러프집합과 퍼지 네이브 베이스 이론을 이용한 효율적인 추론 방법

The Method of Effective Inference Using Rough Set and Fuzzy Naive Bayes Theory

황정식, 손창식, 정환묵
대구가톨릭대학교 컴퓨터정보통신공학부

Jeong-Sik Hwang, Chang-Sik Son, Hwan-Mook Chung
Faculty of Computer and Information Communication Engineering
Catholic University of Daegu
E-mail : icsman@cu.ac.kr

요 약

퍼지 규칙 기반 시스템에서 분류 및 경계를 결정하기 위한 방법으로 퍼지 규칙을 학습하는 다양한 방법들이 제안되고 있다. 그리고 추론 규칙간의 상관성을 고려하여 불필요한 속성을 제거함으로써 좀 더 효율적인 추론 결과를 얻을 수 있다.

따라서 본 논문에서는 퍼지 규칙 기반 시스템에서 각 규칙에 따른 결정 테이블를 작성하고 러프집합을 이용하여 불필요한 속성을 제거하였으며 규칙의 확신도에 퍼지 네이브 베이스 이론을 적용한 추론 방법을 제안한다.

keywords: fuzzy inference, rough set, fuzzy naive bayes

1. 서론

최근 퍼지규칙 기반 시스템은 주어진 입력 벡터를 특정 클래스 영역으로 분류하기 위해 수치적인 데이터로부터 퍼지 IF-THEN 규칙을 학습하는 다양한 방법들이 제안되었다[1-2]. 이들 제안된 방법들의 목적은 추론규칙을 학습함으로써 분류 경계를 결정하는데 있다. 그러나 분류 경계를 결정하기 위한 수단으로 학습에 의한 방법은 다소 많은 시간이 소비된다. 또한 추론 규칙간의 상관성을 고려하지 않는다면, 불분명한 경계를 포함하게 되어 명확한 추론결과를 얻을 수 없다.

이러한 문제점들을 개선하기 위한 방법으로 분류문제에서 발생되는 여러 가지 유형의 불완전한 정보의 의존성과 속성의 수를 감축하기 위한 선택 도구로서 러프집합을 이용한 방법들과 퍼지규칙에서 전건부와 후건부의 소속함수를 학습하지 않고 규칙의 확신도만을 고려하여 분류 경계를 결정할 수 있는 방법이 제안되었다[3-6].

따라서 본 논문에서는 러프집합에 특성과 퍼지 네이브 베이지 이론을 이용한 퍼지추론 방법을 제안한다. 여기서 러프집합은 규칙간의 상관성을 고려하여 코어가 되는 규칙을 선택하기 위해 사용되었고, 퍼지 네이브 베이지 이론은 규칙의 확신도를 계산하기 위해 사용되었다. 그리고 모의 실험에서는 규칙 감축 전과 후의 확신도를 반영한 추론결과를 비교하였다.

2. 관련연구

2.1 러프집합

러프집합은 단순히 구조적인 방법들로서 데이터-집합(dataset)에 포함된 속성들의 수를 감축하고 데이터의 의존성을 찾기 위한 선택 도구이다. 러프집합의 성질 중 지식 감축에는 2가지('리덕트', '코어') 개념들이 중요한 역할을 한다. 지식의 리덕트는 지식의 본질적인 부분으로 지식에 나타나는 모든 기본적 개념들을 정의하기에 충분하고, 반면에 코어는 어떤 의미에서 가장 핵심적인 부분이다. 러프집합의 지식표현에서 식별가능 행렬은 코어, 리덕트로 다른 개념들을 간단히 계산할 수 있게 해 준다[7,8].

[정의 2.1] 만약 $U=\{x_1,x_2,\ldots,x_n\}$인 지식 표현 시스템에서 $S=(U,A)$에서, M(S)로 표기되는 지식 표현 시스템 S의 식별행렬은 다음과 같은

$n \times n$ 행렬이다.

$$(c_{ij}) = \{a \in A : a(x_i) \neq a(x_j)\} \quad (i,j = 1,2,\ldots,n)$$

여기서 성분 c_{ij}는 x_i와 x_j을 구별하게 하는 모든 속성들의 집합이다.

