

The Institute of Mathematics, Leiden University, and the Dept. of Medical Statistics and Bioinformatics, LUMC, announce a Statistical Science Seminar on Wednesday afternoon, March 2, at 15:30 pm, at the LUMC Building 2, Lecture Room 3, Einthovenweg 20, Leiden. Drinks afterwards at 17:00.

JEROME FRIEDMAN
Stanford University

Predictive Learning via Rule Ensembles

Abstract:

General regression and classification models are constructed as linear combinations of simple rules derived from the data. Each rule consists of a conjunction of a small number of simple statements concerning the values of individual input variables. These rule ensembles are shown to produce predictive accuracy comparable to the best methods. However, their principal advantage lies in interpretation. Because of its simple form, each rule is easy to understand, as is its influence on individual predictions, selected subsets of predictions, or globally over the entire space of joint input variable values. Similarly, the degree of relevance of the respective input variables can be assessed globally, locally in different regions of the input space, or at individual prediction points. Techniques are presented for automatically identifying those variables that are involved in interactions with other variables, the strength and degree of those interactions, as well as the identities of the other variables with which they interact. Graphical representations are used to visualize both main and interaction effects.

Pdf of slides: <http://www-stat.stanford.edu/~jhf/talks/wald3.pdf>

Jerome Friedman is plenary speaker at the Dutch Statistics+-OR Meeting on March 3 at the Jaarbeurs (Beatrixgebouw, 4th floor) in Utrecht.

Biographical Sketch

Jerome H. Friedman is Professor Emeritus of Statistics, Stanford University. He received both bachelor's and Ph.D degrees in physics from the University of California, Berkeley. He was leader of the Computation Research Group at the Stanford Linear Accelerator Center from 1972 through 2006. He was Professor of Statistics, Stanford University, from 1982 through 2006, and served as Department Chair from 1988 through 1991. His primary interests center on machine learning and data mining. He has authored or coauthored over 100 papers in major statistical journals as well as three books on Data Mining, and has invented or co-invented several widely used data mining procedures. He has been awarded several honors including member of the National Academy of Sciences, Fellow of the American Academy of Arts and Sciences, Fellow of America Statistical Association, American Statistical Association Statistician of the year(1999), Association for Computing Machinery Data Mining Lifetime Innovation Award, the Emanuel and Carol Parzen Prize for Statistical Innovation, the Institute of Mathematical Statistics Reitz and Wald Lectures, American Statistical Association Noether Senior Lecturer, and paper of the year JASA 1980 and 1985, and Technometrics 1988 and 1992.

For more information, contact: jmeulman@math.leidenuniv.nl