Instructor Information

Instructor: Roselyne Barreto Tchoua (T is silent!)
Office: CDM 702
Office Hours: **online – posted on BLUESTAR on *Tuesdays 4:00-5:30* pm** via zoom and other times by appointments
Link: **https://depaul.zoom.us/j/96600255743**
Phone: 312-362-6796
Email: rtchoua@depaul.edu
Put DSC 478 in your subject – I should reply within 48 hours (or 2 business days)

Course Information

DSC478
Programming Machine Learning Applications
Winter 2021-2022 (Tuesday January 4th, 2022 – Tuesday March 8th, 2022)
**Tuesdays 5:45PM - 9:00PM in Lewis Center Room 1508
 (also Online Asynchronous)**
Note: The first two weeks are online synchronous for everyone. Recordings will be available to online asynchronous students.

Instructor Information

Instructor: Roselyne Barreto Tchoua (T is silent!)
Office: CDM 702
Office Hours (online): Tuesdays between 4:00 and 5:30pm via Zoom and other times by appointments
Link: Join URL: https://depaul.zoom.us/j/96600255743
Other times by appointment
Phone: 312-362-6796
Email: rtchoua@depaul.edu
Put DSC 478 in your subject – I should reply within 48 hours

Course Description

The course will focus on the implementations of various data science fundamental techniques using a high-level programming language. Students will have hands on experience developing both supervised and unsupervised machine learning algorithms (regression, classification and clustering) and will learn how to employ these techniques in the context of popular applications including recommender systems and text mining.

Prerequisite(s): DSC 441 (previously IS 467) and (DSC 430 or CSC 401)

Textbook & Couse Material

The required and recommended textbooks are listed below. The resources directly relevant to topics covered in the course are listed in Course Materials module. Check the DePaul Library for online versions.

Required: Machine Learning in Action, by Peter Harrington, Manning Publications, 2012.

Suggested:
- Python Data Science Essentials - Learn the fundamentals of Data Science with Python, by Alberto Boschetti and Luca Massaron, Packt Publishing, 2015

- Python for Data Analysis, by Wes McKinney, O'Reilly, 2012.

- Data Mining: Practical Machine Learning Tools and Techniques, by Ian Witten and Eibe Frank, 3rd Ed., Morgan Kaufmann, 2011.

Tentative List of Topics (Tentative schedule below):

The following issues and topics will be covered throughout the course. Many of these topics will be revisited several times during the course in a variety of contexts.

- **Data Mining and Knowledge Discovery**
  - The KDD process and methodology
  - Data preparation for knowledge discovery
  - Overview of data mining and Machine Learning techniques
  - Review of Python and overview of Python tools for Data Analysis
- **Supervised Techniques**
  - Classification and Prediction using K-Nearest-Neighbor
  - Classifying with Probability Theory; **_Naïve Bayes_**
  - Building Decision Trees
  - Forecasting and Regression models
  - Evaluating predictive models
- **Unsupervised Learning**
  - Clustering using K-Means
  - Principal Component Analysis and Dimensionality Reduction
  - **_Singular Value Decomposition_**
- **Advanced Topic**
  - **Support Vector Machines**
- **Possible Applications (covered throughout the course)**
  - **_Collaborative Recommender Systems_**
  - **_Concept Discovery from Documents, Blogs, Social Annotations_**

Tentative Schedule

Tuesday 1/4/22
- Introductions
- Overview of Data Mining and Knowledge Discovery Process
- Introduction to Jupyter notebooks
- Introduction to Numpy and Pandas

Tuesday 1/11/22 **(Assign Homework #1)**
- Discuss assignment schedule and final project
- Understanding Characteristics of Data
- Introduction to Numpy and Pandas (continued)
- Data Preparation

Tuesday 1/18/22 **(Open Quiz #1)**
- Distances, similarities and KNN
- Begin Classification & Prediction – Review Basic Concept

Sunday 1/23/22 Homework #1 due

Tuesday 1/25/22 **(Quiz #1 due before class – Assign Homework #2)**
- Supervised Learning
    - Classification & Prediction – Review Basic Concept (continued)
    - Text Classification

Tuesday 2/1/22 **(Open quiz #2)**
- Supervised Learning
    - Classification & Prediction – Review Basic Concept (continued)
    - Notes on Personalization and Recommender Systems