[정의 2.2] 코어는 식별행렬의 모든 단위 원소 엔트리의 집합이다.

$$CORE(A) = \{a \in A : c_{ij} = (a),\ \exists\, i,j\}$$

[정의 2.3] B가 "M(A)의 임의의 엔트리 c(c $\neq \varnothing$)에 대해서, $B \cap c \neq \phi$"인 조건을 만족하는 A의 최소 부분집합이면, $B \subseteq A$인 B는 A의 리덕트이다. 즉, A가 속성들의 전체 집합이고, $B \subseteq A$인 B가 A의 리덕트이면, B는 A와 동일한 분류를 제공하므로, 객체들은 A을 이용하여 구별하는 것과 B을 이용하여 구별하는 것은 동일하다.

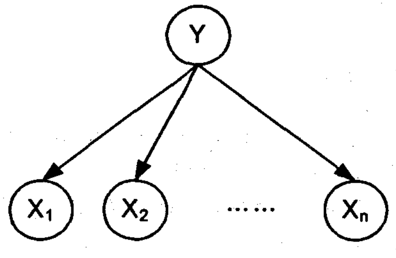
[정의 2.4] 모든 식별 가능 행렬 $M(A)$는 식별가능(부울)함수 $f(A)$로 정의된다.

$$f(A) = \prod_{(x,y) \in U^2} \{\Sigma\delta(x,y) : (x,y) \in U^2 \text{ and } \delta(x,y) \neq 0\}$$

여기서 $\Sigma\delta(x,y)$는 속성들의 집합 $\delta(x,y)$에 할당된 모든 부울 변수들의 부울 합이다.

2.2 퍼지 네이브 베이스

퍼지 네이브 베이스에서는 각 노드를 언어변수에 대응시키고 각 변수의 값은 퍼지 집합을 적용한다. 퍼지 네이브 베이스 이론의 일반적인 구조는 다음과 같다[9].



[그림 1] 퍼지 네이브 베이스

[그림 1]에서 $X_1, X_2, \ldots, X_n$은 전건부 변수, Y는 후건부 변수를 나타낸다.

그리고 후건부 변수 Y에 할당된 사전 확률(prior probabilities)을 계산하기 위한 식은 다음과 같다.

$$P(Y = B_k) = \frac{\sum_{X \in D} B_k(y)}{\sum_{k=1}^{m} \sum_{X \in D} B_k(y)} \tag{1}$$

여기서 B_k는 k번째 규칙의 후건부 변수의 값을 나타내고, $\overline{X}$와 D는 각각 관측된 데이터-집합(data-set)에 포함된 원소를 의미한다.

사전 확률을 제외한 다른 노드들 간의 조건부 확률은 다음 식으로 계산된다.

$$P(X_i = A_j^i \mid Y = B_k) = \frac{\sum_{X \in D} A_j^i(x_i) B_k(y)}{\sum_{j=1}^{k_i} \sum_{X \in D} A_j^i(x_i) B_k(y)} \tag{2}$$

$$(A_j^i : j = 1,2,..,k_i,\ B_k : k = 1,2,..,m)$$

여기서 $P(X_i = A_j^i \mid Y = B_k)$는 k번째 후건부 변수의 소속함수가 주어졌을 때, j번째 규칙에서 i번째 전건부 변수의 소속함수의 값을 나타내고, x_i는 벡터 $\overline{X} \in D$ ($\overline{X} = [X;y]$)의 i번째 원소를 의미한다.

여기서 X와 y는 각각 전체 전건부 변수 벡터와 후건부 변수의 원소를 나타낸다. 즉, 관측된 데이터-집합에서 후건부의 변수의 소속함수가 주어졌을 때, i번째 전건부 변수의 소속함수의 평가치를 나타낸다.

3. 러프집합과 퍼지 네이브 베이스 이론을 이용한 추론 방법

본 논문에서는 퍼지 규칙 내에 포함된 유사 속성을 러프집합의 특성을 이용하여 코어가 되는 속성을 추출하고 퍼지 네이브 베이스 이론을 사용하여 확신도(certainty factor)를 계산한 후, 불필요한 속성을 제거한 추론 방법을 제안한다. 제안된 방법의 처리 단계는 다음과 같다.

[단계 1] 퍼지추론 규칙을 구성한다.
[단계 2] 러프집합의 성질을 이용하여 코어가 되는 규칙을 추출한다.
[단계 3] 퍼지 네이브 베이스를 사용하여 [단계 1]과 [단계 2]에 대한 추론규칙의 확신도를 계산한다.
[단계 4] [단계 3]에 각 확신도의 차이점을 분석한다.
[단계 5] 감축 전과 후의 규칙과 확신도를 바탕으로 추론한다.
[단계 6] 각 추론된 결과값을 비교 분석한다.

