DSO 562 Fraud Analytics

Project 3

May 9th, 2021

Transaction Fraud Detection

Team Member

GU, Xinxue (Sue) **SUN, Yang** **WANG, Bufan**

XIE, Binghong **YANG, Siqin** **ZHAO, Zihao**

Table of Content

1. Executive Summary

This project is aimed at developing a machine learning algorithm to identify credit card transaction fraud. Transaction fraud is fraud occurring during the account usage process. Apart from credit card fraud, which is our focus of this project, common transaction fraud includes insurance claim fraud, tax return fraud and account takeover, etc. Credit card account can be compromised by counterfeit card, online hack, card stolen, or other kinds of account takeover.

The United States is one of the countries that suffers the most from credit card fraud. 47% of consumers have experienced card fraud within the past 5 years.[1] Since the introduction of chips (EVM), fraud has migrated from counterfeit to Card not Present (CNP) fraud, such as via internet, phone, and mail. This increasing CNP fraud has caused tremendous financial losses, which reached $7.2 billion in the US by 2020[2] and is still on the rise.

To fight credit card fraud, we developed a fraud scoring model based on real credit transaction data in 2010, purchased from a US government organization. The dataset was modified by Dr. Stephen Coggeshall based on his expert knowledge on fraud analytics. It has 96,753 records and 10 fields.

After implementing data quality check, exploratory data analysis and data cleaning on this dataset, we created 517 variables based on expert knowledge. Then, we selected 30 variables that have the highest predictive power using filter and backwards stepwise selection. To optimize the model performance, we attempted 6 machine learning algorithms including logistic regression, boosted trees, random forests, and neural network. Given the data size is small, we also addressed some issues caused by the sensitivity of the data.

Rigorously tuning hyperparameters, we eventually achieved 84.29% fraud detection rate (FDR) at 5% on training data, 83.33% on testing data and 59.78% on out of time data with a neural network model. With this model and a FDR cut off at 5%, rejecting the transactions with top 5% fraud scores is expected to catch 59.78% of the real fraud and avoid $1,131,600 financial loss annually.,

[1] Source: Aite Group, ACI Worldwide, n = 6169 (2016)
[2] Source: Aite Group

2. Description of Data

The data has 96,753 credit card transaction records, with 9 fields and 1 fraud label. Synthetically generated from real data distribution, the dataset is built by an identity fraud prevention company to reproduce the important univariate and multivariate field distributions of real data.

The dataset consists of 8 categorical fields, 1 numerical field, and 1 datetime field.

Table 2.1 Summary Table for Categorical Fields

Column Name	# of Records	# Missing	% populated	# of Unique Values	Most Common Field Value
Recnum	96,753	0	100	96,753	2047
Cardnum	96,753	0	100	1,645	5142148452
Merchmun	93,378	3,375	96.51	13,091	930090121224
Merch description	96,753	0	100	13,126	GSA-FSS-ADV
Merch state	95,558	1,195	98.76	227	TN
Merch zip	92,097	4,656	95.19	4,567	38118
Transtype	96,753	0	100	4	P
Fraud	96,753	0	100	2	0

Table 2.2 Summary Table for Numerical Field

Column Name	# of Records	% Populated	Mean	Minimum	Maximum	Standard Deviation
Amount	96,753	100	427.89	0.01	3,102,045.53	10,006.14

Table 2.3 Summary Table for Datetime Field

Column Name	# of Records	% Populated	# of Unique Values	Minimum	Maximum	Most Common Field Value
Date	96,753	100	365	2010-01-01	2010-12-31	2010-02-28

We examined each field of the dataset and found some interesting facts below.

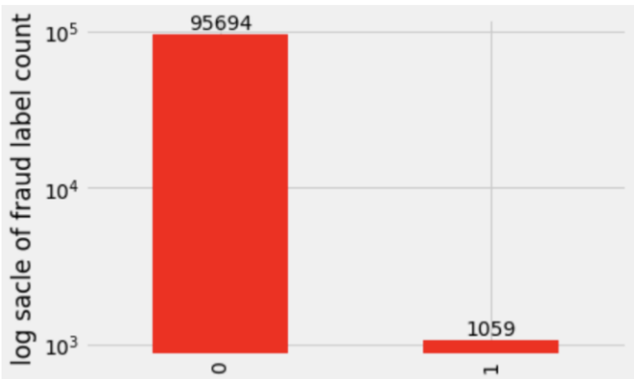


Figure 2.1 Distribution of Fraud Label

In the "Fraud" field, only 1,059 records are labeled as fraudulent, with a proportion of 1.1%.

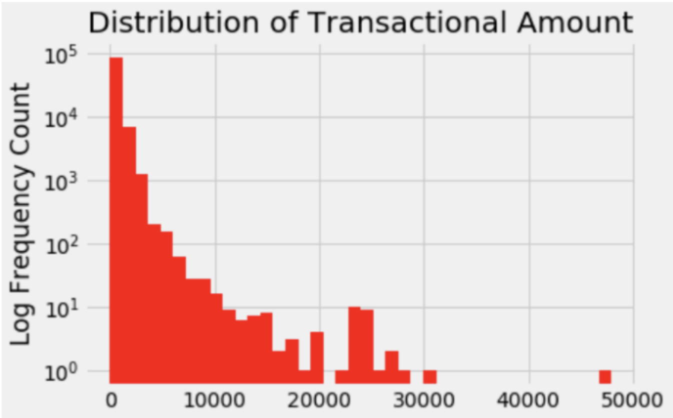


Figure 2.2 Distribution of Amount with the Largest Outlier Removed

Figure 2.2 is the distribution of the field "Amount" with the largest outlier (Record number 52715) removed. The variable amount is positively distributed, with 99% of records falling in the range of 0 to 2500.

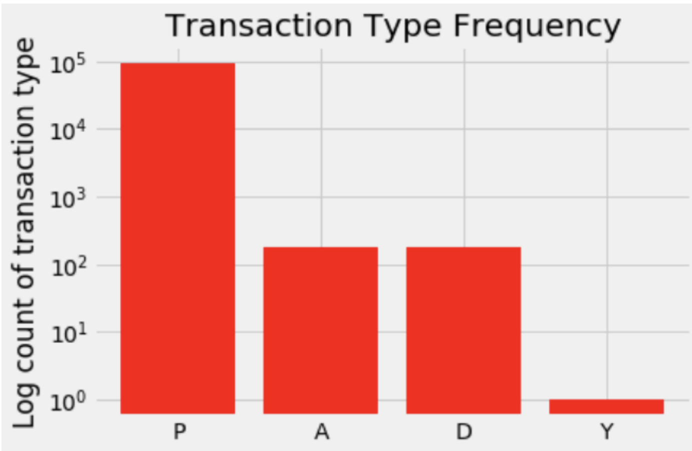


Figure 2.3 Number of Transactions by Type

For "Transtype" field, the bar chart on the left suggests over 99% of records are purchase transactions.

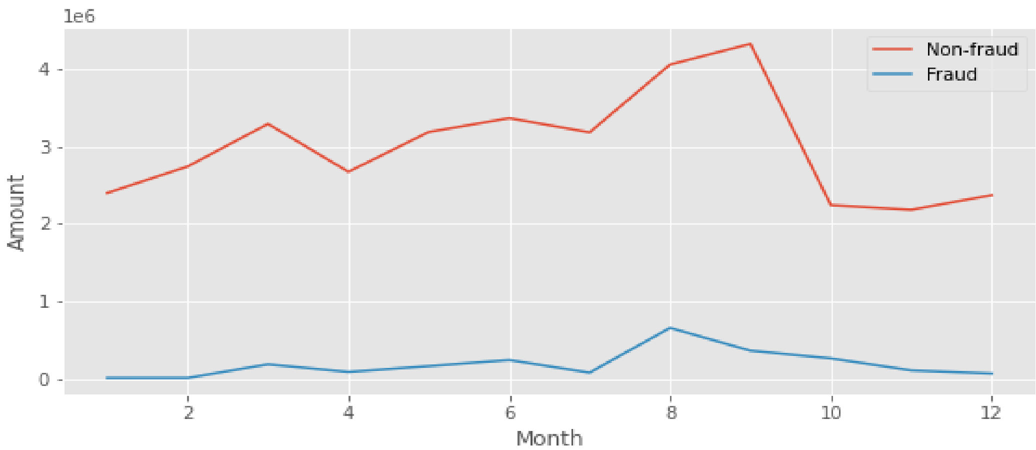


Figure 2.4 Transaction Amount by Month

In terms of date, number of transactions reaches its peak in August then drops quickly after September. Overall, it ranges from 5109 to 11050.

Fraud transaction amount varies consistently with non-fraud transaction amount over time. They both spike in August and drop after September.

More detailed data description can be found in Appendix 9.1.

3. Data Cleaning

We have determined in the data quality checking stage that there are missing values and outlier in the dataset. Besides, there exists format inconsistency in the merchant zip field. Hence, we fixed them before doing further analysis.

Inconsistent format: Normally, zip codes are 5-digit numbers. However, in our dataset, there are a few merchant zip codes that have less than 5 digits. To fix this format inconsistency, we added zeros to the beginning of zip codes that have 4 digits.

Removing outlier: According to our exploratory data analysis, there is only one record with unexpectedly high amount value. Considering removing it does not hurt the continuity of this dataset, we directly dropped that record.

Filling in missing values: There are 3198 missing values in "Merchnum" field, 1020 missing values in "Merch state" field, and 4300 missing values in "Merch zip" field. Our logic of filling in missing values is that for each field that contains missing value(s), we used the most common value of that field over a relevant subset of records (in this case, we tend to use one to two relevant fields) as a replacement. Once we determined which fields to use, we grouped them into categories and replaced the missing field with the most common value for its appropriate group. If there were still missing values after that, we would replace them with the most common value of that field.

The table below displays the process of filling in missing values.

Table 3.1 Summary Table for Filling in Missing Values

Action	Result
"Merchnum" - 3198 missing values in total	
Use "Merch description" to fill NAs	2038 missing "Merchnum" values left
Use "Cardnum" to fill NAs	57 missing "Merchnum" values left
Use global mode to fill NAs	0 missing "Merch zip" values left
"Merch state" - 1020 missing values in total	
Use "Merch description" to fill NAs	363 missing "Merch state" values left
Use global mode to fill NAs	0 missing "Merch state" values left
"Merch state" - 4300 missing values in total	
Use "Merch description" to fill NAs	2043 missing "Merch zip" values left
Use "Cardnum" to fill NAs	42 missing "Merch zip" values left
Use global mode to fill NAs	0 missing "Merch zip" values left

4. Candidate Variables

4.1 Feature Engineering

We used 5 fields to perform feature engineering, including "Cardnum", "Merchnum", "Merch description", "Merch zip", and "Merch state". We wanted to find the frequency of each record based on these categorical variables or the combination of them. For instance, if a particular credit card has on average 20 transactions over 3 days, but suddenly we see 100 transactions over the past 3 days, then it is reasonable to consider that these credit card transactions are more suspicious. Having that logic in mind, we started our variable creation in this direction.

4.2 Variable Creation Process

The entire process of variable creation can be broken down into two steps.

(1) Two Identifiers

We created 4 new fields by concatenating "Cardnum" with 4 other fields (i.e., "Merchnum", "Merch description", "Merch zip", and "Merch state").

1. Cardnum + merchnum
2. Cardnum + merch_description
3. Cardnum + zip
4. Cardnum + state

As a result, we got 6 entities:

1. Cardnum
2. Merchnum
3. Cardnum_merch_description
4. Cardnum_merchnum
5. Cardnum_zip
6. Cardnum_state

(2) Five Series of Variables

Day-since variables are a series of variables defined by how many days since the last time the entity or combination group was seen. 6 variables were created in total using this formula. The alias for variables in this series have the following structure: [day_since _{entity}].

Frequency variables are a series of variables defined by the number of records with same attributes over the last {0, 1, 3, 7, 14, 30} days. 36 variables were created in total **(**6 attributes x 6 different day periods) using this formula. The alias for variables in this series have the following structure: [count_of_{entity}_over_{day}_days].

Relative velocity variables are a series of variables defined by the number or the amount of transaction with attributes seen in the past {0, 1} day divided by the number of applications with the same entity seen in past {7, 14 ,30} days. 72 variables were created in total (6 entities x 2 days x 3 days x 2 calculation methods (number of transactions or amount of transactions)) with this formula. The alias for variables in this series have the following structure: [velocity_{days}_ over days (7, 14 30) on_{entity}].

Amount variables are a series of variables defined by the statistical calculations (e.g., average, median, total, minimum, maximum, maximum – minimum, (maximum – minimum)/ minimum, actual amount/average, actual amount/median, actual amount/maximum, actual amount/total amount) on the transaction amount by a particular entity over the past {0, 1, 3, 7 14 and 30} days. 396 variables were created in total (11 statistical calculation on transaction amount x 6 entities x 6 days) using this formula. The alias for variables in this series have the following structure [{calculations}_over_{days}_{entity}].

Uniqueness variables are a series of variables defined by the unique number of "Merchnum" over the past {0, 1, 3, 7, 14, 30} days. 6 variables were created in total (6 days) using this formula. The alias for variables in this series have following structure: [unique_count_merchnum_each_cardnum_over_{day}].

4.3 Summary of Candidate Variables

We created 517 variables in this stage, which can be divided in 5 categories: 6 day-since variables, 36 frequency variables, 72 relative velocity variables, 6 uniqueness merchnum variables, 396 amount variables, and 1 random variable that is designed to examine the effectiveness of all variables created. With all these carefully created variables, we were ready to go to the next step — feature selection.

For more details and a full list of 517 variables, please refer to Appendix 9.2.

5. Feature Selection

5.1 Methodology for Feature Selection

Feature selection is the process of selecting a subset of candidate variables that contribute the most to model performance. Curse of dimensionality is detrimental to model generalization, so we wanted to exclude any redundant or irrelevant features. In this section, we used two feature selection methods, filter and wrapper, to select 30 top performance variables out of 517 features.

For feature selection, we only used modeling data (transactions happened before 11/01/2010) since OOT data is assumed as future data that is not available for building the model. We also excluded transaction data happened in first two weeks (i.e. before 01/15/2010) for two reasons. First, their features are not well built and thus not reliable. When creating features, we built on past information such as the count of "Cardnum" over past 7 days. In these cases, candidate variables of first-two-week records are not mature enough to be used. Second, fraud labels of first-two-week records are not reliable as they usually come a few months after transactions happened.

5.2 Filter Method

The filter method evaluates the predictive power of each feature using various kinds of univariate measures (e.g. Pearson Correlation, Fisher Score, Kolmogorov-Smirnov, and mutual information). In this project, we chose two measures: Kolmogorov–Smirnov test (KS) and Fraud Detection Rate (FDR). KS is a statistical measure of how well two distributions are separated. In our case, it calculated the maximum difference of cumulative distributions of normal transactions and fraudulent transactions. FDR is a specially designed measure for fraud models that is commonly used and robust. We calculated the percentage of all the fraud caught at a score cutoff as our FDR. For instance, 50% FDR at 3% cutoff means that we caught 50% of all the frauds in the top 3% applications ranked by predicted fraud score (% Frauds caught = # frauds caught / total # frauds).

We first used the filter method as it is useful and easy to implement. Searching among 517 candidate variables is a high dimensionality problem and using univariate measures could bypass this problem tactfully.

Specifically, we first split the data into "Goods" and "Bads" according to the fraud label and then used the "scipy.stats.ks_2samp" function to calculate the KS score for each candidate variable. We also calculated the FDR at 3% for all our candidate variables as a complementary measure. For implementation, we sorted the data according to each variable and used 3% as the cutoff to count the number of "Bads" in the top 3% records. The final FDR is the ratio of "Bads" caught in each top 3% records to the total number of "Bads".

After calculating KS and FDR for each candidate variable, we ranked them accordingly and used the averaged rank as our final score to sort all candidate variables. As a higher rank indicates better prediction power for fraud, we eventually selected 80 variables out of 517 candidates and used wrapper to carry on feature selection.

5.3 Wrapper Method

Different from the filter method that focuses on the performance of each variable, the wrapper method provides information on the joint performance of variables. As variables do interact with each other in the model, the wrapper method would be more helpful in selecting our final variables.

Forward selection and backward selection are two common wrappers that are efficient and effective. Forward selection starts from one variable and adds one at a time while backward selection starts from all variables and removes one at a time. The criterion of adding or removing a variable is its marginal contribution to model performance. In this project, we used backward selection as it outperforms forward selection in empirical experience. To further targeting at fraud detection, we used FDR at 3% as our criterion of model performance.

For implementation, we used the "SequentialFeatureSelector"(SFS) function from "sklearn.feature_selection" package to conduct backward selection. Though backward selection is much more efficient than best subset selection, it still involves $O(80^2)$ models and can be extremely time-consuming. Thus, we used logistic regression as our base model for better efficiency. Besides, we also rescaled the data to accelerate model convergence. With processed data and 80 selected variables by the filter, the backward wrapper provided 30 further selected variables as our final variables for modeling.

