

Course Outline

1. COURSE INFORMATION

Session Offered	Fall 2015	
Course Name	Data Mining	
Course Code	COMP TECH 4DM3	
Date(s) and Time(s) of lectures	Monday September 14 – Monday December 7 at 6:30 pm	
Program Name	Software Engineering Technology	
Calendar Description	Classification, association, prediction and clustering of data. Decision Trees. Bayesian probability. Supervised and unsupervised learning.	
Instructor(s)	Jeff Fortuna	E-Mail: fortunjj@mcmaster.ca Office Hours & Location: Office hours by request – ETB 211

2. COURSE SPECIFICS

Course Description			
Instruction Type	Code	Type	Hours per term
	C	Classroom instruction	36
	L	Laboratory, workshop or fieldwork	
	T	Tutorial	
	DE	Distance education	
	Total Hours		36
Resources	ISBN	Textbook Title & Edition	Author & Publisher
	ISBN: 9781617290183	Required: Machine Learning in Action	Author: Peter Harrington Published by Manning
	ISBN: 9780387848570	Supplemental: The Elements of Statistical Learning	Authors: Trevor Hastie, Robert Tibshirani, Jerome Friedman Published by Springer
	Other Supplies	Source	
	Python interpreter		
Prerequisite(s)	<i>None</i>		
Corequisite(s)	<i>None</i>		
Antirequisite(s)	<i>None</i>		
Course Specific Policies	<p>The basic website for the textbook, which includes all of the source code for each chapter is at:</p> <p>http://www.manning.com/pharrington/</p> <p>There are number of minor errors in the source code. Therefore it is important to look through the forums dedicated to the text:</p> <p>Textbook Forums</p> <p>Projects must be completed individually. A write-up, along with source code, must be submitted by each student. Particular emphasis will be made on the</p>		

	<p>ability of the student to show why each of the three algorithms behaved in the way it did relative to the others used.</p> <p>Brief assignments will be provided following each lecture covering the theory component covered in the lecture. These assignments will be graded on a 3 point scale – 0 for not submitted, 1 for partial completion, 2 for full completion. The completion of these assignments is critical in order to ensure success on the midterm and final exams.</p>
Departmental Policies	<p>Students must maintain a GPA of 3.5/12 to continue in the program.</p> <p>In order to achieve the required learning objectives, on average, B.Tech. students can expect to do at least 3 hours of “out-of-class” work for every scheduled hour in class. “Out-of-class” work includes reading, research, assignments and preparation for tests and examinations.</p> <p>Where group work is indicated in the course outline, such collaborative work is mandatory.</p> <p>The use of cell phones, iPods, laptops and other personal electronic devices are prohibited from the classroom during the class time, unless the instructor makes an explicit exception.</p> <p>Announcements made in class or placed on Avenue are considered to have been communicated to all students including those individuals that are not in class.</p> <p>Instructor has the right to submit work to software to identify plagiarism.</p>
3. SUB TOPIC(S)	
Week 1	N/A
Week 2	Data descriptions, visualization of data, measurements of data similarity and dissimilarity
Week 3	K Nearest Neighbors Classification
Week 4	Decision Trees
Week 5	Bayesian Classifiers
<i>Mid-term recess (Monday, October 12 to Saturday, October 17)</i>	
Week 6	Logistic Regression Classifier
Week 7	Midterm Exam
Week 8	Support Vector Machines
Week 9	Boosting and AdaBoost
Week 10	Data Forecasting with Regression
Week 11	Unsupervised Learning
Week 12	Classifier Evaluation
Week 13	Review
<p>Classes end – Tuesday December 8, 2015</p> <p>Final examination period: Wednesday December, 9, 2015 to Tuesday, December 22, 2015</p> <p>All examinations MUST BE written during the scheduled examination period.</p>	
<p>Note that this structure represents a plan and is subject to adjustment term by term. The instructor and the University reserve the right to modify elements of the course during the term. The University may change the dates and deadlines for any or all courses in extreme circumstances. If either</p>	

type of modification becomes necessary, reasonable notice and communication with the students will be given with explanation and the opportunity to comment on changes.

4. ASSESSMENT OF LEARNING *including dates*	Weight
Assignments	10%
Mid-term test	30%
Project	10%
Final examination (tests cumulative knowledge)	50%
TOTAL	100%

Percentage grades will be converted to letter grades and grade points per the University calendar.

5. LEARNING OUTCOMES

1. Perform statistical analyses to discover patterns in data sets
2. Apply theoretical algorithm descriptions to simple classification problems
3. Compare the behavior of the most popular data mining algorithms
4. Design multiple solutions to a data mining problem using different approaches

6. POLICIES

Anti-Discrimination

The Faculty of Engineering is concerned with ensuring an environment that is free of all discrimination. If there is a problem, individuals are reminded that they should contact the Department Chair, the Sexual Harassment Officer or the Human Rights Consultant, as soon as possible.

<http://www.mcmaster.ca/policy/General/HR/Anti-Discrimination%20policy.pdf>

Academic Integrity

You are required to exhibit honestly and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity.

Academic dishonesty is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage. This behaviour can result in serious consequences e.g. the grade of zero on an assignment, loss of credit with a notation on the transcript (notation reads: "Grade of F assigned for academic dishonesty"), and/or suspension or expulsion from the university.

