Lecture 1 - Introduction to Machine Learning; Getting Started with Git and Python Demonstrations

EEE 4773 Fundamentals of Machine Learning

Course Description: (3 credits) Overview of machine intelligence and the role of machine learning in a variety of real-world problems. Probability and statistics to handle uncertain data. Topics covered include learning models from data in both a supervised and unsupervised fashion, linear models and non-linear models for classification, and linear dimensionality reduction.

- **This course relies mostly on foundational Machine Learning math but also programming!**
- We will use Python for in-class algorithm experiments, class assignments and final project

Instructor: Dr. Catia S. Silva

Instructional Assistant Professor @ ECE, UF

- Office: NEB 467
- Phone: (352) 392-6502
- Email: catiaspsilva@ece.ufl.edu
- Office hours: Mondays 3:00 PM - 4:00 PM, Wednesdays and Thursdays 10:00 AM - 11 AM
  - Temporary arrangements: **all** office hours will be held via Zoom, **Mondays** will have an **in-person** option with limited seating

About Me

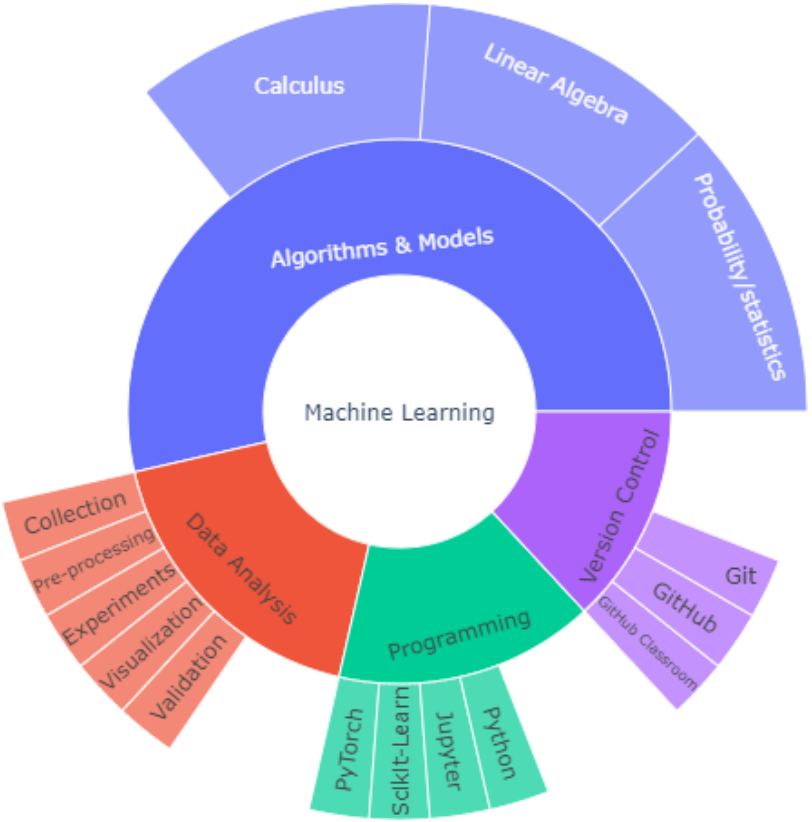
- My pronous are she/her/hers
- The phonetic spelling of my name, Cátia Silva, is Ka-tee-uh SIL-vuh
- PhD in Electrical and Computer Engineering, University of Florida, 2018
- Research interests: machine learning with applications, patter recognition, computational neuroscience, natural language processing, data science, engineering education
- I have industry experience in a large institutition (INESC, Portugal) and small startup company (Boca Raton, Florida)
- I have been working on ML and ML applications for about 10 years
- I was born and raised in Portugal

- I speak Portuguese (native), English, French, Spanish and limited Italian
- I have a 6-year-old Cavalier King Charles Spaniel dog, her name is Maggie
- Enjoy spend time doing fitness activities (currently training for my best 10K), hiking with Maggie, solve puzzles and games, try new cuisines, finding new music

What tools will we use to perform Machine Learning?

In [2]:
```
from IPython.display import Image
Image('figures/tools.png', width=800)
```

Out[2]:



Why should you care?

The biggest advantage to machine learning is that it allows us to do things much more quickly than we would be able to do otherwise.

