Course Content

AI in Biology

Summary: Course website for the Spring 2022 edition of Zoo4926 (section 4G55) / Zoo6927 (section 5F55), AI in Biology. Covering applications of AI in Biology.

Edit in GitHub

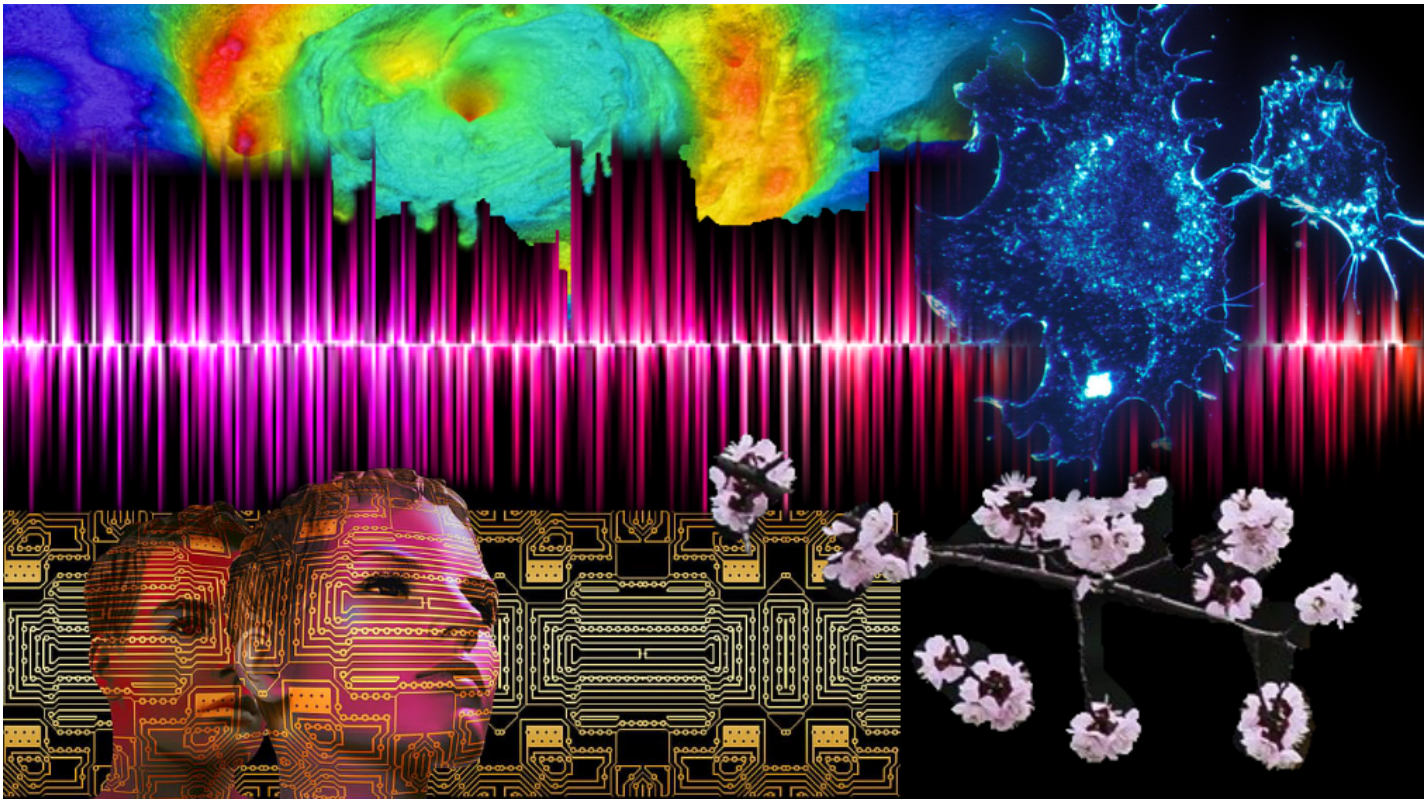
Course Description

Examines how AI has rapidly become ubiquitous in daily life and been applied to diverse areas of Biology. Focuses on machine learning approaches as well as deep learning methods, including transformers. Covers machine learning methods for tabular data, computer vision, transfer learning, natural language processing, and transformer-based architectures. Classes typically applied coding with Jupyter Notebooks on HiPerGator. Prior Python coding experience required.

