Course Syllabus

Course Title
Machine Learning Pipelines Using Cloud-Based Platforms

Semester
Summer 2024

Course Number
IS597-MLC

Instructors
Kevin Trainor (trainor1@illinois.edu)
Jenna Kim (jkim682@illinois.edu)

Teaching Assistant
TBA

Regular Class Sessions
Regular class sessions will be held on _Wednesdays_ from 6:00 PM till 8:00 PM online via Zoom. We look forward to joining you there.

Optional Lab Sessions
We hold optional lab sessions _two times per week_**.** Please join us to ask a question, to discuss solutions to previous assignments, to get help with the current assignment, to discuss the final project, or just to say hello. We will work with students on a first come, first served basis.

The _first optional lab session_ of the week takes place _immediately following our regular class session_. Since class sessions typically end earlier than their scheduled two-hours, we devote the remainder of the time available to helping students. Often, we can stay beyond the two-hour period if many students need help.

On _Friday afternoons from 2:00 PM till 3:00 PM_, we will be holding an _optional online lab session using Zoom_. This represents a good opportunity to get help with your weekly assignments _before the submission deadline on Tuesday night_. Please use a headset when joining the online lab session.

Office Hours
We do not hold conventional open office hours. Students needing help with assignments are encouraged to join in the optional lab sessions (see above). Students who need to discuss confidential matters can arrange for an individual meeting with us on Zoom. If you need to arrange an individual meeting with us, please send us a request via the HelpDesk.

Contacting Instructors or TA
Please do not send requests or other messages to us using Canvas or our regular email addresses. We will not be responding to messages sent to us on those platforms.

Instead, please contact us by creating a ticket using the [Trainor HelpDesk.](#) The HelpDesk system allows you to open a ticket and correspond back-and-forth with us by sending and receiving emails to/from the Trainor HelpDesk system.

For an introduction to the HelpDesk system, please see *[About Trainor HelpDesk.](#)*

For assistance in creating a ticket related to this course (IS597 MLC-Section), please see *[IS597-MLC Section – Compose Email](#)*.

Catalog Description
This is a graduate-level data science course for students who need practical experience in applying machine learning approaches using cloud computing to address real world problems. This course provides hands-on experience in identifying problems suitable for applying machine learning techniques, designing pipeline-based solutions using Python programming, and implementing these solutions on cloud-based platforms.

Course Topics

- What is machine learning?
- Introduction to cloud computing
- Python for machine learning crash course – Part 1
- Python for machine learning crash course – Part 2
- Data pre-processing using cloud computing
- Building pipelines for machine learning
- Training models on cloud-based platforms
- Testing and performance evaluation using cloud computing
- Dimensionality reduction
- Model tuning

Prerequisites
Students must have completed IS597-PR or IS430. Students who have completed a different but equivalent Python programming course may enroll with permission of the instructor.

Credit Hours
4

Learning Outcomes
After completing this course, you should be able to:

- Identify a problem that is suitable for machine learning application.
- Understand how to use cloud computing for machine learning implementation.
- Design a pipeline-based machine learning workflow.
- Build a machine learning model for training and test the model for evaluation.

Required Texts
The following textbooks are available for free to Illinois students on the [oreilly.com site](#):

- McKinney, W. (2022). *Python for Data Analysis, 3rd Edition*. O'Reilly Media. ISBN: 978-1098104030.

- Aurélien Géron (2019). *Hands-On Machine Learning with Scikit-Learn, Keras & Tensorflow, 2nd Edition*. O'Reilly Media. ISBN: 978-1492032649.

- Sebastian Raschka & Vahid Mirjalili (2019). *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and Tensorflow 2, 3rd Edition*. O'Reilly Media. ISBN: 978-1789955750.

Help for finding and using E-Books at the O'Reilly site is available at:

- https://guides.library.illinois.edu/ebook/providers

Technology Requirements
We will be using cloud-based platforms during class sessions, lab sessions, and while completing assignments. *Students are required to bring their own laptop computers to class.* Laptops should be recent model machines running either Windows or macOS that are capable of logging into the Illinois network.

Cloud computing work in this course will take place using Amazon Web Services. We will be using the Learner Lab feature of the AWS Academy program. More details will be made available to students shortly after the beginning of the semester.

Early in the semester, we will be covering a unit called *Python for ML Crash Course*. For this unit, you will be working on your own computer using Anaconda and PyCharm Professional. Links to tutorial videos will be provided in our weekly schedule for help on downloading, installing, and configuring these tools on macOS or Windows.

