


CIS 635 Knowledge Discovery & Data Mining

Week 1: Introduction to CIS 635



Week 1: Module - Introduction to CIS 635

Hidden from students ▾

...



This course introduces students to computational methods for knowledge discovery and data mining. Topics covered include Python programming (specialized on data science), data preparation, data visualization, predictive modelling, model evaluation, clustering, and association analysis techniques.



Week Overview

Hidden from students ▾

...



Read: Some background materials; plus take a survey before.

Hidden from students ▾

...



EXPLORE: Test your Google Colab(oratory) setup

Hidden from students ▾

...



Week overview

This week we will:

- Get to know each other (networking)
- Set up our course objective, guidelines, and evaluation procedure.
- Set up our programming development environment(s), more specifically, **Google Colab**(oratory) on your **Google drive**.
- Test source code and report delivery methods.

Due:

- We don't have any programming assignments for this week.
- What you have to do is to make sure your **Jupyter Notebook** on your **Google Colab**(oratory) environment is ready to write and test code.



Networking!

About myself

- My name is Dr. Kamrul Hasan; you can call me as **Kamrul**
- Pronoun: he/him/his
- Born and raised in a tiny south Asian country, **Bangladesh**



About myself

- Born and raised in a tiny south Asian country, Bangladesh
- **Graduated from University of Montreal, 2014**
 - Multi-media data mining



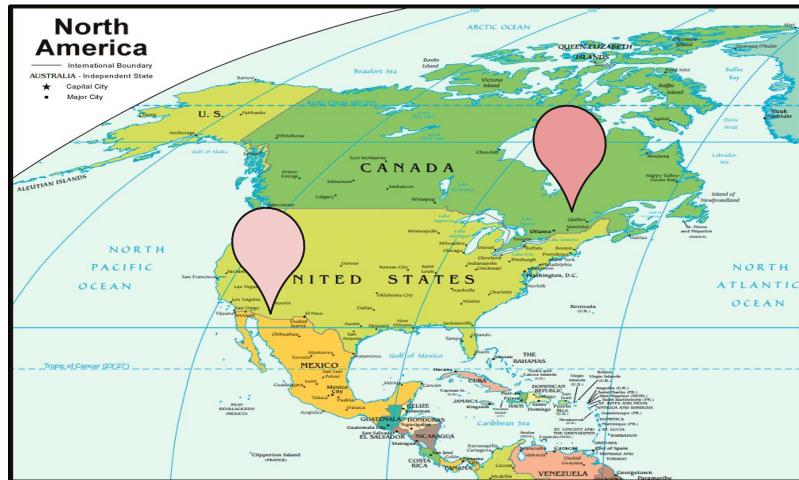
About myself

- Born and raised in a tiny south Asian country, Bangladesh
- Graduated from University of Montreal, 2014
 - Multi-media data mining
- **Probabilistic ML**
 - **Semi-supervised Learning**
 - **Generative AI** (example: LLMs ,ChatGPT)



About myself

- Born and raised in a tiny south Asian country, Bangladesh
- Graduated from University of Montreal, 2014
- **Last ~9 years as a Data Scientist; delivered several AI products in multiple domains**



