

TIMG 5303A

Machine Learning for Technology Entrepreneurship Problem-Solving

[0.5 credit]

Winter 2024

Institute of Technology Entrepreneurship and Commercialization

TIME AND PLACE:

Jan 8 to Apr 10, 2024, Wednesday, 2:35-5:25 pm Eastern time (Ottawa time)

In person: Nicol Building 3020

Online (synchronous portion): Zoom link available on TIMG 5303 course site within Brightspace.

DELIVERY MODE:

HyFlex: students can attend in-person, online or a combination of in-person and online. All course participants must use CU credentials and CU email accounts for communications and login. Students should engage in class regardless of in-person or online attendance.

INSTRUCTOR:

Mika Westerlund, mika.westerlund@carleton.ca

Office: Nicol Building 5029

Office hours: Email is the preferred mode of communication because there is a record of content exchanged. The instructor will be available after class and by appointment depending on availability. Email the instructor to schedule an on-campus or virtual appointment.

COURSE MATERIALS:

Access to online course sessions, course materials, list of readings and recorded videos will be provided through the CU Brightspace system: <https://carleton.ca/brightspace/>.

To access Brightspace and course materials you should use your CU credentials and select the TIMG 5303A ML for Tech Entrepr Probl Solv.

Readings (articles) can be accessed and downloaded through the CU Library Online service using student's CU access privilege. A required textbook is not assigned for this course. Please note there can be changes to the material during the course. An up-to-date list of course readings will be available through the course site within Brightspace.

Lecture slides will be available within the course site on the day of each lecture.

COURSE DESCRIPTION:

Calendar description: Application of machine learning tools to co-create solutions to entrepreneurial problems, with an emphasis on unstructured text analytics. Topics include machine learning tools, application of topic modeling and text analytics, generation of practical competitive insights for managers, and analysis of publicly available sources including websites.

Please note that some topics covered on the course may be different from the calendar description.

OBJECTIVES AND LEARNING OUTCOMES:

To help students acquire the collaborative skills and business analytics expertise needed to co-create solutions for entrepreneurial problems and generate valuable business insights for companies, organizations supporting local companies and the TIM program ecosystem.

The specific objectives are to acquire proficiency in:

- Applying topic modeling, sentiment analysis and other text mining machine learning techniques to generate practical insights from online textual data such as social media comments, customer- and product reviews, and websites.
- Applying text mining machine learning tools and generative AI techniques to generate practical insights from collections of text documents such as journal articles, news articles, open data, and company documentation.

Upon course completion, participants will be able to:

- Apply machine learning tools to solve relevant entrepreneurial and innovation problems.
- Prepare and present machine learning outcomes in a way that advances our knowledge.
- Synthesize steps from data collection, analysis, and interpretation into a coherent report.

CLASS SESSIONS:

Class sessions will include a combination of online lectures and interactive sessions that focus on using various machine learning tools. For each class, read the assigned readings and be prepared to:

- Use machine learning tools to solve real problems in organizations.
- Apply and present key learnings from the assigned course material.
- Generate actionable insights and propose solutions to real problems.

Video recordings of class sessions will be available on the day after lecture within Brightspace.

STUDENT GROUPS

The class will be split into 14 groups of 7 students each. Each student will work individually and together with the group members to perform tasks in class and outside the class hours, participate in informal group meetings, and contribute to delivering the assignments. Constructive collaboration among group members is part of the group assignments. Leaving and changing groups for any reason will not be tolerated, because the course schedule and grading mechanism does not support such alterations in group structure. Once formed, each group should establish a project management structure that will help collaboration between group members and maximize the value of the deliverables. Group members should meet on a weekly basis.

GROUP WORK AND FREE LOADERS

There will be zero tolerance for free loaders. A freeloader is someone who exploits the efforts of team members without making a proportional contribution in return. Group work is an important component of this course. Group conflicts are to be dealt with by the group in a way that is fair, respectful, and fast. In case a non-contributing student is excluded from a group, that student may not be able to complete the course or will need to deliver the assignment individually if the course schedule allows it. Any exclusions from a group must be reported to the instructor immediately.

ADMINISTRATIVE DETAILS:

Missing classes

If a student misses a class, it is the student's responsibility to find out what was covered, what assignments were made and what handouts were distributed. It is strongly recommended to watch the video recording of each missed class.

