

## 규칙 성장 기반 퍼지 분류기의 설계

김옥동\*, 오성권\*, 김현기  
수원대학교\*

### Design of Growing Rule-based Fuzzy Classifier

Wook-Dong Kim\*, Sung-Kwon Oh\* and Hyun-Ki Kim  
The University of Suwon\*

**Abstract** - 본 논문은 퍼지 클러스터링을 이용한 규칙 성장 기반 퍼지 분류기의 설계에 대해서 소개한다. 본 논문의 목적은 퍼지 클러스터링을 통해 형성된 증가된 퍼지 규칙을 이용한 새로운 설계 방법론을 개발하는 것이다. 제안된 분류기는 네개의 기능적인 부분으로 구성된다. 퍼지 규칙의 전반부는 퍼지 클러스터링 알고리즘을 이용해 구성된 멤버쉽 함수를 나타낸다. 후반부는 지역 모델을 구성한다. 지역 모델의 파라미터는 가중 최소 자승법에 의해 추정된다. 추론부에서는, 각 퍼지 규칙의 에러 측정후, 가장 높은 에러를 갖는 하나의 퍼지 규칙이 선택된다. 규칙성장 부분에서는, 네트워크의 강화를 위해 규칙의 성장 과정이 이루어지며, 선택된 규칙은 제안된 분류기에서 더 나은 성능을 위해 두 개 또는 세 개의 세분화된 퍼지 규칙으로 나누어진다. 이러한 새로운 규칙은 context 기반 Fuzzy C-Means 클러스터링에 의해서 형성된다. 제안된 규칙 기반 분류기의 효용성을 토론하며, 머신 러닝 데이터를 이용하여 실험을 수행하였다.

$$f_j(x_1, \dots, x_k) = a_{j0} + \sum_{i=1}^k a_{ji}x_i \quad (2)$$

각 퍼지 규칙 중에서 성능이 나쁜 퍼지 규칙을 선택하여 성장부로 이동한다. 성장부에서는 선택된 퍼지 규칙을 세분화하여 새로운 다수의 퍼지 규칙으로 표현한다. 결과적으로 최종적으로 결정되는 퍼지 규칙 수는 식(3)과 같다.

$$Final\ fuzzy\ rule = (Initial\ fuzzy\ rule - 1) + New\ fuzzy\ rule \quad (3)$$

이 때 적합도는 Context FCM을 이용하여 새로운 적합도를 생성한다. 기존의 FCM에서 구한 소속행렬(적합도)과 Context FCM에서 구한 소속행렬의 합은 항상 1을 만족해야 되며, 식(4)과 같다.

$$\sum_{i=1}^{j-1} u_{ik} + \sum_{q=1}^p t_{qk} + \sum_{i=j+1}^c u_{ik} = 1 \quad (4)$$

