CS 4713/5713 - Computational Learning Theory

Dimitris Diochnos
The University of Oklahoma, School of Computer Science

Fall 2024

Time and Location

The course will take place at Carson Engineering Center 0441. Tuesdays & Thursdays, 10:30am-11:45am.

Canvas: Canvas will be used in order to distribute homework assignments and potentially other reading material.

Instructor

Dimitris Diochnos, 230 Devon Energy Hall (DEH), diochnos@ou.edu .

Teaching Assistants

No teaching assistants are available for this course.

Office Hours

Office hours will be held:

- on **Tuesdays** between **3:00pm-3:50pm**,
- on **Thursdays** between **3:00pm-3:50pm**,
- as well as on **Fridays** between **1:30pm-2:15pm**, or
- **by appointment**.

Prerequisite Background

Design and analysis of algorithms, basic computational complexity theory, mathematical maturity. Tools from probability theory will be discussed (briefly) on demand as they arise.
Prerequisites: CS 4413 or DSA 4413. You can also enroll with the permission of the instructor.

Course Catalog Description

Learning using membership queries, equivalence queries, version spaces, decision trees, linear models. Probably approximately correct (PAC) learning, VC-theory, distribution-independent learning. Representation issues and intractability. Noise models, statistical queries, PAC learning under noise, poisoning attacks, adversarial examples. Distribution-specific learning and evolvability. Online learning and mistake bounds. Weak and strong learning (boosting). No student may receive credit for 4713 and 5713.

Schedule of Classes

The syllabus is continuously updated and subject to change. We will cover the material at a pace that is comforable. Our **first meeting** is on **Tuesday, August 20, 2024** and our **last meeting** is on **Thursday, December 5, 2024**.

A **rough outline** for the course, which is subject to change slightly depending on our pace, is shown in Table 1.

Table 1: Tentative Course Schedule

Period	**Topics**
Week 1	Syllabus, expectations, mathematical background
Week 2	Concept learning using queries. Introduction to version spaces.
Week 3	Conclusion of version spaces. Introduction to decision trees.
Week 4	Conclusion of decision trees. Linear models for classification.
Week 5	Introduction to PAC learning.
Week 6	Learning in the realizable case using a finite hypothesis space. Improving explainability by learning hypotheses with few relevant variables.
Week 7	Reviewing material and preparation for the midterm. Midterm.
Week 8	Agnostic learning using finite hypotheses spaces and empirical risk minimization. Intractability of learning 3-term DNF formulae.
Week 9	VC-theory. Upper bounds and lower bounds for distribution-independent learning. Noise models. Malicious noise, induced distributions.
Week 10	Random misclassification noise. Hardness of learning conjunctions under random misclassification noise.
Week 11	PAC learning under class imbalance. Statistical queries.
Week 12	Poisoning attacks. Evasion attacks (adversarial examples). Other topics on trustworthy machine learning.
Week 13	Distribution-specific learning, evolvability.
Week 14	Online learning. Halving, randomized halving, weighted majority algorithm, randomized weighted majority algorithm, winnow.
Week 15	Weak and strong learning (boosting).
Week 16	Student presentations.

The **midterm exam** is **in-class** and will take place near the 7th week of the classes.

There will be **no final exam** for this class. Instead, we will have a **semester-long project** with presentations occuring during the last week.

Textbook, Notes and Related Reading Material

The class will rely to a large extent on papers as well as on handouts by the instructor. A very good book for introduction to machine learning that we will be using for some of the topics that we plan to cover in this class, is the book *Machine Learning*, by Tom Mitchell [4]. The book is available online for free by the author, at:

- https://www.cs.cmu.edu/~tom/mlbook.html .

Having said that, we plan to cover material that is not available in the above book, and is instead covered by the more recent book **Understanding Machine Learning**, by Shai Shalev-Schwartz and Shai Ben-David [8], or the book **Foundations of Machine Learning**, by Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar [5]. Both of these books are available online for free by the authors at the following addresses respectively:

- http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/ and
- https://cs.nyu.edu/~mohri/mlbook/ .

Furthermore, a classic resource for covering aspects relevant to the PAC model of learning is [3] and we will blend discussions from this book together with the more recent treatment found in [8] or [5].

It is advised that the students take notes from the material that is covered in class.

In order to dive deeper into topics related to trustworthy supervised learning we will be using references such as:

- **Interpretable Machine Learning** [6],
- **Trustworthy Machine Learning** [10],
- **Fairness and Machine Learning: Limitations and Opportunities** [1].