퍼지추론 규칙은 다음과 같다.

$$\begin{aligned} Rule(i):\ & If\ X_1\ is\ A_i^1 \text{ and } X_2\ is\ A_i^2, \ldots, X_n\ is\ A_i^n \\ & Then\ Y\ is\ B_i\ with\ CF_i \\ & \text{for } i = 1,2,\ldots,M,\ CF_i \in (0,1] \end{aligned} \tag{6}$$

여기서 $X_1, X_2, \ldots, X_n$은 전건부 변수, Y는 후

건부 변수, 그리고 $A_i^1, A_i^2, \ldots, A_i^n$, B_i , CF_i는 각각 i번째 전건부와 후건부의 소속함수, 확신도를 나타낸다.

3.1 러프집합에 의한 규칙의 간략화

본 논문에서 사용한 퍼지추론 규칙은 다음과 같다.

$$If\ X_1\ is\ A_{ij}\ \text{and}\ X_2\ is\ B_{ij}\ Then\ Y\ is\ C_{ij}\ with\ CF_{ij} \quad (i,j=1,2,3,4)$$

X_2 \ X_1	A_1	A_2	A_3	A_4
B_1	α	γ	β	δ
B_2	β	α	γ	δ
B_3	γ	α	γ	β
B_4	β	δ	α	δ

[표 1] 추론 규칙의 결정 테이블

[정의 2.1]을 이용하여 지식 표현 시스템에서 식별 행렬을 계산하고 [정의 2.2] ~ [정의 2.4]를 이용하여 최종 식별 가능 함수 $f(A)$을 계산하면 다음과 같다.

$$f(A) = A_1A_2 + A_1A_3 + A_2A_4$$

여기서 최종 식별 가능 함수 $f(A)$는 3개의 리덕트 $\{A_1, A_2\}$, $\{A_1, A_3\}$, $\{A_2, A_4\}$가 생성됨을 알 수 있다.

이와 같이 3개의 각 리덕트에 대한 식별 가능 함수를 계산하면 다음과 같다.

X_2 \ X_1	A_1	A_2	A_1	A_3	A_2	A_4
B_1	α	γ	α	β	γ	ϕ
B_2	β	α	β	γ	α	δ
B_3	γ	ϕ	γ	ϕ	ϕ	β
B_4	ϕ	δ	ϕ	α	δ	ϕ

[표 2] 최종 식별 가능 함수 $\{A_1, A_2\}, \{A_1, A_3\}, \{A_2, A_4\}$

[표 2]에서 ϕ는 러프집합을 사용함으로써 원래 추론 규칙에서 감축된 규칙을 의미한다.

3.2 확신도 계산

퍼지 네이브 베이스 이론에 따라 [표 1]의 16개의 추론규칙에 대한 각각의 확신도를 계산할 수 있다. 예를 들어, 첫 번째 규칙의 확신도를 계산하면 다음과 같다.

$$\begin{aligned} & P(Y=\alpha \mid X_1 = A_1, X_2 = B_1) \\ &= \frac{P(X_1 = A_1, X_2 = B_1 \mid Y=\alpha)}{P(X_1 = A_1, X_2 = B_1)} \\ &= \frac{P(X_1 = A_1, X_2 = B_1 \mid Y=\alpha) P(Y=\alpha)}{P(X_1 = A_1, X_2 = B_1)} \\ &= \frac{P(X_1 = A_1, X_2 = B_1 \mid Y=\alpha) P(Y=\alpha)}{\sum_{C_i=\alpha,\beta,\gamma,\delta} P(X_1 = A_1, X_2 = B_1, Y=C_i)} \\ &= \frac{P(X_1 = A_1, X_2 = B_1 \mid Y=\alpha) P(Y=\alpha)}{\sum_{C_i=\alpha,\beta,\gamma,\delta} P(X_1 = A_1, X_2 = B_1 \mid Y=C_i) P(Y=C_i)} \end{aligned}$$

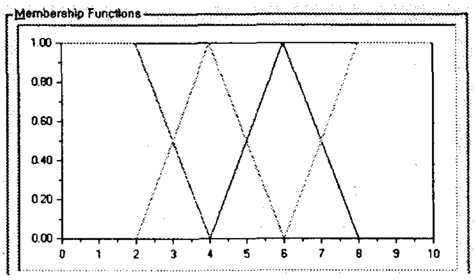
이러한 방법으로 각 추론규칙에 대한 확신도를 계산하면 다음과 같다.

