대한산업공학회/한국경영과학회 2002 춘계공동학술대회
한국과학기술원(KAIST) 2002년 5월 3일~4일

자기연상 다층퍼셉트론의 이상 탐지 성능에 대한 실험

Experiments on the Novelty Detection Capability of Auto-Associative Multi-Layer Perceptron

이형주[1] • 황병호[2] • 조성준[1]
E-mail : {impatton@snu.ac.kr, hwangbh@lge.com, zoon@snu.ac.kr}

[1]서울대학교 공과 대학 산업공학과, 서울시 관악구 신림 9동 산56-1 151-744
[2]LG전자 Digital Media 연구소 Diznet그룹, 서울시 서초구 우면동 16 137-724

Abstract In novelty detection, one attempts to discriminate abnormal patterns from normal ones. Novelty detection is quite difficult since, unlike usual two class classification problems, only normal patterns are available for training. Auto-Associative Multi-Layer Perceptron (AAMLP) has been shown to provide a good performance based upon the property that novel patterns usually have larger auto-associative errors. In this paper, we give a mathematical analysis of 2-layer AAMLP's output characteristics and empirical results of 2-layer and 4-layer AAMLPs. Various activation functions such as linear, saturated linear and sigmoid are compared. The 2-layer AAMLPs cannot identify non-linear boundaries while the 4-layer ones can. When the data distribution is multi-modal, then an ensemble of AAMLPs, each of which is trained with pre-clustered data is required. This paper contributes to understanding of AAMLP networks and leads to practical recommendations regarding its use.

Keywords Novelty Detection, Auto-Associative Multi-Layer Perceptron, Principal Component Analysis, Non-linear Boundary

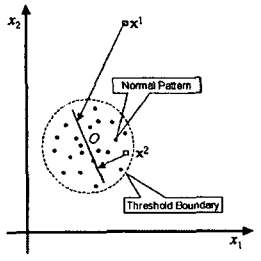
1. 서론

이상 탐지는 일반적이고 정상적인 것에서 벗어나는 이상 패턴을 구분해내는 작업이다. 컴퓨터 시스템 사용자 인증, 음성 식별, 위폐 식별과 같은 인증 문제나 전동기 모터의 오류 예측과 같은 모니터링 문제에 주로 적용된다(Frosini *et al.*, 1996; Obaidat & Sadoun, 1997; Ikbal *et al.*, 1999; Cho *et al.*, 2000a). 일반적인 이진 분류 문제에서는 두 가지 범주의 패턴을 모두 학습한 후, 새로운 데이터가 주어지면 두 가지 범주 중 하나로 구분한다. 그러나, 실제 문제에서는 한 쪽 범주에 속하는 패턴의 수가 극히 적거나 수집이 어려운 부분 노출 환경(partial exposure environment)인 경우가 많다. 그래서, 두 가지 범주 중에서 한 쪽의 패턴만 학습한 후(정상 패턴), 학습한 패턴과 성격이 다른 패턴을 구분하는(이상 패턴) 이상 탐지 방법을 적용하게 된다.

이상 탐지 방법으로는 신경망을 활용한 방법들이 사용되고 있다. 그 중에서도 자기연상 다층퍼셉트론(Auto-Associative Multi-Layer Perceptron ; AAMLP), SOM(Self-Organizing Map) 등이 주로 사용된다(Baldi & Hornik, 1989; Frosini *et al.*, 1996; Kaski & Lagus, 1996; Cho & Han, 2000b).

AAMLP는 학습 패턴의 입력 벡터가 목표 벡터로 사용되어 입력 벡터가 출력되도록, 즉 자기연상(auto-association)되도록 학습되는 MLP이다. 따라서 입력 노드 수와 출력 노드 수는 같고 은닉층 중에서 적어도 하나는 입력/출력 노드의 수보다 적은 수의 노드를 갖는다. 즉, 병목 구간(bottleneck)에서 차원이 축소되었다가 출력층에서 차원이 확장되는 형태를 띤다. 그래서 중복되는 정보는 압축하고 중복되지 않는 정보는 구분하는 특성(redundancy compression and non-redundancy differentiation)을 갖게 된다(Japkowicz *et al.*, 1995). 그리고, 보통의 MLP와 마찬가지로 오류역전파(error back-propagation) 방법을 사용하여 네트워의 가중치 벡터를 조정한다.

여기에서 입력 벡터와 출력 벡터의 유클리드 거리가 자기연상 오류가 된다. 이상 탐지에 AAMLP를 사용하는 근거는 학습 후의 네트워이 학습 패턴과 유사한 입력 벡터(정상 벡터)에 대해서는 낮은 자기연상 오류를, 학습 패턴과 유사하지 않은 입력 벡터(이상 벡터)에 대해서는 높은 오류를 낸다는 것이다(Bianchini *et al.*, 1995). [그림 1]은 AAMLP에 의한 이상 탐지의 예를 보여준다. 점선의 원 내부에 있는 점들이 학습에 이용된 정상 패턴이다. AAMLP는 R^2-공간 상의 모든 점을 원 내부에 있는 선분 위로 사영하고 그 때의 거리를 기준으로 정상/이상으로 분류한다.



