

협력적 추천을 위한 사용자와 항목 모델의 효율적인 통합 방법

(Efficient Methods for Combining User and Article Models for Collaborative Recommendation)

도 영 아 [†] 김 중 수 [†] 류 정 우 ^{**} 김 명 원 ^{***}
(Young A Do) (Jong Su Kim) (Joung Woo Ryu) (Myung Won Kim)

요 약 협력적 추천에서는 일반적으로 사용자 모델과 항목 모델이 사용되어진다. 사용자 모델은 사용자들간의 선호도 상관관계를 학습하고, 추천하고자 하는 항목에 대한 다른 사용자들의 선호도를 기반으로 그 항목을 추천한다. 이와 유사한 방식으로 항목 모델은 항목들간의 선호도 상관관계를 학습하고, 다른 항목들간의 선호도를 기반으로 추천 받는 사용자에게 항목을 추천한다.

본 논문에서는 추천 성능의 향상을 위해서 사용자 모델과 항목 모델간의 다양한 통합 방법을 제안한다. 제안하는 통합 방법으로는 순차적, 병렬적 통합 방법, 퍼셉트론 또는 다층 퍼셉트론을 이용한 통합 방법, 퍼지 규칙을 이용한 통합 방법 그리고 BKS를 적용한 방법이다. 본 실험에서는 통합 모델을 위해서 다층 퍼셉트론을 이용하여 사용자와 항목 모델을 각각 학습한다. 다층 퍼셉트론은 최근접 이웃방법이나 연관 규칙을 이용한 방법과 같은 기존의 추천 방법보다 연관된 항목들간의 가중치를 학습할 수 있고, 기호 데이터와 수치 데이터를 쉽게 처리할 수 있는 장점이 있다. 본 논문에서는 통합된 모델이 어떠한 단일 모델보다도 우수하고, 실험을 통하여 다층 퍼셉트론을 이용한 통합 방법이 다른 통합 방법보다 효율적인 통합 방법임을 보여주고 있다.

키워드 : 추천 시스템, 협력적 추천, 신경망, 개인화, BKS

Abstract In collaborative recommendation two models are generally used: the user model and the article model. A user model learns correlation between users preferences and recommends an article based on other users preferences for the article. Similarly, an article model learns correlation between preferences for articles and recommends an article based on the target user's preference for other articles. In this paper, we investigate various combination methods of the user model and the article model for better recommendation performance. They include simple sequential and parallel methods, perceptron, multi-layer perceptron, fuzzy rules, and BKS. We adopt the multi-layer perceptron for training each of the user and article models. The multi-layer perceptron has several advantages over other methods such as the nearest neighbor method and the association rule method. It can learn weights between correlated items and it can handle easily both of symbolic and numeric data. The combined models outperform any of the basic models and our experiments show that the multi-layer perceptron is the most efficient combination method among them.

Key words : Recommender Systems, Collaborative Recommendation, Neural Network, Personalization, Behavior Knowledge Space

· 본 연구는 한국 과학기술부에서 지원하는 뇌신경정보학 연구사업으로 수행되었음.

[†] 비 회 원 : M플러스택(주)

npng1@naver.com

kjongsu@orgio.net

^{**} 학생회원 : 숭실대학교 대학원 컴퓨터학과

ryu0914@orgio.net

^{***} 종신회원 : 숭실대학교 컴퓨터학부 교수

mkim@comp.ssu.ac.kr

논문접수 : 2001년 10월 9일

심사완료 : 2003년 3월 11일

1. 서론

최근에는 사용자 개인의 취향이나 특성에 맞는 정보를 자동으로 찾거나 추천해 주는 개인화(personalization) 기술로서 추천 시스템이 크게 요구되고 있다 [1,2]. 추천시스템은 고객들의 상품에 대한 선호도와 구매기록을 분석하여 개인별로 차별화된 정보를 제공하기

위한 자동화된 정보필터링 시스템이다. 즉, 취향에 맞는 항목 및 정보를 사용자에게 제공하는 시스템이라 할 수 있다. 이러한 추천시스템에서 사용하고 있는 추천 기술은 크게 협력적 추천(collaborative recommendation), 내용기반 추천(content-based recommendation), 인구통계학적 추천(demographic recommendation)을 들 수 있다. 협력적 추천은 특정 사용자와 유사한 선호도를 갖는 다른 사용자들의 목표 항목에 대한 선호도를 바탕으로 그 항목의 선호도를 추정하여 추천하는 기술이며, 내용기반 추천은 사용자에게 있어 이전에 선호한 항목과 비슷한 속성(특성)을 갖는 항목은 사용자가 선호할 가능성이 높다고 보고 선호도가 표시된 항목들의 속성 정보를 이용하여 추천하는 기술이다. 이에 비해 인구통계적 추천은 사용자의 나이, 성별 등과 같은 인구통계적 정보를 바탕으로 항목이나 정보의 선호도를 추정하여 추천하는 기술이다.

기존 협력적 추천 중에서 대표적인 방법으로는 최근접 이웃 방법(Nearest Neighbor Method), 연관규칙 방법(Association Rule Method)[3] 그리고 신경망 접근 방법[4] 등이 있다. [4]에서는 개인화 된 효율적인 추천을 위해서 연관 규칙을 이용한 협력적 추천기술[3]과 신경망을 이용한 협력적 추천기술의 성능을 비교함으로써 신경망을 이용한 추천기술이 더욱 향상된 추천기술임을 제시하고 있다. 그러나 사용자 신경망 모델은 사용자간의 연관성을 통해서 선호도를 예측하고, 항목 신경망 모델은 항목간의 연관성을 통해서 선호도를 예측하므로 단일 신경망 모델로만 추천함으로써 즉, 사용자 또는 항목 신경망 모델로만 추천함으로써 두 신경망 모델이 서로 다른 선호도를 제시할 경우, 어느 신경망 모델의 선호도 결과로 추천해야하는지에 따라서 추천 성능이 좌우된다.

따라서, 본 논문에서는 단일 신경망 모델이 각각의 다른 특징과 다른 관점으로 예측한 선호도를 통합하여 단일 신경망 모델의 추천기술이나 기존의 추천기술 보다 성능이 향상됨을 보이고 있다. 또한 단일 신경망 모델에 대한 통합 방법으로는 단일 신경망 모델 중 우선 순위를 적용하는 순차적 통합 방법, 동시에 고려하는 병렬적 통합 방법, 단일 신경망 모델의 출력 값을 입력으로 하는 신경망(퍼셉트론, 다층 퍼셉트론)을 이용한 통합 방법, 퍼지 추론을 이용한 통합 방법을 제안한다. 또한 패턴 인식에서 사용되는 CME 방법 중 BKS방법을 통합 방법으로 적용하여 비교하고 있다.