Plagiarism

Plagiarism is a serious instructional offense that will not be tolerated. A case of plagiarism in assignments or the take-home final exam will be referred to the Chair of the Department and the Carleton University Ethics Committee. The instructor will not deal with the matter directly. See: <http://carleton.ca/registrar/academic-integrity/>

Software & IT requirements

This course requires that at least one person in each group installs the open-source software -based Orange machine learning and data mining toolkit, available at <https://orangedatamining.com/> with its text widget (Option → Add-ons → Text). However, it is recommended for all students to have it. The Documentation section on the Orange’s website provides some help and tutorials. Installing and dealing with any compatibility issues or problems regarding operationality is each student’s own responsibility. The instructor does not have resources to help students with their IT problems.

Also, spreadsheet software (preferably Excel) is necessary to handle and edit data sets. Computer programming skills are not required in this course but may be useful, and any student interested in becoming proficient in data analytics, and especially those in the MABA option, are highly recommended to develop and advance their programming skills. The course website within Brightspace provides a link to a book on text analytics using Python machine learning solutions.

Course grading

Marks will be assigned as follows:

Course requirement	Mark (max.)
Assignment 1 (group effort) – Literature review	15
Assignment 2 (group effort) – Data collection & analysis	40
Assignment 3 (group effort) – Validation & interpretation	15
Assignment 4 (group effort) – Finalized report	10
Take-home final exam (individual effort)	20
Total (max.)	100

Assignments 1, 2 and 3 have two parts: 1) a presentation in class using slides and 2) a written document. The maximum time allocated for presentation is A1) 5 minutes + 5 minutes for feedback, and A2 & A3) 10 minutes + 10 minutes for feedback. The maximum length for A1-A3 documents is 3,000 words each. Assignment 4 consists of a written document only (max. 8,000 words). All assignments and the final exam should be submitted through Brightspace and are mandatory to pass the course. See below for details and deadlines.

The take-home final exam will be graded on a competitive basis. The best solution(s) in class will receive the maximum available points. All other solutions will be graded, firstly, in comparison with the top solution(s) and, secondly, in comparison with other students’ solutions. Assessment of the take-home final exam will emphasize originality, novelty, relevance, and depth of the solution, as well as quality and compliance with the provided guidelines. The maximum points any one student can receive is 20 points. See below for details.

In accordance with the Carleton University Graduate Grading System, the letter grades assigned in this course will have the following percentage equivalents:

A+ = 90-100 | A = 85-89 | A- = 80-84 | B+ = 77-79 | B = 73-76 | B- = 70-72 | F = Fail

The minimum passing grade for the course is B-.

Grades entered by Registrar: WDN = Withdrawn from the course | DEF = Deferred

ASSIGNMENTS:

The students must successfully complete 4 mandatory assignments (group work only). When put together, these assignments result in a coherent report. The instructor will mark all assignments and exams himself. Late assignments are not acceptable without permission from the instructor.

Assignments 1, 2 and 3 involve two parts: 1) a presentation in class using slides and 2) a written document. Written documents for A1-A3 should be max. 3,000 words each. Assignment 4 consists of a written document only (max. 8,000 words). Slides and documents should be submitted in PDF format through the course site within Brightspace by the deadline (details available in Brightspace). At least one person from each group will need to present each assignment in class (in the classroom or online). Any feedback received after the presentation in class should be taken into consideration and applied to improve the written document.

Further details and instructions regarding the assignments will be delivered and discussed in class. This information will also be available on the course site within Brightspace.

Deadlines for assignments: 1) Slides are due on the day of presentation by noon (see Class schedule); 2) Reports are due four days after the presentation (see Class schedule and Summary table which also lists the key points in grading of assignments. Slides and reports in PDF format.

