

Child Support Case Management Predictive Analytics



Issues & Objectives

- Child support cases - involving the payment of support by one parent - are ranked on a scale of 1 (full payment) to 5 (no payment)
- The project objective is to predict the collection category of cases based on its past payment pattern and various attributes



Solution

- Multinomial Logistics Regression technique was used to build the predictive models
- Models were developed for 4 major states in the United States.



Challenges

- Extremely large data – Approx. 300,000 cases per month
- Available data and relevant variables differ from state to state. Predictive models built for four states so far
- Collection categories definition



Results

- The accuracy of prediction for a dataset of 12 months was 71–83%.

Sr. No.	Month	Correct Classification
1	Feb-13	83%
2	Mar-13	77%
3	Apr-13	77%
4	May-13	71%
5	Jun-13	80%
6	Jul-13	74%

The project was implemented for 4 major US states. There are two parents in every child support case. One is the Custodial Parent (CP) – the parent who lives with the child the majority of the time and has the primary day-to-day responsibility; the other is the Non-Custodial Parent (NCP) who also has important responsibilities. An aggrieved CP may appeal to the state to enforce child support by the NCP.

Smart Analysis on Bus Transportation System



Issues & Objectives

- Transport regulatory authority in Singapore commissioned a system to:
 - ❖ Automatically discover wrong fare incidents and flag commuter's cards affected by wrong fare charging; and
 - ❖ Detect emerging fault trends in fare collection equipment so that corrective action could be taken in a timely manner



Solution

- Data Storage: Hadoop and MySQL
- Query Tools: Hive and SQL
- Algorithms: rmr (Parallel versions of R) and Java
- Reporting and Dashboards: Pentaho



Challenges

- Large data 15 million transactions per day, which translates to more than 5 billion historical transactions in a year needs to be processed to identify fault patterns and trends
- The data consisted of financial, operations, transit and events data of buses



Results

Robust solution in use for over 2 years allows pro-active rather than reactive maintenance

The screenshot shows a web application interface for 'DIFAULTS - Wrong Fare Occurrences'. The interface includes a navigation menu with 'Home', 'Wrong Fare Occurrences', 'Emerging Fault Trends', 'Reports', and 'System Parameters'. There are also links for 'My Account', 'Settings', 'Help', 'Contact Us', and 'Logout'. The main content area is titled 'Selection Criteria' and contains several input fields: 'Business Date' (From: 15-09-2014, To: 15-09-2014), 'Time duration' (From: 00:00 AM, To: 00:00 PM), 'EZlink Card No' (GLLLGCG), 'Bus Service No' (241,901), 'Bus Service Provider' (All), and 'Failure Mode' (All). A 'Generate Report' button is at the bottom. The right side of the interface shows 'User Selection Criteria was' with the same values entered. Below this is a table with columns: 'Ezlink card No.', 'Transaction type', 'Failure Mode', 'Transaction date & time', 'Mark', 'Bus Trip No', and 'Debit/Type Code'. The table contains several rows of data, including entries for 'Failure to Exit', 'Double Deduction (Entry Follows Entry)', 'Out of sequence Marker Ids', 'Timing Mismatch', and 'Out of sequence Marker Ids'. A blue tooltip is visible over the table, containing links: 'Show other patrons for the same bus service?', 'Show all transactions for this patron for selected bus trip?', 'Show bus operations for current row selection?', and 'Show commuter travel patron for current row selection?'. The table also shows a 'Show 10 entries' dropdown and a '2 3 4 5 ... 4681 Next' pagination control.