Artificial Intelligence and Machine Learning Applications in the Physical Sciences

I. Course Information

AST 4930
Semester 2021 Fall
Meeting Day/Time: Tuesday 10:40 am – 11:30 am, Thursday 10:40 am – 12:35 pm
Location: BRT 0003

Instructor

Prof. Jaehan Bae
Email: jbae@ufl.edu
Office location: BRT 316
Office hours: Tuesday 2 pm – 4pm
Office hours location: BRT 316. Alternatively, we can meet remotely on zoom. If you choose the remote option, send the instructor a canvas message before Tuesday 2 pm to set up a meeting.

The preferred method for contacting the instructor outside of class and office hours is via the canvas messaging system.

Course Description

Modern science is increasingly enabled by efficient, careful, and detailed analysis of big data. This course will aim to (i) introduce students to applied AI/ML analysis techniques, (ii) apply those techniques in class to real world datasets, and (iii) ask students to apply the techniques to cutting-edge research-quality datasets via coding-based assignments and term project. The course would be Python-based and would make use of common Python packages including Scikit-learn, TensorFlow, Keras, and PyTorch. Many of these packages can be run with and without GPU support, allowing students to explore the importance (and limitations/complications) of GPU computing.

The course will be structured with 7 modules (see Section II. Schedule & Coursework). Each module will last two weeks and will both introduce a series of related topics and then ask students to apply those methods through homework assignments. Homework assignments will employ real world datasets.

Course Materials

The course will be taught based on course notes and some combination of the following books:

- *Introduction to Machine Learning with Python*, O'Reilly Media, Müller & Guido, ISBN: 9781449369415
    - This will be the main textbook of the course.
    - It provides an in-depth introduction to Machine Learning with Python including heavy use of Scikit-learn, but with a more general (i.e. less physical) focus.
    - Accompanied with a series of Jupyter notebooks that give good examples that are significantly broader in nature than those provided with the *Python Data Science Handbook* (see below).
- *Python Data Science Handbook,* O'Reilly Media, VanderPlas, ISBN: 9781491912058
    - If you are not familiar with Python this book offers a great place to start.
    - The full text of this book is freely available online at https://jakevdp.github.io/PythonDataScienceHandbook/
- *Statistics, Data Mining & Machine Learning in Astronomy*, Princeton University Press, Ivezić, Connolly, VanderPlas, & Gray, ISBN: 9780691198309
    - Part of the book (and lots of useful resources) is available online at http://www.astroml.org/
    - This book is probably (on average) a bit above the level of this course. However, the text is very useful in that it is both Python-based and it provides abundant examples about how a wide range of techniques can be applied to astrophysical survey (i.e. very large) datasets.
    - The full text of this book is available to read online at UF libraries.
- *Machine Learning techniques for Physics and Astronomy*, Princeton University Press, Acquaviva
    - This book offers probably the most relevant examples to this course as the author is an Astrophysicist. However, this book is under development at the time this course is offered and is not available for purchase. Some astrophysical examples from this book will be introduced during the course.

Statement on Materials and Supplies Fees

N/A

II. Schedule & Coursework

Weekly Course Schedule

Week	Topic
Week 1	Introduction to Python and machine learning
Week 2	Module 1: Supervised learning for classification: k-nearest neighbors and decision trees
Week 3	Module 1: Applications of supervised learning for classification
Week 4	Module 2: Evaluation metrics; cross validation; diagnostics
Week 5	Module 2: Applications of evaluation metrics; cross validation; diagnostics
Week 6	Module 3: Support vector machines; parameter optimization
Week 7	Module 3: Applications of support vector machines; parameter optimization
Week 8	Module 4: Regression methods
Week 9	Module 4: Applications of regression methods
Week 10	Module 5: Ensemble methods for classification and regression
Week 11	Module 5: Applications of ensemble methods for classification and regression
Week 12	Module 6: Clustering and dimensionality reduction
Week 13	Module 6: Applications of clustering and dimensionality reduction
Week 14	Module 7: Deep learning methods
Week 15	Module 7: Applications of deep learning methods
Week 16	Term project presentations

List of Graded Work

Work	Description	Points
Homework Assignments	Homework assignments consist of comprehension questions and mini-coding problems. One homework assignment will be assigned for each module. Homework assignments will ask students to apply the concepts and techniques from the lecture and readings, with a goal of assessing student comprehension. Homework assignment sets will be graded for accuracy. All homework assignments must be submitted through the canvas website.	50
Exams	There will be two exams during the course: a midterm exam and a final exam. The exams will be closed book and closed notes. The tests are designed to assess student comprehension of the concepts covered in the course, and will feature topics from lecture and homework assignments. Exams will be graded for accuracy.	30
Term Project & Presentation	The term project will be used for students to apply the concepts and techniques from the course to real world astronomical datasets in a more in-depth fashion. Students are welcome to consult with the instructor on the topic of their term project. The term project will consist of submitted code, a write-up, and an oral presentation to the class. The term project will be graded based on the accuracy and efficiency of the submitted code, the submitted written project description, and the final oral presentation.	20

The course canvas site will make clear all assignment dates and deadlines. Any questions about deadlines should be directed to the instructor, ideally through the course Discussion pages.

III. Grading

Statement on Attendance and Participation

Attendance and Participation:

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

- Attendance itself is not a graded component of this class. Nevertheless, critical information will be disseminated through our class meetings. Thus, students are expected to either attend class, or otherwise obtain the material discussed during class.
- All assignments are due at 5 p.m. on the assigned due date unless noted otherwise. Students must submit completed assignments via canvas in the format specified in the assignment. Late assignments will generally not be accepted unless a documented reason is provided that qualifies under UF's approved/excused absences.

Grading Scale

For information on how UF assigns grade points, visit: https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/

A	90 – 100%		C	70 – 73.99%
A-	87 – 89.99%		C-	67 – 69.99%
B+	84 – 86.99%		D+	64 – 66.99%
B	80 – 83.99%		D	60 – 63.99%
B-	77 – 79.99%		D-	57 – 59.99%
C+	74 – 76.99%		F	<56.99

IV. Using HiPerGator Resources

Students may use HiPerGator resources for homework and term projects. Students that need to access HiPerGator for course-related activities should ensure that they are properly registered in their course. Those who already have HiPerGator accounts will be added to the class group; the rest will have a temporary HiPerGator account created for them to use for the class. Students who join the class late should remind their instructor to send Research Computing a request to add them to the class. Note that class accounts will expire (and any associated data deleted) two weeks after the "Classes End" date listed for the semester: UF Catalog: Dates & Deadlines.

V. Required Policies

Students Requiring Accommodation

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://disability.ufl.edu/) 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.

UF Evaluations Process

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 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/.

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://www.dso.ufl.edu/sccr/process/student-conduct-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.

Counseling and Wellness Center

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

The Writing Studio

The writing studio is committed to helping University of Florida students meet their academic and professional goals by becoming better writers. Visit the writing studio online at

http://writing.ufl.edu/writing-studio/ or in 2215 Turlington Hall for one-on-one consultations and workshops.

Privacy Considerations for Remote/Recorded Lectures

Should there be circumstances remote/recorded lectures are given, the following policy will be applied.

When course meetings are being recorded, 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 un-mute 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.

As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited. Specifically, you may not video record, audio record, screen shot, or otherwise record any course meetings. Nor may you share any recorded material from class (legitimate course recordings, or otherwise). Uniform adherence to this policy is critical to ensuring a safe and academically engaging environment. Violations of this policy will be immediately escalated to the Dean of Student Affair's office.