It can't solve problems that a human being couldn't also solve, but it can take in a huge amount of data and very quickly build connections and predictions based on it. That becomes even more important as we continue to expand the amount of data we are generating through IoT and connected devices.

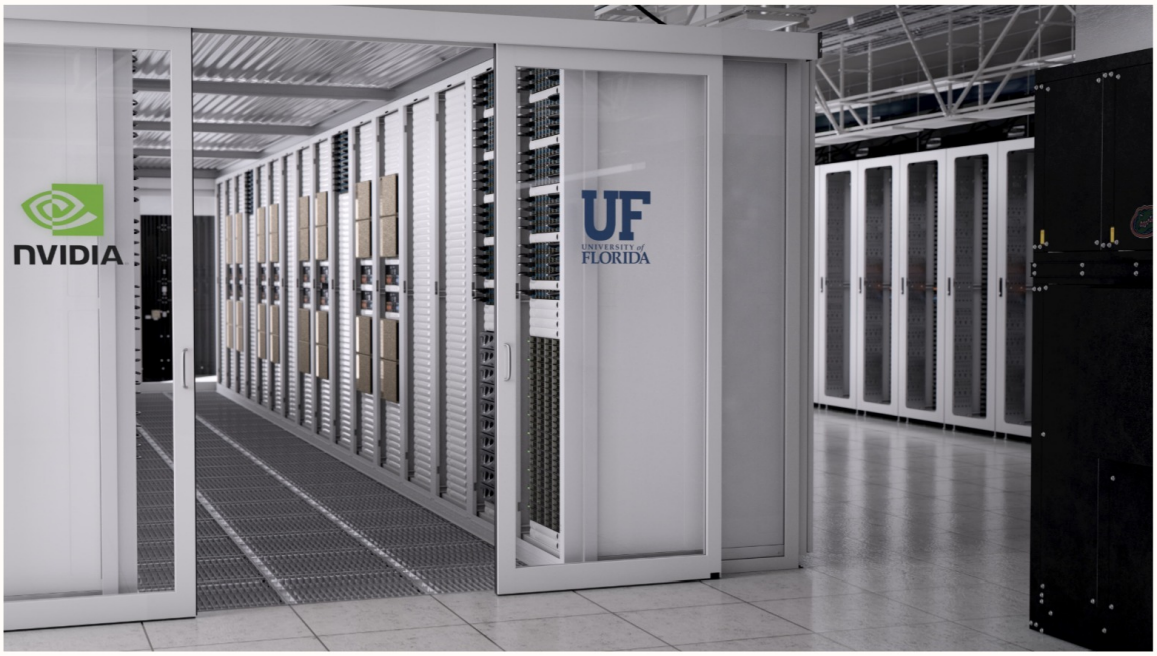
Market Examples

Any industry with access to data can benefit from a greater understanding about what that data means -- whether that's a manufacturing plant trying to anticipate repairs, or the makers of a driverless car. Here's how some industries are using machine learning:

- Healthcare - making medical diagnoses more quickly and more accurate
- Social Media - customizing content based on user behavior
- Manufacturing - antecipating repairs and improving preventive maintenance
- Security - fending off cyber attacks based on anomalous behavior
- Transportation - analyzing and responding to the real world environment in the development of driverless cars
- Agriculture - accurate prediction and estimation of farming parameters to optimize the economic efficiency of livestock production systems
- Gaming - analyzing competitive play to anticipate moves and create more challenging enemies
- Retail - recommending products based on past behavior and similar customers
- Marketing - analyzing customer responses to ads
- Customer service - creating intelligent virtual bots to manage customer interactions
- Real estate - generating property recommendations
- and many more...

UF-NVIDIA Partnership

UF announces $70 million artificial intelligence partnership with NVIDIA



Artist's rendering of University of Florida's new AI supercomputer based on NVIDIA DGX SuperPOD architecture.

Thanks to ECE alumnus Chris Malachowsky!

UF Supecomputer: Hipergator

We will have computing time for the entire semester in Hipergator!

So... what is Artificial Intelligence?