Expanded Description

Artificial Intelligence (AI) as a field of research has existed since at least the 1950s. After initial enthusiasm, the gains of early years slowed and AI entered what has been referred to as an AI winter. Modern computing hardware, rapid growth in data collection and availability, and advances in algorithms have renewed interest in AI and revolutionized the field. AI is rapidly becoming ubiquitous in daily life and in diverse academic fields. This course will examine the applications of AI with particular focus on applications in biology. We will address the topics of what AI is, how intelligent computers really are and may become, where limitations still exist, and how AI technologies can be used to advance biological research.

The course will attempt to provide sufficient background and foundations so that students understand AI algorithms at a conceptual level, but will not focus on the mathematical details. This is not a computer science or mathematics course.



Classes will have some lecture, though most classes will consist of live coding demos and hands-on exercises.

Instructor

Matt Gitzendanner
Email: magitz@ufl.edu (mailto:magitz@ufl.edu)
Office: Dickinson Hall, stop at front desk and they will call me

About: Dr Gitzendanner's background is in plant evolutionary genetics and genomics where he uses genetic tools to study the conservation, evolution, and diversity of plants. The field is generally computationally intensive, and Matt has worked for 10 years training users how to use HiPerGator and other high-performance computing systems to do the amazing research that is done across the University of Florida campus.

The initial version of this course, taught in Spring 2021, was co-developed with Brian Stucky.

In addition to this course, Matt teaches

- Computational Tools for Research in Biology (BSC4452/6451)
- Frontiers in AI (EGN1935, Fall 2021. Will likely be offered in Fall 2022 as an IDS class!)
- Practicum AI
- Research Computing Trainings
- NVIDIA DLI Fundamentals of Deep Learning certified instructor

Matt enjoys spending time outdoors, hiking, backpacking and kayaking.

Prerequisites

BSC4452 or BSC6451 or BSC2891 or permission of instructor based on demonstration of prior Python programming experience. (These formal prerequisites will be effective once the course is officially listed with it's own catalog number).

Computer programming

The course assumes a **basic understanding of computer programming** in general, and Python in particular.

If you have not taken a programming course or are new to Python, there are several LinkedIn Learning courses that will give you sufficient background to be ready for this course (these are free for UF Students):

> **Note:** Note, you do not need to do all of these. Any **one** would get you at a good place to start the semester. We will review the basics of Python in the first weeks.

- Programming Foundations: Fundamentals
  - This course is best for people with no coding experience.
  - Most of the hands-on examples are in Python
- Python Essential Training
  - A good review or introduction for people who know some programming but not Python
- Learning Python
  - Another option for those with some coding experience

Math

You should have a **general understanding of probability and statistics** at the level of a first applied statistics course.

Knowledge of basic calculus and, to a lesser extent, linear algebra, can be helpful. We won't focus on the math, but having a conceptual understanding of derivatives, function optimization, and matrix math will be useful.

If you are unsure, contact the instructor.

Meeting Times

- **Mon, Wed, Fri from 1:55pm - 2:45pm in Bartram 211**

  ⚠ **Important:** You should make every effort to attend class synchronously. While I will record class sessions, **during class is the best opportunity to ask questions and get help from the instructors and others in the class.**

- I understand that some students will need to miss classes sometimes. That is fine and I will do my best to help you catch up, but regular attendance is the best way to learn.

