



Jerome Friedman's revolutionary new contribution to data mining, an alternative to Neural Nets and Support Vector Machines

TreeNet is Jerome Friedman's latest award winning contribution to data mining. In a broadrange of real world applications from credit risk scoring, targeted marketing, fraud detection, document classification, response modeling, and bioinformatics, TreeNet has been shown to be considerably more accurate than its competitors, and faster and easier to use as well.

TreeNet makes use of a revolutionary new concept of "ultra slow learning" in which layers of information are gradually peeled off to reveal structure in data. TreeNet models are typically composed of hundreds of small trees, each of which contributes just a tiny adjustment to the overall model. In spite of the slow learning strategy TreeNet is blazingly fast and generally produces exceptional models within minutes on today's desktops and servers.

TreeNet automates the entire modeling process and frees the modeler from having to worry about many typical data problems. TreeNet is capable of dealing with dirty and flawed data and is highly resistant to data coding errors, even errors in the target variable. TreeNet automates missing values handling and predictor selection, is impervious to outliers and self tests to prevent over-fitting.

TreeNet helps to automate deployment by exporting model source code usable in major database management and statistical packages. Besides producing high performance models, TreeNet delivers a full range of performance and diagnostic reports and summaries, including gains charts, variable importance rankings, overall classification and prediction accuracy, and graphs plotting the impact of every important predictor on the target.

ABOUT SALFORD SYSTEMS

Salford Systems is an award winning data mining software development and consulting company with a proven record of technical and practical excellence. In 2003 we won the Duke/Teradata Churn modeling competition beating all other entrants in all four award categories. In 2000 we won the KDD2000 data mining competition sponsored by the Association for Computing Machinery (ACM) beating every major data mining software developer in 2 out of 5 categories (including most accurate models). In 1999 we were awarded Japan's Nikkei Prize for demonstrated excellence applying CART to quality control problems.

We have been developing advanced analytical software since 1983, and data mining software since 1990. Working closely with the world's leading data miners at UC Berkley, Stanford University, we are committed to delivering best-of-breed state-of-the-art methodology embedded in first class user assisting environments.

Salford Systems also provides training and consultancy in all phases of data mining and offers an ongoing series of data mining training seminars in the US, Europe, Asia, and Australia.



8880 Rio San Diego Drive, #1045
San Diego, CA 92108
619.543.8880 tel
619.543.8888 fax
info@salford-systems.com
www.salford-systems.com

MARS is a trademark of JerlInc. CART is a registered trademark of California Statistical Software, Inc. and is exclusively licensed by Salford Systems. All other trademarks mentioned herein are the property of their respective owners.
© 1999 by Salford Systems. All rights reserved.



TreeNet APPLICATIONS

Credit Risk Scoring (Probability of Default, Loss Given Default) • Fraud Detection • Targeted Marketing (new customer acquisition, cross-sell, upsell) • Churn Modeling (and related CRM) • Document Classification • Microarray Data Analysis • Genomics, Proteomics • Manufacturing and production line quality control

Features

Scientific Pedigree

Robust and data flow resistant

Missing value handling

Automatic predictor selection

Built-in model validation

Simplified model deployment

Benefits

Based on Jerome Friedman's original proprietary code, licensed exclusively to Salford Systems. Salford continues to extend and enhance TreeNet in collaboration with Dr. Friedman.

TreeNet is capable of detecting true data structure even when presented with highly flawed and dirty data. A major innovation is TreeNet's ability to detect errors in the coding of the target variable, as can often happen in cases of fraud and medical diagnosis. The software is also unaffected by outliers.

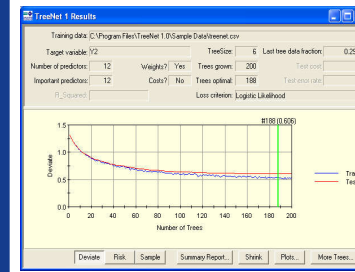
Missing values are handled automatically. There is no need to fix, repair, impute or otherwise deal with missings prior to using the software.

TreeNet can process very large numbers of candidate predictors and rapidly separate those that matter from the rest. Whether your data contains the hundreds or thousands of predictors common in financial services, or the tens of thousands common in bioinformatics, and chemistry, TreeNet will correctly assess the true value of any variable in its importance rankings.

TreeNet contains a number of innovative techniques to prevent overfitting resulting in models that perform accurately when applied to new data. The models are also self-tested internally against a holdout sample and all performance reports are based on hold-out sample results.

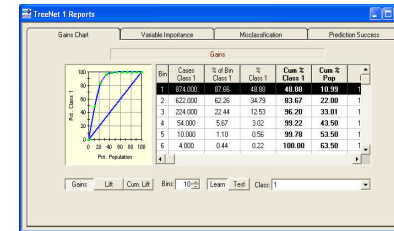
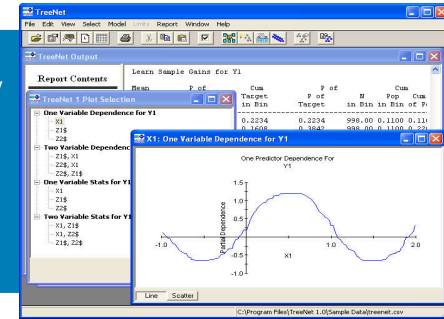
TreeNet will function as a callable scoring engine that can be embedded in a heterogeneous computing environment. In addition TreeNet will optionally output source code in major programming languages, and code compatible with major databases and statistical data processing systems.

SAS and SPSS are trademarks of their respected owners.



TreeNet selects and transforms predictors automatically. The model summary display performance as the number of trees increases.

TreeNet displays how each important predictor affects the target variable.



Summary results include gains charts, variable importance ranking, and confusion matrices for classification problems.

```

M Tree 1 of 200 * /
M % terminal nodes = 6, depth = 4 * /
response = 0.0;
NO_1:
  IF 22 <= ln ("2990")
  then goto NO_2;
  else IF 22 <= ln ("08")
  then goto NO_4;
NO_2:
  IF 4 <= -0.137018 then goto NO_3;
  else goto NO_3;
NO_3:
  response = 0.136414;
  goto DO;
NO_4:
  IF 4 <= 1.07893 then goto NO_5;
  else goto NO_5;
NO_5:
  response = 0.233926;
  goto DO;
NO_6:
  response = 0.348224;
  goto DO;
NO_7:
  IF 21 <= ln ("2990")
  then goto NO_8;
  else IF 22 <= ln ("08")
  then goto NO_9;
  else goto NO_5;
  
```

Export models as source code for major database management systems.

Data Formats

The TreeNet data-translation engine supports data conversion from over 80 file formats, including:

- Statistical analysis packages-SAS® and SPSS
- Databases-Oracle and Informix
- Spreadsheets-Microsoft Excel and Lotus

FREE DEMO
www.salford-systems.com