COURSE INFORMATION

Course title:	Business Applications of Machine Learning		
Course code:	BAIT 509	Credits:	1.5
Session, term, period:	2023W2, Period 3	Class location:	HA 435
Section(s):	BA1	Class times:	BA1 Mon & Wed 10am-12pm BA2 Mon & Wed 2pm-4pm
Course duration:	Jan 8 – Feb 17	Pre-requisites:	n/a
Division:	n/a	Co-requisites:	n/a
Program:	MBAN		

INSTRUCTOR INFORMATION

Instructor:	Quan Nguyen		
Phone:	n/a	Office location:	Zoom
Email:	quan.nguyen@ubc.ca	Office hours:	Thursdays 12:00-13:00

Teaching Assistant	**Email**
Armin Saadat Boroujeni	armin2saadat@gmail.com
Ailar Mahdizadeh	ailar.mahdizadeh@ubc.ca
Mohammad Mahdi Asmae	m.mahdi.asmae@gmail.com
Harsh Sharma	harsh1792@gmail.com
Julian Camilo Becerra Leon	jucble@student.ubc.ca
Meltem Omur	momur@cmmt.ubc.ca

COURSE DESCRIPTION

Introduction to machine learning concepts, such as model training, model testing, generalization error and overfitting. Exposure to a variety of machine learning techniques, with deeper exploration of a few chosen techniques. Forming good scientific questions to address business objectives with machine learning. Python will be the primary programming language used.

COURSE FORMAT

Class time will be used for a combination of lectures, discussion, demonstrations and exercises that students will work through individually or in groups.

LEARNING OBJECTIVES

By the end of this course, students will be able to:

1. Describe fundamental machine learning concepts such as: supervised and unsupervised learning, regression and classification, overfitting, training/validation/testing error, parameters and hyperparameters, and the golden rule.
2. Broadly explain how common machine learning algorithms work, including: naïve Bayes, k-nearest neighbors, decision trees, support vector machines, and logistic regression.
3. Identify when and why to apply data pre-processing techniques such as scaling and one-hot encoding.
4. Use Python and the scikit-learn package to develop an end-to-end supervised machine learning pipeline.

5. Apply and interpret machine learning methods to carry out supervised learning projects and to answer business objectives.

ASSESSMENTS

Summary

Component	Weight
Assignments	60%
Group project	30%
Quiz	10%
Total	100%

Details of Assessments

Assignments

During the term, there will be three individual assignments, each worth 20%. Each assignment will focus on a combination of theory and application and will be distributed as a Jupyter Notebook. Each assignment will require the analysis of a data set. The assignments will be distributed on Canvas with a test file which SHOULD NOT be tampered with. This test file will check certain questions and give feedback on if the written code is on the right track. This however does not guarantee full marks. Students will need to submit the assignment in the form of a report and marks will be based on the depth of the analysis and the presentation.

Group Project

The group assignment will involve the analysis of a more complex data set. The format and submission requirements will be similar to the individual assignment, except that instead of simply answering the specified questions, you will be required to perform a thorough analysis of the case and submit a report summarizing your main findings. Groups will consist of two or three students and will be allocated randomly. All group members will receive the same mark: it is each student's responsibility to ensure that all group members contribute equally to the assignment. In case of any group related issues, please discuss with the instructor.

Quiz

There will be one quiz in the middle of the course that should be completed individually. You will have a 24 hours window to complete it but only 45mins to write it. It will be a mix of multiple choice, true or false questions and long answer questions and it will be hosted on Canvas.

LEARNING MATERIALS

Reading Materials:

The following are recommended (i.e., not mandatory) reading material to supplement the course.

- "Python Data Science Handbook" – Jake VanderPlas. Freely available at: https://jakevdp.github.io/PythonDataScienceHandbook/
- "Data Mining: Practical Machine Learning Tools and Techniques" – Ian Witten, Eibe Frank, Mark Hall, Christopher Pal. Freely available at: https://www.cs.waikato.ac.nz/~ml/weka/book.html
- Scikit-learn Python package documentation. Freely available at: http://scikit-learn.org/stable/documentation.html

Technology Requirements:

- Laptop
- Python Version 3.6 or above (it is recommended to install Python using the Anaconda distribution https://www.anaconda.com/distribution/#download-section)

COURSE-SPECIFIC POLICIES AND RESOURCES

Missed or late assignments, and regrading of assessments

Late submissions will not be accepted and will receive a grade of zero.

Academic Concessions

If extenuating circumstances arise, please contact the RHL Graduate School program office as early as reasonably possible, and submit an Academic Concession Request & Declaration Form https://webforms.sauder.ubc.ca/academic-concession-rhlee. If an academic concession is granted during the course, the student will be provided options by RHL, or by the instructor in consultation with RHL, per UBC's policy on Academic Concession.

Other Course Policies and Resources

[Include the following policies and resources as appropriate.]

