Saint Louis University

Department of Computer Science

Spring 2024

CSCI 1070-1: Introduction to Computer Science-Taming Big Data

3 Credit Hours

Instructor Contact Information

Name: Alexis McKenzie, M. Eng

E-mail: alexis.mckenzie@slu.edu (please start emails with CSCI_1070)

Office Hours: After class until there are no questions remaining

Meeting Time/Location

Lectures: M/W 6:15-7:30 PM, Ritter Hall 115

TEXTBOOKS

"Data Science from Scratch: First Principles with Python (2nd Edition)" by Joel Grus https://a.co/d/03N8TYR

Instructor Communication

I work full-time as a Solution Architect, working on machine learning and data engineering projects, so I am generally not available prior to 4:30 CT. However, I care deeply about you and your understanding of the material covered in this course so I will respond to your questions and emails in the evenings. I will also be available to meet with you preferably after class during my office hours or by appointment throughout the week over Zoom.

COURSE DESCRIPTION

This course is an introduction to data science and machine learning. Fundamentals of software engineering, data visualization, databases, and data analysis will also be covered. There will be a focus on mastering the basic concepts and algorithms in data science and applying them to real-world problems. **The first half of this course will feel a lot more difficult than the second, especially if this is your first CS class. I have been there, I have had students where you are, and with the right effort, you can do this!**

Prerequisites: None.

OUTCOMES OF THE COURSE:

After successfully completing this course, students will be able to:

● use flow control structures to process large datasets in a high-level programming language

● use git for version control

● apply basic concepts of Software Engineering in the implementation of a computer program that satisfies a set of requirements, including thorough testing and iteration until requirements are met

● select and apply appropriate machine learning algorithms to real-world datasets

● create database tables, define a database schema, and read and write data from a database

● formulate and solve real-world problems in data science, including an appropriate evaluation of the method chosen, and present results in both written and presentation form

COURSE CONTENT

Date	**Subject**	**Reading**
01/17/2024	Syllabus + Jupyter and Python Fundamentals	
01/22/2024	Flow Control, Functions, and Classes	Chap. 1-2
01/24/2024	Flow Control, Functions, and Classes 2	
01/29/2024	Getting Data	Chap. 9-10
01/31/2024	Data Visualization	Chap. 3
02/05/2024	Statistics and Probability	Chap. 5-7
02/07/2024	Statistics and Probability 2	
02/12/2024	Intro to Machine Learning	Chap. 8, 11, 25
02/14/2024	Supervised Learning and Pre-processing Data	
02/19/2024	KNN and Linear Regression	Chap. 12
02/21/2024	Multiple Linear Regression	Chap. 14
02/26/2024	Logistic Regression	Chap. 15-16
02/28/2024	SQL	Chap. 24
03/04/2024	Review	
03/06/2024	Midterm	---------
03/11/2024	Fall Break, No Class	

03/13/2024		
03/18/2024 03/20/2024	SQL 2 Decision Trees	 Chap. 17
03/25/2024 03/27/2024	Linear Algebra and PCA Unsupervised Learning	Chap. 4 & 10 Chap. 20
04/03/2024	SVM, oversampling, and undersampling	
04/08/2024 04/10/2024	Naive Bayes Neural Networks	Chap. 13 Chap. 18-19
04/15/2024 04/17/2024	Network Analysis Recommendation Systems	Chap. 22 Chap. 23
04/22/2024 04/24/2024	Natural Language Processing Prompt Engineering and AI	Chap. 21
04/29/2024 05/01/2024	Project & Presentation Project & Presentation	---------

Please note: readings should be completed PRIOR to the class for which they are assigned.

EVALUATION AND GRADING

Exams: There will be a **one-hour fifteen-minute long midterm** and a **final project**. There will be no final exam. The final project will serve as a confirmation you have learned all the material in the learning objectives for this course. The midterm will have a combination of conceptual questions and pseudocoding.

