2004－ICS－138 (7)
2004／12／4

Disease Definition Based on Spiral Discovery of Exceptions

MASATOSHI JUMI,† EINOSHIN SUZUKI,† MUNEAKI OHSHIMA,††
NING ZHONG,†† HIDETO YOKOI††† and KATSUHIKO TAKABAYASHI †††

In this paper, we show our endeavor for defining a disease by removing exceptional patients in a spiral manner. The removal is based on a likelihood-based criterion and can be supported by our previously developed data mining methods and medical experts. Two series of experiments with the chronic hepatitis data showed that our proposed method is effective and promising from various viewpoints.

1. Introduction

In medical treatment, a case is subject to multiple diseases and anomalies. For instance, in the chronic hepatitis data[2], several cases exhibit symptoms of acute hepatitis or nefroze. In diagnosis, a physician understands the status of a case by inferring and verifying various assumptions. On the other hand, in data mining, we try to obtain knowledge based on a hypothesis on the definition of the disease. Such a definition, which can be obtained by removing inappropriate cases, is useful in diagnosis, treatment, and medicine.

Special attention is required for removing cases with symptoms dissimilar to typical cases. Each physician has his/her own image of a disease and removing atypical cases might just correspond to confirming the image. In such a case, the obtained definition cannot deal with the cases outside the image. Selection of typical cases represents a highly intellectual activity in which various factors should be considered. As far as we know, deletion of such cases is left to the experts under the name of data preprocessing. Few attempts try to systematically support such activities and fewer reports are found in the literature.

2. Proposed Method

2.1 Overall Architecture

In order to circumvent the problem in the previous section, we propose a method based on a probabilistic approach. The approach can be supported by our peculiarity-oriented mining method[12], our time-series decision tree induction method[11], and our PrototypeLines visualization method[10]. These mainly consider peculiar data, time-series form, and related blood tests respectively and expected to contribute to detection of exceptional cases from different perspectives.

Our method classifies input cases into typical cases and exceptional cases. The set of typical cases is employed in defining a disease, which, in this paper, corresponds to liver cirrhosis (LC) in the chronic hepatitis data. Since two of us, Takabayashi and Yokoi, who are physicians, feel natural to use a probabilistic definition for a disease, the judgment is based on the naive Bayes method[4]. A naive Bayes classifier predicts a class $\hat{c}_{NBayes,i}$ of an example e_i assuming that each attribute a_j is independent. Here v_{ij} represents the value for a_j of e_i.

$$\hat{c}_{NBayes,i} = \mathrm{argmax}_c \Pr(c) \prod_{j=1}^{m} \Pr(a_j = v_{ij} \mid c) \quad (1)$$

Another reason for employing the naive Bayes method is that it enables a use of an estimated conditional probability $\hat{\Pr}(c|p)$ of a class c given a case p. By comparing $\hat{\Pr}(c|p)$ with his/her actual class c_p, we can detect candidates of cases with symptoms dissimilar to typical cases and obtain a more precise definition of the disease.

From above, the overall process is determined as follows. As data preprocessing, we can use the peculiarity-oriented mining method, the time-series decision tree induction method, and the PrototypeLines visualization method in order to detect candidates of exceptional cases. These candidates are shown to the medical experts, who select exceptional cases among

† Yokohama National University
†† Maebashi Institute of Technology
††† Chiba University Hospital

them. Then based on a criterion[6] defined in section 2.2, we detect candidates of exceptional cases, who are either all removed or removed by medical experts. The last process is iterated until the result of the naive Bayes method converges.

