P556: Applied Machine Learning

Fall 2022, Bloomington (Residential)

Instructor   Prof. Xuhong Zhang
zhangxuh@iu.edu

Associate Instructor   Keith Qingyang Xiao (lead)
xiaoq@indiana.edu

Graders   Please see Canvas for up-to-date information

We understand that the Covid-19 pandemic will continue to affect us all in many ways. The health and safety of you, your family, and our community is of paramount importance. Do not attend in-person class if you contract Covid or feel sick. If you encounter unexpected challenges in taking this class, for example due to health, family care, technology access, or other emergency, please contact the instructor as soon as possible. We will work together to find alternative solutions that still maintain the rigor of the course and fairness to all students.

If you have to quarantine yourself during our lecture time, please email the instructor at least 12 hours prior to the course. The instructor will do a hybrid format with additional live zoom lecture.

Contacting course staff and other students: Communication between you, us, and other students is really important. Help us make this more efficient by following these guidelines!

- For general questions about course material, assignments, etc., please use the Q&A Community tool on Canvas so that others can benefit from the answers. You can expect a response within 24 hours. We encourage you to actively monitor and participate in Q&A Community discussions.

- For more informal interactions with other students in the class, e.g. for coordinating group projects, sharing interesting AI news items, etc., you can use Slack, which we will send out an invitation for you to join. Using Slack is optional; please do not use it for communicating with the course staff.

- For questions and issues that are specific to you, that involve course logistics or administration, or that do not fit into any other category, email us at csp556@indiana.edu. Expect a reply from someone on the course staff within 24 hours; re-send your email if you do not get one.

- For more sensitive personal matters, email the instructor directly (zhangxuh@iu.edu).

- For interactive discussions, each of the instructors will hold open office hours either in person or by video conference; see updated office hours on Canvas.

- For alerting us to problems or concerns with grading, see Regrade Requests, below.

Meetings and Sections: This semester there are two residential sections of CSCI P556: same day, different time slots and locations. We will have a physical seat reserved for students to attend the in-person lectures on Mondays and Wednesdays 1:15-2:30pm/4:45-6:00pm at LI 028/MO 007. Both sections share the same

course staff and office hours, and are very similar in terms of course content and rigor. The programming assignments will be identical across the two sections and share the same deadlines, so students may form teams (see Groups, below) that include students from both sections. **We will follow any Covid-related IU policies as we go through the entire semester for both sections**.

Course Overview: This course focuses on: introducing theoretical concepts and algorithms in a step-by-step manner, while infusing them with intuition, examples and python jupyter notebooks. In this spirit you will study core ML algorithms, while also working through numerous example applications of machine learning. Concrete examples help illustrate the broader concepts by putting the learned material directly into action. This combination of theory and hand-ons will help you master core ML concepts and algorithms that are used, not only in Silicon Valley but, throughout the world, while also offering intuitive yet informative explanations of how machine learning algorithms work, how to use them, and most importantly, how to avoid the most common pitfalls. For those with a stronger interest in ML theory and development this course will provide an optional track that will focus on delving into the theory a little more deeply, and that culminates in coding up core ML algorithms from scratch and possible extending them.

If you want to become a machine learning practitioner, a better problem solver, or maybe even consider a career in machine learning research, then this course is for you. However, for a novice, the theoretical concepts behind machine learning can be quite overwhelming.

Topics and schedules are on the next page.

Tools: Tools are an important part for this course. For hardware tools, like mentioned above, you need a laptop or tablet with keyboard input for on-site programming and quizzes. For class interaction, you can use your laptop, tablets, or smartphones. Related works you need to pay attention and accomplish:

- **Top Hat registration**: Top Hat is the tool that we use for in-class interaction. The registration is posted as an assignment due on the Friday of the first week. Once you registered, you can get access to Top Hat via your Canvas account and respond to our in-class questions. This is located at the assignment module in Canvas.

- **Important– 1st assignment group preference**: Please indicate how many members you prefer to work with for the 1st assignment or you prefer to work alone before the due date. Otherwise we will randomly assign you to work alone, or with 1 or 2 other people. It is located at the quiz module in Canvas. For the 1st assignment we assign group members randomly. See details about groups below.

- **Slack communication**: Slack will be used for us to discuss informal questions. The invitation of Slack will be sent out to you, so you can click and join.

