CSE 291 Syllabus

Winter 2020

**Lecture:** Tuesday, Thursday 12:30-1:50 EBU3B 4140
**Course Webpage:** [http://cseweb.ucsd.edu/~dakane/CSE291/](http://cseweb.ucsd.edu/~dakane/CSE291/)

**Professor:** Daniel Kane
Email: dakane "at" ucsd.edu
Office Hours: Thursdays 2:00-3:00pm in CSE 4212 or by appointment.

**Course Description:** CSE 291 will focus on covering recent results in computational statistics and machine learning, focusing on problems relating to the learning and testing of distributions.

**Prerequisites:** I will try to make this course accessible to anyone with a solid background in linear algebra (including vector spaces and subspaces, inner products, eigenvalues, positive definite matrices and the spectral theorem for self-adjoint operators), probability theory (including probability distributions, Poisson and Gaussian distributions, expectation, variance, Chebyshev and Chernoff bounds) and algorithms. That said, the course will require a fairly high level of mathematical sophistication and will move fairly quickly.

**Textbook:** There will be no textbook for this course, but relevant readings will be linked to from the course webpage.

**Grades:**

*Scribe Notes:* All students taking the course for a grade will be asked to produce scribe notes for one lecture (I will try to arrange a schedule for this early in the quarter). These should be clean **\LaTeX** notes of the material covered in class that day. Please submit your scribe notes no later than a week after the lecture for which the scribe notes are about.

You may submit scribe notes jointly with a partner, but if you do, my standards for grading it will be higher.

*Homeworks:* Students taking the class for 4 units will also be asked to complete a couple of homework assignments. These will likely be short.

*Reading Project:* Students taking the class for 4 units will also be asked to complete a reading project in which they read some recent research paper in the area and write a short report on it. Students should arrange a topic with me by February 7th, and should submit a first draft by March 3rd.

Students signed up to take the course for 1 unit will just be required to do scribe notes. Students signed up for 4 units will be expected to do scribe notes, homeworks and the reading
project. Please make sure that you are signed up for either 1 unit or 4 units, as there will not be intermediate requirements for those signed up for 2 units.

**Schedule:** Below is a rough schedule for topics covered in the class (some may be skipped depending on timing):

*Preliminaries:* Basic setup for distribution learning and testing, basic information theory.

*Unstructured distributions:* Learning complexity, $L^2$ testing algorithm, identity testing, closeness testing, instance optimal testing, independence testing, robust testing, lower bounds.

*One dimensional families:* VC-dimension, $A_k$-distance, effective dimension of standard families, testing algorithms for $A_k$-metric, covers and metric entropy.

*High dimensional families:* Product distributions, Gaussian distributions, graphical models, mixtures of Gaussians and products, robust statistics, SQ-algorithms and lower bounds.