Course Content

Introduction to AI and Python																▼

Machine Learning																		▼

Artificial Intelligence																	▼

Additional Resources																	▼

AI in Biology

Summary: Course website for the Spring 2024 edition of AI in Biology (BSC4892/6895). Covering applications of AI in Biology.

[Edit in GitHub ↗]

Course Description

Examines how Artificial Intelligence 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 are 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.
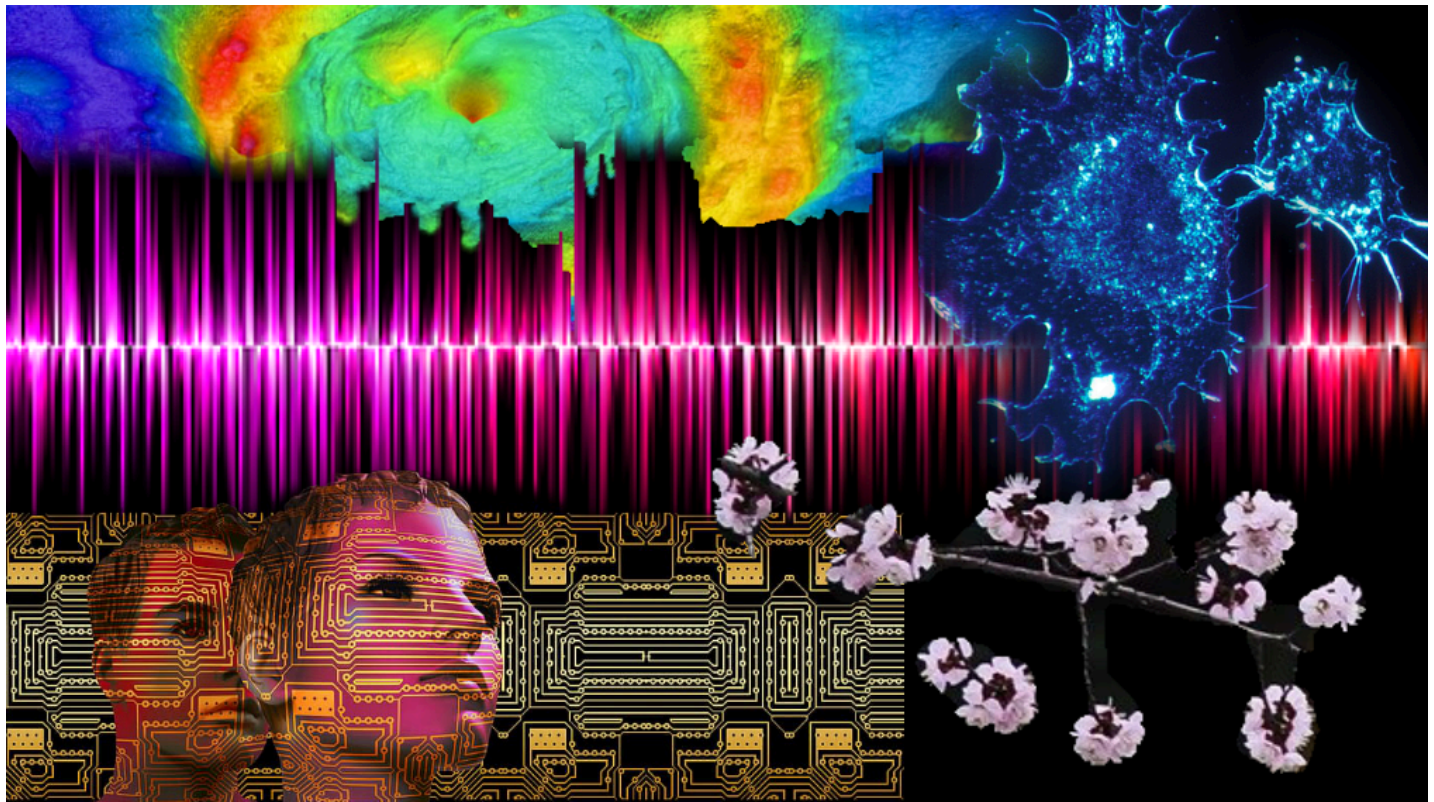
AI Content

Build-AI: Evaluate & Create AI: Higher-order thinking skills (e.g., evaluate, appraise, predict, design) with AI applications. AI course content is over 50%.

This course accomplishes the AI Designation objectives of the subject areas listed above. The course explores the application of AI to biological data. Students build AI systems using Python code and evaluate the performance of these on data.

