

소속 함수 학습을 이용한 퍼지 분류의 성능 개선

Improving the Performance of Fuzzy Classification Using Membership Function Learning

곽동헌⁰, 김명원^{*}

인포아이⁰, 송실대학교 컴퓨터학과^{*}

Dong-Heon Kwak⁰, Myung-Won Kim^{*}

INFOEYE Co., Ltd.⁰ Dept. of Computing, Soongsil University^{*}

E-mail : lovekumdo@infoeye.co.kr

요 약

수치적인 데이터를 분류하기 위한 대표적인 방법은 퍼지 규칙을 사용하는 것이다. 하지만, 이러한 방법은 퍼지 소속 함수를 어떻게 정의하느냐에 따라 퍼지 분류의 성능이 크게 영향을 받는다는 문제점과 퍼지 규칙을 쉽게 이해하기 위해 가능한 퍼지 규칙의 수를 적게 유지해야 한다는 문제점이 있다.

본 논문에서는 효과적이며 이해하기 쉬운 퍼지 규칙을 생성하기 위해 기울기 강하법을 기반으로 하는 소속 함수 학습 방법을 제안한다. 어려움을 감소하기 위해 Penalty 연산과 Reward 연산을 통해 소속 함수가 반복적으로 조절된다. 새로운 소속 함수는 Coverage 연산에 의해 생성된다. 또한 이해하기 쉬운 퍼지 규칙을 최적화하기 위해 학습된 소속 함수를 퍼지 결정 트리에 적용한다. 본 논문에서 제안한 알고리즘의 타당성을 확인하기 위해 벤치 마크 데이터인 Iris, Wisconsin Breast Cancer, Pima, Bupa 데이터를 이용하여 실험 결과를 보인다. 실험 결과를 통해 제안한 알고리즘이 기존의 C4.5와 FID 3.1 알고리즘보다 더 효과적이거나 비슷한 성능을 보임을 알 수 있다.

1. 서론

수치적인 데이터를 분류하기 위한 방법으로는 퍼지 결정 트리와 퍼지 규칙에 의한 분류 방법을 들 수 있다. 퍼지 결정 트리는 ID3, C4.5[1][2]의 이해성과 퍼지 이론의 추론과 표현력을 결합한 방법이다. 퍼지 결정 트리[3]에 의해 생성된 퍼지 규칙은 속성 축에 평행하게 판단 경계선을 결정하는 방법으로는 어려운 패턴을 효율적으로 분류하는 특징이 있다. 퍼지 규칙에 의한 분류 방법은 속성수와 각 속성을 표현하는 소속 함수 개수를 이용하여 퍼지 규칙을 생성하고 새로운 데이터를 분류하는 방법이다. 퍼지 규칙은 입력 데이터의 속성수와 각 속성을 표현하는 소속 함수 개수의 조합 수만큼 생성되고, 각 퍼지 규칙에 대

한 클래스와 신뢰도가 계산된다[4][5]. 생성된 퍼지 규칙으로부터 결론을 도출하기 위해 대표규칙 (Single Winner Rule)을 선택하며, 선택된 대표 규칙을 통해 새로운 데이터를 분류한다[4][5]. 하지만 이러한 분류 방법은 소속 함수의 모양과 개수에 따라 패턴 분류 시 결과가 민감해진다는 문제점이 있다. 또한 퍼지 규칙에 의한 분류 방법은 속성수와 각 속성을 정의하는 소속 함수 개수의 조합 수만큼 퍼지 규칙이 생성됨으로 생성된 퍼지 규칙을 이해하기 어려운 문제점이 있다.

본 논문에서는 이러한 문제점을 해결하기 위해 소속 함수의 모양과 개수를 자동적으로 조절하는 퍼지 규칙의 소속 함수 학습 방법(PRC Learning)을 제안한다. PRC Learning 방법은

Penalty 연산과 Reward 연산을 통해 속성에 대한 소속 함수를 학습시키며, Coverage 연산을 통해 속성에 대한 소속 함수의 개수를 자동적으로 조절하는 방법이다. 또한 이해하기 쉬운 퍼지 규칙을 최적화하기 위해 학습된 소속 함수를 퍼지 결정 트리에 적용한다. 본 논문에서 제안한 알고리즘의 타당성을 확인하기 위해 벤치 마크 데이터인 Iris, Wisconsin Breast Cancer, Pima, Bupa 데이터를 이용하여 실험 결과를 보인다.

