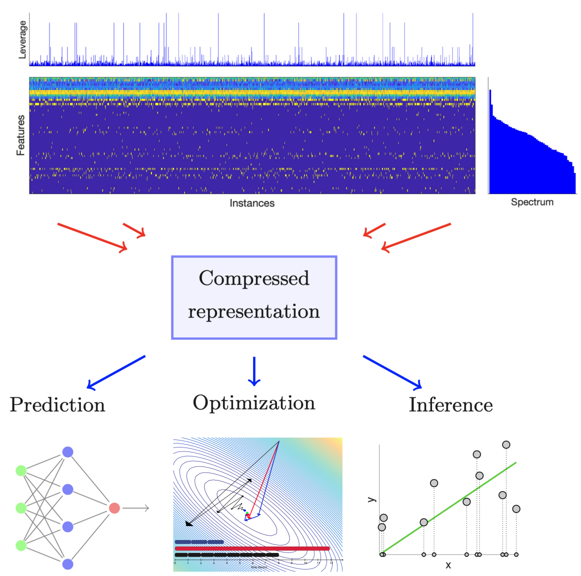
COLLEGE OF ENGINEERING
COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

Algorithms for Machine Learning and Data Science

Instructor: ***Michał Dereziński*** **EECS 498-005/598-005 Special Topics, Fall 2024**

Overview:
This course will introduce algorithmic and theoretical foundations of data science, with a focus on applications to machine learning. With the emergence of machine learning and data science, as well as the ever-increasing data sizes, providing theoretical foundations for the computational complexity of data manipulation algorithms is becoming increasingly important. The course will cover several important algorithms which use randomization and linear algebra to construct compressed data representations and efficiently perform prediction, optimization and inference tasks for big data and ML. We will introduce randomized dimensionality reduction, sketching, and stochastic optimization, among others, and provide the tools needed to design and analyze algorithms that rely on these concepts.

Algorithmic paradigms for big data



The course will delve into topics in linear algebra, such as matrix multiplication and singular value decomposition, as well as in probability, such as expectation, independence and concentration of random variables.

Prerequisites: Linear Algebra (MATH 214 or 217 or 296 or 417 or 419) and Probability (STATS 250 or 280 or 412 or 425 or MATH 425 or EECS 301 or IOE 265)

A tentative list of topics:
Randomized methods for big data (dimensionality reduction, sketching), computational linear algebra (fast matrix multiplication, singular value decomposition, linear regression, principal component analysis), convex optimization (gradient descent, Newton's method, stochastic gradient), and spectral methods (spectral graph theory and spectral clustering).