2.2 Detection Criteria of Exceptional Cases

Here we present our likelihood-based criterion[6]. In a classification problem in general, an example can be intuitively regarded as "typical" or "atypical". The typicalness $\Phi(p)$, which is based on the right-hand side of Eq. (1), of a case p represents a degree to which p belongs to its class c compared with the other class $\bar{c}$. Here v_{ij} represents the value for an attribute a_j of p.

$$\Phi(p) = \frac{\Pr(c)\prod_{j=1}^{m}\Pr(a_j = v_{ij} \mid c)}{\Pr(\bar{c})\prod_{j=1}^{m}\Pr(a_j = v_{ij} \mid \bar{c})} \quad (2)$$

The larger $\Phi(p)$ is, the more certain that p belongs to its class c.

If a naive Bayes classifier is relatively accurate for the class c, the typicalness $\Phi(p)$ tends to be large and vice versa. Thus the typicalness $\Phi(p)$ is relative since it depends on the preciseness of a naive Bayes classifier in terms of the class c. The degree how a naive Bayes classifier is precise for c can be measured by the degree of correct classification and the degree of incorrect classification. For c and $\bar{c}$, we represent the number of correctly-predicted examples by the naive Bayes method μ and ν respectively. Likewise, for c and $\bar{c}$, we represent the number of incorrectly-predicted examples by the naive Bayes method μ_e and ν_e respectively. The preciseness $\Psi(\hat{c})$ of a naive Bayes classifier of the estimated class $\hat{c}$ represents the ratio of the precision for c and the precision for $\bar{c}$, where we use Laplace correction in order to cope with the 0-occurrence problem[4].

$$\Psi(\hat{c}) = \frac{(\mu+1)(\nu+\nu_e+2)}{(\mu+\mu_e+2)(\nu_e+1)} \quad (3)$$

For our LC prediction problem, we define a degree of exception $E(p,\hat{c})$ for a case p and his/her estimated class $\hat{c}$ as follows in order to discriminate exceptional cases from typical cases. In the definition, $\lceil x \rceil$ represents x if x is an integer or $x+1$ otherwise.

$$E(p,\hat{c}) = \lceil -\log_{\Psi(\hat{c})}\Phi(p) \rceil \quad (4)$$

Intuitively, $E(p,\hat{c})$ represents an evaluation index which is equal to the number of upvaluated digits below the decimal point when we measure the typicalness $\Phi(p)$ of a case p in terms of the preciseness $\Psi(\hat{c})$ of a naive Bayes classifier of the estimated class $\hat{c}$. When prediction of the naive Bayes method is accurate, $\Psi(\hat{c})$ tends to be large, and the absolute value of $E(p,\hat{c})$ is relatively small even if the absolute value of $\Phi(p)$ is large. This fits our intuition that certain information rarely leads to an extreme degree of exception.

In each application of the naive Bayes method, cases whose degrees of exception are no less than a user-specified threshold are detected as exceptional. The loop continues until there are no exceptional cases among the detected examples. As the result, we obtain, for typical cases, conditional probabilities, which correspond to the definition of the disease.

3. Experimental Evaluation

3.1 Data Preparation and Initial Experiments

In the experiments, we used data from 180 days before the first biopsy to the day of the first biopsy following advice of medical experts. In the first series of our experiment, we used 46 LC cases and 55 non-LC cases each of whom has test values for all of 14 blood tests in Table 1. In the second series of our experiment, we used 140 LC cases and 160 non-LC cases and used 35 attributes in Table 2.

For the classifier, we first averaged each time sequence then discretized each value following advice of a domain expert. We use for each attribute value U: extremely high, V: very high, H: high, N: normal, L: low, v: very low, and u: extremely low. Each conditional probability is estimated using Laplace correction. A missing value is ignored both in estimating probabilities and in classifying an example.

We initially performed experiments by employing Yokoi as the medical expert and obtained preliminary results[6]. This time, another expert Takabayashi joined Yokoi in the first series of experiments.

