**Multimodal Data Mining**

**BME 6938 Sections: 31B1**

**Class Dates: 1/08/24 - 4/25/24**

**Class Period: T| Period 7 (1:55 PM-2:45 PM)**

**R | Period 7-8 (1:55 PM – 3:50 PM)**

**Location: Keene-Flint Hall, Room 0121**

**Academic Term: Spring 2024**

1. **Instructor:**

* Peng Liu, PhD
* E-mail address: [pliu1@ufl.edu](mailto:pliu1@ufl.edu)
* Office hours: Thursday 2:45 PM – 3:50 PM

1. **Description:** (3 credit hours) – Multimodal data mining, machine learning, and data integration course using computer programming languages for multimodal biomedical data analysis, including medical images, clinical natural language processing, genomics, and other clinical data.
2. **Pre-requisites**: Foundational knowledge in MATLAB or python and computer programming is needed to be successful in this course. If you are only familiar with MATLAB, be prepared to learn Python from scratch in this course, or take Python course first.
3. **Course Objectives**:

* Understand multimodal data mining in the biomedical domain.
* Understand the concept, approaches, and limitations in analyzing different modalities of biomedical data.
* Learn to use biomedical data processing and machine learning techniques to analyze multimodal biomedical data.

1. **Contribution of course to meeting the professional component:** 3 credits of engineering topics (no design component)
2. **Class schedule**: This is also a seminar-style course including expert lectures and an NIH-style grant proposal with a demo as the final project. Each week, you will have one to two ~50-minute recorded video lectures (recommend viewing twice at least) released on *Monday at 10 AM*. The office hour will be following the in-person meeting on *Thursday from 1:55-2:45 PM* in the classroom. On each Thursday, we will host a guest lecture or a hand-one session, followed by critiques and discussion.
3. **Material and Supply Fees**: N/A
4. **Textbooks and Software Required**

* Textbooks required for this course are freely available online at the links below. The acronyms for each textbook will be referred to in Section 10 Course Outline Readings.

1. [**DM**] Data Mining: Concepts and Techniques, 3rd ed. Jiawei Han, Micheline Kamber, and Jian Pei. Morgan Kaufmann Publisher, July 2011. [[Link](https://hanj.cs.illinois.edu/bk3/bk3_slidesindex.htm)]
2. [**DL**] Deep Learning, MIT Press, Ian Goodfellow, Yoshua Bengio, Aaron Courville. [[Link](https://www.deeplearningbook.org/)]

* **Software:** MATLAB (info.apps.ufl.edu or in CSE Active Learning Lab) or Python (recommend Anaconda) (free)

1. **Recommended Reading**:

* [**PML**] Python Machine Learning, Sebastian Raschka, Packet Publisher. 3rd ed. [[Code](https://github.com/rasbt/python-machine-learning-book-3rd-edition)]
* **[D2L]** Dive into Deep Learning (interactive book with code) [[Link](https://d2l.ai/)]
* Standard cs231n. <http://cs231n.stanford.edu/>.
* Stanford cs244d. <http://cs224d.stanford.edu/>

1. **Course Outline:** tentative schedule (subject to change)