Other Books of Interest. Domingos in *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* [2], has a very nice (high-level) description, sometimes accompanied by historical anecdotes, on different aspects of machine learning. Valiant's book *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World* [9] provides a good narrative for connections that we discuss between evolution and learning. Finally, an unfortunate side of the ever-increasing influence that machine learning algorithms have in our daily lives is discussed in *Weapons of math destruction: How big data increases inequality and threatens democracy*, by Cathy O' Neil [7]. There are many more books along these lines and will be mentioned in some preliminary slides in class. These issues motivate the urgency of developing machine learning mechanisms that are **trustworthy**.

Project

Students should form groups of 2 and work on the semester-long project. No group can be of 1 person, or of more than 2 persons. The work of the students will be presented near the end of the course and will also have a written component. Projects can be **implementation-based** or **theory-based**. In either case, the students will have to present what they read, and potentially the findings of an implementation, at the end of the semester. The **students enrolled in CS 5713** will have to do a more extensive literature review by citing and explaining **two more related papers per person** in the group compared to the load that undergraduate students have.

Please see the document **Project Description** for more information about the write-ups and the deadlines.

Examinations

The midterm exam will be a closed-book written exam in the class where we have our regular meetings.

There is no final exam for this class. Your final written report as well as your presentation serve this purpose.

Homework Assignments

Undergraduate students will have five homework assignments whereas graduate students will have six homework assignments. There will be 5-6 homework assignments; most likely six, with a tentative schedule as shown in Table 2.

Table 2: Tentative Homework Schedule

Homework	**Announced**	**Due**	**Assigned To**
1	Week 1 (end)	Week 3 (beginning)	All students
2	Week 3 (beginning)	Week 4 (end)	All students
3	Week 4 (end)	Week 6 (beginning)	All students
4	Week 8 (beginning)	Week 10 (beginning)	All students
5	Week 10 (beginning)	Week 12 (beginning)	All students
6	Week 12 (end)	Week 15 (end)	Graduate students only

The contribution for your grade based on homework will be computed by adding up all the points that you receive from the individual homework assignments and then dividing by the maximum amount of points that you could gather from all these assignments. I expect the assignments to be weighted roughly evenly, except apart from the first one (so that we can start working on homework problems earlier rather than later).

Grading

Grading will be based on the following:

- **40% homework assignments**,
- **40% semester-long project**,
- **20% in-class midterm exam**.

Grades may also be adjusted slightly upward or downward depending on class participation. I expect grading to be along the lines shown in the table below:

Percentage	Grade
$\geq 90\%$	A
$\geq 80\%$	B
$\geq 70\%$	C
$\geq 60\%$	D
otherwise	F

Grades may be curved at the end of the semester, but grade thresholds will never be higher than those shown above.

Additional Workload for Students Enrolled in CS 5713

The additional workload for students enrolled in CS 5713 has two components. The first component is an additional individual homework assignment near the end of the semester, which will allow these students to understand how research is being performed in machine learning theory. This is an effort of more than 20% on the homework assignments compared to students enrolled in CS 4713, which corresponds to more than 8% of the total final grade. Furthermore, the students enrolled in CS 5713 need to perform a more extensive review of the related work in their write-ups for the semester-long project and discuss such broader connections with the main paper that they are presenting during the final presentations. This combined effort of increased literature review and discussion during the final presentation corresponds to approximately 10% increase of load compared to students enrolled in CS 4713 for their projects, thus, overall, this second component corresponds to an additional 4% effort that students enrolled in CS 5713 need to accomplish for the same final grade as those enrolled in CS 4713.

As a summary, students enrolled in CS 5713 need to put about 12%-13% more effort compared to students enrolled in CS 4713, in order to achieve the same numerical (weighted) grade.

Course Policies

Collaboration Policy

Regarding homework assignments, unless otherwise specified, students may discuss problem sets with one another. However, students should afterward write the solutions on their own. Collaborators (people you speak to about an assignment) must be named at the top of the assignment (together with their university IDs). No collaboration will be allowed on exams. Students **must** form **groups of two (2) people** and work together on the semester-long project. You cannot work alone. You cannot work in a group of three (3) or more people.

General Remarks. Please note the following two.

- **If you are unsure if something is permitted, consult with me before doing it.**
- **For the (midterm) exam students are required to work alone** and follow the stated rules exactly.