Instructors

Matt Gitzendanner

- **Email:** magitz@ufl.edu (mailto:magitz@ufl.edu)

- **Office:** Dickinson Hall, stop at front desk and they will call me

- **Phone:** 352-273-1960

- **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.

Arthur Porto

- **Email:** arthur.porto@ufl.edu (mailto:arthur.porto@ufl.edu)

- **Office:** Dickinson Hall, stop at front desk and they will call me

- **Phone:** 352-273-1939

- **About:** Dr Porto is a computational biologist with strong focus on applied computer vision. You can find more about his research at the BioVisionLab ⬀ website.

Note: The initial version of this course, taught in Spring 2021, was co-developed with Brian Stucky ⬀. Previous versions of the course are archived in GitHub branches in the website's repository ⬀.

Prerequisites

BSC4452 or BSC6451 or BSC2891 or permission of instructor based on demonstration of prior Python programming experience.

Computer programming

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

If you need a quick refresher, there are several LinkedIn Learning courses that will give you sufficient background to be ready for this course (these are free for UF Students):

> ℹ **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 we 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 we will do our best to help you catch up, but regular attendance is the best way to learn.

Help Session Times

We are happy to meet in-person or via Zoom.

- Matt's help time schedule:

  - Wednesdays from 10:00am to 11:00

  - Book a time that works for you ↗

  - Email Matt (mailto:magitz@ufl.edu?
    subject=AI%20in%20Biology%20help%20request) to setup a
    different time

- Arthur's help time schedule:

  - Fridays from 10:00am to 11:00

  - Email Arthur (mailto:arthur.porto@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.
> **We expect that you will need help. You should expect that you will need help.**
> We want to help you! We 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 ↗

- We likely won't use this much, but another good resource–especially if you want more of the mathematical background–is: Dive into Deep Learning by Aston Zhang, Zachary Lipton, Mu Li and Alexander Smola ↗

Student Learning Outcomes

Note: For each SLO listed, topics are addressed in lectures and hands-on, interactive coding exercises in class throughout the semester. Assessment of SLOs is noted in the corresponding assignments in the calendar.

Know-AI: Know & Understand

- **SLO1.** Identify, describe, and/or explain the components, requirements, and/or characteristics of AI.

- **SLO2.** Recognize, identify, describe, define, and/or explain applications of AI in multiple domains.

Use-AI: Use & Apply

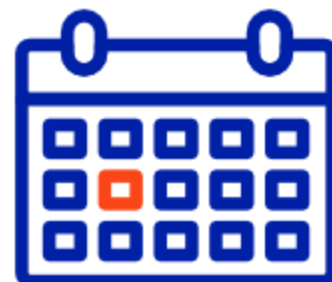- **SLO3.** Select and/or utilize AI tools and techniques appropriate to a specific context and application.

Build-AI: Evaluate & Create

- **SLO5.** Assess the context-specific value or quality of AI tools and applications.

- **SLO6.** Conceptualize and/or develop tools, hardware, data, and/or algorithms utilized in AI solutions.

———

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	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 1	Intro and Python review	3 hours	**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 ↗) 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! GitHub setup

Week	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 2	**Python review, including Pandas. AI for Coding.**	3 hours	Ch 1 of PDSH: IPython: Beyond Normal Python ↗ Introduction to Jupyter (/jupyter_intro.html) and the notebook Introduction to Jupyter ↗ **GitHub Assignment due** (SLO6.) Introduction to Python ↗ Intro Python and Coding Assistants ↗
Week 3	**More AI for Coding, and NumPy**	3 hours	Intro Python and Coding Assistants ↗ Ch 2 of PDSH: Introduction to NumPy ↗ Introduction to NumPy ↗ Ch 3 of PDSH: Data Manipulation with Pandas ↗ Introduction to Pandas ↗ Data Visualization in Pandas ↗

Week	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 4	Machine Leaning Introduction	3 hours	Section 5.01 of PDSH: What is Machine Learning? ↗ What is Machine Learning slides ↗ Introducing Scikit-Learn ↗ **Problem set 1 is due** Basic Python coding assessment (SLO3, SLO6.) Section 5.02 of PDSH: Introducing Scikit-Learn ↗ Introducing Scikit-Learn ↗ Section 5.03 of PDSH: Hyperparameters and Model Validation ↗ Bias/variance tradeoff, model validation, cross-validation, and hyperparameters ↗ See also slides ↗

