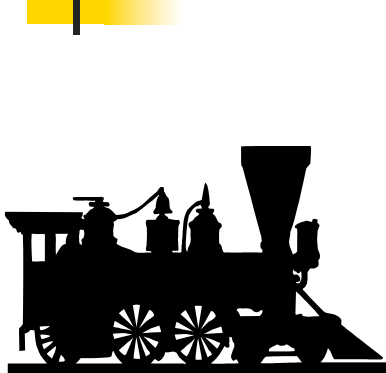


Applying Salford Systems' data mining technology CART® to transportation operations or How do trees and transportation go together?



Bradley Utz

+1 619.543.8880 x104, bradleyu@salford-systems.com



Outline of this presentation

- The transportation industry
- The “transportation company”
- Business processes of the “transportation company”
- Business problems for in-depth analysis
- CART®
- Applying CART® to a typical business problem of the “transportation company”



The transportation industry

- Many national / international companies
- Mobile workforce, globalization, declining barriers to trade
- Once highly regulated, now deregulated in most sectors
- The practice of logistics
 - The science and art of procuring, maintaining, and transporting items



The “transportation company”

- Delivers service for others; manufactures no product
- The “load” -- the property of others, the source of demand, unit of commerce
 - Physical in nature – passengers or freight
 - Requires safe physical handling
- The customer
 - owns the load
 - has inherent goodwill for the transportation company based on its quality of service

The “transportation company”

(cont' d)

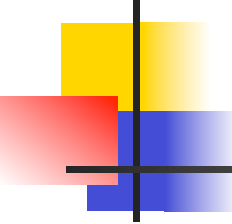
- Operates for profit; has competition
- Runs on a time schedule – responsible for creation, execution and updates
- Acquires, uses, and maintains resources – the workforce, the power unit, the carriage, the fuel
- Works over a geographic network
 - owned or leased or shared
 - Nodes where transport activity is done
 - Directed paths connecting these nodes together tightly or loosely
- Thrives on accurate communications and technology



Business processes of the “transportation company”

- Common to other industries and companies
 - Enterprise strategic planning
 - Financial budgeting and accounting
 - Marketing and sales
 - Customer Service
 - Technology
 - Human Resources and Legal

Business processes of the “transportation company” (cont’ d)



- Differentiator: Operations
 - Fulfilling the demand of the load with the available supply of company resources
 - Allocating the resources
 - Dispatching the resources
 - Training the workforce of operations
 - Moving the load
 - Handling contingent, abnormal events



Business problems for in-depth data analysis

- Goals of any in-depth analysis
 - To explain what happened
 - To predict what may happen
- Technology to apply
 - data mining analysis
 - mathematical modeling for optimization and simulation
 - To show optimal effectiveness
 - To know and understand the business well...better



Business problems for in-depth analysis - examples

(cont' d)

- What are the root causes for the failure to deliver the load to the customer on-time? (metric -- % on-time delivery success)
- Create a set of characteristics of safe crews (metric -- % weekly safety levels)
- Does the seasonality of movements affect the earnings in that customer market; is last quarter highest? (metrics – earnings level, return-on-investment)

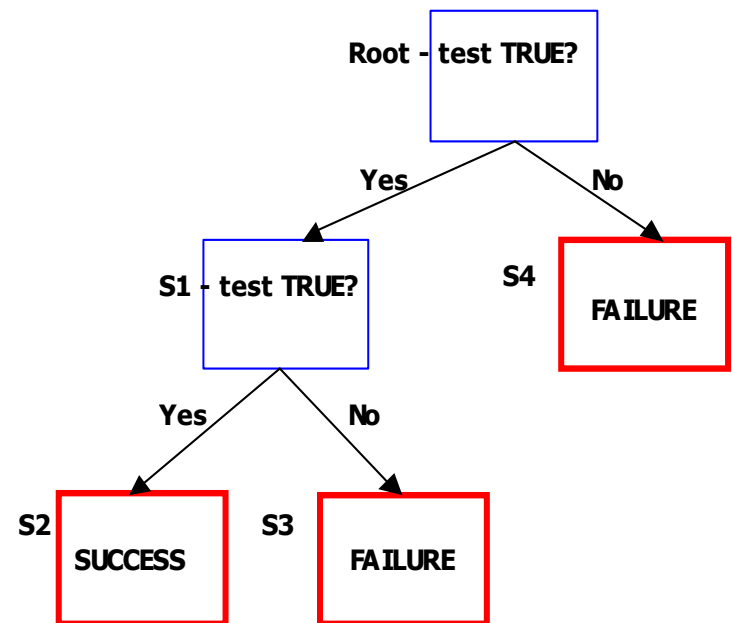
CART® -- Classification and Regression Trees