While a number of definitions of artificial intelligence (AI) have surfaced over the last few decades, John McCarthy offers the following definition in this 2004 paper,

> "*It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.*"

However, decades before this definition, the birth of the artificial intelligence conversation was denoted by Alan Turing's seminal work, "Computing Machinery and Intelligence", which was published in 1950. In this paper, Turing, often referred to as the "father of computer science", asks the following question, **"Can machines think?"**

From there, he offers a test, now famously known as the **"Turing Test"**, where a human interrogator would try to distinguish between a computer and human text response. While this test has undergone much scrutiny since its publish, it remains an important part of the history of AI as well as an ongoing concept within philosophy as it utilizes ideas around linguistics.

But intelligence doesn't have to depend on relating it to human intelligence. But we cannot yet characterize in general what kinds of computational procedures we want to call intelligent, without without comparing it with human intelligence.

The ideal approach is:

- Systems that think rationally
- Systems that act rationally

Whereas the human approach is:

- Systems that think like humans
- Systems that act like humans

Alan Turing's definition falls under the category of "systems that act like humans."

The paper "A Collection of Definitions of Intelligence" (2007) lists **53 definitions of Intelligence**.

Examples: Intelligence is the ability to

1. Learn or understand or to deal with new or trying situations
2. Acquire and apply knowledge
3. Apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (such as tests)
4. Learn facts and skills and apply them...

At its simplest form, artificial intelligence is a field, which combines computer science and robust datasets, to enable problem-solving. It also encompasses sub-fields of **machine learning** and **deep learning**, which are frequently mentioned in conjunction with artificial intelligence. These disciplines

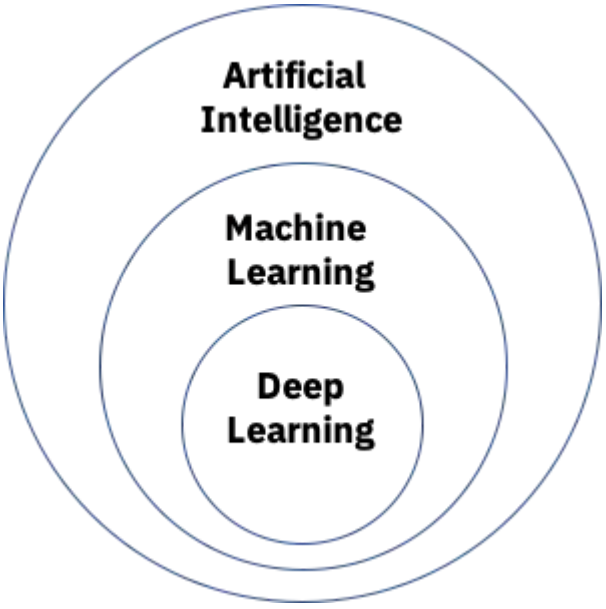
are comprised of AI algorithms which seek to create expert systems which make predictions or classifications based on input data.

Types of artificial intelligence

1. **Weak AI** or **Narrow AI** is AI trained and focused to perform specific tasks. Narrow AI drives most of the AI that surrounds us today. Examples: Amazon's Alexa, IBM Watson, autonomous vehicles

2. **Strong AI** is made up of **Artificial General Intelligence (AGI)** and **Artificial Super Intelligence (ASI)**. Artificial general intelligence (AGI), or general AI, is a theoretical form of AI where a machine would have an intelligence equaled to humans; it would have a self-aware consciousness that has the ability to solve problems, learn, and plan for the future. Artificial Super Intelligence (ASI)—also known as superintelligence—would surpass the intelligence and ability of the human brain. While strong AI is still entirely theoretical with no practical examples in use today, that doesn't mean AI researchers aren't also exploring its development. In the meantime, the best examples of ASI might be from science fiction, such as HAL, the superhuman, rogue computer assistant in *2001: A Space Odyssey*.

Machine Learning



What is Machine Learning? www.wooclap.com/FIZGEH

Machine Learning is a subset of Artificial Intelligence.

Machine Learning can be defined as the machine general ability to solve intelligent tasks by learning from experience/data without being explicitly programmed.

Deep Learning

If deep learning is a subset of machine learning, how do they differ?

Deep learning is actually comprised of neural networks. "Deep" in deep learning refers to a neural network comprised of more than three layers. Deep learning distinguishes itself from classical machine learning by the type of data that it works with and the methods in which it learns.

Machine learning algorithms leverage structured, labeled data to make predictions—meaning that specific features are defined from the input data for the model and organized into tables. This doesn't necessarily mean that it doesn't use unstructured data; it just means that if it does, it generally goes through some pre-processing to organize it into a structured format.