5.4 Selected Variables

Table 5.1 Top 30 Variables with Highest Predictive Power

No.	Variable Name	Description
1	total_amount_over_3_for_cardnum_zip	the number of same combinations of cardnum and zip appeared over last 3 days
2	total_amount_over_14_for_cardnum_merchnum	the number of same combinations of cardnum and merchnum appeared over last 14 days
3	total_amount_over_7_for_cardnum_merchdescription	the number of same combinations of cardnum and merchdescription appeared over last 7 days
4	total_amount_over_3_for_cardnum_merchdescription	the number of same combinations of cardnum and merchdescription appeared over last 3 days
5	total_amount_over_3_for_cardnum_merchnum	the number of same combinations of cardnum and merchnum appeared over last 3 days
6	total_amount_over_1_for_cardnum_state	the number of same combinations of cardnum and state appeared over the last day
7	max_amount_over_14_for_cardnum_merchdescription	the max number of appearances of same combination of cardnum and merchdescription per day over last 14 days
8	total_amount_over_0_for_cardnum_zip	the number of same combinations of cardnum and zip appeared in the same day
9	total_amount_over_0_for_cardnum_merchdescription	the number of same combinations of cardnum and merchdescription appeared in the same day
10	max_amount_over_30_for_cardnum_zip	the max number of appearances of same combination of cardnum and zip per day over last 30 days
11	max_amount_over_3_for_cardnum_ merchdescription	the max number of appearances of same combination of cardnum and merchdescription per day over last 3 days
12	max_amount_over_14_for_cardnum_state	the max number of appearances of same combination of cardnum and state per day over last 14 days
13	max_amount_over_3_for_cardnum_merchnum	the max number of appearances of same combination of cardnum and merchnum per day over last 3 days
14	total_amount_over_1_for_Merchnum	the number of same merchnum appeared over the last day

15	max_amount_over_30_for_cardnum_state	the max number of appearances of same combination of cardnum and state per day over last 30 days
16	total_amount_over_0_for_Merchnum	the number of same merchnum appeared in the same day
17	total_amount_over_1_for_Cardnum	the number of same cardnum appeared over the last day
18	total_amount_over_3_for_Merchnum	the number of same merchnum appeared over the last 3 days
19	total_amount_over_0_for_Cardnum	the number of same cardnum appeared in the same day
20	max_amount_over_0_for_cardnum_merchdescription	the max number of appearances of same combination of cardnum and merchdescription per day in the same day
21	max_amount_over_0_for_cardnum_zip	the max number of appearances of same combination of cardnum and zip per day in the same day
22	max_amount_over_1_for_Merchnum	the max number of appearances of same merchnum per day over the last day
23	max_amount_over_0_for_Cardnum	the max number of appearances of same cardnum per day in the same day
24	max_amount_over_7_for_Cardnum	the max number of appearances of same cardnum per day over the last 7 days
25	total_amount_over_7_for_Merchnum	the number of same merchnum appeared over the last 7 days
26	total_amount_over_14_for_Cardnum	the number of same cardnum appeared over the last 14 days
27	range_amount_over_14_for_Cardnum	the range of number of same cardnum appeared over the last 14 days
28	range_amount_over_7_for_Cardnum	the range of number of same cardnum appeared over the last 7 days
29	mean_amount_over_0_for_Merchnum	the mean of number of same merchnum appeared in the same day
30	mean_amount_over_0_for_Cardnum	the mean of number of same cardnum appeared in the same day

6. Model Algorithms

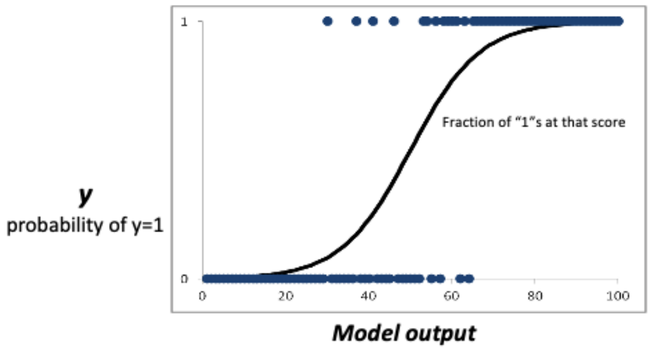
6.1 Algorithm Description

After feature engineering and feature selection, we got down to 30 high performance features, and we were ready to build a binary classification model using a wide range of machine learning algorithms. In this section, we will include a brief description for each algorithm used, treatment for imbalanced dataset, a high-level summary of performance for each model (FDR for training, testing, and OOT), and our final model choice.

We built the following machine learning algorithms.

(1) Logistic Regression

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable (i.e., fraud label). Logistic regression is estimating the parameters of logistic model given all the X parameters for binary classification problem with logit function.



$$y = \frac{1}{1 + e^{-\sum b_i x_i}}$$

Figure 6.1 Logistic Regression

We used logistic regression as our base model, which meant all other models should perform better than logistic regression. We created two versions of logistic regression. For the first version, we used Ridge Regularization to minimize the mean square error, and for the second version, we did not apply any regularization. For other model parameters, we kept the same for both model with default values.

(2) Gradient Boosted Trees

Gradient Boosting is a machine learning technique for regression and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. Decision tree divides the independent variable (x's) space into boxes and places a step above each box at the height of the average of the dependent variable y in that box. A two-dimensional decision tree is shown below:

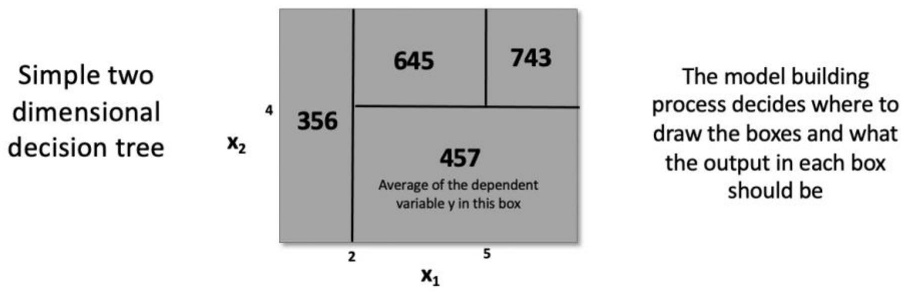


Figure 6.2 Decision Tree

When a decision tree is the weak learner, the resulting algorithm is called the gradient boosted tree. A boosted tree starts with a very simply model, then keep adding in weak models to get closer to the right answer. Each weak model is trained to predict the residual error of the current sum, and each added model makes the overall model slightly better.

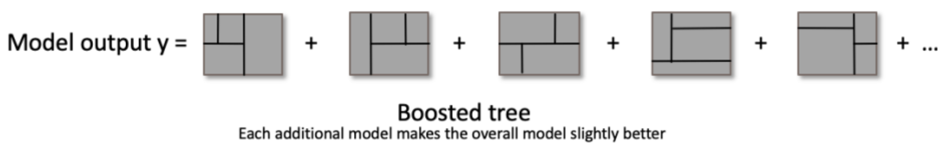


Figure 6.3 Boosted Tree

We tried three versions of boosted trees.

Gradient Boosting Classifier **(GB)** builds an additive model in a forward stage-wise fashion; it allows for the optimization of arbitrary differentiable loss functions, and it is capable of multiclass classification. For each class, a tree is fit on the negative gradient of the binomial or multinomial deviance loss function. Binary classification is a special case where only a single regression tree is induced.

Extreme gradient boosting (**XGBoost)** is another ensemble model, which is constructed from decision tree. We wanted to use this method to compare with random forest because XGboost uses the method called gradient boosting as the loss gradient. Specifically, models are fit using any arbitrary differentiable loss function and gradient descent optimization algorithm. We focused on couple parameters, such as learning rate, number of estimators to see the model performance by adjusting different parameters.

Light Gradient Boosting Machine **(LightGBM)** is an open source distributed gradient boosting framework for machine learning originally developed by Microsoft. It is based on decision tree algorithms and used for ranking, classification, and other machine learning tasks. Instead of growing a tree level-wise, LightGBM grows a tree leaf-wise. It chooses the leaf it believes will yield the largest decrease in loss. Besides, LightGBM does not use the widely used sorted-based decision tree learning algorithm, which searches the best split point on sorted features values. Instead, LightGBM implements a highly optimized histogram-based decision tree learning algorithm, which yields great advantages on both efficiency and memory consumption.

(3) Random Forest

Random forest is an ensemble learning method for classification that operate by constructing multitude of decision trees. We used random forest to detect possible non-linear boundary in the feature space. One advantage of random forest is the ensembles of the model. It created many trees and then output the mode of classes for different tree algorithms.

We created 7 random forest models with different parameters. We specifically selected and modified number of estimators, the depth of each estimator and the smallest sample needed to split a node. Generally, we wanted a deep tree for the random forest because we can avoid overfitting by creating N different trees.

(4) Neural Network

A neural network is a machine learning algorithm construct that maps an input vector to an output scalar, or typically a vector of axes into a single dependent variable. It is inspired by the biological neural networks that constitute human brains. A typical neural network consists of an input layer, some hidden layers and an output layer.

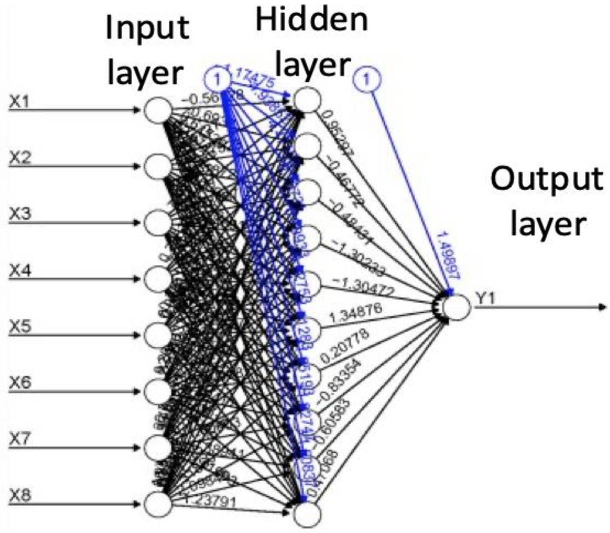


Figure 6.4 Neural Network

The input layer is formed by all independent variables(x's), so it has as many nodes as the input vector dimension. The output layer, which is the last layer, has as many nodes as the output dependent variable dimension, which is typically one. We also have many hidden layers in between the input layer and the output layer. Each node in the hidden layer receives weighted signals from all nodes in the previous layer and then go through some kinds of activation/transformation function, such as logistic function and sigmoid function, to output a single dependent variable (y).

The node weight is trained by backpropagating the error. For each training record shown to the neural net, an error is calculated and propagates backward to each node. Then, neural net would calculate the gradient of that error, with respect to the weight in that node. The gradient, associated with a learning rate, is used to slightly adjust each node weight. Each record is passed through many times as the weights settle into a local optimum.

6.2 Treatment for Imbalanced Dataset

For training datasets, we implemented Synthetic Minority Oversampling Technique (SMOTE) as an oversampling method to create artificial "bad" record.

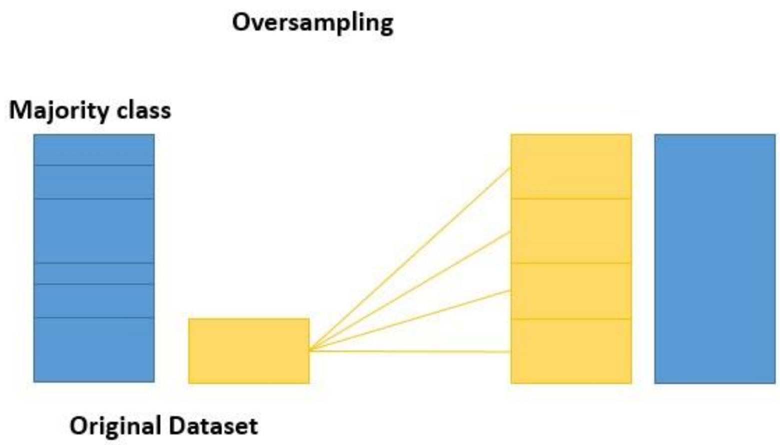


Figure 6.5 Oversampling

Specifically, SMOTE takes a "bad" record from the dataset and consider its k nearest neighbors (in the feature space). Then, it invents new "bad" records that lie somewhere between neighboring know "bad" records. Thus, for each training dataset, we were able to get equal numbers of fraud and non-fraud records.

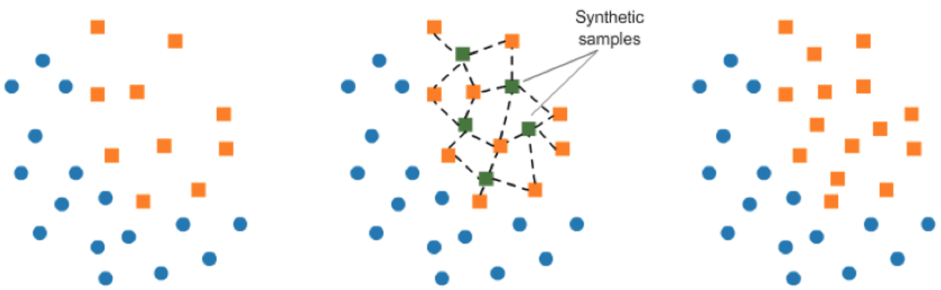


Figure 6.6 Synthetic Minority Oversampling Technique

6.3 Summary of Model Exploration

For each set of hyperparameters of each model, we conducted 15 times cross-validation and calculated the average FDR at 3% of population for training, testing, and OOT dataset. For each validation, we randomly divided our dataset (excluded OOT and first two-weeks of data) into training and testing. We implemented SMOTE on training dataset and then fit our model. Then, we predicted on training, testing, and OOT datasets and calculated FDR for all of them. After 15 times cross-validation, we calculated our average FDR by taking the mean of 15 FDRs.

Table 6.1 Model Summary

Model	Hyperparameters					Average FDR@3%		
Logistic Regression	penalty					Train	Test	OOT
1	l2					64.0%	64.0%	36.0%
2	None					64.0%	64.0%	36.0%
Neural Network	hidden_layer_sizes	activation	learning_rate	learning_rate_init	alpha	Train	Test	OOT
1	(20,)	relu	constant	0.0001	1e-5	69.8%	69.8%	50.3%
2	(5,10)	tanh	invscaling	0.001	8e-4	74.7%	74.4%	54.7%
3	(10,5)	tanh	invscaling	0.001	8e-4	76.5%	74.9%	55.1%
4	(10,10)	tanh	invscaling	0.001	8e-4	76.4%	75.2%	56.7%
5	(10, 10)	relu	constant	0.001	1e-4	78.4%	77.3%	57.9%
6	(10,20)	relu	constant	0.001	1e-4	78.4%	76.3%	57.5%
Random Forest	n_estimators	max_depth	min_samples_split			Train	Test	OOT
1	100	3	Default(1)			68.7%	67.3%	44.5%
2	150	3	Default(1)			69.4%	68.3%	45.1%
3	200	4	5			67.8%	67.2%	44.2%
4	250	4	30			70.2%	68.8%	47.6%
5	300	5	200			67.8%	67.3%	44.4%
6	350	5	20			70.2%	68.8%	48.8%
7	400	6	50			67.9%	67.3%	44.2%
Gradient Boosting Classifier	n_estimators	learning_rate	max_depth	min_samples_split		Train	Test	OOT
1	150	0.01	3	100		71.2%	63.3%	30.5%
2	150	0.025	4	80		83.5%	65.7%	33.4%
3	250	0.025	4	70		85.3%	66.3%	37.6%
4	250	0.025	5	50		88.4%	65.4%	33.5%
5	300	0.025	6	40		90.3%	67.2%	32.8%
XGBoost	n_estimators	learning_rate	min_child_weight			Train	Test	OOT
1	100	0.1	20			85.6%	76.6%	44.7%
2	150	0.3	50			86.8%	76.8%	46.4%
3	200	0.4	40			89.6%	77.6%	45.8%
4	300	0.4	60			87.9%	75.5%	49.7%
5	300	0.3	80			82.8%	74.3%	49.2%
6	400	0.5	50			92.1%	75.8%	48.6%
LightGBM	n_estimators	learning_rate	max_depth	num_leaves		Train	Test	OOT
1	200	0.05	3	6		75.4%	50.2%	32.5%
2	300	0.01	4	12		76.5%	65.2%	39.5%
3	400	0.01	4	12		80.7%	73.5%	39.8%
4	500	0.01	5	8		87.0%	82.3%	42.6%
5	600	0.01	5	30		93.5%	89.3%	39.3%
6	700	0.01	6	30		95.8%	87.7%	34.2%
7	800	0.05	4	18		96.5%	88.5%	37.1%

The table above is our high-level summary of hyperparameter tuning and model performance. We decided to use **Neural Network,** which has the highest average OOT FDR, as our final model.

Table 6.2 Hyperparameters of Final Model

Final Model	hidden_layer	activation	learning_rate	learning_rate_init	alpha
Neural Network	(10,10)	relu	constant	0.001	0.0001

7. Final Model Results

7.1 Result Tables

Our model results are displayed in the three tables below. The first two tables are the training and testing results, which only covers the period from 01/15/2010 to 11/01/2010. The first two weeks' records are excluded because they are not reliable. The third table is OOT population result for the validation purpose.

For each population table, we assigned a fraud score to each record predicted using our best model. Then we sorted the table based on this fraud score in descending order. After sorting the table, we separated the population equally into 100 bins, each covering 1% of the table population. For each bin, bins statistics and cumulated statistics are calculated. Lastly, for details of the statistics calculated, please refer to the tables below, which show the top 20 bins' results. For the full result tables, kindly refer to Appendix 9.3.

Table 7.1 Training Population Results (top 20%)

Training	# Records		# Goods		# Bads		Fraud Rate					
	64658		63964		694		0.010733397					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	647	222	425	34.31%	65.69%	647	222	425	0.35%	61.24%	60.89	0.52
2	647	558	89	86.24%	13.76%	1294	780	514	1.22%	74.06%	72.84	1.52
3	647	607	40	93.82%	6.18%	1941	1387	554	2.17%	79.83%	77.66	2.50
4	647	631	16	97.53%	2.47%	2588	2018	570	3.15%	82.13%	78.98	3.54
5	647	632	15	97.68%	2.32%	3235	2650	585	4.14%	84.29%	80.15	4.53
6	647	632	15	97.68%	2.32%	3882	3282	600	5.13%	86.46%	81.32	5.47
7	647	639	8	98.76%	1.24%	4529	3921	608	6.13%	87.61%	81.48	6.45
8	647	633	14	97.84%	2.16%	5176	4554	622	7.12%	89.63%	82.51	7.32
9	647	641	6	99.07%	0.93%	5823	5195	628	8.12%	90.49%	82.37	8.27
10	647	642	5	99.23%	0.77%	6470	5837	633	9.13%	91.21%	82.08	9.22
11	647	645	2	99.69%	0.31%	7117	6482	635	10.13%	91.50%	81.36	10.21
12	647	642	5	99.23%	0.77%	7764	7124	640	11.14%	92.22%	81.08	11.13
13	647	645	2	99.69%	0.31%	8411	7769	642	12.15%	92.51%	80.36	12.10
14	647	642	5	99.23%	0.77%	9058	8411	647	13.15%	93.23%	80.08	13.00
15	647	645	2	99.69%	0.31%	9705	9056	649	14.16%	93.52%	79.36	13.95
16	647	646	1	99.85%	0.15%	10352	9702	650	15.17%	93.66%	78.49	14.93
17	647	645	2	99.69%	0.31%	10999	10347	652	16.18%	93.95%	77.77	15.87
18	647	645	2	99.69%	0.31%	11646	10992	654	17.18%	94.24%	77.05	16.81
19	647	646	1	99.85%	0.15%	12293	11638	655	18.19%	94.38%	76.19	17.77
20	647	644	3	99.54%	0.46%	12940	12282	658	19.20%	94.81%	75.61	18.67

Table 7.2 Testing Population Results (top 20%)