Late Work Policy

You can postpone once your homework or project submission by 24 hours without any penalty. After the first time that you have a late submission, a 10% (of the maximum possible grade) penalty will be applied for every day that is late – the maximum delay can be 3 days (including the first time that you have a late submission). This penalty is applied of course to every member of the group where you belong to.

We will be using an electronic system (Canvas) for the students' submissions and therefore it is your responsibility to turn in your homework (or work for the semester-long project, should this be the case) on time. **Please coordinate within your group and make one submission per group if you are submitting a semester-long project checkpoint or the final write-up for the semester-long project.**

Chegg and Other Online Tutoring Sources

There are a wide variety of tutoring resources available through paid websites. Many of these sites have students upload assignments and solutions and surreptitiously provide these documents to other students. What appears to be a session with a tutor may be, behind the scenes, the tutor doing a search of their company database of solutions to share. By using these sites you risk being charged with academic misconduct, either by supplying other students with answers they did not author or by receiving someone else's answer that you did not author. Since these companies are not open with students about their practices, you cannot know whether a tutor is providing meaningful support (for example, identifying misunderstandings of content and explaining them) or simply feeding you someone else's solution a bit at a time. The tutor's actions can result in different students submitting answers that are identical, which may be flagged as academic misconduct during grading.

Make-Up Midterm

In some rare cases I can offer a makeup midterm to a student (subject to my schedule and availability as well). However, if the student misses their rescheduled midterm the student will receive a zero (0).

Classroom Conduct

Disruptions of class will not be permitted. Examples of disruptive behavior include:

- Allowing a cell phone or pager to repeatedly beep audibly.
- Playing music or computer games during class in such a way that they are visible or audible to other class members.
- Exhibiting erratic or irrational behavior.
- Behavior that distracts the class from the subject matter or discussion.
- Making physical or verbal threats to a faculty member, teaching assistant, or class member.
- Refusal to comply with faculty direction.

In the case of disruptive behavior, I may ask that you leave the classroom and may charge you with a violation of the Student Code of Responsibilities and Conduct.

Class Web Page

I intend to maintain a web page for the class under

https://www.diochnos.com/teaching

The exact url will be provided to the students by the beginning of the semester. We will also be using Canvas for announcements, homework assignments, and for the semester-long project. Furthermore, Canvas will also provide the students with the exact url of the webpage that I plan to maintain under my website.

Login to the Canvas website using your 4+4 (first four letters of your last name followed by the last four digits of your student number), using your standard OU password. If you have difficulty logging in, call 325-HELP. This software provides a number of useful features, including a list of assignments and announcements, an electronic mailing list, newsgroups, and grade book. All handouts are available from Canvas. You should check the site daily. When I update the site, I will post an announcement telling you what has been added and where it is located.You are responsible for things posted on the site with a 24 hour delay.

Student's Feedback for the Course

The College of Engineering utilizes students' feedback as one of the bases for evaluating the teaching effectiveness of each of its faculty members. The results of these forms are important data used in the process of awarding tenure, making promotions, and giving salary increases. In addition, the faculty uses these forms to improve their own teaching effectiveness. The original request for the use of these forms came from students, and it is students who eventually benefit most from their use. Please take this task seriously and respond as honestly and precisely as possible, both to the machine-scored items and to the open-ended questions.

Class Email Alias

Urgent announcements will be sent through email. It is your responsibility to:

- Have your university supplied email account properly forwarded to the location where you read email.
- Make sure that your email address in Canvas is correct, and forwards email to the place where you read it. I'll send out a test message during the first week of class. If you do not receive this message, it is your responsibility to get the problem resolved immediately.
- Have your email program set up properly so that replying to your email will work correctly the first time. You can send email to yourself and reply to yourself to test this.

If you need assistance in accomplishing any of these tasks, contact 325-HELP.

Newsgroups and Email

The newsgroup on Canvas should be the primary method of communication, outside of class. This allows everyone in the class to benefit from the answer to your question. If you email me a question of general interest, I may post your question and my answer to the newsgroup. Matters of personal interest should be directed to email instead of to the newsgroup, e.g. informing me of an extended personal illness. Posting guidelines for the newsgroup are available on Canvas.

Incompletes

The grade of I is intended for the rare circumstance when a student who has been successful in a class has an unexpected event occur shortly before the end of the class. I will not consider giving a student a grade of I unless the following three conditions have been met.

1. It is within two weeks of the end of the semester.
2. The student has a grade of C or better in the class.
3. The reason that the student cannot complete the class is properly documented and compelling.