Code Plagiarism

Code plagiarism falls under the UBC policy for Academic Misconduct. Students must correctly cite any code that has been authored by someone else or by the student themselves for other assignments. Cases of "reuse" may include, but are not limited to:

- the reproduction (copying and pasting) of code with none or minimal reformatting (e.g., changing the name of the variables)
- the translation of an algorithm or a script from a language to another
- the generation of code by automatic code-generations software

An "adequate acknowledgement" requires a detailed identification of the (parts of the) code reused and a full citation of the original source code that has been reused.
Students are responsible for ensuring that any work submitted does not constitute plagiarism. Students who are in any doubt as to what constitutes plagiarism should consult their instructor before handing in any assignments.

POLICIES APPLICABLE TO COURSES IN THE ROBERT H. LEE GRADUATE SCHOOL

Attendance

Excepting extenuating circumstances, students are expected to attend 100% of their scheduled class hours. Absent students limit their own academic potential, and that of their classmates, and cause unnecessary disruption to the learning environment. Students missing more than 20% of the total scheduled class hours for a course (including classes held during the add/drop period) without having received an academic concession will be withdrawn from that course. Withdrawals, depending on timing, could result in a "W" or an "F" standing on the transcript.

Punctuality

Students are expected to arrive for classes and activities on time and fully prepared to engage. Late arrivals may be refused entry at the discretion of the instructor or activity lead. Students arriving later than halfway through a scheduled class will be treated as absent for that class.

Electronic Devices

During online lectures, students are not permitted to use any electronic devices other than the primary one used for attending the online lecture (e.g. laptop or desktop). Only Zoom should be open during the online lecture unless an instructor advises the use of another program/website for an in-class activity. Feedback from students indicates that personal devices are the number one distraction from effective learning and participation in the online learning environment.

Citation Style

Please use the American Psychological Association (APA) reference style to cite your sources.

Details of the above policies and other RHL Policies are available at:
http://www.calendar.ubc.ca/vancouver/index.cfm?tree=12,199,506,1625

UNIVERSITY POLICIES AND RESOURCES

UBC provides resources to support student learning and to maintain healthy lifestyles but recognizes that sometimes crises arise and so there are additional resources to access including those for survivors of sexual violence. UBC values respect for the person and ideas of all members of the academic community. Harassment and discrimination are not tolerated nor is suppression of academic freedom. UBC provides appropriate accommodation for students with disabilities and for religious observances. UBC values academic honesty and students are expected to acknowledge the ideas generated by others and to uphold the highest academic standards in all of their actions. Details of the policies and how to access support are available on the UBC Senate website at https://senate.ubc.ca/policies-resources-support-student-success.

Respect for Equity, Diversity, and Inclusion

The UBC Sauder School of Business strives to promote an intellectual community that is enhanced by diversity along various dimensions including status as a First Nation, Metis, Inuit, or Indigenous person, race, ethnicity, gender identity, sexual orientation, religion, political beliefs, social class, and/or disability. It is critical that students from diverse backgrounds and perspectives be valued in and well-served by their courses. Furthermore, the diversity that students bring to the classroom should be viewed as a resource, benefit, and source of strength for your learning experience. It is expected that all students and members of our community conduct themselves with empathy and respect for others.

Academic Integrity

The academic enterprise is founded on honesty, civility, and integrity. As members of this enterprise, all students are expected to know, understand, and follow the codes of conduct regarding academic integrity. At the most basic level, this means submitting only original work done by you and acknowledging all sources of information or ideas and attributing them to others as required. This also means you should not cheat, copy, or mislead others about what is your work. Violations of academic integrity (i.e., misconduct) lead to the breakdown of the academic enterprise, and therefore serious consequences arise and harsh sanctions are imposed. For example, incidences of plagiarism or cheating may result in a mark of zero on the assignment or exam and more serious consequences may apply if the matter is referred to the President's Advisory Committee on Student Discipline. Careful records are kept in order to monitor and prevent recurrences.

Academic Freedom and Students Studying from Outside Canada

During this pandemic, the shift to online learning has greatly altered teaching and studying at UBC, including changes to health and safety considerations. Keep in mind that some UBC courses might cover topics that are censored or considered illegal by non-Canadian governments. This may include, but is not limited to, human rights, representative government, defamation, obscenity, gender or sexuality, and historical or current geopolitical controversies. If you are a student living abroad, you will be subject to the laws of your local jurisdiction, and your local authorities might limit your access to course material or take punitive action against you. UBC is strongly committed to academic freedom, but has no control over foreign authorities (please visit http://www.calendar.ubc.ca/vancouver/index.cfm?tree=3,33,86,0 for an articulation of the values of the University conveyed in the Senate Statement on Academic Freedom). Thus, we recognize that students will have legitimate reason to exercise caution in studying certain subjects. If you have concerns regarding your personal situation, consider postponing taking a course with manifest risks, until you are back on campus or reach out to your academic advisor to find substitute courses. For further information and support, please visit: http://academic.ubc.ca/support-resources/freedom-expression

Use of Artificial Intelligence

Generative AI (Including ChatGPT) Not Permitted

Any work submitted must be your own original work, written without outside assistance or collaboration. Any use of generative artificial intelligence (AI), including ChatGPT, is not permitted and constitutes academic misconduct. Any student suspected of submitting work that includes AI generated content may be asked for preliminary work or other materials to evidence the student's original and unaided authorship. The student may also be asked to separately explain or support their work. AI identification methods may also be employed by the instructor. After review, if it is determined by the instructor that submitted work likely contains AI generated content, the work may receive a zero and may be subject to further misconduct measures set out in the UBC Academic Calendar.