3.2 Results with the Data Preprocessing and Domain Experts

In the data preprocessing, we first applied our peculiarity-oriented mining method to the whole data. 13 cases with no less than 4 pe-

culiar attributes were shown to the medical experts, and they removed 4 cases. Second, the medical expert investigated display result of PrototypeLines from 500 days before the first biopsy to 500 days after the first biopsy, and removed four cases among six cases. Third, we applied our peculiarity-oriented mining method to non-LC cases, and the experts removed five cases. Fourth, the medical experts inspected misclassified cases from the time-series decision tree. As the result, two cases were removed as exceptions. Fifth, the medical experts investigated six cases detected by our peculiarity-oriented mining method applied to LC cases and removed two cases.

In the spiral detection, we used 3 as the value of the threshold. In the first three spirals, the experts removed 4, 2, 1 cases out of detected 5, 3, 2 cases respectively. Since only one case who had not been removed was detected in the fourth spiral, the procedure terminated at this point.

We show the final conditional probabilities of the naive Bayes classifier, which corresponds to the disease definition, in Table 1. In the Table, for each blood test a and a category v, "$\hat{\Pr}(a{=}v|\text{non-LC})$ $(n(a{=}v|\text{non-LC}))$ | $\hat{\Pr}(a{=}v|\text{LC})$ $(n(a{=}v|\text{LC}))$" are shown, where $n(\cdot)$ represents the corresponding number of examples in the data set. For instance, there are 6 non-LC cases and 2 LC cases for ZTT=N. The probabilities are obtained using Laplace correction since there are 49 non-LC cases and 33 LC cases in the data set which corresponds to the Table. In the Table, we emphasize categories each of which shows more than 3 times of difference and no smaller than 10 % with boldface and with underline for non-LC predominant and LC predominant respectively. Since each blood test is assumed to represent an ordinal scale, we marked higher/lower categories of the categories appropriately. A marked category will be called a discriminative condition in the rest of this paper.

3.3 Analysis of the First Series of Experimental Results

In data mining, discovered knowledge is typically more important than high accuracy. Previous experiments have revealed that detection of exceptional cases could be done by searching for asynchronism in blood tests ALB, CHE, T-CHO, WBC, PLT; and HGB might be ignored. This piece of knowledge was elaborated in the experiments and the experts first check whether at least two of CHE, ALB, PLT are low then ascertain their decisions with results on T-CHO and WBC. According to them, PLT is the most important blood test and the value 15 K is important as a threshold value. They consider that ALB is also important and T-CHO is relatively unreliable since it tends to be influenced by meals. These facts realized them that traditional rules of thumb are highly valuable.

They felt as if they were educated by the cases shown by our data mining method and even called the process "expert learning/training" instead of machine learning. The process gave them novel hypotheses on the LC prediction problem, but evaluating validness of the hypotheses requires systematic experiments followed by a traditional statistical approach. Anyway they were pleased to sharpen their capability for this problem and have the new hypotheses.

After the experiments, the medical experts inspected the obtained set of typical cases and found six cases who were recognized as exceptional cases. Five of them are due to effect of relatively unimportant blood tests: similar stage (i.e. F3) to LC; and effect of inadequate thresholds for ALB and T-CHO. The last case shows limitation of using average values: s/he had known a period with high PLT thus was recognized as non-LC. We believe that these mistakes are due to data handling, which is not directly related with our proposed method. Anyway, seeing the number of exceptional cases that our method detected, these results are considered to confirm effectiveness of our approach.

The separation into typical and exceptional cases can be also validated by measuring predictive accuracy with cross validation. For the prediction problem, we have considered a two-step process, which we call a separate prediction model[6]. Given a novel case, the process judges him/her as an exception if a similar case exists using a 1-nearest neighbor (1-NN) method for time-series classification[11]. Otherwise, the naive Bayes method is applied to predict his/her class.

