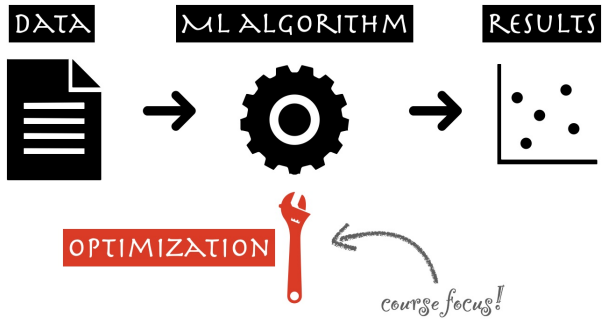
ECGR 6090/8090: Optimization for Machine Learning

Instructor: Prof. Ahmed Arafa, Fall 2023 (TR: 10:00–11:15 am, EPIC 2220)

Suppose we are training a machine learning (ML) algorithm to solve an optimization problem, e.g., training a classifier on a certain dataset.

- *Can we provide a guarantee that such algorithm gives a good solution?*
- *If so, how good is "good"?*
- *Is there a way to quantify how fast the algorithm converges to a required performance?*

If you are interested in finding out answers for these (and related) questions,
$\Rightarrow$ then this course is for you.



What does the course provide?

This course provides an overview of optimization methods used in modern ML systems. By the end of the course, students will be able to (a) assess the performance of popular ML algorithms on various function (model) classes, (b) accurately implement *optimized* ML algorithms for different applications, and (c) characterize time-accuracy tradeoffs for ML training.

What are the course contents?

Topics include: convex optimization, (stochastic) gradient methods, second-order methods, non-convex optimization, distributed optimization, and federated learning.

Grading:

- homework (30%): consisting of theoretical + practical (implementation) components
- exam (20%): after about 2/3 of the semester
- project (40%): in groups, to explore how the studied algorithms and methods perform in real-world example settings
- class participation (10%): attending and engaging in discussions

Are there any prereqs?

Good command of linear algebra (there will be a review), calculus, and programming (e.g., Python). Suggested coreq: *ECGR 4115/5115: Convex Optimization and AI Applications*, offered by Prof. Ahmed Arafa, (TR: 1:00–2:15 pm, EPIC 2220).