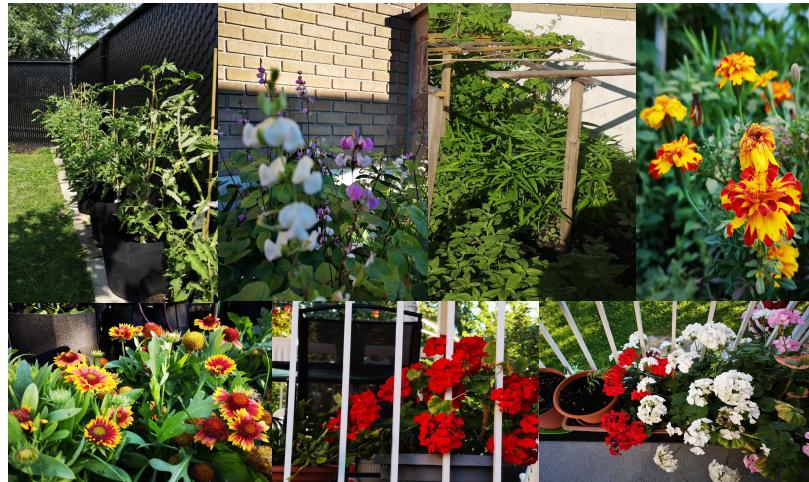
About myself

- Born and raised in a tiny south asian country, Bangladesh
- Graduated from University of Montreal, 2014
- Last ~9 years as a Data Scientist; delivered several AI products in multiple domains
- **Travelling**



About myself

- Born and raised in a tiny south asian country, Bangladesh
- Graduated from University of Montreal, 2014
- Last ~9 years as a Data Scientist; delivered several AI products in multiple domains
- Travelling
- **Gardening**



About you!

- I hope I would get to know each of you as we progress
- We will meet in person and/or virtually as per our availability
- You will collaborate as groups
 - For discussion different topics: course specific and beyond
 - Final project
- I will seek your suggestions

CIS 635 Objective, guidelines, and evaluation procedure

General information (about the course)

Description: This course introduces students to

- Computational methods for knowledge discovery and data mining.
- Topics covered include
 - Python programming (specialized on data science),
 - EDA: data preparation, analysis, and visualization
 - Predictive modeling, model evaluation, clustering, and association analysis techniques.

Objective: Completing this course, you should be able to:

- Wrangle, analyze, and visualize data.
- Exercise and apply supervised and unsupervised knowledge discovery techniques.
- Identify and build models using common data mining techniques.
- Evaluate models for their effectiveness and appropriateness.
- Communicate findings using effective visualizations.

General information (about the course)

Description: This course introduces students to

- Computational methods for knowledge discovery and data mining.
- Topics covered include
 - Python programming (specialized on data science),
 - EDA: data preparation, analysis, and visualization
 - Predictive modeling, model evaluation, clustering, and association analysis techniques.

Objective: Completing this course, you should be able to:

- Wrangle, analyze, and visualize data.
- Exercise and apply supervised and unsupervised knowledge discovery techniques.
- Identify and build models using common data mining techniques.
- Evaluate models for their effectiveness and appropriateness.
- Communicate findings using effective visualizations.



General information (about the course)

Textbook(s): There are no required textbooks for this course. Some good options include:

- **Data Mining** - A Knowledge Discovery Approach by Cis Pedrycz, and Swiniarski (available online through the GVSU library)
- Introduction to Data Mining Concepts and Techniques, 2nd Edition by Tan, Steinbach, and Kumar
- **Pattern Recognition** and Machine Learning by Christopher M. Bishop, 2016
- **Machine Learning with Python Cookbook:** Practical Solutions from Preprocessing to Deep Learning Paperback – April 17 2018 by Chris Albon

- Fundamentals of **Data Visualization** by Claus Wilke (available online)
- **Python for Data Analysis** by Wes McKinney (accessible online through library)
- Automate the Boring Stuff with Python: Practical Programming for Total Beginners Paperback, by Al Sweigart (accessible online through library)
- Python Data Visualization Essentials Guide: Become a Data Visualization expert by building strong proficiency in Pandas, Matplotlib, Seaborn, Plotly, Numpy, and Bokeh (English Edition) Paperback – July 30 2021, by Kalilur Rahman

Requirements: **Blackboard**, **Computer** access, ability to run **Python** (and install any requisite Python packages). If you have trouble, talk to the instructor, and we will work out a solution together.



Tentative Schedule

General information:

- See **Blackboard** for detailed schedule
- Planning and delivery per week
- **Course materials and assignments** will be available on Blackboard **mostly** at the beginning of each week; sometimes as we progress.

Update your Blackboard settings so you receive your emails, messages, and notifications in time.

