicate research papers on topics related to the course content.

Skills:

4. The student should be able to effectively communicate the concepts covered in the course.
5. The student will develop problem solving skills, show creativity and critical thinking working on the assignments.
6. The student will develop skills to implement and apply the techniques learnt in class to real world applications.

University Policies

Academic Consideration

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons, please advise the course instructor in writing, with your name, id#, and e-mail contact. See the academic calendar for information on regulations and procedures for Academic Consideration: <http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml>

Academic Misconduct

The University of Guelph is committed to upholding the highest standards of academic integrity and it is the responsibility of all members of the University community, faculty, staff, and students to be aware of what constitutes academic misconduct and to do as much as possible to prevent academic offences from occurring. University of Guelph students have the responsibility of abiding by the University's policy on academic misconduct regardless of their location of study; faculty, staff and students have the responsibility of supporting an environment that discourages misconduct. Students need to remain aware that instructors have access to and the right to use electronic and other means of detection. Please note: Whether or not a student intended to commit academic misconduct is not relevant for a finding of guilt. Hurried or careless submission of assignments does not excuse students from responsibility for verifying the academic integrity of their work before submitting it. Students who are in any doubt as to whether an action on their part could be construed as an academic offence should consult with a faculty member or faculty advisor. The Academic Misconduct Policy is detailed in the Undergraduate Calendar: <https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml>

Accessibility

The University of Guelph is committed to creating a barrier-free environment. Providing services

for students is a shared responsibility among students, faculty and administrators. This relationship is based on respect of individual rights, the dignity of the individual and the University community's shared commitment to an open and supportive learning environment. Students requiring service or accommodation, whether due to an identified, ongoing disability or a short-term disability should contact the Centre for Students with Disabilities as soon as possible. For more information, contact CSD at 519-824-4120 ext. 56208 or email csd@uoguelph.ca or see the website: <http://www.csd.uoguelph.ca/csd/>

Drop date

The last date to drop one-semester courses, without academic penalty, is Friday, November 3, 2017. For regulations and procedures for Dropping Courses, see the Academic Calendar:

<https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c03/c03-fallsem.shtml>

Policy on Re-grade Request

Any request to remark an exam, assignment or quiz must be addressed to your instructor in writing (hard copy) within one week and must include the following (i) your name, contact information (telephone and email), and signature, (ii) a clear description of where and why you feel that you were graded in error, and (iii) the following statement exactly as it appears here "In requesting a re-grade I promise that I have not written on, erased, or in any way changed my copy of the exam/assignment since it was handed back to me. I understand that to do so would constitute a serious academic offence. I also understand that my entire exam/assignment will be re-graded (with particular attention paid to the points that I have brought up) and understand that as a result my score on the exam/assignment could fall as well as rise."

E-mail Communication

As per university regulations, all students are required to check their *uoguelph.ca* e-mail account regularly: e-mail is the official route of communication between the University and its students. When contacting the TAs and instructor by e-mail, please be sure to use your University of Guelph e-mail address. We apologize not to be able to reply to e-mail inquiries sent from outside the university.

Recording of Materials: Presentations which are made in relation to course work, including lectures, cannot be recorded or copied without the permission of the presenter, whether the instructor, a classmate or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.