본 논문은 다음과 같은 순서로 구성되어 있다. 2장에서는 패턴 인식 분야에서 쓰이는 기존의 통합 방법들을 기술한다. 3장에서는 신경망을 이용한 협력적 추천에 대

해 기술하고, 4장에서는 단일 신경망 모델을 통합하는 방법들을 기술한다. 5장에서는 단일 신경망 모델의 매개 변수와 실험을 통한 결과를 제시하고, 6장에서는 결론을 맺고, 향후연구를 제시한다.

2. 기존의 통합 방법

CME(Combination of Multiple Experts)는 패턴 인식 분야에서 사용되는 방법으로 다른 방법론이나 다른 특성을 갖고 있는 전문가들(classifiers 혹은 experts)이 서로의 의견을 통합하기 위해 사용하는 방법이다. CME 방법을 이용함으로써 집단의 의사 결정이나 문제 해결에 있어서 높은 성능 향상과 더불어 오차를 급속히 줄일 수 있다[5].

본 논문에서는 [4]에서 제안한 단일 신경망 모델인 사용자와 항목 신경망 모델 각각을 추천하고자 하는 관점이 다른 전문가로 고려하고 각각의 의견을 통합함으로써 추천의 성능을 향상시킬 수 있다. 따라서 CME방법들 중 대표적인 방법들을 적용해 보고자 한다.

2.1 Voting

다른 전문가들의 의견을 통합하기 위해 사용되는 대표적인 방법이고, 간단한 방법으로 voting by majority 방법[6]이 있다. 이 방법은 반수 이상의 전문가들이 같은 의견으로 동의하면 그 의견으로 집단의 전체 의사를 결정하는 방법이다. conservative voting rule 방법은 같은 의견으로 모든 전문가들이 동의한 경우 그 의견으로 의사 결정을 하는 방법이다. 또 다른 방법은 comparative majority-voting 방법이다. 이 방법은 가장 많이 동의한 의견과 두 번째로 많이 동의한 의견의 전문가 수 차이가 주어진 임계값보다 클 경우에 가장 많이 동의한 의견이 집단 전체의 의사 결정으로 반영되는 방법이다. 임계값을 고려하는 이유는 최종적으로 선택한 의견의 수가 충분히 커야하기 때문이다. 만약 가장 많은 의견의 수가 두 번째로 많은 의견의 수보다 충분히 크지 않다면 가장 많은 의견의 수가 아주 클지라도 그 결정은 신뢰할 수 없게 된다.

2.2 The Borda Count Method

1770년에 Jean-Charles de Borda에 의해 제시된 방법으로 전문가들의 의견을 상충하는 문제에 쉽게 적용되는 방법 중의 하나이다[7,8]. 이 방법은 전문가들이 의견을 제시할 때 우선 순위를 제시하는 방법으로 가장 높은 우선 순위를 갖는 의견은 가장 높은 가중치를 줌으로써 집단의 의사를 결정한다. 예를 들어 각 아이스크림에 대한 선호도가 같은 100명의 투표자의 우선순위가 <표 1>과 같을 때,

표 1 아이스크림 선호도에 대한 우선순위

| 투표자수 | 27명 | 7명 | 20명 | 10명 | 3명 | 33명 |
|------|-----|----|-----|-----|----|-----|
| 초코렛 | 1 | 1 | 2 | 3 | 2 | 3 |
| 바닐라 | 2 | 3 | 1 | 1 | 3 | 2 |
| 딸기 | 3 | 2 | 3 | 2 | 1 | 1 |

가장 많이 선호하는 아이스크림은 <식1>과 같이 계산된다. 여기서, 우선순위가 높다는 것은 가중치가 높다는 것을 의미하기 때문에 <표 1>에서 우선순위가 1인 경우 가중치를 3으로 생각하여 계산하게 된다. 즉, 초코렛 아이스크림 같은 경우 34명(27+7)의 투표자에 의해 우선순위 1로 설정되었으므로 BordaCount 계산시 그 항목의 가중치는 3이 되어 계산된다.

$$\text{초코렛의 BordaCount} = [(27+7) \times 3] + [(20+3) \times 2] + [(10+33) \times 1] = 191$$

$$\text{바닐라의 BordaCount} = [(20+10) \times 3] + [(27+33) \times 2] + [(7+3) \times 1] = 220$$

$$\text{딸기의 BordaCount} = [(3+33) \times 3] + [(7+10) \times 2] + [(27+20) \times 1] = 189 \quad (1)$$

따라서 BordaCount가 가장 높은 바닐라 아이스크림을 가장 많이 선호하는 것으로 결정할 수 있다. 이와 같이 2.1절에서 서술한 Voting 방법처럼 전문가가 한가지 의견만 제시할 수 있는 것이 아니라 우선순위를 명시하여 여러 의견들도 표현할 수 있는 장점이 있다.

2.3 Behavior Knowledge Space (BKS)

앞에서 기술한 CME 방법들은 각 전문가들을 동등하게 취급하고, 한 전문가가 잘못된 의견을 제시해도 모든 전문가의 의견을 통합하는 데에는 큰 영향을 미치지 않는다. 그 이유는 CME방법들이 전문가의 독립성을 가정하고 있기 때문이다. BKS [5]에서는 이러한 가정을 두지 않기 위해 모든 전문가들의 의견을 동시에 수용하는 지식 공간(knowledge space)으로부터 의사 결정을 도출한다. 지식 공간에서 전문가의 모든 의사 결정을 수용하기 때문에 이러한 방법을 행동 지식 공간(Behavior-Knowledge Space)이라고 부른다. 간단하게 말해서 BKS 방법은 행동 지식 공간으로부터 그 집단의 최종 결정을 내리게 된다.

예를 들어 K차원의 행동 지식 공간을 형성한다면 각 차원은 전문가(expert)로 대응되고, 다시 차원은 전문가가 결정할 수 있는 의견들로 구성되어 각각 전문가들의 의견들의 교차(intersection)는 BKS의 셀(cell/unit)로 형성된다[5,9]. 각각의 셀은 세 가지의 정보를 내포하고 있다: 1) 셀에 포함된 입력데이터 개수, 2) 대표 의견,

3) 셀에 포함된 각 클래스에 대한 입력데이터 개수.