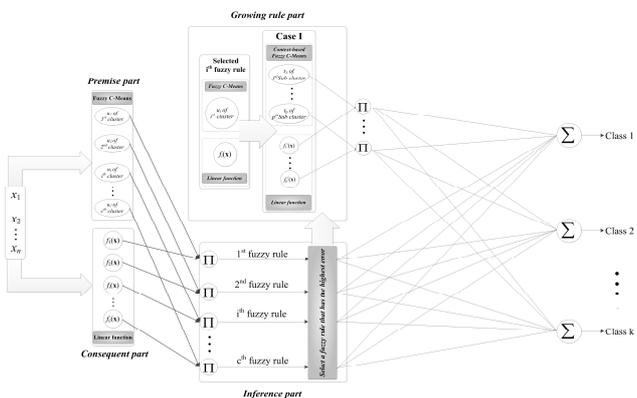
## 1. 서 론

패턴인식을 위해 다양한 형태의 분류기가 제안되고 있으며 그 중에서도 SVM 과 Neural networks가 대표적으로 많이 사용되고 있다. 또한 기존의 제안된 모델의 성능 개선 및 학습 속도 향상 등을 위해 다양한 방향으로 연구가 진행되고 있다. 본 논문에서는 클러스터링 알고리즘을 결합한 새로운 형태의 퍼지 분류기를 설계한다. 일반적으로 퍼지 모델에서 기존 클러스터링의 역할은 멤버쉽 함수의 중심점을 결정에 관여하거나 또는 직접적으로 퍼지수를 정하기 위해 사용되었다. 이와 달리, 본 논문에서 제안한 규칙 성장 기반 퍼지 분류기는 퍼지 클러스터링을 통해 퍼지 규칙의 생성 및 규칙 수 증가를 결정하여 분류기의 구조를 생성한다.

## 2. 본 론

### 2.1 규칙 성장 기반 퍼지 분류기의 구조

제안된 분류기는 전반부, 후반부, 추론부, 성장부와 같이 네 부분으로 정의된다. 전체 적인 규칙 성장 기반 퍼지 분류기의 구조는 그림 1과 같다.



〈그림 1〉 규칙 성장 기반 퍼지 분류기

전반부 및 후반부에 의해 정의되는 퍼지 규칙은 식(1)과 같고, then절은 1차 선형식을 사용하였으며 식 (2)와 같다.

$$R_i : \text{If } \mathbf{x} \text{ is } u_i \text{ with } \mathbf{v}_i \text{ center of cluster then } f_i(\mathbf{x}) \quad (1)$$

### 2.2 규칙 성장 기반 퍼지 분류기의 학습 방법

전반부에서는 Fuzzy C-Means 클러스터링을 이용하여 입력공간을 주어진 클러스터의 수로 그룹화하고 각 그룹의 중심점을 이용하여 각 입력들을 퍼지 소속정도로 표현한다. 후반부에서는 전반부에서 구한 퍼지 소속정도와 선형 함수를 이용하여 퍼지 규칙을 생성한다. 퍼지 규칙의 후반부는 1차 선형 식을 사용하며 함수의 파라미터 계수는 가중 최소자승법(Weighted Least Square Estimation; WLSE)을 사용하여 동조한다. 성장부에서는 후반부에서 정의한 퍼지 규칙들 중 실제 출력과 가장 오차가 큰 퍼지 규칙을 성장부로 보내서 새로운 퍼지 규칙을 생성하며, 여기서는 Context Fuzzy C-Means 클러스터링을 사용하여 전반부에서 구한 퍼지 소속정도를 다시 한번 더 클러스터링을 통해 세분화 한다.

#### 2.2.1 Fuzzy C-Means 클러스터링

FCM 클러스터링은  $m$ 개의 벡터  $x_i(i=1, \dots, n)$  집합을  $c$ 개의 클러스터로 분할하고, 목적함수가 최소가 일 때 생성된 각 클러스터에서 중심 값을 찾는다. FCM 클러스터링 방법은 0과 1사이의 소속정도에 의해서 나타난 소속감의 정도를 가지고, 주어진 데이터들을 몇 개의 클러스터로 나눌 수 있는지를 판별하기 위해 퍼지 분할을 사용한다는 점이다. 즉 퍼지 분할을 적용하기 위해서, 소속함수  $U$ 는 0과 1사이의 값을 가진 요소들로 나타낸다. 그러나 데이터 집합에 대한 소속감 정도의 합은 식(5)과 같이 항상 1이다.

$$\sum_{i=1}^c u_{ik} = 1, \forall k=1, \dots, n \quad (0 < \sum_{i=1}^c u_{ik} < 1) \quad (5)$$

FCM 클러스터링에 대한 목적함수는 다음과 같은 형태를 가지고 있다.

$$J(u_{ik}, \mathbf{v}_i) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m (d_{ik})^2 \quad (6)$$

여기서,  $u_{ik}$ 는 0과 1사이의 수적인 값으로  $i$ 번째 클러스터에 속해져 있는  $x_k$ 의  $k$ 번째 데이터의 소속정도를 나타낸다.  $\mathbf{v}_i$ 는  $i$ 번째 클러스터 중심 벡터이다.  $j=1, 2, \dots, s)$ 는 특성 공간상의 변수이다.  $m$ 는 소속 함수의 애매함의 정도에 대한 영향을 나타내는 퍼지화 계수이다. 이 값은 일반적으로 2로 설정한다.