For the experimental results in the previous Section, the accuracy of the separate prediction

Table 1 Conditional probabilities (%) and numbers of examples for typical cases. where each category shows "$\hat{\Pr}(a{=}v|\text{non-LC})$ $(n(a{=}v|\text{non-LC}))$ | $\hat{\Pr}(a{=}v|\text{LC})$ $(n(a{=}v|\text{LC}))$"

GOT	**N: 35.8(18) | 2.7(0)**	H: 39.6(20) | 37.8(13)	V: 20.8(10) | 54.1(19)	U: 3.8(1) | 5.4(1)
GPT	N: 15.1(7) | 5.4(1)	H: 50.9(26) | 21.6(7)	V: 22.6(11) | 62.2(22)	U: 11.3(5) | 10.8(3)
TTT	N: 39.6(20) | 24.3(8)	H: 28.3(14) | 45.9(16)	V: 28.3(14) | 10.8(3)	U: 3.8(1) | 18.9(6)
ZTT	N: 13.2(6) | 8.1(2)	H: 69.8(36) | 70.3(25)	V: 13.2(6) | 18.9(6)	U: 3.8(1) | 2.7(0)
D-BIL	N: 88.7(46) | 40.5(14)	H: 7.5(3) | 43.2(15)	V: 1.9(0) | 10.8(3)	U: 1.9(0) | 5.4(1)
I-BIL	N: 96.2(49) | 75.0(26)	H: 1.9(0) | 19.4(6)	V: 1.9(0) | 5.6(1)	
T-BIL	N: 96.2(49) | 69.4(24)	H: 1.9(0) | 25.0(8)	V: 1.9(0) | 5.6(1)	
ALB	L: 13.7(6) | 54.3(18)	N: 86.3(43) | 45.7(15)		
CHE	v: 1.9(0) | 8.1(2)	L: 3.8(1) | 45.9(16)	N: 90.6(47) | 43.2(15)	H: 3.8(1) | 2.7(0)
TP	L: 3.8(1) | 2.8(0)	N: 92.3(47) | 83.3(29)	H: 3.8(1) | 13.9(4)	
T-CHO	L: 3.8(1) | 13.5(4)	N: 86.5(44) | 81.1(29)	H: 5.8(2) | 2.7(0)	V: 3.8(1) | 2.7(0)
WBC	u: 1.9(0) | 5.3(1)	v: 1.9(0) | 5.3(1)	L: 9.3(4) | 13.2(4)	N: 83.3(44) | 71.1(26)
				H: 3.7(1) | 5.3(1)
PLT	u: 1.9(0) | 8.1(2)	v: 5.7(2) | 45.9(16)	L: 28.3(14) | 35.1(12)	**N: 64.2(33) | 10.8(3)**
HGB	L: 3.9(1) | 17.1(5)	N: 96.1(48) | 82.9(28)		

model is 73.7 %. More precisely, the accuracies for exceptional cases and typical cases were 41.2 % and 80.5 % respectively. The overall accuracy is similar to that of a conventional naive Bayes classifier, but it should be noted that the predictive accuracy for LC cases is higher than the predictive accuracy obtained by a conventional naive Bayes classifier. The accuracies for LC and non-LC cases were 76.1 % and 71.7 % respectively mostly because the correctly predicted cases with the 1-NN method were all LC. Our separate prediction model, which first applies the 1-NN method for predicting the class of exceptional cases, regards LC cases important since it is adequate in predicting exceptional LC cases due to the nature of its dissimilarity measure[☆]. This fits the nature of the LC prediction problem in which overlooking of LC cases costs more than misprediction of non-LC cases.

Although it is impossible to detect an exceptional LC case who shows good results for all blood tests, our 1-NN method could detect exceptional LC cases relatively accurately. A partial LC case who shows partial aggravation of blood tests was known among medical experts only empirically, but we have succeeded in detecting several of them. Such cases might suffer from genetic problems, and detailed inspection can be expected to reveal their true causes.

[☆] Exceptional LC cases are relatively stable in their time sequences and are more easily predicted with the 1-NN method.