- **IU Github**: We'll be using IU Github for managing your source code. It's okay if you haven't used github before – we'll give you enough information to get started later on. For now, please create your github account by simply logging in (with your IU credentials) to IU GitHub. (Note that this is a different account than the main GitHub site.

Office Hours: The instructor and each of the AIs will hold weekly open office hours, either in person or online via the Zoom video conferencing tool. Please see Canvas for up-to-date times and URLs. If you cannot make any of these office hours, contact an instructor to make a separate appointment.

Prerequisties: Experience with programming, data structures, and algorithms will be assumed (CS C343 or equivalent). Assignments will involve substantial amount of programming in Python. For grading convenience, your homework needs to be submitted in Python (Jupyter Notebook). Submissions in other language will be be accepted or graded. If you do not know Python, we will expect you to learn it outside of class; Python has a reputation for being easy to learn, and those with programming background in another

general-purpose programming language (like C/C++, Java, Ruby, etc.) can usually learn it within a few days. Tutorial resources about Python will be provided on our first hands-on programming class. Please consult the instructor if you have concerns about your programming background. In addition, we will encounter math of various kinds, including linear algebra, probability theory, and basic calculus. We will review the key mathematical concepts as we go, but students with limited math backgrounds may need to do additional reading outside of class.

Date	Topic	Notes
Aug, 22	Introduction	
Aug, 24	Input: concept, instances and attributes (1)	
Aug, 29	Input: concept, instances and attributes (2)	
Aug, 31	Data Transformation	
Sep, 5	No class	Labor day
Sep, 7	Hands-on programming skills (1)	1st in-class quiz
Sep, 12	Output knowledge representation	
Sep, 14	Basic models: Linear models (1)	
Sep, 19	Basic models: Linear models (2)	
Sep, 21	Basic models: Instance based models and Clustering	2nd in-class quiz; 1st home
Sep, 26	Onsite programming (1)	
Sep, 28	Credibility: evaluating model performance	
Oct, 3	Credibility: understanding model performance	
Oct, 5	Hands-on programming skills (2)	3rd in-class quiz
Oct, 10	Basic models: Naive Bayes	
Oct, 12	Basic models: SVM	2nd homework due
Oct, 17	Advanced models: Kernel methods	
Oct, 19	Advanced models: Trees and rules	4th in-class quiz
Oct, 24	Advanced models: Ensemble learning	
Oct, 26	Hands-on programming skills (3)	
Oct, 31	Written Midterm Exam	
Nov, 2	Onsite programming (2)	5th in-class quiz
Nov, 7	Advanced models : Gaussian process	
Nov, 9	Advanced models : Perceptron and Deep Learning basic	3rd homework due
Nov, 14	Advanced models: Deep learning (1)	
Nov, 16	Advanced models: Deep learning (2)	6th in-class quiz
Nov, 21	No class	Thanks giving break
Nov, 23	No class	Thanks giving break
Nov, 28	Hands-on programming skills (4)	
Nov, 30	Machine Learning Applications	4th homework due
Dec, 5	Onsite programming (3)	
Dec, 7	Course review	
Dec, 12	Final exam for course 36860; Natural Language Processing for course 9100	
Dec, 16	Final exam for course 9100	

Textbook: There are several books can be good resources for further reading. Norvig and Russell, *Artificial Intelligence: A Modern Approach*;Trevor Hastie, Robert Tibshirani, Jerome Friedman, *The elements of statistic learning*, 2nd edition; Goodfellow, Bengio and Courville, *Deep Learning*. You can check if there is an updated new pdf version available online. These books are not required and you will *not* need the book for the exams, so it is practical to share a book with other students in the class.

Resources: Indiana University Luddy School has partnered with Knack to provide students with access to campus tutors. Students looking for additional assistance outside of the classroom are advised to consider

working with a peer tutor through Knack. To view available tutors, visit indiana.joinknack.com (Links to an external site.) and sign in with your student account.

If you have any questions about the partnership, please feel free to reply to this email or contact Knack's Campus Success Lead Priya Thomas at priya@joinknack.com.

Grading and Requirements

- Assignments (40%): 4 assignments, most of which will require substantial programming projects. These are typically group assignments (see groups, below).

- In-class assignments (15%), and bi-weekly quizzes (10%): These are not meant to be stressful events, but ways to give you practice understanding course concepts while rewarding regular class participation and attendance. They also give instructors feedback on how well the class is progressing. We will have 3 in-class programming assignments and each will be counted for 5%. The quiz that you get lowest score will be dropped and not counted for your final grade. So in total your top 5 quiz will be counted.