Help Session Times

I am happy to meet in-person or via Zoom. The Zoom link will be in Canvas

- Tuesdays from 2:30 pm to 3:30 pm
- Wednesdays from 11:00am to noon
- Email Matt (mailto:magitz@ufl.edu?subject=AI%20in%20Biology%20help%20request) to setup a different time

☑ **Tip:** Coding is not always easy. Simple solutions are not always obvious. There will be some frustration.
I expect that you will need help. You should expect that you will need help.
I want to help you! I cannot always help if you do not ask for the help you need. ***Please ask for help!***

Course Textbooks

While we will not use any one text for the course, we will use sections of these books and other free resources. All will be free online resources.

- 

Python Data Science Handbook by Jake VanderPlas

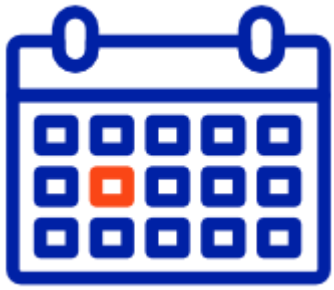
- 

Dive into Deep Learning by Aston Zhang, Zachary Lipton, Mu Li and Alexander Smola

---

Course Calendar

> **Important:** This is subject to change, please check back frequently.



For readings, there may be links to pages with my notes and additional explanations on the content from the texts.

Week	Reading/Assignment	Topic
	Week 1: Intro and Python review	
1 (Week_01.html)	**Not required** reading, but a fun intro to AI: People's Guide to AI by Mimi Ọnụọha and Mother Cyborg (Diana Nucera)	Course introduction (slides)
1 (Week_01.html)	Take the HiPerGator Account training	Brief intro and history of AI slides * Origins of AI as an academic discipline. * A repeating pattern: major hype and enthusiasm followed by an AI "winter". * Where are we now? * Constant need to question!

Week	Reading/Assignment	Topic
	Week 2: More Python review, including Pandas. Git and Github.com	
2 (Week_02.html)	 Ch 1 of PDSH: IPython: Beyond Normal Python	Introduction to Jupyter (/jupyter_intro.html) and the notebook Introduction to Jupyter
2 (Week_02.html)		Introduction to Python
2 (Week_02.html)		Git and Github Finishing Introduction to Python
	Week 3: Git and Github.com and NumPy	
3 (Week_03.html)		MLK: No Class
3 (Week_03.html)	 Ch 2 of PDSH: Introduction to NumPy	Introduction to NumPy
3 (Week_03.html)	 Ch 3 of PDSH: Data Manipulation with Pandas	Introduction to Pandas Data Visualization in Pandas

Week	Reading/Assignment	Topic
	Week 4: Machine Leaning Introduction	
4 (Week_04.html)	 Section 5.01 of PDSH: What is Machine Learning?	What is Machine Learning slides Introducing Scikit-Learn
4 (Week_04.html)	 Section 5.02 of PDSH: Introducing Scikit-Learn	Introducing Scikit-Learn
4 (Week_04.html)	 Section 5.03 of PDSH: Hyperparameters and Model Validation	Bias/variance tradeoff, model validation, cross-validation, and hyperparameters See also slides
	Week 5: Bias-variance trade-off, linear regression	
5 (Week_05.html)		Bias/variance tradeoff, model validation, cross-validation, and hyperparameters See also slides
5 (Week_05.html)	Skim  Section 5.04 of PDSH: Feature Engineering	Bias/variance tradeoff, model validation, cross-validation, and hyperparameters See also slides

Week	Reading/Assignment	Topic
5 (Week_05.html)	 Section 5.06 of PDSH: In Depth: Linear Regression Problem set 3 available, due Wednesday, February 16	Linear Regression–Lasso and Ridge Regression
	Week 6: Classification, logistic regression, SVMs	
6 (Week_06.html)		Hands-on data analysis, problem set help
6 (Week_06.html)		Classification, logistic regression
6 (Week_06.html)	 Section 5.07 of PDSH: In-Depth: Support Vector Machines)	Support vector machines
	Week 7: SVMs continued, Random Forests	
7 (Week_07.html)		Hands-on SVMs Work through either: - Tutorial: image classification with scikit-learn Remote Sensed Hyperspectral Image Classification With The Extended Morphological Profiles and Support Vector Machines

