



Course Syllabus: Special Topic in Artificial Intelligence - CS 290E

| | |
|--|--|
| Division | Computer, Electrical and Mathematical Sciences & Engineering |
| Course Number | CS 290E |
| Course Title | Special Topic in Artificial Intelligence |
| Academic Semester | Fall |
| Academic Year | 2019/2020 |
| Semester Start Date | 08/25/2019 |
| Semester End Date | 12/10/2019 |
| Class Schedule (Days & Time) | 01:00 PM - 02:30 PM Mon Thu |

Instructor(s)

| Name | Email | Phone | Office Location | Office Hours |
|--------------|--------------------|---------------|----------------------------------|----------------------|
| Marco Canini | marco@kaust.edu.sa | +966128080489 | 4418, 1, AI-Khawarizmi (bldg. 1) | By appointment only. |

Teaching Assistant(s)

| Name | Email |
|------|-------|
| TBA | TBA |

Course Information

| | |
|--|---|
| Comprehensive Course Description | The recent successes of AI/ML owe a great debt to the rapid innovations in hardware and software systems and to the broad accessibility to this technology. These systems have enabled training increasingly complex models on ever larger datasets. In the process, these systems have also simplified model development, enabling the rapid growth in the machine learning community. On the one hand, innovations in AI/ML have the potential to improve the systems infrastructure itself by applying AI to improve system design, job scheduling, video streaming or simplify network management. This class covers a broad array of research topics and latest trends in systems designs to better support the next generation of AI applications, and applications of AI to optimize the architecture and the performance of systems. |
| Course Description from Program Guide | |
| Goals and Objectives | <p>The goal of this class is to cover a broad array of research topics and latest trends in systems designs to better support the next generation of AI applications, and applications of AI to optimize the architecture and the performance of systems. The format of this course will be a mix of lectures, seminar-style discussions, and student presentations. Students will be responsible for paper readings, and completing a hands-on project. Readings will be selected from recent conference proceedings and journals. Suggested project ideas will be provided by the instructor, but students are strongly encouraged to come up with their own project ideas.</p> <p>On completion of this course students should have a broad understanding of some key papers and concepts in research at the intersection of computer systems and AI/ML, as well as an appreciation of how to argue for or against any particular idea. By the end of the course, students should be able to:</p> <ul style="list-style-type: none"> -Comprehend and critique relevant research papers in the area of “systems for AI” and “AI for systems”. -Present their research both orally in a concise way and within the allotted time as well as in writing. -Defend the research approach, design decisions, and the evaluation methods in a discussion. -Moderate a discussion after a research presentation. |

| | |
|----------------------------------|--|
| Required Knowledge | Students are expected to have good programming skills and must have taken at least one systems-related course (e.g., from operating systems/CS140, distributed systems/CS240, networking/CS244, or databases/CS245), and one AI/ML-related course (e.g., machine learning/CS229) or instructor consent, is required. Students must have experience using a machine learning framework such as TensorFlow, PyTorch, Ray, etc. |
| Reference Texts | There is no textbook for the course. You will be expected to read 1-2 papers per class. The papers are available electronically through the course website. |
| Method of evaluation | 15.00% - Attendance and Participation 20.00% - Scientific review article presentation 50.00% - Research Project 15.00% - Presentation |
| Nature of the assignments | <p>The class consists of two main activities:</p> <ol style="list-style-type: none"> 1. Papers and discussions: We will read 2-4 papers per week, and discuss them in class. You will need to spend about 2 hours reading each paper, and making notes, to prepare you to discuss them in class. This only works if you come to class ready to discuss the papers in detail, which is why 15% of your grade is for in-class participation. There will be as much as 10-15 hours of reading per week. Do not take this course unless you are willing and able to do a lot of reading. 2. Project: You will do a semester-long open-ended project. |

Course Policies

Groups

It is highly recommended that you work in groups. Teams are suggested to be 2-3 students.