Deep learning eliminates some of data pre-processing that is typically involved with machine learning. These algorithms can ingest and process unstructured data, like text and images, and it automates feature extraction, removing some of the dependency on human experts. For example, let's say that we had a set of photos of different pets, and we wanted to categorize by "cat", "dog", "hamster", etc. Deep learning algorithms can determine which features (e.g. ears) are most important to distinguish each animal from another. In machine learning, this hierarchy of features is established manually by a human expert.

Food for Thought for Next Class: what is the *learning* in Machine/Deep Learning?

Programming in Python

- You are not expected to already know Python
- However, you do need good basic programming skills to do well in this course
- You should also have a solid programming experience in Matlab from EEL 3135:
  - Check out: NumPy for Matlab users

I have uploaded a video on "Programming with Python" for those of you that have not programed with Python. This video introduces the basic syntax and main characteristics of programing in Jupyter Notebook; view it here.

Analytical work

Pre-requisite: EEL3135 Signals & Systems, EEL3850 Data Science for ECE

I have created review material for you to review as needed on:

- Calculus
- Linear Algebra
- Probability and Statistics

See Lecture 0 in our Lectures repository on GitHub.

How does a typical lecture look like?

A typical lecture will be presented in Jupyter Notebook accompanied by whiteboard illustrations and live coding or running code examples from lecture notes.

- I will publish the class notes (Jupyter notebooks) **before** every lecture
- I will share the notebook with edits after class
- I will use my iPad using the Explain Everything app as a virtual whiteboard and the classroom board
- I will share handwritten whiteboard pages after class as well

Course Homepage

In this course we will use two main *households*:

1. Canvas page: announcements, send/receive emails, participation assignments through discussion boards, question and answering on assignment issues through discussion boards.

2. GitHub Organization: I will post all lecture notes in this private organization. You will complete all assignments in a private repository and send its URL to Canvas as the assignment submission.

    - In order to receive an invite to join the organization, join the GitHub Classroom first by accepting to create a Short Assignment 0 repository.
    - **Clone (at least) the "Lectures" repository to your local machine and pull from that repository before class**
    - Be sure to download Git
    - Complete one (or a few) introductory tutorials:
        - Git bootcamp: https://help.github.com/categories/bootcamp/
        - Tutorials: https://www.atlassian.com/git/tutorials/
        - Interactive Introduction: https://try.github.io/

Course Objectives (as time allows)

Upon completion of this course, the student will be able to:

- Identify relevant real-world problems as instances of canonical machine learning problems
- Design and implement effective strategies for data preprocessing
- Explain and utilize concepts of machine learning for data science and electrical engineering
- Compare and contrast evaluation metrics
- Foresee and mitigate human-based liabilities of machine learning algorithms
- General level of competency in critical questioning and analysis
- Students will know how to make connections between different fields of machine learning

The main goal of this course is to equip the students with a machine learning mindset for successful practical implementations, in particular: understand, analyze and design an approach to work with a data science or electrical engineering problem.

Time commitment

This is an estimate:

Work	Hours/Week
Attend lectures, ask questions	2.5
Study/Read ~20 pages of lecture notes and code	3
Reading assignments	1.5
Homework exercises	2.5
Total	**9.5**

Software Required

1. **Anaconda Distribution**

- Includes Python 3.8
- It includes all libraries, modules and tools we will use: Jupyter notebooks, `NumPy`, `Matplotlib`, `SciPy`, `Pandas`, `scikit-learn`, `random`, `PyTorch`
- Download and install it before next class

Some popular libraries in Anaconda

You have 2 options to manage your packages and virtual environment/s:

1. using `pip`. System that manages Python packages.

2. using `conda`. System that manages packages that may be written in any programming language.

Since we will use Python packages, you can use either one of these systems to manage your virtual environment. Which one to use typically comes with your specific needs. I typically use `conda` and that would be sufficient for this course.