<식 2>에서 $R_{e(1), \dots, e(K)}$ 는 임의의 셀($BKS(e(1), \dots, e(K))$)에서의 대표 의견(j)을 나타내고 있다. $e(i), (1 \leq i \leq K)$ 은 i번째 전문가에 의한 입력데이터의 의견을 의미하며, M은 클래스 개수를 나타내고 $n_{e(1), \dots, e(K)}(m)$ 는 클래스(m)에 대한 입력데이터 개수를 의미한다. 따라서 대표의견(j)은 셀에 포함된 각 클래스에 대한 입력데이터 개수 중 가장 많은 클래스로 정의하게 된다.

$$R_{e(1), \dots, e(K)} = \{j \mid n_{e(1), \dots, e(K)}(j) = \max_{1 \leq m \leq M} n_{e(1), \dots, e(K)}(m)\} \quad (2)$$

테스트 데이터가 들어오면 전문가들은 이러한 셀의 데이터 정보를 이용하여 즉, 행동 지식 공간의 정보를 이용하여 의견을 통합 도출하게 된다.

예를 들어 같은 입력데이터에 대해서 첫 번째 전문가는 4, 두 번째 전문가는 9라는 의견을 제시했을 때 $e(1)=4, e(2)=9$ 가 된다. BKS(4,9) 셀에 포함된 데이터 수를 클래스 별로 합하면 <식 3>과 같다.

$$\begin{cases} n_{4,9}(4) = 15, \\ n_{4,9}(9) = 5, \\ n_{4,9}(m) = 0, \text{ 여기서 } m \notin \{4, 9\}, m \in A \end{cases} \quad (3)$$

여기서 $A = \{1, \dots, M\}$ 이다. 즉, $n_{4,9}(4) = 15$ 는 BKS(4,9) 셀에 포함된 데이터 중 클래스가 4인 데이터 개수가 15개라는 것을 의미하고 따라서 BKS(4,9)의 대표의견(R4,9)은 4가 된다.

2.4 순차적 통합 방법(Sequential Method)

순차적 통합 방법은 전문가들에게 우선 순위를 주어진 의견을 순차적으로 적용하는 방법이다. 연관 규칙을 이용한 추천 방법을 제안하고 있는 [3]에서는 사용자 연관 규칙과 항목 연관 규칙을 통합하는 방법으로 임계값을 이용한 통합 방법을 기술하고 있다. 사용자 연관 규칙의 최소 지지도가 임계값보다 크면 사용자 연관 규칙을 사용하고, 작으면 항목 연관 규칙을 사용하는 방법이다. 본 논문에서 적용하는 순차적 통합 방법은 4.1절에서 자세히 다루고 있다.

Voting방법과 Borda Count 방법은 셋 이상의 전문가가 필요하므로 본 논문에서 통합하려는 방법에 적합하지 않다. 그러므로 사용자와 항목 신경망 모델을 통합하는데 있어 기존 통합 방법 중 BKS와 순차적 통합 방법을 적용한다.

3. 신경망을 이용한 추천

신경망은 고등동물 두뇌의 정보처리 메커니즘을 모방한 새로운 계산 방법으로 학습기능을 가짐으로써 기존의 컴퓨터가 해결하기 어려운 문제, 특히 효율적인 알고

리즘이 없는 비 알고리즘 문제를 용이하게 해결할 수 있고, 잡음에 강해 패턴인식과 시스템 인식 및 제어 등 여러 문제에 많이 응용되고 있다. 다층 퍼셉트론(Multi-Layer Perceptron)은 입력층과 출력층 사이에 하나 이상의 은닉층(중간층)이 존재하는 신경망 모델이다. 전형적인 다층 퍼셉트론은 [그림 1]과 같이 X 라는 입력벡터를 갖는 입력층과 O 라는 출력벡터를 갖는 출력층으로 이루어진다. 일반적인 다층 퍼셉트론의 학습방법은 오류역전파(Error Back-Propagation) 알고리즘을 적용하여 입력층의 각 노드에 입력 데이터를 제시하면 이 신호는 각 노드에서 변환되어 은닉층에 전달되고 최종적으로 출력층으로 나오게 된다. 이 출력값과 바라는 출력값(Desired Output Value)을 비교하여 그 차이를 감소시키는 방향으로 연결강도를 조정하는 것이다.

[4]에서는 항목들간 또는 사용자들간의 선호 상관관계를 다층 퍼셉트론으로 학습시킴으로써 모델을 생성하고, 그 모델을 사용하여 선호도를 예측한다. 신경망을 이용하면 항목이나 사용자간의 가중치를 학습할 수 있고, 연속 수치형, 이진 논리형, 범주형 등의 자료 유형에 상관없이 데이터 처리가 용이하다. 또한 다른 이질적(내용, 인구통계학적 정보 및 항목이나 사용자들 간의 상관관계)인 종류의 데이터 및 정보를 통합하기가 용이하다. 그러므로 [4]에서는 추천 기술로써 신경망을 이용한 추천 기술을 제안하고, 기존의 추천 기술인 연관 규칙[3]을 이용한 추천기술보다 성능의 우수함을 제시하고 있다.

사용자 신경망 모델은 다른 사용자들간의 연관성으로 항목의 선호도를 예측한다. [그림 2]에서 보는 바와 같이 사용자(A)모델은 항목(B)에 대한 다른 사용자들의 선호도 정보를 이용하여 항목(B)에 대한 사용자(A)의 선호도를 예측하는 신경망 모델이다. 반면, 항목 신경망 모델은 항목들간의 연관성으로 사용자에게 추천할 항목의 선호도를 예측한다. [그림 3]에서와 같이 항목(B)모델은 사용자(A)의 다른 항목들의 선호도 정보를 가지고 사용자(A)에게 항목(B)의 선호도를 예측하는 신경망 모델이다.

