The bump hunting and its application to the customer data

Bump hunting とその顧客データへの応用

Hideo Hirose
Department of Systems Design and Informatics,
Kyushu Institute of Technology,
Fukuoka 820-8502, Japan
Email: hirose@ces.kyutech.ac.jp

Abstract: In difficult classification problems of the z-dimensional points into two groups having 0-1 responses due to the messy data structure, it is more favorable to search for the denser regions for the response 1 assigned points than to find the boundaries to separate the two groups. To such problems often seen in customer databases, we have developed a bump hunting method using probabilistic and statistical methods. By specifying a pureness rate in advance, a maximum capture rate will be obtained. Then, a trade-off curve between the pureness rate and the capture rate can be constructed. In finding the maximum capture rate, we have used the decision tree method combined with the genetic algorithm. We first explain a brief introduction of our research: what the bump hunting is, the trade-off curve between the pureness rate and the capture rate, the bump hunting using the tree genetic algorithm, the upper bounds for the trade-off curve using the extreme-value statistics. Then, the assessment for the accuracy of the trade-off curve is tackled from the genetic algorithm procedure viewpoint. Using the new genetic algorithm procedure proposed, we can obtain the upper bound accuracy for the trade-off curve. Then, we may expect the actually attainable trade-off curve upper bound. The bootstrapped hold-out method is used in assessing the accuracy of the trade-off curve, as well as the cross validation method.

Keywords: decision tree, genetic algorithm, extreme-value statistics, trade-off curve, bootstrapped hold-out, new GA tree.

*This report is presented at the symposium '高度情報抽出のための統計理論・方法論とその応用，九州大学附属図書館視聴覚ホール，11/20-11/22，2008

1 Introduction

We review our study so far briefly in this section: these are, 1) what the bump hunting is, 2) the trade-off curve between the pureness rate and the capture rate, 3) the bump hunting using the decision tree, 4) the genetic algorithm adapted to the tree structure, 5) the upper bound for the trade-off curve by using the extreme-value statistics, 6) actual lower bound for the trade-off curve, and 7) their summary. Then following this section, we discuss the new method, the new GA tree, and its application to real customer data.

1.1 What is the bump hunting?

Suppose that we are interested in classifying n points in a z-dimensional feature variable space into two groups according to their responses, where each point is assigned response 1 or response 0 as its target variable. For example, if a customer makes a decision to act a certain way, then we assign response 1 to this customer, and assign response 0, otherwise. We want to know the customers' preferences presenting response 1. We assume that their personal features, such as gender, age, living district, education, family profile, etc., are already obtained.

Many classification problems have been dealt with elsewhere to rather simpler cases using the methods of the linear discrimination analysis, the nearest neighbor, logistic regression, decision tree, neural networks, support vector machine, boosting, bagging, and etc. (see [9], e.g.) as fundamental classification problems. In some

real data cases in customer classification, a difficulty has been raised; since many response 1 points and 0 points are closely located in the feature variable space, response 1 points are hardly separable from response 0 points [10, 11]; therefore, it is difficult to find the favorable customers. In such a case, to find the denser regions to the favorable customers could be an alternative. Such regions are called the bumps, and finding them is called the bump hunting; see Figure1. The bump hunting has been studied in the fields of statistics, data mining, and machine learning [1, 2, 7, 8].

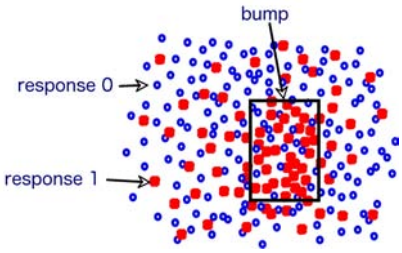

Figure 1: The bump hunting for the denser regions to response 1 points which are hardly separable from response 0 points.

1.2 Trade-off curve between the pureness rate and the capture rate

By specifying a pureness rate p_0 in advance, where the pureness rate p is the ratio of the number of points of assigned response 1 to the total number of points assigned responses 0 and 1 in the target region, a maximum capture rate c_m will be obtained, where the capture rate c is the ratio of the number of points assigned response 1 to the number of points assigned responses 0 and 1 in the total regions. Then a trade-off curve between the pre-specified pureness rate p_0 and the maximum capture rate c_m can be constructed; see Figure 2.

