Course Outline

1. COURSE INFORMATION

Session Offered	Winter 2021
Course Name	Data Mining
Course Code	SFWR TECH 4DM3
Program Name	Computer and Information Technology
Calendar Description	Classification, association, prediction and clustering of data. Decision Trees. Bayesian probability. Supervised and unsupervised learning.
Instructor	Dr. Jeff Fortuna — E-Mail: fortunjj@mcmaster.ca / Office Hours: zoom by request

2. COURSE SPECIFICS

Course Description	The course gives students the basic theory to understand the operation of the most common algorithms employed in data mining. The student's theoretical understanding will be tested on the midterm and final exams through the use of small problems that can be worked out by hand. Each algorithm that is discussed is exemplified by Python code in the course text. A course project allows the student to compare the results of 3 of the classification algorithms applied to a problem of their choosing.

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	Peter Harrington Published by Manning
ISBN: 9780387848570	Supplemental: The Elements of Statistical Learning	Trevor Hastie, Robert Tibshirani, Jerome Friedman Published by Springer
ISBN: 9781420089646	Supplemental: The Top Ten Algorithms in Data Mining	Edited by Xindong Wu, Vipin Kumar Published by CRC Press
ISBN: 9780123814791	Supplemental: Data Mining Concepts and Techniques	Jiawei Han, Micheline Kamber, Jian Pei Published by Morgan Kauffman

Other Supplies	Source

Prerequisite(s)	ENG TECH 3ST3
Corequisite(s)	none
Antirequisite(s)	none
Course Specific Policies	

	Brief assignments will be provided following each lecture covering the theory component covered in the lecture. These **assignments** will not be graded, but **must** be completed in order to ensure success on the midterm and final exams.
Departmental Policies	Students must maintain a GPA of 3.5 on a 12 point scale to continue in the program.
	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.
	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.
	Announcements made in class or placed on Avenue are considered to have been communicated to all students including those not in class.

3. SUB TOPIC(S)		
Week 1	Data descriptions, visualization of data, measurements of data similarity and dissimilarity	
Week 2	K Nearest Neighbors Classification	
Week 3	Bayesian Classifiers - Review of Bayes' Theorem - Naïve Bayes' Classifier	
Week 4	Decision Trees	
Week 5	Boosting and AdaBoost - Combining Multiple Similar Classifiers - Weak Learners - Decision Stumps - Adaptive Boosting	
Week 6	Logistic Regression - Optimization of regression coefficients - Logistic Regression Classification	
Week 8	Midterm Exam	
Week 9	Support Vector Machines - Maximum Margin Classification - Kernels	
Week 11	Classifier Performance	
Week 12	Unsupervised Learning	

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 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	Weight
Assignments	10 %
Project – Solve a classification problem using 3 of the approaches covered in the course using Python.	20 %
Midterm Exam	30 %
Final Exam	40 %

	TOTAL	100%

Course results determined on a percentage scale will be converted to an official letter grade, as indicated in the Undergraduate Calendar. The results of all courses attempted will appear on your transcript as letter grades.

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. COURSE OUTLINE – APPROVED ADVISORY STATEMENTS

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/Discrimination_Harassment_Sexual_Harassment-Prevention&Response.pdf

ACADEMIC INTEGRITY

You are expected to exhibit honesty and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity. It is your responsibility to understand what constitutes academic dishonesty.

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. For information on the various types of academic dishonesty please refer to the Academic Integrity Policy, located at https://secretariat.mcmaster.ca/university-policies-procedures-guidelines/

The following illustrates only three forms of academic dishonesty: The following illustrates only three forms of academic dishonesty:
• plagiarism, e.g. the submission of work that is not one's own or for which other credit has been obtained.
• improper collaboration in group work.
• copying or using unauthorized aids in tests and examinations.

AUTHENTICITY / PLAGIARISM DETECTION

Some courses may use a web-based service (Turnitin.com) to reveal authenticity and ownership of student submitted work. For courses using such software, students will be expected to submit their work electronically either directly to Turnitin.com or via an online learning platform (e.g. A2L, etc.) using plagiarism detection (a service supported by Turnitin.com) so it can be checked for academic dishonesty.