Sunday 2/6/22 Homework #2 due

Tuesday 2/8/22 **(Quiz #2 due before class) (Proposal due)**
- Supervised Learning
    - Personalization & Recommender Systems (continued)
    - Basic Regression Analysis
    - Model Selection & Optimization
        - Feature Selection
        - Parameter Selection

Tuesday 2/15/22 **(Assign Homework #3)** (Receive feedback on proposal)
- Unsupervised Learning
    - Clustering

Tuesday 2/22/22 **(Open Quiz #3)**
- Principle Component Analysis

- SVD (Singular Value Decomposition) and Matrix Factorization

Sunday 2/27/22 Homework #3 due

Tuesday 3/1/22 **(Quiz #3 due before class) (Assign Homework #4)**
- Final project deliverable discussion
- Support Vector Machines
- Testing models – Model selection

Tuesday 3/8/22
- Ensemble Methods
- Questions about assignments/projects
- Real-world applications
- Brief Course Summary

Sunday 3/13/22 Assignment #4 due

Sunday 3/20/22 Final project due

Grading

The structure and grading in the class will be centered around 4-5 assignments and a final project. The assignments will involve Python implementations of selected data mining techniques and their applications in various domains. The assignments will typically involve both programming components as well as problems related to the material covered in class. Some assignments may also involve the use of other open source data mining tools. These assignments must be done individually, unless otherwise specified. Late assignments will be penalized 5% per day. No late work will be accepted after three days since the assignment was due.

The final project will be a more complex programming/implementation assignment that will involve integrating multiple concepts and techniques. Student will be able to choose from among several possible projects ideas or propose their own. More details on the final project are available in the Project module.

The final grade will be determined (tentatively) based on the following components:
Assignments = 50%
Quizzes = 15%
Final Project = 35%

The general grading scheme will be based on a curve. At the end of the quarter, some adjustments may be made based on overall class performance as well as signs of individual effort. Plusses and minuses will be given at the high/low ends of each grade range.

The final grade will be assigned according to the following scale:

Percentage Grade	Letter Grade	Manner of fulfillment
95-100	A	Excellent
90-94	A-	
85-89	B+	Very Good
80-84	B	
75-79	B-	
70-74	C+	Satisfactory
65-69	C	
60-64	C-	Poor
55-59	D+	
50-54	D	
0 – 50	F	

IMPORTANT - **Graduate students taking courses under the purview of the School of Computing will be graded using A/B/C/D/F (no option for Pass/D/F).**

Homework Assignments and Final Exam Policies

The assignments must be submitted online on the D2L site at https://D2L.depaul.edu. Only legible, organized homework which shows your work will be graded. Include your name, section number, date, and homework number on the first page of your assignment. It is your responsibility to check that your files are uploaded correctly on D2L; you should always keep a copy of your submission. Extra credit points will be given for, active participation in the lectures and Discussion Forum (may round up your grade for participation).

Expectations

Content: This course contains reviews of DSC 441, but we cover more topics and at a slightly faster pace. We insist on programming the techniques we review and learn. **This does not mean you can ignore the concepts**. The quizzes will be my way to check that you are retaining fundamentals.

Homework: This is a graduate course, do not expect everything to be laid out of you, though the notebooks from in-class coding sessions will help. The assignments take some time, do not delay getting started because it "looks like you have time". Coding and debugging will take you sometimes looking things up, discussing things on D2L, *pulling your hair*, *banging your head against the wall*, that is the way you learn this material.

Do not ask for solutions on D2L. Do not post answers on D2L. Do discuss errors for example. If this happens, we will just have to close the platform.

Debugging by email does not work. Do not send me code and say, "what's wrong?". I do not debug code for students. That said, I am here to help, if you still have problems after searching for a solution, come to office hours and be ready to tell me what you have tried.

Preemptive piece of advice: get the examples (from class or other tutorials) working and change only one thing at the time.

The codes we run together in class give you most of what you need for your homework. On the one hand the homework is a lot of work, on the other hand you can reuse code from class. So, if you find yourself reinventing the wheel from scratch, you may be going the wrong track.

Attendance (Not directly applicable for synchronous course – but join Zoom call if possible!)