Course Schedule
The schedule for this course will be available via our Weekly Schedule at:

https://courseinfo.ligent.net/2024su/_illinois/is597_mlc/index.html

The course schedule is always subject to reasonable change to account for changes in circumstances and to correct errors. When we make changes to the schedule, we will announce them via our Canvas Announcements Forum. Postings to this forum should result in you being sent an email copy of the announcement as well.

Course Elements

1. Readings

Required readings will be assigned from the textbooks listed above and from other resources that will be identified in the course schedule. Generally, readings are chosen to accompany any lecture videos or tutorial videos provided for the week. *So, you should expect to complete the readings before playing the videos.*

2. Lecture Videos

Some lectures will take place during class. Others have been recorded as videos. *If there is a recorded lecture video for you to play, it will be indicated in our Weekly Schedule. In that case, you will be expected to play these videos before our class session for that week.*

3. Tutorial Videos

We have created tutorial videos for the following use cases:

- There are a number of tutorials that explain how to do activities necessary for the course. These include installing software, reading grading rubrics, submitting assignments, and related activities.

- Many tutorials are coordinated with the exercises in your coding assignments. They represent a demonstration of the skill you will be expected to use when doing the exercise.

4. Weekly Assignments

There will be weekly assignments. As mentioned above, assignments will often be paired with tutorial videos. These should allow you to complete your assignment using the same general approach that has been demonstrated in the tutorial video.

5. Final Project

The Final Project is an individual assignment. Each student will define their own research question, find a data set that is related to the question, and build a machine learning pipeline to analyze these data by training and evaluating their model.

The Final Project assignment will include two separate submissions. The Final Project Proposal will be a planning document that will be due before the mid-point of the semester. *The Final Project Delivery* will include all parts of the implemented project and will be due at the end of the semester. Each of these submissions will be based on an instruction document and grading rubric that will provide more details of the requirements.

6. Attendance

The iSchool expects students to attend all classes except in cases of emergency. See *Student Code on Attendance:* http://studentcode.illinois.edu/article1/part5/1-501/.

7. *Participation*
   Your participation in the course is an important element of the course. Accordingly, a *significant portion of your grade for this course (15%)* will be determined by your participation. Students will earn participation credit for:

   - Introducing yourself during class in Week 1.
   - Making speaking contributions during class.
   - Active participation in a group breakout session during class.
   - Presenting for your group during debrief of a group breakout session during class.

Course Grading Policies

1. *iSchool and University Grading Policies Apply*
   Many iSchool and University Academic Policies have grade implications. Please see *iSchool and University Academic Policies* below.

2. *Careful Attention to Detail is Required*
   One important goal of this course is to train you to become a responsible information professional. The work of information professionals is highly detail oriented. Clients rely on information professionals to deliver a correct work product that conforms to stated requirements and best practices.

   When your work is graded, deductions will be made for all deviations from the assignment instructions. Some of these deductions will be made for small deviations that may seem insignificant to you. *So, it is a good practice to carefully check your work against all instructions before submitting*.

3. *Assignment Resubmissions are Not Permitted After the Assignment Deadline*
   While assignments may be resubmitted before the assignment deadline, they may not be resubmitted after the assignment deadline. If you should accidentally make a submission that is somehow defective, you must discover and correct this error before the assignment deadline. Resubmission of assignments after the deadline will not be permitted for any reason.

4. *Deadline Extensions Must be Requested Before the Deadline*
   If you believe that you have a valid reason for a deadline extension, please submit a Deadline Extension Request before the deadline using the HelpDesk. We have a practice of granting reasonable extension requests. We will only grant extensions beyond the beginning of our next class session in very limited circumstances.

5. *Deductions Will be Made for Late Submissions*
   The grading rubrics for Weekly Assignments and for the Final Project include substantial deductions for late submissions. Please see the assignment grading rubrics for more details.

6. *Assignments Submitted Too Late Will Not be Graded*
   Weekly Assignments submitted more than 14 days late will be considered too late. Final Projects submitted more than 7 days late will be considered too late.

Assignments that are submitted too late will not be graded. These submissions will earn a grade of zero. If you are in danger of missing the *too late* deadline, and you believe that you have a valid reason for an extension, please submit a *Deadline Extension Request* before the deadline via the HelpDesk.

7. *Grade Adjustments Will Be Limited to Automatic Rounding*
   All grades will be awarded on 0 to 100-point scale. Fractional values will be rounded automatically. Fractional portions of grades ending in .0 through .4 will be rounded down. Fractional portions of grades ending in .5 through .9 will be rounded up.

   No further adjustments will be made to grades. This policy applies even in situations in which increasing a grade by just 1 point would cause a student's final letter grade for the course to cross a threshold (i.e., from B+ to A-). Regardless of the potential consequences, grade adjustments will be limited to automatic rounding.