Key dates:

- Jan 08: Class begins (weekly planning & delivery)
- March 03-10: March break
- April 20: Class Ends
- April 27: Semester ends

Office: MAK D-2-216 (by appointment),
DCIH 530H (Thursday: 2:00- 4:00 pm)

Meeting Times: Online or in person

Evaluation

Grading distribution:

- Weekly reflections (including class & project meeting attendance & participation): 10%
- Class tests: 50% (best 3 out of total 6)
- Final Project (group + individual performance): 40%
 - *Assignments: 10% (group and/or individual)*
 - *Final Project: 10% (group)*
 - *Documentation: 10% (group)*
 - *Presentation: 10% (group + individual)*

Grade points:

A	93%	C	73%
A-	90%	C-	70%
B+	87%	D+	67%
B	83%	D	60%
B-	80%	F	Below 60%
C+	77%		

Note: Your final grade percentage will be rounded to the next integer percentage value. For example, an **89.1%** will round up to a **90%**.



Policy & expectation

Expectation: I expect the following to ensure your success in this course:

- check Blackboard on a regular basis for announcements, course material, and assignments
- stay up to date with required course materials.
- let me know how the class and my teaching can be improved
- adhere to the **GVSU policy of Academic Honesty** <http://www.gvsu.edu/coursepolicies/>

Course policy:

- Weekly reflections and homework assignments are to be completed **individually**.
- Due dates: All assignments will be due at 11:59pm Michigan time on the due date (unless otherwise stated, the first Monday following the assignment week).
- Late policy: You will lose 10% off of your maximum grade per day late, to a cap of five days (50% off), after which the assignment will not be accepted.



Policy & expectation

Course policy (cont.):

- Homework extensions: Life happens, and so I provide **5 free no-questions-asked late days** that can be used on a single homework assignment or spread across multiple assignments without penalty. These days cannot be applied to project deadlines.
- The instructor reserves the right to modify course policies, the course calendar, and due dates.

Academic Honesty: All students are expected to adhere to the academic honesty standards set forth by GVSU. In addition, students in this course are expected to adhere to the academic honesty guidelines as set forth by the school of computing. Details can be found at <https://www.gvsu.edu/computing/academic-honesty-30.htm>

I believe that you can learn a lot from your peers, both in the class and in the broader community. Therefore, I encourage collaboration with both. However, do not mistake this as a license to cheat. *Learning* from and with your peers is encouraged, but passing their work off as your own is prohibited.



Policy & expectation

Academic Honesty (cont.): With respect to all individual assignments in this course:

- Document collaboration; **no electronic transfer of code** between students is permitted.
- Code used from internet (including tools such as **ChatGPT**) must be cited, an active link to that code. Ref. code should not solve the **entirety of an assigned** problem/project (i.e., don't have someone else do your project).
- You are encouraged to engage in conversations in **online forums**, but do not post solutions or solicit others to complete your work for you.
- You are encouraged to talk about problems with each other in **non-technical terms** (i.e., not code).
- Ultimately, **you are responsible** for all aspects of your submissions. You should be able to explain and defend your submission if the work is entirely your own.



Useful resources

Blackboard: Course materials, assignments, grades, and announcements will be posted to Blackboard (<https://lms.gvsu.edu/>). It is your responsibility to stay informed.

Other academic resources: GVSU also provides opportunities for students to improve your academic skills through resources, such as:

- [The writing center](#)
- [Speech lab](#)
- [Research consultants](#)
- [Library liaisons](#)

Disability support : If you are in need of accommodations due to disability you must present a memo to me from Disability Support Resources (DSR), indicating the existence of a disability and the suggested reasonable accommodations. If you have not already done so, please contact the Disability Support Resources office (215 CON) by calling 331-2490 or email to dsrgvsu@gvsu.edu. Please note that I cannot provide accommodations based upon disability until I have received a copy of the DSR issued memo. All discussions will remain confidential. For more information, see <https://www.gvsu.edu/dsr/>

Religious observance: Many University students, staff, and faculty observe religious traditions from a variety of religions. The [Religious Inclusion Policy](#) acknowledges the right of students, staff, and faculty to engage in religious observances. If you need special accommodations to observe a religious holiday, make arrangements in advance with the instructor.