식(6)의 목적함수를 최소화하기 위해서 먼저  $\mathbf{v}_i$ 와  $u_{ik}$ 와 관련된 식의 목

적합수를 분화시켜야 한다.

$$v_{ij} = \frac{\sum_{k=1}^n (u_{ik})^m x_{ki}}{\sum_{j=1}^n (u_{ik})^m} \quad (7)$$

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}}\right)^{2/m-1}} \quad (8)$$

식 (7) 와 (8)에 의해서 묘사되어진 시스템은 분석적으로 해결할 수 없다. 그러나 FCM 클러스터링 알고리즘은 주어진 위치로부터 시작해서 목적함수의 최소로 접근하기 위해서 반복적으로 접근한다.

### 2.2.2 가중 최소 자승법

본 논문에서는, 후반부 다항식의 파라미터 계수를 가중최소자승법 (Weighted Least Square Estimator; WLSE)을 사용하였으며 성능평가 함수  $Q_w$ 을 행렬형식으로 표현하면 식 (9)과 같다.

$$Q_w = \sum_{j=1}^c (Y - X_j a_j)^T U_j (Y - X_j a_j) \quad (9)$$

여기서,  $n$ 은 추정하고자 하는  $j$ 번째 다항식의 계수,  $Y$ 는 출력데이터,  $U_j$ 는  $j$ 번째 입력공간에 대한 입력 데이터들의 소속 값을 의미한다.  $X_j$ 는  $j$ 번째 로컬모델의 계수를 추정하기 위한 입력데이터 행렬을 의미하며 로컬모델이 선형일 경우 다음처럼 정의된다.

$$X_j = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m1} & \cdots & x_{mk} \end{bmatrix} \quad U_j = \begin{bmatrix} u_{j1} & 0 & \cdots & 0 \\ 0 & u_{j2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{jm} \end{bmatrix}$$

여기서  $m$ 은 데이터의 수이다.  $j$ 번째 규칙에 대한 로컬모델인 다항식의 계수는 식(10)에 의해 구해진다.

$$a_j = (X_j^T U_j X_j)^{-1} X_j^T U_j Y \quad (10)$$

### 2.2.3 Context Fuzzy C-Means 클러스터링

Context FCM 클러스터링은 FCM과 동일과 연산과정을 거친다. 다만 차이가 있다며, FCM의 소속정도의 합은 항상 1이지만 Context FCM의 경우에는 항상 1이 아니다. 식(11)와 같이  $u_k$ 는 context의 값을 의미하여 Context FCM의 소속정도의 합은  $u_k$ 와 같다.

$$\sum_{i=1}^{s_c} t_{ik} = u_k, \forall k=1, \dots, n \quad (0 < \sum_{i=1}^{s_c} t_{ik} < u_k) \quad (11)$$

따라서, context FCM의 소속행렬은 식(13)과 같다. 일반적으로 context의 값은 다양한 방법으로 정의되지만, 본 논문에서는 전반부에 FCM을 통해 얻어진 소속행렬을 context로 사용하였다.

$$s_{vij} = \frac{\sum_{k=1}^n (t_{ik})^m x_{ki}}{\sum_{j=1}^n (t_{ik})^m} \quad (12)$$

$$t_{ik} \equiv \frac{u_k}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}}\right)^{2/m-1}} \quad (13)$$

### 2.3 실험결과 및 고찰

제안된 분류기의 성능 평가를 위해 대표적인 데이터를 사용하였으며 표 1과 같이 전체 데이터의 수 및 입력변수의 수 그리고 클래스의 수를 나타내었다. 또한 통계적인 성능평가를 위해 5 fold cross validation (5fcv)을 사용하였다.