Now, we let TP be true positive, TN be true negative, FP be false positive, and FN be false negative. Since a response 1 point in or outside the bump regions is considered to be TP or FN, respectively, and a response 0 in or outside the bumps is FP or TN, the pureness rate p can be

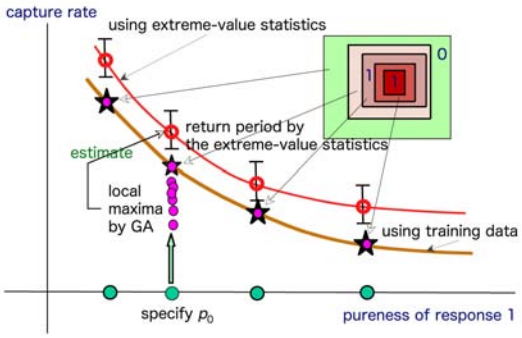

Figure 2: Three trade-off curves between the pureness rate and the capture rate: 1) using the training data, 2) using the extreme-value statistics.

defined by

$$p = \frac{\#\mathrm{TP}}{\#\mathrm{TP} + \#\mathrm{FP}}$$

in the bump regions; the capture rate c can also be defined by

$$c = \frac{\#\mathrm{TP}}{\#\mathrm{TP} + \#\mathrm{FN}}$$

in the total region [16], where "#" expresses the size of the samples; see Figure 3.

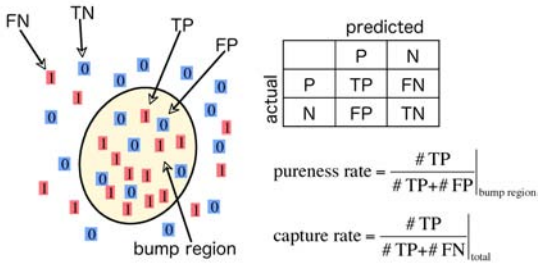

Figure 3: The pureness rate and the capture rate.

In a recall-precision curve, recall is defined by #TP/#(TP+FN) which is identical to the capture rate, and precision is defined by #TP/#(TP+FP) which is identical to the pureness rate; thus, a trade-off curve between the capture rate and the pureness rate seems to be equivalent to a recall-precision curve superficially [4, 6], e.g. However, we should note that these two are totally different from each other. As is seen in Figure 4, it can be considered that our trade-off curve is constructed by collecting the skyline points consisting of many trade-off

curves where each curve is corresponding to one classifier [14].

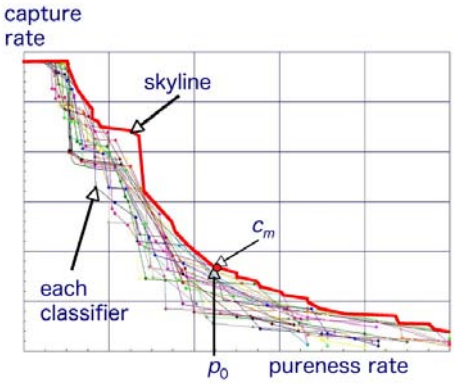

Figure 4: Trade-off curve as a skyline curve consisting of many classifiers.

1.3 Bump hunting using the decision tree

In order to make future actions easier, we adopt simpler boundary shapes such as the union of z-dimensional boxes located parallel to some explanation variable axes for the bumps as shown in Figures 1 and 2; that is, it would be convenient to adopt the binary decision tree. However, can the decision tree find the bumps?

The decision tree primarily tries to make some region classify into much purer subregions. Usually, the purer regions are much concerned with as the target point regions (the response 1 points), and the decision tree works in such a situation. However, if we are not interested in response 0 point regions where the decision tree intended to find the purer regions, we may discard such regions and expect much denser regions for response 1 to the rest of the regions. In a messy data case as shown in Figure 1, the decision tree also can do this; that is, it can find the boundaries for the bumps. Figure 5 shows how the decision tree finds the boundaries for the bumps; this is a typical example in one dimensional case, however the similar treatment can also be realized in higher dimensions, if the feature variables are not highly correlated to each other [13].

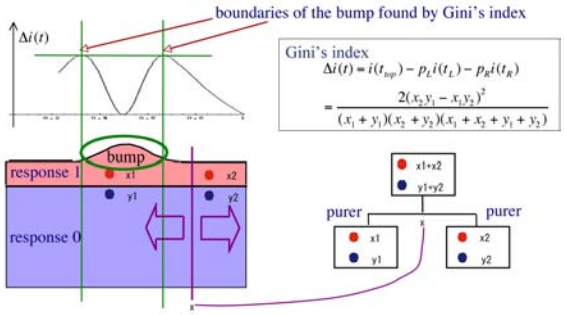

Figure 5: The decision tree finds the the boundary for the bumps.