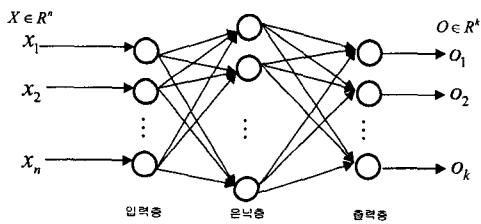


그림 1 3층 구조를 가진 다층 퍼셉트론

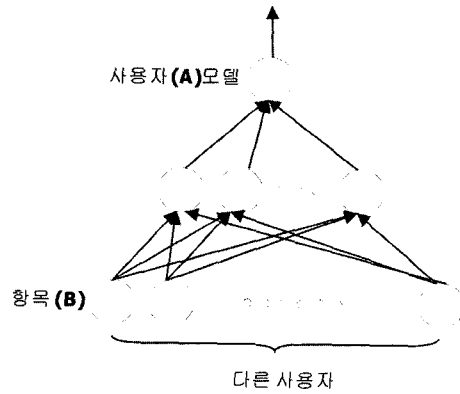


그림 2 사용자(A)에 대한 신경망 모델

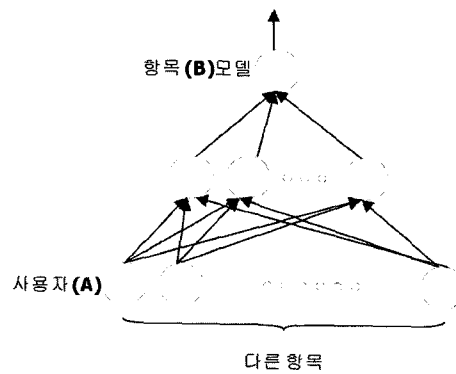


그림 3 항목(B)에 대한 신경망 모델

4. 사용자와 항목 신경망 모델의 통합 방법

[4]에서는 단일 신경망 모델로만 추천함으로써, 다음과 같은 문제가 발생한다. 사용자와 항목 신경망 모델이 한 항목에 대해 동시에 같은 선호도로 동의하면 즉, [그림 4]의 A영역인 경우에는 어느 모델의 선호도 결과에 따라도 상관없다. 그러나 두 모델이 서로 다른 선호도로 동의하지 않을 때, [그림 4]의 D영역일 경우, 선택한 모델의 결과에 따라 성능이 좌우된다. 그러므로 D영역은 문제에 따라서 어떻게 선호도 영역으로 분할하느냐에 따라 추천 성능에 영향을 미치게 된다. 그러므로 서로 다른 관점을 통해 예측된 선호도를 통합하여 항목에 대한 추천이 이루어지면, 단일 신경망 모델에 의한 추천이나 기존의 추천 방법보다 유연성과 효율성 모두 향상될 수 있다.

본 논문에서는 이러한 점을 이용하여 [그림 5]와 같이 사용자와 항목 신경망 모델을 통합한 통합모델을 제안

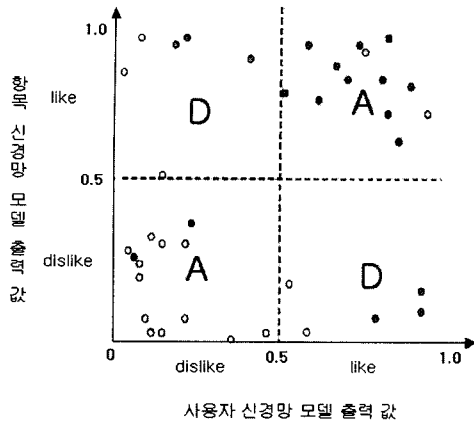


그림 4 사용자와 항목 신경망 모델의 선호도 분포

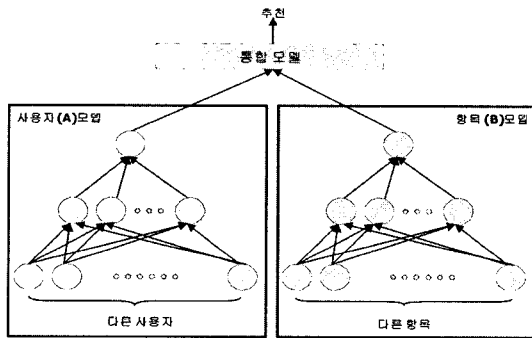


그림 5 제안한 통합 모델

하고 이에 따른 여러 가지 통합방법들에 대해 제안하고 있다.

4.1 순차적 통합 방법(Sequential Method)

순차적 통합 방법은 어떤 한 모델을 우선적으로 채택하여 선호도를 결정되되 확실한 결론을 내리기 어려울 때 다른 모델의 결과로 선호도를 판정하는 방법이다. 만약 개인의 취향이 강한 항목 예를 들면 의류, 화장품 같은 경우에는 항목 신경망 모델을 먼저 적용하고, 그렇지 않은 영화, 독서, 음반인 경우에는 사용자 신경망 모델을 먼저 적용한다. 본 논문에서는 사용자 신경망 모델의 실제 출력 값이 0.3보다 작으면 dislike, 0.7보다 크면 like로 선호도를 판정하고, 그렇지 않으면 항목 신경망 모델의 결과로 선호도를 판정한다. 이때 항목 신경망 모델이 0.5보다 크면 like, 작으면 dislike로 선호도를 판정한다.

[그림 6]의 (B)는 순차적 통합 방법을 사용할 경우 입력 공간에서 like, dislike 선호도의 분할 영역을 보여주고 있다.

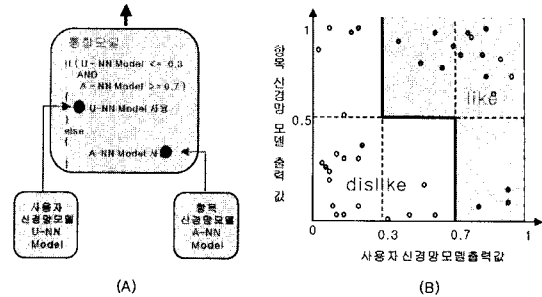


그림 6 순차적 통합 방법과 선호도 경계선

4.2 병렬적 통합 방법(Parallel Method)

병렬적 통합 방법은 두 모델의 결과를 동시에 고려하여 최종 결론을 내리는 방법이다. 즉, 두 모델의 선호도 결과 중에 가장 확실한 선호도로 판정하는 방법이다. 예를 들어 어떤 항목에 대해 사용자 신경망 모델의 선호도가 0.9이고, 항목 신경망 모델의 선호도가 0.2일 경우 사용자 신경망 모델이 긍정적인 선호도에 더 가깝기 때문에 그 항목을 사용자에게 추천한다.

[그림 7]의 (B)는 병렬적 통합 방법을 사용할 경우 입력 공간에서 선호도의 분할 영역을 보여주고 있다. 두 모델의 선호도 결과 중 가장 확실한 선호도로 판정하는 방법이므로 (B)와 같은 like, dislike 선호도의 경계선이 생기게 된다.