*Notation:* ***L****:**Lecture,* ***S****: Hands-on Session*, ***G****: Guest Lectures,* ***Bold****:* *In-Person Session.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Week** | **Day** | **Date** | **Sec** | **Topic** | **Reading** | **﻿Quiz/HW** | **Project** |
| Module |  |  |  | **Part 1: Multimodal** |  | |  |
| 1 | T | 9-Jan | **L0** | **Introduction & Course Logistics** |  |  | Release |
| R | 11-Jan | **L1** | **Welcome! How to Read Papers** |  |  |  |
| S1 | Python: Introduction |  | ﻿Quiz 1 release |
| 2 | T | 16-Jan | L2 | Biomedical Image Analysis |  | ﻿Quiz 1 due |  |
| R | 18-Jan | **G1** | **Goolge Colab and HiperGator Tutorial** | HPG account |  |  |
| S2 | Python: Numpy |  | ﻿Quiz 2 release | ﻿Group |
| 3 | T | 23-Jan | L3 | Biomedical Image Filtering |  | Quiz 2 due |  |
| R | 25-Jan | **G2** | **Image Processing and Machine Learning on HPG-AI** ***(*Yunchao Yang)** |  | HW1 release |  |
| S3 | Python: Pandas |  | Quiz 3 release |  |
| 4 | T | 30-Jan | L4 | Edge Detection |  | ﻿Quiz 3 due |  |
| R | 1-Feb | **G3** | **GitHub Tutorial and Demo (Hao Ye)** |  | HW1 due |  |
| L5 | Morphological Operations |  | Quiz 4 release |  |
|  |  |
| 5 | T | 6-Feb | L6 | Neuroimage Analysis (Joseph Gullett) |  | Quiz 4 due |  |
| R | 8-Feb | **S4** | **Neuroimage Hands-on Session** | [Install FSL](https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT) |  |  |
|  | ﻿Project teamwork and discussion |  | No Quiz |  |
| 6 | T | 13-Feb | L7 | Natural Language Processing (Yonghui Wu) |  | HW2 release |  |
| R | 15-Feb | **G4** | **NLP Hands-on Session (Borui Zhang)** |  |  |  |
| S5 | NLTK Toolbox Tutorial |  | Quiz 6 release |  |
| 7 | T | 20-Feb | L8 | Genomic Medicine (Yan Gong) |  | Quiz 6 due | Milestone 1: Topic |
| R | 22-Feb | **S6** | **GWAS Hands-On (Yan Gong)** | [Install PLink](https://zzz.bwh.harvard.edu/plink/) | HW2 due |
|  | ﻿Project teamwork and discussion |  |  |  |
|  |  |  |  | **Part 2: Data Mining** | | |  |
|  | T | 27-Feb | L9 | Machine Learning (DM 8.1) | |  |  |
| 8 | R | 29-Feb | **L10A** | **Feature Engineering** | |  |  |
|  | L10B | Evaluation (DM 8.5) | | Quiz 8 release |  |
| 9 | T | 5-Mar | L11 | Supervised Learning: KNN, Naïve Bayes, SVM (DM 8.3, 9.3, 9.5) | | Quiz 8 due | Milestone 2: Specific Aims |
| R | 7-Mar | **S7** | **SVM Hands-on** | | HW3 release |
| L12 | Decision Tree, Neural Networks (DM 8.2, 8.6) | | Quiz 9 release |
|  |  | 16-Mar |  | Spring Break (no class) | |  |  |
|  | T | 19-Mar | L13 | Unsupervised Learning: Partition Clustering (DM 10) | | Quiz 9 due |  |
| 10 | R | 21-Mar | **S8** | **Partition Clustering in Python** | | HW3 due |  |
|  | L14 | Unsupervised Learning: Hierarchical Clustering (DM 10) | | Quiz 10 release |  |
| 11 | T | 26-Mar | L15 | Deep Learning & CNN (DL 9, PML 15) | | Quiz 10 due | Milestone 3: Research Strategy |
| R | 28-Mar | **S9** | **PyTorch, Tensorflow, Keras, and CNN** | |  |
|  | **﻿Project teamwork and discussion** | | Quiz 11 release |
| 12 | T | 2-Apr | L16 | Recurrent Neural Networks (RNN) (DL 10, PML 16) | | Quiz 11 due |  |
| R | 4-Apr | **G5** | **AI for Nursing EHR** **(Tamara Marciera & Ragnhildur Bjarnadottir)** | |  |  |
| S10 | RNN Programming Tutorial | | Quiz 12 release |  |
| 13 | T | 9-Apr | L17 | Generative Adversarial Networks (GAN) (DL 11) | | Quiz 12 due | Milestone 4: Preliminary Results |
| R | 11-Apr | **S11A** | **Diffusion Generative Models** | |  |
| S11B | ﻿GAN Programming Tutorial | | Quiz 13 release |
| 14 | T | 16-Apr |  | ﻿Project teamwork and discussion | | Quiz 13 due |  |
|  | R | 18-Apr | **L18** | **﻿Project Presentation & Demo -A** | |  |  |
|  |  |  |  |  |
| 15 | T | 23-Apr | **L19** | **﻿﻿Project Presentation & Demo -B & Closing Remarks** | | Report due 4/28 |  |