[그림 1] AAMLP에 의한 이상 탐지의 예

이 논문은 과학기술부 뇌공학/뇌과학 연구사업, 2001년도 두뇌한국 21 사업 핵심분야 및 서울대학교 발전기금의 지원으로 수행되었음

즉, x^1은 공간 상에서 정상 패턴과 거리가 멀기 때문에 이상 패턴으로 분류되고, 거리가 가까운 x^2는 정상 패턴으로 분류된다.

기존의 연구에서 AAMLP의 은닉층에 선형 활성화 함수를 사용하면, 선형 주성분분석과 성능 면에서 차이가 없으며 국소 최소치(local minima)를 회피하고 전역적 최소치(global minima)에 도달할 수 있다는 점이 밝혀져 있다(Baldi & Hornik, 1989). 또한, 기존의 연구들에서는 AAMLP에 비선형 함수를 사용하면 선형 함수를 사용할 때보다 성능이 향상된다는 것을 실증적으로 보여주었다. 그러나, 이상 탐지에서 AAMLP의 학습 특성과 분류 성능에 있어서 비선형 은닉층의 영향에 대한 분석과 해석은 저자들이 아는 한 아직 이루어진 바가 없다. 본 연구에서는 AAMLP의 학습 특성과 그 한계에 대하여 알아보았다.

본 논문의 구성은 다음과 같다. 2장에서는 2층 AAMLP의 학습 특성에 대하여 설명한다. 3장에서는 2층 AAMLP의 한계에 대한 해결 방안으로서 4층 AAMLP의 학습 특성과 분류 성능을 실증적으로 분석한다. 4장에서는 본 연구의 결론을 요약·정리하고 연구의 한계와 향후 연구과제에 대하여 토의한다.

2. 2층 AAMLP의 학습 특성

2.1 2층 AAMLP

2층 AAMLP의 구조는 [그림 2]와 같다. 2층 AAMLP에서 입력층은 $l=0$, 은닉층은 $l=1$, 그리고 출력층은 $l=2$에 해당하며 l층에는 총 n_l개의 노드 $(1(l),\ldots,n(l))$가 있다. 학습에 이용되는 입력 벡터 $(x_1,\ldots,x_N)$는 모두 N개이며 각각의 입력 벡터 (x_k)에 대해 네트워크의 출력 벡터 (y_k)가 생성된다. 노드 $j(l-1)$과 노드 $i(l)$ 사이의 가중치는 $w_{i(l),j(l-1)}$이며 노드 $i(l)$ 상의 bias는 $w_{i(l)}$이다. 은닉층 활성화 함수 (f_1)는 출력값의 변역에 제한이 있는 비선형 함수인 시그모이드 함수이고 출력층 활성화 함수 (f_2)는 선형 함수인 항등 함수이다. 입력 벡터 (x_k)에 대해 노드 $i(l)$의 활성값 $a_{i(l)}(x_k)$은 다음과 같이 계산된다.

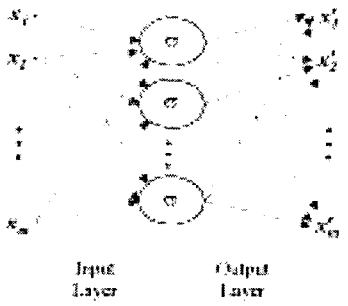
$$a_{i(l)}(x_k) = w_{i(l)} + \sum_{j=1}^{n_{l-1}} w_{i(l),j(l-1)} x_{j(l-1)}(x_k) \quad (1)$$

그리고 입력 벡터 x_k에 대한 노드 $i(l)$의 출력값 $x_{i(l)}(x_k)$은 활성값 $a_{i(l)}(x_k)$를 활성화 함수 f_l에 적용한 결과다.

$$x_{i(l)}(x_k) = f_l(a_{i(l)}(x_k)) \quad (2)$$

오류 함수 J는 학습 벡터의 평균제곱오차(mean of squared error, MSE)이다.

$$J = \frac{1}{N}\sum_{k=1}^{N} \| x_k - y_k \|^2 \quad (3)$$



[그림 2] 2층 AAMLP의 구조

2.2 2층 AAMLP의 특성

2층 AAMLP는 다음과 같은 특성을 지닌다.

① 다차원 평면의 존재 : 2층 AAMLP에서는 출력 벡터가 반드시 만족하는 선형 벡터 방정식이 정의된다. 선형 벡터 방정식은 다차원 평면을 정의하는데 모든 출력 벡터가 그 위에 위치하므로 이 다차원 평면을 출력 제한 다차원 평면이라 부른다.

② 출력 제한 다차원 평면의 유한성 : 은닉층 출력값의 변역이 제한되어 있으므로 출력층 출력값의 변역도 제한된다.