Find help for creating and managing your **virtual environments**:

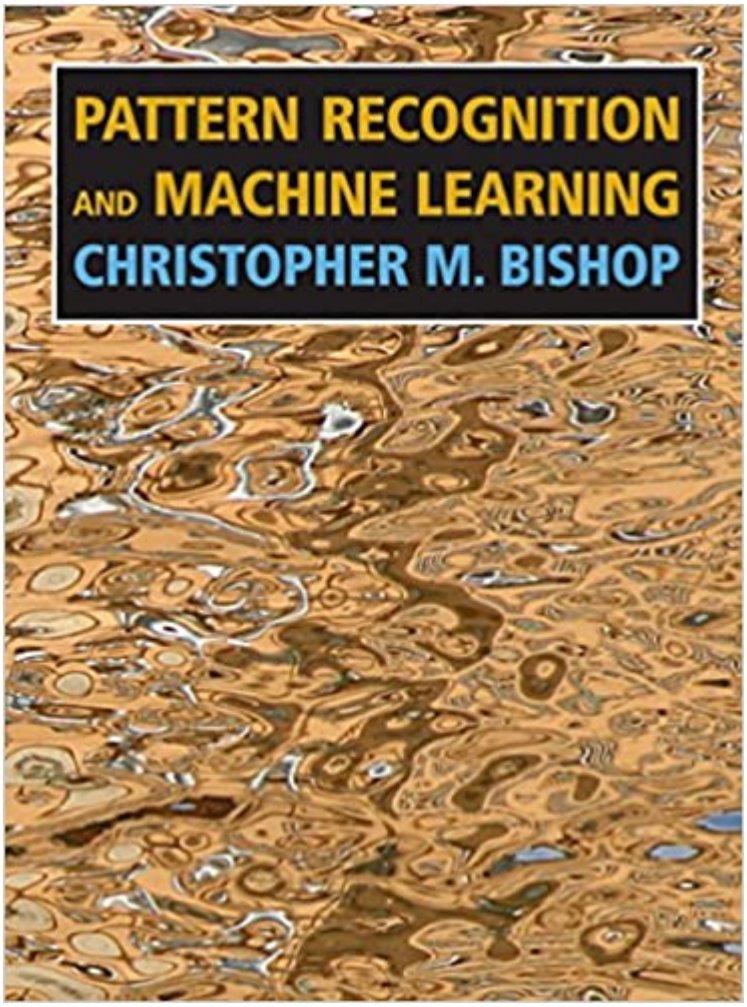
- using `conda` to manage virtual environments.
- using `pip` to manage virtual environments.

Textbooks - Required

1. **Pattern Recognition and Machine Learning**

- Author: Christopher Bishop
- Edition: 1st
- Publisher: Springer
- Year: 2006
- ISBN: 978-0-38731-073-2

An **digital version** (PDF) is freely available and is perfectly fine for this course: you can download it here



Library Course Reserves

- Additional readings will be listed in our Canvas page
- All reading materials are available on Course Reserves with both hard copy (at Marston't library) and electronic access

Access Course Reserves in Canvas, but you must open it with any browser other than Chrome.

Course Schedule

A complete course schedule can be found in our Syllabus.

Zoom Settings and Expectations

1. Choose to turn on your camera. It will help me *read your faces* and identify clarity issues.
2. I will ask you to mute your microphones through the lesson unless you have a question/comment.
3. Ask (loads of) questions! To ask a question in class, either:
    - Unmute your microphone and speak up

- Type your question in the chat box
- Raise your hand, using the *raise hand* feature (under center-low bar --> People --> Raise Hand)

4. All lectures will be **recorded** (audio and video). All videos will be available to you in our Canvas page under "Zoom Conferences" icon.

    - Students who participate with their camera engaged or utilize a profile image are agreeing to have their video or image recorded.
    - If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image.
    - Likewise, students who unmute during class and participate orally are agreeing to have their voices recorded. If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live.
    - The chat will not be recorded or shared.

Course Policies

Please read the syllabus carefully.

1. **How to get help:** office hours, email (via Canvas or at catiaspsilva@ece.ufl.edu), telephone, or Slack.

    - Slack channel: https://uf-eee4773-fall2021.slack.com/

2. **Attendance:** attendance is not required though summative and cumulative assessments. I will prepare course materials with the expectation that students will attend class synchronously. If anything schedule conflict comes up, please email me so I'm informed and we can work together for alternative arrangements.

3. **Grading:** make sure you submissions are carefully completed with clean and well documented code. Make full use of Jupyter features, such as markdown language. Individual assignments will **not** be curved. Final grades **will** be curved.

4. **Late Work:** I will accept all assignment submissions as long as solutions have not yet been released, but you will lose the **on-time** points listed in the rubric. Solutions will typically be released 1 week after the assignment is due.