- Data mining software of Salford Systems based on the CART algorithm by Breiman, Friedman, Olshen, and Stone (1984)
- Classification – to describe separate categories of objects
- Regression – to fit general equations of independent variables to a dependent

CART® -- Classification and Regression Trees (cont' d)

- A classification tree
 - Structure of general knowledge from data
 - Predictor of discrete categories, binary or greater
 - A root node, successor nodes, and branches to successor nodes
 - Successor nodes either terminate or reproduce others in the process



CART® -- Classification and Regression Trees (cont' d)



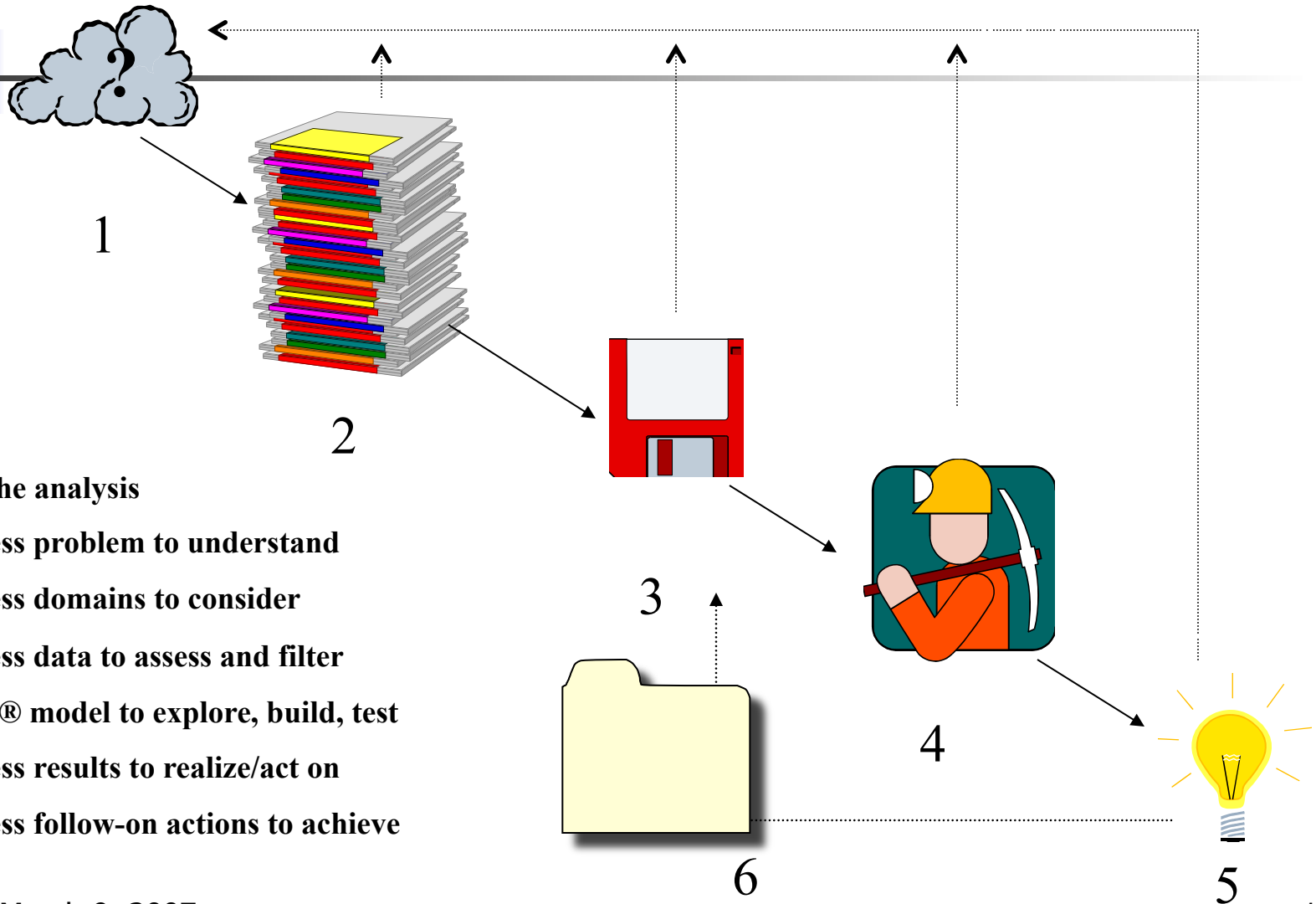
- The process
 - Accepts a large data set of discrete target-class values and independents at the root node
 - Partitions the set successively from root to successor nodes
 - Ranks all possible independents at each split
 - Selects one appropriate, influential independent to split on
 - Stops along a branch when
 - A successor has records of only one target-class value
 - A successor has reached a stopping condition
 - Stops overall when all branch paths stop