8. *Re-Grading Requests Made Using the HelpDesk Will be Given Fair Consideration*
   It is possible for one of your assignment submissions to be missed during the grading process. This is especially true for assignments that are submitted late. If this happens to you, please submit a Re-Grading Request to the HelpDesk to remind us that your submission still needs grading. Make sure to fully identify the assignment that needs attention.

   Each assignment that we grade is accompanied by a grading feedback form. Please read this feedback to understand our grading decisions. If, after reading the grading feedback form, you believe that our grading decisions are somehow unfair, please submit a Re-Grading Request to the HelpDesk. Include details in your request that identify the assignment and your rationale for the re-grading request. We will give these requests fair consideration and inform you of our determination by posting back to the HelpDesk ticket.

9. *Extra Credit Opportunities are Not Available*
   We do not offer any opportunities to submit work for extra credit.

10. *The Expected Grade for Assignment Submissions That Meet All Expectations is 95.*
    The grading rubrics for all assignments have been designed such that submissions that meet all expectations for the *Regular Exercises* will score 95 points. Assignments will also include a *Challenge Exercise*. Students who choose to submit the Challenge Exercise may earn from 0 to 5 additional points. Challenge Exercises are designed to be quite difficult and the amount of help available from the instructors for Challenge Exercises will be limited to hints.

11. *Weekly Assignment Submissions that Meet Certain Criteria are Subject to a Minimum Score Guarantee*

The grading rubric for Weekly Assignments has been designed to promote two important behaviors:

- Submitting your work in a properly named and formatted file. This helps substantially with grading workflow.

- Submitting your work by the assignment deadline. This assures that you will get the benefit of having tried to solve the problem on your own before seeing the solutions of others.

While separate grading rubric and assignment submission instructions documents will be published, the following is a summary of the coding assignment grading rubric features:

- 10 points will be awarded for submitting a single, properly named and properly formatted file to the proper Canvas assignment submission activity.

- A minimum of 75 points will be awarded for submissions that are submitted on time, and that demonstrate a good faith effort on all Regular Exercises. Late submissions will be awarded 74 points or fewer in this category.

The implication of this grading scheme is that you can expect a score of 85 or higher on all Weekly Assignment submissions that meet both criteria.

12. *Your Participation Grade Will Be Based Upon Participation Points Earned Throughout the Semester*

The table below lists activities for which you may earn participation points and the points earned for each instance.

Activity	Points Earned
Introducing yourself during class in Week 1	10
Speaking contributions during class	2
Actively participating in a group breakout session during class	5
Speaking for your group during debrief of group breakout session during class	2

Your participation grade for the course will be calculated at the end of the semester based upon the total number of participation points earned. Participation grades will

be calculated using a curve. The details of this curve-based grading are presented below:

Choice of Grading Scheme

Participation Points Earned	Applicable Grading Scheme
>= 51	Percentile-Based Scheme
<= 50	Low-Participation Scheme

Percentile-Based Scheme

Participation Points Percentile	Participation Grade
100	100
75	95
50	90
25	80
0	75

Low-Participation Scheme

Participation Points Earned	Participation Grade
50	50
35	35
25	25
<= 24	0

13. *Attendance at Class Sessions May Affect Your Grade*
    While attendance is not graded directly, it will have a significant impact on your participation grade. *Nearly all activities that earn participation points occur during class.*

Academic Integrity Policies and Practices Specific to This Course
There are iSchool and University policies regarding academic integrity. Please see the *iSchool Academic Integrity Statement* under *iSchool and University Academic Policies* (below).

Designing, coding, and testing your own programs is the primary learning activity in this course. Most of the learning value of these activities is the result of doing them on your own. Coding your programs independently leads to expected differences between the program files submitted by different students. These differences are not only expected, but they are often celebrated, as we explore the boundaries of good programming style and good programming practice.

The program files that you submit as your own work in this course will be compared to those submitted by other students in your section and by students in other sections. Your submitted work will also be compared to the official class solution as well as solutions generated by AI tools such as ChatGPT. Instances of high correlation between the work that you submit, and the work of others will be investigated as possible instances of plagiarism or cheating. Both plagiarism and cheating are serious academic integrity infractions. See https://studentcode.illinois.edu/article1/part4/1-402/ .

The submission of highly correlated work represents an academic integrity violation for students who copy the work of others, for students whose work is copied by others, and for students who create and submit a collaborative work when independent work is expected. *So, please do your work independently. Never put yourself in a position where your work could be copied and submitted by another student.*

Our work in checking your submissions for academic integrity violations will continue throughout the course. File comparison is a time-consuming activity and late submissions create new opportunities to discover highly correlated work. So, investigations can be initiated at any time during the course – even after an assignment has been graded.

