

References

- [1] E. D. Madyatmadja and M. Aryuni, "Comparative Study of Data Mining Model for Credit Card Application Scoring in Bank," *J. Theor. Appl. Inf. Technol.*, vol. 59, no. 2, 2014.
- [2] M. Aryuni and E. D. Madyatmadja, "Feature Selection in Credit Scoring Model for Credit Card Applicants in XYZ Bank: A Comparative Study," *Int. J. Multimed. Ubiquitous Eng.*, vol. 10, no. 5, pp. 17–24, May 2015.
- [3] M. Nithya, S. Suba, and B. Vaishnavi, "Predict the usage of laptops among students in rural areas using weka tool," *Int. J. Adv. Technol. Eng. Sci.*, vol. Vol. No.3, no. 01 September 2015, p. 6.
- [4] N. Padhy, "Data Mining: A prediction Technique for the workers in the PR Department of Orissa (Block and Panchayat)," *Int. J. Comput. Sci. Eng. Inf. Technol.*, vol. 2, no. 5, pp. 197–36, Oct. 2012.
- [5] D. Kumar Thiwary, "A Comparative Study Of Classification Algorithms For Credit Card Approval Using WEKA," *GALAXY Int. Interdiscip. Res. J.*, vol. Vol.2, Mar. 2014.
- [6] Plc.lk, 'People's Leasing Company PLC - Annual Report 2013 - 2014', 2015. [Online]. Available: http://www.plc.lk/inpages/annual_report_2013_14/risk-management.html. [Accessed: 03- Jan- 2015].
- [7] Annual Report 2013/14, 1st ed. colombo: Swarnamahala Financial Services PLC, 2015, pp. 23,24.
- [8] Annual Report 2012/13, 1st ed. colombo: Swarnamahala Financial Services PLC, 2014, p. 25

- [9]S. Faizan Ahmed and Q. Ali Malik, 574 *International Journal of Economics and Financial Issues* | Vol 5 • Issue 2 • 2015 *Credit Risk Management and Loan Performance: Empirical Investigation of Micro Finance Banks of Pakistan*, 1st ed. International Journal of Economics and Financial Issues, 2015.
- [10]D. Foust and A. Pressman, 'Not-So-Magic Numbers', *Business Week*, 2008.
- [11]*Foundations of Banking Risk: An Overview of Banking, Banking Risks, and Risk-Based Banking Regulation.*, 1st ed. New Jersey: John Wiley& Sons, 2009.
- [12]T. Gestel and B. Baesens, *Credit Risk Management: Basic Concepts: Financial Risk Components, rating analysis,models,economic and regulary capital*, 1st ed. Oxford University Press, 2008.
- [13]G. Yingjian and W. Chong, *Research on Credit Risk Assessment in Commercial Bank Based on Information Integration*, 1st ed. .



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Appendix A: Categorization

Data sets have been categorized on to following order

- Dependent Variable: Status - Ceased, Closed

- Other Variables:

Factor 1 - Profession - Businessmen, Doctor, Engineer, Executive, Farmer, Lawyer, Lecturer, Manager, Other, Self Employee, Small Businessmen, Teacher

Factor 2 - Brand Name - Bajaj, Chevrolet, Isuzu, Mahindra & Mahindra, Maruti Suzuki, Mazda, Mitsubishi Motors, Nissan, Suzuki, Tafe, Tata Motors, Toyota

Factor 3 - Model Name - 800, 45D, ACE, ALTO, AR4S-UG, ATLAS, AUTO 4S, AUTO AR4S, AUTO RE 2 STROKE, BAJAJ 4 S, BAJAJ AUTO, BOLERO MAXI TRUCK, CONDOR, COROLLA, DBA-NZE141, DYNA, ELF 350, HIACE, KF-GM70-HALF-BODY, KG-LH172, KG-VWE25, LA-HA238, LP7155, PAJERO JEEP, RE 205, TITAN, TOWNACE, UA-HR528, VANATTE

Factor 4 - Manufacture Year - 1993 to 2012

Factor 5 - Vehicle Class - Dual purpose Motor vehicle, Farm vehicle, Heavy Motor Lorry, Light Motor Lorry, Motor Car, Motor Coach, Motor Lorry UP 1700kg, Motor Tricycle, Motorcycles UP 100CC

Factor 6 - Fuel Type - Diesel, Petrol

Factor 7 - Actual Value - <5,00,000(P), 5,00,000-9,99,999(Q), 10,00,000-14,99,999(R), 15,00,000-19,99,999(S), 20,00,000-24,99,999(T), 25,00,000-29,99,999(U), 30,00,000-34,99,999(V), >35,00,000(W)