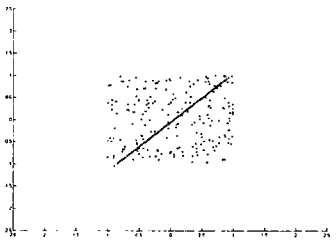
③ 학습 진행 방향 : 학습 벡터들과 출력 제한 다차원 평면은 공간 상에서 유클리드 거리가 가깝게 된다. 즉, $\overline{y_i}$는 학습 벡터 x_k를 출력 제한 다차원 평면으로 직교사영한 점이 된다.

④ 입력 공간 분할 : 출력 제한 다차원평면과 직교하는 평면 위에 위치하는 입력 벡터들은 네트워크 상에서 동일한 출력 벡터를 생성한다.

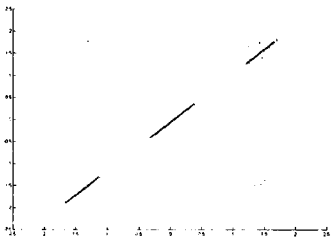
2층 AAMLP에서 학습 오류가 최소화되기 위해서 다음의 사항들이 만족되어야 한다.

- 학습 벡터 분할 원소들의 내부 산포도가 최소가 되어야 한다.
- 학습 벡터들과 출력 제한 다차원 평면이 공간 상에서 유클리드 거리가 가까워야 한다.

2층 AAMLP의 학습은 출력 제한 다차원 평면이 학습 벡터들과 공간적으로 유클리드 거리가 가까운 위치에 놓이도록 그 위상을 조정해 가는 과정으로 이해될 수 있다. 그 결과 네트워에 이상 패턴이 주어졌을 경우 그 출력 벡터와의 차이가 커서 자기 연상 오류가 크고, 정상 패턴이 주어졌을 경우 그 출력 벡터와의 차이가 작아서 자기 연상 오류가 작다. 그러므로 네트워크 상의 자기 연상 오류를 보고 특정 패턴이 정상 패턴인지 이상 패턴인지를 구분할 수 있다. 일정한 임계값을 정해 놓고 자기 연상 오류가 임계값 이하이면 정상 패턴으로, 임계값 이상이면 이상 패턴으로 분류한다. 이상의 출력 특성을 볼 때 2층 AAMLP가 이상 탐지에 있어 효과적인 해결 방안이 될 수 있음을 알 수 있다.

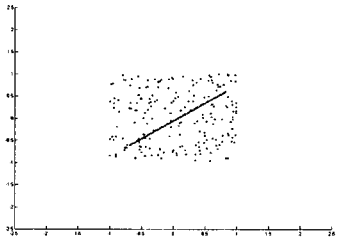


(a) 정상 학습 데이터

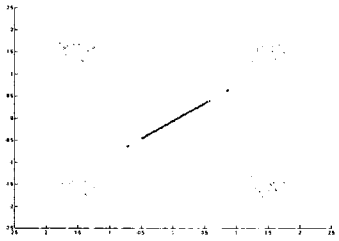


(b) 이상 테스트 데이터

[그림 3] 주성분분석의 학습 결과 (일양분포 데이터)



(a) 정상 학습 데이터



(b) 이상 테스트 데이터

[그림 4] 2층 AAMLP의 학습 결과 (일양분포 데이터)

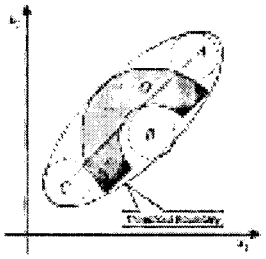
2.4 주성분분석과의 비교

주성분분석은 은닉층에 선형 활성화 함수를 사용하는 2층 AAMLP와 동일하다고 알려져 있다. 따라서, 2층 AAMLP와 주성분분석의 비교는 은닉층의 활성화 함수의 차이에 대한 비교라고 할 수 있다.

[그림 3]과 [그림 4]는 동일한 학습 벡터들로부터 학습된 두 모델이 생성한 다차원 평면의 차이를 보여주고 있다. 학습 데이터는 일양분포를 따르는 2차원 벡터들을 사용하였다. [그림 3]은 선형 활성화 함수가 은닉층에 사용되는 2층 AAMLP, 즉 주성분분석의 결과이다. 그리고, [그림 4]는 시그모이드 활성화 함수가 은닉층에 사용되는 2층 AAMLP의 결과이다. 각각의 그림에서 (a)는 정상 학습 패턴에 대한 네트워 출력 결과이고, (b)는 이상 테스트 패턴에 대한 출력 결과이다. 주성분분석의 경우에 무한 다차원 평면을 생성하기 때문에, 정상이 아닌 패턴에 대해서도 작은 오류를 낸다는 것을 알 수 있다[그림 3-(b)]. 반면, 2층 AAMLP의 경우에는 유한 다차원 평면을 생성하기 때문에, 주성분분석에서 발생할 수 있는 문제점을 피할 수 있다[그림 4-(b)]. 선형 함수를 사용하는 주성분분석에서는 네트워의 출력값을 제한하는 것이 불가능하기 때문에 [그림 3]과 같은 문제점이 발생할 수 있다. 2층 AAMLP에서는 (0,1) 범위의 출력만을 산출하는 시그모이드 함수를 사용한다. 따라서, AAMLP는 학습되면서 학습 벡터 영역 내부에 다차원 평면이 위치하도록 가중치를 조정해간다.