Summary table for assignments:

Max. points	Assignment	Deadline (for reports)
15 p.	A1: Literature review (article corpus 5 p., topic identification 10 p.)	Sun Feb 11 (all groups) <i>Extended deadline</i>
40 p.	A2: Data collection & analysis (collection 10 p., (pre)processing 10 p., analysis 20p.)	Wed Mar 20 (all groups) <i>Extended deadline</i>
15 p.	A3: Validation & interpretation (validation 5 p., interpretation 10 p.)	Sun Apr 7 (all groups) <i>Extended deadline</i>
10 p.	A4: Final report (revision 5 p., story 5 p.)	Sun Apr 14 (all groups) <i>Extended deadline</i>

TAKE-HOME FINAL EXAM:

Take-home final exam is an individual effort, and group effort is considered an act of plagiarism. The examination questions along with detailed instructions and guidelines will be introduced and discussed in class and distributed through the course site within Brightspace during Session 13.

Your examination report with the solutions is **due on Thursday April 25, 2024**, and it must be submitted in PDF format through the course site within Brightspace by the deadline. Late submission is not acceptable without permission from the instructor.

There are 2 questions in the final take-home exam. The maximum length of the exam report is 6 pages (3 pages per question). Use 1.5 line spacing. Font type: Times New Roman or Arial. Font size: 10 pts or 12 pts. Use margins of one inch (minimum) all around. PDF. The take-home final exam will be marked based on a comparison with other students' exam reports. See above for criteria.

Summary table for the take-home final exam:

Max. points	Take-home final exam	Deadline
20 p.	2 questions, max. a'10 pts, totaling max. 20 points	Thu April 25

CHATGPT:

Students are encouraged to use ChatGPT (<https://chat.openai.com/chat>) or other GPT tools to:

- Improve the language, flow and content of your assignments and reports.
- Align parts of an assignment to make a whole.
- Acquire and synthesize information.
- Complement machine learning analyses.
- Create notes and learning diaries of the course content.
- Eliminate errors in grammar, spelling, and capitalization as well as citations and references.
- Provide a skeletal structure for assignments.
- Generate ideas, make recommendations, and extract what is most important.

Students are discouraged to use ChatGPT to:

- Provide unedited AI-generated output as a solution to assignments and exam questions.
- Produce content for assignments/exams without double-checking citations and references.

CLASS SCHEDULE:

This calendar is tentative; it may be modified as the course proceeds. Please consult the course site within Brightspace for up-to-date information on the course schedule and readings.

Session	Date	Readings	Details / Other
<i>Introduction to text analytics</i>			
#01	Jan 10	• Ch. 1-3 / Anandarajan et. (2019)	• Course outline
<i>Foundations of text processing</i>			
#02	Jan 17	• Ch. 4 / Anandarajan et. (2019)	• Data analysis in class • AirBnB literature review data available on the course site
<i>Topic modeling: LSI & Literature reviews</i>			
#03	Jan 24	• Ch. 5-7 / Anandarajan et. (2019)	• Data analysis in class • AirBnB literature review data available on the course site
<i>Presentation of A1 – Literature review</i>			
#04	Jan 31	• Presentations (all groups 1-14)	Deadlines for all groups: • A1 slides: Jan 31 by Noon • A1 report: Sun Feb 11
<i>Sentiment analysis & Product reviews</i>			
#05	Feb 7	Ch. 10 / Anandarajan et. (2019)	• Data analysis in class • Googleplay data available on the course site
<i>Topic modeling: LDA & User comments</i>			
#06	Feb 14	• Ch. 8 / Anandarajan et. (2019)	• Data analysis in class • Googleplay data available on the course site
<i>*Winter break – No classes*</i>			
	<i>Feb 21</i>	• <i>No classes</i>	• <i>No classes</i>
<i>Information visualization & interpretation</i>			
#07	Feb 28	Ch. 12 / Anandarajan et. (2019)	• Googleplay data available on the course site
<i>Presentation of A2 – Day 1</i>			
#08	Mar 6	• Presentations (groups 1-7)	Deadlines for groups 1-7: • A2 slides: Mar 6 by noon • A2 report: Mar 20
<i>Presentation of A2 – Day 2</i>			
#09	Mar 13	• Presentations (groups 8-14)	Deadlines for groups 8-14: • A2 slides: Mar 13 by noon • A2 report: Mar 20
<i>Validation & storytelling</i>			
#10	Mar 20	Ch. 11 / Anandarajan et. (2019)	• Googleplay data available on

			the course site
<i>Presentation of A3 – Day 1</i>			
#11	Mar 27	<ul style="list-style-type: none"> • Presentations (groups 1-7) 	Deadlines for groups 1-7: <ul style="list-style-type: none"> • A3 slides: Mar 27 by noon • A3 report: Apr 7
<i>Presentation of A3 – Day 2</i>			
#12	Apr 3	<ul style="list-style-type: none"> • Presentations (groups 8-14) • Delivery of the take-home final exam 	Deadlines for groups 8-14: <ul style="list-style-type: none"> • A3 slides: Apr 3 by noon • A3 report: Apr 7 Other deadlines: <ul style="list-style-type: none"> • A4 report: Apr 14 (all groups) • Take-home final exam deadline: April 25