1.4 The GA tree algorithm

In the decision tree, by selecting optimal explanation variables and splitting points to split the z-dimensional explanation variable subspaces into two regions from the top node to downward using the Gini's index as in the conventional method, we may obtain the number of response 1 points by collecting nodes where the pureness rates p are satisfying to be larger than the pre-specified pureness rate p_0. However, much response 1 points could be obtained if we locate appropriate explanation variables to each branching knot. This is because the conventional algorithm has a property of the local optimizer. Thus, we have developed a new decision tree method with the assistance of the random search methods such as the genetic algorithm (GA) specified to the tree structure, where the most adequate explanation variables are selected by using the GA, but the best splitting points are determined by using the Gini's index [20]; see Figure 6. We call this the GA tree.

The mutation can be done in the same manner to the standard genetic algorithms, however, the crossover should be different from those used in common because we are dealing with the tree structures. To preserve good inheritance in the tree structures, we have designed our crossover method as shown in Figure 7; we will know later that this causes many local maxima for the capture rates.

So far, we have been using the following evolution procedure in the GA tree:

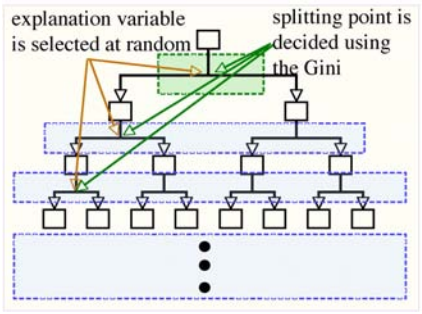

Figure 6: The bump hunting procedure using the decision tree with the genetic algorithm.

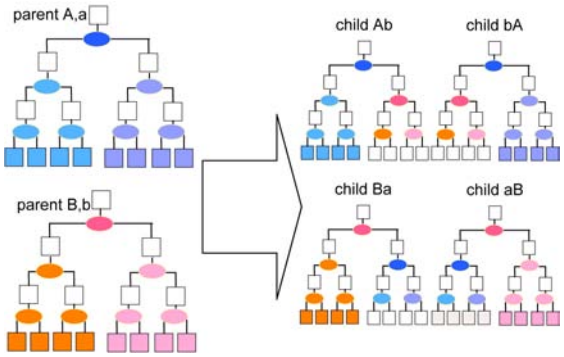

Figure 7: Crossover in the GA tree.

1) the number of initial seeds is set to 30; here, the initial seeds mean the trees where the explanation variables to be allocated to each branch are randomly selected,
2) obtain the capture rate to each seed tree, and select the top 20 best trees,
3) in the next generation, divide each tree to the left wing with or without the top node and the right wing with or without the top node, and combine the left wing and right wing trees of different parents to produce children trees (see Figure 7); this is a new crossover procedure in the GA tree; 40 children are then delivered, and select the top 20 best trees,
4) this evolution procedure is repeated by the 20th generation,
5) at the final stage, select the best one rule to apply the future data,
6) the mutation rate is set to around 5%,
7) select the best one tree in the final stage.

1.5 Upper bound for the trade-off curve by using the extreme-value statistics

The GA tree algorithm have a strong inclination of searching for the local maxima because we are intended to preserve a good inheritance in evolution procedure. Solutions obtained by the GA tree primarily are not the global optimal; this is a drawback of the algorithm. However, we have observed the existence of many local maxima with each starting point in the GA tree procedure. This turns out to become an advantage; the use of the extreme-value statistics [3] can then be used to estimate the return period (expected global maximum capture rate), and the method did work successfully when the shape of the marginal density function of an explanation variable is simple, such as monotonic or unimodal [13, 20]. This property is also observed in a real customer database [14].

Thus, we add a function of
8) estimating the upper bound capture rate by using the best 20 trees in each final stage of the evolution to our GA tree procedure; that is, we do procedures 1) - 7) for 20 cases, and estimate the upper bound using these 20 local maxima.

In Figure 1, how we have obtained the trade-off curve for the upper bound is shown. The procedures explained above is, however, applied only to the training data.