<표 1> 실험에 사용된 데이터 종류

Data sets	Number of data	No. of input variables	No. of classes
Iris	150	4	3
WDBC	569	30	2

식(14)는 분류기의 성능의 척도가 되는 패턴 분류률을 나타낸다.

$$PCR[\%] = \frac{True}{N_{tr \text{ or } te}} \times 100 \quad (14)$$

여기서,  $tr$ 은 학습데이터의 수를 나타내며,  $te$ 는 테스트 데이터의 수를 나타낸다.

표2는 제안된 분류기에서 규칙 성장 전과 후의 패턴 분류률을 나타낸다.

<표 2> 제안된 분류기와 기존 분류기의 성능비교

Iris					
Before growing rules			After growing rules		
Rule	TR[%] (PI)	TE[%] (EPI)	Rule	TR[%] (PI)	TE[%] (EPI)
2	97.00±1.12 (0.1342)	96.00±3.65 (0.169)	-	-	-
3	97.17±0.95 (0.1285)	95.33±1.83 (0.1686)	(2-1)+2= 3	97.17±0.95 (0.1189)	96.00±2.79 (0.1523)
4	97.17±0.95 (0.1075)	96.00±2.79 (0.1521)	(3-1)+2= 4	97.00±1.39 (0.1185)	95.33±1.83 (0.1586)
5	97.33±1.09 (0.1063)	96.67±2.36 (0.1530)	(4-1)+2= 5	97.50±1.02 (0.1030)	95.33±1.83 (0.1468)
6	97.67±1.49 (0.1028)	95.33±2.98 (0.1470)	(5-1)+2= 6	97.33±1.09 (0.1021)	95.33±3.80 (0.1493)
WDBC					
Before growing rules			After growing rules		
Rule	TR[%] (PI)	TE[%] (EPI)	Rule	TR[%] (PI)	TE[%] (EPI)
2	96.79±0.26 (0.1550)	96.47±1.11 (0.026)	-	-	-
3	97.80±0.31 (0.1400)	97.18±1.17 (0.2025)	(2-1)+2= 3	97.80±0.16 (0.1270)	96.30±0.77 (0.1900)
4	97.67±0.19 (0.1245)	96.48±0.63 (0.1973)	(3-1)+2= 4	97.89±0.20 (0.1284)	97.00±0.80 (0.1931)
5	97.71±0.30 (0.1188)	96.66±0.76 (0.2011)	(4-1)+2= 5	98.37±0.25 (0.1115)	96.31±0.41 (0.1953)
6	98.07±0.09 (0.1076)	96.65±1.17 (0.1993)	(5-1)+2= 6	98.20±0.18 (0.1064)	96.66±0.76 (0.1950)

### 3. 결 론

본 논문에서는 규칙성장 기반 퍼지 분류기를 제안하였다. 규칙 성장을 위해 기존의 단일 클러스터링을 사용하여 설계된 네트워크의 퍼지 규칙을 이용한다. 다수의 퍼지 규칙들 중 성능에 기여하는 정도가 부족한 규칙을 선택하여 성장부에서 새로운 규칙으로 세분화하여 초기모델의 구조를 최대한 유지하면서 성능 개선을 개선했다.

### 감사의 글

본 연구는 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음(GRRR 수원2015-B2, U-city 보안감시 기술협력센터) 그리고 한국산업단지공단의 10차년도 산업집적지경쟁력강화산업계획의 생산기술사업화 지원사업으로 연구를 수행하였음(NTIS-1415136442)

### [참 고 문 헌]

- [1] S.K. Oh, W.D. Kim, W. Pedrycz, and B.J. Park, "Polynomial-based radial basis function neural networks(P-RBF NNs) realized with the aid of particle swarm optimization," Fuzzy Sets and Systems, vol. 163, no. 1, pp. 54-77, January 2011.
- [2] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
- [3] W. Pedrycz, "Conditional fuzzy c-means," Pattern Recognition Letters, vol. 17, no. 6, pp. 625-631, May 1996.