Testing	# Records		# Goods		# Bads		Fraud Rate					
	16165		15991		174		0.010763996					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	162	58	104	35.80%	64.20%	162	58	104	0.36%	59.77%	59.41	0.56
2	162	140	22	86.42%	13.58%	324	198	126	1.24%	72.41%	71.18	1.57
3	162	152	10	93.83%	6.17%	486	350	136	2.19%	78.16%	75.97	2.57
4	162	156	6	96.30%	3.70%	648	506	142	3.16%	81.61%	78.44	3.56
5	162	159	3	98.15%	1.85%	810	665	145	4.16%	83.33%	79.17	4.59
6	162	156	6	96.30%	3.70%	972	821	151	5.13%	86.78%	81.65	5.44
7	162	157	5	96.91%	3.09%	1134	978	156	6.12%	89.66%	83.54	6.27
8	162	159	3	98.15%	1.85%	1296	1137	159	7.11%	91.38%	84.27	7.15
9	162	159	3	98.15%	1.85%	1458	1296	162	8.10%	93.10%	85.00	8.00
10	162	161	1	99.38%	0.62%	1620	1457	163	9.11%	93.68%	84.57	8.94
11	162	161	1	99.38%	0.62%	1782	1618	164	10.12%	94.25%	84.13	9.87
12	162	162	0	100.00%	0.00%	1944	1780	164	11.13%	94.25%	83.12	10.85
13	162	162	0	100.00%	0.00%	2106	1942	164	12.14%	94.25%	82.11	11.84
14	162	162	0	100.00%	0.00%	2268	2104	164	13.16%	94.25%	81.10	12.83
15	162	161	1	99.38%	0.62%	2430	2265	165	14.16%	94.83%	80.66	13.73
16	162	161	1	99.38%	0.62%	2592	2426	166	15.17%	95.40%	80.23	14.61
17	162	161	1	99.38%	0.62%	2754	2587	167	16.18%	95.98%	79.80	15.49
18	162	162	0	100.00%	0.00%	2916	2749	167	17.19%	95.98%	78.79	16.46
19	162	159	3	98.15%	1.85%	3078	2908	170	18.19%	97.70%	79.52	17.11
20	162	162	0	100.00%	0.00%	3240	3070	170	19.20%	97.70%	78.50	18.06

Table 7.3 Out Of Time Population Results (top 20%)

OOT	# Records		# Goods		# Bads		Fraud Rate					
	12236		12057		179		0.014628964					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	123	65	58	52.85%	47.15%	123	65	58	0.54%	32.40%	31.86	1.12
2	123	96	27	78.05%	21.95%	246	161	85	1.34%	47.49%	46.15	1.89
3	123	110	13	89.43%	10.57%	369	271	98	2.25%	54.75%	52.50	2.77
4	123	118	5	95.93%	4.07%	492	389	103	3.23%	57.54%	54.32	3.78
5	123	119	4	96.75%	3.25%	615	508	107	4.21%	59.78%	55.56	4.75
6	123	118	5	95.93%	4.07%	738	626	112	5.19%	62.57%	57.38	5.59
7	123	122	1	99.19%	0.81%	861	748	113	6.20%	63.13%	56.92	6.62
8	123	122	1	99.19%	0.81%	984	870	114	7.22%	63.69%	56.47	7.63
9	123	120	3	97.56%	2.44%	1107	990	117	8.21%	65.36%	57.15	8.46
10	123	122	1	99.19%	0.81%	1230	1112	118	9.22%	65.92%	56.70	9.42
11	123	118	5	95.93%	4.07%	1353	1230	123	10.20%	68.72%	58.51	10.00
12	123	122	1	99.19%	0.81%	1476	1352	124	11.21%	69.27%	58.06	10.90
13	123	122	1	99.19%	0.81%	1599	1474	125	12.23%	69.83%	57.61	11.79
14	123	122	1	99.19%	0.81%	1722	1596	126	13.24%	70.39%	57.15	12.67
15	123	119	4	96.75%	3.25%	1845	1715	130	14.22%	72.63%	58.40	13.19
16	123	121	2	98.37%	1.63%	1968	1836	132	15.23%	73.74%	58.52	13.91
17	123	122	1	99.19%	0.81%	2091	1958	133	16.24%	74.30%	58.06	14.72
18	123	123	0	100.00%	0.00%	2214	2081	133	17.26%	74.30%	57.04	15.65
19	123	122	1	99.19%	0.81%	2337	2203	134	18.27%	74.86%	56.59	16.44
20	123	121	2	98.37%	1.63%	2460	2324	136	19.28%	75.98%	56.70	17.09

7.2 Fraud Scores Over Time

We can validate our predicted fraud scores by checking how they change over time and with increasing number of transactions. Normally, when the number of activities associated with an entity increases, the fraud score would increase if abnormal or suspicious events happen. In this section, we examined one credit card and one merchant as examples.

First, we found an abnormal transaction pattern associated with card ***5142235211***. There were 34 transactions made by this card in November and December 2010. Among these 34 transactions, 17 happened on November 25^{th} and 15 happened on November 26^{th}. The graphs below show how our predicted fraud scores change over time and over transaction count respectively. They show that our predicted fraud scores increased dramatically within the aforementioned two days when the number of transactions were extremely high compared to other days.

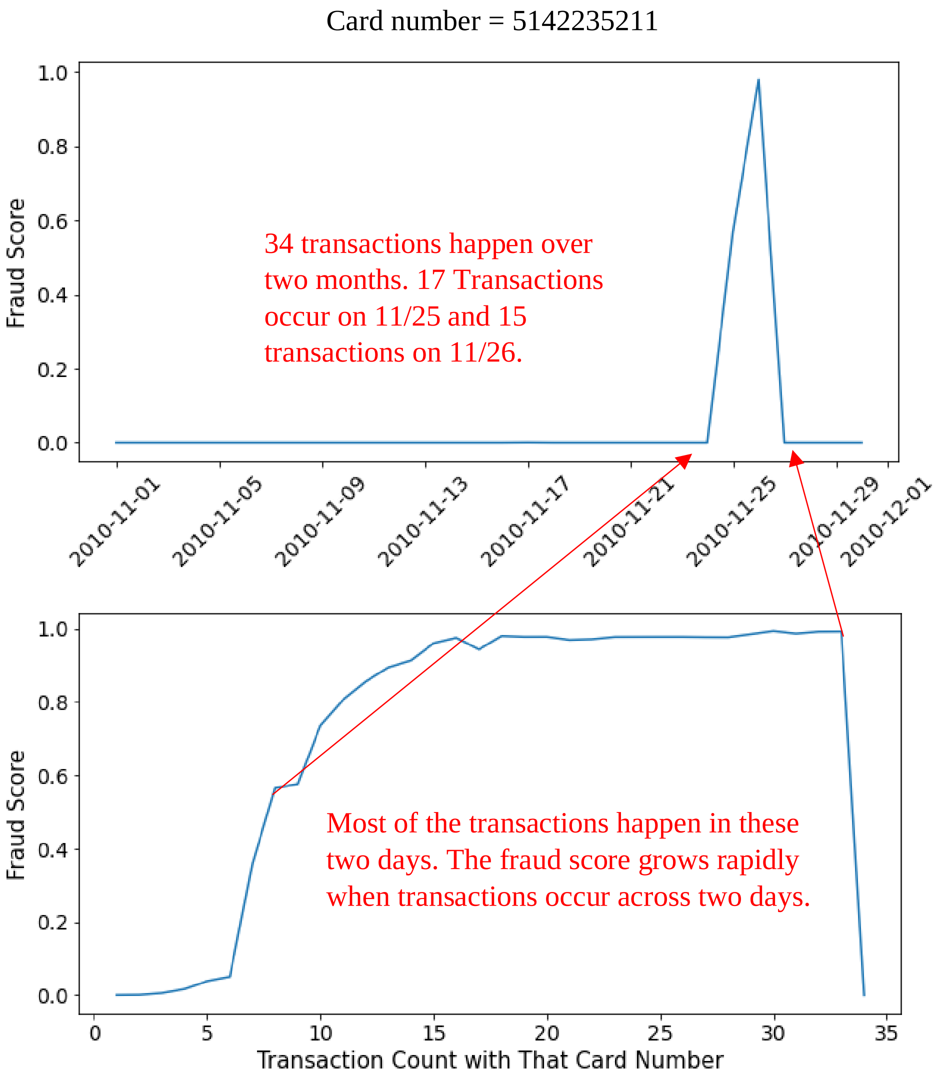


Figure 7.1 Fraud Score and Card Activities

Second, we found similarly abnormal transaction patterns associated with merchant ***4353000719908***. There were 90 purchases made from this merchant in November 2010. However, 32 transactions happened in two days – 17 on November 25th and 15 on November 26th. The graphs below show that large bunching of transactions led to very high predicted fraud scores within these two days. Besides, there was a small rise of fraud score at the beginning of the month when there was a small bunching of transactions.

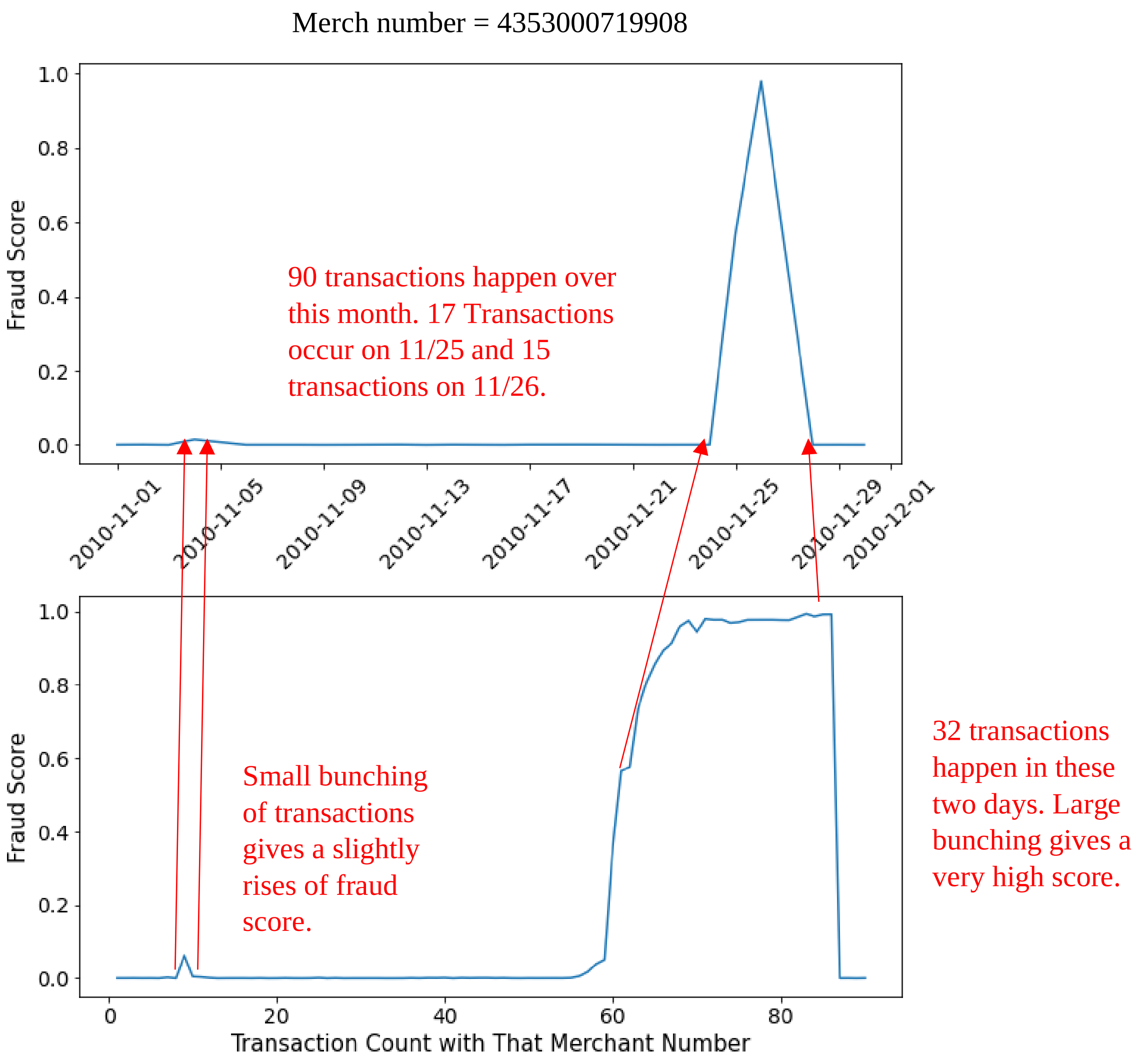


Figure 7.2 Fraud Score and Merchant Activities

7.3 Score Cut-off

After sorting the population according to predicted fraud scores in descending order, we needed to decide the optimal cut-off score. All transactions whose fraud scores are higher than this score shall be declined. Companies would save money from fraudulent transactions caught and lose money on genuine transactions declined. The optimal cut-off score would bring us the highest overall savings while declining the lowest percentage of transactions.

To find out the cut-off score, we calculated the overall savings at each population percentile. Our assumptions were as follows: First, there would be a $2000 saving for each actual fraud caught; Second, a $50 loss would be incurred for each false positive transaction that is declined; Third, overall savings equal to fraud savings minus lost sales incurred. With these assumptions, we calculated the total fraud savings, lost sales, and overall savings at each percentile.

As shown by the graph below, when we declined more percentage of transactions, the total fraud savings (represented by the blue line) kept increasing while more frauds were caught, but at a decreasing rate. Total lost sales (represented by the orange line) increased at a relatively stable rate. As a result, the overall savings (represented by the green line) would first increase and then decrease, with a peak at around 5-7%, where the overall savings during the last 2 months were over $188,000. While maintaining high overall savings, we also wanted to minimize the number of false positive for customer retention purpose. Hence, our team recommended a score cut-off at 5%, which would bring an annual saving of $1,131,600.

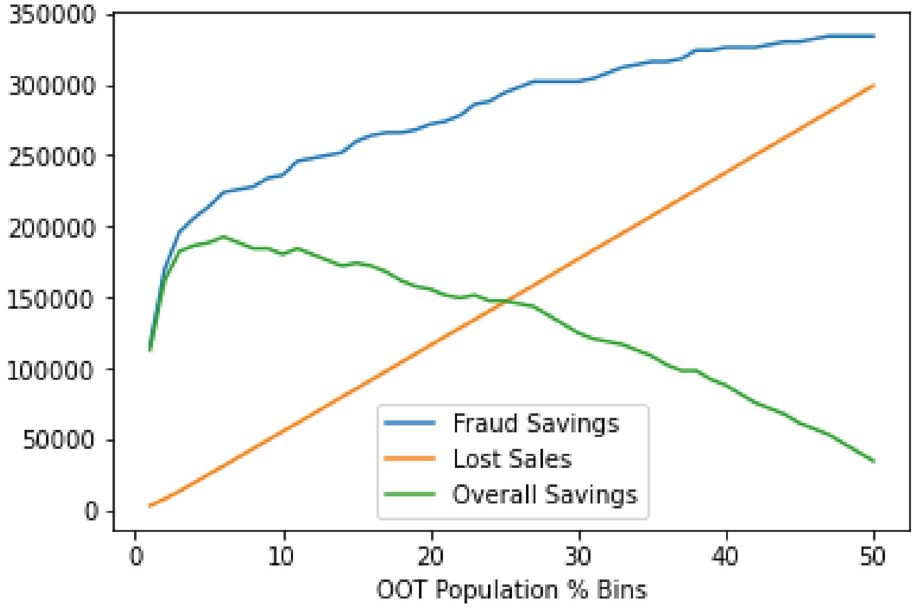


Figure 7.3 Score Cut-off

8. Conclusions

Throughout the project, we have built a real time fraud algorithm to identify card transaction fraud. Given the 96,753 records of card transaction data, we delicately checked data quality. We then carefully dealt with the outlier and imputed missing values. With cleaned data, we created 517 candidate variables, which took time flow characteristics into consideration. Then, we selected the top 30 variables in terms of predictive power for model building. During the model building phase, we began with building baseline logistic regression models with different combination of parameters. Overall, these models performed well on modelling data but not on OOT data, which suggests an issue of overfitting. We also tried a wide range of non-linear models (Gradient Boosted Trees, Random Forest, and Neural Network) and tuned the hyperparameters to achieve the best model performance. With the performance results for all these models, we finally selected a neural network model with all parameters set out in section 6 and 7. Eventually, we were able to achieve a fraud detection rate of 59.78% at 5% population on OOT data.

We could make further improvements to our model in several aspects if given more time. First, we can create more candidate variables beyond our 517 existing ones. If our computer capacity and time allow, we can build as many as thousands of variables. Also, if time permits, we can consult with several domain experts for advice in building expert variables. These might change our best 30 variables used for model building and consequently the results. Moreover, if we were able to collect more data records, we can implement subsampling on the training data to have a higher bad/good ratio. We can run our model for several times, with each time removing the goods with low fraud scores from the training data. In this way, we will have a higher proportion of hard-to-classify records in the training set, which allows the model to focus more on these to improve performance.

9. Appendix

9.1 Data Quality Report for Application Fraud

(1) Description

Dataset Name: Credit Card Transaction Data
Dataset Purpose: Real-time credit card transaction data to identify fraud.
Data Source: The dataset was modified by Dr. Stephen Coggeshall based on his expert knowledge on fraud analytics.
Time Period: 01/01/2010 – 12/31/2010
Number of Fields: 10
Number of Records: 96,753

(2) Summary Tables

Table 9.1 Summary Table for Categorical Fields

Column Name	# of Records	# Missing	% populated	# of Unique Values	Most Common Field Value
Recnum	96,753	0	100	96,753	2047
Cardnum	96,753	0	100	1,645	5142148452
Merchmun	93,378	3,375	96.51	13,091	930090121224
Merch description	96,753	0	100	13,126	GSA-FSS-ADV
Merch state	95,558	1,195	98.76	227	TN
Merch zip	92,097	4,656	95.19	4,567	38118
Transtype	96,753	0	100	4	P
Fraud	96,753	0	100	2	0

Table 9.2 Summary Table for Numerical Field

Column Name	# of Records	% Populated	Mean	Minimum	Maximum	Standard Deviation
Amount	96,753	100	427.89	0.01	3,102,045.53	10,006.14

Table 9.3 Summary Table for Datetime Field

Column Name	# of Records	% Populated	# of Unique Values	Minimum	Maximum	Most Common Field Value
Date	96,753	100	365	2010-01-01	2010-12-31	2010-02-28

(3) Data Field Exploration

Field 1:
Name: Recnum
Description: Unique identifier of each transaction in the data.

Field 2:
Name: Cardnum
Description: Credit card number that associates with the transaction records.