1.6 Actual lower bound for the trade-off curve

The solutions mentioned in 1.4 and 1.5 are, however, the best fitted solutions [9]; that is, the rules are constructed by using the training data and the evaluations are performed also by using the same training data [5, 19]; so, the solutions could be optimistic. If we apply the rules obtained by the training data to a new test data case having the same data structure, we may no longer expect the same performance in the new data case. We have been aware that we should pay much attention to this kind of problems even though the size of the explanation variable is small [17].

The bootstrapped hold-out method, the BHO, proposed in [15, 21] or the cross validation method has been used to estimate the biases between the results using the training data and those using the test data [14, 16]. The BHO produces the tree by using the training data of randomly selected half size of the sample data without replacement, and we evaluate the performance of the rule by using the rest of the half size of the samples. The typical cross validation of 10-fold cross validation requires the 10 performance evaluations. Thus, we also evaluate the 10 cases of the GA tree procedure in the BHO procedure [20]. In Figure 8, we show a typical cases for the cross validation results and the BHO results in a real data case where p_0 is pre-specified to 0.45 ; the real data case will be explained later. In the figure, we can see the biases between the results using the training data and those using the test data are not so different from each other between the cross validation method and the BHO method. We may use the BHO method when the sample size is small to some extent.

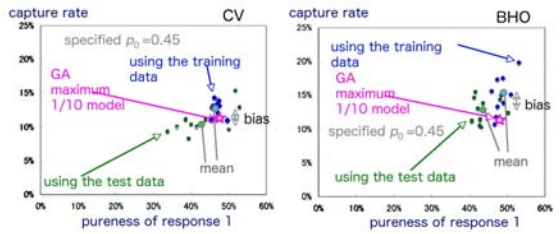

Figure 8: The cross validation results and the bootstrapped hold-out results using the trainig and test data in a real data case.

1.7 Summarizing

Summarizing the above, the trade-off curve we are dealing with have three aspects. The first is the curve obtained by using the training data; we can apply the if-then-rules to the future data only to this curve. The second is the return period curve obtained by using the extreme-value statistics; by estimating the upper bound for the capture rate, we can know where we are. The third is the trade-off curve obtained by using the test data; we can expect the actual capture rate for response 1. These three are indispensable like the Trinity to comprehend the whole figure of the trade-off curve between the pureness rate and the capture rate. In our previous study, however, we have not paid attention much to the genetic algorithm procedure itself; the results obtained from the previous study could be optimistic because they are using only the training data in the evolution procedure. In this paper, we tackle this point.

2 Customer Data

A real customer data case we are dealing with is taken from a corresponding course in Japan [14, 16]. The number of customers is very large, say 160,000; thus, we will not use all these data because of the high computing cost. Therefore, we will treat 15,870 samples, randomly selected from the original database, where the number of response 1 (customers we are interested in) is 2,863; thus the mean pureness rate becomes 18.0%. The number of features of the customers is more than 60, but we will use 41 variables; the variables are both continuous and discrete. We call this 1/10 model here. A much smaller case consisting of 1,635 samples was also investigated, where the number of response 1 is 290; the mean pureness rate is 17.7%. The number of variables is 44. We call this 1/100 model. Our primary objective here is how many response 1 samples can be captured if we require at least 40-50% pureness rate, p_0, from a practical viewpoint using these two smaller models.

3 Extreme-Value Statistics Approach

As mentioned before, the GA tree has a strong inclination of searching for the local maxima because we are using the tree structure in evolution procedure. This property turns out to be a merit to know the upper bound for the trade-off curve, although, in general, the genetic algorithm will not guarantee the global maximum. A set of samples collected from the local maxima could be samples for the extreme-value statistics for maxima. If the mother distribution function is a normal, exponential, log-normal, gamma, Gumbel, or Rayleigh type distribution, then the limit-

ing distribution of the maximum values from the mother distribution follows the Gumbel distribution (see [3] e.g.). Thus, we apply the Gumbel distribution to the local maxima samples. In the following, we investigated the two cases, 1/100 model and 1/10 model, to assess the reliability for the trade-off curve due to the limited number of samples.

3.1 Sample size effect

1/100 model real data case

For example in a data case where samples are drawn at random with 1/100 probability from a real customer data case, i.e., 1/100 model mentioned above, we have 20 local maxima of 48, 45, 48, 39, 56, 44, 32, 41, 56, 70, 40, 49, 42, 52, 38, 53, 47, 55, 34, 45, for the number of captures when we specify the pureness rate of 50%. If we fit the Gumbel distribution to the data, we can estimate the shape and scale parameters as 7.38 and 42.6. Then, the return capture rate (return period) for 500 trials is estimated to be 88.5. In Figure 9, we can see that the histogram of the sampled data well expresses the fitted Gumbel distribution density function. Using this result, we can guess the number of return captures in the real full data case, which could be 8,642 and this corresponds to 30.5% capture rate. However, the results by applying the test data to the rules obtained by the training data was very pessimistic. The bias between the training data trade-off curve and test data one shown in Figure 10 becomes very large because of large number of explanation variables, resulting that the rules obtained by the training data are not applicable to decide future action. In the figure, each point shows the mean value of the 10 cases BHO results.