Paper discussions

The course is organized around topics in ML/AI systems, each of which generally comprises of an overview lecture by the instructor and a discussion section. During the discussion, we will dissect the papers' key ideas, approaches, contributions, positioning against related work, and evaluation methods and results. It is expected that students would have read and submitted their paper reviews in advance. To organize the discussion, students will play different roles and rotate through these.

-Reviewers: every student will review each paper (see below)

-Detailed Reviewers: three students will write detailed reviews; these are longer reviews that carry a substantially greater level of detail.

-Area Chair: one student per topic serves as the area chair; this student collects the reviews by the day before discussion at 1:00PM and summarizes them for everyone's benefit. They posts the summary publicly on Piazza by 8:00PM the day before discussion.

-Champion: the role of the Champion is to advocate in favor of the paper; the Champion is chosen by the Area Chair in agreement with the instructor based on the reviews; there will be one Champion for each paper.

-Detractor: the role of the Detractor is to argument against the paper; the Detractor is also chosen by the Area Chair as the Champion; there will be one Detractor for each paper.

-During the discussion, the purpose of the Area Chair's summary will be to capture the paper's key technical idea and summarize the reviews.

The Area Chair ideally should create a written summary using the following structure:

1. This paper solves problem X using method Y, demonstrated with {proofs, models, simulations, experiments, bold assertions, etc.}.
2. Three reasons to accept it are x, y and z. (Should be from reviews)
3. Three concerns that came up about it are x, y, and z. (Should be from reviews)
4. My personal recommendation is {accept, reject, not sure}.

In class, the Area Chair will lead the paper discussion. They will start by presenting their summary of the paper (taking up just a few minutes; it's not a formal paper presentation), then calling up to the Champion who will champion the paper, and then calling up to the Detractor to hear the concerns. Then they will enlarge the discussion to involve other Reviewers and moderate the discussion.

Paper reviews

Read all the papers of each day carefully and write your own paper review. Being able to critically judge others' work is crucial for your understanding. Each critique should be about a half a page and should roughly cover the following questions:

-What is the problem addressed by the paper? Is the problem real? Why is this problem important?

-What is the hypothesis of the work?

-What is the proposed solution's main idea, and what key insight guides their solution? Why is the solution different from previous work?

› Is problem new?

› Are assumptions different?

› Is workload different?

-Does the paper (or do **you**) identify any fundamental/hard trade-offs?

-What is one (or more) drawback or limitation of the proposal, and how will you improve it?

-Do you think the work will be influential in 10 years?

› Why or why not?

The paper reviews of each day must be submitted electronically at least 24 hours prior to the corresponding class. Late reviews will not be counted. You can miss up to 4 paper reviews during the term without penalty. Each missing review beyond the fourth one will result in 25% decrease in grade for paper reviews. Meaning, failing to turn in 8 or more reviews on time will result in a zero for all paper reviews. The grade for this part will be assigned as follows: at the end of the term, 2 of your reviews will be randomly selected and graded. The higher grade of the two will be used for grading.

Detailed Reviewers

When a student is assigned to do a detailed review, the critique will follow the format of an actual review and is expected to enter the paper is much more depth. Every student will do 2 or 3 detailed reviews.

Submission

The paper reviews of each day must be submitted electronically at least 24 hours prior to the corresponding class. **Late reviews will not be counted.** You can miss up to 4 paper regular reviews during the term without penalty. Each missing review beyond the fourth one will result in 25% decrease in grade for paper reviews. Meaning, failing to turn in 8 or more reviews on time will result in a zero for all paper reviews.

Grading by peers

The Area Chairs, Champions and Detractors identify the top 3 regular reviews on each paper and submit this information privately to the instructors.

Grading

At the end of term, the grade for this part will be assigned as follows:

1. two of your reviews (not including detailed) will be randomly selected and graded. The higher grade of the two will be used for grading.
2. one of your detailed reviews chosen at random will be graded.

The final grade will be the average between the two parts above.