2.5 2층 AAMLP의 한계

2층 AAMLP에서 시그모이드는 출력값의 범위를 제한시킬 뿐 데이터의 비선형성을 반영할 수는 없다. 주성분분석과 마찬가지로 모든 입력 벡터 x_k에 대한 출력값 $\overline{y_i}$는 다차원 평면 위에 위치한다. 2층 AAMLP의 이러한 특성은 심각한 문제를 발생시킬 위험이 있다. [그림 5]에는 2층 AAMLP에서 발생할 수 있는 문제점의 예가 도시되어 있다. 어둡게 칠해진 활모양의 ***O*** 부분이 학습 데이터로 주어졌다. 2층 AAMLP는 그림과 같은 직선을 정의하고 타원형의 경계가 이상 패턴을 분류하는 임계값이 된다. 여기에서, 만약 ***A, B, C*** 부분이 이상 패턴으로 주어진다면,



[그림 5] 2층 AAMLP의 한계

AAMLP는 이 패턴들을 정상 패턴으로 오분류한다.

3. 4층 AAMLP의 학습 특성과 실험 결과

3.1 4층 AAMLP

[그림 5]와 같은 2층 AAMLP의 문제점 때문에 4층 AAMLP를 이상 탐지기로 사용한다. 4층 AAMLP는 입

력 패턴의 비선형성을 반영할 수 있다고 알려져 있다. 4층 AAMLP의 구조는 [그림 6]과 같다. Mapping layer와 de-mapping layer에는 시그모이드 활성화 함수를 사용하고, 나머지 층에는 선형 활성화 함수를 사용한다. 그리고, 학습 과정에서 mapping/de-mapping layer를 거치면서 "redundancy compression and non-redundancy differentiation" 현상이 일어난다. 이 때, 은닉층의 시그모이드 함수가 비선형 학습 데이터의 학습을 가능하

[그림 6] 4층 AAMLP의 구조

게 해 준다.

기존의 연구들에서 4층 AAMLP는 주로 차원 축소(dimensionality reduction)를 위한 NLPCA(Non-linear Principal Component Analysis)로서 사용되었다(Kramer, 1991). 입력층의 노드들은 각각 초기에 주어진 변수들을 의미하고, 병목 구간의 노드들은 NLPCA를 통해 추출된 변수들을 의미한다. 4층 AAMLP의 학습이 끝나면, 병목 구간의 출력값을 새롭게 생성된 변수들의 값으로 사용한다. 그러나, 4층 AAMLP가 비선형성을 완벽히 학습할 수 있는 것은 아니다. 기존의 연구들에서는 차원 축소 작업에서의 4층 AAMLP의 단점을 지적하였다(Reddy *et al.*, 1996; Malthouse, 1998).

① 최적의 사영이 결정되는 것이 아니기 때문에 큰 오류가 발생할 수 있다.
② 서로 교차하는 곡선 또는 곡면을 학습하지 못한다.
③ 불연속곡선 또는 곡면을 학습하지 못한다.
④ 내삽(interpolation) 성능은 좋지만, 외삽(extrapolation) 성능이 떨어진다.

여기에서, ④는 4층 AAMLP가 정상 패턴에 대해서는 작은 오류를 발생시키지만, 이상 패턴에 대한 오류는 크다는 것을 의미한다. 따라서, 충분한 수의 정상 패턴이 주어진다면, 이상 탐지 성능은 상당히 좋을 것이라고 기대할 수 있다. 따라서, 차원 축소 작업에서는 단점으로 지적된 4층 AAMLP의 특성이 이상 탐지 작업에서는 핵심적인 역할을 할 수 있다. 본 절에서는 다음과 같은 특성을 실험적으로 보임으로써 4층 AAMLP를 분석하였다.

- Saturated linear 함수를 은닉층에 사용한 4층 AAMLP는 학습 데이터 내부에 여러 개의 선분을 생성한다. 즉, 생성되는 선분의 수를 제외하면 2층 AAMLP의 학습 특성과 유사하다.
- 시그모이드 함수와 saturated linear 함수를 4층 AAMLP의 은닉층에 적용한 학습 결과는 유사하다.
- 시그모이드를 은닉층 활성화 함수로 사용하는 4층 AAMLP는 saturated linear 함수를 은닉층 활성화 함수로 사용하는 4층 AAMLP와 같이 비선형 데이터를 학습할 수 있다.

3.2 비선형 이상 탐지

4층 AAMLP가 2층 AAMLP보다 비선형 학습 데이터에 대한 성능이 우수하다는 사실은 실험적으로 알려져 있다. 그러나, 4층 AAMLP는 알고리듬의 복잡성 때문에 분석에 어려움이 많아 그 학습 특성과 이상 탐지 성능에 대한 명확한 설명은 되어 있지 않다. 본 연구에서는 실험적인 방법을 사용하여 4층 AAMLP가 어떻게 동작하는지에 대한 설명을 시도하였다.