1/10 model real data case

A much larger case of 1/10 model mentioned above is also investigated. The 20 local maxima are 207, 230, 251, 258, 255, 238, 170, 229, 204, 292, 247, 218, 281, 237, 230, 206, 195, 208, 193, 147, for the number of captures by the half training data, and the number of return capture is estimated as 425.5. Using this result, we can

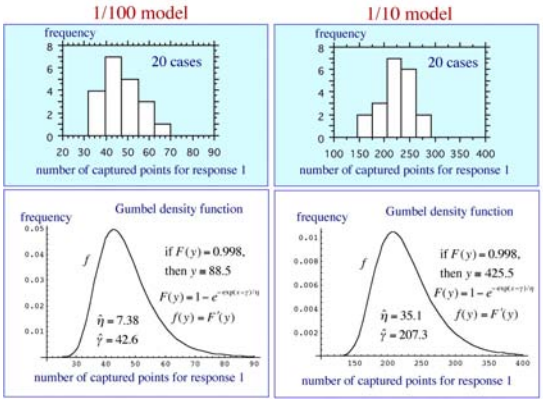

Figure 9: Gumbel distribution fit to the 20 local maxima using the GA tree.

guess the number of return captures in the real full data case, which could be 4,290 and this corresponds to 14.9% capture rate. The results in this case are considered to be reliable because of small biases (see Figures 9 and 10); however, the result obtained by using the 1/100 model is considerably optimistic. About two times of the response 1 points are obtained in the 1/100 model. We should, thus, pay much attention to the number of samples if the number of explanation variable is large comparing to the samples, which is well known.

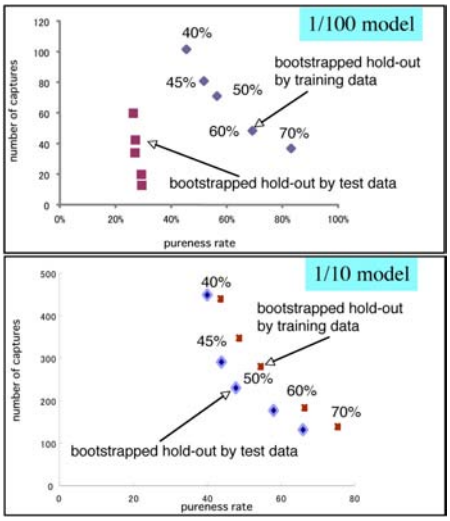

Figure 10: Trade-off curve assessment using the test data.

Table 1: Estimated captures by two different model sizes.

	1/100 model	1/10 model
computed return capture	88.5	425.5
return capture rate	30.5%	14.9%
full model return capture	8,642	4,290

3.2 Trade-off curve assessing using the bias between the training data results and the test data results

As explained before, we can construct the three kinds of trade-off curves. First, we simply obtain the optimal tree with a pre-specified pureness rate using all the sampled data as the training data. By connecting these obtained points, we can construct the trade-off curve. We have to use these rules (trees) actually. This curve is, however, optimistic.

Therefore, to know the actual capture rate, we have made the second trade-off curve. By dividing the sample data into two groups, training data and the test data, which is determined by the accuracy evaluation methods such as the cross validation and the BHO, a number of points of pureness rate and capture rate are obtained for the training data set and the test data set. By averaging these points, we can obtain the bias between the training data results and the test data results. Then, we can apply this bias to the very original trade-off curve which is obtained by using all the sampled data. This estimated trade-off curve indicates the actual capture rate when we specify the pureness rate. In that sense, this curve can be a lower bound for the trade-off curve.

All the points in the trade-off curve are obtained by using the direct results from the genetic algorithm procedure. The maximum value obtained for the capture rate cannot necessarily be the real global maximum. However, using the extreme-value statistics, we can estimate the return period (return capture rate) as the global maximum value. Therefore, we are able to know where we are, i.e., we can estimate the upper bound ceiling for the trade-off curve.