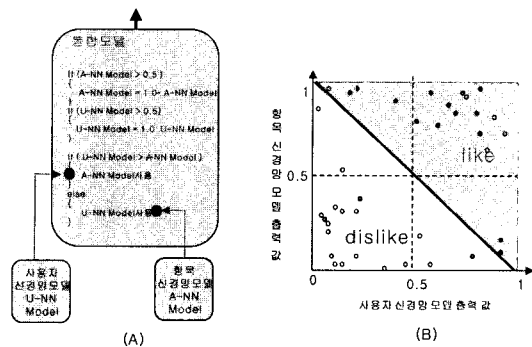


그림 7 병렬적 통합 방법과 선호도 경계선

4.3 신경망을 이용한 통합 방법

신경망을 이용한 통합 방법은 [10]에서 사용하고 있다. [10]에서는 패턴인식에 있어 단일모델의 통합방법으로 다층퍼셉트론을 이용하고 있다. 통합모델은 3개의 단일모델로 구성되어 있으면 2개의 단일모델은 베이지안 모델이고 1개의 단일모델은 다층퍼셉트 모델로 이루어

져 있다.

본 논문에서는 4.1과 4.2에서 제시한 순차적 통합 방법과 병렬적 통합 방법은 선호도 예측 기준에 있어서 정적이고, 문제에 대해 적응성이 없다. 또한, 이들 방법은 기본적으로 선형적인 통합 방법으로서 통합 효과가 그다지 크지 않다. 따라서 [10]과 같이 단일 모델들(사용자 신경망 모델, 항목 신경망 모델)의 출력을 입력으로 하는 신경망 모델의 학습 과정을 적용하여 보다 유연하고 효율적인 통합 방법을 적용한다.

4.3.1 퍼셉트론(Perceptron)을 이용한 통합 방법

1957년 로젠블랫(Rosenblatt)에 의해 처음 소개된 퍼셉트론은 비교적 정확히 기술된 최초의 신경망 모델이었으며 여러 분야에 걸쳐 커다란 영향을 끼쳤다. 또한 바라는 출력을 생성하기 위해 가중치(weight)를 학습시킬 수 있으며, [그림 8]에서는 퍼셉트론의 대표적인 구조를 나타내고 있다. <식 4>는 [그림 8]에 표시된 각각의 기호를 정의하고 있다. 퍼셉트론의 학습 목적은 신경망이 학습을 통해서 실제 출력 O와 바라는 출력 D가 같아지는 것을 원한다. 입력 벡터와 바라는 출력 벡터가 주어지고, 가중치와 학습율을 초기화 한다. 초기화 과정을 수행하고 난 뒤에 실제 신경망의 출력을 계산한다. 이러한 계산과정에 대한 수행을 마친 뒤, 가중치를 갱신한다. 학습과정 동안 발생하는 에러 값을 누적시킨 후 모든 패턴의 입력벡터가 다 주어진 뒤에 바라는 에러범위에 들어오면 학습 알고리즘의 수행을 끝마친다. 만약, 바라는 에러범위에 들어오지 않을 경우 학습 알고리즘은 다시 수행된다.

$$\begin{aligned}
 \text{입력 벡터: } X &= [x_1, \dots, x_i, \dots, x_n]^T \in R^n \\
 \text{가중치: } W &= [w_{11}, w_{12}, \dots, w_i, \dots, w_{kn}] \in R^k \\
 \text{바라는 출력: } D &= [d_1, d_2, \dots, d_i, \dots, d_k]^T \in R^k \\
 \text{실제 출력: } O &= [o_1, o_2, \dots, o_i, \dots, o_k]^T \in R^k
 \end{aligned} \quad (4)$$

퍼셉트론은 선형 분리 가능한 패턴들만을 분류할 수가 있다. 선형 분리 가능이란 데이터 클래스가 하나의 직선에 의해 두 개의 영역으로 나뉘어지는 것을 말한다. [그림 9]의 (B)는 입력 공간에서 선호도의 분할 영역이

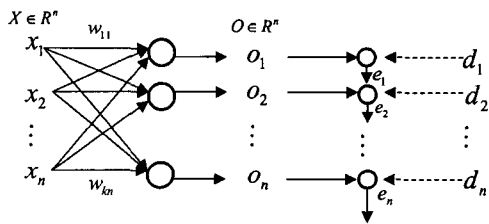


그림 8 퍼셉트론의 구조

퍼셉트론에 의해 입력 데이터를 선형 분리하는 것을 보여주고 있다. 다시 말하면, 두 모델의 출력 값에 따라 적절히 선형 분리함으로써 통합 모델에 의한 정확도를 향상시킬 수 있다. 그러나 선형 분리 가능한 데이터만을 분리할 수 있으므로 복잡한 경계선은 분리할 수 없다. 본 논문에서는 초기 가중치는 [-0.5, 0.5], 학습율은 0.05로 초기화하여 학습한다.

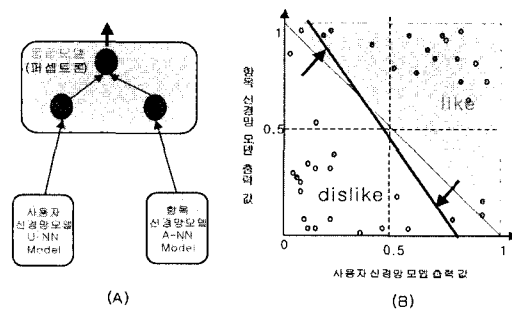


그림 9 퍼셉트론을 이용한 통합 방법과 선호도 경계선

4.3.2 다층 퍼셉트론을 이용한 통합 방법

다층 퍼셉트론은 단층 퍼셉트론과 유사한 구조를 가지고 있지만 중간층과 각 노드의 입출력 특성을 비선형으로 함으로써 네트워크의 능력을 향상시켜 단층 퍼셉트론의 여러 가지 단점들을 극복할 수 있다. 다층 퍼셉트론에서 층의 개수가 증가할수록 퍼셉트론이 형성하는 결정 구역의 특성은 더욱 고급화된다. 따라서 문제에 대한 적응성을 높이기 위해 사용자와 항목 신경망 모델의 출력을 입력으로 하는 다층 퍼셉트론을 이용한 통합 방법을 제안한다. 다층 퍼셉트론은 임의의 복잡한 경계선을 근사할 수 있으므로 보다 성능이 향상된 통합 효과를 기대할 수 있다. [그림 10]의 (B)는 퍼셉트론을 이용한 통합 방법보다 다층 퍼셉트론에 의해 더욱 정교한 경계

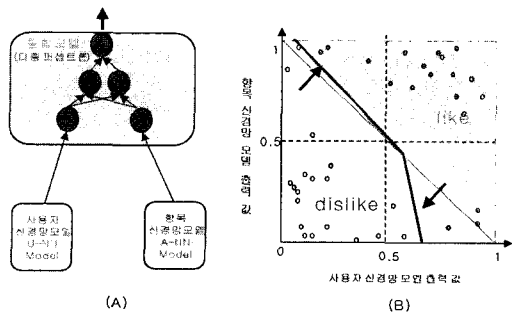


그림 10 다층 퍼셉트론(MLP)을 이용한 통합 방법과 선호도 경계선

계선으로 분리하는 것을 알 수 있다. 본 논문에서는 두 개의 은닉 노드를 갖는 다층 퍼셉트론으로 초기 가중치는 [-0.5, 0.5], 학습율은 0.05로 초기화하여 학습한다.