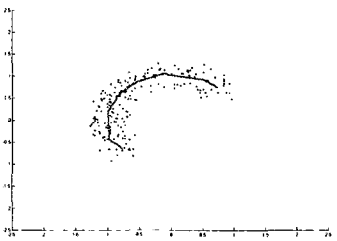
[그림 7]은 2.5에서 지적된 2층 AAMLP의 한계를 4층 AAMLP로 해결한 예를 보여준다. 4층 AAMLP의 출력값은 곡선형의 학습 데이터 내부에 위치한 곡선 또는 다차원곡면이 된다. 4층 AAMLP는 정상 패턴을 정확히 학습함을 알 수 있다. [그림 5]와 같은 이상 탐지 문제가 주어진다면 4층 AAMLP는 효과적인 이상 탐지기로 사용될 수 있다.

Saturated linear 함수를 사용한 4층 AAMLP을 동일한 데이터에 적용하였다. 정상 패턴을 학습한 결과는 [그림 8]과 같다. AAMLP의 출력값은 서로 연결된 선분들로 이루어진 출력 공간을 생성하였다. Saturated linear 함수는 식(5)와 같다.

$$f(x)=\begin{cases} 0 & (x\le 0) \\ x & (0<x<1) \\ 1 & (1\le x) \end{cases} \qquad (5)$$

[그림 8]에서 4층 AAMLP는 saturated linear 함수를

[그림 7] 시그모이드 함수를 은닉층에 사용한 4층 AAMLP의 학습 결과(비선형 데이터)



[그림 8] Saturated linear 함수를 은닉층에 사용한 4층 AAMLP의 학습 결과(비선형 데이터)

은닉층에 사용하였다. 따라서, 각각의 선분은 2층 AAMLP가 정의하는 직선과 동일하다. 각 선분에 사영되는 학습 데이터에 국한시켜 생각해 보자. 그러면, AAMLP는 학습 벡터 분할 원소 S_i들의 내부 산포도를 최소화시키면서, 학습 벡터들과 유클리드 거리가 가까워지도록 선분들의 위상을 조정한다. 따라서, 학습이 끝난 후 AAMLP가 생성하는 출력 제한 다차원 평면은 S_i에 속하는 학습 벡터들이 직교사영된 값이고, 그 결과 다차원 평면은 입력 벡터 내부에 위치한다. 2층 AAMLP는 하나의 선분으로 정상 학습 데이터를 학습하는 반면, 4층 AAMLP는 여러 개의 선분을 사용하여 정상 학습 데이터를 학습하는 것이다.

여기에서 눈여겨볼 점은 [그림 7]과 [그림 8]의 형태상 유사성이다. 시그모이드가 부드러운 곡선이고, saturated linear는 $x=0$과 $x=1$의 점에서 꺾인다는 점을 제외하면 두 함수는 비교적 유사한 형태를 지니고 있다. [그림 7]과 [그림 8]에서 AAMLP 출력 형태의 유사성도 이러한 활성화 함수의 특성에 기인한 것으로 보인다. Saturated linear 함수를 은닉층에 사용한 AAMLP의 출력은 시그모이드를 사용한 AAMLP의 출력의 근사값이며, 두 가지 AAMLP는 모두 정상 학습 벡터들의 자기연상 오류를 최소화시키는 방향으로 학습이 진행된다고 생각할 수 있다.

Saturated linear 함수를 은닉층에 사용했을 때 2층과 4층 AAMLP는 유사한 학습 특성을 보였고, 시그모이드 함수와 saturated linear 함수를 4층 AAMLP의 은닉층에 적용했을 때 유사한 학습 결과를 보였다. 따라서, 4층 AAMLP에서는 정상 학습 벡터들의 자기연상 오류를 최소화시키는 방향으로 학습이 진행되고, 출력 벡터는 정상 학습 데이터의 내부에 위치하며, 그 결과 4층 AAMLP는 비선형 데이터를 학습할 수 있다.

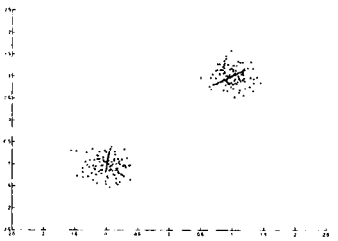
3.3 Multi-modal 데이터 학습

4층 AAMLP를 이상 탐지에 적용할 때 전제조건은 데이터의 분포가 uni-modal이어야 한다는 것이다. Multi-modal 분포에 4층 AAMLP를 적용할 때의 문제점은 [그림 9]에 잘 나타나 있다.

