

# 客観的指標群に基づくクラスタリングと 主観的評価基準の比較

## A Comparison between Clustering Algorithms based on Objective Rule Evaluation Indices and Human Subjective Criteria

阿部 秀尚<sup>\*1</sup>  
Hidenao Abe

津本 周作<sup>\*1</sup>  
Shusaku Tsumoto

横井 英人<sup>\*2</sup>  
Hideto Yokoi

大崎 美穂<sup>\*3</sup>  
Miho Ohsaki

山口 高平<sup>\*4</sup>  
Takahira Yamaguchi

<sup>\*1</sup> 島根大学医学部  
Shimane University, School of Medicine

<sup>\*2</sup> 香川大学医学部附属病院  
Kagawa University Hospital

<sup>\*3</sup> 同志社大学工学部  
Faculty of Engineering, Doshisha University

<sup>\*4</sup> 慶應義塾大学理工学部  
Faculty of Science and Technology, Keio University

This paper presents a comparison between clusterings based on objective rule evaluation indices and human subjective criterion to support an evaluation procedure of a data mining process with a human-system interaction. Post-processing of mined results is one of the key issues to make a data mining process successful. However, it is difficult for human experts to evaluate many thousands of rules from a large dataset with noises completely. As an active support method, we have designed a prediction method based on rule evaluation models, which are learned from objective measure values and evaluations of a human expert for each rule. To evaluate performances of clustering algorithms, we have done a case study with meningitis and chronic hepatitis data mining results. Then, we discuss about the applicability of clustering for our rule evaluation support method.

### 1. はじめに

データマイニングにおけるマイニング結果の後処理では、専門家によるマイニング結果の評価手続きがデータマイニングプロジェクト全体の成否を左右する重要な手続きとして位置づけられている [CRISP-DM]。しかし、専門家が介入する手続きには、専門家の養成や評価の獲得に多大なコストがかかる問題がある。さらに、通常、データマイニングプロセスは1つのデータマイニングプロジェクトの中で繰り返し実行される必要がある。また、その繰り返しの中で各手続きも反復して実行される。そのため、コストのかかるマイニング結果の評価作業も繰り返し実行する必要がある、データマイニングプロジェクトを実行するための全体のコストを大きく引き上げる。

以上の問題に対し、我々は、マイニング結果のうち IF-THEN ルールを対象にしたルール評価支援手法を開発してきた [Abe 05]。本手法は、専門家のルール評価作業を客観的指標群に基づくルール評価モデルを用いて反復的な作業を支援することを目的としている。ルール評価モデルは、各ルールを評価する客観的指標値と専門家の評価（教師信号）を基にルール評価モデルを学習アルゴリズムにより獲得し、新規ルールの評価ラベルを予測することにより専門家のルール評価作業を支援する（図1参照）。

本手法は、ルール評価作業の繰り返しを想定し、専門家による所与のルール集合の一部の評価からルール評価モデルを構築して評価ラベルの予測を行うことを可能としている。しかしながら、ルール評価モデルを客観的指標群の値に基づく分類モデルとして学習し、予測を提示するだけでは専門家の評価支援として不十分である。特に、専門家の主観的基準が確立されていない状態では、分類モデルの学習が困難であり、データセット上での集まりを見出すクラスタリングが適していると考えら

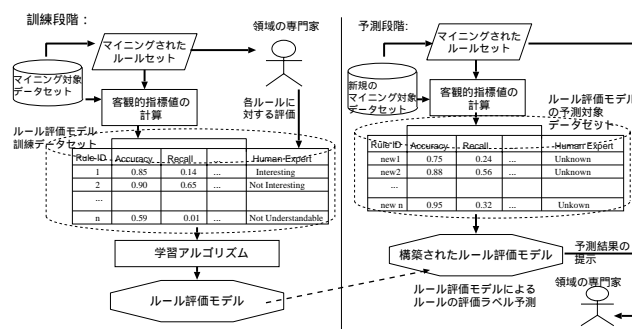


図 1: 客観的指標群に基づくルール評価支援手法の概観。

れる。

本稿では、クラスタリングアルゴリズムによって生成されたクラスと専門家が主観的評価基準によって与えたクラスとの比較を行う。この結果から、客観的指標群によって表現されたクラスが専門家の主観的評価基準をどの程度推測可能かを評価する。このため、専門家の主観的評価基準の確立度合いが異なる5種類のルール評価結果を用い、クラスタリングによるクラスと評価ラベルの一致度を基に本手法へのクラスタリングの適用可能性を議論する。

### 2. 客観的指標群に基づくクラスタリングアルゴリズムの性能評価実験

本実験では、各ルールに対し、確率、統計検定、情報量、事例数、属性値の差異の各観点に基づく39種類の客観的指標 [Ohsaki 04] を算出し、データセットを作成する。なお、各客観的指標は  $0 \leq x \leq 1$  で正規化されている。本実験で利用するデータマイニング結果は、髄膜炎データセットに対する先行研究 [Hatazawa 00] の評価結果と、慢性肝炎データに対

表 1: 6 種類の髄膜炎データセットとそれらに対するマイニング結果 .

