CMPT 300 Data Mining
Syllabus Fall 2017

Place & Time
Tuesdays and Thursdays (2:00 - 3:50 PM), Gore 227.

Course Website
Materials for this course are available at
https://CMPT300.github.io

Instructor
Jingsai Liang — Malouf 214 — JLiang@westminstercollege.edu
Office 801-832-2421, Cell 615-877-4259

Office Hours
I am available in my office during the following times:

Day	Time
Mondays & Wednesdays	10:00 - 11:30 AM
Mondays & Wednesdays	1:00 - 3:30 PM

If you cannot make any of the below times, you can contact me for a more convenient appointment.

I also regularly check email and make an attempt to respond to all email messages within a few hours on weekdays, and within a day on weekends.

Course Description
In many ways you may consider this class a logical extension of CMPT 202/306 – further data structures and algorithms. However, we will look at advanced data structures and the corresponding algorithms used to implement and manipulate them. Inherent in the study of algorithms is the careful study of the performance characteristics of such algorithms. Therefore, we will also look at the mathematical analysis and behavior of algorithms. Such a study will allow us to understand what an algorithm (and ultimately a computer) can and cannot do, and how to distinguish between two or more algorithms that solve the same task.

We will use Python for labs and homework assignments. No prior experience in Python is expected. Fortunately, learning Python is quite simple and it provides a rich library of common data structures that you can leverage for many programming tasks.

Prerequisites

CMPT 202, and MATH 201 and 210. From CMPT 202 you learned about object-oriented programming in Java, data structures, and algorithm analysis; skills you will be expected to use in this class.

Many topics covered in MATH 210 – notably recurrence relations – will be used as we study the analysis of algorithms.

Textbook

The recommended, but not required, text for this class is *Introduction to Data Mining*, by Pang-Ning Tan, etc.

Course Outline

The general topics we will cover this semester (although in no particular order) include:

- Data: types, quality, preprocessing, similarity.

- Classification

- Association Analysis

- Clustering Analysis

- Regression Analysis

- PCA

- SVM

Programming Labs

To reinforce theoretical concepts of data structures and algorithm analysis, there will be several programming assignments this semester using both Python and Java. We will spend two or three hours per week in a lab setting. Because the labs will involve pair programming, attendance during labs is mandatory. Labs will be graded credit / no credit for each step of the lab, adding up to a numeric score. You must demonstrate your completed lab to me by the end of the lab period for your grade. You should still upload your solutions to Canvas. If you do not finish your program during the lab period, you may choose to complete the lab separately or work as a team. You still must demonstrate your lab to me before the next lab period.

Homework

We will have regular homework assignments. These assignments will not be weighted equally. You are welcome to work together in designing a solution to a homework problem, but all answers and programming code should be your own work. (For labs, everything should be your and your partner's own work.) Please list all students you worked with on each assignment.

Exams

The midterm and final exams will be distributed as take-home exams. They will involve some programming. Exams are open-textbook (only the assigned textbooks from CMPT 202, MATH 210, and CMPT 306 courses) and notes, but you may not consult other books, the Internet, or people other than your instructor.

Late Policy

Each homework and lab assignment will have a particular weight and due date. It is expected that you complete all assignments and turn in all necessary information upon the due date. My policy will be to deduct 5% for each day that an assignment is turned in late (weekends count as one day.) Such a policy is intended to encourage you to plan accordingly and reward those who do.

Grading

Your grade will be assigned according to the following percentages:

$93+ \Rightarrow A$	$73..77.9 \Rightarrow C$
$90..92.9 \Rightarrow A-$	$70..72.9 \Rightarrow C-$
$88..89.9 \Rightarrow B+$	$65..69.9 \Rightarrow D+$
$83..87.9 \Rightarrow B$	$60..64.9 \Rightarrow D$
$80..82.9 \Rightarrow B-$	$55..59.9 \Rightarrow D-$
$78..79.9 \Rightarrow C+$	$0..54.9 \Rightarrow F$

Your course grade will be determined according to the following policy:

Topic	Weight
Labs	25%
Individual Homeworks	35%
Take-home midterm exam	20%
Take-home final exam	20%

*** The instructor reserves the right to alter the above grading scheme.

In Class Computer Use

Please use computers responsibly during class - do not write personal emails, web surf for non-academic purposes, play computer games, etc. while in class.

Academic Honesty

Cheating is not tolerated. A first cheating offense will earn a grade of zero on the assignment or exam, and a second offense will receive a failing grade for the course. Any cheating offense will be reported to the Dean of Students. Please refer to the Student Handbook for the College's statement on academic honesty.

Communication

This syllabus outlines course policies, and of course such policies can either be modified. But things work best when there is an open, two-way communication between both you and me. If you are having issues with the class, an assignment, etc. please talk to me. I value frank and open communication far more than any written policy.