X_2 \ X_1	A_1	A_2	A_3	A_4
B_1	0.25	0.25	0.25	0.75
B_2	0.5	0.5	0.5	0.75
B_3	0.4	0.5	0.67	1
B_4	0.67	0.5	0.5	0.86

[표 3] 규칙 감축 전 확신도

4. 모의실험

예를 들면, 전건부 변수와 후건부 변수의 소속함수와 추론 연산자는 다음과 같다고 가정한다.



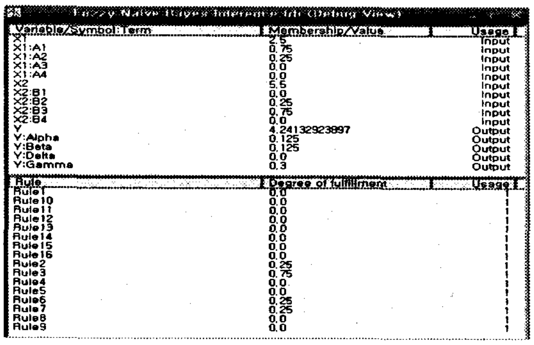
[그림 2] 전건부와 후건부 소속함수

4.1 규칙 감축 전 추론 결과 (확신도 포함)

만약 입력이 $X_1 = 2.5$, $X_2 = 5.5$ 로 주어졌다면 [표 1]에 사상되는 규칙과 확신도는 다음과 같다.

$R_2 : If\ X_1\ is\ A_1\ \text{and}\ X_2\ is\ B_2\ Then\ Y\ is\ \beta\ with\ 0.5$
$R_3 : If\ X_1\ is\ A_1\ \text{and}\ X_2\ is\ B_3\ Then\ Y\ is\ \gamma\ with\ 0.4$
$R_6 : If\ X_1\ is\ A_2\ \text{and}\ X_2\ is\ B_2\ Then\ Y\ is\ \alpha\ with\ 0.5$
$R_7 : If\ X_1\ is\ A_2\ \text{and}\ X_2\ is\ B_3\ Then\ Y\ is\ \alpha\ with\ 0.5$

위의 추론규칙과 확신도를 바탕으로 추론한 결과는 아래와 같다.



Variable/Symbol:Term	Membership/Value	Usage
X1	2.5	Input
X1:A1	0.75	Input
X1:A2	0.25	Input
X1:A3	0.0	Input
X1:A4	0.0	Input
X2	5.5	Input
X2:B1	0.0	Input
X2:B2	0.25	Input
X2:B3	0.75	Input
X2:B4	0.0	Input
Y	4.24132923897	Output
Y:Alpha	0.125	Output
Y:Beta	0.125	Output
Y:Delta	0.0	Output
Y:Gamma	0.3	Output

Rule	Degree of fulfillment	Usage
Rule1	0.0	1
Rule10	0.0	1
Rule11	0.0	1
Rule12	0.0	1
Rule13	0.0	1
Rule14	0.0	1
Rule15	0.0	1
Rule16	0.0	1
Rule2	0.25	1
Rule3	0.75	1
Rule4	0.0	1
Rule5	0.0	1
Rule6	0.25	1
Rule7	0.25	1
Rule8	0.0	1
Rule9	0.0	1

[그림 3] 규칙 감축 전 확신도를 고려한 결과

입력값에 따른 추론 결과값은 약 4.24이고 후건부 소속함수의 적합도는 α와 β구간에서 0.125, γ구간에서 0.3임을 알 수 있다.

4.2 규칙 감축한 후 추론 결과 (확신도 포함)

여기서는 위에서 구한 추론 결과값과 비교하기 위해서 동일한 입력값을 주었다. [표 1]추론규칙의 결정 테이블을 러프집합의 특성을 이용하여 3가지 형태의 추론규칙을 추출하였다([표 2]참조).

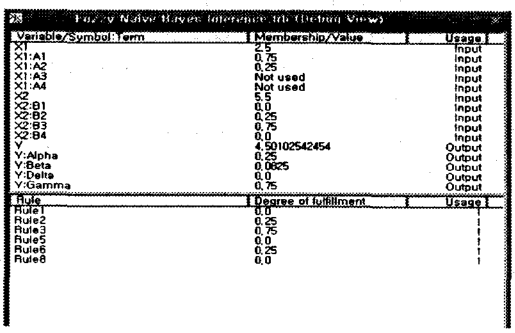
추출된 규칙은 어떤 추론규칙을 사용하여도 상관없으므로 실험에서는 $\{A_1, A_2\}$의 추론규칙을 사용한다. $\{A_1, A_2\}$의 추론규칙에 대한 확신도를 계산하면 다음과 같다.