It is expected that you will attend every class and remain for the duration; it is the single most important action you can take in mastering the course objectives. Coming 15 minutes late or leaving 15 minutes constitutes an absence for the student. You are responsible for all material covered, assignments delivered or received, and announcements made in class sessions that you miss. For distance learning students, if you can join the zoom call it would be best, but recordings will be posted on D2L following the class session. This means that attendance translates into viewing the material in a timely manner, participating in the discussion forum, and being sure to email or call in any questions that you have.

For online students:
Normally, recordings of each lecture will be available a few hours after the "live" class and can be found at the course website https://d2l.depaul.edu. Online students and those who missed the zoom call will have access to the recording of the call. Students are expected to watch the lectures every week and to keep up with the course information posted on the course website.

Email

Email is the primary means of communication between faculty and students enrolled in this course outside of class time. Students should be sure their email listed under "demographic information" at http://campusconnect.depaul.edu is correct. **Please always put "DSC478" in the subject line**.

Changes to Syllabus

This syllabus is subject to change as necessary to better meet the needs of the students. Significant changes are unlikely and will be thoroughly addressed in class. If a change occurs, it will be thoroughly addressed during class and posted under Announcements in D2L.

Attitude – Also valid for online discussions

A professional and academic attitude is expected throughout this course. Measurable examples of non-academic or unprofessional attitude include but are not limited to: talking to others when the instructor is speaking, mocking another's opinion, cell phones ringing, emailing, texting or using the internet whether on a phone or computer. If any issues arise

a student may be asked to leave the classroom. The professor will work with the Dean of Students Office to navigate such student issues.

Civil Discourse

DePaul University is a community that thrives on open discourse that challenges students, both intellectually and personally, to be Socially Responsible Leaders. It is the expectation that all dialogue in this course is civil and respectful of the dignity of each student. Any instances of disrespect or hostility can jeopardize a student's ability to be successful in the course. The professor will partner with the Dean of Students Office to assist in managing such issues.

Cell Phones/On Call

(This is also not directly applicable for synchronous class, **however** mute yourself during the zoom call if you are not speaking)
If you bring a cell phone to class, it must be off or set to a silent mode. Should you need to answer a call during class, students must leave the room in an undisruptive manner (make sure to mute yourself). Out of respect to fellow students and the professor, texting is never allowable in class. If you are required to be on call as part of your job, please advise me at the start of the course.

School Policies

Online Course Evaluations

Evaluations are a way for students to provide valuable feedback regarding their instructor and the course. Detailed feedback will enable the instructor to continuously tailor teaching methods and course content to meet the learning goals of the course and the academic needs of the students. They are a requirement of the course and are key to continue to provide you with the highest quality of teaching. The evaluations are anonymous; the instructor and administration do not track who entered what responses. A program is used to check if the student completed the evaluations, but the evaluation is completely separate from the student's identity. Since 100% participation is our goal, students are sent periodic reminders over three weeks. Students do not receive reminders once they complete the evaluation. Please see https://resources.depaul.edu/teaching-commons/teaching/Pages/online-teaching-evaluations.aspx for additional information.

Mental Health and Academic Assistance

Balancing the hard work of achieving your educational goals with the other demands of life is difficult at the best of times. For many of us, for a variety of reasons, things are all the more difficult now. I want to make sure you feel comfortable, not embarrassed, reaching out to me for support. I will also point out where the University has great resources just a phone call or email away. These have been created and maintained for you, so use them. Sometimes people feel like their situation isn't the worst possible, so they assume they do not need help, but don't let that prevent you from reaching out.

- DePaul University Counseling Services – mental health is as important as physical health, and we have professionals just a phone call away:

https://offices.depaul.edu/student-affairs/about/departments/Pages/ucs.aspx (call (773) 325-7779 or 911 for emergency).
- The kind people at the Office of the Dean of Students can help you with a wide range of topics, including figuring out if you should withdraw or apply for an incomplete: https://offices.depaul.edu/student-affairs/about/departments/Pages/dos.aspx .
- There are lots of additional, more specific resources listed here with the Office of Student Affairs, including crisis hotlines and sexual assault resources (note Title IX refers to a law protecting you from sex discrimination, including harassment and assault): https://offices.depaul.edu/student-affairs/support-services/counseling/Pages/Crisis-Hotlines.aspx .
- Finally, the associate Dean Dr. Lucia Dettori has offered to be a resource who can direct you to the right office if and when necessary (ldettori@cdm.depaul.edu).