Factor 8 - Lease Percentage - 60% to 100%

Factor 9 - Req. Amount - <5,00,000(p), 5,00,000-9,99,999(q), 10,00,000-14,99,999(r), 15,00,000-19,99,999(s), 20,00,000-24,99,999(t), 25,00,000-29,99,999(u), 30,00,000-34,99,999(v), >35,00,000(w)

Factor 10 - Monthly Income - <50,000(I), 50,000-99,999(II), 1,00,000-1,49,999(III),
1,50,000-1,99,999(IV), >2,00,000(V)

Factor 11 - Interest - 9 to 15

Factor 12 - Number of rentals - 12 to 60

Factor 13 - Installment - <20,000(i), 20,000-39,999(ii), 40,000-59,999(iii), 60,000-
79,999(iv), 80,000-99,999(v), >1,00,000(vi)



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Appendix B: Variable Selection Procedure (Technology: Chi – Squared)

Status * Gender

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.691 ^a	1	.101		
Continuity Correction ^b	2.572	1	.109		
Likelihood Ratio	2.700	1	.100		
Fisher's Exact Test				.107	.054
N of Valid Cases	6000				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 393.38.

b. Computed only for a 2x2 table

Status * Age

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.878 ^a	6	.693
Likelihood Ratio	3.877	6	.693
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.91.

Status * Profession

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	679.354 ^a	11	.000
Likelihood Ratio	955.310	11	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 29.82.

Status * District

Chi-Square Tests



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4.115 ^a	7	.766
Likelihood Ratio	4.109	7	.767
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 44.06.

Status * BrandName

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	157.472 ^a	11	.000
Likelihood Ratio	167.321	11	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.35.

Status * ModelName

Chi-Square Tests



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	392.574 ^a	28	.000
Likelihood Ratio	457.831	28	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 12.02.

Status * Manuf.Year

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	41.545 ^a	19	.002
Likelihood Ratio	41.803	19	.002
N of Valid Cases	6000		

a. 1 cells (2.5%) have expected count less than 5. The minimum expected count is 4.90.

Status * Vehicle Class

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	252.964 ^a	8	.000
Likelihood Ratio	269.882	8	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.35.

Status * Fuel Type

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	37.032 ^a	1	.000		
Continuity Correction ^b	36.705	1	.000		
Likelihood Ratio	36.967	1	.000		
Fisher's Exact Test				.000	.000
N of Valid Cases	6000				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 987.90.

b. Computed only for a 2x2 table



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Status * ActualValue

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	311.275 ^a	7	.000
Likelihood Ratio	367.826	7	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 99.68.

Status * Lease Percentage

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	27.255 ^a	4	.000
Likelihood Ratio	27.262	4	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 494.40.

Status * Req.Amount

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	197.350 ^a	7	.000
Likelihood Ratio	208.573	7	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 86.33.



Status * Monthly Income

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.921 ^a	1	.048		
Continuity Correction ^b	3.571	1	.059		
Likelihood Ratio	3.892	1	.049		
Fisher's Exact Test				.057	.030
N of Valid Cases	6000				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 56.07.

b. Computed only for a 2x2 table



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Status * Interest

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	134.384 ^a	2	.000
Likelihood Ratio	138.243	2	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 561.59.

Status * No.of Rentals

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	88.824 ^a	4	.000
Likelihood Ratio	90.264	4	.000
N of Valid Cases	6000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 425.42.

Status * Installment

Chi-Square Tests



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	85.487 ^a	4	.000
Likelihood Ratio	87.334	4	.000
N of Valid Cases	6000		

Appendix C: Best Model Selection

Classifier
Choose **LMT -I -1 -M 15 -W 0.0**

Test options
 Use training set
 Supplied test set (Set...)
 Cross-validation (Folds: 10)
 Percentage split (%: 66)
 More options...

Classifier output
 Legend: -ve = Cease, +ve = Close
 Tree size (total number of nodes): 25
 Leaves (number of predictor nodes): 17
 Time taken to build model: 0.22 seconds
 === Evaluation on training set ===
 === Summary ===
 Correctly Classified Instances 3502 58.3667 %
 Incorrectly Classified Instances 2498 41.6333 %
 Kappa statistic 0.1548
 Mean absolute error 0.4395
 Root mean squared error 0.4678
 Relative absolute error 88.9747 %
 Root relative squared error 94.1275 %
 Total Number of Instances 6000
 === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.518	0.364	0.533	0.518	0.526	0.633	Cease
	0.636	0.482	0.622	0.636	0.629	0.633	close
Weighted Avg.	0.584	0.429	0.583	0.584	0.583	0.633	