2. 관련 연구

분류에 있어서 수치적인 속성인 경우 구간 값을 어떻게 정하는지에 따라 결과가 민감해 진다는 문제점이 있다. 이러한 문제점은 불확실하거나 애매한 개념을 처리할 때 사용되는 퍼지 이론으로 보완할 수 있다. 하지만 일반적으로 퍼지 규칙을 유도할 때, 직관적이고 경험적으로 소속 함수를 생성하기 때문에 소속 함수를 어떻게 정의하느냐에 따라 퍼지 시스템의 성능이 크게 영향을 받는다. 패턴 분류 문제에서 이러한 문제를 해결하기 위한 기존의 소속 함수 학습 방법으로는 유전자 알고리즘을 통해 최적의 소속 함수 모양과 개수를 결정하는 방법[3]과 온라인 학습을 통한 오류를 보정하는 방법[4]이 있다.

히스토그램을 기반으로 하는 소속 함수 학습 방법은 초기 소속 함수를 생성하고, 유전자 알고리즘을 통해 생성된 초기 소속 함수를 보정하여 최적의 소속 함수 모양과 개수를 찾아 주지만, 학습 시간이 오래 걸린다는 문제점과 학습 변수에 성능이 민감하다는 문제점을 가지고 있다. 또한 삼각 소속 함수만을 생성하기 때문에 일반성이 결여되어 있다. 반면 오류 보정 학습 방법은 퍼지 규칙의 분류 능력을 향상시키기 위해 잘못 분류하는 퍼지 규칙의 소속 함수에 대해서만 학습을 수행하기 때문에 수행 속도가 빠르다는 장점이 있지만, 소속 함수 개수를 각 속성별로 사용자가 정의해 주어야 한다는 문제점을 가지고 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 소속 함수 학습 방법(PRC Learning)과 학습된 소속 함수를 퍼지 결정 트리에 적용하여 이해하기 쉬우며 정확도가 높은 퍼지 규칙을 생성하는 알고리즘을 제안한다.

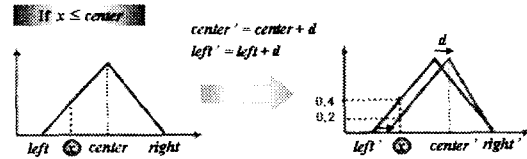
3. PRC Learning과 퍼지 결정 트리를 이용한 지식 추출

3.1 소속 함수 학습 방법(PRC Learning)

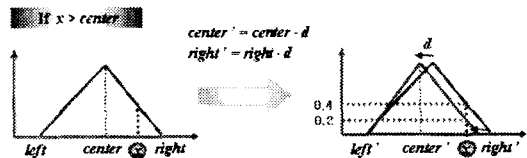
3.1.1 Penalty 연산

Penalty 연산은 소속 함수의 소속 정도를 낮춰 주는 연산이다. [그림1], [그림2]와 같이 $left$, $center$, $right$ 세 개의 x 좌표 값을 갖는 삼각형 소속 함수가 주어졌을 때, 소속 함수는 다음과 같은 과정에 의해 학습이 이루어진다. d 는 학습 매개변수를 의미한다. ($d = 0.005$)

- (단계1) 만약, 입력값 x 에 대해 $x \leq center$ 이면,
 $center' = center + d$, $left' = left + d$
- (단계2) 만약, 입력값 x 에 대해 $x > center$ 이면,
 $center' = center - d$, $right' = right - d$



[그림1] 삼각 소속 함수의 Penalty 연산 적용 (1)

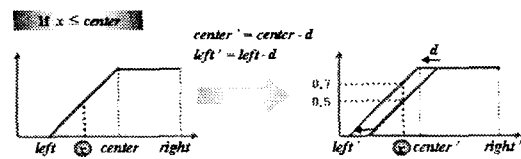


[그림2] 삼각 소속 함수의 Penalty 연산 적용 (2)

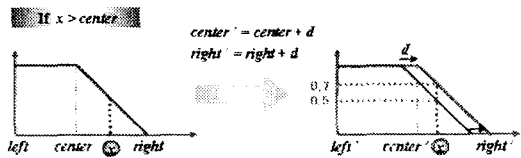
3.1.2 Reward 연산

Reward 연산은 소속 함수의 소속 정도를 높여 주는 연산이다. [그림3], [그림4]와 같이 $left$, $center$, $right$ 세 개의 x 좌표 값을 갖는 사다리꼴 소속 함수가 주어졌을 때, 소속 함수는 다음과 같은 과정에 의해 학습이 이루어진다.