5. **Make-Up Policy:** If you feel that any assignment needs to be re-graded, you must discuss this with me within 1 week of grades being posted. If approved, the entire assignment will be subject to complete evaluation. Excused absences must be consistent with university policies in the graduate catalog and require appropriate documentation

6. **Collaboration:** healthy collaboration is encouraged. If another student contributes substantially to your understanding of a problem, you should cite this student. You will not be negatively judged for citing another student.

7. **Cheating and Plagiarism:** you are expected to submit your own work. If you are suspected of dishonest academic activity, I will invite you to discuss it further in private. Academic dishonesty will likely result in grade reduction, with severity depending on the nature of the dishonest activity. I am obligated to report on academic misconduct with a letter to the department, college and/or university leadership. Repeat offenses will be treated with significantly greater severity.

Grading

Grading will be based on:

Assignment	Total	Percentage Final Grade
Exams	2	20% each
Homework	~ 5	20%
Short Assignments	~ 7	10%
Participation	~5 + extra	10%
Final Project	1 (group project)	20%

Homeworks will have 2 parts: (1) quiz with analytical exercises, typically solved on paper. (2) practical problems to be implemented in Python.

Exams will be drawn from lectures and readings. Practice exams will be provided.

Short Assignments will typically consist of short problems (with shorter turnaround time) to help consolidate and retain the information learned in class.

Participation will be in the form of discussion boards participation and class participation

Final Project is a group assignment. The objective of this project is to implement an end-to-end Machine Learning/Deep Learning model using a data set collected from students in the class. The outcomes of the final project include working code and a report.

Mark your calendars!

- **Midterm Exam Date: Thursday, October 07**

Use Canvas to select all time slots that work for you: Canvas Survey

- **Final Project due: Friday, December 10 @ 11:59 PM**
- **Final Exam Date (scheduled): Wednesday, December 15 @ 3:00 PM – 5:00 PM**

Grading Scale

Percent	Grade	Grade Points
93.4 - 100	A	4.00
90.0 - 93.3	A-	3.67
86.7 - 89.9	B+	3.33
83.4 - 86.6	B	3.00
80.0 - 83.3	B-	2.67
76.7 - 79.9	C+	2.33
73.4 - 76.6	C	2.00
70.0 - 73.3	C-	1.67
66.7 - 69.9	D+	1.33
63.4 - 66.6	D	1.00
60.0 - 63.3	D-	0.67
0 - 59.9	E	0.00

Students Requiring Accommodations

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting https://disability.ufl.edu/students/get-started/.

- Please make sure you share your accommodation letter with me as soon as you have it, so we can discuss your access needs.

Course Evaluations

You are expected to provide professional and respectful feedback on the quality of instruction in this course by completing the course evaluations online via GatorEvals.

- The University is using a relatively new evaluation system, and evaluation results are now publicly available here: https://gatorevals.aa.ufl.edu/public-results/
- Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/.
- You will be notified when the evaluation period opens, and can complete evaluations through the email you receive from GatorEvals, in the Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/.

University Honesty Policy

All UF students are bound by The Honor Pledge which states:

We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment."

The Honor Code specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Commitment to a Safe and Inclusive Learning Environment

The Herbert Wertheim College of Engineering values broad diversity within our community and is committed to individual and group empowerment, inclusion, and the elimination of discrimination. It is expected that every person in this class will treat one another with dignity and respect regardless of gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture.

If you feel like your performance in class is being impacted by discrimination or harassment of any kind, please contact your instructor or any of the following:

- Your academic advisor or Graduate Program Coordinator
- Robin Bielling, Director of Human Resources, 352-392-0903, rbielling@eng.ufl.edu
- Curtis Taylor, Associate Dean of Student Affairs, 352-392-2177, taylor@eng.ufl.edu
- Toshikazu Nishida, Associate Dean of Academic Affairs, 352-392-0943, nishida@eng.ufl.edu

Software Use

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: https://registrar.ufl.edu/ferpa.html

Health and Wellness

Covid-19 Protocols:

- You are expected to wear approved face coverings at all times during class and within buildings even if you are vaccinated. Please continue to follow healthy habits, including best practices like frequent hand washing. Following these practices is our responsibility as Gators.