COURSE SHARING WEBSITES:

Materials created for this course (including presentations, slides, case studies, assignments, and exams) remain the intellectual property of the author(s). They are intended for personal use and may not be reproduced or redistributed without prior written consent of the author(s).

REQUESTS FOR ACADEMIC ACCOMMODATION:

You may need special arrangements to meet your academic obligations during the term. For an accommodation request, the processes are as follows:

Academic accommodations for students with disabilities

Students with disabilities requiring academic accommodations in this course are encouraged to contact a coordinator at the Paul Menton Centre for Students with Disabilities (PMC) at 613-520-6608 or pmc@carleton.ca for a formal evaluation or a Letter of Accommodation at the beginning of the term. You must also contact the PMC (carleton.ca/pmc/) no later than two weeks before the assignment or exam requiring accommodation. After requesting accommodation from PMC, meet with your instructor to ensure accommodation arrangements are made.

Pregnancy obligation

Please contact your instructor with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, visit the Equity Services website: carleton.ca/equity/wp-content/uploads/Student-Guide-to-Academic-Accommodation.pdf

Religious obligation

Please contact your instructor with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, visit the Equity Services website: carleton.ca/equity/wp-content/uploads/Student-Guide-to-Academic-Accommodation.pdf

Survivors of Sexual Violence

As a community, Carleton University is committed to maintaining a positive learning, working and living environment where sexual violence will not be tolerated, and its survivors are supported through academic accommodations as per Carleton's Sexual Violence Policy. For more information about the services available at the university and to obtain information about sexual violence and/or support, visit: carleton.ca/sexual-violence-support

Other accommodation

For information on other academic accommodation please contact the departmental administrator timprogram@CUNET.Carleton.ca or visit: <https://students.carleton.ca/course-outline/>

READINGS:

The following is a preliminary list of course articles and may change as the course progresses. An up-to-date list of readings is available on the course site within Brightspace. To access the required journal articles in electronic format, click the DOI or PERMALINK hyperlinks for each article listed on the course site or go to CU Library's website (<https://library.carleton.ca/>). Log in using your CU credentials and use the library's search engine to access and download the material(s).

Course book:

- Anandarajan, M., Hill, C., & Nolan, T. (2019): Practical Text Analytics. Maximizing the value of text data. Springer Nature: Switzerland.

The course is based on this book; it is a recommended reading for all students, as it provides basic understanding of text analytics. See the course website "Readings" section for more information.

Complementary readings:

- Albrecht, J., Ramachandran, S., & Winkler, C. (2020): Blueprints for Text Analytics Using Python. Machine Learning-Based Solutions for Common Real World (NLP) Applications. O'Reilly: USA.

This additional book is recommended especially for MABA students; it provides more advanced options to perform text analytics. See the course website "Readings" section for more information.

Readings for Session 1 – Introduction to text analytics:

- Chapters 1-3 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 2 – Foundation of text processing:

- Chapter 4 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 3 – Latent Semantic Indexing (LSI) & Literature reviews:

- Chapters 5-7 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 4 – Presentation of A1 – Literature review:

- N/A

Readings for Session 5 – Sentiment analysis & Product reviews:

- Chapter 10 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 6 – Probabilistic topic models (LDA) & User comments:

- Chapter 8 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 7 – Information visualization & interpretation:

- Chapter 12 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 8 – Presentation of A2 – Day 1:

- N/A

Readings for Session 9 – Presentation of A2 – Day 2:

- N/A

Readings for Session 10 – Visualization & storytelling:

- Chapter 11 of Anandarajan et al. (2019): Practical Text Analytics

Readings for Session 11 – Presentation of A3 – Day 1:

- N/A

Readings for Session 12 – Presentation of A2 – Day 2:

- N/A