- In-class Midterm exam (15%): *If you have an unavoidable conflict with the midterm day or time, you must notify us at least 2 weeks before the midterm date. We will not be able to schedule make-up or conflict exams after that day.*

- In-class Final exam (20%): *If you have an unavoidable conflict with the final exam day or time, you must notify us at least 2 weeks before the exam date. We will not be able to schedule make-up or conflict exams after that day.*

Final Course Letter Grades: Students whose final course scores (reflecting all assignments, exams, etc.) are above 90% are guaranteed at least an A-, above 80% are guaranteed at least a B-, and above 70% are guaranteed at least a C-. However, other than these guaranteed minimum scores, we do not use a predefined mapping of final course scores to letter grades. Instead, we map final course scores to letter grades such that we roughly match the grade distributions of this course in past semesters (which are available on the Registrar's website). This means that the thresholds for final grades are typically significantly lower than the guaranteed thresholds above (e.g., required score for an A- is typically below 90%. Students should thus view their grades not as absolute numbers but relative to the performance of the rest of the class; we will publish the median and mean scores for assignments and exams to facilitate this. Students with concerns about their grades are welcome to consult with the instructor during the office hours.

Groups: We encourage students to work in small groups for the programming assignments, because: (1) in a good group, students bring complementary strengths and weaknesses, allowing them to learn from one another, (2) a good group can accomplish much more in a short period of time than a single person can, and (3) working in groups is a fact of life in both industry and academia, and developing good teamwork skills is crucial to succeeding in most computing-related careers. At the same time, we understand that finding and building good teams is difficult. Our teamwork policy tries to balance these considerations, as follows:

1. All students working in a team for an assignment receive the same grade on that assignment. Each group makes up to 3 people.

2. After each assignment, we collect confidential feedback from each student about the other members of their team. We use this information to monitor for problematic teams and students. In exceptional cases, we may also use this information to adjust final course grades (up or down).

3. For the first assignment, we only ask you if you want to work alone, with another one or with two people, to encourage students to meet new colleagues. Please fill out your first group assignment preference before the due date.

4. For each subsequent assignment, we collect confidential teammate preferences from each student. These preferences can include requesting a particular team, requesting not to be assigned to a particular person, or requesting to work alone (in a team of 1).

5. We will assign students to teams according to their preferences. Students whose preferences are compatible (e.g., three students all request one another) are typically assigned their preferred team; otherwise, we attempt to satisfy as many preferences as possible. Students who do not fill out the preference form correctly or on time may be assigned to an arbitrary team.

6. Teams are final once they are assigned (but may be changed for the next programming assignment).

Late Submission Policy: Assignments will be accepted up to 48 hours after the due date, but a 10% late penalty will be assessed. Assignments received more than 48 hours late will not be accepted and will receive a grade of 0. Note: All assignments must be submitted according to the instructions given on the assignment hand-outs, which will typically require electronic (online) submission. We use the time that submissions are received by the servers to judge timeliness. It is the student's responsibility to upload submissions well ahead of the deadline to avoid last minute problems with network connectivity, browser crashes, etc. It is a very good idea to make early submissions and then upload updates as the deadline approaches; we will grade the last submission received before the deadline.

Attendance Policy: Students are expected to attend all class sessions. To allow for an occasional unavoidable exception, we will drop the lowest in-class quiz. Additional grades will only be dropped if (1) a student is missing class for a valid academic or religious event (such as a job interview or conference) and notifies the instructor at least 1 week ahead of time, or (2) a student misses class for health reasons (involving themselves or a dependent child or quarantine), notifies the instructor within 24 hours of missing class. In case of quarantine, notifies the instructor at least 12 hours prior to the course so the instructor can set up zoom live course if necessary.

