지능정보연구 제15권 제4호 2009년 12월(pp.123~140)

ANFIS에서 생성된 규칙의 해석용이성 평가

송희석
한남대학교 경영정보학과 부교수
(hssong@hnu.kr)

김재경
경희대학교 경영대학 & 경영연구원 교수
(jaek@khu.ac.kr)

퍼지신경망 모형은 인공신경망의 네트워크 구조 표현방법 및 학습알고리듬과 퍼지시스템의 추론방법을 통합한 모형으로 제어 및 예측분야에 성공적으로 적용되고 있다. 본 연구에서는 퍼지신경망 모형 중 우수한 예측정확도로 인해 최근 각광받고 있는 ANFIS (Adaptive Network-based Fuzzy Inference System)모형에서 생성된 퍼지규칙의 해석용이성을 평가하였다. ANFIS모형은 인간 전문가와 상호작용하면서 규칙을 정제해 나갈 수 있다. 특히 인간전문가의 사전지식을 이용하여 초기 퍼지규칙을 만들고 난 후 모형을 학습하면 최적에 수렴하는 시간을 단축할 뿐 아니라, 전역 최적치 도달가능성이 높아진다고 보고되고 있다. 이러한 관점에서 볼 때 규칙의 해석용이성은 인간 전문가와의 상호작용을 위해 매우 중요한 이슈가 될 수 있다. 본 연구에서는 ANFIS모형과 의사결정나무 모형에서 생성된 규칙을 해석용이성 관점에서 비교하기 위한 측도를 제안하고 각 규칙들을 비교하였다. 본 연구에서 제안된 해석용이성 측도들은 규칙을 생성하는 다양한 기계학습 모형의 규칙생성 능력을 평가하는 기준으로도 활용될 수 있을 것이다.

논문접수일 : 2009년 11월 17일 게재확정일 : 2009년 12월 09일 교신저자 : 김재경

1. 서론

기계학습 분야의 여러 모형 중 퍼지시스템(Fuzzy System)은 퍼지규칙을 사용함으로써 정확한 계량적인 분석과정 없이 인간 전문가의 정성적인 지식이나 추론과정을 대신할 수 있다는 점에서 각광받고 있다(Jang, 1993; Nauck and Kruse, 1999). 특히 대상시스템이 수리적으로 모형화 하기 어렵고, 비선형성이 내재되어 있는 복잡한 함수의 근사화 또는 제어 및 예측문제에서 퍼지시스템은 성공적으로 적용되어 왔다. 퍼지시스템을 구현하기 위해서는 인간전문가로부터 지식을 추출하여 규칙베이스를 구축할 필요가 있으며, 이러한 지식획득 과정은 퍼지시스템 구축의 장애요인으로 인식되어 왔다. 그러나 인공신경망과 퍼지시스템이 결합된 퍼지신경망의 출현으로 지식획득 과정이 자동화됨으로써 퍼지시스템의 연구가 확대되고 있다(Song and Kim, 2009). 퍼지신경망 모형은 인공신경망의 네트워크 구조 표현방법과 학습알고리듬을 사용하여 퍼지 if-then 규칙을 생성하기 때문에 블랙박스 모형인 신경망모형에서 제공하지 못하는 대상도메인에 대한 지식을 의사결정자에게 제공할 수 있다(Babuška and Verbruggen, 2003). 또한 퍼지 if-then 규칙은 “높다”, “크다”, “따뜻하다” 등 사람이 주로 사용하는 퍼지용어로 기술되기 때문에 인간 전문가가 이해하기 용이하고, 전문

* 이 논문은 2009학년도 한남대학교 학술연구조성비 지원에 의하여 연구되었음.

가 지식과의 통합이 가능하다. 이는 퍼지신경망이 미지의 시스템에 대한 지식을 획득하는 용이한 도구로 사용될 수 있음을 의미 한다(Jin et al., 1999). 그럼에도 불구하고 퍼지신경망 연구는 예측 및 제어 성능 향상 관점에서 주로 이루어졌으며 해당 도메인에 대한 지식획득의 수단으로 활용된 사례는 드물다. 퍼지신경망을 이용하여 도메인에 대한 지식을 용이하게 획득하기 위해서는 퍼지신경망 학습을 통해 생성된 퍼지규칙이 간결하고, 일관성이 있으며(Mutually exclusive), 빠짐없이 대상도메인을 표현(Completely exhaustive)할 수 있어야 한다. 따라서 생성된 퍼지규칙을 간결성, 일관성, 중복성, 완전성 등의 해석용이성(Interpretability) 관점에서 평가하는 것은 지식획득 수단으로써의 퍼지신경망 연구에 중요한 기반이 될 수 있다.

본 연구에서는 퍼지신경망의 일종인 ANFIS를 기반으로 한 분류모형을 설계하고 이를 기존의 의사결정나무분석모형과 비교하여 해석용이성을 평가한다. ANFIS(Adaptive Network-based Fuzzy Inference System)모형은 Jang(1993)에 의해 제안된 모형으로 주어진 입력과 출력데이터를 사용하여 최소자승법(Least Square Method)과 결합된 역전파학습(Back Propagation)을 통해 입력과 출력 데이터의 함수 근사를 가능하게 하는 최적의 네트워크 파라미터를 발견한다. ANFIS 모형은 탁월한 함수근사능력을 갖추어 최근 시계열 예측 및 시스템컨트롤 분야에 폭넓게 적용되고 있는 모형이다. 이와 같은 폭넓은 활용에도 불구하고 ANFIS 모형에서 생성된 퍼지규칙의 해석용이성에 초점을 맞춘 연구는 찾아볼 수 없다. 따라서 ANFIS에서 생성된 규칙이 기존의 분류기법에서 생성된 규칙과 비교하여 해석용이성 측면에서 어떠한지를 구체적인 측도를 제시하고 측정하는 연구가 필요하다. 이러한 연구는 기존의 분류기법과 비교하여 퍼지신경망 모형에서 생성된 규칙의 해석용이성 수준을 평가하는데 만 사용될 뿐 아니라, 규칙을 생성하는 다양한 기계학습 모형의 규칙생성 능력을 평가하는 기준으로도 활용될 수 있을 것이다.

논문의 구성은 다음과 같다. 먼저 제 2장에서 지식획득 수단으로 퍼지신경망을 활용한 연구들을 소개하고 해석용이성에 대한 기존 연구를 요약한다. 제 3장에서 퍼지규칙 생성을 위한 ANFIS 모형을 설계하며 제 4장에서는 해석용이성을 평가하기 위한 측도를 제시한다. 제 5장에서는 사례연구를 통해 ANFIS 모형에서 생성된 퍼지규칙의 해석용이성을 평가하고 의사결정나무에서 생성된 규칙과 비교하며 마지막으로 제 6장에서 연구의 한계와 향후연구 내용을 기술한다.

2. 관련 연구

블랙박스모형인 신경망모형과는 달리 퍼지신경망에서는 If-Then 퍼지규칙을 생성해주며 신경망의 학습 알고리듬을 적용하여 퍼지 규칙에 사용된 파라미터를 자동으로 조정해 줌으로써 예측성능을 최적화 한다. 이러한 장점으로 인해 퍼지신경망은 제어, 패턴인식, 고장진단 등의 분야에 폭넓게 사용되고 있다. 본 장에서는 퍼지시스템을 예측 또는 제어 정확도관점 보다는 도메인에 대한 지식획득관점에서 다루고 있는 연구들을 소개하기로 한다. 도메인에 대한 지식획득은 일반적으로 퍼지신경망 학습을 통해 생성된 퍼지규칙을 해석함으로써 이루어진다. Mikut et al.(2005)는 해석용이성을 “사람이 퍼지 규칙베이스를 조사함으로써 퍼지시스템의 행위를 이해할 수 있는 정도”로 정의하고 있으며, 퍼지신경망이 다른 통계모형이나 인공신경망 등의 대안적인 모델과 비교하여 가지는 가장 중요한 장점으로 해석용이성을 지적 하였다. 기존

연구(Bodenhofer and Bauer, 2002; Cordon and Herrera, 2000; Jin et al., 1998)를 조사해보면 해석용이성이 높은 퍼지규칙의 특성을 다음과 같이 다섯 가지로 요약할 수 있다. 첫째, 직관적인 이해가 가능하도록 규칙의 숫자가 충분히 작아야 한다. 둘째, 규칙에 일관성이 있어야 한다. 이는 규칙의 조건부가 유사하다면 결론부도 유사해야함을 의미한다. 셋째, 규칙의 조건부에는 작은 수의 입력변수만이 포함되도록 하여 그 구조가 간단해야 한다. 넷째, 퍼지용어가 직관적으로 이해하기 쉬워야 하며, 해당 용어의 소속함수 형태가 언어적 표현과 일치하여야 한다. 마지막으로 추론 메카니즘이 정확한 결과를 유도해 낼 수 있도록 퍼지 규칙이 생성되어야 한다.