- (단계1) 만약, 입력값 x 에 대해 $x \leq center$ 이면,
 $center' = center - d$, $left' = left - d$
- (단계2) 만약, 입력값 x 에 대해 $x > center$ 이면,
 $center' = center + d$, $right' = right + d$



[그림3] 사다리꼴 소속 함수의 Reward 연산 적용 (1)



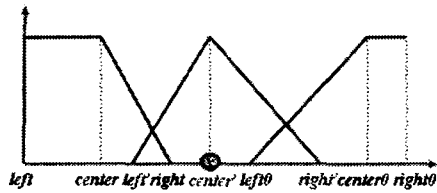
[그림4] 사다리꼴 소속 함수의 Reward 연산 적용 (2)

3.1.3 Coverage 연산

소속 함수에 대한 학습 수행 시, 주어진 입력 데이터를 표현할 수 있는 퍼지 규칙이 존재하지 않는 경우가 발생할 수 있다. [그림5]와 같이 입력 데이터를 표현할 수 있는 소속 함수가 존재하지 않을 경우에 Coverage 연산을 수행한다.

Coverage 연산은 새로운 소속 함수를 생성하는 연산으로, 입력 데이터를 새로운 소속 함수의 중앙 좌표 값으로 하여 삼각 소속 함수를 생성한다. Coverage 연산을 수행함으로써 속성별 소속 함수의 개수를 사용자가 직접 정의해주어야 하는 문제점을 해결할 수 있다. Coverage 연산에 의해 생성되는 삼각 소속 함수의 세 좌표 *left'*, *center'*, *right'* 는 다음과 같이 구한다.

$$- \text{center}' = x, \text{left}' = (\text{center} + \text{right}) / 2, \text{right}' = (\text{left} + \text{center}) / 2$$



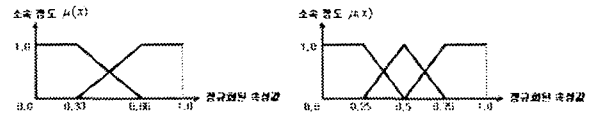
[그림 5] Coverage 연산 적용

3.2 퍼지 결정 트리를 이용한 지식 추출

퍼지 결정 트리(Fuzzy Decision Tree : FDT)는 간결한 규칙을 생성하는 결정 트리 생성 알고리즘인 ID3 방법에 퍼지의 추론과 표현력을 결합하여 트리를 생성할 수 있도록 확장한 방법으로 규칙의 형태, 생성되는 규칙의 수, 생성된 각 규칙의 조건절의 함수가 간결한 것이 특징이다 [3]. 퍼지 결정 트리는 비단말 노드(nonterminal node), 단말 노드(terminal node), 링크(link)로 구성되어 있다. 비단말 노드는 분기를 위한 속성을 가지고 있으며, 단말 노드는 클래스 명과 CF를 가지고 있다. 노드와 노드 사이를 연결하는 링크는 속성에 대한 소속 함수를 나타낸다. 본 논문에서는 적은 수의 이해하기 쉽고 정확도가 높은 퍼지 규칙을 추출하기 위해 PRC Learning에 의해 학습된 소속 함수를 퍼지 결정 트리에 적용한다.

4. 실험 및 결과

본 논문에서 제안한 방법의 타당성을 확인하기 위해 Iris, Wisconsin Breast Cancer, Pima, Bupa 데이터를 사용하여 실험하였다. 입력 데이터는 [0,1]의 값으로 정규화 하였으며, [그림6]과 같이 삼각형, 사다리꼴 소속 함수의 조합을 통해 수치적인 속성을 표현하는 소속 함수를 사용하였다. 또한, Penalty 연산과 Reward 연산을 적용시 소속 함수를 학습하기 위해 필요한 학습 매개 변수(*d*)는 실험을 통하여 0.005 값을 사용하였다.



