

*Salford Systems: Data Mining and predictive  
Analytics for Credit Risk*



## *Salford Systems: Company OverView*

- Founded 1983 as a SAS® procedures developers
  - First products were : PROC MLOGIT, PROC MPROBIT
- 2009 Series of products
  - CART® 6.0 in Standard, EX, and PRO versions
  - TreeNet® 2.0 , PRO EX
  - MARS 3.0
  - RandomForests® 1.0
  - Salford Predictive Modeling Suite (SPM integrates all tools above)
  - Client/Server versions of all products
  - SCORING Server: For large scale high speed scoring
- Services: credit risk scorecards, consultation, training

## *Salford Systems: R&D Resources*

- Dan Steinberg, CEO. Ph.D Econometrics, Harvard
- N. Scott Cardell, Director, R&D PhD Econometrics, Harvard
- *Technical Advisers and collaborators*
- Jerome H. Friedman, Stanford University ( CART, MARS, Treenet)
- Leo Breiman\*, UC Berkeley (RandomForests)
- Richard Olshen, Stanford University (Survival CART, Bioinformatics)
- Charles Stone, UC Berkeley (CART, MARS large sample theory)
- Rob Tibshirani , Stanford (modern statistical methods)
- Richard Carson, UC San Diego, Visualization, SuperComputer methods
- \* Leo Breiman passed away in July 2005

## *Salford Systems Honors and Awards: Wins based on Salford Tools*

- 2009 KDDCup 1<sup>st</sup> Runner Up IDAnalytics\*, and FEG Japan\*
- 2008 DMA Direct Marketing Association. 2<sup>nd</sup> Place\*
- 2007 Pacific Asia PAKDD: Credit Card Cross Sell. 1<sup>st</sup> place
- 2006 DMA Direct Marketing Association: Predictive Modeling\*
- 2006 PAKDD Pacific Asia KDD: Telco Customer Type Profiling
- 2005 BI-Cup Latin America: Predictive Modeling E-commerce\*
- 2004 KDDCup: Predictive Modeling ‘Most Accurate’\*
- 2002 NCR/Teradata Duke University: Predictive Modeling-Churn
- 2000 KDDCup: Predictive Modeling- Online behavior
- 2000 KDDCup: Best CRM Analysis
  - \* Winner or honors awarded to a Salford Client
- This record far exceeds that of any other software developer

## *Credit Risk and Related Experience -1*

- Chase Bank: Mortgage Default model (joint with mortgage pre-pay model). System deployed to guide buying and selling of mortgages
- American Express: Bankruptcy prediction model. Objective included creating a bankruptcy score and identification of groups of extreme risk credit card holders (12-month ahead forecasts)
- Corporate Risk models: Alliance Bank, Malaysia. Primarily small and medium sized corporates
- Corporate Risk and Rating Models: Rating Agency Malaysia (RAM) Large Corporates. Default Risk and Rating Prediction

## *Credit Risk and Related Experience -2*

- Corporate Risk including real estate and unique projects: MayBank, Malaysia. Five categories of risk model.
- Consumer Risk: New Auto Loans. Hong Leong Finance, Singapore.
- Consumer Risk: Sub-prime finance of electronics stores purchases. Pacific Retail Finance, New Zealand (GE Money)

## *Key Features of CART, MARS, and TreeNet*

- Very high degree of automation
- Minimal data preparation required (no prep is an option)
- Automatic selection of relevant fields for prediction
  - TreeNet is an especially capable variable selection tool
- Automatic handling of missing values
- CART, TreeNet, robust to outliers and dirty data
- MARS: outliers clearly identified
- Accurate models that hold up on new data
- Optionally controlled by scripts;
  - complete audit trail generated for any interactive session ensuring ability to reproduce analysis
- New automation built into CART 6.0 PRO EX to search for best models

## *Key Advantages of Salford Technology*

- Robust, reliable, capable of handling huge modeling tasks.
  - Used in e-commerce to rebuild models automatically every 24 hours.
- Capable of dealing with substantially flawed data including errors in the target (dependent variable)
- Superb accuracy proven repeatedly in open competitions
- Ultra-High speed performance
- Develop either easily interpretable or more complex more accurate models
- Modeling automation to dramatically accelerate scorecard development
- Patent pending innovations to ensure models are stable and will perform out of sample
- Designed to operate seamlessly in a SAS environment, read, write, deploy
  - Deploy models as SAS, SQL, C, Java, and PMML/XML for in database

## *ROI of Salford Tools*

- In the 2002 Teradata/Duke churn modeling competition Salford models were:
  - 10% better than 2<sup>nd</sup> best model in top decile lift and area under ROC
  - 40% better than median model in top decile lift
- A major credit card issuer has conducted a year long test of TreeNet models for targeted marketing
  - TreeNet models typically 4%-8% better than best in-house models
  - TreeNet models built in less time

## *Select Key Banking Software Customers*



- American Express
- Citibank
- Capital One
- VISA Credit Card
- Bank Negara (Central Bank, Malaysia)
- World Bank

## *Core Banking Applications*



- Credit Risk Scorecard Development (Consumer)
- Model quality assessment (fast review of existing scorecards)
- Collections and Bad Debt Recovery
- Corporate Risk Scorecard Development
- Automatic Ratings Systems (Corporate Bonds)

## *Salford Presentations of Interest*

- Application of TreeNet to reverse engineering of agency ratings of commercial banks in Europe (Study for European Central Bank)
- Small Sample Modeling with TreeNet (Paper presented at 2005 Edinburgh Conference of Credit Risk and Credit Control)
- Interaction Detection with TreeNet (Paper presented at 2009 Edinburgh Conference of Credit Risk and Credit Control)