한편, 해석용이성이 높은 퍼지규칙을 생성하기 위한 연구가 다양한 관점에서 시도 되었다. 해석용이성은 정확성과 상충관계에 있기 때문에 이 연구들은 정확성을 일정수준으로 유지하면서 해석용이성을 높이도록 시도하고 있다. 해석용이성이 높은 퍼지규칙을 생성하는 연구는 퍼지신경망의 학습알고리듬을 새로 개발하여 퍼지규칙 생성단계에서 해석용이성을 높이고자하는 접근법과 생성된 퍼지 규칙에 대해 사후적인 처리를 통해 해석용이성을 높이고자하는 접근법의 두 가지 접근법으로 구분된다. 첫 번째 접근법을 이용하여 해석용이성을 높이는 방법으로는 학습알고리듬에 소속함수와 파라미터에 대한 제약사항을 포함하여 학습 시 소속함수와 파라미터가 일정한 범위 내에서만 조정되도록 하는 방법(Bodenhofer and Bauer, 2002; Valente, 1999)과 퍼지규칙의 구조를 사전에 정해두어 복잡한 구조의 퍼지규칙이 생성되지 못하도록 하는 방법(Cordon and Herrera, 2000) 등이 있다. 사후적인 처리를 통해 퍼지 규칙의 해석용이성을 높이는 방법으로는 유사한 퍼지 규칙들을 발견하여 통합하거나 모순이 되는 퍼지규칙들을 발견하여 이를 해결하는 방식(Chen and Linkens, 2004; Jin, 2000)이 있다. 이들 연구에서는 제시한 방법의 성능 평가를 목적으로 해석용이성을 제 각기 정의하여 측정하고 있기 때문에 연구마다 다른 해석용이성 평가기준을 사용하고 있어 어떤 방법을 적용하여 생성한 퍼지규칙이 해석용이성이 높은지에 대한 상호 비교가 어렵다.

본 연구에서는 퍼지신경망 모형 중 제어 및 예측분야에 성공적으로 적용한 사례가 가장 많은 ANFIS 모형의 지식획득 능력을 평가하고자 한다. ANFIS 모형에서 생성된 퍼지규칙의 해석용이성 수준을 평가한 연구는 아직 발견되지 않고 있다. 이를 위해 분류문제를 대상으로 ANFIS 모형을 구축하고(Song and Kim, 2009) 간결성, 완전성, 중복성 등 다양한 관점에서 해석용이성 평가 측도를 정의하여 의사결정나무기법에서 얻어진 규칙과 비교하여 지식획득 수단으로써의 퍼지규칙의 장단점을 파악하고자 한다.

3. ANFIS 모형 설계

3.1 ANFIS 네트워크 구조

퍼지로직과 신경망이 결합되는 방식에 따라 퍼지신경망은 크게 두 가지 형태로 나누어진다. 첫 번째 형태는 신경망이 퍼지논리나 퍼지연산을 통합하는 방법이고 두 번째 형태는 퍼지추론에 기반한 신경망을 구현하는 방법이다(Han et al., 2008). 이중 퍼지추론에 기반 한 신경망의 구현이 더 폭넓게 연구되고 있으며 이는 다시 Mamdani 모형과 Takagi-Sugeno 모형으로 구분된다. Mamdani 모형은 규칙의 결과부가 퍼지집합인 모형이고 Takagi-Sugeno 모형은 규칙의 결과부가 입력변수의 일차

선형함수로 주어지는 모형이다. Takagi-Sugeno 모형의 퍼지추론 방법은 출력값이 일차선형함수 또는 상수라는 점에서 맘다니 퍼지추론과 디퍼지화 방법이 다르다(Takagi and Sugeno, 1983). Takagi-Sugeno 퍼지추론 방법은 계산 면에서 효율적일 뿐 아니라 인공신경망 등의 최적화 기법과 결합하여 규칙을 추출하는 적응형 기법에 유리하며, 출력 표면의 연속성이 보장되는 등의 장점 때문에 ANFIS에서는 이 모형을 채택하여 규칙을 표현하고 추론 방식을 설계하고 있다.

퍼지추론 과정은 입력변수로 구성된 다차원의 입력공간을 세부영역으로 구분하여 각 세부영역에서 추론의 결과 값을 산출하도록 하는 "분할 후 정복(Divide and Conquer)" 방식으로 이루어진다. 즉, 퍼지규칙의 조건부는 입력공간을 세분영역으로 분할하게 되며, 퍼지규칙의 결론부는 각 영역에서의 추론결과를 기술하게 된다. 따라서 퍼지규칙의 구조를 설계한다는 것은 입력공간을 어떻게 분할하는가와 밀접히 관계되어 있다. ANFIS에서 입력변수들로 구성된 다차원공간을 분할하는 방법으로 그리드분할(Grid Partitioning)과 차감클러스터링(Subtractive Clustering)방법을 주로 사용한다. 그리드 분할 (Grid partitioning)은 입력공간을 중복이 생기지 않도록 격자구조로 공간을 분할하는 것이다. 일반적으로 그리드 분할법을 적용하면 분할된 퍼지영역이 동질적으로 존재하기 때문에 퍼지규칙의 해석용이성(Interpretability)이 높아진다(Bersini and Bontempi, 1997). 이 방법은 입력변수의 수가 작을 때(입력공간의 차원이 낮을 때) 주로 사용되는 방법이다. 예를 들면, 입력변수가 10개이고, 각 입력변수가 2개의 소속함수로 구분이 되는 경우, $2^{10}=1024$개의 세분영역이 생기게 되며, 각 영역마다 1개의 규칙씩 총 1024개의 퍼지규칙이 만들어지게 되어 매우 복잡한 시스템이 된다. 이 경우 다른 공간분할 방법을 사용하는 것이 효과적이다. 한편, 차감 클러스터링(Subtractive Clustering) 기법은 다차원의 입력데이터를 분석하여 이를 군집화 함으로써 입력공간을 몇 개의 세분된 퍼지공간으로 구분하는 방법이다. 그리드 분할의 경우, 입력변수가 증가함에 따라 규칙이 기하급수적으로 증가하는데 이러한 단점을 해결하기 위해 차감 클러스터링이 사용된다. 또한 차감 클러스터링은 퍼지 군집의 개수를 지정하지 않더라도 입력데이터의 공간을 적절한 갯수로 분할한다(Chiu, 1994). 차감 클러스터링에서는 군집의 반경을 지정 하게 되는데 이는 데이터 공간을 단위 하이퍼큐브로 간주할 때 군집 중심으로 부터의 영향범위를 나타내며 0에서 1사이의 값을 가진다. 너무 작은 군집반경을 지정하면 작은 군집이 많이 생기게 되고 따라서 퍼지규칙의 수가 많아진다. 반대로 군집반경을 크게 설정하면 군집의 수가 작아져서 규칙의 수도 작아지게 된다.