The only exceptions to this general prohibition on submission of highly correlated work come from two situations where we give you code:

1. *Starter Files*
    As part of assignments, we sometimes give you starter files. Typically, these are test data files. But, sometimes, these are programs that we expect you to include in your project before submission. You are permitted to submit any programs that we provide to you as starter files without further modification.

2. *Tutorial Examples*
    During tutorial videos, we create programs with the expectation that you may follow along. Often, these tutorial programs can serve as a great starting point for the program that you are expected to submit for the exercise. You are permitted to copy and adapt code from our tutorial programs, provided that you type the code into your computer yourself.

Basis for Determining Overall Course Grade

The various components of student work will contribute to the overall course grade based upon the following percentages:

Category	Weight
Participation	15%
Weekly Assignments	40%
Final Project Proposal	15%
Final Project Delivery	30%

Letter grades will be determined as follows:

- A+: 97 - 100%;
- A : 93 - 96%;
- A- : 90 - 92%;
- B+: 87 - 89%;
- B : 83 - 86%;
- B- : 80 - 82%;
- C+: 77 - 79%;
- C : 73 - 76%;
- C- : 70 - 72%;
- D+: 67 - 69%;
- D : 63 - 66%;
- D- : 60 - 62%;
- F :   0 - 59%;

Please note that when converting overall course number grades to letter grades, simple rounding of number grades will be used. Please see *Grade Adjustments Will Be Limited to Automatic Rounding* under *Grading Policies* above.

ISCHOOL AND UNIVERSITY ACADEMIC POLICIES

Incomplete Grades
An exceptional request for an incomplete grade is most often granted to students encountering a medical emergency or other extraordinary circumstances beyond their control. Students must request an incomplete grade from the instructor. The instructor and student will agree on a due date for completion of coursework. The student must submit an Incomplete Form signed by the student, the instructor, and the student's academic advisor to the front office: https://uofi.app.box.com/s/sx7arobhr0gfw12teaetmp1qq32ifdrd
Please see the Student Code for full details:
http://studentcode.illinois.edu/article3/part1/3-104/

iSchool Academic Integrity Statement
The iSchool has the responsibility for maintaining academic integrity so as to protect the quality of education and research in our school and to protect those who depend on our integrity. Consequences of academic integrity infractions may be serious, ranging from a written warning to a failing grade for the course or dismissal from the University.
See the student code for academic integrity requirements:
http://studentcode.illinois.edu/article1/part4/1-401/

Statement of Inclusion
https://diversity.illinois.edu/about/senate-diversity-resolution/
As the state's premier public university, the University of Illinois at Urbana-Champaign's core mission is to serve the interests of the diverse people of the state of Illinois and beyond. The institution thus values inclusion and a pluralistic learning and research environment, one which we respect the varied perspectives and lived experiences of a diverse community and global workforce. We support diversity of worldviews, histories, and cultural knowledge across a range of social groups including race, ethnicity, gender identity, sexual orientation, abilities, economic class, religion, and their intersections.

Religious Observances
In keeping with our Statement of Inclusion and Illinois law, the University is required to reasonably accommodate its students' religious beliefs, observances, and practices in regard to admissions, class attendance, and the scheduling of examinations and work requirements.

If you anticipate the need for an accommodation, please communicate with your instructor in the first two weeks of class. If you are an undergraduate student and your instructor requires an absence letter, you must fill out the Religious Observance Accommodation Request form: https://cm.maxient.com/reportingform.php?UnivofIllinois&layout_id=19 . Other accommodations may be available.

Accessibility Statement
To obtain accessibility-related academic adjustments and/or auxiliary aids, students with disabilities must contact the course instructor and the Disability Resources and Educational Services (DRES) as soon as possible. To contact DRES you may visit 1207 S. Oak St., Champaign, call 333-4603 (V/TTY), or e-mail a message to disability@uiuc.edu .

iSchool COVID-19 Statement
In keeping with university and iSchool policy, all students are required to engage in appropriate behavior to protect the health and safety of our community. If you are on campus, this includes being fully vaccinated, wearing a facial covering properly when required, maintaining social distance, if requested, and using hand sanitizer as needed.

If you feel ill or are unable to come to class or complete class assignments due to issues related to COVID-19, including but not limited to testing positive yourself, feeling ill, caring for a family member with COVID-19, or having unexpected child-care obligations, you should contact the instructor immediately and cc your advisor.

Contact Hours
This course will require approximately 54 contact hours.

Last Revised
2024-05-08