Add/Drop/Withdrawal Deadlines. Please consult the OU academic calendar (as well as the policies of the School of Engineering) for the following deadlines:

- **Add a course**
- **Drop a course without penalty (course removed from transcript)**
- **Drop a course with a W on transcript**

University Policies

The instructor reserves the right to add, remove, or change any element of class policy at any time and for any reason, within the limits of University policy.

OU is committed to creating a learning environment that meets the needs of its diverse student body. If you anticipate or experience any barriers to learning in this course, please feel welcome to discuss your concerns with me.

Academic Integrity

Academic misconduct hurts everyone but particularly the student who does not learn the material. All work submitted for an individual grade should be the work of that single individual and not his/her friends. It is fine to ask a fellow student for help as long as that help does not consist of copying any computer code, or solutions to other assignments. Students working on joint projects (e.g., the groups you form for the semester-long project) may certainly help one another and are expected to share ideas, code, and solutions within their group. However, they may not share beyond their group.

1. Do not show another student (or group) a copy of your projects or homework before the submission deadline. The penalties for permitting your work to be copied are the same as the penalties for copying someone else's work.
2. Make sure that your computer account is properly protected. Use a good password, and do not give your friends access to your account or your computer system. Do not leave printouts or thumb drives around a laboratory where others might access them.

Upon the first documented occurrence of academic misconduct, I will report it to the Campus Judicial Coordinator. The procedure to be followed is documented in the University of Oklahoma Academic Misconduct Code. In the unlikely event that I elect to admonish the student, the appeals process is described in `http://www.ou.edu/integrity`. For specific definitions on what constitutes cheating, review the Student's Guide to Academic Integrity (https://www.ou.edu/integrity/students).

Religious Observance

It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays, without penalty. [See Faculty Handbook 3.15.2]
(https://apps.hr.ou.edu/FacultyHandbook/#3.15.2).

Reasonable Accommodation Policy

The Accessibility and Disability Resource Center is committed to supporting students with disabilities to ensure that they are able to enjoy equal access to all components of their education. This includes your academics, housing, and community events. If you are experiencing a disability, a mental/medical health condition that has a significant impact on one or more life functions, you can receive accommodations to provide equal access. Possible disabilities include, but are not limited to, learning disabilities, AD(H)D, mental health, and chronic health. Additionally, we support students with temporary medical conditions (broken wrist, shoulder surgery, etc.) and pregnancy. To discuss potential accommodations, please contact the ADRC at 730 College Avenue, (ph.) 405.325.3852, or adrc@ou.edu.

Title IX Resources and Reporting Requirement

Anyone who has been impacted by gender-based violence, including dating violence, domestic violence, stalking, harassment, and sexual assault, deserves access to resources so that they are supported personally and academically. The University of Oklahoma is committed to offering resources to those impacted, including: speaking with someone confidentially about your options, medical attention, counseling, reporting, academic support, and safety plans. If you would like to speak with someone confidentially, please contact OU Advocates (available 24/7 at 405-615-0013; see https://www.ou.edu/gec/gender-based-violence/advocates) or another confidential resource (see "Can I make an anonymous report?"; https://www.ou.edu/gec/gender-based-violence/learn-more). You may also choose to report gender-based violence and discrimination through other means, including by contacting the Institutional Equity Office (ieo@ou.edu, 405-325-3546; https://www.ou.edu/eoo) or police (911). Because the University of Oklahoma is committed to the safety of you and other students, I, as well as other faculty, Graduate Assistants, and Teaching Assistants, are mandatory reporters. This means that we are obligated to report gender-based violence that has been disclosed to us to the Institutional Equity Office. This includes disclosures that occur in: class discussion, writing assignments, discussion boards, emails and during Student/Office Hours. For more information, please visit the Institutional Equity Office (https://www.ou.edu/eoo).

Adjustments for Pregnancy/Childbirth Related Issues

Should you need modifications or adjustments to your course requirements because of documented pregnancy-related or childbirth-related issues, please contact your professor or the Accessibility and Disability Resource Center at 405/325-3852 as soon as possible. Also, see the Institutional Equity Office FAQ on Pregnant and Parenting Students' Rights (https://www.ou.edu/content/dam/eoo/documents/faqs/faqs-pregnant-and-parenting-students.pdf) for answers to commonly asked questions.

