John Elder
elder@dataminginglab.com

Worst

- Short-cuts on sampling / cross-validation
  Using all the data more than we realize
- Focus on algorithms
  to detriment of full ecosystem in which our craft thrives
- Privacy and Security:
  Failure here endangers our whole field!

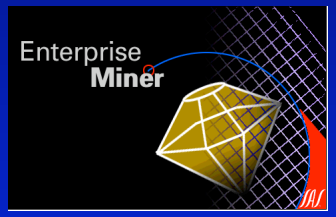
Best

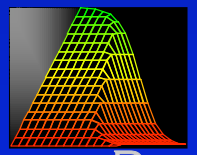
- Scoring: e.g., credit scoring and fraud detection
- Software tools getting better and better
- Recommendation systems:
  even early versions are well-accepted
- Ensembles

"Of course machines can think. After all, humans are just machines made of meat."
- MIT CS professor

Human and computer strengths are more complementary than alike.

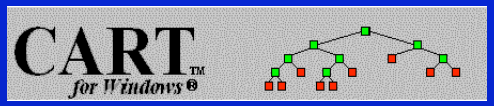
Data Mining Products

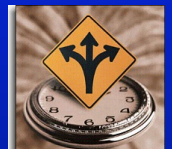


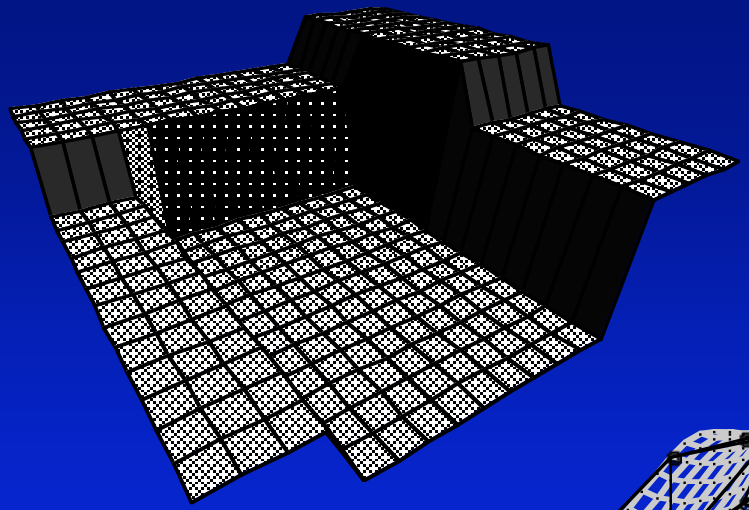


PREDICTIVE DYNAMIX

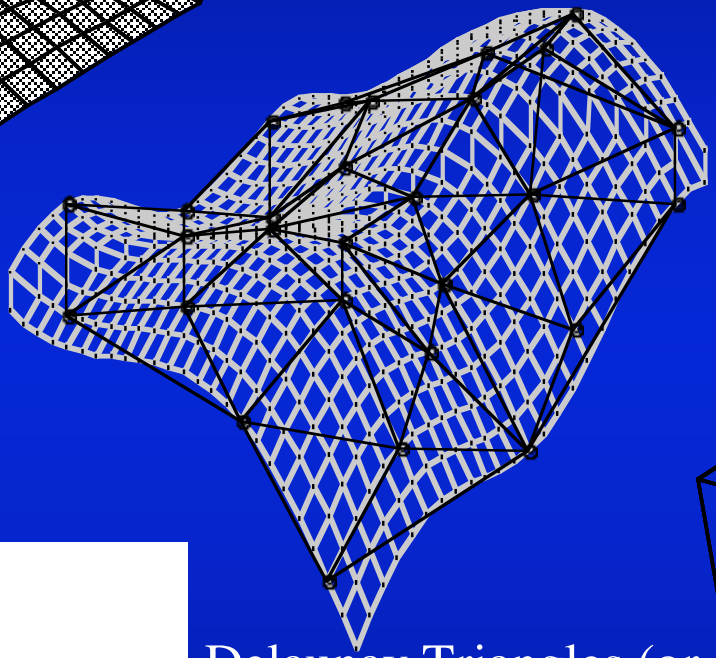
Model 1



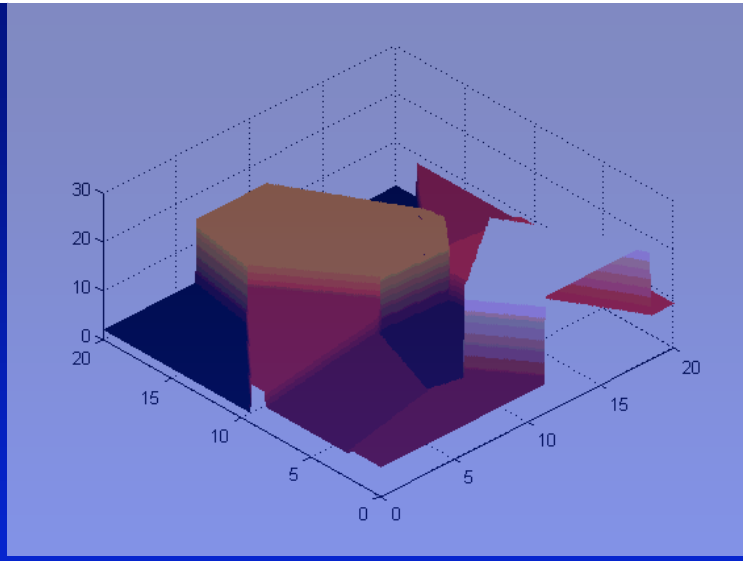




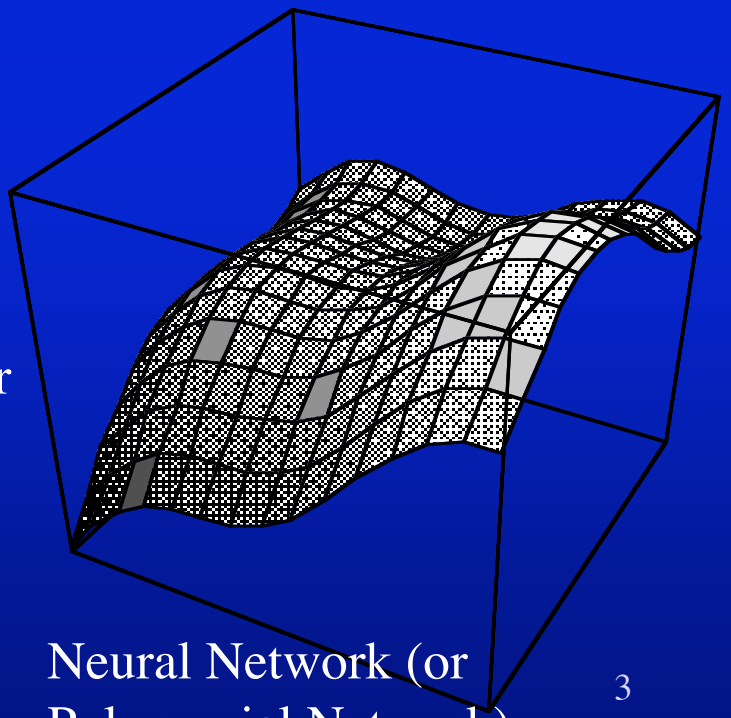
Decision Tree



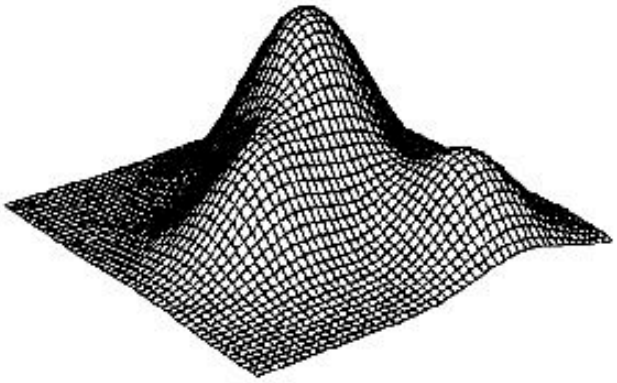
Delaunay Triangles (or Hinging Hyperplanes)