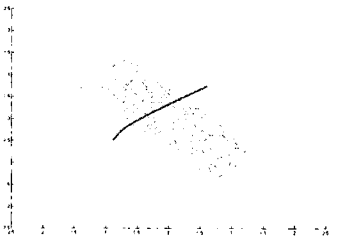
[그림 9]에서 (a)와 (b)는 각각 정상 학습 패턴과 이상 테스트 패턴에 대한 출력 결과이다. (a)에서 AAMLP는 학습 패턴들을 비교적 정확하게 학습하였다. 그러나, (b)에서 이상 패턴을 탐지하지 못하는 것을 볼 수 있다. AAMLP는 시그모이드를 은닉층 활성화 함수로 사용하기 때문에, 그 출력값은 연속함수의 형태로 나타날 수밖에 없다. 그 결과 multi-modal 분포의 학습 데이터를 학습했을 때에는 2개 분포 사이의 공간까지 임의로 "학습"되는 결과가 발생한다. 따라서, 학습 데이터를 정확히 학습했다 하더라도, 그 이외의 영역에 대하여는 상당한 위험이 존재한다.

실제 데이터가 multi-modal 분포를 가지고 있는 문제의 경우 이러한 AAMLP의 특성은 매우 큰 문제를 초래할 수 있다. 따라서, AAMLP를 적용하기 위해서는 "modality"에 대한 검증이 선행되어야 한다. 이러한 문제를 해결하기 위해 데이터의 modality를 추정하는 방법으로 군집화 기법을 사용할 수 있다. 군집화를 구현하는 방법으로는 EM 알고리듬, Gaussian kernel, SOM 등이 사용된다(Xu & Jordan, 1993; Bishop, 1994; Yin & Allinson, 2001).

[그림 9]의 학습 데이터가 두 개의 군집으로 이루어졌다는 사실을 이미 알고 있다면, 각각의 군집에 대하여 독립적인 4층 AAMLP를 학습시켜 앙상블 방법을 이용할 수 있다. 이것은 학습 벡터의 군집과 같은 수(P)개의 AAMLP를 학습시키는 것으로 일반화할 수 있다. 각각의 군집 p에 대하여 독립적으로 학습된

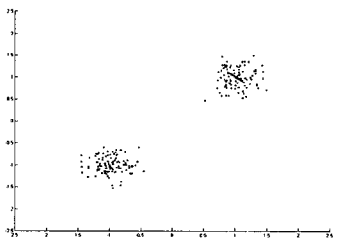


(a) 정상 학습 데이터

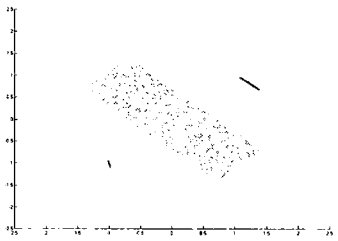


(b) 이상 테스트 데이터

[그림 9] 4층 AAMLP의 학습 결과 (Multi-modal 데이터)



(a) 정상 학습 데이터



(b) 이상 테스트 데이터

[그림 10] 4층 AAMLP 앙상블의 학습 결과 (Multi-modal 데이터)

AAMLP$_p$가 있다고 하자. 그리고, 입력 벡터 x에 대한 AAMLP$_p$의 출력 벡터를 y_p라고 하면, 입력 벡터 x에 대한 AAMLP 앙상블의 출력값 y_{com}는 다음과 같다.

$$y_{com} = \min_{y_p}[\ \| x - y_p \|^2\ ] \quad (p=1,\ldots,P) \quad (6)$$

즉, 새로운 입력 벡터가 주어지면 P개의 AAMLP$_p$로부터 출력값을 계산한 후, 그 중에서 가장 작은 오류를 발생시키는 AAMLP$_p$의 출력값을 앙상블의 출력값으로 사용하게 된다.

[그림 10]은 [그림 9]의 정상 학습 데이터를 두 개의 군집으로 나누어 독립적인 두 개의 4층 AAMLP를 학습시킨 결과이다. 정상 패턴에 대한 학습 결과 (a)는 [그림 9]와 큰 차이가 없다. 그러나 독립적인 두 개의 4층 AAMLP를 학습시켰기 때문에 출력값은 학습 데이터의 내부에만 위치하게 된다. 그 결과, (b)에서 이상 패턴들이 정상 패턴들과 같은 출력값들을 산출하고, 다라서 더 큰 오류를 발생시켰다. 따라서, 2개의 군집에 하나의 AAMLP를 적용했을 때보다 더 좋은 이상 탐지 성능을 가질 수 있다.

각 군집에 독립적인 AAMLP를 학습시켜 앙상블 방법을 사용했을 때가 단일 AAMLP를 사용했을 때보다 더 좋은 성능을 보임을 알 수 있다. 앙상블 방법은 정상 패턴에 대하여 더 작은 오류를 발생시키면서 이상 패턴에 대해서도 더 큰 오류를 발생시켰다. 실험 결과 효과적인 군집화 기법과 결합하여 modality를 정확히 추정한다면 AAMLP의 앙상블로서 이상 탐지 성능을 향상시킬 수 있음을 알 수 있었다.

