

ECE657A: Data and Knowledge Modeling and Analysis

Winter 2018

University of Waterloo, Faculty of Engineering
Department of Electrical and Computer Engineering

Lectures:

January 3 – March 28, 2018

Wednesdays, 11:30am – 2:20pm in room E5 5106/5128

Course TAs:

Instructor:

Prof. Mark Crowley, E5 4114

mcrowley@uwaterloo.ca

Office hours: arrange for appointment by email

Sriram Ganapathi Subramanian email

s2ganapa@uwaterloo.ca

Office hours: by appointment

Course Description

Engineers encounter data in many of their tasks. Whether the sources of this data may be experiments, databases, computer files, or the Internet, there is a dire need for effective methods to model and analyze the data and extract useful knowledge and information from it. This course aims to provide engineering graduate students with essential knowledge of data representation, grouping, mining and knowledge discovery.

Recommended background: *Data Structures and Algorithms, Probability*, or consent of instructor.

Evaluation:

5% Homeworks (Weekly, short question, due night of lecture)

15% Assignments (two assignments, in groups of up to 3 students)

30% Group Project: (5% Proposal, 8 % presentation, 2% proposal/presentation peer reviews, 15% report)

50% Final Exam

Major Topics:

1. Data types, sources, nature, scales and distributions
2. Data representations, transformation, feature extraction and selection, dimensionality reduction and normalization
3. Classification: Distance based, Decision Tree based, Statistical based, Deep Learning based.
4. Clustering: Partition-based, Hierarchical, Model and Density based.
5. Reinforcement Learning: Classification, Control and learning patterns over time.
6. Retrieval and Mining: Similarity measures and matching techniques.
7. Knowledge discovery in data: Association rules mining, web mining, text mining.
8. Performance measures and tools: Statistical Analysis, Validity and Assessment Measures.

Additional Resources and Links:

- See resources posted to LEARN
- See resources updated and posted to external course website:
<https://markcrowley.ca/teaching/ece-657a-data-and-knowledge-modeling-and-analysis/>

Textbooks:

There is no required textbook. But most of the course is based on the following books and will be useful to take a look at them.

1. K. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press, 2012.
2. I. Goodfellow, Y. Bengio and A. Courville. Deep Learning. MIT Press, 2016.
Online for free at <http://www.deeplearningbook.org> (first half covers many of basics of this course, second half focusses on Deep Learning which we will talk about for a couple lectures briefly.)
3. M. Dunham. Data Mining Introductory and Advanced Topics, ISBN: 0130888923, Prentice Hall, 2003.
4. R. O. Duda, P. E. Hart and D. G. Stork. Pattern Classification (2nd ed.), John Wiley and Sons, 2001.

Papers and electronic references will be made available on the course website which is on LEARN (go to <http://learn.uwaterloo.ca> to log in).

Recipe for success:

Attend lectures. Do weekly homeworks in class. Do complementary work at home. Ask questions. Most of all, have fun.

Policy and Rules

Academic Integrity:

In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. [Check www.uwaterloo.ca/academicintegrity/ for more information.]

Grievance:

A student who believes that a decision affecting some aspect of his/her university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70, Student Petitions and Grievances, Section 4, <http://www.adm.uwaterloo.ca/infosec/Policies/policy70.htm>. When in doubt please be certain to contact the department's administrative assistant who will provide further assistance.

Discipline:

A student is expected to know what constitutes academic integrity to avoid committing academic offenses and to take responsibility for his/her actions. A student who is unsure whether an action constitutes an offense, or who needs help in learning how to avoid offenses (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course professor, academic advisor, or the undergraduate associate dean. For information on categories of offenses and types of penalties, students should refer to Policy 71, Student Discipline, <http://www.adm.uwaterloo.ca/infosec/Policies/policy71.htm>. For typical penalties check Guidelines for the Assessment of Penalties, <http://www.adm.uwaterloo.ca/infosec/guidelines/penaltyguidelines.htm>.

Plagiarism-detection software may be used on any submitted work.

Appeals:

A decision made or penalty imposed under Policy 70, Student Petitions and Grievances (other than a petition) or Policy 71, Student Discipline may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72, Student Appeals, <http://www.adm.uwaterloo.ca/infosec/Policies/policy72.htm>.

Note for students with disabilities:

The Office for Persons with Disabilities (OPD), located in Needles Hall, Room 1132, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with the OPD at the beginning of each academic term.