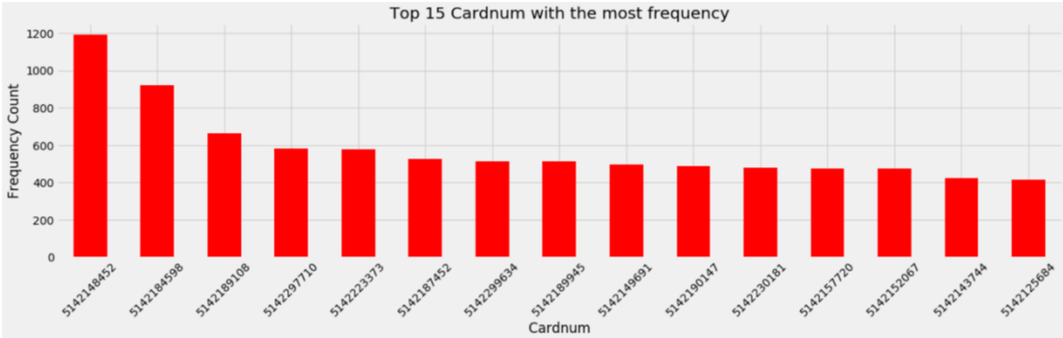


Figure 9.1 Histogram of Top 15 Most Common Cardnum

Field 3:
Name: Merchnum
Description: The number of the merchants that associates with the transaction records.

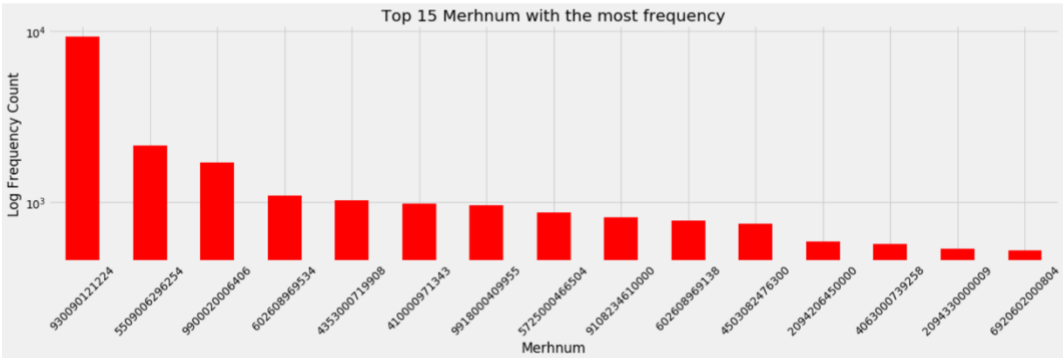


Figure 9.2 Histogram of Top 15 Most Common Merchnum

Field 4:
Name: Merch Description
Description: Brief description of the merchants that associates with the transactions.

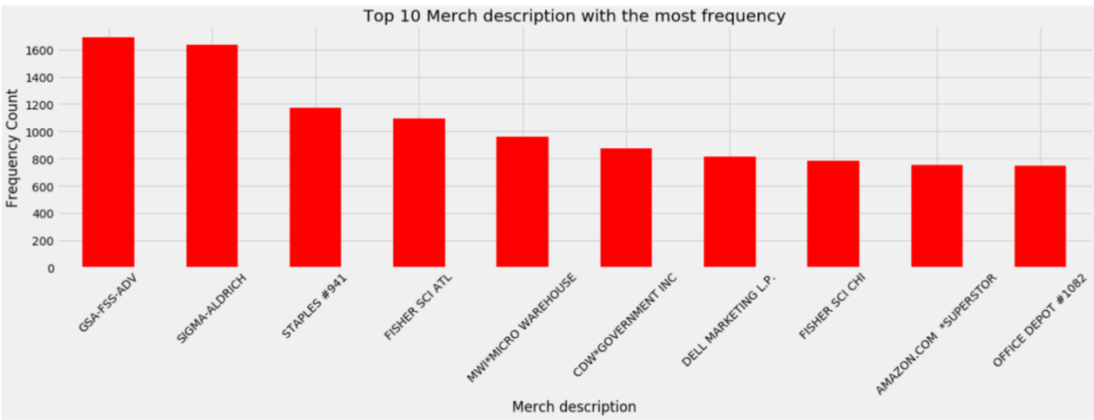


Figure 9.3 Histogram of Top 10 Most Common Merch Description

Field 5:
Name: Merch state
Description: States that associate with the transactions.

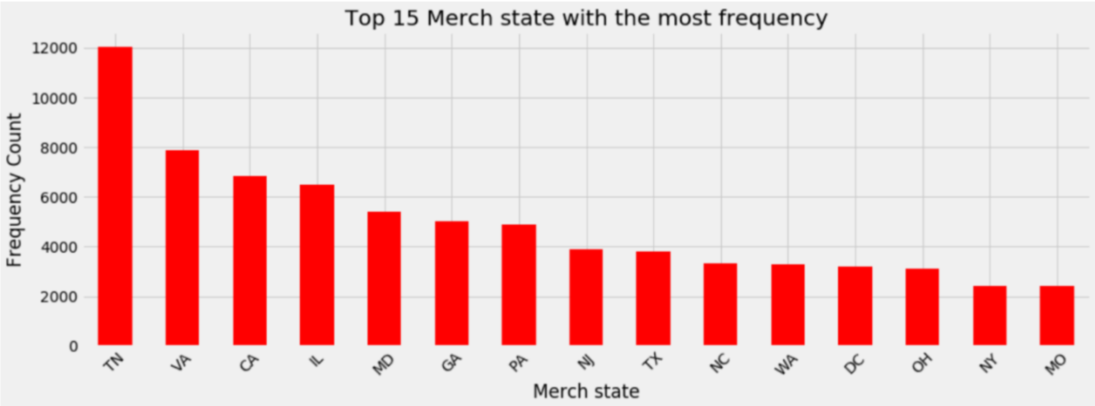


Figure 9.4 Histogram of Top 15 Most Common Merch State

Field 6:
Name: Merch zip
Description: Zip codes that associate with the transactions.

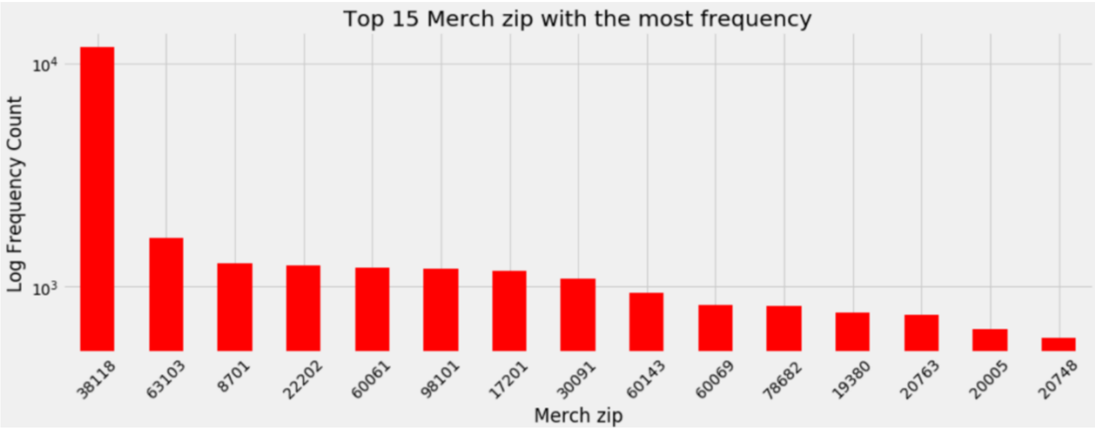


Figure 9.5 Histogram of Top 15 Most Common Merch zip

Field 7:
Name: Transtype
Description: The type of the transactions.

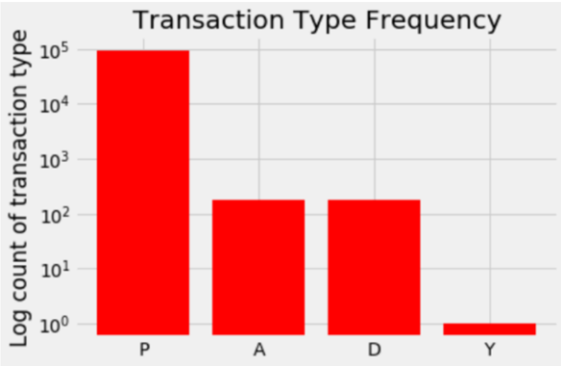


Figure 9.6 Histogram of Transaction Type Count

Field 8:
Name: Date
Description: Date of the transaction.

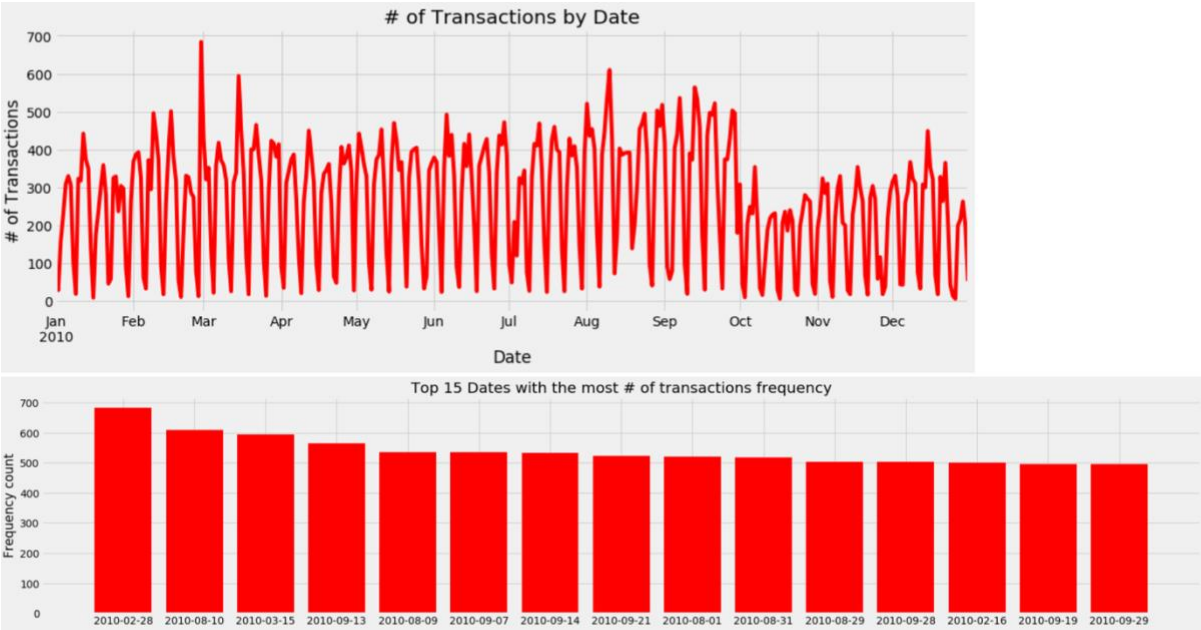


Figure 9.7 Number of Transaction by date and Count of Top 15 Most Common Date

Field 9:
Name: Amount
Description: Transaction Amount. The graph below excludes amount greater than 50,000.

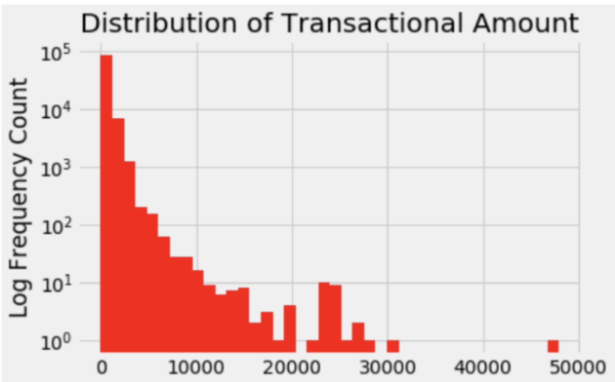


Figure 9.8 Distribution of Amount with the Largest Outlier Removed

Field 10:
Name: Fraud
Description: Fraud label for each transaction.

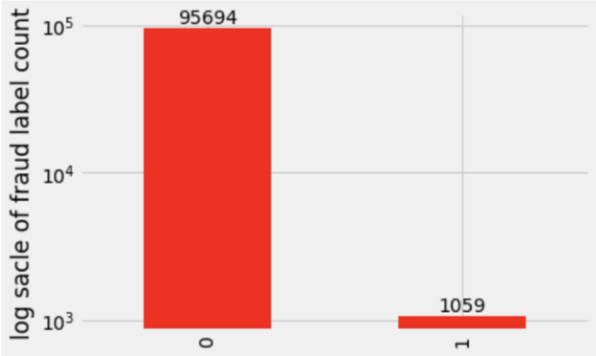


Figure 9.9 Histograms of Fraud Count

9.2 Feature Engineering Result

Table 9.3 Statistics of Candidate Variables

No.	Variable name	Mean	Std.	Min	Median	Max
1	day_since_Cardnum	-234.14	3441.22	-95582	1	350
2	mean_amount_over_0_for_Cardnum	393.56	726.85	0.01	160	28392.84
3	median_amount_over_0_for_Cardnum	381.35	718.73	0.01	149.65	28392.84
4	max_amount_over_0_for_Cardnum	498.21	1030.96	0.01	191.93	47900
5	min_amount_over_0_for_Cardnum	314.06	637.42	0.01	99.99	28392.84
6	range_amount_over_0_for_Cardnum	184.14	807.82	0	0	44800
7	range_percent_amount_over_0_for_Cardnum	9.63	258.61	0	0	70618
8	total_amount_over_0_for_Cardnum	741.65	3431.45	0.01	220	218301.8
9	actual_over_mean_over_0_Cardnum	1	0.44	0	1	23.79
10	actual_over_median_over_0_Cardnum	1.41	9.9	0	1	657.89
11	actual_over_max_over_0_Cardnum	0.88	0.28	0	1	1
12	actual_over_total_over_0_Cardnum	0.77	0.35	0	1	1
13	count_of_Cardnum_over_0_days	2.47	6	1	1	146
14	unique_count_merchnum_each_cardnum_over_0	1.34	0.83	1	1	24
15	mean_amount_over_1_for_Cardnum	395.45	675.83	0.01	188.5	28392.84
16	median_amount_over_1_for_Cardnum	364.22	658.72	0.01	157.7	28392.84
17	max_amount_over_1_for_Cardnum	610.87	1212.86	0.01	264.98	47900
18	min_amount_over_1_for_Cardnum	251.13	524.7	0.01	75	28392.84
19	range_amount_over_1_for_Cardnum	359.75	1105.71	0	3	44800
20	range_percent_amount_over_1_for_Cardnum	21.73	462.57	0	0.1	117049
21	total_amount_over_1_for_Cardnum	1110.05	5669.43	0.01	339.16	307468.1
22	actual_over_mean_over_1_Cardnum	1	0.64	0	1	23.79
23	actual_over_median_over_1_Cardnum	1.83	13.35	0	1	674.76
24	actual_over_max_over_1_Cardnum	0.77	0.36	0	1	1
25	actual_over_total_over_1_Cardnum	0.64	0.39	0	0.83	1
26	count_of_Cardnum_over_1_days	3.37	7.94	1	2	177
27	unique_count_merchnum_each_cardnum_over_1	1.8	1.45	1	1	37
28	mean_amount_over_3_for_Cardnum	395.85	629.37	0.01	213.66	28392.84
29	median_amount_over_3_for_Cardnum	341.36	592.65	0.01	160	28392.84
30	max_amount_over_3_for_Cardnum	739.79	1367.58	0.01	359.06	47900
31	min_amount_over_3_for_Cardnum	198.46	454.48	0.01	54.99	28392.84
32	range_amount_over_3_for_Cardnum	541.33	1309.32	0	113.37	44800
33	range_percent_amount_over_3_for_Cardnum	38.61	473.41	0	2.32	117049
34	total_amount_over_3_for_Cardnum	1512.93	6115.51	0.01	518.61	310843.1
35	actual_over_mean_over_3_Cardnum	1	0.84	0	1	38
36	actual_over_median_over_3_Cardnum	2.34	17.12	0	1	1570.86
37	actual_over_max_over_3_Cardnum	0.67	0.4	0	1	1
38	actual_over_total_over_3_Cardnum	0.51	0.4	0	0.45	1
39	count_of_Cardnum_over_3_days	4.79	11.45	1	2	251
40	unique_count_merchnum_each_cardnum_over_3	2.43	2.08	1	2	39
41	mean_amount_over_7_for_Cardnum	397.19	560.08	0.14	251.83	25500

42	median_amount_over_7_for_Cardnum	307.16	501.73	0.14	161.25	25500
43	max_amount_over_7_for_Cardnum	960.43	1603.13	0.14	553.18	47900
44	min_amount_over_7_for_Cardnum	135.23	351.36	0.01	35	23940
45	range_amount_over_7_for_Cardnum	825.2	1587.09	0	396	44800
46	range_percent_amount_over_7_for_Cardnum	79.24	1076.61	0	10.1	169499
47	total_amount_over_7_for_Cardnum	2384.04	7158.5	0.14	986.14	312616.1
48	actual_over_mean_over_7_Cardnum	0.99	1.08	0	0.95	59.87
49	actual_over_median_over_7_Cardnum	2.99	27.09	0	1	5747.54
50	actual_over_max_over_7_Cardnum	0.54	0.41	0	0.48	1
51	actual_over_total_over_7_Cardnum	0.36	0.37	0	0.19	1
52	count_of_Cardnum_over_7_days	7.63	16.61	1	4	369
53	unique_count_merchnum_each_cardnum_over_7	3.72	3.27	1	3	39
54	mean_amount_over_14_for_Cardnum	396.99	522.92	0.14	279.45	25500
55	median_amount_over_14_for_Cardnum	279.02	456.12	0.14	160	25500
56	max_amount_over_14_for_Cardnum	1188.64	1829.5	0.14	797.92	47900
57	min_amount_over_14_for_Cardnum	93.91	288.96	0.01	23.98	23940
58	range_amount_over_14_for_Cardnum	1094.73	1824.66	0	695.5	44800
59	range_percent_amount_over_14_for_Cardnum	131.12	1242.62	0	25.55	169499
60	total_amount_over_14_for_Cardnum	3768.18	9421.92	0.14	1723.64	313995.1
61	actual_over_mean_over_14_Cardnum	1	1.27	0	0.75	71.33
62	actual_over_median_over_14_Cardnum	3.47	29.25	0	1	6145.64
63	actual_over_max_over_14_Cardnum	0.44	0.4	0	0.27	1
64	actual_over_total_over_14_Cardnum	0.25	0.32	0	0.1	1
65	count_of_Cardnum_over_14_days	11.8	20.72	1	7	380
66	unique_count_merchnum_each_cardnum_over_14	5.55	5.04	1	4	52
67	mean_amount_over_30_for_Cardnum	396.58	479.34	0.17	304	25500
68	median_amount_over_30_for_Cardnum	250.94	402.43	0.17	159	25500
69	max_amount_over_30_for_Cardnum	1482.17	2076.88	0.17	1175.81	47900
70	min_amount_over_30_for_Cardnum	59.19	212.4	0.01	15.5	23911
71	range_amount_over_30_for_Cardnum	1422.99	2077.51	0	1114.2	47691.17
72	range_percent_amount_over_30_for_Cardnum	240.33	2140.58	0	62	169499
73	total_amount_over_30_for_Cardnum	6675.65	14591.2	0.17	3238.63	353997.3
74	actual_over_mean_over_30_Cardnum	1.01	1.61	0	0.6	137.99
75	actual_over_median_over_30_Cardnum	3.95	32.61	0	1	6288.56
76	actual_over_max_over_30_Cardnum	0.34	0.37	0	0.17	1
77	actual_over_total_over_30_Cardnum	0.16	0.25	0	0.05	1
78	count_of_Cardnum_over_30_days	20.36	30.91	1	12	426
79	unique_count_merchnum_each_cardnum_over_30	8.97	8.33	1	7	83
80	day_since_Merchnum	-5151.51	16779.5	-96749	0	358
81	mean_amount_over_0_for_Merchnum	395.56	764.16	0.01	158.97	28392.84
82	median_amount_over_0_for_Merchnum	381.94	755.73	0.01	147	28392.84
83	max_amount_over_0_for_Merchnum	503.73	998.74	0.01	193.7	47900
84	min_amount_over_0_for_Merchnum	326.89	714.43	0.01	100	28392.84
85	range_amount_over_0_for_Merchnum	176.84	712.8	0	0	44800