These three kinds of the trade-off curves are, in a sense, like the Trinity to comprehend the whole figure of the trade-off curve. However, we do not know yet the accuracy for the trade-off curve estimated by the extreme-value statistics.

3.3 Reliability for the return period due to the number of initial seeds

All the initial seeds have been set to 30 so far. Considering that the property of the local convergence of the GA procedure, it would be better to provide much larger number of seeds to verify if the extreme-value statistics works well. Figure 11 shows the results of the number of captures in a data case resampled from the real customer data case when the number of seeds and the successive number of iterations are set to larger values. Here, the pre-specified pureness rate is 50% and the model is 1/100 scale. We can see that the extreme-value statistics work very well, and we can use this method even if the number of samples to the Gumbel distribution is small such as 20; we can see that the predicted number of captures by the extreme-value statistics preserve almost a constant value even though the converged local maxima are gradually becoming larger as the number of seeds becomes larger.

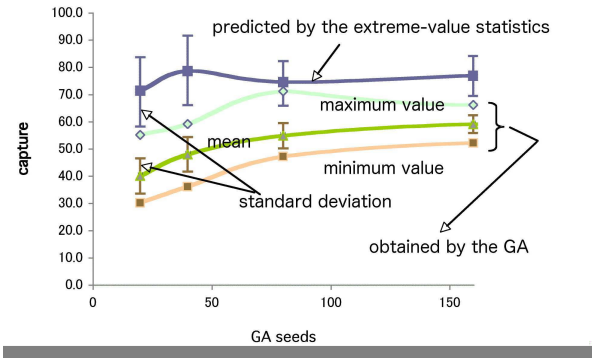

Figure 11: Number of captures of response 1 versus number of the seeds in the genetic algorithm.

4 Upper Bound Accuracy Assessment Using the New GA Tree

Since we know that the trade-off curve to actually applicable rules is optimistic if we use the training data only, the upper bound curve will

also be optimistic, i.e., be conservative. Here, we want to know a much more accurate upper bound for the trade-off curve. The accuracy for the trade-off curve to actually applicable rules is assessed by using the cross validation or the BHO method. In the GA procedures, however, the optimal tree is constructed by using only the training data; that is, 20 rules after 20 evolution generations to each procedure are obtained by using the training data only. Then, only the very last generation rule is assessed by using the test data. This is performed by the assessment method such as the cross validation or the BHO method. However, such a method cannot be applicable to assess the accuracy for the upper bound curve obtained by using the extreme-value statistics because the test data results do not necessarily follow the extrem-value distributions even though the training data results do. The upper bound capture rates using the extreme-value statistics would no longer be obtained because the results by using the test data would not necessarily have the property of local maximum. Even if we can estimate the optimistic upper bound for the trade-off curve, we could not obtain the accuracy of the curve if we insist on using the cross validation method or the BHO method as they are.

4.1 Applying the GA tree to the test data

We, thus, propose here a new genetic algorithm procedure adapted to the tree structure.
1) the number of initial seeds is set to 30,
2) obtain the capture rate to each seed, and select the top 20 best trees,
3) in the next generation, select the top 20 best trees from 40 children trees by the evaluation results *using the test data*
4) this evolution procedure is repeated by the 20th generation,
5) at the final stage, select the best one rule to apply the future data,
6) the mutation rate is set to around 5%,
7) select the best one tree in the final stage.
8) estimate the upper bound capture rate using the etreme-value statistics using the 20 local maxima.

So, each evolution generation stage, we produce the trees, and select the best trees using the test data. Then, we can expect that the final stage results could be local maxima for the test data, and we may apply the extreme-value statistics to these final results, as we applied the extreme-value statistics to optimal results obtained by using the training data. Figure 12 shows one of the new GA tree procedure results using the BHO method; the capture rates for the training data and the test data are not necessarily monotonically increasing because the optimality in the training data case is not succeeded to the next generation. We can see that the capture rates are stabilized within 10 generations, and this tendency is observed in simulation data and real data in common.

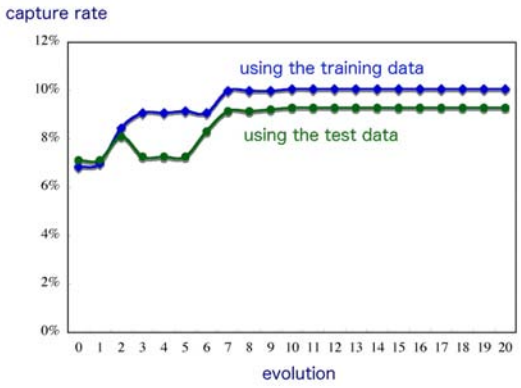

Figure 12: Capture rate convergence of response 1 as the evolution generation increases.