Quizzes: There will be 12 quizzes (10 counted) throughout the course that will occur at the end of class. They are designed to be completed in class. Quiz questions will come from assigned readings and previous or current lectures. You will often be asked to pseudocode or provide an outline of the code you would write to solve the problem in the question. We will discuss this more in class. Several quiz questions will reappear on the midterm. If you miss a quiz there will be no makeup.

Homework: You will have 13 homework assignments of which I will count the highest 12 grades. The first two homework assignments will not require you to use git, but after that, you should submit a direct Github link to your homework assignment through Canvas. If you submit your first two weeks of homework with a Github link you can earn a point of extra credit per assignment.

It is possible for everyone in the class to earn an A. No make-ups will be allowed for any reason for the only exam in this course (the midterm), other than a verifiable, written doctor's note or permission granted by the instructor prior to the date of the exam. Be forewarned, it is HIGHLY UNLIKELY that I will grant that extension. In addition, I will make every attempt to have the midterm and quizzes graded within 1 week of the dates administered.

No late assignments will be accepted. All homework must be submitted online to the instructor. All assignments are due on the day indicated by 11:59 pm Central Time. You have one week to submit assigned homework, though I will make every attempt to post assignments early so you have more than a week to work on them. Saint Louis University is in the Central Time Zone and all times noted are relative to that time zone, no matter the location of the student. Adjust your submission time accordingly. You will have to work on the assignments outside of class on your own time.

Saint Louis University uses the +/- system from A to C, with no plus or minus for grades of D or F. Plus and minus will be apportioned with a "+" assigned for the upper third of the letter grade range and a "-" for the lower third of the letter grade range.

I will apply the following scaled grading curve:

Point Range	Grade
≥ 93%	A
≥90%	A-
≥87%	B+

≥84%	B
≥80%	B-
≥77%	C+
≥74%	C
≥70%	C-
≥60%	D
Below 60%	F

Grading Scale

Grades will be made available on Canvas. I will push to post grades within two weeks of being turned in. Your grade will be calculated in the following manner:

- Midterm Exam	1 x 75 pts each	75 pts
- Quizzes*	10 x 5 pts each	50 pts
- Homework Assignments*	12 x 25 pts each	300 pts
- Final Project and Presentation	1 x 75 pts each	75 pts
Total		500 pts

*Although the above shows you will only be graded on 10 quizzes and 12 homework assignments, you will have 12 quizzes and 13 homework assignments in this course. I will drop the lowest two quiz grades and the lowest homework grade.

Attendance

You are responsible for all the material and administrative announcements presented during class. Note that while attendance is not strictly mandatory, there is a quiz component of your grade that depends on being present to take it. Should you choose not to attend class, you will not earn points for the quiz, should it occur on that day.

Collaboration Policy

Unless explicitly stated otherwise, all work that you submit should be the result of your own effort. For ALL course work, you ARE NOT permitted to consult the solutions from another student (former or current), copy/consult the provided solutions from previous years, or look online for exact problem solutions. We will monitor results from ChatGPT and other AIs, and if you produce an answer that is the same as or nearly the same as what is produced, you will get a zero for that assignment. AI is a useful tool but it is often wrong, and your ability to critically think through hard problems will always be valuable no matter how good AI gets.

When working on course materials, you may discuss approaches to solving the problems with your classmates. However, you must work out all details of any solutions discussed and write up the solution completely on your own. In particular, when working with a student on an assigned homework problem you should do so verbally -- nothing should be written down. This will keep your discussion at a high-level so that everyone can work out the details on their own.

Your final project may be completed in groups of two or it may be completed on your own. I will check git to see that both students contributed roughly evenly to the project. There will be additional requirements for the project if you choose to work on a team, but it is designed to be less work for each person if you decide to partner. In industry and academia, collaboration is critical and I want to foster that where it makes sense in this course.

Changes to Course Policies

The standards and requirements set forth in this syllabus may be modified at any time by the course instructor. Notice of such changes will be announced in class, via email, or posted to the course website.

College Policies

Computing Resources

For those of you who choose to use the lab computers, please read the department and university policies (Links to an external site.) on the appropriate use of computer systems.

