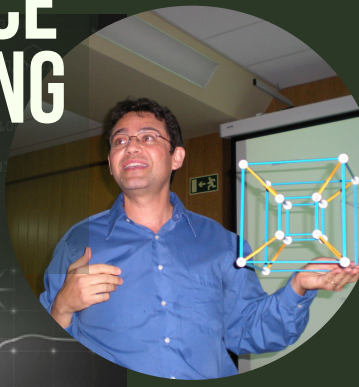
CAL POLY HUMBOLDT
MATHEMATICS DEPARTMENT
COLLOQUIUM SERIES PRESENTS

BSS 166
APR. 07, 2022
4PM

A GEOMETER'S VIEW OF STATISTICAL INFERENCE AND MACHINE LEARNING

DR. JESÚS DE LOERA, UC DAVIS



In statistical inference, we wish to learn the properties or parameters of distribution through sufficiently many samples. A famous example is logistic regression, a popular non-linear model in multivariate statistics and supervised learning.

Similarly, for unsupervised machine learning, one wishes to uncover the shape and patterns from samples of a distribution. We use only the intrinsic geometry and topology of the sample. A famous example of this type of method is the k-means clustering algorithm. A fascinating challenge is to explain the variability of behavior of k-means algorithms with distinct random initializations and the shapes of the clusters.

In this talk, we explain the relation of inference and learning to combinatorial geometry. Along the way, we will see fascinating connections to the coupon collector problem, topological data analysis, and the computation of Tukey centerpoints of data clouds (a high-dimensional generalization of median). All new theorems are joint work with T. Hogan, R. D. Oliveros, E. Jaramillo-Rodriguez, and A. Torres-Hernandez.

Members of the community are invited to attend this free lecture. In order to ensure the continued safety of our campus community, Cal Poly Humboldt asks all visitors to wear masks. Please be prepared to present a vaccination card with proof of vaccination or a negative COVID-19 test result from the last 72 hours, if requested by Humboldt staff.