Regrade Requests: We follow a formal process for requesting regrades to ensure fairness to all students. Do not contact AIs directly about your grading; they are instructed to ignore regrade requests that do not follow the procedure below. If you have a question about the grading of an assignment, or believe you have discovered an error in our grading of one of your assignments, then, within one week of when the grade was released: (1) write a specific, precise, self-contained, well-reasoned explanation, including evidence (e.g. screenshots of program output, etc.) if appropriate; (2) submit this write-up in TWO places: (a) as a private Q&A Community post using the "regrade-request" tag, and (b) as a comment on your Assignment in Canvas; (3) within two weeks, a grader will examine your request and make any corrections; (4) if you are not satisfied, you may appeal by posting a follow-up comment on QA Community, and the lead instructor will regrade the entire assignment from scratch to determine a final grade. Note that regrade requests may result in either positive or negative grade adjustments because we might conclude that we were too generous in grading and reduce your grade instead of increasing it. For this reason, do not submit regrade requests for minor grade problems (e.g. 1-2 points); the grading system for this class is designed so that your final grade will not be affected by a few points here or there. Your time (and our time) is too valuable to argue about small grade adjustments; reserve your requests for cases where a grader made a mistake or significantly misunderstood.

Technology requirements: You will likely also want a Python development environment on your local computer. There are many free and commercial options, and you can choose whichever you'd like. See https://wiki.python.org/moin/IntegratedDevelopmentEnvironments for a list; popular choices include PyCharm (which is free for academic use), IDLE (a simple IDE that comes built-in to most Python distributions), and XCode (which is free but available only for Macs).

Available software: Although not required, other software may be helpful (e.g. graphing programs, word processors, etc.). Get no-cost access to hundreds of software programs and applications through IUware

and IUanyWare. All you need is your IU email address. Use IUware to install software directly onto your hard drive. Use IUanyWare to stream 400+ apps on your desktop or through the mobile app with your IU login. Visit iuware.iu.edu and iuanyware.iu.edu, or contact the UITS Support Center to learn more. Students in this class are invited to use Boost, a free smartphone app developed at IU that provides notifications and reminders about schoolwork in Canvas. It is designed to help students keep track of assignment deadlines, important announcements, and course events all in one easy-to-use app. For more information, see https://kb.iu.edu/d/atud or https://boost.iu.edu.

Academic Integrity Policy: We take academic integrity very seriously. You are required to abide by the Indiana University policy on academic integrity, as described in the Code of Student Rights, Responsibilities, and Conduct, as well as the Computer Science Statement on Academic Integrity (http://www.soic.indiana.edu/ doc/graduate/ graduate-forms/Academic-Integrity-Guideline-FINAL-2015.pdf). It is your responsibility to understand these poli- cies. Briefly summarized, the work you submit for course assignments, projects, quizzes, and exams must be entirely your own (or entirely that of your group, if groupwork is permitted). If you use the ideas (including text, source code, algorithms, concepts, diagrams, slides, etc.) of others, you must give proper credit with a prominent citation and an explicit indication of which idea(s) or material(s) you borrowed so that another person (e.g. a grader) can easily separate your contribution from the work of others. You may discuss assignments with other students (or students in other groups) at a high level, by for example discussing general methods or strategies to solve a problem, but you must cite the other student in your submission. Looking at someone else's code related to an assignment, whether online or from another student, will almost certainly lead to academic dishonesty. Sharing your assignment code with another student also almost certainly constitutes academic dishonesty. The consequences of academic dishonesty are extremely serious. We respond to plagiarism and academic miscon- duct according to university policy. In assigning sanctions, we follow CS Program policy: "The ordinary departmental level penalty for cheating is failure in the course" but "in all cases, the penalty will be more severe than not turning in the assignment." In addition, "the student will no longer be eligible for the guaranteed financial aid." Moreover, University policy requires us to report the incident to the Dean of Students, who may apply additional sanctions, including expulsion from the university. Students agree that by taking this course, papers and source code submitted to us may be subject to textual similarity review, for example by Turnitin.com. These submissions may be included as source documents in reference databases solely for the purpose of detecting plagiarism of such papers or codes.

Religious Holidays: Indiana University respects the right of all students to observe religious holidays and will make reasonable accommodation, upon request, for such observances. Students missing class for a religious observance can find the officially approved accommodation form by going to the Vice Provost for Faculty and Academic Affairs webpage for religious accommodations. The form must be submitted at least 2 weeks prior to the anticipated absence.