Final Exam Preparation Period

Pre-finals week will be defined as the seven calendar days before the first day of finals. Faculty may cover new course material throughout this week. For specific provisions of the policy please refer to OU's Final Exam Preparation Period policy.
Please see https://apps.hr.ou.edu/FacultyHandbook#4.10

Emergency Protocol

During an emergency, there are official university procedures (https://www.ou.edu/campussafety/policy-and-procedures) that will maximize your safety.

Severe Weather: If you receive an OU Alert to seek refuge or hear a tornado siren that signals severe weather.

1. Look for severe weather refuge location maps located inside most OU buildings near the entrances.
2. Seek refuge inside a building. Do not leave one building to seek shelter in another building that you deem safer. If outside, get into the nearest building.
3. Go to the building's severe weather refuge location. If you do not know where that is, go to the lowest level possible and seek refuge in an innermost room. Avoid outside doors and windows.
4. Get in, Get Down, Cover Up
5. Wait for official notice to resume normal activities.

Additional Weather Safety Information is available through the Department of Campus Safety.

Armed Subject/Campus Intruder

If you receive an OU Alert to shelter-in-place due to an active shooter or armed intruder situation or you hear what you perceive to be gunshots: 1. Avoid: If you believe you can get out of the area WITHOUT encountering the armed individual, move quickly towards the nearest building exit, move away from the building, and call 911. 2. Deny: If you cannot flee, move to an area that can be locked or barricaded, turn off lights, silence devices, spread out, and formulate a plan of attack if the shooter enters the room. 3. Defend: As a last resort fight to defend yourself. For more information, visit OU's Active Shooter page (https://ou.edu/police/psafe/active-shooter-training). Shots Fired on Campus Procedure – Video: https://www.youtube.com/watch?v=BsEOhGJIdI8

Fire Alarm/General Emergency

If you receive an OU Alert that there is danger inside or near the building, or the fire alarm inside the building activates: 1. LEAVE the building. Do not use the elevators. 2. KNOW at least two building exits 3. ASSIST those that may need help 4. PROCEED to the emergency assembly area 5 ONCE safely outside, NOTIFY first responders of anyone that may still be inside building due to mobility issues. 6. WAIT for official notice before attempting to re-enter the building. OU Fire Safety on Campus: https://vimeo.com/125093634

Mental Health Support Services

If you are experiencing any mental health issues that are impacting your academic performance, counseling is available at the University Counseling Center (UCC). The Center is located on the second floor of the Goddard Health Center, at 620 Elm Rm. 201, Norman, OK 73019. To schedule an appointment call (405) 325-2911. For more information, please visit University Counseling Center (https://www.ou.edu/ucc).

Pre-Finals Week Policies

During pre-finals week, all normal class activities will continue; however, no assignment, test, or examination accounting for more than 3% of the course grade may be assigned, unless it is assigned in advance of pre-finals week and worth less than 10%, or scheduled at least 30 days prior if worth more than 10%. No activity or field trip may be scheduled that conflicts with another class. There are some exceptions and nuances, so please review the Final Exam Policies (https://www.ou.edu/registrar/academic-records/academic-calendars/final-exam-schedule/final-exam-policies) prior to designing your course schedule.

Acknowledgements

I would like to thank, in alphabetical order, Professors Le Gruenwald, Dean Hougen, Amy McGovern, and Deborah Trytten, for providing valuable feedback on a preliminary version of the syllabus when I first arrived at the University of Oklahoma and proposed this new course.

References

[1] Solon Barocas, Moritz Hardt, and Arvind Narayanan. *Fairness and Machine Learning: Limitations and Opportunities.* fairmlbook.org, 2019. `http://www.fairmlbook.org`.

[2] Pedro Domingos. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, Inc., USA, 2018. ISBN 0465094279.

[3] Michael J. Kearns and Umesh V. Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, 1994. ISBN 0-262-11193-4.

[4] Tom M. Mitchell. *Machine Learning*. McGraw Hill Series in Computer Science. McGraw-Hill, 1997. ISBN 978-0-07-042807-2.

[5] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. Adaptive computation and machine learning. MIT Press, 2012.

[6] Christoph Molnar. *Interpretable Machine Learning*. Independently Published, 2 edition, 2022. URL https://christophm.github.io/interpretable-ml-book.

[7] Cathy O'Neil. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books, 2017.

[8] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, New York, NY, USA, 2014. ISBN 1107057132, 9781107057135.

[9] Leslie Valiant. *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*. Basic Books, Inc., New York, NY, USA, 2013.

[10] Kush R. Varshney. *Trustworthy Machine Learning*. Independently Published, Chappaqua, NY, USA, 2022.