4.4 퍼지 추론을 이용한 통합 방법

퍼지이론은 현상의 불확실한 상태를 그대로 표현해주는 방법으로 1965년 미국 버클리 대학의 자데(Lofti A. Zadeh)교수에 의해서 처음 소개되었다. 퍼지이론은 애매하게 표현된 자료를 우리에게 유용한 자료로 만들기 위하여 퍼지집합(fuzzy set), 퍼지논리(fuzzy logic), 퍼지숫자(fuzzy number)등의 개념을 포함하고 있으며 수학적인 계산방법도 잘 개발되어 있다. 각 퍼지 규칙은 <식 5>와 같이 다수 개의 퍼지 집합으로 결합되어 있는 조건절과 결론을 가지며 각 규칙에는 규칙에 대한 확신도(CF : Certainty Factor)을 가지고 있다.

$$\text{규칙 } i: \text{if } a_{i1} \text{ is } U_{i1} \text{ and } a_{i2} \text{ is } U_{i2} \text{ and } a_{in} \text{ is } U_{in} \text{ then Class } k \text{ (CF}_i) \quad (5)$$

<식 5>에서 a_{ik} 는 I번째 규칙의 k번째 속성을 나타내며 U_{ik} 는 속성 a_{ik} 의 속성 값을 나타내는 소속함수를 의미한다.

즉, 주어진 데이터 $X = (x_1, x_2, \dots, x_n)$ 에 대하여 각 클래스에 대한 확신도는 <식 6>과 같이 계산된다.

$$\text{Conf}_k(X) = \max_{i \in R(k)} (\min_j (\mu_{U_{ij}}(x_j)) \cdot F_i) \quad (6)$$

여기서 $R(k)$ 는 결론부가 'Class is k'를 나타내는 모든 규칙들을 의미하고, $\mu_{U_{ij}}(x_j)$ 은 i번째 규칙에서 j번째 속성에 대한 입력 값(x_j)의 소속정도를 나타내고 있다. 따라서 $\text{Conf}_k(X)$ 가 가장 큰 클래스를 최종 결론으로 한다.

4.3절에서 기술한 신경망을 이용한 통합 방법은 두 모델의 결과 값을 바탕으로 학습을 통해 최종 결과 값의 정확도를 향상시키는 반면, 퍼지 규칙 방법은 퍼지 추론을 통해 정확도를 향상시킬 수 있다. 본 논문에서는 퍼지 소속 함수로 각 신경망 모델의 출력에 대해 세 개의 퍼지 집합(Dislike, Middle, Like)으로 나타내며, <표 2>의 퍼지 추론연산을 적용한다. 각 모델의 출력 값 영역은 [그림 11]과 같이 분할되고, 각 영역에 대응되는 퍼지 규칙들은 <표 3>과 같이 생성되고 퍼지 추론을 통하여 최종 결과 값을 산출함으로써 정확도를 향상시킨다.

표 2 규칙에 적용되는 추론연산

| | |
|---------|--------------------------------------------|
| min | 각 조건절의 만족도를 합성하는 연산 |
| product | 각 규칙의 조건부의 만족도를 합성한 결과와 그 규칙의 확신도를 합성하는 연산 |
| max | 각 규칙의 결과를 합성하는 연산 |

표 3 퍼지규칙

1. if User_Model is dislike and Movie_Model is dislike then dislike (CF:0.89)
2. if User_Model is like and Movie_Model is dislike then like (CF:0.88)
3. if User_Model is dislike and Movie_Model is like then like (CF:0.75)
4. if User_Model is like and Movie_Model is like then like (CF:0.93)
5. if User_Model is middle and Movie_Model is like then like (CF:0.67)
6. if User_Model is middle and Movie_Model is dislike then dislike (CF:0.67)
7. if User_Model is middle and Movie_Model is middle then like (CF:0.63)
8. if User_Model is dislike and Movie_Model is middle then dislike (CF:0.6)
9. if User_Model is like and Movie_Model is middle then like (CF:0.91)

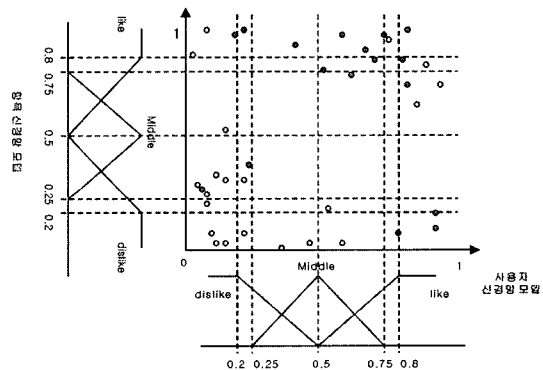


그림 11 퍼지규칙에 의한 분할 영역

4.5 BKS를 이용한 통합 방법

본 논문에서는 BKS 행동 지식 공간을 [그림 4]와 같이 첫 번째 전문가를 사용자 신경망 모델로 두 번째 전문가를 항목 신경망 모델로 정의하며 셀은 BKS(dislike,dislike), BKS(like,dislike), BKS(dislike, like), BKS(like,like)와 같이 네 개로 구성된다. 여기서 BKS(like,like) 셀을 살펴보면 <식 7>과 같다.

$$\begin{cases} n_{\text{like, like}}(\text{like}) = 11, \\ n_{\text{like, like}}(\text{dislike}) = 2, \end{cases} \quad (7)$$

따라서 BKS(like,like) 셀에서 대표의견(Rlike,like)은 like로 정의할 수 있다.