입력공간의 분할이 이루어지고 나면 소속함수의 형태를 정의해야 한다. 소속함수의 형태로 주로 사용되는 것은 삼각형(Triangular)함수, 사다리꼴(Trapezoidal)함수, 가우시안(Gaussian)함수, 종형(bell)함수, 시그모이달(sigmoidal)함수 등이 있다. 이중 가우시안함수와 종형함수는 부드러운 곡선 형태를 취하며 작은 수의 파라미터와 간략한 표현법에 의해 퍼지집합을 묘사하기 때문에 가장 폭넓게 사용되고 있다. 한편 시그모이달함수는 오른편과 왼편이 개방된 형태의 함수로써 "매우크다", "매우높다" 등의 극단적인 퍼지집합을 표현하는데 자주 사용된다. 두 개의 입력과 한 개의 출력을 가지며, 일차선형함수를 출력 값으로 가지는 두 개의 Takagi-Sugeno 규칙으로 구성된 간단한 모형을 예로 들어 퍼지추론 과정을 설명하면 다음과 같다(Efendigil et al., 2009, Jang, 1993). 여기서 A_1과

B_1은 퍼지집합을 나타내며, f_1과 f_2는 출력함수를, p_1, q_1, r_1, p_2, q_2, r_2 는 출력함수의 파라미터이다.

규칙 1 : 만일 x가 A_1이고,
y가 B_1이면, $f_1 = p_1x + q_1y + r_1$
규칙 2 : 만일 x가 A_2이고, y가 B_2이면,
$f_2 = p_2x + q_2y + r_2$

먼저 입력변수 x와 y값이 입력되면 소속함수를 호출하여 소속도를 결정한다. 입력변수가 두 개이므로 두 개의 소속도가 min연산에 의해 결합되어 규칙의 조건부의 결합소속도가 결정된다. 결합소속도는 해당규칙의 적합도라 볼 수 있으며, 이를 가중치로 하여 각 결론부의 함수 값을 가중평균함으로써 최종 출력값을 산출한다. 이러한 ANFIS의 추론방식을 분류문제에 적용하기 위해서는 연속형의 추론 결과 값이 특정 계급값으로 도출되어야 하므로 디퍼지화 단계의 마지막 부분에 계급값을 할당하는 과정을 추가한다.

ANFIS 기반 분류모형은 다섯계층의 전방향 연결 신경망 구조로 되어 있다(Song and Kim, 2009). ANFIS는 주어진 훈련데이터에 대한 학습이 각 계층별 노드함수의 파라미터를 조정함으로써 이루어지기 때문에 적응형 네트워크(Adaptive network)라 불려진다. ANFIS에서 각 계층별 노드는 서로 다른 노드함수를 가지며, 이 노드함수는 학습과정을 통해 조정이 되는 파라미터를 가진다. 그러나 노드를 연결하는 링크는 노드 간 흐름의 방향을 나타낼 뿐 어떠한 가중치나 파라미터도 가지지 않는다. 각 계층별 구조와 노드함수를 설명하면 다음과 같다.

계층 1 : 계층 1의 각 노드 i는 노드 함수 $O_i^1 = \mu_{A_i}(x)$, $i = 1, 2$를 가진다. 여기서 x는 노드 i의 입력이며 A_i는 이 노드함수와 연관된 퍼지집합이다. O_i^1는 입력 x가 퍼지집합 A_i에 소속된 정도를 나타내는 소속함수이다. 여기서 소속함수는 여러 가지 형태로 정의될 수 있으며, 종형함수를 예로 들면 다음과 같은 함수로 표현된다.

$$\mu_{A_i}(\chi) = \frac{1}{1 + \left[\left(\frac{\chi - c_i}{a_i}\right)^2\right]^{b_i}},\ a_i, b_i, c_i\text{는 소속함수의 파라미터} \quad (1)$$

종형 함수의 각 파라미터 값이 달라짐에 따라 소속함수의 모양이 달라지며, 학습과정을 통해 오차를 최소로 하는 파라미터 값을 조정해 나가게 된다. 소속함수의 파라미터는 퍼지규칙의 조건부에 나타나므로 조건부 파라미터라고도 부른다.

계층 2 : 이 계층에 있는 노드들은 퍼지 규칙의 조건부에 나타난 각 퍼지집합의 소속도(계층 1에서 산출된)를 입력받아 T-norm 연산을 수행하게 되는데 본 연구에서는 각 소속도를 곱하여 출력을 하게 되며, 노드 출력은 다음 식으로 표현된다.

$$O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y),\ i = 1, 2. \quad (2)$$

이 노드의 출력은 퍼지 규칙의 적합도를 나타낸다.

계층 3 : 이 계층의 각 노드에서는 i번째 규칙의 적합도를 정규화 하게 되며 다음과 같은 연산을 수행한다.

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2},\ i = 1, 2. \quad (3)$$

계층 4 : 이 계층의 각 노드에서는 각 규칙의 결

론부의 출력함수에 표준화된 적합도를 곱하는 연산을 수행하게 되며 계산식은 다음과 같다.

$$O_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i),\ i = 1, 2. \quad (4)$$

여기서 $\overline{w_i}$는 계층 3의 출력이며, 출력함수의 파라미터인 p_i, q_i, r_i는 결론부 파라미터라고 부르며, 학습과정을 통해 오차를 최소화 하도록 그 값이 결정된다.

계층 5 : 이 계층의 노드는 단일노드로 구성되며, 하위 계층의 모든 입력값을 바탕으로 다음과 같은 출력값을 계산한다. 출력값은 퍼지집합이 아닌 연속형값을 가지게 된다.

$$O_i^5 = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (5)$$

3.2 ANFIS 학습

ANFIS의 학습과정은 주어진 입력과 출력데이터를 사용하여 출력오차를 최소화하도록 규칙의 조건부 및 결론부 파라미터를 조정해 나가는 것이다. 출력오차 최소화를 위한 기본 학습로직은 다음과 같다(Jang, 1993). 주어진 적응형 네트워크가 L개의 계층으로 구성되어 있고 k번째 계층이 $N(k)$개의 노드를 가진다고 가정하고, k번째 계층의 i번째 노드를(k, i), 그리고 이 노드의 노드함수를 O_i^k로 표현하면, 노드출력은 식 (6)과 같이 노드입력 신호와 파라미터의 함수로 나타난다.

$$O_i^k = O_i^k(O_i^{k-1}, \cdots, O_{N(k-1)}^{k-1}, a, b, c, \cdots) \quad (6)$$

주어진 훈련데이터 세트가 P개의 사례를 가진다면, p번째 훈련데이터 사례에 대한 최종 출력오차는 식 (7)와 같은 오차자승합으로 표현된다.

$$E_p = \sum_{m=1}^{N(L)} (T_{m,p} - O_{m,p}^L)^2 \quad (7)$$

식 (7)에서 $T_{m,p}$는 p번째 목표출력 벡터의 m번째 요소이고, $O_{m,p}^L$는 p번째 입력벡터에 따라 산출된 출력벡터의 m번째 요소이다. 따라서 훈련데이터에 대한 전체 오차는 $E = \sum_{p=1}^{P} E_p$로 표시된다. 오차최소화를 위한 경사하강법 기반의 학습절차를 개발하기 위하여 p번째 입력벡터에 대해 각 출력노드에서의 오차율을 계산할 필요가 있다. 출력노드 (L, i)에서의 오차율은 식 (7)로부터 다음과 같이 도출된다.

$$\frac{\delta E_p}{\delta O_{i,p}^L} = -2(T_{i,p} - O_{i,p}^L) \quad (8)$$

또한, 적응형 네트워크의 각 내부노드인 (k, i)에서의 오차율은 체인규칙에 의해 식 (9)와 같이 도출된다.

$$\frac{\delta E_p}{\delta O_{i,p}^k} = \sum_{m=1}^{N(k+1)} \frac{\delta E_p}{\delta O_{m,p}^{k+1}} \frac{\delta O_{m,p}^{k+1}}{\delta O_{i,p}^k}, \quad 1 \le k \le L-1 \quad (9)$$

즉, 내부노드에서의 오차율은 다음 계층에 있는 노드들의 오차율의 선형결합으로 표현될 수 있다. 만일 적응형 네트워크의 파라미터를 α라 한다면, p번째 훈련데이터 사례의 오차를 최소로 하는 파라미터 조정식은 식 (10)로부터 도출될 수 있다.

