COURSE INFORMATION

Course title:	Business Applications of Machine Learning		
Course code:	BAIT 509	Credits:	1.5
Session, term, period:	2020W2, Period 3	Class location:	HA 337
Section(s):	BA1	Class times:	Mon & Wed 16:00-18:00
Course duration:	Jan 6 to Feb 15, 2020	Pre-requisites:	n/a
Division:	n/a	Co-requisites:	n/a
Program:	MBAN		

INSTRUCTOR INFORMATION

Instructor:	Tomas Beuzen		
Phone:	n/a	Office location:	ESB 1045
Email:	tomas.beuzen@stat.ubc.ca	Office hours:	Thursday 13:00-14:00
Teaching assistant 1:	Andy Tai		
Email:	andy.tai@mail.utoronto.ca		
Teaching assistant 2:	Aarti Kashyap		
Email:	kaarti.sr@gmail.com		

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 and testing error;
2. Broadly explain how common machine learning algorithms work, including: naïve Bayes, k-nearest neighbors, decision trees, support vector machines, and ensemble methods;
3. Implement a machine learning pipeline (i.e., loading data, creating a machine learning model, testing the model) in Python; and,
4. 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%
Class participation	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. Each assignment will require the analysis of a data set. You will be provided with the data, and a set of questions. You will need to submit the assignment in the form of a report. Your marks will be based on the depth of the analysis and the presentation in the form of a report.

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.

Class Participation

We all bring experience and knowledge into the classroom, and all class participants should share this and benefit by it. Effective class participation includes:

- being prepared for class participation by reading the assigned materials;
- asking questions about concepts from lectures or readings that you agree or disagree with;
- sharing your experience or point of view with the class;
- building on points raised by others;
- clarifying issues; or,
- relating topics discussed to previous class discussions.

Direct student-student interaction is encouraged. Such interaction should be both positive and courteous even when your opinions differ. Class attendance is important and will be recorded. Failing to attend class will decrease your participation score.

LEARNING MATERIALS

Reading Materials:

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

- "An Introduction to Statistical Learning: with Applications in R" – Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. Freely available at: http://www-bcf.usc.edu/~gareth/ISL/
- "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

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

Devices such as laptops, tablets, and cell phones are not permitted to be used in class unless directed by the instructor for in-class activities. Students who do not follow the School's policy in this regard may be required to leave the room for the remainder of the class, so that they do not distract others. Research shows that students' use of laptops in class has negative implications for the learning environment, including reducing their own grades and the grades of those sitting around them.

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.

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.

COPYRIGHT

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.

COURSE SCHEDULE

(Subject to change with class consultation)

Class	Date	Topic	Readings or Activities	Assessments due
1	Jan 6, 2020 (Mon)	Introduction to machine learning and decision trees	In-class exercises	
2	Jan 8, 2020 (Wed)	Fundamentals of machine learning and error	In-class exercises	
3	Jan 13, 2020 (Mon)	Cross-validation, kNN and loess	In-class exercises	
4	Jan 15, 2020 (Wed)	Feature pre-processing	In-class exercises	
5	Jan 20, 2020 (Mon)	Naïve Bayes and logistic regression	In-class exercises	Assignment 1 due at 6pm Monday 20th Jan
6	Jan 22, 2020 (Wed)	Model and feature selection	In-class exercises	
7	Jan 27, 2020 (Mon)	Workflow and forming good machine learning questions from business questions	In-class exercises	Assignment 2 due at 6pm Monday 27th Jan
8	Jan 29, 2020 (Wed)	Support Vector Machines	In-class exercises	
9	Feb 3, 2020 (Mon)	Advanced ML techniques	In-class exercises	
10	Feb 5, 2020 (Wed)	Topics related to the group project	Work on group project	Assignment 3 due at 6pm Friday 7th Feb
	Feb 10-15, 2020	Exam Week		Final Assignment (group project) due