5. 실험

EachMovie[11] 데이터는 최근 추천시스템의 타당성

을 검증하기 위해 많이 사용되는 데이터로서 72,916명의 사용자와 1,628편의 영화로 구성되어 있다. 사용자가 본 영화에 대한 선호도(rating) 정보는 0(dislike), 0.2, 0.4, 0.6, 0.8, 1.0(like)의 수치로 표현되어져 있고, 각 사용자의 나이와 성별, 각 영화의 장르 정보가 포함되어 있다. 따라서 제안하는 통합 방법들의 성능을 확인하기 위해 EachMovie데이터를 적용한 기존 추천 기술의 성능과 비교 평가한다. 본 실험은 기존의 추천기술인 연관규칙과 제안한 통합 모델의 성능을 비교하기 위해 [3]에서 사용된 데이터를 적용한다. 즉, 영화 100편 이상 본 사용자 1000명을 모델 생성을 위해 학습 데이터로 생성하고, 사용자 아이디가 70000번 이상인 사용자 중 영화 100편 이상 본 사용자 100명을 모델 평가를 위한 테스트 데이터로 생성한다. 생성된 학습데이터와 테스트 데이터를 이용하여 통합모델을 위한 단일 신경망 모델을 생성한다. 본 실험에서는 사용자 아이디 70000번 이상의 사용자 30명과 영화 30편을 임의적으로 선택하여 총 60개의 단일 신경망 모델을 생성하였다.

5.1 신경망 모델의 매개 변수

단일 신경망 모델의 학습 매개 변수는 <표 4>에서 보는 바와 같은 조건으로 모델을 생성한다. 일반적으로 신경망을 학습할 경우 MSE(Mean Square Error)가 0에 충분히 수렴할 때까지 학습시키는 것이 보통이다. 그러나 협력적 추천의 경우 사용자들 또는 항목들 간의 선호도의 평균적 경향을 학습하는 것이 중요하며 소수 특정 데이터에 대하여 정확히 학습시킬 필요는 없다. 왜냐하면 어떤 특정 선호도 정보가 절대적이지 않기 때문이다.

실험데이터는 <표 4>와 같이 [4]에서 가장 높은 인식율을 보인 방법으로 정량화 한다. 또한 생성된 사용자와 항목 신경망 모델로 통합 모델을 생성하여 제안한 통합방법들에 대한 성능을 확인한다.

표 4 사용자와 항목 신경망 모델 학습 매개 변수

| | | | | | |
|----------|------------|-----|-----|-----------|-------------|
| 매개 변수 | 값 | | | | |
| 학습 종료 조건 | MSE ≒ 0.04 | | | | |
| 초기 가중치 | -0.5 ~ 0.5 | | | | |
| 학습율 | 0.05 | | | | |
| 입력노드수 | 100 | | | | |
| 은닉노드수 | 5 | | | | |
| 선호도 | 0.0 | 0.2 | 0.4 | no rating | 0.6 0.8 1.0 |
| 정량화된 선호도 | -1 | 0 | | | 1 |

5.2 실험 결과

각각 30개의 사용자와 항목(영화) 신경망 모델들 중

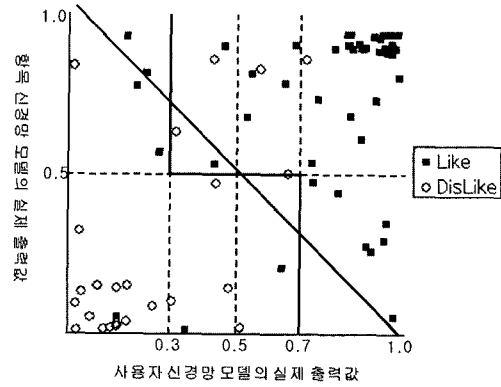


그림 12 통합 모델의 데이터 분포

사용자와 관련된 항목(영화) 신경망 모델과 통합하여 총 59개의 통합모델을 생성하였다. 통합모델에 대한 데이터 분포는 [그림 12]와 같으며 두 모델의 선호도 동의 영역에 데이터가 많이 위치하고 있음을 알 수 있다.

순차적 통합 방법에서는 우선, 사용자 신경망 모델의 실제 출력 값이 0.7보다 크면 like로 선호도를 판정하고, 0.3보다 작으면 dislike로 선호도를 판정한다. 만약 사용자 신경망 모델의 실제 출력 값이 0.3과 0.7사이의 출력 값이라면 항목 신경망 모델로 선호도를 판정한다. 병렬적 통합 방법에서는 두 신경망 모델의 실제 출력 값 중에서 가장 확실한 선호도로 판정하게 된다. 신경망(퍼셉트론과 다층 퍼셉트론)을 이용한 통합 방법에서는 사용자와 항목 신경망 모델의 실제 출력 값을 입력으로 하여 통합을 위한 신경망 모델을 생성한다.

<표 5>에서 보는 바와 같이 단일 신경망 모델만을 사용하는 추천 기술보다 사용자와 항목 신경망 모델을 통합한 통합모델이 추천 기술의 성능이 높음을 알 수 있다. 그 이유는 사용자 신경망 모델은 사용자간의 연관성으로, 항목 신경망 모델은 항목간의 연관성으로 선호도를 예측한다. 즉, 두 개의 단일 신경망 모델이 서로 다른 관점에서 선호도를 예측하기 때문에 두 개의 단일 신경망 모델의 선호도를 통합함으로써 보다 성능이 우

표 5 각 통합 방법의 성능 비교

| 연관 규칙 | 단일 신경망 | 단일 신경망 통합 모델 (사용자와 항목 신경망 모델의 통합) | | | | | | | |
|---------|---------|-----------------------------------|------|------|------|------|------|------|------|
| | | 항목 | 사용자 | 순차적 | 병렬적 | 신경망 | | BKS | 퍼지추론 |
| 퍼셉트론 | 다층 퍼셉트론 | | | | | | | | |
| 인식률 (%) | 71.7 | 78.9 | 84.2 | 84.2 | 88.2 | 88.2 | 88.2 | 86.8 | 86.8 |

수합을 알 수 있다. 통합 방법에 있어 큰 차이는 없었으나 유연성과 적응성이 높은 신경망(퍼셉트론, 다층 퍼셉트론)을 이용한 통합 방법이 일반적으로 성능이 높다는 것을 알 수 있다. 사용자와 항목을 혼합한 연관 규칙 추천 기술[3]과 비교하여도 본 논문에서 제안하고 있는 여러 가지 통합 방법들이 추천 기술에 있어서 우수함을 알 수 있다.