86	range_percent_amount_over_0_for_Merchnum	10.21	122.98	0	0	22352.5
87	total_amount_over_0_for_Merchnum	781.28	2844.81	0.01	275.44	217467.2
88	actual_over_mean_over_0_Merchnum	1	0.63	0	1	37.96
89	actual_over_median_over_0_Merchnum	1.25	3.63	0	1	543.6
90	actual_over_max_over_0_Merchnum	0.81	0.34	0	1	1
91	actual_over_total_over_0_Merchnum	0.72	0.39	0	1	1
92	count_of_Merchnum_over_0_days	6.81	19	1	1	260
93	mean_amount_over_1_for_Merchnum	397.17	751.05	0.01	174.6	28392.84
94	median_amount_over_1_for_Merchnum	371.92	740.6	0.01	149	28392.84
95	max_amount_over_1_for_Merchnum	586.79	1132.55	0.01	245	47900
96	min_amount_over_1_for_Merchnum	294.24	680.41	0.01	83.75	28392.84
97	range_amount_over_1_for_Merchnum	292.55	936.44	0	0	44800
98	range_percent_amount_over_1_for_Merchnum	19.58	143.91	0	0	22352.5
99	total_amount_over_1_for_Merchnum	1128.33	4359.44	0.01	396.53	306633.4
100	actual_over_mean_over_1_Merchnum	1	0.78	0	1	43.43
101	actual_over_median_over_1_Merchnum	1.39	4.73	0	1	533.42
102	actual_over_max_over_1_Merchnum	0.74	0.38	0	1	1
103	actual_over_total_over_1_Merchnum	0.64	0.42	0	1	1
104	count_of_Merchnum_over_1_days	11.58	31.66	1	1	327
105	mean_amount_over_3_for_Merchnum	397.26	732.06	0.01	188.26	28392.84
106	median_amount_over_3_for_Merchnum	362.34	720.48	0.01	150	28392.84
107	max_amount_over_3_for_Merchnum	671.74	1212.22	0.01	300	47900
108	min_amount_over_3_for_Merchnum	268.91	650.89	0.01	70	28392.84
109	range_amount_over_3_for_Merchnum	402.83	1066.4	0	25.25	44800
110	range_percent_amount_over_3_for_Merchnum	33.58	225.73	0	0.41	22352.5
111	total_amount_over_3_for_Merchnum	1585.04	5170.02	0.01	569.06	307302.6
112	actual_over_mean_over_3_Merchnum	1	0.98	0	1	77.68
113	actual_over_median_over_3_Merchnum	1.5	7.32	0	1	1553.38
114	actual_over_max_over_3_Merchnum	0.69	0.41	0	1	1
115	actual_over_total_over_3_Merchnum	0.57	0.43	0	0.66	1
116	count_of_Merchnum_over_3_days	21.03	55.25	1	2	466
117	mean_amount_over_7_for_Merchnum	396.21	690.02	0.01	205.91	27218
118	median_amount_over_7_for_Merchnum	346.72	677.25	0.01	153.41	27218
119	max_amount_over_7_for_Merchnum	823.39	1348.95	0.01	401.5	47900
120	min_amount_over_7_for_Merchnum	233.64	595.87	0.01	53	27218
121	range_amount_over_7_for_Merchnum	589.75	1264.54	0	122.87	44800
122	range_percent_amount_over_7_for_Merchnum	69.05	397.61	0	2.74	51119
123	total_amount_over_7_for_Merchnum	2617.12	6327.07	0.01	960	313984.6
124	actual_over_mean_over_7_Merchnum	1	1.28	0	1	106.21
125	actual_over_median_over_7_Merchnum	1.68	21.2	0	1	6145.64
126	actual_over_max_over_7_Merchnum	0.61	0.43	0	0.96	1
127	actual_over_total_over_7_Merchnum	0.48	0.43	0	0.34	1
128	count_of_Merchnum_over_7_days	42.18	106.32	1	3	762
129	mean_amount_over_14_for_Merchnum	397.34	661.3	0.01	223.9	27218

130	median_amount_over_14_for_Merchnum	335.97	643.4	0.01	155.43	27218
131	max_amount_over_14_for_Merchnum	983.88	1533.81	0.01	510.71	47900
132	min_amount_over_14_for_Merchnum	205.52	546.91	0.01	41.8	27218
133	range_amount_over_14_for_Merchnum	778.37	1490.83	0	253.37	44800
134	range_percent_amount_over_14_for_Merchnum	126.79	724.54	0	6.65	51119
135	total_amount_over_14_for_Merchnum	4307.64	8764.84	0.01	1435.2	319334.7
136	actual_over_mean_over_14_Merchnum	1	1.52	0	1	182.22
137	actual_over_median_over_14_Merchnum	1.77	26.08	0	1	7736.47
138	actual_over_max_over_14_Merchnum	0.56	0.43	0	0.59	1
139	actual_over_total_over_14_Merchnum	0.41	0.42	0	0.2	1
140	count_of_Merchnum_over_14_days	76.08	192.23	1	4	1091
141	mean_amount_over_30_for_Merchnum	396.84	618.56	0.01	240.2	27218
142	median_amount_over_30_for_Merchnum	321.94	588.59	0.01	157.5	27218
143	max_amount_over_30_for_Merchnum	1216.13	1829.81	0.01	750	47900
144	min_amount_over_30_for_Merchnum	176.06	501.21	0.01	30.4	27218
145	range_amount_over_30_for_Merchnum	1040.07	1818.12	0	453.07	44800
146	range_percent_amount_over_30_for_Merchnum	246.6	1282.44	0	14.11	54500
147	total_amount_over_30_for_Merchnum	7867.53	13989.5	0.01	2237.56	320373
148	actual_over_mean_over_30_Merchnum	1.01	2.99	0	0.99	786.56
149	actual_over_median_over_30_Merchnum	1.92	28.6	0	1	7736.47
150	actual_over_max_over_30_Merchnum	0.5	0.43	0	0.37	1
151	actual_over_total_over_30_Merchnum	0.34	0.41	0	0.1	1
152	count_of_Merchnum_over_30_days	146.69	377.03	1	6	1828
153	day_since_cardnum_merchdescription	-18947.5	28553.6	-96751	0	362
154	mean_amount_over_0_for_cardnum_merchdescription	395.69	795.36	0.01	141.3	28392.84
155	median_amount_over_0_for_cardnum_merchdescription	393.35	788.97	0.01	140	28392.84
156	max_amount_over_0_for_cardnum_merchdescription	421.19	935.85	0.01	148	47900
157	min_amount_over_0_for_cardnum_merchdescription	376.6	765.24	0.01	129.95	28392.84
158	range_amount_over_0_for_cardnum_merchdescription	44.59	522.78	0	0	44800
159	range_percent_amount_over_0_for_cardnum_merchdescription	1.78	106.63	0	0	22352.5
160	total_amount_over_0_for_cardnum_merchdescription	526.06	2618.5	0.01	155.57	217467.2
161	actual_over_mean_over_0_cardnum_merchdescription	1	0.18	0	1	7.64
162	actual_over_median_over_0_cardnum_merchdescription	1.02	0.6	0	1	100
163	actual_over_max_over_0_cardnum_merchdescription	0.96	0.15	0	1	1
164	actual_over_total_over_0_cardnum_merchdescription	0.9	0.26	0	1	1
165	count_of_cardnum_merchdescription_over_0_days	1.7	3.16	1	1	85
166	mean_amount_over_1_for_cardnum_merchdescription	397.21	798.08	0.01	144.37	28392.84
167	median_amount_over_1_for_cardnum_merchdescription	394.44	792.43	0.01	142.28	28392.84
168	max_amount_over_1_for_cardnum_merchdescription	432.63	1010.33	0.01	150.76	47900

169	min_amount_over_1_for_cardnum_merchdescription	371.24	760.44	0.01	126.9	28392.84
170	range_amount_over_1_for_cardnum_merchdescription	61.4	656.54	0	0	44800
171	range_percent_amount_over_1_for_cardnum_merchdescription	2.22	107.68	0	0	22352.5
172	total_amount_over_1_for_cardnum_merchdescription	595.01	4010.95	0.01	163	306633.4
173	actual_over_mean_over_1_cardnum_merchdescription	1	0.22	0	1	7.64
174	actual_over_median_over_1_cardnum_merchdescription	1.02	0.55	0	1	50.91
175	actual_over_max_over_1_cardnum_merchdescription	0.95	0.18	0	1	1
176	actual_over_total_over_1_cardnum_merchdescription	0.88	0.28	0	1	1
177	count_of_cardnum_merchdescription_over_1_days	1.81	3.41	1	1	85
178	mean_amount_over_3_for_cardnum_merchdescription	398.1	795.23	0.01	147.05	28392.84
179	median_amount_over_3_for_cardnum_merchdescription	394.58	789.94	0.01	144.66	28392.84
180	max_amount_over_3_for_cardnum_merchdescription	441.72	1014.99	0.01	157.5	47900
181	min_amount_over_3_for_cardnum_merchdescription	365.15	754.94	0.01	123.92	28392.84
182	range_amount_over_3_for_cardnum_merchdescription	76.57	671.78	0	0	44800
183	range_percent_amount_over_3_for_cardnum_merchdescription	2.71	108.2	0	0	22352.5
184	total_amount_over_3_for_cardnum_merchdescription	625.05	4054.24	0.01	171.99	306633.4
185	actual_over_mean_over_3_cardnum_merchdescription	0.99	0.25	0	1	7.64
186	actual_over_median_over_3_cardnum_merchdescription	1.03	1.17	0	1	301.1
187	actual_over_max_over_3_cardnum_merchdescription	0.94	0.2	0	1	1
188	actual_over_total_over_3_cardnum_merchdescription	0.85	0.3	0	1	1
189	count_of_cardnum_merchdescription_over_3_days	1.92	3.57	1	1	85
190	mean_amount_over_7_for_cardnum_merchdescription	399.91	790.52	0.01	151.47	28392.84
191	median_amount_over_7_for_cardnum_merchdescription	394.66	785.79	0.01	147.93	28392.84
192	max_amount_over_7_for_cardnum_merchdescription	459.6	1024.16	0.01	169	47900
193	min_amount_over_7_for_cardnum_merchdescription	354.62	746.41	0.01	117.16	28392.84
194	range_amount_over_7_for_cardnum_merchdescription	104.98	698.76	0	0	44800
195	range_percent_amount_over_7_for_cardnum_merchdescription	3.93	120.66	0	0	22352.5
196	total_amount_over_7_for_cardnum_merchdescription	681.79	4097.68	0.01	190	306633.4
197	actual_over_mean_over_7_cardnum_merchdescription	0.99	0.3	0	1	7.97
198	actual_over_median_over_7_cardnum_merchdescription	1.05	1.96	0	1	442.87
199	actual_over_max_over_7_cardnum_merchdescription	0.91	0.24	0	1	1
200	actual_over_total_over_7_cardnum_merchdescription	0.81	0.33	0	1	1

201	count_of_cardnum_merchdescription_over_7_days	2.1	3.84	1	1	90
202	mean_amount_over_14_for_cardnum_merchdescription	402.02	788.09	0.01	158.42	28392.84
203	median_amount_over_14_for_cardnum_merchdescription	394.58	783.26	0.01	149.99	28392.84
204	max_amount_over_14_for_cardnum_merchdescription	481.42	1046.77	0.01	182.06	47900
205	min_amount_over_14_for_cardnum_merchdescription	341.67	733.75	0.01	109.36	28392.84
206	range_amount_over_14_for_cardnum_merchdescription	139.75	743.68	0	0	44800
207	range_percent_amount_over_14_for_cardnum_merchdescription	5.32	125.34	0	0	22352.5
208	total_amount_over_14_for_cardnum_merchdescription	759.46	4167.47	0.01	210	306633.4
209	actual_over_mean_over_14_cardnum_merchdescription	0.98	0.34	0	1	8.94
210	actual_over_median_over_14_cardnum_merchdescription	1.07	2.1	0	1	400
211	actual_over_max_over_14_cardnum_merchdescription	0.89	0.26	0	1	1
212	actual_over_total_over_14_cardnum_merchdescription	0.77	0.35	0	1	1
213	count_of_cardnum_merchdescription_over_14_days	2.3	4.11	1	1	97
214	mean_amount_over_30_for_cardnum_merchdescription	404.2	780.77	0.01	166.89	28392.84
215	median_amount_over_30_for_cardnum_merchdescription	393.05	778.59	0.01	152	28392.84
216	max_amount_over_30_for_cardnum_merchdescription	514.94	1074.89	0.01	200.78	47900
217	min_amount_over_30_for_cardnum_merchdescription	322.71	713.89	0.01	100	28392.84
218	range_amount_over_30_for_cardnum_merchdescription	192.23	803.16	0	0	44800
219	range_percent_amount_over_30_for_cardnum_merchdescription	8.16	151.6	0	0	22352.5
220	total_amount_over_30_for_cardnum_merchdescription	908.66	4314.58	0.01	249	306633.4
221	actual_over_mean_over_30_cardnum_merchdescription	0.98	0.4	0	1	11.64
222	actual_over_median_over_30_cardnum_merchdescription	1.09	1.64	0	1	225.07
223	actual_over_max_over_30_cardnum_merchdescription	0.85	0.3	0	1	1
224	actual_over_total_over_30_cardnum_merchdescription	0.72	0.37	0	1	1
225	count_of_cardnum_merchdescription_over_30_days	2.7	4.77	1	1	99
226	day_since_cardnum_merchnum	-16969.1	27645.1	-96749	0	362
227	mean_amount_over_0_for_cardnum_merchnum	395.76	796.81	0.01	141.25	28392.84
228	median_amount_over_0_for_cardnum_merchnum	393.38	790.52	0.01	140	28392.84
229	max_amount_over_0_for_cardnum_merchnum	421.47	935.89	0.01	148	47900
230	min_amount_over_0_for_cardnum_merchnum	376.76	767.62	0.01	129.95	28392.84
231	range_amount_over_0_for_cardnum_merchnum	44.71	520.69	0	0	44800
232	range_percent_amount_over_0_for_cardnum_merchnum	1.87	106.66	0	0	22352.5
233	total_amount_over_0_for_cardnum_merchnum	529.23	2622.14	0.01	159.5	217467.2
234	actual_over_mean_over_0_cardnum_merchnum	1	0.22	0	1	20.24

235	actual_over_median_over_0_cardnum_merchnum	1.03	0.84	0	1	149.23
236	actual_over_max_over_0_cardnum_merchnum	0.96	0.17	0	1	1
237	actual_over_total_over_0_cardnum_merchnum	0.88	0.27	0	1	1
238	count_of_cardnum_merchnum_over_0_days	2.1	5.91	1	1	145
239	mean_amount_over_1_for_cardnum_merchnum	397.01	798.44	0.01	144.03	28392.84
240	median_amount_over_1_for_cardnum_merchnum	394.16	792.88	0.01	142	28392.84
241	max_amount_over_1_for_cardnum_merchnum	432.87	1010.41	0.01	150.53	47900
242	min_amount_over_1_for_cardnum_merchnum	371.13	761.51	0.01	126.87	28392.84
243	range_amount_over_1_for_cardnum_merchnum	61.74	656.46	0	0	44800
244	range_percent_amount_over_1_for_cardnum_merchnum	2.37	107.84	0	0	22352.5
245	total_amount_over_1_for_cardnum_merchnum	600.15	4020.57	0.01	168	306633.4
246	actual_over_mean_over_1_cardnum_merchnum	1	0.26	0	1	20.24
247	actual_over_median_over_1_cardnum_merchnum	1.04	1.26	0	1	254.05
248	actual_over_max_over_1_cardnum_merchnum	0.94	0.19	0	1	1
249	actual_over_total_over_1_cardnum_merchnum	0.86	0.3	0	1	1
250	count_of_cardnum_merchnum_over_1_days	2.42	7.59	1	1	177
251	mean_amount_over_3_for_cardnum_merchnum	398.1	797.31	0.01	146.97	28392.84
252	median_amount_over_3_for_cardnum_merchnum	394.53	792.06	0.01	144.07	28392.84
253	max_amount_over_3_for_cardnum_merchnum	442.21	1018.85	0.01	157.25	47900
254	min_amount_over_3_for_cardnum_merchnum	365.26	756.59	0.01	124	28392.84
255	range_amount_over_3_for_cardnum_merchnum	76.95	676.55	0	0	44800
256	range_percent_amount_over_3_for_cardnum_merchnum	3	110.09	0	0	22352.5
257	total_amount_over_3_for_cardnum_merchnum	632.57	4064.33	0.01	179.96	306633.4
258	actual_over_mean_over_3_cardnum_merchnum	0.99	0.3	0	1	20.24
259	actual_over_median_over_3_cardnum_merchnum	1.08	5.33	0	1	1553.38
260	actual_over_max_over_3_cardnum_merchnum	0.92	0.22	0	1	1
261	actual_over_total_over_3_cardnum_merchnum	0.83	0.33	0	1	1
262	count_of_cardnum_merchnum_over_3_days	3.03	10.98	1	1	248
263	mean_amount_over_7_for_cardnum_merchnum	399.9	788.36	0.01	151	27218
264	median_amount_over_7_for_cardnum_merchnum	394.12	781.6	0.01	147.4	27218
265	max_amount_over_7_for_cardnum_merchnum	461.23	1036.1	0.01	169	47900
266	min_amount_over_7_for_cardnum_merchnum	354.45	743.13	0.01	117.36	27218
267	range_amount_over_7_for_cardnum_merchnum	106.78	721.66	0	0	44800
268	range_percent_amount_over_7_for_cardnum_merchnum	4.68	128.83	0	0	22352.5
269	total_amount_over_7_for_cardnum_merchnum	694.1	4108.14	0.01	204	306633.4
270	actual_over_mean_over_7_cardnum_merchnum	0.99	0.37	0	1	20.24
271	actual_over_median_over_7_cardnum_merchnum	1.17	19.3	0	1	5747.54
272	actual_over_max_over_7_cardnum_merchnum	0.89	0.26	0	1	1
273	actual_over_total_over_7_cardnum_merchnum	0.78	0.36	0	1	1
274	count_of_cardnum_merchnum_over_7_days	4.05	15.65	1	1	358
275	mean_amount_over_14_for_cardnum_merchnum	401.66	783.92	0.01	158	27218