Learning Goals

Westminster College has identified several learning goals that all graduates will achieve. In addition, the computer science program also has developed program learning goals that are specific for all computer science graduates. CMPT 306 Algorithms and Data Structures satisfies the following college and program-specific learning goals:

College-Wide Learning Goals

- *Critical, analytical and integrative thinking.* This is a focus of this course and is emphasized throughout! This course focuses on an evaluation and selection of appropriate data mining techniques that solve a variety of problems.

- *Creative and reflective capacities.* You will have to develop creative solutions to several problems presented to you this semester.

- *Leadership, collaboration, and teamwork.* Each week you will work with a partner in a pair programming lab exercise.

- *Writing and other communication skills.* There will be several written homework exercises this semester.

Pronouns, Correct Names, and Inclusion It is your right to be identified by your correct name and pronouns. I support people of all gender expressions and gender identities and welcome students to use whichever pronouns or names that best reflect who they are. In this spirit, I expect all students to also use the correct pronouns and names of classmates. Please inform me if my documentation reflects a name different than what you use and if you have any questions or concerns please contact me after class, by email, or during office hours.

Your Rights Under Federal Laws

Section 504 of Rehabilitation Act of 1973/ADA
Westminster College seeks to provide equal access in higher education to academically qualified students with physical, learning, and psychiatric disabilities. If you need disability-related accommodations in this class, have emergency medical information you wish to share with me, or need special arrangements in case the building must be evacuated, please inform me immediately. Please see me privately after class or in my office. Disability Services authorizes disability-related academic accommodations in cooperation with the students themselves and their instructors. Students who need academic accommodations or have questions about their eligibility should contact Karen Hicks, Disability Services Coordinator, in the START Center (801-832-2280) or email `disabilityservices@westminstercollege.edu` .

Title IX of the Education Amendments of 1972 prohibits sex discrimination against any participant in an educational program or activity that receives federal funds. Westminster is committed to providing a safe and non-discriminatory learning, living, and working environment to all members of the Westminster community and does not discriminate on the basis of sex. This includes on the basis of gender, gender identity, gender expression, nonconformity with gender stereotypes, or sexual orientation. The College?s Title IX policy strictly prohibits sexual assault, sexual harassment, gender-based harassment, gender-based discrimination, sexual exploitation, interpersonal violence (dating violence, domestic violence, stalking), and retaliation for making a good faith report of prohibited conduct or participating in any proceeding under the policy. The policy and accompanying procedures are available at www.westminstercollege.edu/about/resources/title-ix and discuss prohibited conduct, resources, reporting, supportive measures, rights, investigations, and sanctions for violations of the policy. If you want to make a report of prohibited conduct, you may contact Westminster?s Title IX Coordinator, Jason Schwartz-Johnson, or report an incident online. Jason can be reached at `jsj@westminstercollege.edu`, 801-832-2262, or in Malouf 107. You can also contact Deputy Coordinator Traci Siriprathane at `tsiriprathane@westminstercollege.edu` or 801-832-2862 or in HWAC 215. Please note that to the extent permitted by law, the College aims to protect the privacy of all parties involved in the investigation and resolution of reported violations of the policy. However, the College has a duty to look into and take actions in response to reports and cannot guarantee confidentiality or that an investigation will not be pursued. The Counseling Center is a confidential resource, and by law the counselors who work there cannot reveal confidential information to any third party without express permission unless there is an imminent threat of harm to self or others. **As an instructor I am a responsible employee and am required to report any information I obtain regarding conduct that may violate the policy to the Title IX Coordinator so that students can receive supportive measures and referrals to resources, they are aware of their options, and the safety of the campus community can be ensured.** If you begin to disclose an incident of prohibited conduct, I may interrupt you because I want to make sure that you have had the opportunity to

discuss the incident with confidential resources on and off campus first. If you need supportive measures inside or outside the classroom because of an incident of prohibited conduct, please reach out to the Title IX Coordinator for assistance.

Title VI of the Civil Rights Act of 1964 prohibits discrimination based on race, color, or national origin in any program or activity receiving federal financial assistance. In addition to these, Westminster?s Equal Opportunity policy prohibits discrimination or harassment based on ethnicity, age, religion, military status, or genetic information in any of its programs or activities. If you encounter this type of discrimination or harassment, or feel that you have been retaliated against for reporting prohibited conduct or participating in any related proceeding, you can contact the Equal Opportunity Officer, Jason Schwartz-Johnson. He can be reached at `jsj@westminstercollege.edu`, 801-832-2262, or in Malouf 107. You can also contact Julie Freestone, Equal Opportunity Administrator, at `jfreestone@westminsetercollege.edu`, 801-832-2573, or in Bamberger 106. The equal opportunity policy and procedures can be accessed from the Student Life webpage. **As an instructor, just as with Title IX, I am a responsible employee and am required to report any information I obtain regarding discrimination or harassment to the Equal Opportunity Officer for further review.**