Students who do not wish their work to be submitted through the plagiarism detection software must inform the Instructor before the assignment is due. No penalty will be assigned to a student who does not submit work to the plagiarism detection software. All submitted work is subject to normal verification that standards of academic integrity have been upheld (e.g., on-line search, other software, etc.).   For   more   details   about   McMaster's   use   of   Turnitin.com   please   go   to www.mcmaster.ca/academicintegrity.

COURSES WITH AN ON-LINE ELEMENT

Some courses may use on-line elements (e.g. e-mail, Avenue to Learn (A2L), LearnLink, web pages, capa, Moodle, ThinkingCap, etc.).   Students should be aware that, when they access the electronic components of a course using these elements, private information such as first and last names, user names for the McMaster e-mail accounts, and program affiliation may become apparent to all other students in the same course. The available information is dependent on the technology used.

Continuation in a course that uses on-line elements will be deemed consent to this disclosure. If you have any questions or concerns about such disclosure please discuss this with the course instructor.

ONLINE PROCTORING

Some courses may use online proctoring software for tests and exams. This software may require students to turn on their video camera, present identification, monitor and record their computer activities, and/or lock/restrict their browser or other applications/software during tests or exams. This software may be required to be installed before the test/exam begins.

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.

CONDUCT EXPECTATIONS

As a McMaster student, you have the right to experience, and the responsibility to demonstrate, respectful and dignified interactions within all of our living, learning and working communities. These expectations are described in the Code of Student Rights & Responsibilities (the "Code"). All students share the responsibility of maintaining a positive environment for the academic and personal growth of all McMaster community members, whether in person or online.

It is essential that students be mindful of their interactions online, as the Code remains in effect in virtual learning environments. The Code applies to any interactions that adversely affect, disrupt, or interfere with reasonable participation in University activities. Student disruptions or behaviours that interfere with university functions on online platforms (e.g. use of Avenue 2 Learn, WebEx or Zoom for delivery), will be taken very seriously and will be investigated. Outcomes may include restriction or removal of the involved students' access to these platforms.

ACADEMIC ACCOMMODATION OF STUDENTS WITH DISABILITIES

Students with disabilities who require academic accommodation must contact Student Accessibility Services (SAS) at 905-525-9140 ext. 28652 or sas@mcmaster.ca to make arrangements with a Program Coordinator. For further information, consult McMaster University's Academic Accommodation of Students with Disabilities policy.

REQUESTS FOR RELIEF FOR MISSED ACADEMIC TERM WORK

McMaster Student Absence Form (MSAF): In the event of an absence for medical or other reasons, students should review and follow the Academic Regulation in the Undergraduate Calendar "Requests for Relief for Missed Academic Term Work".

ACADEMIC ACCOMMODATION FOR RELIGIOUS, INDIGENOUS OR SPIRITUAL OBSERVANCES (RISO)

Students requiring academic accommodation based on religious, indigenous or spiritual observances should follow the procedures set out in the RISO policy. Students should submit their request to their Faculty Office normally within 10 working days of the beginning of term in which they anticipate a need for accommodation or to the Registrar's Office prior to their examinations. Students should also contact their instructors as soon as possible to make alternative arrangements for classes, assignments, and tests. http://www.mcmaster.ca/policy/Students-AcademicStudies/Studentcode.pdf

COPYRIGHT AND RECORDING

Students are advised that lectures, demonstrations, performances, and any other course material provided by an instructor include copyright protected works. The Copyright Act and copyright law protect every original literary, dramatic, musical and artistic work, including lectures by University instructors

The recording of lectures, tutorials, or other methods of instruction may occur during a course. Recording may be done by either the instructor for the purpose of authorized distribution, or by a student for the purpose of personal study. Students should be aware that their voice and/or image may be recorded by others during the class. Please speak with the instructor if this is a concern for you.

EXTREME CIRCUMSTANCES

The University reserves the right to change the dates and deadlines for any or all courses in extreme circumstances (e.g., severe weather, labour disruptions, etc.). Changes will be communicated through regular McMaster communication channels, such as McMaster Daily News, A2L and/or McMaster email.