276	median_amount_over_14_for_cardnum_merchnum	393.79	779.03	0.01	149.78	27218
277	max_amount_over_14_for_cardnum_merchnum	483.99	1064.96	0.01	182.5	47900
278	min_amount_over_14_for_cardnum_merchnum	341.12	728.71	0.01	109.35	27218
279	range_amount_over_14_for_cardnum_merchnum	142.87	775.89	0	0	44800
280	range_percent_amount_over_14_for_cardnum_merchnum	6.46	138.58	0	0	22352.5
281	total_amount_over_14_for_cardnum_merchnum	778.3	4178.13	0.01	239.68	306633.4
282	actual_over_mean_over_14_cardnum_merchnum	0.99	0.44	0	1	23.11
283	actual_over_median_over_14_cardnum_merchnum	1.21	20.58	0	1	6145.64
284	actual_over_max_over_14_cardnum_merchnum	0.85	0.3	0	1	1
285	actual_over_total_over_14_cardnum_merchnum	0.73	0.38	0	1	1
286	count_of_cardnum_merchnum_over_14_days	5.35	19.04	1	1	369
287	mean_amount_over_30_for_cardnum_merchnum	403.76	777.44	0.01	166.15	27218
288	median_amount_over_30_for_cardnum_merchnum	392.27	776.24	0.01	151.37	27218
289	max_amount_over_30_for_cardnum_merchnum	519.55	1105.57	0.01	202	47900
290	min_amount_over_30_for_cardnum_merchnum	322.37	710.95	0.01	100	27218
291	range_amount_over_30_for_cardnum_merchnum	197.18	848.99	0	0	44800
292	range_percent_amount_over_30_for_cardnum_merchnum	10.21	180.11	0	0	22352.5
293	total_amount_over_30_for_cardnum_merchnum	938.99	4322.25	0.01	294.51	306633.4
294	actual_over_mean_over_30_cardnum_merchnum	0.98	0.53	0	1	25.03
295	actual_over_median_over_30_cardnum_merchnum	1.26	21.13	0	1	6288.56
296	actual_over_max_over_30_cardnum_merchnum	0.81	0.33	0	1	1
297	actual_over_total_over_30_cardnum_merchnum	0.67	0.4	0	1	1
298	count_of_cardnum_merchnum_over_30_days	7.76	27.43	1	1	409
299	day_since_cardnum_zip	-13819.5	25559.5	-96741	0	362
300	mean_amount_over_0_for_cardnum_zip	395.61	794.65	0.01	141.71	28392.84
301	median_amount_over_0_for_cardnum_zip	393.06	788.31	0.01	140	28392.84
302	max_amount_over_0_for_cardnum_zip	422.91	936.68	0.01	149	47900
303	min_amount_over_0_for_cardnum_zip	375.25	764.41	0.01	129	28392.84
304	range_amount_over_0_for_cardnum_zip	47.66	526.66	0	0	44800
305	range_percent_amount_over_0_for_cardnum_zip	2.01	106.78	0	0	22352.5
306	total_amount_over_0_for_cardnum_zip	532.32	2624.12	0.01	160.95	217467.2
307	actual_over_mean_over_0_cardnum_zip	1	0.23	0	1	20.24
308	actual_over_median_over_0_cardnum_zip	1.03	1.11	0	1	234.79
309	actual_over_max_over_0_cardnum_zip	0.95	0.17	0	1	1
310	actual_over_total_over_0_cardnum_zip	0.88	0.28	0	1	1
311	count_of_cardnum_zip_over_0_days	2.11	5.93	1	1	146
312	mean_amount_over_1_for_cardnum_zip	397.23	797.03	0.01	145	28392.84
313	median_amount_over_1_for_cardnum_zip	394.1	791.53	0.01	142.62	28392.84
314	max_amount_over_1_for_cardnum_zip	436.01	1012.47	0.01	153.73	47900
315	min_amount_over_1_for_cardnum_zip	368.91	758.81	0.01	125	28392.84
316	range_amount_over_1_for_cardnum_zip	67.09	662.97	0	0	44800
317	range_percent_amount_over_1_for_cardnum_zip	2.65	108.01	0	0	22352.5

318	total_amount_over_1_for_cardnum_zip	607.19	4024.02	0.01	171.95	306633.4
319	actual_over_mean_over_1_cardnum_zip	1	0.28	0	1	20.17
320	actual_over_median_over_1_cardnum_zip	1.06	2.27	0	1	543.6
321	actual_over_max_over_1_cardnum_zip	0.93	0.21	0	1	1
322	actual_over_total_over_1_cardnum_zip	0.85	0.31	0	1	1
323	count_of_cardnum_zip_over_1_days	2.47	7.82	1	1	177
324	mean_amount_over_3_for_cardnum_zip	398.11	793.61	0.01	148.74	28392.84
325	median_amount_over_3_for_cardnum_zip	393.91	788.56	0.01	144.98	28392.84
326	max_amount_over_3_for_cardnum_zip	447.52	1018.95	0.01	160.62	47900
327	min_amount_over_3_for_cardnum_zip	361.65	752.56	0.01	120.8	28392.84
328	range_amount_over_3_for_cardnum_zip	85.87	682.74	0	0	44800
329	range_percent_amount_over_3_for_cardnum_zip	3.5	108.8	0	0	22352.5
330	total_amount_over_3_for_cardnum_zip	644.74	4069.19	0.01	186.4	306633.4
331	actual_over_mean_over_3_cardnum_zip	0.99	0.33	0	1	20.17
332	actual_over_median_over_3_cardnum_zip	1.08	2.38	0	1	387.76
333	actual_over_max_over_3_cardnum_zip	0.91	0.24	0	1	1
334	actual_over_total_over_3_cardnum_zip	0.81	0.33	0	1	1
335	count_of_cardnum_zip_over_3_days	3.13	11.28	1	1	251
336	mean_amount_over_7_for_cardnum_zip	400.03	787.37	0.01	154.08	28392.84
337	median_amount_over_7_for_cardnum_zip	393.52	782.91	0.01	148.45	28392.84
338	max_amount_over_7_for_cardnum_zip	469.6	1030.15	0.01	175.97	47900
339	min_amount_over_7_for_cardnum_zip	349.15	741.42	0.01	113	28392.84
340	range_amount_over_7_for_cardnum_zip	120.44	715.49	0	0	44800
341	range_percent_amount_over_7_for_cardnum_zip	5.29	121.63	0	0	22352.5
342	total_amount_over_7_for_cardnum_zip	715.3	4115.82	0.01	216.9	306633.4
343	actual_over_mean_over_7_cardnum_zip	0.99	0.41	0	1	32.76
344	actual_over_median_over_7_cardnum_zip	1.13	3.42	0	1	442.87
345	actual_over_max_over_7_cardnum_zip	0.87	0.28	0	1	1
346	actual_over_total_over_7_cardnum_zip	0.76	0.37	0	1	1
347	count_of_cardnum_zip_over_7_days	4.25	16.31	1	1	369
348	mean_amount_over_14_for_cardnum_zip	402.25	784.94	0.01	160.89	28392.84
349	median_amount_over_14_for_cardnum_zip	392.79	780.21	0.01	150	28392.84
350	max_amount_over_14_for_cardnum_zip	497.08	1056.84	0.01	194.8	47900
351	min_amount_over_14_for_cardnum_zip	334.03	728.25	0.01	104.9	28392.84
352	range_amount_over_14_for_cardnum_zip	163.05	767.26	0	0	44800
353	range_percent_amount_over_14_for_cardnum_zip	7.55	127.47	0	0	22352.5
354	total_amount_over_14_for_cardnum_zip	812.72	4191.54	0.01	259.97	306633.4
355	actual_over_mean_over_14_cardnum_zip	0.98	0.49	0	1	32.76
356	actual_over_median_over_14_cardnum_zip	1.18	3.82	0	1	543.6
357	actual_over_max_over_14_cardnum_zip	0.83	0.32	0	1	1
358	actual_over_total_over_14_cardnum_zip	0.71	0.39	0	1	1
359	count_of_cardnum_zip_over_14_days	5.71	19.98	1	1	380
360	mean_amount_over_30_for_cardnum_zip	404.28	775.42	0.01	171.25	28392.84
361	median_amount_over_30_for_cardnum_zip	389.53	773.66	0.01	152.1	28392.84

362	max_amount_over_30_for_cardnum_zip	540.35	1088.98	0.01	220.89	47900
363	min_amount_over_30_for_cardnum_zip	311.11	705.03	0.01	94.88	28392.84
364	range_amount_over_30_for_cardnum_zip	229.24	836.73	0	0	44800
365	range_percent_amount_over_30_for_cardnum_zip	12.1	155.06	0	0	22352.5
366	total_amount_over_30_for_cardnum_zip	1003.04	4355.88	0.01	329	306633.4
367	actual_over_mean_over_30_cardnum_zip	0.98	0.59	0	1	32.76
368	actual_over_median_over_30_cardnum_zip	1.28	8.51	0	1	2248.7
369	actual_over_max_over_30_cardnum_zip	0.77	0.35	0	1	1
370	actual_over_total_over_30_cardnum_zip	0.64	0.41	0	1	1
371	count_of_cardnum_zip_over_30_days	8.44	28.95	1	1	425
372	day_since_cardnum_state	-5563.69	17079.4	-96741	3	356
373	mean_amount_over_0_for_cardnum_state	395.51	797.75	0.01	144.2	47900
374	median_amount_over_0_for_cardnum_state	392.17	791.54	0.01	141.71	47900
375	max_amount_over_0_for_cardnum_state	432.22	944.24	0.01	152.51	47900
376	min_amount_over_0_for_cardnum_state	366.87	767	0.01	124	47900
377	range_amount_over_0_for_cardnum_state	65.35	540.93	0	0	29807.71
378	range_percent_amount_over_0_for_cardnum_state	2.57	107.95	0	0	22352.5
379	total_amount_over_0_for_cardnum_state	553.64	2640.1	0.01	169	217467.2
380	actual_over_mean_over_0_cardnum_state	1	0.26	0	1	20.24
381	actual_over_median_over_0_cardnum_state	1.05	1.55	0	1	234.79
382	actual_over_max_over_0_cardnum_state	0.94	0.19	0	1	1
383	actual_over_total_over_0_cardnum_state	0.86	0.29	0	1	1
384	count_of_cardnum_state_over_0_days	2.16	5.94	1	1	146
385	mean_amount_over_1_for_cardnum_state	397.06	795.35	0.01	150	47900
386	median_amount_over_1_for_cardnum_state	391.45	789.76	0.01	146	47900
387	max_amount_over_1_for_cardnum_state	457.49	1029.97	0.01	165.46	47900
388	min_amount_over_1_for_cardnum_state	350.92	751.58	0.01	114.93	47900
389	range_amount_over_1_for_cardnum_state	106.57	702.63	0	0	30199.27
390	range_percent_amount_over_1_for_cardnum_state	3.93	110.2	0	0	22352.5
391	total_amount_over_1_for_cardnum_state	658.47	4053.65	0.01	190	306633.4
392	actual_over_mean_over_1_cardnum_state	1	0.33	0	1	20.17
393	actual_over_median_over_1_cardnum_state	1.08	1.73	0	1	231.59
394	actual_over_max_over_1_cardnum_state	0.91	0.24	0	1	1
395	actual_over_total_over_1_cardnum_state	0.81	0.33	0	1	1
396	count_of_cardnum_state_over_1_days	2.59	7.82	1	1	177
397	mean_amount_over_3_for_cardnum_state	397.79	781.98	0.01	157.51	47900
398	median_amount_over_3_for_cardnum_state	388.09	775.43	0.01	149	47900
399	max_amount_over_3_for_cardnum_state	485.92	1055.71	0.01	181.82	47900
400	min_amount_over_3_for_cardnum_state	331.72	727.65	0.01	104.9	47900
401	range_amount_over_3_for_cardnum_state	154.2	770.6	0	0	30199.27
402	range_percent_amount_over_3_for_cardnum_state	6.03	112.54	0	0	22352.5
403	total_amount_over_3_for_cardnum_state	738.12	4120.51	0.01	216.75	306633.4
404	actual_over_mean_over_3_cardnum_state	0.99	0.4	0	1	20.17

405	actual_over_median_over_3_cardnum_state	1.13	2.38	0	1	301.1
406	actual_over_max_over_3_cardnum_state	0.87	0.28	0	1	1
407	actual_over_total_over_3_cardnum_state	0.76	0.36	0	1	1
408	count_of_cardnum_state_over_3_days	3.33	11.27	1	1	251
409	mean_amount_over_7_for_cardnum_state	399.86	765.85	0.01	171.85	47900
410	median_amount_over_7_for_cardnum_state	381.39	755.61	0.01	152.89	47900
411	max_amount_over_7_for_cardnum_state	541.12	1122.34	0.01	215	47900
412	min_amount_over_7_for_cardnum_state	301.53	691.91	0.01	91.16	47900
413	range_amount_over_7_for_cardnum_state	239.58	894.58	0	0	30199.27
414	range_percent_amount_over_7_for_cardnum_state	10.4	128.78	0	0	22352.5
415	total_amount_over_7_for_cardnum_state	904.11	4247.12	0.01	280.56	306633.4
416	actual_over_mean_over_7_cardnum_state	0.99	0.52	0	1	32.76
417	actual_over_median_over_7_cardnum_state	1.23	3.74	0	1	442.87
418	actual_over_max_over_7_cardnum_state	0.8	0.33	0	1	1
419	actual_over_total_over_7_cardnum_state	0.67	0.39	0	1	1
420	count_of_cardnum_state_over_7_days	4.67	16.29	1	1	369
421	mean_amount_over_14_for_cardnum_state	401.85	745.8	0.01	187.75	47900
422	median_amount_over_14_for_cardnum_state	373.26	737.61	0.01	156.47	47900
423	max_amount_over_14_for_cardnum_state	606.28	1191.47	0.01	259.2	47900
424	min_amount_over_14_for_cardnum_state	270.46	645.37	0.01	76.9	47900
425	range_amount_over_14_for_cardnum_state	335.82	1016.74	0	6.38	30199.27
426	range_percent_amount_over_14_for_cardnum_state	16.73	141.44	0	0.35	22352.5
427	total_amount_over_14_for_cardnum_state	1154.78	4510.89	0.01	367.06	306633.4
428	actual_over_mean_over_14_cardnum_state	0.99	0.63	0	1	32.76
429	actual_over_median_over_14_cardnum_state	1.34	4.25	0	1	416.25
430	actual_over_max_over_14_cardnum_state	0.74	0.37	0	1	1
431	actual_over_total_over_14_cardnum_state	0.59	0.41	0	0.64	1
432	count_of_cardnum_state_over_14_days	6.46	19.96	1	2	380
433	mean_amount_over_30_for_cardnum_state	402.94	703.28	0.01	207.75	47900
434	median_amount_over_30_for_cardnum_state	358.28	689.17	0.01	158.9	47900
435	max_amount_over_30_for_cardnum_state	711.31	1301.05	0.01	327.8	47900
436	min_amount_over_30_for_cardnum_state	229.38	573.81	0.01	59.55	47900
437	range_amount_over_30_for_cardnum_state	481.92	1187.69	0	55.28	30199.27
438	range_percent_amount_over_30_for_cardnum_state	31.62	392.13	0	1.99	106735.5
439	total_amount_over_30_for_cardnum_state	1676.04	5194.15	0.01	542.02	306633.4
440	actual_over_mean_over_30_cardnum_state	0.99	0.76	0	1	32.76
441	actual_over_median_over_30_cardnum_state	1.48	4.93	0	1	452.86
442	actual_over_max_over_30_cardnum_state	0.66	0.39	0	1	1
443	actual_over_total_over_30_cardnum_state	0.49	0.41	0	0.37	1
444	count_of_cardnum_state_over_30_days	9.9	28.92	1	3	425
445	velocity_0_over_7_on_Cardnum_amount	0.99	0.92	0	1	59.87
446	velocity_0_over_7_on_Cardnum_count	0.47	0.32	0	0.37	1
447	velocity_0_over_14_on_Cardnum_amount	0.99	1.1	0	0.86	65.53

448	velocity_0_over_14_on_Cardnum_count	0.34	0.29	0	0.25	1
449	velocity_0_over_30_on_Cardnum_amount	1	1.33	0	0.7	70.96
450	velocity_0_over_30_on_Cardnum_count	0.22	0.24	0	0.13	1
451	velocity_1_over_7_on_Cardnum_amount	1	0.75	0	1	59.87
452	velocity_1_over_7_on_Cardnum_count	0.57	0.3	0	0.5	1
453	velocity_1_over_14_on_Cardnum_amount	1	0.93	0	0.97	50.2
454	velocity_1_over_14_on_Cardnum_count	0.4	0.29	0	0.33	1
455	velocity_1_over_30_on_Cardnum_amount	1.01	1.19	0	0.81	70.96
456	velocity_1_over_30_on_Cardnum_count	0.26	0.25	0	0.17	1
457	velocity_0_over_7_on_Merchnum_amount	1	0.76	0	1	62.88
458	velocity_0_over_7_on_Merchnum_count	0.58	0.38	0	0.5	1
459	velocity_0_over_14_on_Merchnum_amount	1	1.06	0	1	182.22
460	velocity_0_over_14_on_Merchnum_count	0.49	0.39	0	0.33	1
461	velocity_0_over_30_on_Merchnum_amount	1.01	2.73	0	1	786.56
462	velocity_0_over_30_on_Merchnum_count	0.4	0.39	0	0.22	1
463	velocity_1_over_7_on_Merchnum_amount	1.01	0.62	0	1	51.94
464	velocity_1_over_7_on_Merchnum_count	0.65	0.34	0	0.67	1
465	velocity_1_over_14_on_Merchnum_amount	1.01	0.92	0	1	182.22
466	velocity_1_over_14_on_Merchnum_count	0.53	0.37	0	0.5	1
467	velocity_1_over_30_on_Merchnum_amount	1.02	2.66	0	1	786.56
468	velocity_1_over_30_on_Merchnum_count	0.43	0.38	0	0.25	1
469	velocity_0_over_7_on_cardnum_merchdescription_amount	0.99	0.22	0	1	6.21
470	velocity_0_over_7_on_cardnum_merchdescription_count	0.91	0.21	0.02	1	1
471	velocity_0_over_14_on_cardnum_merchdescription_amount	0.98	0.28	0	1	8.94
472	velocity_0_over_14_on_cardnum_merchdescription_count	0.87	0.25	0.02	1	1
473	velocity_0_over_30_on_cardnum_merchdescription_amount	0.98	0.35	0	1	10.65
474	velocity_0_over_30_on_cardnum_merchdescription_count	0.81	0.29	0.02	1	1
475	velocity_1_over_7_on_cardnum_merchdescription_amount	0.99	0.19	0	1	6.02
476	velocity_1_over_7_on_cardnum_merchdescription_count	0.93	0.18	0.02	1	1
477	velocity_1_over_14_on_cardnum_merchdescription_amount	0.99	0.26	0	1	8.94
478	velocity_1_over_14_on_cardnum_merchdescription_count	0.89	0.23	0.02	1	1
479	velocity_1_over_30_on_cardnum_merchdescription_amount	0.98	0.32	0	1	8.94
480	velocity_1_over_30_on_cardnum_merchdescription_count	0.83	0.28	0.02	1	1
481	velocity_0_over_7_on_cardnum_merchnum_amount	0.99	0.27	0	1	13.72
482	velocity_0_over_7_on_cardnum_merchnum_count	0.87	0.26	0	1	1
483	velocity_0_over_14_on_cardnum_merchnum_amount	0.99	0.35	0	1	20.83
484	velocity_0_over_14_on_cardnum_merchnum_count	0.81	0.3	0	1	1
485	velocity_0_over_30_on_cardnum_merchnum_amount	0.98	0.45	0	1	24.06