 === Confusion Matrix ===

a \ b	classified as	
1212	2118	1
		b = close

 Result list (right-click for options):
 20:57:11 - trees.J48
20:58:49 - trees.ADTree
 20:58:55 - trees.BFTree
 20:59:39 - trees.FT
 21:00:45 - trees.J48graft
 21:02:45 - trees.LADTree
 21:03:24 - trees.SimpleCart
 21:05:46 - trees.RandomForest
 21:06:17 - trees.RandomTree
 21:12:43 - trees.LMT

Status: OK

ADTree

Correctly Classified Instances	Time Taken to Build the Model
58.3%	0.22 Seconds

Classifier

Choose LMT -I -1 -M 15 -W 0.0

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) Status

Start Stop

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Classifier output

Size of the tree: 7

Number of Leaf Nodes: 4

Time taken to build model: 31.04 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	3555	59.25 %
Incorrectly Classified Instances	2445	40.75 %
Kappa statistic	0.2096	
Mean absolute error	0.4362	
Root mean squared error	0.467	
Relative absolute error	88.3177 %	
Root relative squared error	93.9777 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.763	0.544	0.529	0.763	0.625	0.637	Cease
	0.456	0.237	0.706	0.456	0.554	0.637	close
Weighted Avg.	0.593	0.374	0.627	0.593	0.586	0.637	

=== Confusion Matrix ===

a b <-- classified as

2037 633 | a = Cease

1812 1518 | b = close

Status OK



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

BFTree

Correctly Classified Instances	Time Taken to Build the Model
59.25%	31.04 Seconds

Classifier

Choose LMT -I-1 -M 15 -W 0.0

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) Status

Start Stop

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Classifier output

```
[Installation=i1] * -0.34 +
[Installation=v] * 0.5
```

Time taken to build model: 12.61 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	4539	75.65 %
Incorrectly Classified Instances	1461	24.35 %
Kappa statistic	0.5082	
Mean absolute error	0.3044	
Root mean squared error	0.419	
Relative absolute error	61.6239 %	
Root relative squared error	84.3042 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.738	0.229	0.721	0.738	0.73	0.821	Cease
	0.771	0.262	0.786	0.771	0.779	0.821	close
Weighted Avg.	0.757	0.247	0.757	0.757	0.757	0.821	

=== Confusion Matrix ===

```

a   b  <-- classified as
1971 699 | a = Cease
762 2568 | b = close

```

Status OK



University of Moratuwa, Sri Lanka.
 Electronic Theses & Dissertations
www.lib.mrt.ac.lk

FT Tree

Correctly Classified Instances	Time Taken to Build the Model
75.65 %	12.61 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0.0**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Status

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Classifier output

Size of the tree : 4826

Time taken to build model: 86.52 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	4706	78.4333 %
Incorrectly Classified Instances	1294	21.5667 %
Kappa statistic	0.5682	
Mean absolute error	0.2798	
Root mean squared error	0.374	
Relative absolute error	56.6426 %	
Root relative squared error	75.2614 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.813	0.239	0.732	0.813	0.77	0.879	Cease
	0.761	0.187	0.836	0.761	0.797	0.879	close
Weighted Avg.	0.784	0.21	0.789	0.784	0.785	0.879	

=== Confusion Matrix ===

```

a   b  <-- classified as
2172 498 | a = Cease
796 2534 | b = close

```

Status OK



University of Moratuwa, Sri Lanka.
 Electronic Theses & Dissertations
www.lib.mrt.ac.lk

J48

Correctly Classified Instances	Time Taken to Build the Model
78.4%	86.52 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0,0**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Status

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft**
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Status
OK

Classifier output

Size of the tree : 5684

Time taken to build model: 98.1 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	4706	78.4333 %
Incorrectly Classified Instances	1294	21.5667 %
Kappa statistic	0.5682	
Mean absolute error	0.2798	
Root mean squared error	0.374	
Relative absolute error	56.6426 %	
Root relative squared error	75.2614 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.813	0.239	0.732	0.813	0.77	0.879	Cease
	0.761	0.187	0.836	0.761	0.797	0.879	close
Weighted Avg.	0.784	0.21	0.789	0.784	0.785	0.879	

=== Confusion Matrix ===

a	b	<-- classified as
2172	498	a = Cease
796	2534	b = close



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

J48graft

Correctly Classified Instances	Time Taken to Build the Model
78.4%	98.1 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0.0**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Status

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree**
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Classifier output

#expanded nodes: 60

#Processed examples: 199524

#Ratio e/n: 2494.05

Time taken to build model: 5.35 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	3486	58.1 %
Incorrectly Classified Instances	2514	41.9 %
Kappa statistic	0.1851	
Mean absolute error	0.4405	
Root mean squared error	0.468	
Relative absolute error	89.1866 %	
Root relative squared error	94.1653 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===


	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.736	0.543	0.521	0.736	0.61	0.625	Cease
	0.457	0.264	0.683	0.457	0.548	0.625	close
Weighted Avg.	0.581	0.388	0.611	0.581	0.575	0.625	