- If you are sick, stay home and self-quarantine. Please visit the UF Health Screen, Test & Protect website about next steps, retake the questionnaire and schedule your test for no sooner than 24 hours after your symptoms began. Please call your primary care provider if you are ill and need immediate care or the UF Student Health Care Center at 352-392-1161 (or email covid@shcc.ufl.edu) to be evaluated for testing and to receive further instructions about returning to campus. UF Health Screen, Test & Protect offers guidance when you are sick, have been exposed to someone who has tested positive or have tested positive yourself. Visit the UF Health Screen, Test & Protect website for more information.

U Matter, We Care:

Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

Counseling and Wellness Center: http://www.counseling.ufl.edu/cwc, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Sexual Discrimination, Harassment, Assault, or Violence If you or a friend has been subjected to sexual discrimination, sexual harassment, sexual assault, or violence contact the Office of Title IX Compliance, located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, title-ix@ufl.edu

Sexual Assault Recovery Services (SARS), Student Health Care Center, 392-1161.

University Police Department at 392-1111 (or 9-1-1 for emergencies), or http://www.police.ufl.edu/.

Academic Resources

E-learning technical support, 352-392-4357 (select option 2) or e-mail to Learning-support@ufl.edu. https://lss.at.ufl.edu/help.shtml.

Career Resource Center, Reitz Union, 392-1601. Career assistance and counseling. https://www.crc.ufl.edu/.

Library Support, http://cms.uflib.ufl.edu/ask. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. https://teachingcenter.ufl.edu/.

Writing Studio, 302 Tigert Hall, 846-1138. Help brainstorming, formatting, and writing papers. https://writing.ufl.edu/writing-studio/.

Student Complaints Campus: https://care.dso.ufl.edu.

On-Line Students Complaints: http://www.distance.ufl.edu/student-complaint-process.

Any Questions?

Demonstrations

Git Demonstration

Install and download Git and GitHub Desktop.

1. How to `clone` a repository

You can use **Git Bash** to clone a repo or use GitHub Desktop. **I will demonstrate how to do it using Git Bash.** But you can also use the GitHub Desktop interface.

For example, let's create "Short Assignment 0" repository and clone it.

2. Getting the latest edits from a repository - use `git pull`

To `pull` from a repository, simply call `git pull` using Git Bash.

3. How to manage files within a repo

The 3 most used Git commands are: `git pull`, `git add`, `git commit` and `git push`. You can call these commands directly on the **Git Bash** console within the cloned repository on your machine.

This should be sufficient to get you started with Git and GitHub in this course. To learn more, watch the tutorials below:

- Git bootcamp: https://help.github.com/categories/bootcamp/
- Tutorials: https://www.atlassian.com/git/tutorials/
- Interactive Introduction: https://try.github.io/

The Curious git is also a great resource.

Python Environment Demonstration

Step 1: Download and install Anaconda with Python 3.8 (default). If you are installing Anaconda for the first time, this will create a *base* environment with all the Anaconda libraries installed and **ready to run**.

Step 2: (optional) Create a new environment and install all libraries.

> *conda create --name eee4773*
>
> *conda activate eee4773*
>
> *conda install anaconda*

Step 3: (optional) Customize your environment by installing Jupyter Notebook extensions and RISE (Jupyter Notebook slideshow extensions).

> *conda install -c conda-forge jupyter_contrib_nbextensions*
>
> *conda install -c conda-forge rise*

Launching Jupyter Notebooks

Using Conda command line:

> *conda activate eee4773*
>
> *jupyter notebook*

Or launch it with Anaconda navigator interface.

To prepare for next class

1. Download and install Anaconda
2. Download and install Git
3. (Optional) Download and install GitHub Desktop
4. Join GitHub organization
5. Create your repository for Short Assignment 0. You will be automaticallt added as a collaborator. I will send you an invitation to join organization as a member, only then you will have access to important repositories, such as Lectures.
6. Clone the following repositories to your machine: "Lectures" and "Assignment-Solutions".
7. Watch the **"Programming with Python - the basics" video** and review all the pre-requisite materials I pushed to Lecture 0's folder. Follow along with the Notebook.
8. Take a look at the **Modules** page in Canvas and get familiar with a typical lecture layout: it includes readings and activities to help you study.

Any Questions?

Feel free to email me afterwards or come talk with me during office hours: Mondays 3-4 PM, Wednesdays and Thursdays 10-11 AM