3.4 Results and Analysis of the Automatic Approach

In the second series of experiments, we used neither data preprocessing nor domain experts. The value of the threshold was first settled to 3, and was decremented in each subsequent spiral. The stopping condition was defined as a decrease of predictive accuracy based on a 10-fold cross validation or a reachment of the predictive accuracy to 100 %. We show the final conditional probabilities, which correspond to the disease definition, in Table 2.

We asked Yokoi, the domain expert, about the value of discriminative conditions. He classified most of the conditions as valid and known though he had never thought about the relation about AMY (amylase) and hepatitis. One of his colleagues, who has a better expertise than him in hepatitis, didn't know about the relation neither. Thus Yokoi searched MEDLINE using amylase and hepatitis as keywords, and investigated about 40 abstracts among 150 abstracts which were hit. One of the abstracts (PMID: 10063922) actually confirms our finding, especially its main objective is to report an interpretation of the finding. Such a piece of knowledge in MEDLINE can be considered as reliable and advanced. Yokoi considers that we might have succeeded in discovering a clue for a piece of knowledge which goes beyond the textbook level. It should be particularly noted that our removal of exceptional cases made the finding possible since it does not appear in the original definition.

Table 2 Disease definition of the automatic approach

GOT	**N: 28.3(27) | 4.7(4)**	H: 50.5(49) | 40.6(42)	V: 19.2(18) | 42.5(44)	U: 2.0(1) | 12.3(12)
GPT	N: 11.1(10) | 7.5(7)	H: 44.4(43) | 33.0(34)	V: 30.3(29) | 43.4(45)	U: 14.1(13) | 16.0(16)
TTT	N: 32.3(31) | 44.3(46)	H: 37.4(36) | 34.9(36)	V: 24.2(23) | 13.2(13)	U: 6.1(5) | 7.5(7)
ZTT	N: 14.1(13) | 28.3(29)	H: 70.7(69) | 57.5(60)	V: 12.1(11) | 11.3(11)	U: 3.0(2) | 2.8(2)
D-BIL	N: 85.9(84) | 32.1(33)	H: 12.1(11) | 41.5(43)	V: 1.0(0) | 14.2(14)	U: 1.0(0) | 12.3(12)
I-BIL	N: 96.0(94) | 72.6(76)	H: 2.0(1) | 18.9(19)	V: 1.0(0) | 3.8(3)	U: 1.0(0) | 4.7(4)