1. **Attendance Policy, Class Expectations, and Make-Up Policy.**

**Attendance**:

1. View the recorded lecture videos (recommended twice) on time.
2. Attend on-time attendance to in-person sessions.
3. Attend office hours if having questions.

**Expectations:**

* Schedule video lecture “class times” for yourself.
* Complete the course activities (lecture videos, quizzes, homework) on time.
* Every week’s module will be available every **Monday at 10 AM** via Canvas.
* Quizzes are due **the coming** **Monday by 11.59 PM.**
* Homework assignments 1, 2, and 3 are due **the coming** **Saturday by 11.59 PM.**
* Write your code. Do not copy code from others (Code plagiarism will be checked)
* Practice implementing newly learned concepts will make you learn better.
* Think creatively for final projects.
* Better late than never.
* Ask for help if you need it (the instructor holds office hour every Thursday)

1. **Evaluation of Grades**:

|  |  |
| --- | --- |
| **Assignments** | **Percentage of Final Grade** |
| Quiz | 20% |
| Programming Assignments | 35% |
| Admin + Attendance | 5% |
| Final Project | 40 % |
| Total | 100% |

1. **Quiz**

Quizzes contribute 20 points to your grade. Quiz questions will be asked at the end of each module to help students consolidate their knowledge.

1. **Programming Assignments**

Programming assignments will contribute 35 points to your final grade. It will be Jupiter Notebook or python programming assignments.

1. **Admin & Attendance**

Administrative assignments (e.g., Introduce yourself discussion, Microsoft Teams), attendance to in-person class sessions, and evaluations at the end of each module will contribute 5 points to your final grade.

1. **Final Project**

The Final Project will contribute 40 points to your grade. This project report will be an NIH R21/R03 style research project proposal including preliminary results (6 pages single-spaced) with a one-page specific aims page (7 pages total). Students must also include a bibliography of citations referenced in the text, but this text does not count toward the 7-page document length. The font must be Arial 11 single-spaced with no more than 1-inch margins on all sides.

This proposal can focus on the student’s specific area of graduate research and will be expected to integrate two or more methods as a central feature of the research proposal and demonstrate the integration of core knowledge of multiple modalities and machine learning/deep learning in biomedical engineering. This is intended to demonstrate the student’s mastery of the conceptual and practical application of methods and theoretical content covered during the course in their research program.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Category | Requirements | | Percentage | Points |
| Report (50%) | Specific Aims: with at least 2 specific aims and 1 specified hypothesis per specific aim /  Abstract: 1 paragraph abstract of the MICCAI paper | R21: 1 page. Paper: ~0.5 pg | 10% | 4 |
| Research Strategy: Significance, Innovation, Design, and Methods sections / introduction and methodology | R21:6 pages  Paper:  8 pages | 15% | 6 |
| Preliminary Results / Results & Analysis | 15% | 6 |
| Bibliography | | 5% | 2 |
| Language & Format | | 5% | 2 |
| Code (10%) | The program runs well and produces the expected results | | 10% | 4 |
| Present (30%) | Final Project Presentation | | 20% | 8 |
| Demo: Live demo of applying your system/method | | 10% | 4 |
| Milestone (10%) | Milestone 1: Topic | | 2.5% | 1 |
| Milestone 2: Specific Aims | | 2.5% | 1 |
| Milestone 3: Research Strategy | | 2.5% | 1 |
| Milestone 4: Preliminary Results | | 2.5% | 1 |
| Award | (Extra Credit) Best and runner-up in Oral Presentation | | +5%/+2.5% | +2/+1 |
| Score | Total Points | | (100+5)% | 40+2 |