=== Confusion Matrix ===

a b <-- classified as

1965	705	a = Cease
1809	1521	b = close

Status
OK



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

LAD Tree

Correctly Classified Instances	Time Taken to Build the Model
58.1%	5.35 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0.0**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Status

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT**

Classifier output

[Actual value=v] * 0.63 +
[Monthly Income=II] * 0.09

Time taken to build model: 89.46 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	3516	58.6 %
Incorrectly Classified Instances	2484	41.4 %
Kappa statistic	0.1925	
Mean absolute error	0.4391	
Root mean squared error	0.4684	
Relative absolute error	88.8885 %	
Root relative squared error	94.2469 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.726	0.526	0.525	0.726	0.61	0.634	Cease
	0.474	0.274	0.683	0.474	0.559	0.634	close
Weighted Avg.	0.586	0.386	0.613	0.586	0.582	0.634	

=== Confusion Matrix ===

a	b	<-- classified as
1939	731	a = Cease
1753	1577	b = close

Status OK



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

LMT

Correctly Classified Instances	Time Taken to Build the Model
58.6%	89.46 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0.0**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Status

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree**
- 21:12:43 - trees.LMT

Classifier output

Monthly income = v : Cease (0/0)

Size of the tree : 10900

Time taken to build model: 0.07 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	5019	83.65 %
Incorrectly Classified Instances	981	16.35 %
Kappa statistic	0.6757	
Mean absolute error	0.1892	
Root mean squared error	0.3076	
Relative absolute error	38.304 %	
Root relative squared error	61.8904 %	
Total Number of Instances	6000	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.92	0.23	0.762	0.92	0.834	0.941	Cease
	0.77	0.08	0.923	0.77	0.839	0.941	close
Weighted Avg.	0.837	0.147	0.851	0.837	0.837	0.941	

=== Confusion Matrix ===

a	b	-- classified as	
2456	214	a = Cease	
767	2563	b = close	

Status OK



University of Moratuwa, Sri Lanka.
 Electronic Theses & Dissertations
www.lib.mrt.ac.lk

RandomTree

Correctly Classified Instances	Time Taken to Build the Model
83.65%	0.07 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0.0**

Test options

Use training set

Supplied test set Set...

Cross-validation Folds

Percentage split %

More options...

(Nom) Status

Start Stop

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart**
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Classifier output

```

number of leaf nodes: 7
Size of the Tree: 13
Time taken to build model: 12.02 seconds

=== Evaluation on training set ===
=== Summary ===
Correctly Classified Instances      3612      60.2 %
Incorrectly Classified Instances    2388      39.8 %
Kappa statistic                    0.2105
Mean absolute error                 0.4346
Root mean squared error             0.4661
Relative absolute error             87.976 %
Root relative squared error         93.7957 %
Total Number of Instances          6000

=== Detailed Accuracy By Class ===
                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                0.657   0.442   0.544     0.657   0.595     0.651   Cease
                0.558   0.343   0.67     0.558   0.609     0.651   close
Weighted Avg.   0.602   0.387   0.614     0.602   0.603     0.651

=== Confusion Matrix ===
      a    b  <-- classified as
1753  917 |  a = Cease
1471 1859 |  b = close

```

Status
OK



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

SimpleCart

Correctly Classified Instances	Time Taken to Build the Model
60.2 %	12.02 Seconds

Classifier

Choose **LMT -I -1 -M 15 -W 0.0**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Status

Result list (right-click for options)

- 20:57:11 - trees.J48
- 20:58:49 - trees.ADTTree
- 20:58:55 - trees.BFTree
- 20:59:39 - trees.FT
- 21:00:45 - trees.J48graft
- 21:02:45 - trees.LADTree
- 21:03:24 - trees.SimpleCart
- 21:05:46 - trees.RandomForest
- 21:06:17 - trees.RandomTree
- 21:12:43 - trees.LMT

Classifier output

```

out of bag error: 0.4525

Time taken to build model: 5.42 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances      5019      83.65 %
Incorrectly Classified Instances    981      16.35 %
Kappa statistic                    0.6696
Mean absolute error                 0.2543
Root mean squared error             0.327
Relative absolute error              51.476 %
Root relative squared error         65.7902 %
Total Number of Instances          6000

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                0.826   0.155   0.81       0.826   0.818     0.939   Cease
                0.845   0.174   0.858     0.845   0.852     0.939   close
Weighted Avg.   0.837   0.166   0.837     0.837   0.837     0.939

=== Confusion Matrix ===

  a  b  <-- classified as
2205 465 |  a = Cease
 516 2814 |  b = close

```

Status
OK



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Correctly Classified Instances	Time Taken to Build the Model
83.65 %	5.42 Seconds