486	velocity_0_over_30_on_cardnum_merchnum_count	0.74	0.34	0	1	1
487	velocity_1_over_7_on_cardnum_merchnum_amount	0.99	0.23	0	1	13.72
488	velocity_1_over_7_on_cardnum_merchnum_count	0.89	0.23	0	1	1
489	velocity_1_over_14_on_cardnum_merchnum_amount	0.99	0.32	0	1	20.83
490	velocity_1_over_14_on_cardnum_merchnum_count	0.83	0.29	0	1	1
491	velocity_1_over_30_on_cardnum_merchnum_amount	0.99	0.41	0	1	24.06
492	velocity_1_over_30_on_cardnum_merchnum_count	0.76	0.33	0	1	1
493	velocity_0_over_7_on_cardnum_zip_amount	0.99	0.3	0	1	13.72
494	velocity_0_over_7_on_cardnum_zip_count	0.85	0.27	0	1	1
495	velocity_0_over_14_on_cardnum_zip_amount	0.98	0.39	0	1	21.65
496	velocity_0_over_14_on_cardnum_zip_count	0.79	0.32	0	1	1
497	velocity_0_over_30_on_cardnum_zip_amount	0.98	0.5	0	1	24.85
498	velocity_0_over_30_on_cardnum_zip_count	0.71	0.35	0	1	1
499	velocity_1_over_7_on_cardnum_zip_amount	0.99	0.26	0	1	13.72
500	velocity_1_over_7_on_cardnum_zip_count	0.88	0.25	0	1	1
501	velocity_1_over_14_on_cardnum_zip_amount	0.99	0.35	0	1	21.65
502	velocity_1_over_14_on_cardnum_zip_count	0.81	0.3	0	1	1
503	velocity_1_over_30_on_cardnum_zip_amount	0.98	0.46	0	1	24.85
504	velocity_1_over_30_on_cardnum_zip_count	0.73	0.34	0	1	1
505	velocity_0_over_7_on_cardnum_state_amount	0.99	0.42	0	1	18.39
506	velocity_0_over_7_on_cardnum_state_count	0.77	0.31	0	1	1
507	velocity_0_over_14_on_cardnum_state_amount	0.99	0.54	0	1	22.34
508	velocity_0_over_14_on_cardnum_state_count	0.67	0.35	0	0.75	1
509	velocity_0_over_30_on_cardnum_state_amount	0.99	0.67	0	1	27.26
510	velocity_0_over_30_on_cardnum_state_count	0.56	0.36	0	0.5	1
511	velocity_1_over_7_on_cardnum_state_amount	0.99	0.36	0	1	18.39
512	velocity_1_over_7_on_cardnum_state_count	0.81	0.28	0	1	1
513	velocity_1_over_14_on_cardnum_state_amount	0.99	0.48	0	1	21.7
514	velocity_1_over_14_on_cardnum_state_count	0.71	0.33	0	1	1
515	velocity_1_over_30_on_cardnum_state_amount	0.99	0.62	0	1	27.26
517	velocity_1_over_30_on_cardnum_state_count	0.58	0.36	0	0.5	1

9.3 Full Final Results of Model Selected

Table 9.4 Training Population Results

Training	# Records		# Goods		# Bads		Fraud Rate					
	64658		63964		694		0.010733397					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	647	222	425	34.31%	65.69%	647	222	425	0.35%	61.24%	60.89	0.52
2	647	558	89	86.24%	13.76%	1294	780	514	1.22%	74.06%	72.84	1.52
3	647	607	40	93.82%	6.18%	1941	1387	554	2.17%	79.83%	77.66	2.50
4	647	631	16	97.53%	2.47%	2588	2018	570	3.15%	82.13%	78.98	3.54
5	647	632	15	97.68%	2.32%	3235	2650	585	4.14%	84.29%	80.15	4.53
6	647	632	15	97.68%	2.32%	3882	3282	600	5.13%	86.46%	81.32	5.47
7	647	639	8	98.76%	1.24%	4529	3921	608	6.13%	87.61%	81.48	6.45
8	647	633	14	97.84%	2.16%	5176	4554	622	7.12%	89.63%	82.51	7.32
9	647	641	6	99.07%	0.93%	5823	5195	628	8.12%	90.49%	82.37	8.27
10	647	642	5	99.23%	0.77%	6470	5837	633	9.13%	91.21%	82.08	9.22
11	647	645	2	99.69%	0.31%	7117	6482	635	10.13%	91.50%	81.36	10.21
12	647	642	5	99.23%	0.77%	7764	7124	640	11.14%	92.22%	81.08	11.13
13	647	645	2	99.69%	0.31%	8411	7769	642	12.15%	92.51%	80.36	12.10
14	647	642	5	99.23%	0.77%	9058	8411	647	13.15%	93.23%	80.08	13.00
15	647	645	2	99.69%	0.31%	9705	9056	649	14.16%	93.52%	79.36	13.95
16	647	646	1	99.85%	0.15%	10352	9702	650	15.17%	93.66%	78.49	14.93
17	647	645	2	99.69%	0.31%	10999	10347	652	16.18%	93.95%	77.77	15.87
18	647	645	2	99.69%	0.31%	11646	10992	654	17.18%	94.24%	77.05	16.81
19	647	646	1	99.85%	0.15%	12293	11638	655	18.19%	94.38%	76.19	17.77
20	647	644	3	99.54%	0.46%	12940	12282	658	19.20%	94.81%	75.61	18.67
21	647	644	3	99.54%	0.46%	13587	12926	661	20.21%	95.24%	75.04	19.56
22	647	645	2	99.69%	0.31%	14234	13571	663	21.22%	95.53%	74.32	20.47
23	647	645	2	99.69%	0.31%	14881	14216	665	22.23%	95.82%	73.60	21.38
24	647	647	0	100.00%	0.00%	15528	14863	665	23.24%	95.82%	72.58	22.35
25	647	644	3	99.54%	0.46%	16175	15507	668	24.24%	96.25%	72.01	23.21
26	647	647	0	100.00%	0.00%	16822	16154	668	25.25%	96.25%	71.00	24.18
27	647	646	1	99.85%	0.15%	17469	16800	669	26.26%	96.40%	70.13	25.11
28	647	646	1	99.85%	0.15%	18116	17446	670	27.27%	96.54%	69.27	26.04
29	647	646	1	99.85%	0.15%	18763	18092	671	28.28%	96.69%	68.40	26.96
30	647	645	2	99.69%	0.31%	19410	18737	673	29.29%	96.97%	67.68	27.84
31	647	646	1	99.85%	0.15%	20057	19383	674	30.30%	97.12%	66.82	28.76
32	647	647	0	100.00%	0.00%	20704	20030	674	31.31%	97.12%	65.80	29.72
33	647	645	2	99.69%	0.31%	21351	20675	676	32.32%	97.41%	65.08	30.58
34	647	647	0	100.00%	0.00%	21998	21322	676	33.33%	97.41%	64.07	31.54
35	647	647	0	100.00%	0.00%	22645	21969	676	34.35%	97.41%	63.06	32.50
36	647	647	0	100.00%	0.00%	23292	22616	676	35.36%	97.41%	62.05	33.46
37	647	647	0	100.00%	0.00%	23939	23263	676	36.37%	97.41%	61.04	34.41
38	647	646	1	99.85%	0.15%	24586	23909	677	37.38%	97.55%	60.17	35.32
39	647	646	1	99.85%	0.15%	25233	24555	678	38.39%	97.69%	59.31	36.22
40	647	646	1	99.85%	0.15%	25880	25201	679	39.40%	97.84%	58.44	37.11
41	647	646	1	99.85%	0.15%	26527	25847	680	40.41%	97.98%	57.57	38.01
42	647	647	0	100.00%	0.00%	27174	26494	680	41.42%	97.98%	56.56	38.96
43	647	646	1	99.85%	0.15%	27821	27140	681	42.43%	98.13%	55.70	39.85
44	647	646	1	99.85%	0.15%	28468	27786	682	43.44%	98.27%	54.83	40.74
45	647	646	1	99.85%	0.15%	29115	28432	683	44.45%	98.41%	53.96	41.63
46	647	647	0	100.00%	0.00%	29762	29079	683	45.46%	98.41%	52.95	42.58
47	647	647	0	100.00%	0.00%	30409	29726	683	46.47%	98.41%	51.94	43.52
48	647	646	1	99.85%	0.15%	31056	30372	684	47.48%	98.56%	51.08	44.40
49	647	647	0	100.00%	0.00%	31703	31019	684	48.49%	98.56%	50.06	45.35
50	647	647	0	100.00%	0.00%	32350	31666	684	49.51%	98.56%	49.05	46.30
51	647	647	0	100.00%	0.00%	32997	32313	684	50.52%	98.56%	48.04	47.24
52	647	647	0	100.00%	0.00%	33644	32960	684	51.53%	98.56%	47.03	48.19
53	647	647	0	100.00%	0.00%	34291	33607	684	52.54%	98.56%	46.02	49.13
54	647	646	1	99.85%	0.15%	34938	34253	685	53.55%	98.70%	45.15	50.00
55	647	646	1	99.85%	0.15%	35585	34899	686	54.56%	98.85%	44.29	50.87
56	647	646	1	99.85%	0.15%	36232	35545	687	55.57%	98.99%	43.42	51.74
57	647	647	0	100.00%	0.00%	36879	36192	687	56.58%	98.99%	42.41	52.68
58	647	647	0	100.00%	0.00%	37526	36839	687	57.59%	98.99%	41.40	53.62
59	647	647	0	100.00%	0.00%	38173	37486	687	58.60%	98.99%	40.39	54.56
60	647	647	0	100.00%	0.00%	38820	38133	687	59.62%	98.99%	39.38	55.51

Training	# Records		# Goods		# Bads		Fraud Rate					
	64658		63964		694		0.010733397					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
61	647	646	1	99.85%	0.15%	39467	38779	688	60.63%	99.14%	38.51	56.36
62	647	647	0	100.00%	0.00%	40114	39426	688	61.64%	99.14%	37.50	57.31
63	647	647	0	100.00%	0.00%	40761	40073	688	62.65%	99.14%	36.49	58.25
64	647	645	2	99.69%	0.31%	41408	40718	690	63.66%	99.42%	35.77	59.01
65	647	647	0	100.00%	0.00%	42055	41365	690	64.67%	99.42%	34.75	59.95
66	647	647	0	100.00%	0.00%	42702	42012	690	65.68%	99.42%	33.74	60.89
67	647	647	0	100.00%	0.00%	43349	42659	690	66.69%	99.42%	32.73	61.82
68	647	646	1	99.85%	0.15%	43996	43305	691	67.70%	99.57%	31.87	62.67
69	647	647	0	100.00%	0.00%	44643	43952	691	68.71%	99.57%	30.85	63.61
70	647	647	0	100.00%	0.00%	45290	44599	691	69.73%	99.57%	29.84	64.54
71	647	647	0	100.00%	0.00%	45937	45246	691	70.74%	99.57%	28.83	65.48
72	647	647	0	100.00%	0.00%	46584	45893	691	71.75%	99.57%	27.82	66.42
73	647	646	1	99.85%	0.15%	47231	46539	692	72.76%	99.71%	26.95	67.25
74	647	646	1	99.85%	0.15%	47878	47185	693	73.77%	99.86%	26.09	68.09
75	647	647	0	100.00%	0.00%	48525	47832	693	74.78%	99.86%	25.08	69.02
76	647	647	0	100.00%	0.00%	49172	48479	693	75.79%	99.86%	24.06	69.96
77	647	647	0	100.00%	0.00%	49819	49126	693	76.80%	99.86%	23.05	70.89
78	647	647	0	100.00%	0.00%	50466	49773	693	77.81%	99.86%	22.04	71.82
79	647	647	0	100.00%	0.00%	51113	50420	693	78.83%	99.86%	21.03	72.76
80	647	647	0	100.00%	0.00%	51760	51067	693	79.84%	99.86%	20.02	73.69
81	647	647	0	100.00%	0.00%	52407	51714	693	80.85%	99.86%	19.01	74.62
82	647	647	0	100.00%	0.00%	53054	52361	693	81.86%	99.86%	18.00	75.56
83	647	647	0	100.00%	0.00%	53701	53008	693	82.87%	99.86%	16.98	76.49
84	647	647	0	100.00%	0.00%	54348	53655	693	83.88%	99.86%	15.97	77.42
85	647	647	0	100.00%	0.00%	54995	54302	693	84.89%	99.86%	14.96	78.36
86	647	646	1	99.85%	0.15%	55642	54948	694	85.90%	100.00%	14.10	79.18
87	647	647	0	100.00%	0.00%	56289	55595	694	86.92%	100.00%	13.08	80.11
88	647	647	0	100.00%	0.00%	56936	56242	694	87.93%	100.00%	12.07	81.04
89	647	647	0	100.00%	0.00%	57583	56889	694	88.94%	100.00%	11.06	81.97
90	647	647	0	100.00%	0.00%	58230	57536	694	89.95%	100.00%	10.05	82.90
91	647	647	0	100.00%	0.00%	58877	58183	694	90.96%	100.00%	9.04	83.84
92	647	647	0	100.00%	0.00%	59524	58830	694	91.97%	100.00%	8.03	84.77
93	647	647	0	100.00%	0.00%	60171	59477	694	92.99%	100.00%	7.01	85.70
94	647	647	0	100.00%	0.00%	60818	60124	694	94.00%	100.00%	6.00	86.63
95	647	647	0	100.00%	0.00%	61465	60771	694	95.01%	100.00%	4.99	87.57
96	647	647	0	100.00%	0.00%	62112	61418	694	96.02%	100.00%	3.98	88.50
97	647	647	0	100.00%	0.00%	62759	62065	694	97.03%	100.00%	2.97	89.43
98	647	647	0	100.00%	0.00%	63406	62712	694	98.04%	100.00%	1.96	90.36
99	647	647	0	100.00%	0.00%	64053	63359	694	99.05%	100.00%	0.95	91.30
100	605	605	0	100.00%	0.00%	64658	63964	694	100.00%	100.00%	-	92.17

