

# 데이터마이닝을 위한 경쟁학습 모델과 BP알고리즘을 결합한 하이브리드 알고리즘

강문식, 윤호근, 이상용  
공주대학교 전자계산학과, 공주대학교 정보통신공학부  
{cult, sylee@kongju.ac.kr}

Mun-sik Kang, Sang-yong Lee  
Dept. of Computer Science, Division of Information & Communication  
Engineering, Kongju National University

[314-802] 충남 공주시 신관동 공주대학교 자연과학대학 전자계산학과  
인공지능연구실(326호)  
공주대학교 전자계산학과 T.041)850-8520, FAX.041)856-5790

## 요 약

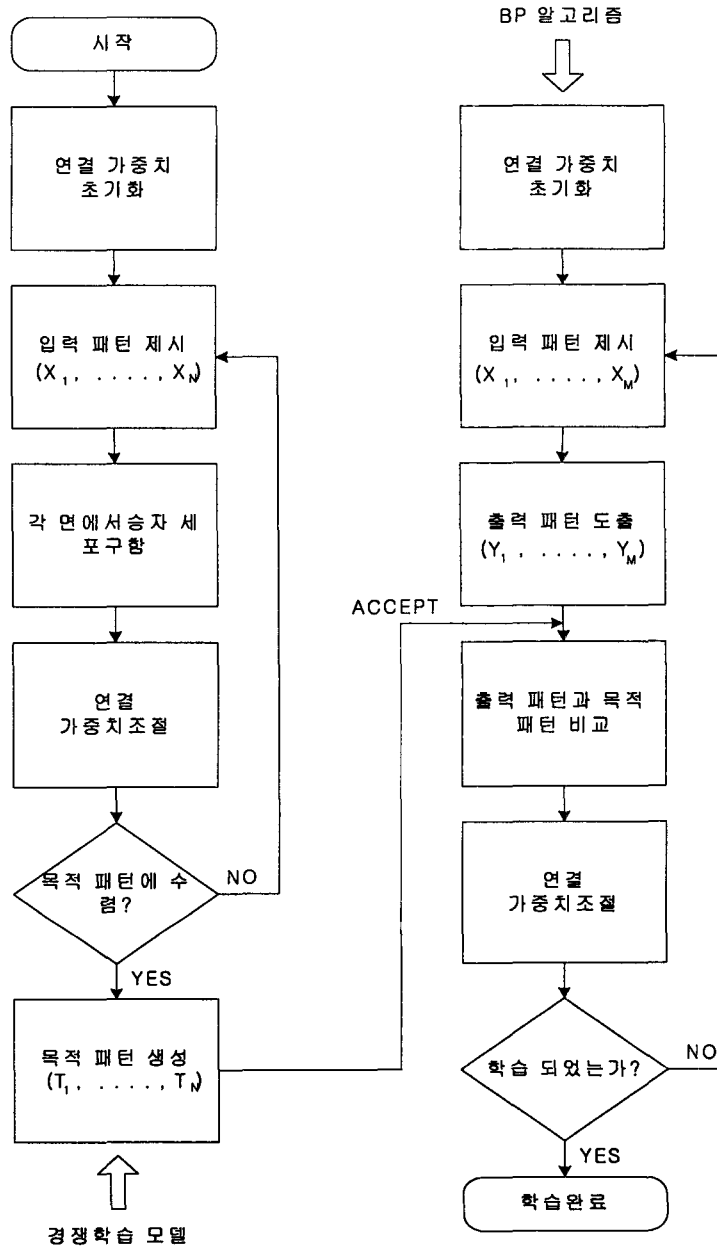
최근에 데이터 마이닝에서 군집화와 분류기법이 중요한 도구로써 많은 분야에서 응용되어 사용되고 있다. 감독 학습 방법인 BP알고리즘과 무감독 학습 방법인 경쟁학습 모델(Competitive learning model)은 데이터 마이닝을 위한 패턴 분류에 많이 쓰이고 있다. 그러나, 두 알고리즘 모두 지역극소에 빠지거나, 정규화 과정을 거치는 등의 문제로 데이터를 마이닝하는 과정에서 중요한 정보가 왜곡되어 잘못된 결과를 도출하는 사례가 종종 발생하고 있다.

감독 학습 방법의 경우에는 목적 패턴을 사용하기 때문에, 신경망의 출력을 인위적으로 제어할 수 있다. 그러나, 무감독 학습 방법의 경우에는 목적 패턴이 없기 때문에, 여러 개의 뉴런이 동시에 반응있다. 따라서, 하나의 신경세포만이 반응할 수 있도록 해주는 승자전취(Competitive learning model) 같은 메커니즘이 필요하게 되는데, 문제점은 신경망 내에서 어떤 연결 가중치나 여러 입력 패턴 중 하나의 입력패턴이 현저하게 클 경우, 모든 입력 패턴에 반응해서 학습이 이루어지지 않는 경우가 있다. 그래서 연결 가중치나 입력 패턴이 정규화 과정을 거쳐서 다른 가중치나 입력 패턴에 비해 현저하게 큰 경우를 방지하는데, 이러한 과정은 가중치나 패턴이 변질되기 때문에 좋은 방법이라고 결코 말할 수가 없다.

본 논문은 경쟁학습 모델과 BP알고리즘을 혼합형으로 만든 하이브리드형 알고리즘을 제안한다. 어떠한 입력 패턴에 대해 목적 패턴을 만들 수 없는 경우에 이 알고리즘을 사용하게 되는데, 경쟁학습 모델을 이용해서 목적 패턴만을 생성함으로써, 정규화의 위험 과정을 줄일 수 있고, 현재 널리 응용되고 있는 BP알고리즘의 목적 패턴으로 이용할 수 있다. 여기에서, 경쟁학습 모델은 학습 규칙으로 그로스버그(Stephen Grossberg)의 인스타 규칙(Instar rule)을 이용하고, BP알고리즘의 학습 규칙으로는 룸멜하트(David E Rumelhart)의 일반화된 델타 규칙(Generalized delta rule)을 이용하였다. 두 알고리즘의 혼합형으로 만든 알고리즘은 <그림 1>과 같다.

본 알고리즘의 타당성을 평가하기 위해 임의의 입력 패턴을 이용하여, 알고리즘의 패턴

분류 효과를 실험하였다. <표 1>의 임의의 입력 패턴으로 아래 알고리즘의 경쟁학습 모델 부분에서 목적패턴을 만든 후에 BP알고리즘에 사용하였다. 그리고 현재는 효과적인 패턴 분류를 할 수 있는 작업중이다.



<그림 17> 하이브리드 알고리즘

<표 2> 학습을 위한 입력 패턴

학습 패턴					
번호	입력패턴		번호	입력패턴	
0	-0.17	0.00	10	0.31	-0.09
1	-0.35	0.18	11	-0.21	-0.1
2	-0.7	-0.8	12	-0.5	0.62
3	-0.65	-0.55	13	-0.49	0.52
4	-0.9	0.19	14	0.61	0
5	0.6	-0.35	15	-1	0.98
6	0.75	0.7	16	0.07	0.72
7	0.23	0.36	17	-0.72	-0.32
8	0.67	0.54	18	0.29	-0.34
9	0.23	0.5	19	-0.61	0.41