X_2 \ X_1	A_1	A_2
B_1	0.5	1.0
B_2	0.33	1.0
B_3	1.0	ϕ
B_4	ϕ	1.0

[표 4] 규칙 감축 후 확신도

R_2 : *If* X_1 *is* A_1 and X_2 *is* B_2 *Then* Y *is* β *with* 0.33
R_3 : *If* X_1 *is* A_1 and X_2 *is* B_3 *Then* Y *is* γ *with* 1.0
R_6 : *If* X_1 *is* A_2 and X_2 *is* B_2 *Then* Y *is* α *with* 1.0

위의 추론규칙과 확신도를 바탕으로 추론한 결과는 아래와 같다.



Variable/Symbol:Term	Membership/Value	Usage
X1	2.5	Input
X1:A1	0.75	Input
X1:A2	0.25	Input
X1:A3	Not used	Input
X1:A4	Not used	Input
X2	5.5	Input
X2:B1	0.0	Input
X2:B2	0.25	Input
X2:B3	0.75	Input
X2:B4	0.0	Input
Y	4.50102542454	Output
Y:Alpha	0.25	Output
Y:Beta	0.0825	Output
Y:Delta	0.0	Output
Y:Gamma	0.75	Output

Rule	Degree of fulfillment	Usage
Rule1	0.0	1
Rule2	0.25	1
Rule3	0.75	1
Rule5	0.0	1
Rule6	0.25	1
Rule8	0.0	1

[그림 4] 규칙 감축 후 확신도를 고려한 결과

규칙 감축 후 추론 결과값은 약 4.5이고 후건부 소속함수의 적합도는 α구간에서 0.25, β구간에서 0.0825, γ구간에서 0.75임을 알 수 있다.

5. 결론 및 향후 연구과제

본 논문에서는 러프집합에 특성과 퍼지 네이브 베이지 이론을 이용한 퍼지추론 방법을 제안하였다. 러프집합은 규칙간의 상관성을 고려하여 코어가 되는 규칙을 선택하기 위해 사용되었고, 퍼지 네이브 베이지 이론은 규칙의 확신도를 계산하기 위해 사용되었다. 또한 실험을 통해 규칙 감축 후 확신도가 더욱 명확하게 나타남을 알 수 있었고 보다 명확한 추론 결과를 얻을 수 있었다.

6. 참고문헌

[1] C. S. Lee and C. Y. Pan, “An Intelligent Fuzzy Agent for Meeting Scheduling Decision Support System”, Fuzzy Sets and Systems, vol. 142, pp. 467-488, 2004.

[2] D. Nauck and R. Kruse, “A Neuro-Fuzzy Method to Learn Fuzzy Classification Rules from Data”, Fuzzy Sets and Systems, vol. 89 pp. 277-288, 1997.

[3] 정구범, 정환묵, “러프집합을 이용한 정보시스템에서의 불완전 데이터 처리에 관한 연구”, 한국 퍼지 및 지능시스템학회 논문지, 제9권 3호, pp. 1225-1127, 1999.

[4] 정환묵, 피수영, 최경옥, “러프-신경망과 χ^2 검정에 의한 효율적인 의사결정지원 시스템”, 한국 퍼지 및 지능시스템학회 논문지, 제6권 8호, pp. 2106-2212, 1999.

[5] 정환묵, “개념 상승과 속성의 최적 감축에 의한 결정 규칙의 생성”, 한국 퍼지 및 지능시스템학회 논문지, 제9권 4호, pp. 367-374, 1999.

[6] H. Ishibuchi, T. Nakashima, “Effective of Rule Weights in Fuzzy Rule-Based Classification Systems”, IEEE Transaction on Fuzzy Systems, vol. 9, 2001.

[7] Z. Pawlak, “Rough Sets : Theoretical Aspects of Reasoning about Data”, Kluwer Academic Publishing, Dordrecht, 1991.

[8] Ewa Orlowska. “Incomplete Information : Rough Set Analysis”. A Springer-Verlag Company. pp. 23-57, 95-108. 1997.

[9] Yongchuan Tang, and Yang Xu, “Application of Fuzzy Naive Bayes and a Real-Valued Genetic Algorithm in Identification of Fuzzy Model”, Journal of Information Sciences, vol. 169, pp. 205-226, 2005.