Academic Integrity and Plagiarism

This course will be subject to the academic integrity policy passed by faculty. More information can be found at https://resources.depaul.edu/teaching-commons/teaching/academic-integrity/Pages/default.aspx
The university and school policy on plagiarism can be summarized as follows: Students in this course should be aware of the strong sanctions that can be imposed against someone guilty of plagiarism. If proven, a charge of plagiarism could result in an automatic F in the course and possible expulsion. The strongest of sanctions will be imposed on anyone who submits as his/her own work any assignment which has been prepared by someone else. If you have any questions or doubts about what plagiarism entails or how to properly acknowledge source materials be sure to consult the instructor.

All students are expected to abide by the University's Academic Integrity Policy which prohibits cheating and other misconduct in student coursework. **Publicly sharing or posting online any prior or current materials from this course (including exam questions or answers), is considered to be providing unauthorized assistance prohibited by the policy. Both students who share/post and students who access or use such materials are considered to be cheating under the Policy and will be subject to sanctions for violations of Academic Integrity**.

This also applies to sharing and copying someone else's work. I will consider the one who shares as having the same fault as the one who copies.

Withdrawal
Students who withdraw from the course do so by using the Campus Connection system (http://campusconnect.depaul.edu). Withdrawals processed via this system are effective the day on which they are made. Simply ceasing to attend, or notifying the instructor, or nonpayment of tuition, does not constitute an official withdrawal from class and will result in academic as well as financial penalty. The dropping dates can be found at: https://academics.depaul.edu/calendar/Pages/default.aspx

Retroactive Withdrawal

This policy exists to assist students for whom extenuating circumstances prevented them from meeting the withdrawal deadline. During their college career students may be allowed one medical/personal administrative withdrawal and one college office administrative withdrawal, each for one or more courses in a single term. Repeated requests will not be considered. Submitting an appeal for retroactive withdrawal does not guarantee approval.

All students are required to manage their class schedules each term in accordance with the deadlines for enrolling and withdrawing as indicated in the University Academic Calendar. Information on enrollment, withdrawal, grading and incompletes can be found at: http://www.cdm.depaul.edu/Current%20Students/Pages/PoliciesandProcedures.aspx

Excused Absence (not applicable for synchronous online course)

In order to petition for an excused absence, students who miss class due to illness or significant personal circumstances should complete the Absence Notification process through the Dean of Students office. The form can be accessed at http://studentaffairs.depaul.edu/dos/forms.html. Students must submit supporting documentation alongside the form. The professor reserves the sole right whether to offer an excused absence and/or academic accommodations for an excused absence.

Incomplete Grades

An incomplete grade is a special, temporary grade that may be assigned by an instructor when unforeseeable circumstances prevent a student from completing course requirements by the end of the term and when otherwise the student had a record of satisfactory progress in the course. All incomplete requests must be approved by the instructor of the course and a CDM Associate Dean. Only exceptions cases will receive such approval. Information about the Incomplete Grades policy can be found at http://www.cdm.depaul.edu/Current%20Students/Pages/Grading-Policies.aspx

Preferred Name & Gender Pronouns

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance, and nationalities. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the quarter so that I may make appropriate changes to my records. Please also note that students may choose to identify within the University community with a preferred first name that differs from their legal name and may also update their gender. The preferred first name will appear in University related systems and documents except where the use of the legal name is necessitated or required by University business or legal need. For more information and instructions on how to do so, please see the Student Preferred Name and Gender Policy at http://policies.depaul.edu/policy/policy.aspx?pid=332

Students with Disabilities

Students seeking disability-related accommodations are required to register with DePaul's Center for Students with Disabilities (CSD) enabling them to access

accommodations and support services to assist with their success. There are two office locations:
• Loop Campus – Lewis Center #1420 – (312) 362-8002
• Lincoln Park Campus – Student Center #370 – (773) 325-1677
Students who register with the Center for Students with Disabilities are also invited to contact Dr. Gregory Moorhead, Director of the Center, privately to discuss how he may assist in facilitating the accommodations to be used in a course. This is best done early in the term. The conversation will remain confidential to the extent possible.

Please see https://offices.depaul.edu/student-affairs/about/departments/Pages/csd.aspx for Services and Contact Information.