Disabilities: Every attempt will be made to accommodate qualified students with disabilities (e.g. mental health, learning, chronic health, physical, hearing, vision neurological, etc.). You must have established your eligibil- ity for support services through the appropriate office that services students with disabilities. Note that ser- vices are confidential, may take time to put into place and are not retroactive; captions and alternate media for print materials may take three or more weeks to get produced. Please contact Disability Services for Students at http://disabilityservices.indiana.edu, 812-855-7578, or in Wells Library Room W302. Walk-ins are welcome 8 AM to 5 PM, Monday through Friday. You can also locate a variety of campus resources for students and visitors that need assistance at: http://www.iu.edu/ ada/index.shtml

Title IX: Our responsibility is to create a positive learning environment for all students. Federal law (Title IX) and IU's Sexual Misconduct Policy prohibit sexual misconduct in any form, including sexual harassment, sexual assault, stalking, and dating and domestic violence. If you have experienced sexual misconduct, or know someone who has, the University can help. If you are seeking help and would like to speak to someone confidentially, you can make an appointment with Sexual Assault Crisis Services (SACS) at (812) 855-8900

(for counseling services), Confidential Victim Advocates (CVA) at (812) 856-2469 (for advocacy and advice services), or IU Health Center at (812) 855-4011 (for health and medical services). Title IX and University policy require us to share any information brought to our attention about potential sexual misconduct with the campus Deputy Title IX Coordinator or IU's Title IX Coordinator. In that event, those individuals work to ensure that appropriate measures are taken and resources are made available. Protecting student privacy is of utmost concern, and information will only be shared with those that need to know to ensure the University can respond and assist. Visit stopsexualviolence.iu.edu to learn more.

Bias-based incidents: Bias incidents (events or comments that target an individual or group based on age, color, religion, disability, race, ethnicity, national origin, sex, gender, gender identity, sexual orientation, marital status or veteran status) are not appropriate in our classroom or on campus. What should you do if you experience, witness, see, or hear a bias incident? Report it by submitting a report online (biasincident.indiana.edu) or calling the Dean of Students Office (812-855-8187).

Emergency Preparedness: Although rare, emergencies can and do occur, and it's important that you know how to handle them. Below is a brief summary; see http://protect.iu.edu/ for more information.

- Tornado: Seek shelter. Move to an interior room on the lowest level. Stay away from windows and exterior doors. Listen to a weather radio for updates. Stay away from hazardous materials.

- Fire: Evacuate. Pull the fire alarm. Call 911. Leave the building, closing doors behind you. If unable to exit go to the nearest stairwell or place of refuge, don't use elevators. Assemble in designated area.

- Medical emergency: Call 911. Do not move the victim unless in immediate danger. If trained, administer first aid, CPR/AED.

- Suspicious activity, e.g. object is out of the ordinary, person is behaving strangely, gut feeling that something is wrong: If you see something suspicious, call 911.

- Hazardous materials: Stay back. If life-threatening, pull the fire alarm, evacuate and call 911. If non-life- threatening, call 911 and provide information on type of incident and location.

- Bomb threat: Remain calm. Get as much information as possible from the caller: location of device, what it looks like, what will cause it to explode. Note background sounds, gender of caller, other notable characteristics of the caller's voice. Call 911.

- Active shooter: Call 911. Leave the building if possible, otherwise hide in a concealed place. Lock and barricade door, turn off lights. Wait for law enforcement. As a last resort, overpower the shooter.

Special considerations during the Covid-19 pandemic. It is important that you follow campus public health guidelines, which we summarize here.

- **Mask requirement**: Please follow IU's policy at: https://www.iu.edu/covid/prevention/masks-and-ppe.html

- **Student Rights**: Any student who believes another person in a class is threatening the safety of the class by not wearing a mask may leave the class without consequence.

- **Attendance**: The CRA requires that you take your temperature every morning and that you refrain from attending class if you have a temperature of 100.4 or other symptoms of illness. In order to ensure that you can do this, attendance will not be a factor in the final grade. Attendance may still be taken to comply with accreditation requirements.

- **Summary Suspension Policy**: A student may be summarily suspended from the university and summarily excluded from university property and programs by the Provost or designee of a university

campus. The Provost or designee may act summarily without following the hearing procedures established by this section if the officer is satisfied that the student's continued presence on the campus constitutes a serious threat of harm to the student or to any other person on the campus or to the property of the university or property of other persons on the university campus.

The Provost has determined that refusal to comply with the public health requirements specified in the Student Responsibility form, including the requirement of wearing a mask in all IU buildings, constitutes "a serious threat of harm to other persons" within the meaning of the summary suspension policy. In addition, the Provost has determined that a person who does not comply with these requirements, as evidenced by three credible violations of the policy reported to the campus from any source, constitutes "a serious threat of harm to other persons" within the meaning of the summary suspension policy.