Week	Reading/Assignment	Topic
7 (Week_07.html)	 Section 5.08 of PDSH: Decision Trees and Random Forests) Problem set 4 available, due Friday, February 28	Decision Trees and Random Forests
7 (Week_07.html)		Random Forests and Ensemble Methods
	Week 8: More Ensemble methods, XGBoost. Intro to neural networks	
8 (Week_08.html)	Random Forests and Ensemble Methods	
8 (Week_08.html)		Intro to Artificial Neural Networks: Lecture 05 and Notebook
8 (Week_08.html)		Multi-Layer Neural Networks
	Week 9: Computer vision with Convolutional Neural Networks	
9 (Week_09.html)		Convolutional Neural Networks and Lect_06
9 (Week_09.html)		Convolutional Neural Networks
9 (Week_09.html)		Transfer Learning

Week	Reading/Assignment	Topic
	Week 10: Time series analysis with RNNs	
10 (Week_10.html)		Time Series Analysis with RNNs: work through this tutorial (data and notebook located at `blue_zoo4926/share/SoyBean_TS/` or work on assignment 5.
10 (Week_10.html)		Time Series Analysis with RNNs
10 (Week_10.html)		Time Series Analysis with RNNs
	Week 11: Natural language processing and the rise of Transformers	
11 (Week_11.html)		Natural Language Processing
11 (Week_11.html)		Natural Language Processing
11 (Week_11.html)		Mamba and Custom Kernels (/vit_mamba_setup.html) Vision Transformers intro and option 1 Vision Transformers on Casava diseases
	Week 12: Transformer architecture in depth	
12 (Week_12.html)		Transformers
12 (Week_12.html)		AlphaFold background

Week	Reading/Assignment	Topic
12 (Week_12.html)		Transformers
	Week 13: Transformers and what's new in AI?	
13 (Week_13.html)		Transformers
13 (Week_13.html)		Topic TBD
13 (Week_13.html)		Topic TBD
	Week 14: Finish up projects	
14 (Week_14.html)		Project time
14 (Week_14.html)		Project time
14 (Week_14.html)		Project time
	Week 15: Present projects	
15 (Week_15.html)		Project Presentations
15 (Week_15.html)		Project Presentations

Software and Hardware

Participants will need a computer with internet connection, webcam and microphone for all classes.

Several free/open source software packages will be used throughout the course, and students will be required to install some of these. Students will use a (free) Research Computing account to access HiPerGator for coursework. Students will be required to apply for a (free) Github.com account for coursework.

If you have technical difficulties with Canvas, please contact the UF Helpdesk at:

- http://helpdesk.ufl.edu
- (352) 392-HELP (4357)
- Walk-in: HUB 132

Any requests for make-ups due to technical issues should be accompanied by the ticket number received from the Help Desk when the problem was reported to them. The ticket number will document the time and date of the problem. Please e-mail the instructor within 24 hours of the technical difficulty if you wish to request a make-up.

All faculty, staff and student 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.

Grading

Assignment Values

See also the List of Graded Work page (Grading.html).

Item	Undergraduate Points	Graduate Points
Problem Sets	5 @ 20 points each: 100 points (77%)	6 @ 30 points each: 180 points (75%)
Class Project	20 points (15%)	40 points (17%)
Project presentation	10 points (8%)	20 points (8%)

Undergraduates will have 5 problem sets worth 20 points each and slightly less weighting on the project. Graduates will have one extra question for each problem set (making each worth 30 points), one extra problem set and slightly higher weighting on the project.