Week	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 5	**Bias-variance trade-off, linear regression**	3 hours	Bias/variance tradeoff, model validation, cross-validation, and hyperparameters ↗ See also slides ↗ Skim Section 5.04 of PDSH: Feature Engineering ↗ Bias/variance tradeoff, model validation, cross-validation, and hyperparameters ↗ See also slides ↗ Section 5.06 of PDSH: In Depth: Linear Regression ↗ Linear Regression–Lasso and Ridge Regression ↗
Week 6	**Classification, logistic regression, SVMs**	3 hours	Hands-on data analysis, problem set help Classification, logistic regression ↗ **Problem set 2 is due** Linear and Logistic regression (SLO1, SLO2, SLO3, SLO5, SLO6) Section 5.07 of PDSH: In-Depth: Support Vector Machines ↗ Support vector machines ↗

Week	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 7	**SVMs continued, Random Forests**	3 hours	Hands-on SVMs Work through one of these: - Tutorial: image classification with scikit-learn ↗ - Remote Sensed Hyperspectral Image Classification With The Extended Morphological Profiles and Support Vector Machines ↗ - Image Classification Using Machine Learning-Support Vector Machine(SVM) ↗ Section 5.08 of PDSH: Decision Trees and Random Forests ↗ Decision Trees ↗ and Random Forests **Problem set 3 is due** Support Vector Machines (SLO1, SLO2, SLO3, SLO5, SLO6) Problem set 4 available, due Friday, March 8 Random Forests and Ensemble Methods ↗

Week	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 8	**More Ensemble methods, XGBoost. Clustering Algorithms.**	3 hours	Papers with[out] code: Kumar *et al.* Sustainability 2022, 14(21), 13998; https://doi.org/10.3390/su142113998 🔗 Discuss and implement the paper. Section 5.11 of PDSH: k-Means Clustering 🔗 Clustering algorithms – K-means 🔗 Section 5.11 of PDSH: k-Means Clustering 🔗 Clustering algorithms – K-means 🔗
Week 9	**Intro to Artificial Neural Networks**	3 hours	Intro to Artificial Neural Networks: Lecture 05 🔗 and Notebook 🔗 Multi-Layer Neural Networks 🔗 **Problem set 4 due** CLustering methods (SLO1, SLO2, SLO3, SLO5, SLO6) Convolutional Neural Networks 🔗 and Lect_06 🔗
Week 10	**Computer vision with Convolutional Neural Networks**	3 hours	More CNNs Transfer Learning 🔗 Generative AI – (e.g., Denoising, novel discovery)

Week	AI-Related Topic	# Contact Hours of AI-related Content	AI-related readings, projects, assignments, etc.
Week 11	**Natural language processing and the rise of Transformers**	3 hours	Natural Language Processing ↗ Natural Language Processing ↗ **Problem Set 5 due** CNN application (SLO1, SLO2, SLO3, SLO5, SLO6) 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**	3 hours	Transformers AlphaFold background ↗ Alphafold in Colab ↗
Week 13	**Transformers and what's new in AI?**	3 hours	Transformers Multimodal AI **Problem set 6 due (graduate students only)** Transformer exercise (SLO1, SLO2, SLO3, SLO5, SLO6) Multimodal AI
Week 14	**Finish up projects**	3 hours	Project time (team designed AI project) (SLO1, SLO2, SLO3, SLO5, SLO6)
Week 15	**Present projects**	2 hours	Project Presentations (SLO1, SLO2, SLO3, SLO5, SLO6)

Software and Hardware

- Participants will need a computer with internet connection 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 (% of total)	Graduate Points (% of total)
Problem Sets	5 @ 20 points each: 100 points (74%)	6 @ 30 points each: 180 points (73%)
GitHub Assignment	5 points (4%)	5 points (2%)
Class Project	20 points (15%)	40 points (16%)

Item	Undergraduate Points (% of total)	Graduate Points (% of total)
Project presentation	10 points (7%)	20 points (8%)
Total	**135 (100%)**	**245 (100%)**

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

By the end of the course, students will be able to:

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, we do not take attendance. You are all adults and we assume you are taking the class the learn. **The best way to learn is to regularly attend class.** We are sure students will miss class for various reasons. We are happy to help you catch up. If you regularly miss class and fall behind, we 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. We 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 instructors of circumstances that lead to late work or missed classes. We 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! **Our goal is for all students to learn the material** and we 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 who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center. Click here to get started with the Disability Resource Center ☒. 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.

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 Conduct Code specifies a number of behaviors that are in violation of this code and the possible sanctions. Click here to read the Conduct Code ↗. If you have any questions or concerns, please consult with the instructor.

Health and Wellness

Site last generated: Dec 4, 2024