It is your responsibility to understand what constitutes academic dishonesty. For information on the various kinds of academic dishonesty please refer to the Academic Integrity Policy, located at: <http://www.mcmaster.ca/policy/Students-AcademicStudies/AcademicIntegrity.pdf>.

The following illustrates only three forms of academic dishonesty:

1. Plagiarism. E.g. the submission of work that is not own or for which other credit has been obtained
2. Improper collaboration in group work
3. Copying or using unauthorized aids in tests and examinations.

Requests for Relief for Missed Academic Term Work (Assignments, Mid-Terms, etc.)

The McMaster Student Absence Form is a self-reporting tool for **Undergraduate Students** to report absences that last up to 3 days and provides the ability to request accommodation for any missed academic work. Please note, this tool cannot be used during any final examination period.

You may submit a maximum of 1 Academic Work Missed requests per term. It is YOUR responsibility to follow up with your Instructor immediately regarding the nature of the accommodation.

If you are absent more than 3 days or exceed 1 request per term you **MUST** visit your Associate Dean's Office (Faculty Office). You may be required to provide supporting documentation.

This form should be filled out immediately when you are about to return to class after your absence.

<http://www.mcmaster.ca/msaf/>

E-Learning Policy

Consistent with the Bachelor of Technology's policy to utilize e-learning as a complement to traditional classroom instruction, students are expected to obtain appropriate passwords and accounts to access Avenue To Learn for this course. Materials will be posted by class for student download. It is expected that

students will avail themselves of these materials prior to class. Students should be aware that, when they access the electronic components of this course, private information such as first and last names, user names for the McMaster e-mail account, and program affiliation may become apparent to all other students in the course. The available information is dependent on the technology used. Continuation in this course will be deemed consent to this disclosure. If you have any questions or concerns about this disclosure please discuss this with the course instructor. Avenue can be accessed via <http://avenue.mcmaster.ca>.

Communications

It is the student's responsibility to:

- Maintain current contact information with the University, including address, phone numbers, and emergency contact information.
- Use the University provided e-mail address or maintain a valid forwarding e-mail address.
- Regularly check the official University communications channels. Official University communications are considered received if sent by postal mail, by fax, or by e-mail to the student's designated primary e-mail account via their @mcmaster.ca alias.
- Accept that forwarded e-mails may be lost and that e-mail is considered received if sent via the student's @mcmaster.ca alias.
- Check the McMaster/Avenue email and course websites on a regular basis during the term.

Turnitin (Optional)

This course will be using a web-based service (Turnitin.com) to reveal plagiarism. Students submit their assignment/work electronically to Turnitin.com where it is checked against the internet, published works and Turnitin's database for similar or identical work. If Turnitin finds similar or identical work that has not been properly cited, a report is sent to the instructor showing the student's work and the original source. The instructor reviews what Turnitin has found and then determines if he/she thinks there is a problem with the work. Students who do not wish to submit their work to Turnitin.com must still submit a copy to the instructor. No penalty will be assigned to a student who does not submit work to Turnitin.com. All submitted work is subject to normal verification that standards of academic integrity have been upheld (e.g., on-line search, etc.). To see the Turnitin.com Policy, please go to

<http://www.mcmaster.ca/academicintegrity/turnitin/students/>

Protection of Privacy Act (FIPPA)

The Freedom of Information and Protection of Privacy Act (FIPPA) applies to universities. Instructors should take care to protect student names, student numbers, grades and all other personal information at all times. For example, the submission and return of assignments and posting of grades must be done in a manner that ensures confidentiality.

<http://www.mcmaster.ca/univsec/fippa/fippa.cfm>

Academic Accommodation of Students with Disabilities Policy

Students who require academic accommodation must contact Student Accessibility Services (SAS) to make arrangements with a Program Coordinator. Academic accommodations must be arranged for each term of study. Student Accessibility Services can be contacted by phone 905-525-9140 ext. 28652 or e-mail sas@mcmaster.ca. For further information consult McMaster's policy for Academic Accommodation of Students with Disabilities

<http://www.mcmaster.ca/policy/Students-AcademicStudies/AcademicAccommodation-StudentsWithDisabilities.pdf>

Students must forward a copy of the SAS accommodation to the instructor of each course and to the Program Administrator of the B.Tech. Program immediately upon receipt. If a student with a disability chooses NOT to take advantage of a SAS accommodation and chooses to sit for a regular exam, a petition for relief may not be filed after the examination is complete. <http://sas.mcmaster.ca>

Student Code of Conduct

The Student Code of Conduct (SCC) exists to promote the safety and security of all the students in the McMaster community and to encourage respect for others, their property and the laws of the land. McMaster University is a community which values mutual respect for the rights, responsibilities, dignity and well-being of others. The purpose of the Student Code of Conduct is to outline accepted standards of behavior that are harmonious with the goals and the well-being of the University community, and to define the procedures to be followed when students fail to meet the accepted standards of behavior. All students have the responsibility to familiarize themselves with the University regulations and the conduct expected of them while studying at McMaster University.

<http://judicialaffairs.mcmaster.ca/pdf/SCC.pdf>