6. 결론 및 향후 연구

추천 시스템에 있어서 신경망을 이용한 기술은 최근에 그 성능의 우수함이 입증되고 있다.[4,12] 본 논문에서는 다른 관점을 갖는 사용자와 항목 신경망 모델을 통합함으로써 단일 신경망 모델이나 연관규칙을 이용한 추천기술보다 성능이 향상됨을 보이고 있다. 모델을 통합하기 위한 방법으로 기존 CME방법 중 BKS방법을 적용하여 본 논문에서 적용하고 있는 통합방법과 비교하였다. 적용하고 있는 통합방법은 생성된 모델의 우선순위를 고려한 순차적 통합방법과 생성된 모델을 동시에 고려하는 병렬적 통합 방법이 있으며 이러한 방법들은 통합하기가 쉬운 반면 선호도 예측 기준에 있어서 정적이므로 문제에 대한 적응성이 결여되는 단점을 가지고 있다. 이러한 단점을 보완하기 위해 본 논문에서는 신경망과 퍼지 추론을 이용한 통합방법을 적용하였다. 신경망 통합 방법 중 퍼셉트론을 이용한 통합 방법은 선형적인 선호도 예측 기준만을 생성함으로써 성능 향상에 대한 한계점을 보인다. 반면 다층 퍼셉트론의 경우 선형적인 선호도 예측 기준 뿐만 아니라 비선형적인 선호도 예측 기준도 생성할 수 있어 퍼셉트론의 단점을 보완하고 있다. 이상 통합방법들은 단일 모델에 대한 성능보다 향상됨을 알 수 있었다.

향후 연구로는 기존의 CME 방법 중 퍼지를 이용한 CME방법을 적용해보고자 한다. 두 신경망 모델의 실제 출력 값은 Like 혹은 Dislike의 선호 정도이기 때문에 이러한 선호도는 퍼지를 이용한 CME 방법을 적용함으로써 성능 향상을 기대할 수 있다. 또한 본 논문에서는 사용자와 항목 신경망 모델만을 통합하여 그 결과를 비교하였으나, 다른 관점을 갖는 여러 신경망 모델 혹은 다른 것을 이용한 모델 예를 들면 퍼지를 이용한 CME 방법 또는 연관규칙 등을 이용한 모델들을 개발하여 통합하는 데에 적용해 보고자 한다.

참 고 문 헌

- [1] Michael J. Pazzani, "A Framework for Collaborative, Content-Based and Demographic Filtering", *Artificial Intelligence Review* 13(5-6), pp.393-408, 1999.
- [2] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl, "Item-based Collaborative Filtering Recommendation Algorithms", *In Proc. of the 10th International World Wide Web Conference*, pp.285-295, 2001.
- [3] Weiyang Lin, Sergio A. Alvarez, Carolina Ruiz, "Collaborative Recommendation via Adaptive Association Rule Mining", *WebKDD-2000*, Boston, MA, Aug, 2000.
- [4] 김종수, 도영아, 류정우, 김명원, "신경망을 이용한 추천시스템의 성능 향상". *한국뇌학회지*, Vol.1, No.2, pp.223-244, 12월, 2001.
- [5] Y. S. Huang and C. Y. Suen, "A Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.17, pp.90-94, 1995.
- [6] C. A. O. Vieira and P. M. Mather, "A Comparative Study of Multiple Classifier Combination Methods in Remote Sensing", *In Proceedings of the IC-AI'2000*, Vol.1, pp.39-46, June, 2000.
- [7] Tin Kam Ho, Jonathan J. Hull and Sargur N. Srihari, "On Multiple Classifier Systems for Pattern Recognition", *Proc. of the 11th IAPR International Conference on Pattern Recognition*, pp.84-87, September, 1992
- [8] Merijn Van Erp and Lambert Schomaker, "Variants of The Borda Count Method for Combining Ranked Classifier Hypotheses", *Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition*, pp.443-453, 2000.
- [9] A. Khotanzad and C. Chung, "Hand Written Digit Recognition using BKS Combination of Neural Network Classifiers", *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp.94-99, 1994.
- [10] Pierre LOONIS, El-Hadi ZAHZAH, Jean-Pierre BONNEFOY, "Multi-classifiers Neural Network Fusion versus Dempster-Shafer's orthogonal rule", *IEEE ICNN/ICEC & ANZIS*, 1995
- [11] P.McJones, Eachmovie collaborative filtering data set, <http://www.research.digital.com/SRC/eachmovie>, 1997, DEC Systems Research Center
- [12] D. Billsus and M. J. Pazzan, "Learning collaborative information filters", *In Proceedings of the Fifteenth International Conference on Machine Learning*, pp.6-54, Madison, WI, 1998.



도 영 아
 1999년 2월 강남대학교 전산학과 졸업(학사). 2002년 2월 숭실대학교대학원 컴퓨터학과 졸업(석사). 2002년 12월 M플러스텍(주) 연구원. 관심분야: 신경망, 유전자알고리즘, 데이터마이닝



김 중 수
 1997년 2월 서울산업대학교 전자계산학과 졸업(학사). 2002년 2월 숭실대학교대학원 컴퓨터학과 졸업(석사). 2002년 12월 M플러스텍(주) 연구원. 관심분야: 신경망, 유전자알고리즘, 데이터마이닝



류 정 우
 1998년 2월 숭실대학교 정보과학대학 인공지능학과 졸업(학사). 2000년 2월 숭실대학교대학원 컴퓨터학과 졸업(석사). 2000년 2월~현재 숭실대학교대학원 컴퓨터학과 박사과정. 관심분야: 신경망, 유전자알고리즘, 퍼지이론, 데이터마이닝, 에이전트



김 명 원
 1972년 서울대학교 응용수학과 졸업. 1981년 University of Massachusetts (Amherst) Computer Science 석사 학위 취득. 1986년 University of Texas (Austin) Computer Science 박사 학위 취득. 1975년~1978년 한국과학기술 연구소 연구원. 1982년~1985년 Institute for Computing Science & Computer Application (Univ. of Texas) 연구원. 1985년~1987년 AT&T Bell Labs. (Naperville) Member of Technical Staff. 1987년~1994년 한국전자통신연구소 책임연구원. 1991년~1993년 충남대학교 전자계산학과 겸임부교수. 2000년~2001년 미국 IBM T.J. Watson 연구소 방문과학자. 1994년~현재 숭실대학교 컴퓨터학부 교수. 2002년~현재 숭실대학교 정보지원처장. 1992년~1993년 한국신경회로망 연구회 회장. 1992년~1993년 한국정보과학회 뉴로 컴퓨팅 연구회 회장. 1993년~1995년 정보과학회 뉴로컴퓨팅 연구회 위원장. 1993년~1995년 IEEE Neural Network Council 한국지부장. 1998년~2000년 한국인지학회 부회장. 1997년~2000년 한국뇌학회 부회장. 2001년~2002년 한국뇌학회 회장. 관심분야 유연추론, 신경회로망, 퍼지시스템, 진화알고리즘, 패턴인식, 자동추론, 기계학습, 데이터마이닝, creativity engineering 등