$$\frac{\delta E_p}{\delta \alpha} = \sum_{O^* \in S} \frac{\delta E_p}{\delta O^*} \frac{\delta O^*}{\delta \alpha},$$

S는 파라미터 α를 포함하는 노드집합 (10)

따라서 파라미터 α에 관한 전체 오차는 식 (11)과 같이 계산된다.

$$\frac{\delta E}{\delta \alpha} = \sum_{p=1}^{P} \frac{\delta E_p}{\delta \alpha} \quad (11)$$

결국 파라미터 α에 관한 조정식은 식 (12)와 같다. 이 식에서 η는 학습률이며, k는 스텝사이즈이다. 스텝사이즈를 조절함으로써 최적해 수렴의 속도를 조절할 수 있다.

$$\Delta \alpha = -\eta \frac{\delta E}{\delta \alpha}, \quad \eta = \frac{k}{\sqrt{\sum_{\alpha} (\frac{\delta E}{\delta \alpha})^2}} \quad (12)$$

ANFIS의 학습 방법은 배치학습과 온라인학습으로 구분될 수 있는데, 배치학습은 매 반복(epoch)이 끝난 후 식 (11)에 의해서 파라미터 α에 관한 조정이 이루어지는 방법이며, 온라인 학습은 식 (10)에 의해서 훈련데이터의 사례가 제시될 때 마다 파라미터 α에 관한 조정이 이루어지는 학습방법이다. 이러한 방식으로 적응형 네트워크에서 경사하강방식을 적용하여 파라미터에 대한 조정을 수행할 수 있으나, 이 방법은 수행속도가 느리고 지역최소화 해에 수렴할 가능성이 존재한다. 따라서 Jang(1993)은 경사하강법과 최소자승법(LSE)을 결합하여 파라미터를 조정해 나가는 혼합학습 알고리듬을 제시하였다. 혼합학습 알고리듬은 전방향 학습과 역방향 학습으로 구성되는데, 전방향 학습에서는 조건부 파라미터를 고정시키고 결론부 파라미터를 최적화하기 위해 최소자승법(Least-squares Method)을 적용한다. 최적의 결론부 파라미터가 발견되면 역방향 학습이 즉시 수행된다. 역방향 학습에서는 소속함수를 정의하는 규칙의 조건부에 있는 파라미터를 조정하기 위해 경사하강(Gradient Descent) 기법을 채택하고 있다. 이외에도 Gauss-Newton 또는 Levenberg-Marquardt 학습방법 등이 ANFIS에서 사용가능하다(Matlab, 2009). 학습은 주어진 오차목표(Error Tolerance; 본 연구에서는 0으로 지정)에 도달하거나 주어진 반복횟수에 도달할 때까지 진행된다. 학습률은 처음에는 증가하다가 오차가 증감을 반복하게 되면 감소하도록 설정하는 것이 일반적이다.

3.3 ANFIS 모형에 의한 분류추론

한편, ANFIS 모형은 최종 결과 값이 계급값이 아닌 연속형 변수 값을 취하게 되므로, 이 모형을 분류문제에 적용하기 위해서는 산출된 결과 값을 이용하여 적절한 계급값을 할당하는 과정이 필요하다(Song and Kim, 2009). 본 연구에서는 의사결정나무기법을 이용하여 계급 값을 할당하는 방식과 군집분석을 이용하여 계급 값을 할당하는 두 가지 방식을 이용하여 그 성능을 각 각 평가한다. 먼저 의사결정나무분석기법을 이용하여 계급 값을 할당하는 방식은 훈련 데이터 세트로부터 분류규칙을 추출하고 이를 검증용 데이터에 적용하여 최종 계급값을 할당하는 방법이다. 이 방법은 검증데이터의 분류 정확성을 최대로 하는 분류규칙은 훈련데이터의 분류정확도를 최대로 하는 분류 규칙과 유사할 것이라는 가정 하에서 만들어진 방법이다. 의사결정나무분석에 의한 계급값 할당 방식은 훈련용 데이터의 대표성이 부족하거나 목표변수에 대한 입력변수의 설명력이 부족한 경우, 분류

정확도를 저하시킬 우려가 있다. 따라서 훈련용 데이터를 이용하지 않고 검증용 데이터만으로 계급값을 할당하는 방식을 고려해 볼 수 있다. 군집분석은 비 지도학습이므로 검증용 데이터만을 이용하여 계급값을 할당하는 것이 가능하다. 즉, 군집의 수를 계급값의 수로 지정하여 두고 군집분석을 실시한 후 각 군집에 적절한 계급값을 부여하는 방식이다.

4. 해석용이성 평가를 위한 측도

일반적으로 규칙의 해석용이성을 높이고자 하면 예측 및 제어 성능은 저하된다. 따라서 본 연구에서는 정확성을 최대로 하는 분류모형을 구축하고, 이 모형에서 생성된 규칙을 기준으로 해석용이성을 평가하는 보수적인 접근법을 활용하고자 한다. Jin (1999)은 퍼지규칙의 해석용이성을 제고하기 위해서는 완전성(Completeness), 일관성(Consistency), 간결성(Compactness)을 갖추어야 한다고 주장하였다. 그의 연구에서 완전성은 “모든 분할된 퍼지공간에 매핑되는 퍼지규칙이 적어도 하나 이상 존재”하는 퍼지공간 상의 완전성을 의미하고 있다. 또한 일관성은 생성된 퍼지규칙 간에 서로 상충되는 경우가 발생되지 않아야 함을 의미한다. 이는 규칙의 구문적 측면에서 체크할 수 있는데 두 규칙의 조건부는 유사한데 결론부가 다르다면 일관성에 문제가 있다는 것이다. 마지막으로 간결성은 규칙의 수가 작고 작은 수의 퍼지집합으로 구성되어 있을수록 해석이 용이한 규칙이라고 하였다. 그러나 이외에도 내용이 유사한 규칙이 여러 개 있다면 규칙의 중복이 발생하여 규칙에 대한 가독성을 저하시킬 수 있다. 또한 Jin이 제시한 일관성의 개념도 결국은 유사한 두 규칙 간에 구문적 또는 의미적인 상충관계의 해소라 본다면 일관성도 규칙의 조건부의 중복으로 기인한 것으로 볼 수 있다. 따라서 본 연구에서는 완전성과 (규칙 조건부의) 중복성, 간결성을 해석용이성을 평가하기 위한 차원으로 보고 이를 측정하기 위한 측도를 제시하기로 한다.

본 연구에서 완전성은 데이터로부터 자동 생성된 규칙이 모든 검증용 데이터를 설명할 수 있는 정도로 정의한다. 예를 들면, 검증용 데이터내의 각 사례들이 모두 최소 한 개 이상의 규칙과 매치될 수 있다면 1의 완전성을 가지고, 모든 사례가 어떤 규칙과도 매치되지 않는다면 0의 완전성을 갖는 것으로 정의한다. 아래 식은 완전성을 측정하기 위한 식이다. 완전성은 검증용 데이터 세트의 전체 사례 중 경계값(t) 이상의 적합도를 갖는 규칙이 존재하는 사례의 비율로 측정한다.

$$CP=\frac{\sum_j \nu_j}{N},\quad \nu_j=\begin{cases}1 & \text{if } \prod_i \mu_{A_i}(x_{ij}) \geq t \text{인 규칙이 최소 한 개 존재}\\ 0 & else\end{cases}$$
$$where\ x_{ij}: j\text{번째 사례의 } i\text{번째 입력변수}$$
$$N: \text{총 사례의 수}$$
$$M: \text{입력변수의 수}$$
$$t: \text{경계 값(0과 1 사이)} \qquad (13)$$

평가대상의 규칙이 퍼지규칙이 아니라 의사결정나무에서 생성된 규칙과 같은 비 퍼지규칙의 경우, $\mu_{A_i}(x_{ij})$는 1 또는 0의 값만을 취하게 된다. 한편, ANFIS에서 생성된 퍼지 규칙은 조건부에 퍼지집합을 포함하고 있기 때문에 본연적으로 규칙간 조건부의 중복이 발생할 가능성이 높다. 본 연구에서는 생성된 규칙의 중복성을 정의하고 해석용이성 평가의 측도로 활용하고자 한다. 중복성은 해당분야의 전문가가 생성된 규칙을 확인하여 판단할 수도 있으나, 전문가의 주관이 개입될 수 있으므로 규칙의 구문적 및 의미적 측면에서 “규칙의 조건부가 유사한 정도”로 정의하는 것이 바람

직하다. 규칙의 조건부가 유사하고 결론부가 다르면 두 규칙 간 일관성에 문제가 있음을 의미하며, 규칙의 조건부가 유사하고 결론부도 유사하면 두 규칙이 중복규칙이 되어 해석용이성이 저하된다. 따라서 규칙의 조건부의 중복성을 해석용이성의 평가 측도로 정의한다. 중복성 측도를 정의하기 위하여 다음과 같은 두 개의 임의의 규칙을 고려해 보자.