Using this new method, 20 local maxima could be obtained, and we may apply the extreme-value statistics to these results to estimate a much more accurate trade-off curve. In addition, we could assess the accuracy for this trade-off curve. An illustrative example is shown in Figure 13.

4.2 The new GA tree

Someone suspects that the accuracy evaluation by using the test data in evolution procedure in the GA tree would still be optimistic. The test data are always treated like the training data. To overcome this problem, we propose here a new method, classifying the original data into three subsets; the first subset is for the training data, the second is for the evaluation data,

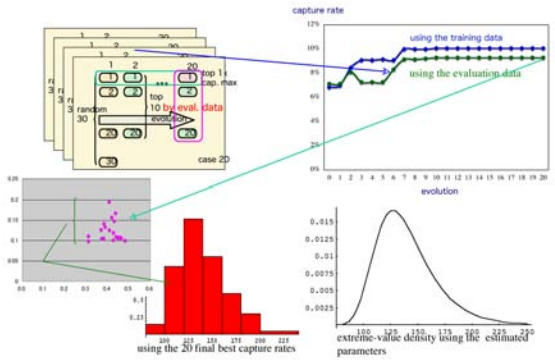

Figure 13: Gumbel distribution fit to the 20 local maxima using the test data.

and the third is for the test data. The GA procedure is performed using the training data and the evaluation data; in each evolution process, the selection is made by using the evaluation data, which enables us to use the extreme-value statistics with validity. At the final stage in the GA procedure, the test data subset is used to assess the accuracy for the upper bound trade-off curve. Figure 14 illustrates this diagram.

We applied this method to the real customer data case of 1/10 model mentioned earlier. We can see that the capture rate results by the evaluation data and those by the test data are highly correlated to each other; see Figure 15. The trade-off curve finally can be constructed as shown in Figure 16.

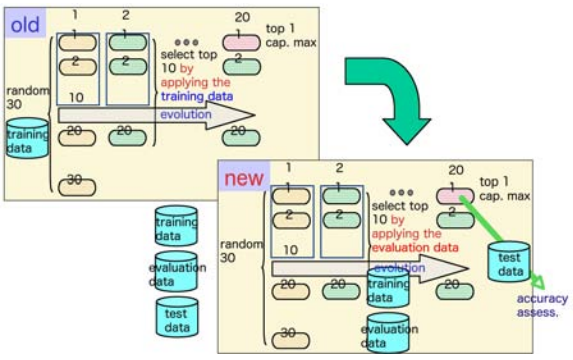

Figure 14: The new GA tree diagram.

5 Concluding Remarks

In difficult classification problems of the z-dimensional points into two groups having 0-1

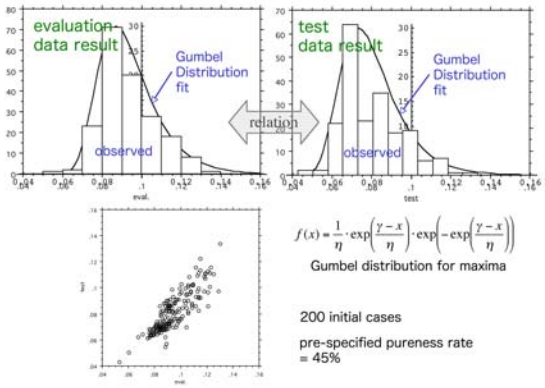

Figure 15: The relation of the capture rates between by the evaluation data and the test data.

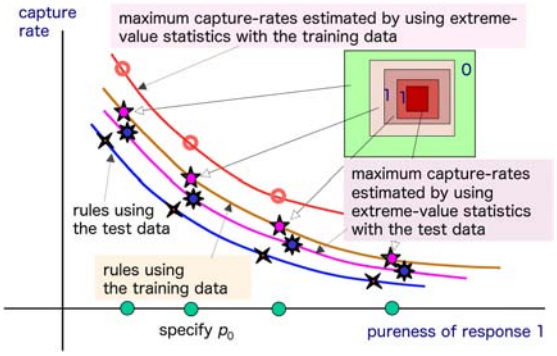

Figure 16: The trade-off curves for the upper bound curve with its accuracy.