Nearest Neighbor



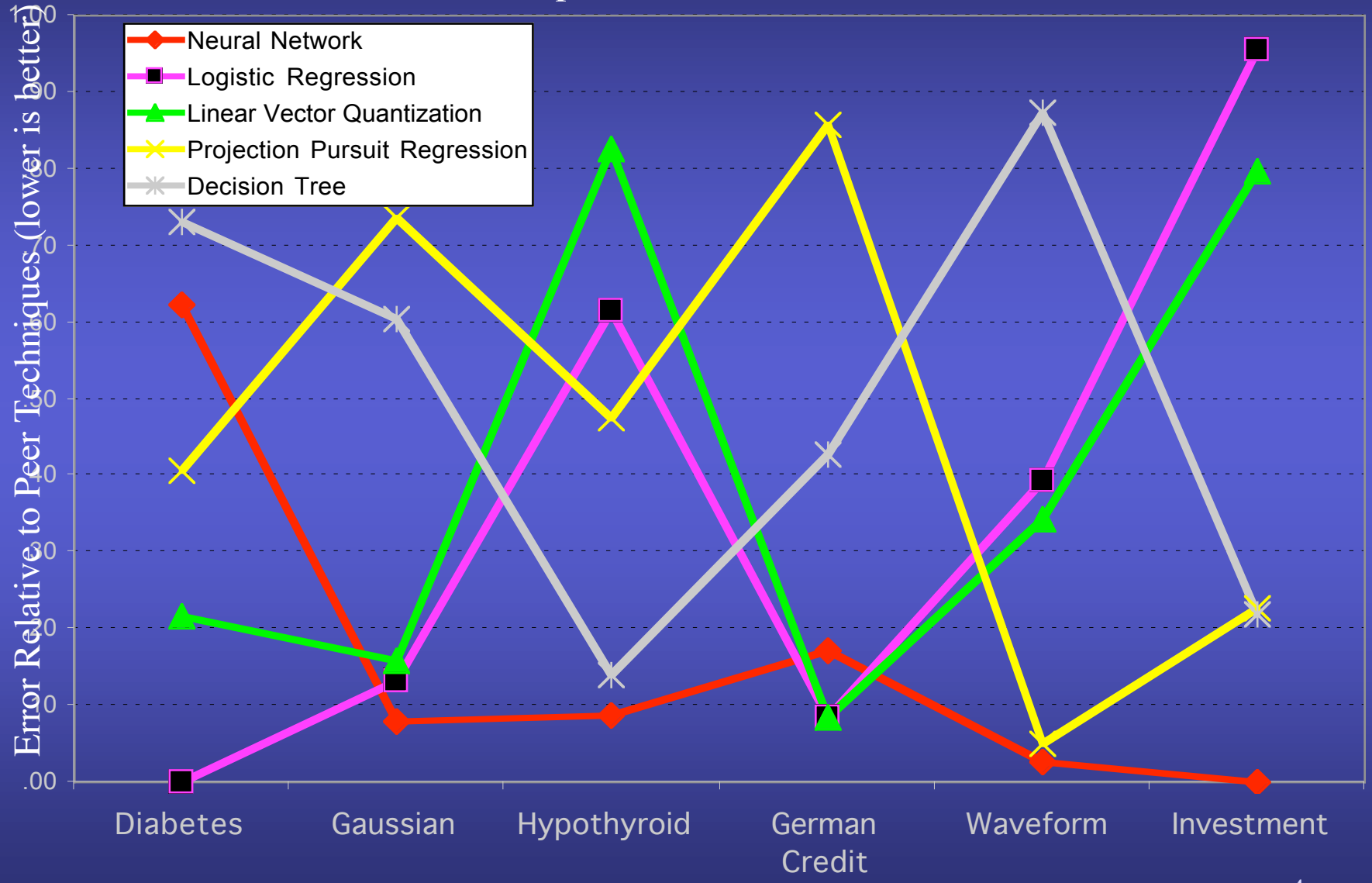
Neural Network (or Polynomial Network)