Title IX

Saint Louis University and its faculty are committed to supporting our students and seeking an environment that is free of bias, discrimination, and harassment. If you have encountered any form of sexual misconduct (e.g., sexual assault, sexual harassment, stalking, domestic or dating violence), we encourage you to report this to the University. If you speak with a faculty member about an incident that involves a Title IX matter, that faculty member must notify SLU's Title IX coordinator (or that person's equivalent on your campus) and share the basic facts of your experience. This is true even if you ask the faculty member not to disclose the incident. The Title IX contact will then be available to assist you in understanding all of your options and in connecting you with all possible resources on and off campus.

For most students on the St. Louis campus, the appropriate contact is Anna R. Kratky (DuBourg Hall, room 36; anna.kratky@slu.edu; 314-9773886). If you wish to speak with a confidential source, you may contact the counselors at the University Counseling Center at 314-977-TALK. To view SLU's sexual misconduct policy, and for resources, please visit the following web addresses: https://www.slu.edu/here4you (Links to an external site.) and https://www.slu.edu/general-counsel (Links to an external site.).

Disability Accommodations

Students with a documented disability who wish to request academic accommodations must formally register their disability with the University. Once successfully registered, students also must notify their course instructor that they wish to use their approved accommodations in the course.

Please contact Disability Services to schedule an appointment to discuss accommodation requests and eligibility requirements. Most students on the St. Louis campus will contact Disability Services, located in the Student Success Center and available by email at Disability_services@slu.edu or by phone at 314.977.3484. Once approved, information about a student's eligibility for academic accommodations will be shared with course instructors by email from Disability Services and within the instructor's official course roster. Students who do not have a documented disability but who think they may have one also are encouraged to contact to Disability Services. Confidentiality will be observed in all inquiries.

Academic Integrity

Academic integrity is honest, truthful and responsible conduct in all academic endeavors. The mission of Saint Louis University is "the pursuit of truth for the greater glory of God and for the service of humanity." Accordingly, all acts of falsehood demean and compromise the corporate endeavors of teaching, research, health care, and community service via which SLU embodies its mission.

The University strives to prepare students for lives of personal and professional integrity, and therefore regards all breaches of academic integrity as matters of serious concern. The governing University-level Academic Integrity Policy was adopted in Spring 2015, and can be accessed on the Provost's Office website (Links to an external site.).

Additionally, each SLU College, School, and Center has adopted its own academic integrity policies, available on their respective websites. All SLU students are expected to know and abide by these policies, which detail definitions of violations, processes for reporting violations, sanctions, and appeals. Please direct questions about any facet of academic integrity to your faculty, the chair of the department of your academic program, or the Dean/Director of the College, School or Center in which your program is housed.

Student Success Center

In recognition that people learn in a variety of ways and that learning is influenced by multiple factors (e.g., prior experience, study skills, learning disability), resources to support student success are available on campus. The Student Success Center assists students with academic-related services and is located in the Busch Student Center (Suite 331). Students can visit https://www.slu.edu/life-at-slu/student-success-center/ (Links to an external site.) to learn more about tutoring services, university writing services, disability services, and academic coaching.

University Writing Services

Students are encouraged to take advantage of University Writing Services in the Student Success Center; getting feedback benefits writers at all skill levels. Trained writing consultants can help

with writing projects, multimedia projects, and oral presentations. University Writing Services offers one-on-one consultations that address everything from brainstorming and developing ideas to crafting strong sentences and documenting sources. For more information, visit https://www.slu.edu/life-at-slu/student-success-center/ (Links to an external site.) or call the Student Success Center at 314-977-3484.

Basic Needs

Security Students in personal or academic distress and/or who may be specifically experiencing challenges such as securing food or difficulty navigating campus resources, and who believe this may affect their performance in the course, are encouraged to contact the Dean of Students Office (deanofstudents@slu.edu or 314-977-9378) for support. Furthermore, please notify the instructor if you are comfortable in doing so, as this will enable them to assist you with finding the resources you may need.