responses due to the messy data structure, we have proposed to use the bump hunting method, and we have shown that there is a trade-off curve between the pureness rate and the capture rate, In finding the maximum capture rate, we have found that the decision tree method combined with the genetic algorithm is useful. In this paper, after explained a brief introduction of the research so far, that is, what the bump hunting is, the trade-off curve between the pureness rate and the capture rate, the bump hunting using the tree genetic algorithm, the upper bounds for the trade-off curve using the extreme-value statistics, then, the assessment method for the accuracy of the trade-off curve is described from the genetic algorithm procedure viewpoint. Using the new genetic algorithm procedure proposed, we can obtain the upper bound accuracy for the trade-off curve. Using this, we will make future de-

cisions by applying the rules obtained by using the training data with the knowledge of how far the rules we are using are located from the optimal points. To do this, we have proposed a new method, the new GA tree, along with the bootstrapped hold-out method in assessing the accuracy of the trade-off curve.

References

[1] Agarwal, D., Phillips, J.M., and Venkatasubramanian, S. :The hunting of the bump: On maximizing statistical discrepancy, SODA'06. (2006) 1137-1146

[2] Becker, U and Fahrmeir, L. :Bump hunting for risk: a new data mining tool and its applications, Computational Statistics, 16. (2001) 373–386

[3] Castillo, E.: Extreme Value Theory in Engineering. Academic Press. 1988

[4] Davis, J., and Goadrich, M.: The relationship between precision-recall and ROC Curves, Proceedings of the 23 International Conference on Machine Learning, 2006

[5] Efron, B.: Estimating the error rate of a prediction rule: improvements in cross-validation. JASA. 78 (1983) 316-331

[6] Fawcett, T. : An introduction to ROC analysis, Pattern Recognition Letters 27 (2006) 861-874.

[7] Friedman, J.H. and Fisher, N.I.: Bump hunting in high-dimensional data. Statistics and Computing. **9** (1999) 123–143.

[8] Gray, J.B. and Fan, G: Target: Tree analysis with randomly generated and evolved trees. Technical report. The University of Alabama (2003).

[9] Hastie, T., Tibshirani, R. and Friedman, J.H.: Elements of Statistical Learning. Springer (2001).

[10] Hirose, H.: A method to discriminate the minor groups from the major groups. Hawaii International Conference on Statistics, Mathematics, and Related Fields, (2005).

[11] Hirose, H.: Optimal boundary finding method for the bumpy regions. IFORS2005 Triennial Conference, (2005).

[12] Hirose, H.: The bump hunting by the decision tree with the genetic algorithm, Advances in Computational Algorithms and Data Analysis, pp.305-318, Springer (2008).

[13] Hirose, H., Yukizane,T. and Miyano, E.: Boundary detection for bumps using the Gini's index in messy classification problems, The 3rd International Conference on Cybernetics and Information Technologies, Systems and Applications, pp.293-298, (2006).

[14] Hirose, H., Yukizane,T. and Deguchi T.: The bump hunting method and its accuracy using the genetic algorithm with application to real customer data, pp.128-132, IEEE 7th International Conference on Computer and Information Technology, (2007).

[15] Hirose, H., Ohi, S. and Yukizane, T.: Assessment of the prediction accuracy in the bump hunting procedure. The 6th Annual Hawaii International Conference on Statistics, Mathematics, and Related Fields, (2007).

[16] Hirose, H. and Yukizane, T.: The bump hunting using the decision tree combined with the genetic algorithm: extreme-value statistics aspect, International Conference on Machine Learning and Data Analysis, (2007).

[17] Hirose, H. and Yukizane, T.: The accuracy of the trade-off curve in the bump hunting, The 7th Annual Hawaii International Conference on Statistics, Mathematics, and Related Fields, (2008).

[18] Hirose, H., Yukizane, T. and Zaman, F. : Accuracy assessment for the trade-off curve and its upper bound curve in the bump hunting using the new tree genetic algorithm, 7th World Congress in Probability and Statistics, (2008).

[19] Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI, 1995.

[20] Yukizane, T., Ohi, S., Miyano, E. and Hirose, H.: The bump hunting method using the genetic algorithm with the extreme-value statistics. IEICE Trans Inf. Syst., E89-D. (2006) 2332-2339.

[21] Yukizane T., Hirose, H, Ohi, S., and Miyano, E.: Accuracy of the solution in the bump hunting. IPSJ MPS SIG report, MPS06-62-04 (2006) 13-16.