1. **Grading Scale**:

|  |  |  |
| --- | --- | --- |
| **Points** | **Grade** | **Grade Points** |
| 93.00 - 100.00 | A | 4.00 |
| 90.00 – 92.99 | A- | 3.67 |
| 87.00 – 89.99 | B+ | 3.33 |
| 83.00 – 86.99 | B | 3.00 |
| 80.00 – 82.99 | B- | 2.67 |
| 77.00 – 79.99 | C+ | 2.33 |
| 73.00 – 76.99 | C | 2.00 |
| 70.00 – 72.99 | C- | 1.67 |
| 67.00 – 69.99 | D+ | 1.33 |
| 63.00 – 66.99 | D | 1.00 |
| 60.00 - 62.99 | D- | 0.67 |
| 0 - 59.99 | E | 0.00 |

More information on UF grading policy may be found at: <http://gradcatalog.ufl.edu/content.php?catoid=10&navoid=2020#grades>

Late Policy:

* This applies to quizzes, aims page, and project milestone reports.
* Does not apply to modality presentations, project final presentations, final project reports, and final project code.
* The first-time late submission will be given a warning only. From the second time on, (number of late minutes \* 0.1 points) will be deducted from the points of the late assignment, e.g., 30 min late = 30\*0.1=3 points if you did not make any mistakes.

1. **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/>. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

1. **Course Evaluation**

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback professionally and respectfully is available at <https://gatorevals.aa.ufl.edu/students/>. Students will be notified when the evaluation period opens and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via <https://ufl.bluera.com/ufl/>. Summaries of course evaluation results are available to students at <https://gatorevals.aa.ufl.edu/public-results/>.

1. **University Honesty Policy**

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 (<https://sccr.dso.ufl.edu/policies/student-honor-code-student-conduct-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.

**Honor Code**

* If you turn in someone else’s work as if it were your own, you are guilty of cheating. This includes homework, codes, projects, quizzes, thinking critiques, presentation slides (without citation or proper credit giving), and any required course turn-in material.
* You are also guilty if you knowingly aid in cheating.
* Software will be used to compare your submitted work to others.
* However, it is okay to discuss with other classmates about homework, idea critiques, and group projects (obviously, okay to work with a project partner). But everyone must turn in their original work.
* Do not post your work on public repositories like GitHub (private repositories are fine)
* If we catch you cheating, you will get negative points on the assignment: It is better to not do the work than to cheat!  If it is a midterm exam, final exam, or final project, you get an E (fail) in the class. All cases of cheating will be reported to the office of student conduct.

1. **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](mailto:rbielling@eng.ufl.edu)
* Curtis Taylor, Associate Dean of Student Affairs, 352-392-2177, [taylor@eng.ufl.edu](mailto:taylor@eng.ufl.edu)
* Toshikazu Nishida, Associate Dean of Academic Affairs, 352-392-0943, [nishida@eng.ufl.edu](mailto:nishida@eng.ufl.edu)

1. **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.

1. **Student Privacy**

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

1. **Campus Resources**

*Health and Wellness*

**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](mailto: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**](https://titleix.ufl.edu/), located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, [title-ix@ufl.edu](mailto:title-ix@ufl.edu)

**Sexual Assault Recovery Services (SARS)**

Student Health Care Center, 392-1161.

**University Police Department**at392-1111 (or 9-1-1 for emergencies), or <http://www.police.ufl.edu/>.

*Academic Resources*

**E-learning technical suppor***t*, 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](https://care.dso.ufl.edu/).

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