Applying CART® to a typical business problem of the “transportation company”

- What are the root causes for a late load arrival at the destination?
 - Root causes for failure involve one or more predictors and their specific values.
 - Dual results – two sets of branches... business rules for failure (“late”) and those for success (“not late”) derive from the same set of predictors.

Applying CART® to a typical business problem of the “transportation company”



Steps in the analysis

1. Business problem to understand
2. Business domains to consider
3. Business data to assess and filter
4. CART® model to explore, build, test
5. Business results to realize/act on
6. Business follow-on actions to achieve

March 9, 2007



Transportation business domains to consider

- Work (Loads and Stops for delivery, pickup, rest, and refuel)
 - Schedule of times and actual times from origin to final destination
 - Count of delivery stops per load
 - Type of load to move
 - Profitability
- Equipment to hold the load (types needed/available)
- Customer requirements and feedback
 - Loading at origin and unloading at destination (yes/no)
 - Hours available / hours open
 - Weekend / weekday delivery
 - Customer Goodwill from start of order onward



Transportation business domains to consider

(cont' d)

- Workforce
 - Concerns: pay, time on duty, and time off duty
 - Geographically based
 - Various levels of tenure and experience
 - Operating in different weather environments
 - Safety profile

- Market Environment
 - Grading the shipment for revenue potential



The CART® model to explore, learn and test

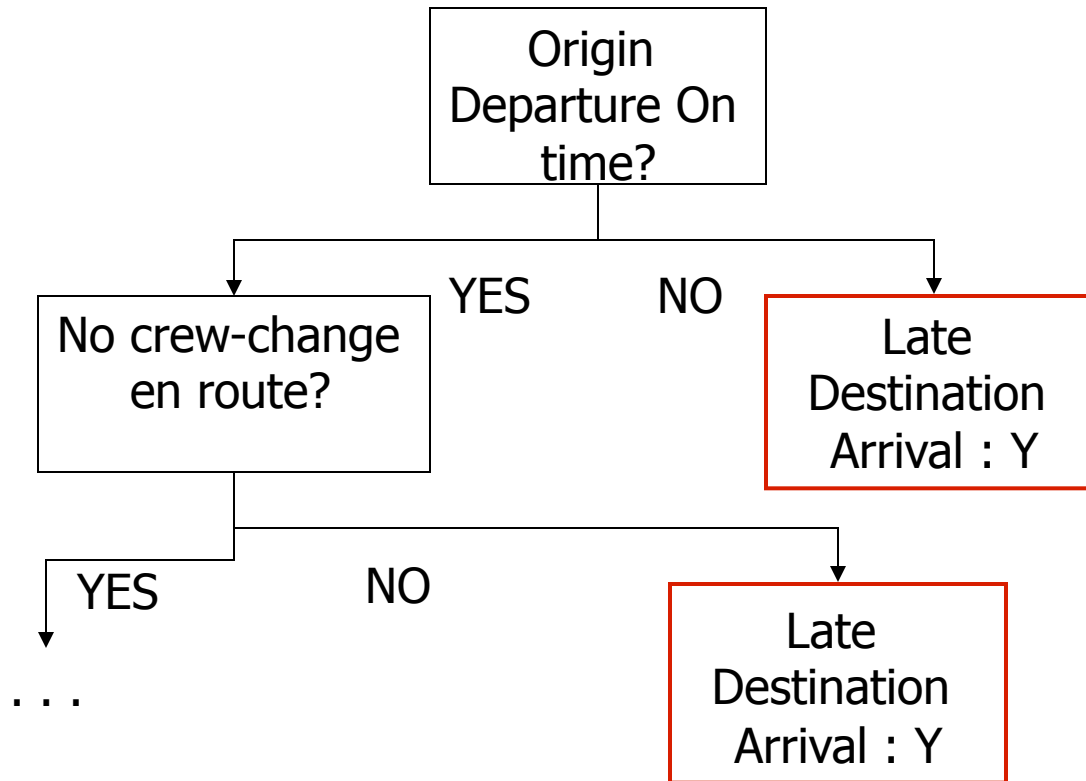
- Business Problem: “What are the root causes for a late load arrival at the destination?”
- Scope:
 - many months of moves throughout the country covering 24 x 7 operations
 - many domains: planning, operations, safety, customer, loads
 - many independent predictors, more categorical than numerical
 - Schedule times, geography, load characteristics, crew descriptors
 - Qualitative or quantitative, extracted and transformed