SUSTAINABLE DEVELOPMENT GOALS (SDGS)

At UBC Sauder, we are committed to responsible business practices that can have transformative impacts on society. One of the ways we are reinforcing our commitment to responsible business is by showcasing relevant content in our courses via the lens of the United Nations Sustainable Development Goals. In this course, we will touch on topics that relate to the following goals:

Sustainable Development Goal	Description of how and when the goal is covered in the course.
GOAL 3: Good Health and Well-being 	We will use a cancer tumour dataset as an example application of machine learning in healthcare

Goal 4: Quality Education 	***Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*** We will use an example of predicting academic retention for first-year undergraduate students.
Goal 10: Reduced Inequality 	We will discuss some fairness metrics in evaluation of machine learning models, that will have impact on the societal inequality.

COPYRIGHT

All materials of this course (course handouts, lecture slides, assessments, course readings, etc.) are the intellectual property of the instructor or licensed to be used in this course by the copyright owner. Redistribution of these materials by any means without permission of the copyright holder(s) constitutes a breach of copyright and may lead to academic discipline and could be subject to legal action. Any lecture recordings are for the sole use of the instructor and students enrolled in the class. In no case may the lecture recording or part of the recording be used by students for any other purpose, either personal or commercial. Further, audio or video recording of classes are not permitted without the prior consent of the instructor. Students may not share class Zoom links or invite others who are not registered to view sessions.

ACKNOWLEDGEMENT

UBC's Point Grey Campus is located on the traditional, ancestral, and unceded territory of the xʷməθkʷəy̓əm (Musqueam) people, who for millennia have passed on their culture, history, and traditions from one generation to the next on this site.

ONLINE TEACHING TOOL & REQUIREMENTS

This course will be taught using Zoom for synchronous classes and office hours.

For this course, you are required to use a Zoom account during synchronous classes and office hours. If you do not have a Zoom account, you can create one here: https://zoom.us/signup. Note: creating a Zoom account requires that you provide a first name, last name, and email address to Zoom. For privacy purposes, you may consent to using your existing email address and your real name. Alternatively, if you prefer, you may sign up using an alternative email address and an anonymized name that does not identify you (i.e. Jane Doe, jane.doe@email.com). If you have trouble creating an account, or accessing a Zoom session, please contact CLCHelp@sauder.ubc.ca. You will be required to provide the email address associated with your Zoom account in a Canvas quiz for identification purposes.

To help replicate the classroom experience, make sessions more dynamic and hold each person accountable, both students and instructors are required to have their cameras on during Zoom sessions. Students who require an accommodation with regard to the "camera on" requirement must contact their instructors in advance of the first class to discuss options. As professional graduate students,

students are expected to conduct themselves professionally by joining sessions on time, muting mics when not speaking, refraining from using any other technology when in-session, attending in business casual dress (at a minimum), and participating from a quiet environment. Content from synchronous sessions will be selectively recorded per instructor discretion and made available to students on Canvas for a maximum duration of the course length. This is done to allow students the opportunity to return to lecture content to solidify learnings.

COURSE SCHEDULE

(Subject to change with class consultation)

Class	Date	Topic	Readings or Activities	Assessments due
1	Jan 8, 2024 (Mon)	Introduction to machine learning and Decision Trees	In-class exercises	
2	Jan 10, 2024 (Wed)	Splitting and cross-validation	In-class exercises	
3	Jan 15, 2024 (Mon)	Baseline models, KNN and SVM	In-class exercises	
4	Jan 17, 2024 (Wed)	Support vector machines and Feature Preprocessing	In-class exercises	Assignment 1 due
5	Jan 22, 2024 (Mon)	Preprocessing Categorical Features and Column Transformer	In-class exercises	
6	Jan 24, 2024 (Wed)	Naive Bayes and Hyperparameter Optimization	In-class exercises	
7	Jan 29, 2024 (Mon)	Linear Models	In-class exercises	Assignment 2 due
8	Jan 31, 2024 (Wed)	Business Objectives/Statistical Questions and Feature Selection	In-class exercises	Quiz
9	Feb 5, 2024 (Mon)	Classification and Regression Metrics	In-class exercises	
10	Feb 7, 2024 (Wed)	Topics related to the group project	Work on group project	Assignment 3 due
	Feb 12-17, 2024 (Exam week)	Exam Week		Group Project – To be scheduled by RHL Office