R_i : 만일 x_1이 A_{i1}이고, x_2 가 A_{i2}이고, …, x_n 이 A_{in}이면, $y = c_i$

R_k : 만일 x_1이 A_{k1}이고, x_2가 A_{k2}이고, …, x_n 이 A_{kn}이면, $y = c_k$

임의의 두 규칙 간 조건부의 유사성을 측정하기 위해, 우선 퍼지집합의 유사성을 다음과 같이 소속함수를 이용하여 정의한다. i번째 규칙의 j번째 퍼지집합과 k번째 규칙의 j번째 퍼지집합의 유사성 $S(A_{ij}, A_{kj})$은 다음 식으로 계산할 수 있다.

$$S(A_{ij},\ A_{kj}) = \frac{\mu_{A_{ij}}(x_j) \cap \mu_{A_{kj}}(x_j)}{\mu_{A_{ij}}(x_j) \cup \mu_{A_{kj}}(x_j)},$$

A_{ij} : i번째 규칙의 j퍼지집합,

$\mu_{A_{ij}}$: A_{ij}의 소속함수 (14)

퍼지집합 간 유사도는 두 퍼지집합의 소속함수가 동일할 때 1이 되고, 소속함수간 겹침이 없이 완전히 다르면 0이 된다. 이 유사성 측도는 퍼지집합이 아닌 비 퍼지규칙에도 적용가능하며, 이때의 소속함수는 모든 x_j 값에 대해 0 또는 1의 값을 가진다. 퍼지집합 간 유사도를 이용하여 두 규칙 i와 k의 조건부 유사도$SP(i, k)$는 다음과 같이 측정할 수 있다.

$$SP(i, k) = \min_{j=1\cdots n} S(A_{ij},\ A_{kj}) \quad (15)$$

따라서 생성된 전체 규칙 집합에 대한 중복성(RD)은 식 (16)으로 측정할 수 있다. 중복성 측도는 0에서 1까지의 값을 가지며, 높은 값일수록 규칙 간 중복이 많이 발생함을 의미한다. 아래 식에서 분모는 규칙간 조건부의 중복성을 비교하기 위해 규칙집합에서 임의의 두 규칙을 선택하는 모든 경우의 수를 나타내며, 분자는 임의의 두 규칙 간 조건부의 유사도의 총합을 나타낸다.

$$RD = \frac{\sum_{i,k \in R, i \neq k} SP(i, k)}{s(s-1)/2},$$

s는 생성된 규칙의 수 (16)

도메인에 대한 이해를 위해 가장 중요한 개념 중 하나가 간결성이다. 간결성은 규칙이 직관적으로 이해하기 쉽도록 간단히 기술된 정도이다. 간결성은 규칙의 수나 소속함수의 수로도 표현할 수 있지만, 규칙의 구조나 의미 있는 퍼지용어의 사용여부 등이 직관적인 이해를 높이는데 중요한 요인이 되므로 별도의 측도를 정의하기보다는 정성적으로 평가하도록 한다.

5. 사례연구

ANFIS 분류모형에서 생성된 규칙의 해석용이성을 평가하기 위해 입력변수가 연속형 값을 취하는 아이리스(Iris) 데이터 세트를 활용하여 사례연구를 수행하였다. 아이리스 데이터 세트는 UCI 기계학습 리파지토리에 등록되어 있는 데이터 세트로써(Asuncion and Newman, 2007) 패턴인식 또는 분류문제에 가장 널리 알려진 데이터 세트이다.

<표 1> 아이리스 데이터의 모형별 분류정확도

ANFIS 모형	반복 횟수	훈련용 RMSE	검증용 RMSE	계급값할당방법	검증데이터 정분류율
SubC(0.6)	20	0.0743	0.0890	의사결정나무	92%
				군집	96%
GriP(2, 2, 2, 2) 종형/선형	90	0.0149	0.1282	의사결정나무	92%
				군집	92%

아이리스 데이터는 세 가지의 아이리스 품종을 계급값으로 가지며, 각 계급값 별로 50개씩 총 150개의 레코드로 구성되어 있다. 입력변수는 꽃받침 길이(SL ; Sepal length), 꽃받침 넓이(SW ; Sepal Width), 꽃잎 길이(PL ; Petal Length), 꽃잎 넓이(PW; Petal Width)의 네 가지 특성으로 구성되어 있다. 전체 데이터 중 훈련용 데이터가 67%, 검증용 데이터가 33%가 되도록 사례를 랜덤하게 추출하여 실험을 진행하였다. 학습은 검증용 데이터의 오차(RMSE ; Root Mean Square Error)를 최소로 할 때 까지 반복 수행하였다. ANFIS 기반 분류모형의 비교대상으로는 현재 분류문제에 가장 폭넓게 적용되고 있는 C5.0을 사용하였다. ANFIS 모형의 설계와 수행을 위해서는 MATLAB 7.6을 사용하였고, 의사결정나무분석과 군집분석을 수행하기 위해서는 클레멘타인 11.0을 활용하였다.

먼저 의사결정나무분석 기법 중 하나인 C5.0을 이용하여 생성된 규칙의 해석용이성을 비교하기 위해 훈련 데이터만을 사용하여 분류 정확도를 최대로 하는 의사결정나무 모형을 구축하였다. 이때, 글로벌 가치치기 방법을 선택하였고, 분기를 위한 최소 레코드 수는 2로 지정하였다. 이렇게 구축된 C5.0 모형을 검증용 데이터에 적용한 결과 정분류율은 92%, 오분류율은 8%로 나타났다. 한편, ANFIS 모형에서 생성된 규칙의 해석용이성을 평가하기 위해 차감클러스터링 모형과 그리드분할 모형 각각에서 검증용 데이터의 분류정확도가 가장 높은 두 개의 모형을 선정하였다. 선정된 두 모형은 차감클러스터링으로 입력공간을 분할하고 영향반경은 0.6으로 설정한 SubC(0.6) 모형과 그리드분할 방식의 입력공간 분할과 조건부의 소속함수를 종형으로 정의하고 결과부의 함수형태를 선형함수로 정의한 GriP(2, 2, 2, 2)모형이다. <표 1>은 각 모형의 RMSE(Root Mean Square Error)와 분류정확도를 나타낸다. 이상의 결과를 토대로 볼 때 분류정확도에 있어서는 ANFIS 기반 분류모형이 C5.0 모형보다 유사하거나 높은 정확도를 나타내고 있다.

생성된 규칙에 대한 해석용이성을 평가하기 위해서 각 모형별 생성 규칙을 조사하고 완전성, 중복성, 간결성 측면에서 비교를 하였다. 그림1에 나타난 각 모형별 규칙들은 동일한 훈련데이터로부터 추출한 것이며, 별도의 규칙에 대한 필터링 없이 모든 생성규칙을 나열한 것이다.