データ名	属性数	クラス数	ルール数	興味深い	妥当	理解不能
Diag	29	6	53	15	38	0
C_Course	40	12	22	3	18	1
Culture+diag	31	12	57	7	48	2
Diag2	29	2	35	8	27	0
Course	40	2	53	12	38	3
Cult_find	29	2	24	3	18	3
計	—	—	244	48	187	9

する 2 種類の問題についてのデータマイニングの評価結果である .

髄膜炎に対するデータマイニングでは, 140 レコードから成る 6 種類の鑑別診断に関するルール計 244 ルールに対し, 専門医のコメントから 3 種類の評価ラベル (I: 興味あり・NI: 妥当・NU: 理解不能) を与えた評価結果である (表 1) .

慢性肝炎データセットについては, 42 種類の検査項目についての任意の期間のパターンを条件部として, GPT のパターン [Ohsaki 04] とインターフェロン (IFN) の検査結果 [Abe 06] を結論部とした 2 種類の IF-THEN ルール集合の生成を行い, 専門家によって 4 種類の評価ラベル (EI: 特に興味あり・I: 興味あり・NI: 妥当・NU: 理解不能) が与えられた評価結果である (表 2) .

表 2: 4 種類の慢性肝炎データマイニング結果と評価ラベルの内訳 .

	#Rules	Class Distribution				%Def class
		EI	I	NI	NU	
GPT						
Phase1(GPT1)	30	3	8	16	3	53.33
Phase2(GPT2)	21	2	6	12	1	57.14
IFN						
First Time(IFN1)	26	4	7	11	7	42.31
Second Time(IFN2)	32	15	5	11	1	46.88

これらのデータセットに対し, Weka[Witten 00] に実装された K-Means, EM クラスタリングの 2 種類のクラスタリングアルゴリズムを適用した . K-Means については, 所与のデータセットでのクラス数をクラスタ数として指定し, EM クラスタリングのパラメータは規定値を用いた . 各クラスタへのクラス値の割り当ては, データセット中で割合の高いクラスから順にクラスタに割り当てられたインスタンス数の多いクラスをクラスタのクラス値として割り当てる .

表 3 に示すように, K-Means では GPT パターンを結論部としたデータマイニング評価において高い精度で専門家の評価基準を推定できた . また, EM アルゴリズムでは全データセットにおいて過半数のデータを正しくクラスタに纏めることが困難であるという結果となった .

表 3: クラスタとクラス値が一致した割合 (%) .

	髄膜炎	GPT1	GPT2	IFN1	IFN2
K-Means	36.1	43.3	71.4	42.3	34.4
EM	19.3	50.0	47.6	42.3	50.0

我々が提案する医療データマイニングにおける仮説生成段

階では, GPT2 は「仮説検証」段階にあたり, 専門家が主観的評価基準を確立しつつある段階である . 一方, クラスタリングの推定精度が特に低かった髄膜炎データマイニング結果や IFN2 は「仮説洗練化」段階にあたり, 専門家の主観的評価基準は確立されている . 以上のことから, クラスタリングによるルール評価モデルに基づくルール評価支援は, 専門家の主観的評価基準が十分に確立されていない段階で有効に機能すると言える .

### 3. おわりに

本稿では, 客観的指標によりインタラクティブに専門家の IF-THEN ルールの評価作業を支援する手法について, より柔軟な支援を可能とするため, ルール評価モデルの構築にクラスタリングを適用し, ルール評価モデルの構築性能について評価実験を行った . この結果, クラスタリングによる専門家の主観的評価基準の推定は, 専門家の主観的評価基準が確立されつつある仮説検証段階において有効であることが示された .

今後は, 主観的ルール評価指標との比較, クラスタリングアルゴリズムの頑強性の評価を行っていく予定である .

### 参考文献

- [Abe 04] Abe, H. and Yamaguchi, T.: Constructive Meta-Learning with Machine Learning Method Repositories, in Proc. of the 17th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE 2004, pp. 502-511 (2004).
- [Abe 05] Abe, H., Tsumoto, S., Ohsaki, M., and Yamaguchi, T.: A Rule Evaluation Support Method with Learning Models Based on Objective Rule Evaluation Indexes, in Proc. of the Fifth International Conference on Data Mining (ICDM2005), pp. 549-552 (2005).
- [Abe 06] Abe, H., Ohsaki, M., Yokoi, H., and Yamaguchi T.: Implementing an Integrated Time-Series Data Mining Environment based on Temporal Pattern Extraction Methods – A Case Study of an Interferon Therapy Risk Mining for Chronic Hepatitis –, JSAI2005 Workshops, LNAI4012, pp. 425-435 (2006).
- [CRISP-DM] CRISP-DM: <http://www.crisp-dm.org/>
- [Hatazawa 00] Hatazawa, H., Negishi, N., Suyama, A., Tsumoto, S., and Yamaguchi, T.: Knowledge Discovery Support from a Meningoencephalitis Database Using an Automatic Composition Tool for Inductive Applications, in Proc. of KDD Challenge 2000 in conjunction with PAKDD2000, pp. 28-33, (2000).
- [Ohsaki 04] Ohsaki, M., Kitaguchi, S., Kume, S., Yokoi, H., and Yamaguchi, T.: Evaluation of Rule Interestingness Measures with a Clinical Dataset on Hepatitis, in Proc. of ECML/PKDD 2004, LNAI3202, pp. 362-373 (2004).
- [Witten 00] Witten, I. H and Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann, (2000).