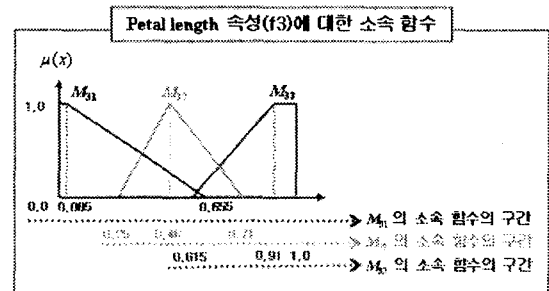
[그림6] 초기 소속 함수가 2개인 경우와 3개인 경우

[표1]은 본 논문에서 제안하는 PRC Learning의 타당성을 검증하기 위해 Iris 데이터를 이용하여 실험한 분류율과 오류 보정 방법[4]에 의한 분류율을 비교한 결과이다. 초기 소속 함수는 2개와 3개를 사용하였으며, 훈련 데이터를 시험 데이터로 사용하여 실험한 경우와 Leaving One Out 방법으로 실험한 경우에 대해 분류율을 비교하였다.

[표1] PRC Learning과 오류 보정방법의 분류율(단위:%)

	Training Data				Test Data			
	Error Correction Method [4]		PRC Learning		Error Correction Method [4]		PRC Learning	
	소속 함수 2개	소속 함수 3개	소속 함수 2개	소속 함수 3개	소속 함수 2개	소속 함수 3개	소속 함수 2개	소속 함수 3개
속성2개	66.7	96.0	95.33	97.33	66.7	95.33	95.33	97.33
속성3개	66.7	97.3	95.33	97.33	66.7	95.33	94.67	96.00
속성4개	69.3	96.7	94.67	97.33	70.7	95.33	93.33	97.33

[표1]을 통해 Iris 데이터에 대하여 PRC Learning 방법이 오류 보정 방법보다 효율적임을 알 수 있었다. [그림7]은 Iris 데이터의 속성 4개를 모두 사용하여 PRC Learning 방법을 적용하였을 때, 꽃잎(petal)의 길이(*f3*) 속성을 표현하는 학습된 소속 함수를 나타낸다.



[그림7] PRC Learning에 의해 학습된 소속 함수

[표2]는 본 논문에서 제안하는 학습된 소속 함수를 퍼지 결정 트리에 적용하는 알고리즘을 Iris, Wisconsin Breast Cancer, Pima, Bupa 데이터에 적용한 실험 결과이다. 초기 소속 함수는 3개를 사용하였으며, 10 fold cross validation으로 실험하였다.

[표2] 학습된 소속함수를 퍼지결정트리에 적용한 실험결과

데이터명	Iris		Wisconsin Breast Cancer		Pima		Bupa	
	분류율 (%)	tree size	분류율 (%)	tree size	분류율 (%)	tree size	분류율 (%)	tree size
Unseen Data	96.20	4.0	95.90	7.0	74.09	10.50	68.20	15.40

[표3]은 Iris, Pima, Bupa 데이터를 이용하여 C4.5, FID 3.1 알고리즘[6]에 적용한 분류율과 tree size를 [표2]에서 실험한 결과와 비교한 것이다. [표3]에서 알 수 있듯이 Iris 데이터에 대해서는 본 논문에서 제안하는 알고리즘이 C4.5, FID 3.1에 비하여 분류율이 높았으며, tree size가 적어 간결하며 이해하기 쉬운 규칙을 생성할 수 있었다. Pima, Bupa 데이터는 C4.5, FID 3.1 알고리즘에 비해 분류율은 조금 낮지만 tree size가 현저하게 줄었다. 즉, 조건부가 짧고 규칙의 수가 적어 간결하며 이해하기 쉬운 퍼지 규칙을 생성할 수 있음을 알 수 있었다.

[표3] C4.5, FID3.1과 제안한 알고리즘의 분류율/tree size 비교

데이터명	C4.5		FID3.1		Proposed Algorithm	
	분류율(%)	tree size	분류율(%)	tree size	분류율(%)	tree size
Iris	94.00	5.0	96.00	5.0	96.20	4.0
Pima	74.70	45.2	75.90	23.1	74.09	10.5
Bupa	67.90	34.4	70.20	28.9	68.20	15.4

다음은 학습된 소속 함수를 퍼지 결정 트리에 적용하는 알고리즘으로 Iris 데이터를 실험하였을 때 생성되는 규칙이다. 생성된 규칙에서 사용된 short, medium, long의 어휘적인 표현은 사용자가 결정한다.