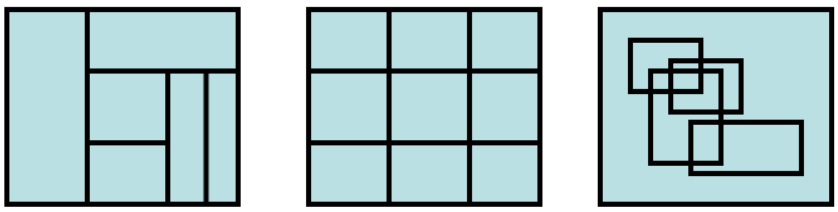
C5.0분류규칙(a)은 예측을 목적으로 의사결정나무로부터 유도된 규칙으로써 특정 사례에 대해 여러 개의 규칙이 매치되거나, 아무 규칙도 매치되지 않을 수도 있다. 만일 여러 개의 규칙이 매치되면 규칙의 신뢰도를 가중치로 삼아 가중치의 합이 큰 계급값을 할당하는 방법으로 최종 결론을 내리게 되며, 아무런 규칙도 매치되지 않는 경우에는 디폴트 값이 적용된다. C5.0에서 생성된 규칙(a)와

(a) C5.0 분류규칙	1. Default : 0 2. IF pl ⇐ 1.900 THEN 0 3. IF pl > 1.900 THEN 0.5 4. IF pl > 5 THEN 1 5. IF pw > 1.500 AND sw ⇐ 2.900 THEN 1
(b) C5.0 의사결정나무 규칙	1. Default : 0 2. IF pl ⇐ 1.900 THEN 0 3. IF pl > 1.900 AND pw ⇐ 1.500 THEN 0.5 4. IF pl > 1.900 AND pw > 1.500 AND pl ⇐ 5 AND sw > 2.900 THEN 0.5 5. IF pl > 1.900 AND pw > 1.500 AND pl ⇐ 5 AND sw ⇐ 2.900 THEN 1 6. IF pl > 1.900 AND pw > 1.500 AND pl > 5 THEN 1
(c) ANFIS 분류모형 GriP (2, 2, 2, 2) 종형 선형	1. IF sl = low AND sw = low AND pl = low AND pw = low THEN -2.589 2. IF sl = low AND sw = low AND pl = low AND pw = high THEN -2.962 3. IF sl = low AND sw = low AND pl = high AND pw = low THEN -7.66 4. IF sl = low AND sw = low AND pl = high AND pw = high THEN 26.9 5. IF sl = low AND sw = high AND pl = low AND pw = low THEN -1.099 6. IF sl = low AND sw = high AND pl = low AND pw = high THEN -0.4362 7. IF sl = low AND sw = high AND pl = high AND pw = low THEN 3.743 8. IF sl = low AND sw = high AND pl = high AND pw = high THEN 7.404 9. IF sl = high AND sw = low AND pl = low AND pw = low THEN -1.57 10. IF sl = high AND sw = low AND pl = low AND pw = high THEN -1.545 11. IF sl = high AND sw = low AND pl = high AND pw = low THEN -26.96 12. IF sl = high AND sw = low AND pl = high AND pw = high THEN 27.3 13. IF sl = high AND sw = high AND pl = low AND pw = low THEN 2.575 14. IF sl = high AND sw = high AND pl = low AND pw = high THEN 3.449 15. IF sl = high AND sw = high AND pl = high AND pw = low THEN 11.96 16. IF sl = high AND sw = high AND pl = high AND pw = high THEN 1.128
(d) ANFIS 분류모형 SubC(0.6)	1. IF sl = low AND sw = mid-high AND pl = low AND pw = low THEN -0.01245 2. IF sl = mid AND sw = mid-low1 AND pl = mid-high AND pw = mid THEN 1.41 3. IF sl = high AND sw = mid-low2 AND pl = high AND pw = high THEN 2.334

<그림 1> 아이리스 데이터에서 생성된 모형별 규칙 비교

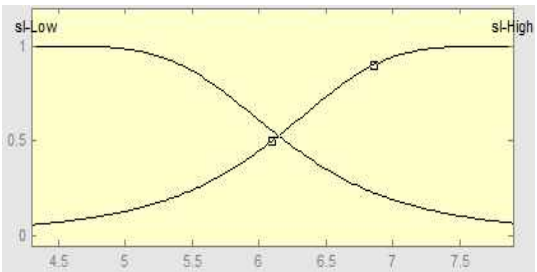
(b)는 규칙의 결론부에 계급값이 나타나는 데 비해, ANFIS의 규칙은 퍼지규칙이므로 연속형 값을 가지며, 특정 입력이 주어지면 이 입력과 매치되는 규칙들의 적합도를 종합하여 최종결론을 내리게 된다. C5.0에서 생성된 규칙이 전역에 걸쳐 일반화된 규칙임에 비해, ANFIS의 퍼지규칙은 로컬매핑(개별 전문가에 의해 기술된) 규칙이기 때문에 개별규칙만으로 결과를 직관적으로 이해하기는 어렵다. 조건부가 유사한 여러 규칙들의 결론부를 가중 평균하는 디퍼지화를 통해 최종결과가 도출됨을 이해해야 한다. 퍼지 규칙의 이러한 로컬매핑 특성은 현재 훈련 중인 학습데이터의 출력 오차를 최소화 할 뿐 아니라 과거에 학습된 데이터의 출력오차도 최소로 유지하도록 함으로써 정확성 향상에 기여하게 된다.

각 기법에서 생성된 규칙을 비교하면, 의사결정나무분석은 입력공간을 나무구조로 세분화하는데 비해, ANFIS 기반의 분류모형은 그리드분할 또는 임의군집화 방법을 사용하여 입력공간을 분할하고 있다. <그림 2>는 각 기법의 입력공간 분할 방법을 이차원 입력공간을 예로 들어 보여주고 있다. <그림 2>의 (a)와 같이 나무구조로 입력공간을 분할 할 경우, 그리드 분할과 달리 규칙의 수가 기하급수적으로 증가하는 것을 방지할 수 있다는 장점이 있지만, 세분된 퍼지영역을 표현하기 위해 더 많은 규칙들이 필요하며, 각 규칙의 조건부에 있는

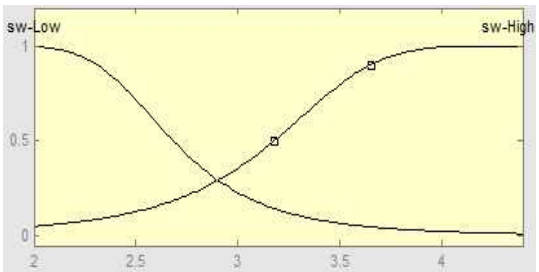


(a) 의사결정나무의 입력공간 분할 (b) 그리드 분할의 입력공간 분할 (c) 차감 클러스터링의 입력공간 분할

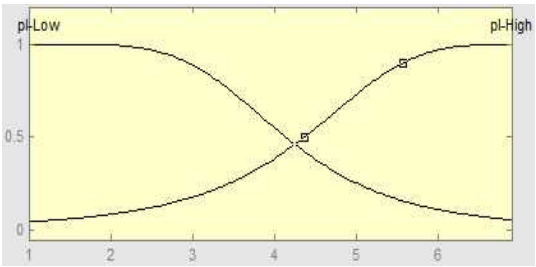
<그림 2> 각 기법의 입력공간 분할 방법(Jang and Sun, 1995)



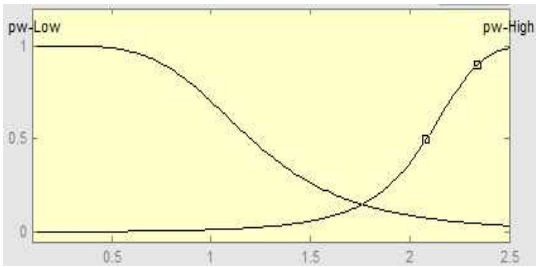
(a) sl 변수의 소속함수



(b) sw 변수의 소속함수



(c) pl 변수의 소속함수

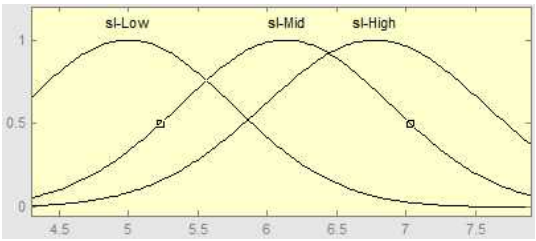


(d) pw 변수의 소속함수

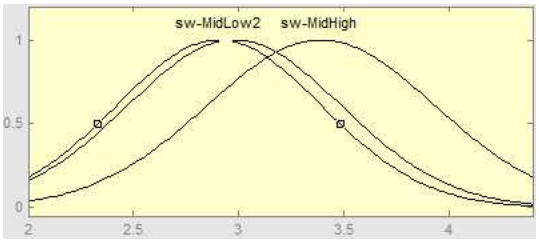
<그림 3> 그리드분할 모형의 입력변수별 소속함수

입력변수가 취하는 값에 “작다”, “크다”와 같은 의미 있는 명칭을 부여하기가 힘들어진다. 차감클러스터링의 경우, 훈련데이터에 대한 군집화를 통해 입력공간을 분할하기 때문에 빈 공간이 생기게 되어 완전성이 다소 저하 될 수 있다.

