Predictive Analytics for Bank Loan Default

CLIENT
Client is a growing Bank that has an attractive portfolio of Loan offerings.

CHALLENGES
Loan Distribution was a principal offering of the Bank. The Bank wanted to reduce the credit risk and the defaults in loan repayment.

TOOLS & TECHNOLOGIES
Apache Flume, Apache SQOOP, Apache Hive, Apache Pig, Cloudera Hadoop, R Studio, Shiny R, dplyr (for data preprocessing)

MSYS VALUE CREATION
MSys Predictive Analytics Approach
- Proposed predictive analytics solution to measure the probability that a debtor will default.
- The objective was to include the score of probability for defaulting as a key component in getting the loans.
- The score was used to as a measure for credit risk of the potential loan customer.
- After analyzing the business problem, our focus was on two model types to measure the credit risk:
  - Logistic Regression
  - Decision Trees

BUSINESS BENEFITS
- The Bank saw the credit risk model as an opportunity to improve the speed of loan processing.
- The loan evaluation and distribution was more scientific than heuristic based.
- The Bank was able to move higher into the technology value chain
- After implementation Bank that would be able to reduce the number of loan defaults

A sample data of Model Scoring and Evaluation to decide which Model to select