T-BIL	N: 96.0(94) | 61.3(64)	H: 2.0(1) | 28.3(29)	V: 1.0(0) | 2.8(2)	U: 1.0(0) | 7.5(7)
ALB	v: 1.0(0) | 3.8(3)	L: 14.3(13) | 46.7(48)	N: 84.7(82) | 49.5(51)	
CHE	v: 1.0(0) | 9.1(3)	L: 5.1(4) | 56.8(24)	N: 91.9(90) | 31.8(13)	H: 2.0(1) | 2.3(0)
TP	v: 1.0(0) | 1.9(1)	L: 4.0(3) | 10.4(10)	N: 91.9(90) | 80.2(84)	H: 3.0(2) | 7.5(7)
T-CHO	v: 1.0(0) | 2.8(2)	L: 6.1(5) | 11.2(11)	N: 87.9(86) | 73.8(78)	H: 3.0(2) | 9.3(9)
				V: 2.0(1) | 2.8(2)
WBC	u: 1.0(0) | 3.8(3)	v: 2.0(1) | 6.6(6)	L: 11.0(10) | 20.8(21)	N: 84.0(83) | 63.2(66)
				H: 2.0(1) | 5.7(5)
PLT	u: 1.0(0) | 6.7(6)	v: 11.1(10) | 34.6(35)	L: 26.3(25) | 40.4(41)	**N: 61.6(60) | 18.3(18)**
HGB	L: 3.1(2) | 26.2(26)	N: 96.9(93) | 73.8(75)		
ALP	L: 1.0(0) | 10.5(10)	N: 59.2(57) | 68.6(71)	H: 39.8(38) | 21.0(21)	
CL	L: 1.1(0) | 2.9(2)	N: 96.6(84) | 92.2(94)	H: 2.3(1) | 4.9(4)	
CRE	**L: 42.9(41) | 6.7(6)**	N: 56.1(54) | 87.6(91)	H: 1.0(0) | 5.7(5)	
K	L: 2.3(1) | 8.7(8)	N: 96.6(84) | 89.3(91)	H: 1.1(0) | 1.9(1)	
LAP	**N: 99.0(95) | 30.7(30)**	H: 1.0(0) | 69.3(69)		
LDH	L: 37.8(36) | 51.4(53)	N: 53.1(51) | 28.6(29)	H: 9.2(8) | 20.0(20)	
NA	L: 1.1(0) | 5.8(5)	N: 96.6(84) | 92.2(94)	H: 2.3(1) | 1.9(1)	
G-GTP	N: 69.1(66) | 52.5(52)	H: 30.9(29) | 47.5(47)		
UN	L: 5.1(4) | 1.0(0)	N: 93.9(91) | 89.5(93)	H: 1.0(0) | 9.5(9)	
F-ALB	L: 43.8(31) | 80.9(75)	N: 56.2(40) | 19.1(17)		
FE	L: 40.7(10) | 14.0(11)	N: 40.7(10) | 58.1(49)	H: 18.5(4) | 27.9(23)	
G.GL	N: 47.9(34) | 18.1(16)	H: 52.1(37) | 81.9(76)		
F-A1.GL	**L: 48.6(35) | 8.4(7)**	N: 50.0(36) | 85.3(80)	H: 1.4(0) | 6.3(5)	
F-A2.GL	L: 18.9(13) | 7.4(6)	N: 77.0(56) | 75.8(71)	H: 4.1(2) | 16.8(15)	
F-B.GL	L: 2.7(1) | 7.4(6)	N: 70.3(51) | 75.8(71)	H: 27.0(19) | 16.8(15)	
F-A/G	L: 4.1(2) | 44.2(41)	N: 60.8(44) | 46.3(43)	**H: 35.1(25) | 9.5(8)**	
UA	L: 4.1(3) | 18.1(18)	N: 92.9(90) | 79.0(82)	H: 3.1(2) | 2.9(2)	
U-PH	L: 4.8(3) | 12.5(11)	N: 83.1(68) | 83.3(79)	H: 12.0(9) | 4.2(3)	
IG-G	N: 37.9(10) | 24.6(16)	H: 62.1(17) | 75.4(51)		
PT	**L: 16.5(12) | 1.0(0)**	N: 74.7(58) | 82.7(80)	H: 8.9(6) | 16.3(15)	
AMY	N: 83.8(56) | 43.3(38)	H: 16.2(10) | 56.7(50)		