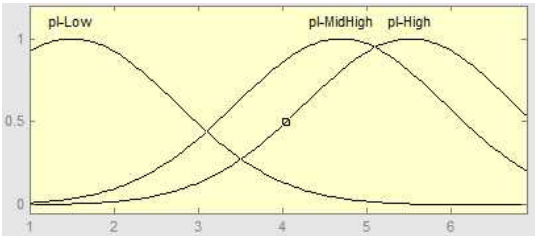
<그림 3>은 아이리스 데이터 세트에 적용된 그리드분할 모형의 각 입력변수별 소속함수 형태와 위치를 나타내고 있으며, <그림 4>는 차감클러스터링 모형의 각 입력변수별 소속함수 형태와 위치를 보여주고 있다. 그리드분할 모형의 경우, 기대했던 것처럼 네 변수 모두에서 각 퍼지집합의 소속함수가 뚜렷이 구분되도록 조정되어 퍼지용어의 할당이 용이한 것으로 나타났다. 그러나 차감클러스터링 모형의 경우, 규칙의 개수는 적으나 퍼지



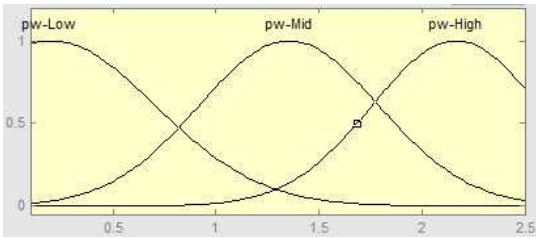
(a) sl 변수의 소속함수



(b) sw 변수의 소속함수



(c) pl 변수의 소속함수



(d) pw 변수의 소속함수

<그림 4> 차감클러스터링 모형의 입력변수별 소속함수

집합 간 소속함수 형태의 겹침 현상이 상대적으로 많이 발생하였다. 특히, sw변수의 경우 MidLow1과 MidLow2 퍼지집합의 소속함수는 겹침이 심하여 유사한 형태를 취하고 있다.

규칙의 완전성과 중복성, 간결성 측면에서 제안된 측도에 의해 각 규칙을 비교하면 <표 2>와 같다. 첫째, 완전성(CP) 측면에서 각 모형의 규칙을 비교하기 위해서 규칙의 적합도가 t = 0.25(경계값) 미만인 사례의 비율이 얼마나 되는지를 비교하였다. C5.0 규칙과 그리드분할모형의 경우, 검증데이터 세트 내의 모든 사례에 대해 적합도 0.25이상으로 매치되는 규칙이 존재하였다. 그러나 차감클러스터링 모형의 경우, 전체 50개 사례 중 2개의 사례는 매치되는 규칙이 없거나 적합도가 0.25미만으로 나타났다. 둘째, 중복성(RD) 측면에서 각 모형의 규칙을 비교하면 그리드분할모형의 규칙이 가장 중복도가 높았고, 다음으로 차감클러스터링 모형의 규칙, C5.0 분류규칙, C5.0의사결정나무규칙의 순서로 나타났다. ANFIS 모형에서 생성된 규칙들은 애매모호하거나 불확실성을 표현하기 위한 퍼지용어를 사용하기 때문에 중복성이 다소 높게 나타난 것으로 보인다. 차감클러스터링이 규칙수가 작음(3개)에도 불구하고 중복성이 상대적으로 높게 나타난 것은 그리드분할 시 입력변수를 두 개의 소속함수로 분할한데 반해 차감클러스터링의 경우 세 개의 소속함수로 분할함으로써, 학습을 통한 파라미터 조정결과 소속함수 간 겹치는 부분이 많았기 때문이다. 퍼지규칙은 일반적으로 최적화를 위해 규칙의 조건부(소속함수)와 결론부에 사용된 파라미터를 조정하기 때문에 규칙 간 중복이 많이 발생할 수 있음을 <그림 4>를 통해서도 알 수 있다. 마지막으로 간결성(SM)측면에서 볼 때, 규칙의 수는 차감클러스터링모형(3개), C5.0 분류규칙(5개), C5.0의사결정나무규칙(6개), 그리드분할모형(16개)의 순으로 증가하였다. 그러나 C5.0의사결정나무규칙은 입력변수가 반복적으로

<표 2> 생성된 규칙의 해석용이성 비교

규칙비교 관점	C5.0 분류규칙	C5.0의사결정 나무규칙	ANFIS 그리드분할 규칙	ANFIS 차감클러스터링 규칙
완전성(CP)	1.0000	1.0000	1.0000	0.9600
중복성(RD)	0.2049	0.0000	0.2600	0.2167
간결성(규칙수)	5개	6개	16개	3개

규칙에 사용되어 가독성이 매우 저하되었다. 한편 ANFIS 기반 분류모형들은 규칙의 조건부가 사람에게 친밀한 퍼지용어로 기술될 수 있어서 규칙에 대한 직관적 이해가 매우 용이하였다. 결과적으로 분류정확도를 최대로 하는 ANFIS 모형을 이용하여 규칙의 해석용이성을 평가 했음에도 불구하고 규칙의 완전성과 중복성 면에서 C5.0분류규칙과 비교하여 큰 차이를 보이지 않았다. 또한 ANFIS 차감클러스터링 규칙은 규칙의 수가 작아서 매우 간결하였고, 그리드분할 규칙의 경우 규칙의 수는 많았으나 규칙의 구조가 모두 일정하고, 퍼지용어로 기술되어 직관적인 이해가 용이하다는 장점이 있었다. 분류정확도가 요구되면서 입력변수의 수가 작은 도메인에 대한 지식획득이 필요한 경우에는 ANFIS 그리드 분할 모형이 대안이 될 수 있으며, 분류정확도가 요구되면서 입력변수의 수가 많은 도메인에 대한 지식획득이 필요한 경우에는 ANFIS 차감클러스터링 모형이 대안이 될 수 있다.

6. 결론

ANFIS 기반의 분류모형은 인간 전문가와 상호작용하면서 규칙을 정제해 나갈 수 있다. 특히 인간전문가가 사전지식을 이용하여 초기 퍼지규칙을 만들고 난 후 모형을 학습하면 최적에 수렴하는 시간을 단축할 뿐 아니라, 전역최적치 도달가능성이 높아진다. 이러한 관점에서 볼 때 규칙의 해석용이성은 매우 중요한 이슈가 될 수 있다. 본 연구에서는 퍼지신경망의 일종인 ANFIS를 분류문제에 적용하여 퍼지규칙을 생성하고 이를 규칙의 해석용이성 관점에서 평가하였다. 이를 위해 각 모형에서 생성된 규칙을 해석용이성 관점에서 비교하기 위한 측도를 제안하고, 사례연구를 통해 ANFIS의 퍼지규칙과 C5.0의사결정나무 규칙을 제안한 해석용이성 측도를 사용하여 비교하였다. 사례연구 결과 분류정확도를 최대로 하는 ANFIS 모형을 이용하여 규칙의 해석용이성을 평가 했음에도 불구하고 규칙의 완전성과 중복성 면에서 C5.0분류규칙과 비교하여 큰 차이를 보이지 않았다. 또한 ANFIS 차감클러스터링 규칙은 규칙의 수가 작아서 매우 간결하였고, 그리드분할 규칙의 경우 규칙의 수는 많았으나 규칙의 구조가 모두 일정하고, 퍼지용어로 기술되어 직관적인 이해가 용이하다는 장점이 있었다. 그러나 이러한 특성이 모든 사례에 적용될 수 있는 것은 아니므로 추가적인 사례에 대한 연구가 필요하다. 본 연구에서 제안된 해석용이성 측도들은 규칙을 생성하는 다양한 기계학습 모형의 규칙생성 능력을 평가하는 기준으로도 활용될 수 있을 것이다.