Grading in this class is consistent with UF policies available at: https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/

Grade Disputes and Revisions

Grading Scale and GPA Equivalent

Student Learning Outcomes

1. Summarize major events in the history of AI from the 1950s to present.
2. Proficiently launching Jupyter Notebooks on HiPerGator, requesting appropriate resources for the task.
3. Analyze and visualize complex tabular data with NumPy, Pandas, and matplotlib
4. Calculate linear regression using machine learning approaches with Scikit-learn
5. Explain the bias/variance tradeoff
6. Assess ML/AI models, conduct cross-validation and tune hyperparameters
7. Apply support vector machines, decision trees, random forests and ensemble methods to analyze data
8. Code a simple single-neuron perceptron from scratch
9. Code multi-layer neural networks using Keras/Tensorflow
10. Conduct computer vision experiments using convolutional neural networks
11. Conduct time series data with recurrent neural networks
12. Apply transfer learning
13. Identify key concepts in natural language processing, including tokenization, word embeddings, and the rise of transformer architectures.
14. Apply transformers to computer vision tasks
15. Conduct protein folding folding prediction using transformer architectures

---

Course Policies

Class Attendance and Makeup Policy

Requirements for class attendance and makeup assignments, and other work in this course are consistent with university policies that can be found in the online catalog at: https://catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/

In general, I do not take attendance. You are all adults and I assume you are taking the class the learn. **The best way to learn is to regularly attend class.** I are sure students will miss class for various reasons. I am happy to help you catch up. If you regularly miss class and fall behind, I may ask that you hold questions on content you have missed until after class, or ask that you coordinate a time to go over the content. I will make every effort to record and post all classes to help those that miss classes.

Assignment Policy

Assignment dates will be announced at least one week in advance and students will have at least three days to complete the assignment. Each assignment will clearly state if it is an individual or group assignment. Individual assignments must be the student's own work, completed without the assistance of others.

All assignments are "open book, open internet", you may use whatever resources you desire to complete the assignment. Though only assignments specifically noted as group assignments should be worked on with other people.

Makeup and Late policy

Please notify the instructor of circumstances that lead to late work or missed classes. I will generally work with you and accept late work. All assignments are designed for both your own learning and my assessment of your efforts. Much of the course builds on previous sections and falling behind on assignments will make it harder to keep up. If you need help, please ask! **My goal is for all students to learn the material** and I understand that some students will need more help than others. The grade is based on the end product, not the amount of time and help needed to get there.

Students Requiring Accommodations

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://disability.ufl.edu/students/get-started/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

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 in a professional and respectful manner is available at 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 ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at gatorevals.aa.ufl.edu/public-results/.

Class Demeanor and Netiquette

Students are expected to be in class on time and behave in a manner that is respectful to the instructors and to fellow students.

Opinions held by other students should be respected in discussion, and conversations that do not contribute to the discussion should be held at minimum, if at all.

Students should be working on course content during class.

Discussion Boards

The GitHub discussion boards can be used to ask for and provide help by all. Students should be supportive and considerate of others at all times. Rude or inappropriate comments will be removed and the poster will be warned.

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 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

Health and Wellness

Inclusive Learning Environment

This course embraces the University of Florida's Non-Discrimination Policy, which reads:

> The University shall actively promote equal opportunity policies and practices conforming to laws against discrimination. The University is committed to nondiscrimination with respect to race, creed, color, religion, age, disability, sex, sexual orientation, gender identity and expression, marital status, national origin, political opinions or affiliations, genetic information and veteran status as protected under the Vietnam Era Veterans' Readjustment Assistance Act.

If you have questions or concerns about your rights and responsibilities for inclusive learning environment, please see the instructor or refer to the Office of Multicultural & Diversity Affairs website: http://multicultural.ufl.edu.

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

Statement Regarding Course Recording

Our class sessions may be audio visually recorded for students in the class to refer back to and for use of enrolled students who are unable to attend live. 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, keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate verbally are agreeing to have their voices recorded. If

you are unwilling to consent to have your voice recorded during class, you will need to keep your mute button activated. As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited.

Additional UF Policies and Resources

Site last generated: Dec 8, 2022