4. 결론 및 향후 과제

이상 탐지는 일반적이고 정상적인 것에서 벗어나는 이상 패턴을 구분해내는 문제이다. 대개의 경우 이상 탐지에서는 정상 패턴만이 주어지고 이상 패턴이 주어지지 않는다. 하나의 클래스의 패턴만 주어지는 등의 제약 조건들로 인해 이상 탐지는 분류 문제 가운데에서도 어려운 문제에 속한다. AAMLP는 이러한 제약 조건들 하에서 효과적으로 이상 탐지를 해결할 수 있는 구조와 성질을 갖고 있다. 즉 정상 패턴만으로 AAMLP를 학습시켰을 때 학습에 이용되지 않은 이상 패턴은 정상 패턴, 즉 학습 벡터와의 차이가 커서 네트웍 상에서 자기 연상 오류가 크게 나타나므로 이 성질을 이용해 구분해낼 수 있다. 본 논문에서는 이상 탐지에 사용되는 2층, 4층 자기연상 다층퍼셉트론(AAMLP)의 성질과 문제점들을 고찰하였다. 결과를 정리하면 다음과 같다.

2층 AAMLP의 경우 출력 벡터의 공간 상의 위치를 제약하는 유한 출력 제한 다차원 평면이 존재한다. 오류 함수를 최소로 하는 2층 AAMLP의 학습은 출력 제한 다차원 평면이 공간 상에서 학습 벡터와의 유클리디언 거리가 가깝게 되는 방향으로 이루어진다. 출력 제한 다차원 평면이 학습 벡터와 공간 상에서 유클리디언 거리가 가까우므로 학습 벡터와 공간 상에서 멀리 떨어져 있는 입력 벡터는 네트웍 상의 출력 벡터와의 차이가 크게 된다. 즉 학습 벡터와 차이가 큰 패턴은 학습된 AAMLP에 입력으로 주어졌을 때 자기 연상 오류가 크게 나타난다.

2층 AAMLP가 생성하는 다차원 평면은 비선형 데이터를 학습할 수 없다는 단점이 있어 그에 대한 해결 방안으로 4층 AAMLP가 제시되었다. 4층 AAMLP는 비선형 데이터를 학습할 수 있다고 알려져 있지만, 모델의 비선형성 때문에 구체적인 학습 특성과 분류 성능에 대해서는 연구된 바가 없다. Saturated linear 함수를 사용한 4층 AAMLP는 여러 개의 선분 또는 유한 다차원 평면들의 집합으로 이루어진 출력값을 생성한다. 이러한 출력값은 시그모이드 함수를 은닉층에 사용한 4층 AAMLP의 학습 결과와 유사하며, saturated linear 함수를 은닉층에 사용한 2층 AAMLP의 학습 특성과 유사하다. 따라서, 시그모이드를 은닉층 활성화 함수로 사용한 4층 AAMLP는 2층 AAMLP와 마찬가지로 자기연상오류를 최소화시키는 방향으로 학습이 진행된다. 그 결과 출력 벡터는 정상 학습 데이터 내부에 위치함을 알 수 있다.

한편, 4층 AAMLP는 연속함수인 시그모이드 함수를 은닉층 활성화 함수로 사용하기 때문에, multi-modal 데이터에 적용할 수 없다는 한계점이 있다. 그에 대한 해결책으로 군집화 방법으로 modality를 추정하여, 정상 학습 데이터의 각 군집마다 개별적인 AAMLP를 학습시켜 앙상블 기법을 적용하는 방법을 생각할 수 있다. 여기에서는 군집화 방법의 성능이 이상 탐지 성능을 좌우하는 중요한 요인이 된다.

4층 AAMLP를 실제 문제에 활용하기 위해서는 몇 가지 문제점이 먼저 해결되어야 한다. 시그모이드를 은닉층 활성화 함수로 사용하면 전역적 최소치가 아닌 국소 최소치에 도달하는 경우가 있기 때문에, 4층 AAMLP에서는 입력 벡터가 반드시 유클리드 거리가 가장 가까운 출력 벡터로 출력되지는 않는 단점이 있다. 기존의 연구에 따르면 4층 AAMLP에서는 모든 입력 벡터가 항상 자기연상 오류를 최소화하고 출력 공간에서 유클리드 거리가 가까운 점으로 출력되는 것은 아니라는 점이 밝혀져 있다(Malthouse, 1998). 선형 주성분분석에서는 목적함수가 2차이기 때문에 최적화 문제를 풀면 항상 전역적 최소치에 도달한다. 반면, 4층 AAMLP에서는 목적함수가 지수함수의 역수인 비선형 형태이기 때문에, 최적화 문제를 풀면 전역 최소치가 아닌 국소 최소치에 도달하는 경우가 있다. 그 결과 4층 AAMLP에서는 입력 벡터가 출력 공간에서 더 멀리 떨어져 있는 점에 위치하는 "fanning" 현상이 일어난다. 그런데, AAMLP의 학습은 비교사학습(unsupervised learning)이기 때문에 국소 최적화를 피하기 위한 검증(validation) 작업이 어렵다. 그리고, 4층 AAMLP는 보통의 MLP보다 더 복잡한 구조를 가진다. 입력 벡터의 차원 수가 같다면 AAMLP는 더 많은 수의 가중치 벡터를 학습해야 하기 때문에, 학습이 오래 걸리고 계산 비용이 많이 든다는 단점이 있다.