Additional, the final grade will be influenced by the grading by peers of your reviews during the semester. Specifically, being graded as a top 3 reviewer increases this part of the grade as follows: * by 2 peers awards 1% * by 4 peers awards 2% * ... * by 10 peers or more awards 5%

Participation

Is class participation based solely on attendance? No. Attendance is a necessary but not sufficient condition for good class participation. You are expected to attend all lectures (you may skip up to 2 lectures due to legitimate reasons), and more importantly, participate in class discussions. If you have any concerns about not being able to regularly attend class (e.g., you will have to miss several classes during the quarter) please discuss this as soon as possible with the instructor. Beyond attendance, we evaluate class participation by observing how prepared students are to discuss the covered paper when they come to class. This is not a trivial requirement because we expect papers to have been read thoroughly prior to lecture.

Project

Research projects are the critical aspect of the course. Your goal is to do publishable quality ML systems research on an instructor-approved problem; that is, to add to our understanding of how to build systems. Research projects must be written up in a term paper.

Exam

This class has no exam.

Late work

There is no policy on late work. If you cannot submit your work by the deadline, it will not be accepted.

Exceptional circumstances

If you have a medical emergency, then email the instructor to request an extension. At times, there may be system problems with the computer infrastructure which may prevent you from submitting your reviews on time. In these cases an extension will be granted to the entire class.

Misconduct policy

We will have zero tolerance for academic misconduct. Cheating, plagiarism, and any form of dishonesty will be handled with maximum severity, according to university regulations. If you are ever in doubt about whether an action on your part may constitute unacceptable behavior, please ask the instructor before proceeding—doing so afterward is too late. Apart from the term project, any work you turn in must be your own and is to be done individually, and the usual code of conduct applies. You must acknowledge any sources of your words, ideas, and software when they are not your own, and you must disclose in advance, without any specific request, any sources you used. Do not use code from a student who took the course in a previous semester.

Additional Information

For a more up to date version of this syllabus and the actual lecture schedule, you are required to see the version of the syllabus on the class website. The class website will be announced on the first lecture. An online discussion forum will be offered on the Piazza system. This instructor does not use Blackboard.

Tentative Course Schedule

(Time, topic/emphasis & resources)

| Week | Lectures | Topic |
|-------------|----------------------------------|--|
| 1 | Mon 08/26/2019 Thu 08/29/2019 | Introduction and Course Overview ML Systems Research |
| 2 | Mon 09/02/2019 Thu 09/05/2019 | Background Review |
| 3 | Mon 09/09/2019 Thu 09/12/2019 | Machine Learning Platforms |
| 4 | Mon 09/16/2019 Thu 09/19/2019 | Model Development and Training Frameworks |
| 5 | Mon 09/23/2019 Thu 09/26/2019 | Saudi National Day Prediction Serving |
| 6 | Mon 09/30/2019 Thu 10/03/2019 | Hyperparameter Tuning and AutoML |
| 7 | Mon 10/07/2019 Thu 10/10/2019 | Distributed Model Training |
| 8 | Mon 10/14/2019 Thu 10/17/2019 | Model Compilation |
| 9 | Mon 10/21/2019 Thu 10/24/2019 | Hardware Acceleration for Machine Learning Mid-semester Project Presentations |
| 10 | Mon 10/28/2019 Thu 10/31/2019 | Mid-semester break Debugging and Interpretability |
| 11 | Mon 11/04/2019 Thu 11/07/2019 | Debugging and interpretability |
| 12 | Mon 11/11/2019 Thu 11/14/2019 | Testing and Verification |
| 13 | Mon 11/18/2019 Thu 11/21/2019 | Machine Learning Applied to Systems |
| 14 | Mon 11/25/2019 Thu 11/28/2019 | Machine Learning Applied to Networking |
| 15 | Mon 12/02/2019 Thu 12/05/2019 | Project Presentations |
| 16 | Mon 12/09/2019 | Exams |

Note

The instructor reserves the right to make changes to this syllabus as necessary.