ANFIS 모형의 단점으로는 입력변수의 수가 많아질수록 입력공간 차원의 증가로 인해 규칙의 수가 갑자기 증가한다는 것이다. 이는 계산시간의 증가와 함께 규칙의 해석용이성을 저하시키는 결과를 가져온다. 또한 ANFIS 모형은 입력공간의

분할, 소속함수의 형태 결정, 규칙의 구조 결정, 추론 연산방법의 결정, 학습방법과 계수의 결정 등 다양한 요소의 조합을 통해 구축된다. 따라서 실제 문제에 적용하기 위해서는 ANFIS 모형의 특성과 해당 도메인의 데이터 특성을 잘 파악하고 있는 전문가의 도움이 필요할 수도 있다. 향후 연구로써 본 연구에서 제안된 평가측도들을 다양한 경영분야의 분류문제에 적용하여 그 유용성을 확인하는 것도 의미 있는 연구가 될 것이다. 또한 연관성규칙 등 분류문제가 아닌 다양한 문제에 해석용이성 평가측도를 적용함으로써 도메인을 다양한 관점에서 이해하는 것도 가능 할 것이다.

참고문헌

Asuncion, A. and D. J. Newman, UCI Machine Learning Repository [http://www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, CA : University of California, School of Information and Computer Science, 2007.

Babuška, R. and H. Verbruggen, "Neuro-fuzzy methods for nonlinear system identification", Annual Reviews in Control, Vol.27(2003), 73~85.

Bersini, H. and G. Bontempi, "Now comes the time to defuzzify neuro-fuzzy models", Fuzzy Sets and Systems, Vol.90(1997), 161~169.

Bodenhofer, U. and P. Bauer, A formal model of interpretability of linguistic variables : Trade-off between Accuracy and Interpretability in Fuzzy Rule-Based Modelling, Studies in Fuzziness and Soft Computing, Physica, Heidelberg, 2002.

Chen, M.-Y. and D. A. Linkens, "Rule-base self-generation and simplification for data-driven fuzzy models", Fuzzy Sets and Systems, Vol.142, No.2(2004), 243~265.

Chiu, S., "Fuzzy Model Identification Based on Cluster Estimation", Journal of Intelligent and Fuzzy Systems, Vol.2, No.3(1994).

Cordon, O. and F. Herrera, "A proposal for improving the accuracy of linguistic modeling", IEEE Trans. Fuzzy Systems Vol.8, No.3 (2000), 335~344.

Efendigil, T., S. Onut and C. Kahraman, "A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models : A comparative analysis", Expert Systems with Applications, Vol.36 (2009), 6697~707.

Han M., Y. Sun and Y. Fan, "An improved fuzzy neural network based on T. S model", Expert Systems with Applications, Vol.34 (2008), 2905~2920.

Jang, J.-S. R., "ANFIS: Adaptive-Network-based Fuzzy Inference Systems", IEEE Transactions on Systems, Man, and Cybernetics, Vol.23, No.3(1993), 665~685.

Jang, J.-S. Roger and C.-T. Sun, "Neuro-Fuzzy Modeling and Control", The Proceedings of the IEEE, Vol.83, No.3(1995), 378~406.

Jin, Y. W., V. Seelen and B. Sendhoff, "An approach to rule-based knowledge extraction", Proceedings of IEEE Conference on Fuzzy Systems, 1998, 1188~1193.

Jin, Y. W., V. Seelen and B. Sendhoff, "On generating flexible, complete, consistens and compact (FC3) fuzzy rules from data using evolution strategies", IEEE Transactions on Systems, Man, and Cybernetics, Vol.29, No.4 (1999), 829~845.

Jin, Y., "Fuzzy modeling of high-dimensional sys-

tems : complexity reduction and interpretability improvement", IEEE Transactions on Fuzzy Systems, Vol.8, No.2(2000), 212~221.

Matlab, Fuzzy logic toolbox 2 user's guide. The Math Works Inc, 2009.

Mikut, R., J. Jakel, and L. Groll, "Interpretability issues in data-based learning of fuzzy systems", Fuzzy Sets and Systems, Vol.150(2005), 179~197.

Nauck, D. and R. Kruse, "Neuro-fuzzy systems for function approximation", Fuzzy Sets and Systems, Vol.101(1999), 261~271.

Song, H. S. and J. K. Kim, "Design and Evaluation of ANFIS-based Classification Model", Journal of Inteligence and Information Systems, Vol.15, No.3(2009), 151~165.

Takagi, T. and M. Sugeno, "Derivation of fuzzy control rules from human operator's control actions", Proceedings of the IFAC symposium on fuzzy information, knowledge representation and decision analysis, 1983, 55~60.

Valente J., "Semantic constraints for membership function optimization", IEEE Trans, Systems Man Cybernetics.Part A : Systems and Humans Vol.29, No.1(1999), 128~138.

Abstract

Evaluation of Interpretability for Generated Rules from ANFIS

Hee Seok Song · Jae Kyeong Kim

Fuzzy neural network is an integrated model of artificial neural network and fuzzy system and it has been successfully applied in control and forecasting area. Recently ANFIS(Adaptive Network-based Fuzzy Inference System) has been noticed widely among various fuzzy neural network models because of outstanding performance of control and forecasting accuracy. ANFIS has capability to refine its fuzzy rules interactively with human expert. In particular, when we use initial rule structure for machine learning which is generated from human expert, it is highly probable to reach global optimum solution as well as shorten time to convergence. We propose metrics to evaluate interpretability of generated rules as a means of acquiring domain knowledge and compare level of interpretability of ANFIS fuzzy rules to those of C5.0 classification rules. The proposed metrics also can be used to evaluate capability of rule generation for the various machine learning methods.

Key Words : Fuzzy system, Fuzzy neural network, ANFIS, Classification rule, Interpretability, Completeness, Redundancy

* Department of Management Inforamtion System, Hannam University
** The School of Management, Kyung Hee University

저 자 소 개

송희석

1987년 고려대학교에서 경영학사, 1989년 한국과학기술원 경영과학 석사, 2003년 한국과학기술원 경영공학 박사학위를 취득하였다. 1989년부터 약 15년 간 대우정보시스템에 근무하였으며, 2003년 이후 한남대학교 경영정보학과 부교수로 재직 중이다. Artificial Intelligence Review, Knowledge Based Systems, Expert Systems, Expert Systems with Applications 등 다수의 저널에 논문을 발표하였고 관심분야는 CRM과 Data Mining, 유비쿼터스 비즈니스 등이다.

김재경

서울대학교에서 산업공학 학사, 한국과학기술원(KAIST)에서 경영정보시스템 전공으로 석사 및 박사학위를 취득하였다. 현재 경희대학교 경영대학 교수로 재직하고 있으며, 한국경영정보학회 부회장, BK21 네트워크 과학에 기반한 경영학 전문연구인력 양성사업팀 단장, 저탄소녹색성장국민포럼산하 그린IT분과위원을 역임중이다. 주요 연구분야로는 비즈니스 인텔리전스, 네트워크 경영, 그린 비즈니스/IT 등이다. Artificial Intelligence Review, Electronic Commerce Research and Applications, European Journal of Operational Research, Expert Systems with Applications, Group Decision and Negotiations, IEEE Transactions on services computing, International Journal of Human-Computer Studies, Technological Forecasting and Social Change, 등 다수의 학술지에 논문을 게재하였으며, 또한 학술지 지능정보연구 편집위원장, Information Technology and Management(SSCI) AE(Associate Editor)를 역임중이다.