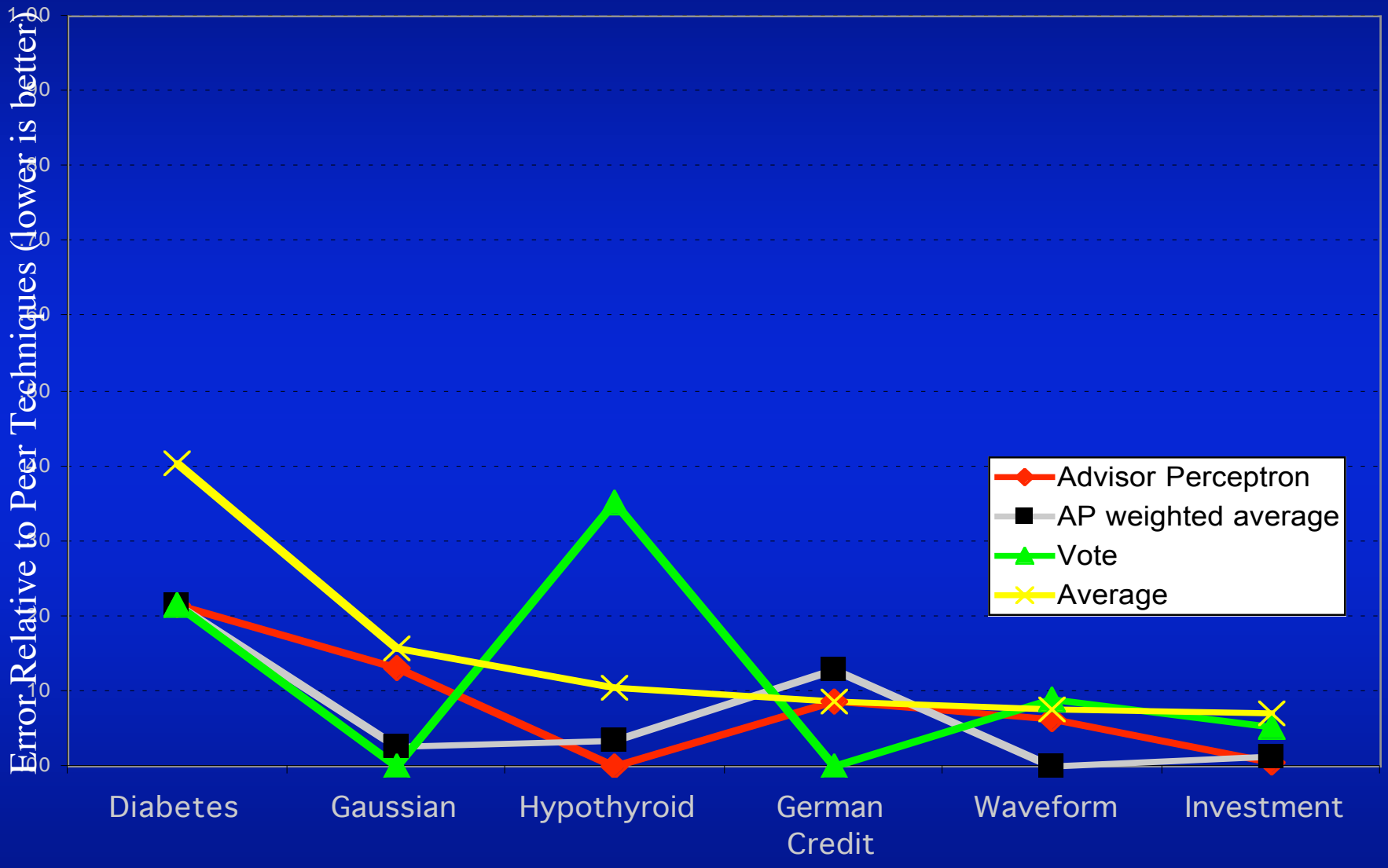
Kernel

Relative Performance Examples: 5 Algorithms on 6 Datasets

(with Stephen Lee, U. Idaho, 1997)



Essentially every Bundling method improves performance



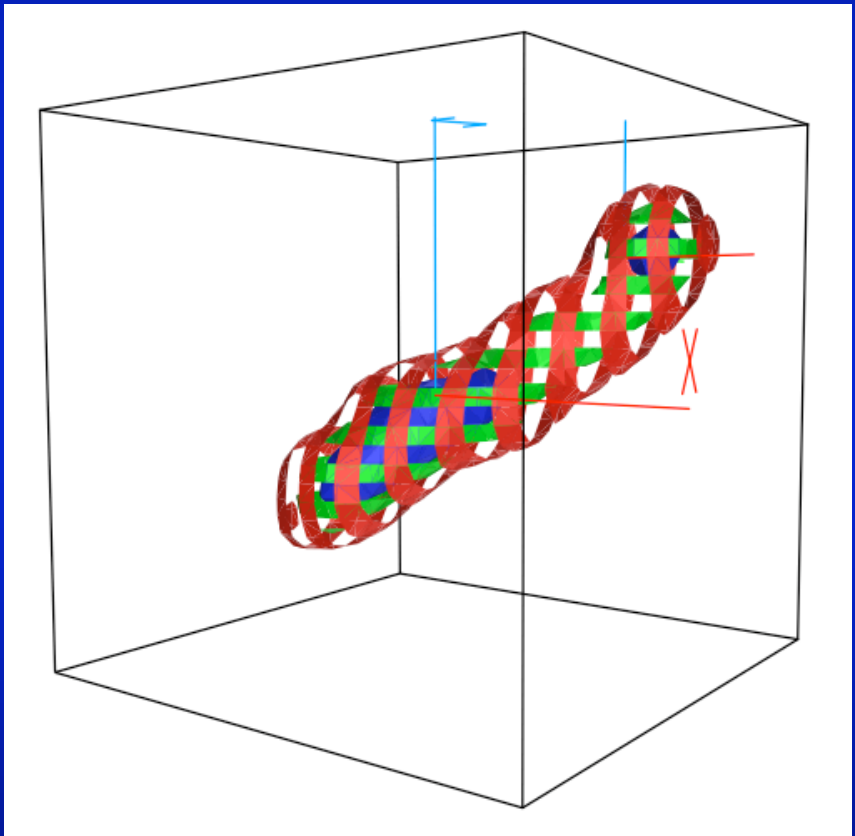
Future

- Visualization
- Medical treatment
  becomes data- and results-driven
- Hope: Data Mining process
  becomes more scientific, repeatable
    - KDD Conference better for industry