산업체에서의 오류 발생을 예측할 수 있는 모니터링 시스템이나 허가받지 않거나 규약에 어긋나는 패턴을 구분해내는 인증 시스템, 그리고 특히 최근에 주목받고 있는 컴퓨터 시스템 및 네트웍 보안 등은 모

두 이상 탐지의 범주에 속하는 것들이다. 이러한 문제들에 대해 AAMLP는 좋은 대안이 될 수 있으며 다양한 응용의 모색이 가능하다.

4층 AAMLP의 학습 특성에 대하여 실증적인 결론을 내렸지만, Saturated linear 함수와 시그모이드 함수의 유사성을 근거로 검증되지 않은 가설을 도출해 내는데 그쳤다. 4층 AAMLP의 학습 진행 방향, 정의되는 다차원 곡면의 형태와 학습 후 출력 특성 등에 대한 분석이 이루어져야 한다. 그리고, 본 논문에서는 시각적인 분석을 위하여 인공적으로 생성된 2차원 데이터셋에만 AAMLP를 적용하였다. 실제 이상 탐지 문제에서는 더 높은 차원의 입력 벡터가 주어지기 때문에, 2차원 데이터에서 볼 수 없었던 많은 상황들이 발생할 수 있다. 향후에는 실제 문제에서 나타날 수 있는 여러 상황 속에서 AAMLP의 행동을 분석해서 그 취약점을 보완하는 방안에 대한 연구가 필요하다. 또한, 입력 벡터의 차원이 높아지면, AAMLP의 복잡도가 증가하고 그에 따라 분류 성능과 계산 비용이 영향을 받을 수 있다. 따라서, 주어진 이상 탐지 문제에 따라 AAMLP의 은닉층 노드 수를 결정하는 것도 연구 대상이 될 수 있다. 그리고, multi-modal 데이터를 학습하기 위한 여러 가지 clustering 방법들에 대한 연구도 필요하다.

참고문헌

Baldi, P. and Hornik, K. (1989) Neural Networks and Principal Component Analysis : Learning from Examples without Local Minima, *Neural Networks, 2,* 53-58.

Bianchini, M., Frasconi, P. and Gori, M. (1995) Learning in Multi-layered Networks used as Autoassociators, *IEEE Transactions on. Neural Networks, 6(2)*, 512-515.

Bishop, C.M. (1994) Novelty Detection and Neural Network Validation, *IEE Proceedings of Visual Image Signal Process, 141(4)*, 217-222

Cho, S., Han, C., Han, D. and Kim, H. (2000a) Web Based Keystroke Dynamics Identity Verification using Neural Networks, *Journal of Organizational Computing and Electronic Commerce, 10(4)*, 295-307

Cho, S. and Han, D., (2000b) Apparatus for Authenticating an Individual Based on a Typing Pattern by Using a Neural Network System, *Patent No. 6,151,593, US Patent and Trademark Office, Washington DC 20231, USA*, Granted on Nov. 21, 2000

Frosini, A., Gori, M. and Priami, P. (1996) A Neural Network-based Model for Paper Currency Recognition and Verification, *IEEE Transactions on Neural Networks 7(6)*, 1482-1490

Ikbal, M., Misra, H. and Yegnanarayana, B. (1999) Analysis of Autoassociative Mapping Neural Networks, *Proceedings of International Joint Conference on Neural Networks,* #854

Japkowicz, N., Myers, C. and Gluck, M. (1995) A Novelty Detection Approach to Classification, *Proceedings of the Fourteenth International Conference on Artificial Intelligence*, 518-523

Kaski, S. & Lagus, K. (1996) Comparing Self-organizing Maps, *Proceedings of International Conference on Artificial Neural Networks*, 809-814

Kramer, M. (1991) Nonlinear Principal Component Analysis using Autoassociative Neural Networks, *AIChE Journal, 37(2)*, 233-243

Malthouse, E.C., (1998) Limitations of Non-linear PCA as Performed with Generic Neural Networks, *IEEE Transactions on Neural Networks, 9(1)*, 165-173

Obaidat, M.S. & Sadoun, B. (1997) Verification of Computer Users using Keystroke Dynamics, *IEEE Transactions on Systems, Man, and Cybernetics, 27(2)*, 261-269

Reddy, V., Riley, P. and Mavrovouniotis, M. (1996) Analysis of Plant Measurements through Input-training Neural Networks, *Computers Chemical Engineering, 20., Suppl.*, 889-894

Xu, L. and Jordan, M.I. (1993) Unsupervised Learning by EM Algorithm based on Finite Mixture of Gaussians, *Proceedings of World Congress of Neural Networks*, 431-434

Yin, H. and Allinson, N.M. (2001) Self-organizing Mixture Networks for Probability Density Estimation, *IEEE Transactions on Neural Networks, 12(2)*, 405-411