Table 9.5 Testing Population Results

Testing	# Records		# Goods		# Bads		Fraud Rate					
	16165		15991		174		0.010763996					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	162	58	104	35.80%	64.20%	162	58	104	0.36%	59.77%	59.41	0.56
2	162	140	22	86.42%	13.58%	324	198	126	1.24%	72.41%	71.18	1.57
3	162	152	10	93.83%	6.17%	486	350	136	2.19%	78.16%	75.97	2.57
4	162	156	6	96.30%	3.70%	648	506	142	3.16%	81.61%	78.44	3.56
5	162	159	3	98.15%	1.85%	810	665	145	4.16%	83.33%	79.17	4.59
6	162	156	6	96.30%	3.70%	972	821	151	5.13%	86.78%	81.65	5.44
7	162	157	5	96.91%	3.09%	1134	978	156	6.12%	89.66%	83.54	6.27
8	162	159	3	98.15%	1.85%	1296	1137	159	7.11%	91.38%	84.27	7.15
9	162	159	3	98.15%	1.85%	1458	1296	162	8.10%	93.10%	85.00	8.00
10	162	161	1	99.38%	0.62%	1620	1457	163	9.11%	93.68%	84.57	8.94
11	162	161	1	99.38%	0.62%	1782	1618	164	10.12%	94.25%	84.13	9.87
12	162	162	0	100.00%	0.00%	1944	1780	164	11.13%	94.25%	83.12	10.85
13	162	162	0	100.00%	0.00%	2106	1942	164	12.14%	94.25%	82.11	11.84
14	162	162	0	100.00%	0.00%	2268	2104	164	13.16%	94.25%	81.10	12.83
15	162	161	1	99.38%	0.62%	2430	2265	165	14.16%	94.83%	80.66	13.73
16	162	161	1	99.38%	0.62%	2592	2426	166	15.17%	95.40%	80.23	14.61
17	162	161	1	99.38%	0.62%	2754	2587	167	16.18%	95.98%	79.80	15.49
18	162	162	0	100.00%	0.00%	2916	2749	167	17.19%	95.98%	78.79	16.46
19	162	159	3	98.15%	1.85%	3078	2908	170	18.19%	97.70%	79.52	17.11
20	162	162	0	100.00%	0.00%	3240	3070	170	19.20%	97.70%	78.50	18.06
21	162	162	0	100.00%	0.00%	3402	3232	170	20.21%	97.70%	77.49	19.01
22	162	162	0	100.00%	0.00%	3564	3394	170	21.22%	97.70%	76.48	19.96
23	162	162	0	100.00%	0.00%	3726	3556	170	22.24%	97.70%	75.46	20.92
24	162	162	0	100.00%	0.00%	3888	3718	170	23.25%	97.70%	74.45	21.87
25	162	161	1	99.38%	0.62%	4050	3879	171	24.26%	98.28%	74.02	22.68
26	162	162	0	100.00%	0.00%	4212	4041	171	25.27%	98.28%	73.01	23.63
27	162	162	0	100.00%	0.00%	4374	4203	171	26.28%	98.28%	71.99	24.58
28	162	162	0	100.00%	0.00%	4536	4365	171	27.30%	98.28%	70.98	25.53
29	162	162	0	100.00%	0.00%	4698	4527	171	28.31%	98.28%	69.97	26.47
30	162	161	1	99.38%	0.62%	4860	4688	172	29.32%	98.85%	69.53	27.26
31	162	162	0	100.00%	0.00%	5022	4850	172	30.33%	98.85%	68.52	28.20
32	162	162	0	100.00%	0.00%	5184	5012	172	31.34%	98.85%	67.51	29.14
33	162	162	0	100.00%	0.00%	5346	5174	172	32.36%	98.85%	66.49	30.08
34	162	162	0	100.00%	0.00%	5508	5336	172	33.37%	98.85%	65.48	31.02
35	162	162	0	100.00%	0.00%	5670	5498	172	34.38%	98.85%	64.47	31.97
36	162	162	0	100.00%	0.00%	5832	5660	172	35.39%	98.85%	63.46	32.91
37	162	161	1	99.38%	0.62%	5994	5821	173	36.40%	99.43%	63.02	33.65
38	162	162	0	100.00%	0.00%	6156	5983	173	37.41%	99.43%	62.01	34.58
39	162	162	0	100.00%	0.00%	6318	6145	173	38.43%	99.43%	61.00	35.52
40	162	162	0	100.00%	0.00%	6480	6307	173	39.44%	99.43%	59.98	36.46
41	162	162	0	100.00%	0.00%	6642	6469	173	40.45%	99.43%	58.97	37.39
42	162	162	0	100.00%	0.00%	6804	6631	173	41.47%	99.43%	57.96	38.33
43	162	162	0	100.00%	0.00%	6966	6793	173	42.48%	99.43%	56.95	39.27
44	162	162	0	100.00%	0.00%	7128	6955	173	43.49%	99.43%	55.93	40.20
45	162	162	0	100.00%	0.00%	7290	7117	173	44.51%	99.43%	54.92	41.14
46	162	162	0	100.00%	0.00%	7452	7279	173	45.52%	99.43%	53.91	42.08
47	162	162	0	100.00%	0.00%	7614	7441	173	46.53%	99.43%	52.89	43.01
48	162	162	0	100.00%	0.00%	7776	7603	173	47.55%	99.43%	51.88	43.95
49	162	162	0	100.00%	0.00%	7938	7765	173	48.56%	99.43%	50.87	44.88
50	162	162	0	100.00%	0.00%	8100	7927	173	49.57%	99.43%	49.85	45.82
51	162	162	0	100.00%	0.00%	8262	8089	173	50.58%	99.43%	48.84	46.76
52	162	162	0	100.00%	0.00%	8424	8251	173	51.60%	99.43%	47.83	47.69
53	162	162	0	100.00%	0.00%	8586	8413	173	52.61%	99.43%	46.81	48.63
54	162	162	0	100.00%	0.00%	8748	8575	173	53.62%	99.43%	45.80	49.57
55	162	162	0	100.00%	0.00%	8910	8737	173	54.64%	99.43%	44.79	50.50
56	162	162	0	100.00%	0.00%	9072	8899	173	55.65%	99.43%	43.78	51.44
57	162	162	0	100.00%	0.00%	9234	9061	173	56.66%	99.43%	42.76	52.38
58	162	162	0	100.00%	0.00%	9396	9223	173	57.68%	99.43%	41.75	53.31
59	162	162	0	100.00%	0.00%	9558	9385	173	58.69%	99.43%	40.74	54.25
60	162	161	1	99.38%	0.62%	9720	9546	174	59.70%	100.00%	40.30	54.86

Testing	# Records		# Goods		# Bads		Fraud Rate					
	16165		15991		174		0.010763996					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
61	162	162	0	100.00%	0.00%	9882	9708	174	60.71%	100.00%	39.29	55.79
62	162	162	0	100.00%	0.00%	10044	9870	174	61.72%	100.00%	38.28	56.72
63	162	162	0	100.00%	0.00%	10206	10032	174	62.74%	100.00%	37.26	57.66
64	162	162	0	100.00%	0.00%	10368	10194	174	63.75%	100.00%	36.25	58.59
65	162	162	0	100.00%	0.00%	10530	10356	174	64.76%	100.00%	35.24	59.52
66	162	162	0	100.00%	0.00%	10692	10518	174	65.77%	100.00%	34.23	60.45
67	162	162	0	100.00%	0.00%	10854	10680	174	66.79%	100.00%	33.21	61.38
68	162	162	0	100.00%	0.00%	11016	10842	174	67.80%	100.00%	32.20	62.31
69	162	162	0	100.00%	0.00%	11178	11004	174	68.81%	100.00%	31.19	63.24
70	162	162	0	100.00%	0.00%	11340	11166	174	69.83%	100.00%	30.17	64.17
71	162	162	0	100.00%	0.00%	11502	11328	174	70.84%	100.00%	29.16	65.10
72	162	162	0	100.00%	0.00%	11664	11490	174	71.85%	100.00%	28.15	66.03
73	162	162	0	100.00%	0.00%	11826	11652	174	72.87%	100.00%	27.13	66.97
74	162	162	0	100.00%	0.00%	11988	11814	174	73.88%	100.00%	26.12	67.90
75	162	162	0	100.00%	0.00%	12150	11976	174	74.89%	100.00%	25.11	68.83
76	162	162	0	100.00%	0.00%	12312	12138	174	75.91%	100.00%	24.09	69.76
77	162	162	0	100.00%	0.00%	12474	12300	174	76.92%	100.00%	23.08	70.69
78	162	162	0	100.00%	0.00%	12636	12462	174	77.93%	100.00%	22.07	71.62
79	162	162	0	100.00%	0.00%	12798	12624	174	78.94%	100.00%	21.06	72.55
80	162	162	0	100.00%	0.00%	12960	12786	174	79.96%	100.00%	20.04	73.48
81	162	162	0	100.00%	0.00%	13122	12948	174	80.97%	100.00%	19.03	74.41
82	162	162	0	100.00%	0.00%	13284	13110	174	81.98%	100.00%	18.02	75.34
83	162	162	0	100.00%	0.00%	13446	13272	174	83.00%	100.00%	17.00	76.28
84	162	162	0	100.00%	0.00%	13608	13434	174	84.01%	100.00%	15.99	77.21
85	162	162	0	100.00%	0.00%	13770	13596	174	85.02%	100.00%	14.98	78.14
86	162	162	0	100.00%	0.00%	13932	13758	174	86.04%	100.00%	13.96	79.07
87	162	162	0	100.00%	0.00%	14094	13920	174	87.05%	100.00%	12.95	80.00
88	162	162	0	100.00%	0.00%	14256	14082	174	88.06%	100.00%	11.94	80.93
89	162	162	0	100.00%	0.00%	14418	14244	174	89.08%	100.00%	10.92	81.86
90	162	162	0	100.00%	0.00%	14580	14406	174	90.09%	100.00%	9.91	82.79
91	162	162	0	100.00%	0.00%	14742	14568	174	91.10%	100.00%	8.90	83.72
92	162	162	0	100.00%	0.00%	14904	14730	174	92.11%	100.00%	7.89	84.66
93	162	162	0	100.00%	0.00%	15066	14892	174	93.13%	100.00%	6.87	85.59
94	162	162	0	100.00%	0.00%	15228	15054	174	94.14%	100.00%	5.86	86.52
95	162	162	0	100.00%	0.00%	15390	15216	174	95.15%	100.00%	4.85	87.45
96	162	162	0	100.00%	0.00%	15552	15378	174	96.17%	100.00%	3.83	88.38
97	162	162	0	100.00%	0.00%	15714	15540	174	97.18%	100.00%	2.82	89.31
98	162	162	0	100.00%	0.00%	15876	15702	174	98.19%	100.00%	1.81	90.24
99	162	162	0	100.00%	0.00%	16038	15864	174	99.21%	100.00%	0.79	91.17
100	127	127	0	100.00%	0.00%	16165	15991	174	100.00%	100.00%	-	91.90

Table 9.6 Out of Time Population Results

OOT	# Records		# Goods		# Bads		Fraud Rate				
	12236		12057		179		0.014628964				
	Bins Statistics					Cumulative Statistics					
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	123	65	58	52.85%	47.15%	123	65	58	0.54%	32.40%	31.86	1.12
2	123	96	27	78.05%	21.95%	246	161	85	1.34%	47.49%	46.15	1.89
3	123	110	13	89.43%	10.57%	369	271	98	2.25%	54.75%	52.50	2.77
4	123	118	5	95.93%	4.07%	492	389	103	3.23%	57.54%	54.32	3.78
5	123	119	4	96.75%	3.25%	615	508	107	4.21%	59.78%	55.56	4.75
6	123	118	5	95.93%	4.07%	738	626	112	5.19%	62.57%	57.38	5.59
7	123	122	1	99.19%	0.81%	861	748	113	6.20%	63.13%	56.92	6.62
8	123	122	1	99.19%	0.81%	984	870	114	7.22%	63.69%	56.47	7.63
9	123	120	3	97.56%	2.44%	1107	990	117	8.21%	65.36%	57.15	8.46
10	123	122	1	99.19%	0.81%	1230	1112	118	9.22%	65.92%	56.70	9.42
11	123	118	5	95.93%	4.07%	1353	1230	123	10.20%	68.72%	58.51	10.00
12	123	122	1	99.19%	0.81%	1476	1352	124	11.21%	69.27%	58.06	10.90
13	123	122	1	99.19%	0.81%	1599	1474	125	12.23%	69.83%	57.61	11.79
14	123	122	1	99.19%	0.81%	1722	1596	126	13.24%	70.39%	57.15	12.67
15	123	119	4	96.75%	3.25%	1845	1715	130	14.22%	72.63%	58.40	13.19
16	123	121	2	98.37%	1.63%	1968	1836	132	15.23%	73.74%	58.52	13.91
17	123	122	1	99.19%	0.81%	2091	1958	133	16.24%	74.30%	58.06	14.72
18	123	123	0	100.00%	0.00%	2214	2081	133	17.26%	74.30%	57.04	15.65
19	123	122	1	99.19%	0.81%	2337	2203	134	18.27%	74.86%	56.59	16.44
20	123	121	2	98.37%	1.63%	2460	2324	136	19.28%	75.98%	56.70	17.09
21	123	122	1	99.19%	0.81%	2583	2446	137	20.29%	76.54%	56.25	17.85
22	123	121	2	98.37%	1.63%	2706	2567	139	21.29%	77.65%	56.36	18.47
23	123	119	4	96.75%	3.25%	2829	2686	143	22.28%	79.89%	57.61	18.78
24	123	122	1	99.19%	0.81%	2952	2808	144	23.29%	80.45%	57.16	19.50
25	123	120	3	97.56%	2.44%	3075	2928	147	24.28%	82.12%	57.84	19.92
26	123	121	2	98.37%	1.63%	3198	3049	149	25.29%	83.24%	57.95	20.46
27	123	121	2	98.37%	1.63%	3321	3170	151	26.29%	84.36%	58.07	20.99
28	123	123	0	100.00%	0.00%	3444	3293	151	27.31%	84.36%	57.05	21.81
29	123	123	0	100.00%	0.00%	3567	3416	151	28.33%	84.36%	56.03	22.62
30	123	123	0	100.00%	0.00%	3690	3539	151	29.35%	84.36%	55.01	23.44
31	123	122	1	99.19%	0.81%	3813	3661	152	30.36%	84.92%	54.55	24.09
32	123	121	2	98.37%	1.63%	3936	3782	154	31.37%	86.03%	54.67	24.56
33	123	121	2	98.37%	1.63%	4059	3903	156	32.37%	87.15%	54.78	25.02
34	123	122	1	99.19%	0.81%	4182	4025	157	33.38%	87.71%	54.33	25.64
35	123	122	1	99.19%	0.81%	4305	4147	158	34.39%	88.27%	53.87	26.25
36	123	123	0	100.00%	0.00%	4428	4270	158	35.42%	88.27%	52.85	27.03
37	123	122	1	99.19%	0.81%	4551	4392	159	36.43%	88.83%	52.40	27.62
38	123	120	3	97.56%	2.44%	4674	4512	162	37.42%	90.50%	53.08	27.85
39	123	123	0	100.00%	0.00%	4797	4635	162	38.44%	90.50%	52.06	28.61
40	123	122	1	99.19%	0.81%	4920	4757	163	39.45%	91.06%	51.61	29.18
41	123	123	0	100.00%	0.00%	5043	4880	163	40.47%	91.06%	50.59	29.94
42	123	123	0	100.00%	0.00%	5166	5003	163	41.49%	91.06%	49.57	30.69
43	123	122	1	99.19%	0.81%	5289	5125	164	42.51%	91.62%	49.11	31.25
44	123	122	1	99.19%	0.81%	5412	5247	165	43.52%	92.18%	48.66	31.80
45	123	123	0	100.00%	0.00%	5535	5370	165	44.54%	92.18%	47.64	32.55
46	123	122	1	99.19%	0.81%	5658	5492	166	45.55%	92.74%	47.19	33.08
47	123	122	1	99.19%	0.81%	5781	5614	167	46.56%	93.30%	46.73	33.62
48	123	123	0	100.00%	0.00%	5904	5737	167	47.58%	93.30%	45.71	34.35
49	123	123	0	100.00%	0.00%	6027	5860	167	48.60%	93.30%	44.69	35.09
50	123	123	0	100.00%	0.00%	6150	5983	167	49.62%	93.30%	43.67	35.83
51	123	121	2	98.37%	1.63%	6273	6104	169	50.63%	94.41%	43.79	36.12
52	123	123	0	100.00%	0.00%	6396	6227	169	51.65%	94.41%	42.77	36.85
53	123	122	1	99.19%	0.81%	6519	6349	170	52.66%	94.97%	42.31	37.35
54	123	123	0	100.00%	0.00%	6642	6472	170	53.68%	94.97%	41.29	38.07
55	123	123	0	100.00%	0.00%	6765	6595	170	54.70%	94.97%	40.27	38.79
56	123	123	0	100.00%	0.00%	6888	6718	170	55.72%	94.97%	39.25	39.52
57	123	123	0	100.00%	0.00%	7011	6841	170	56.74%	94.97%	38.23	40.24
58	123	122	1	99.19%	0.81%	7134	6963	171	57.75%	95.53%	37.78	40.72
59	123	123	0	100.00%	0.00%	7257	7086	171	58.77%	95.53%	36.76	41.44
60	123	121	2	98.37%	1.63%	7380	7207	173	59.77%	96.65%	36.87	41.66

OOT	# Records	# Goods	# Bads	Fraud Rate
	12236	12057	179	0.014628964

	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
61	123	122	1	99.19%	0.81%	7503	7329	174	60.79%	97.21%	36.42	42.12
62	123	123	0	100.00%	0.00%	7626	7452	174	61.81%	97.21%	35.40	42.83
63	123	123	0	100.00%	0.00%	7749	7575	174	62.83%	97.21%	34.38	43.53
64	123	123	0	100.00%	0.00%	7872	7698	174	63.85%	97.21%	33.36	44.24
65	123	123	0	100.00%	0.00%	7995	7821	174	64.87%	97.21%	32.34	44.95
66	123	123	0	100.00%	0.00%	8118	7944	174	65.89%	97.21%	31.32	45.66
67	123	123	0	100.00%	0.00%	8241	8067	174	66.91%	97.21%	30.30	46.36
68	123	123	0	100.00%	0.00%	8364	8190	174	67.93%	97.21%	29.28	47.07
69	123	123	0	100.00%	0.00%	8487	8313	174	68.95%	97.21%	28.26	47.78
70	123	123	0	100.00%	0.00%	8610	8436	174	69.97%	97.21%	27.24	48.48
71	123	122	1	99.19%	0.81%	8733	8558	175	70.98%	97.77%	26.79	48.90
72	123	123	0	100.00%	0.00%	8856	8681	175	72.00%	97.77%	25.77	49.61
73	123	122	1	99.19%	0.81%	8979	8803	176	73.01%	98.32%	25.31	50.02
74	123	123	0	100.00%	0.00%	9102	8926	176	74.03%	98.32%	24.29	50.72
75	123	123	0	100.00%	0.00%	9225	9049	176	75.05%	98.32%	23.27	51.41
76	123	123	0	100.00%	0.00%	9348	9172	176	76.07%	98.32%	22.25	52.11
77	123	123	0	100.00%	0.00%	9471	9295	176	77.09%	98.32%	21.23	52.81
78	123	123	0	100.00%	0.00%	9594	9418	176	78.11%	98.32%	20.21	53.51
79	123	123	0	100.00%	0.00%	9717	9541	176	79.13%	98.32%	19.19	54.21
80	123	123	0	100.00%	0.00%	9840	9664	176	80.15%	98.32%	18.17	54.91
81	123	123	0	100.00%	0.00%	9963	9787	176	81.17%	98.32%	17.15	55.61
82	123	123	0	100.00%	0.00%	10086	9910	176	82.19%	98.32%	16.13	56.31
83	123	123	0	100.00%	0.00%	10209	10033	176	83.21%	98.32%	15.11	57.01
84	123	123	0	100.00%	0.00%	10332	10156	176	84.23%	98.32%	14.09	57.70
85	123	122	1	99.19%	0.81%	10455	10278	177	85.25%	98.88%	13.64	58.07
86	123	122	1	99.19%	0.81%	10578	10400	178	86.26%	99.44%	13.18	58.43
87	123	122	1	99.19%	0.81%	10701	10522	179	87.27%	100.00%	12.73	58.78
88	123	123	0	100.00%	0.00%	10824	10645	179	88.29%	100.00%	11.71	59.47
89	123	123	0	100.00%	0.00%	10947	10768	179	89.31%	100.00%	10.69	60.16
90	123	123	0	100.00%	0.00%	11070	10891	179	90.33%	100.00%	9.67	60.84
91	123	123	0	100.00%	0.00%	11193	11014	179	91.35%	100.00%	8.65	61.53
92	123	123	0	100.00%	0.00%	11316	11137	179	92.37%	100.00%	7.63	62.22
93	123	123	0	100.00%	0.00%	11439	11260	179	93.39%	100.00%	6.61	62.91
94	123	123	0	100.00%	0.00%	11562	11383	179	94.41%	100.00%	5.59	63.59
95	123	123	0	100.00%	0.00%	11685	11506	179	95.43%	100.00%	4.57	64.28
96	123	123	0	100.00%	0.00%	11808	11629	179	96.45%	100.00%	3.55	64.97
97	123	123	0	100.00%	0.00%	11931	11752	179	97.47%	100.00%	2.53	65.65
98	123	123	0	100.00%	0.00%	12054	11875	179	98.49%	100.00%	1.51	66.34
99	123	123	0	100.00%	0.00%	12177	11998	179	99.51%	100.00%	0.49	67.03
100	59	59	0	100.00%	0.00%	12236	12057	179	100.00%	100.00%	-	67.36