- If petal(f3) is short(0.0,0.06,0.63) Then
setosa(c1) CF:1.0
- If petal(f3) is medium(0.225,0.42,0.7) Then
vesicolor(c2) CF:0.97
- If petal(f3) is long(0.6,0.895,1.0) Then
virginica(c3) CF:0.973

PRC Learning에 의해 학습된 소속 함수를 퍼지 결정 트리에 적용할 때, 퍼지 결정 트리의 매개 변수인 θ_s, θ_d 를 조정함으로써 분류율과 퍼지 규칙의 수를 조절할 수 있다. [표4]는 각 실험 데이터에 대해 퍼지 결정 트리 생성 시 실험에 사용된 매개 변수를 보여준다.

[표4] 퍼지 결정 트리 적용 시 매개 변수 값

데이터명	Iris	Wisconsin Breast Cancer	Pima	Bupa
θ_s	0.4	0.1	0.1	0.1
θ_d	0.9	0.9	0.9	0.9

5. 결론 및 향후과제

수치적인 데이터를 효과적으로 분류하기 위해 본 논문에서는 기울기 강하법을 기반으로 하는 소속 함수 학습 방법(PRC Learning)과 학습된 소속 함수를 기반으로 퍼지 결정 트리를 생성하여 이해하기 쉬우며 정확도가 높은 퍼지 규칙을 생성하는 알고리즘을 제안하였다. 제안하는 알고리즘은 첫째, Penalty 연산과 Reward 연산을 통

해 소속 함수를 학습하며, Coverage 연산을 통해 각 속성에 대한 소속 함수의 개수를 조정한다. 둘째, 학습된 소속 함수를 이용하여 퍼지 결정 트리를 생성하고, 적은 수의 간결하며 이해하기 쉬운 퍼지 규칙을 추출한다. Iris, Wisconsin breast cancer, Pima, Bupa 데이터에 대한 실험 결과와 다른 분류 알고리즘의 실험 결과를 비교해 보았을 때, 제안된 알고리즘에 의해 생성되는 퍼지 규칙은 규칙의 수가 적은은 물론 조건부가 짧고 규칙의 형태가 간결하여 이해하기 쉬운 것을 알 수 있었다. 하지만, 데이터에 따라 다른 알고리즘보다 분류율이 저하되는 문제점이 있었다.

향후 연구로는 제안된 알고리즘을 실제 응용문제에 적용하는 것이며, 대용량의 데이터를 효율적으로 처리할 수 있도록 알고리즘을 개선하는 것이다. 또한 속성수와 각 속성을 표현하는 소속 함수의 수만큼 지수적으로 증가하는 초기 규칙의 수를 줄이기 위한 연구가 필요하며, 분류율이 높으면서 tree size를 줄일 수 있도록 알고리즘을 개선하는 것이 필요하다.

6. 참고문헌

- [1] J. Ross Quinlan. C4.5 : programs for machine Learning. Morgan Kaufmann. 1993.
- [2] Ian H. Witten, Eibe Frank. Data Mining : Practical machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann. 1999.
- [3] M.W. Kim, J.G.Lee and C.W.Min, Efficient Fuzzy Rule Generation Based on Fuzzy Decision Tree for Data Mining, IEEE Int. Fuzzy Systems Conf. Proceeding, 1999,pp.1223-1228
- [4] Nakashima, T. Nakai, G. Ishibuchi, H. Improving the performance of fuzzy classification systems by membership function learning and feature selection. Fuzzy System, 2002. FUZZ-IEEE'02. Pro. of the 2002 IEEE Int.l Conf. on, Vol: 1, 2002. pp488-493
- [5] Ishibuchi, H. Nozaki, K. Tanaka, H. Pattern classification by distributed representation of fuzzy rules. Fuzzy Systems, 1992. IEEE Int. Conf. on, 1992. pp.143-650
- [6] Janikow, C.Z., Fajfer, M., Fuzzy partitioning with FID3.1, Fuzzy Information Processing Society, 1999. NAFIPS. 18th International Conference of the North American, 10-12 June 1999 Page(s): 467-471