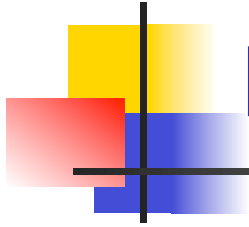
The CART® model to explore, learn and test (cont' d)

- With input ready, let CART divide into Learn and Test subsets (suggested 70%/30%)
- Choose specific set of independent variables rationally
- Set CART®'s process settings reasonably
 - Reduce maximum number of terminal nodes (100 or fewer)
 - Reduce number of successor branches (depth of 8 or fewer)
 - Set standard-error rule setting in range [0.25, 0.5]
 - Set minimum number of node records for splitting
 - For the sake of *rare* target-class values, use PRIORS EQUAL to reduce the influence of more frequent target-class values
- Run CART® and observe the output

Business Results to realize and act on. . . CART® tree excerpt



Business Results to realize and act on . . . Prediction Success and Misclassification Tables

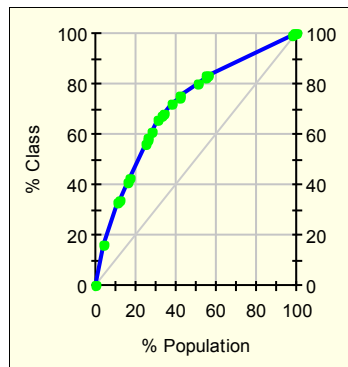


Actual class	Total cases	Percent Correct	N N=161	Y N=183
N	143	97.20	139	4
Y	201	89.05	22	179
Total	344			
Average		93.13		
Overall % correct		92.44		

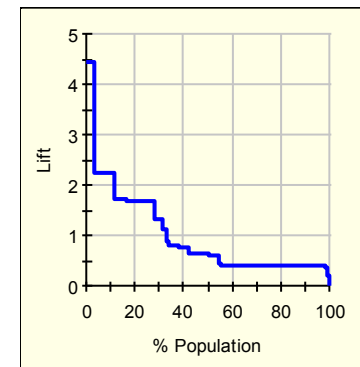
Class	N Cases	N Misclassified	Pct Error	Cost
N	143	4	2.8	0.03
Y	201	22	10.95	0.11

Business Results to realize and act on. . .

- Review the relative importance of the predictors
- Change the mix of independent variables and rerun as needed for better overall accuracy
- Gains possible for Class: Late Destination Arrival = Y



- Lift for Class: Late Destination Arrival = Y



Business Results to realize and act on. . .



- Articulating this business knowledge with confidence
 - e.g., late origin departures lead to late destination arrivals
- Modifying the business process metric
- Re-engineering the business process with new knowledge for improved performance
 - e.g., reduce frequent stops en route
- Investigating relationships to other business processes – finding synergy, leveraging knowledge
- Further investigating the detail level of the independent predictors of this problem for additional phase of analysis
 - e.g., add congestion levels of the geography to the analysis



The Last Slide

Salford Systems

4740 Murphy Canyon Road, Suite 200, San Diego, California 92123

Tel +1.619.543.8880 FAX +1.619.543.8888 www.salford-systems.com