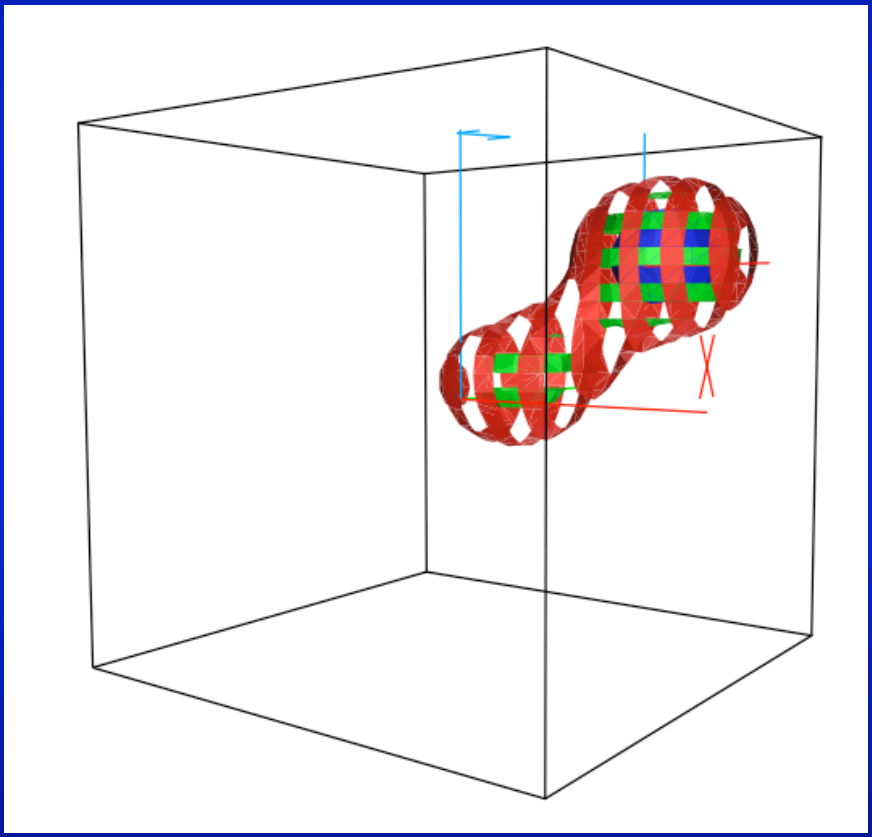
© Hildebrandts

Visualization as Model:
Pharmaceutical application

Placebo



Drug



Density surfaces enclose ascending quartiles of data

Visualization counter-examples…

John F. Elder IV
Chief Scientist, Elder Research, Inc.

Dr. John Elder heads a data mining consulting team with offices in Charlottesville, Virginia and Washington DC (*www.datamininglab.com*). Founded in 1995, Elder Research, Inc. focuses on investment, commercial and security applications of pattern discovery and optimization, including stock selection, image recognition, text mining, process optimization, cross-selling, biometrics, drug efficacy, credit scoring, market timing, and fraud detection.

John obtained a BS and MEE in Electrical Engineering from Rice University, and a PhD in Systems Engineering from the University of Virginia, where he's recently been an adjunct professor, teaching Optimization. Prior to 12 years leading ERI, he spent 5 years in aerospace defense consulting, 4 heading research at an investment management firm, and 2 in Rice's *Computational & Applied Mathematics* department.

Dr. Elder has authored innovative data mining tools, is active on Statistics, Engineering, and Finance conferences and boards, is a frequent keynote conference speaker, and will chair the 2009 *Knowledge Discovery and Data Mining* conference. John's courses on data analysis techniques – taught at dozens of universities, companies, and government labs – are noted for their clarity and effectiveness. For five years, Dr. Elder was honored to serve on a panel appointed by the President to guide technology for National Security.

John is a follower of Christ and the proud father of 5.