In overall, Yokoi considers that removal of exceptional cases seems valid since it strengthens his image of LC. According to him, most of the discriminative conditions cannot be used as predictors of LC since he uses PLT, ALB, and CHE only for most of the cases. He also pointed out that they lack of knowledge in judging LC for the additional 21 attributes compared with the initial 14 attributes. The automatic removal with 35 attributes was not exhaustive but was proved to be highly accurate. Among the 20 cases who were judged as exceptional, Yokoi judged 18 as exceptional.

4. Related Work

Liu and Motoda classified instance selection mainly into sampling, methods associated with classification, methods associated with clustering, and instance labeling[7]. Our criterion-based method in section 2.2 belongs to the second approach. The removal of irrelevant data in classification is mainly pursued in nearest-neighbor methods and instance-based learners[1),3),8)]. However, our objective is separation of irrelevant examples in classification instead of reduction of data.

Various methods for density estimation are based on likelihood, e.g. 5). While these methods assume that examples are identically distributed, we assume that our examples follow either of two distributions. Moreover, instead of assuming density of the example space, we consider to which distribution an example belongs in the context of medical data mining.

Several supports for instance selection in a KDD process are considered and presented in 9) in a general manner. We mainly assume medical data mining and obtained promising results in a specific problem related with the chronic hepatitis data.

5. Conclusions

In this paper, we have described our endeavor for disease definition with chronic hep-

atitis data. The motivation is based from our previous endeavor, from which we had come to believe that exceptional cases exist for physicians. We used typical cases in defining a target disease while we believe that each exceptional case requires individual investigation then might lead to interesting discoveries.

Selection of exceptional cases represents a highly intellectual activity in which various factors should be considered. Thus our use of data preprocessing and domain experts are justified when we try to obtain a precise definition of a disease even with a high cost. The automatic approach, which employs none of them, was proved to be relatively accurate and this shows the effectiveness of our criterion. As we described in section 3.4, Yokoi considers that we might have succeeded in discovering a clue for a piece of knowledge which goes beyond the textbook level thanks to this approach. We expect that further investigation by fine tuning various factors can bring discoveries of similar values.

Acknowledgments This work was partially supported by the grant-in-aid for scientific research on priority area "Active Mining" from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

1) D. W. Aha, D. Kibler, and M. K. Albert", "Instance-Based Learning Algorithms", *Machine Learning*, Vol. 6, No. 1, pp. 37-66, 1991.
2) P. Berka: ECML/PKDD 2003 Discovery Challenge. Download Data about Hepatitis, *http://lisp.vse.cz/challenge/ecmlpkdd2003/* (current April 26th, 2003)
3) H. Brighton and C. Mellish: "Advances in Instance Selection for Instance-Based Learning Algorithms", *Data Mining and Knowledge Discovery*, Vol. 6, No. 2, pp. 131–152, 2002.
4) P. Domingos and M. Pazzani: On the Optimality of the Simple Bayesian Classifier under Zero-One Loss, *Machine Learning*, Vol. 29, No. 2/3, pp. 103–130 (1997).
5) D. Fragoudis, D. Meretakis, and S. Kokothanassis: "Integrating Feature and Instance Selection for Text Classification", *Proc. Eighth ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD)*, pp. 501–506, 2002.
6) M. Jumi, E. Suzuki, M. Ohshima, N. Zhong, H. Yokoi, and K. Takabayashi: "Spiral Discovery of a Separate Prediction Model from Chronic Hepatitis Data", *Proc. Third International Workshop on Active Mining (AM)*, pp.1–10, 2004.
7) H. Liu and H. Motoda: "On Issues of Instance Selection", *Data Mining and Knowledge Discovery*, Vol. 6, No. 2, pp. 115–130, 2002.
8) D. Madigan et al.: "Likelihood-Based Data Squashing: A Modeling Approach to Instance Construction", *Data Mining and Knowledge Discovery*, Vol. 6, No. 2, pp. 173–190, 2002.
9) T. Reinartz: "A Unifying View on Instance Selection", *Data Mining and Knowledge Discovery*, Vol. 6, No. 2, pp. 191–210, 2002.
10) E. Suzuki, T. Watanabe, H. Yokoi, and K. Takabayashi: Detecting Interesting Exceptions from Medical Test Data with Visual Summarization, *Proc. Third IEEE International Conference on Data Mining (ICDM)*, pp. 315-322 (2003).
11) Y. Yamada, E. Suzuki, H. Yokoi, and K. Takabayashi: Decision-tree Induction from Time-series Data Based on a Standard-example Split Test, *Proc. Twentieth International Conference on Machine Learning (ICML)*, pp. 840-847 (erratum http://www.slab.dnj.ynu.ac.jp/ erratumicml2003.pdf) (2003).
12) N. Zhong, Y. Y. Yao, and M. Ohshima: Peculiarity Oriented Multi-Database Mining, *IEEE Transaction on Knowledge and Data Engineering*, Vol. 15, No. 4, pp. 952-960 (2003).