

신용카드 사기 검출을 위한 신경망 분류기의 진화 학습

Evolutionary Learning of Neural Networks Classifiers for Credit Card Fraud Detection

박래정

Lae-Jeong Park

강릉대학교 정보전자공학부

요약

본 논문은 신용카드 사기 검출을 위한 신경망 분류기를 학습하기 위한 효과적인 방식을 제시한다. 제안하는 방식은 신경망 분류기의 전통적인 학습 방식인 출력 오차(MSE)를 기준으로 하지 않고, 진화연산 기법을 사용하여 분류기의 검출율/오검출율을 기준으로, 국소적 영역에서 성능을 최대화하는 학습 방식으로서, 높은 성능뿐만 아니라 실제 사기 검출 시스템 적용 시에 요구되는 다양한 평가 기준이나 조건들을 고려할 수 있는 장점이 있다. 실제 신용카드 거래 데이터에서의 실험을 통하여 제안한 방식이 검출율과 검출효율 및 상대적인 이익측면에서도 만족할 만한 성능을 나타냄을 보인다.

Abstract

This paper addresses an effective approach of training neural networks classifiers for credit card fraud detection. The proposed approach uses evolutionary programming to train the neural networks classifiers based on maximization of the detection rate of fraudulent usages on some ranges of the rejection rate, not minimization of mean square error(MSE) that is a common criterion for neural networks learning. This approach enables us to get classifiers of satisfactory performance and to offer a directive method of handling various conditions and performance measures that are required for real fraud detection applications in the classifier training step. The experimental results on "real" credit card transaction data indicate that the proposed approach produces classifiers of high quality in terms of a relative profit as well as detection rate and efficiency.

Key Words : fraud detection, neural networks, pattern classification, evolutionary programming.

1. 서론

최근 온라인 상거래가 급증하고, 사용의 편리성 및 혜택으로 인해 신용카드의 사용이 급증하고 있는 추세이다. 하지만, 이에 따른 다양한 형태의 부정사용도 증가하고 있다. 분실이나 도난 등으로 인한 부정사용은 그 중의 하나로서, 현재 국내외의 여러 신용카드회사나 은행들은 이러한 사기(fraudulent) 거래로 인한 경제적 손실을 줄이기 위한 사기거래 조기 검출 시스템을 개발하여 운용하고 있거나 도입단계에 있다. 따라서, 효과적인 사기 검출 시스템을 개발하려는 많은 연구들이 있어왔으며, 현재 HNC software 등의 외국 솔루션 회사에서는 상용 사기 검출 시스템을 판매하고 있다. 한편, 기계 학습(machine learning)과 데이터 마이닝(data mining) 연구 분야에서, 사기 검출 알고리즘에서의 여러 기술적인 문제에 대한 연구 결과가 발표되고 있다[1-6].

일반적인 패턴 분류기(classifier) 디자인과 달리, 사기검출 분류기 학습 및 디자인에 있어서 크게 다음과 같은 기술적인 어려움이 존재한다.

- 사기거래 수가 정상(legitimate) 거래 수에 비해 월등히 적으며 또한 데이터 양이 엄청나게 많다. 즉, 클래스 분포가 매우 불균형적(imbalance)이며, 보통 비율이 약 $10^{-3} \sim 10^{-1}$ 정도이다[2,3]. 이러한 불균형적인 클래스 분포로 인해 분류기의 학습을 위한 학습 데이터 결정이 용이하지 않은 문제점이 있다[7].
- 둘째로, 사기와 정상거래가 구분하기 힘든 유사한 패턴을 가지므로 클래스 분포가 상당히 겹쳐져(overlapping) 있다. 따라서, 분류의 정확도(accuracy)를 최대한 하도록 분류기의 분류 경계선(decision boundary)을 형성하기 위한 학습 과정이 일반적인 경우처럼 용이하지 않다.
- 마지막으로, 각 거래의 오분류 비용(misclassification cost)이 일정하지 않다. 정상(사기) 거래를 사기(정상) 거래로 분류할 때의 비용이 일정한 고정값을 가지는 것이 아니고, 거래의 승인 요청 금액에 따라 달라지므로 각 거래마다의 비용을 고려하여야 한다.

기존의 관련 연구를 살펴보면, Fawcett와 Provost는 휴대전화의 사기 검출 문제에서 사용자의 규칙적인 이용 패턴을 기준으로 IF-THEN 규칙들을 추출하고, 이 규칙들과 일치하지 않는 정도를 입력으로 하는 선형 분류기를 학습하는 방법을 사용하였다[1]. Dorrnsoro 등은 신용카드 사기 검출에서 불균형적이고 겹쳐있는 클래스 분포로 인한 신경망 분류기

접수일자 : 2001년 5월 25일

완료일자 : 2001년 9월 18일

이 논문은 2001년도 두뇌한국 21사업에 의하여 지원되었음.

학습시의 문제를 해결하기 위하여 비선형 discriminant analysis 기법을 기반으로 하는 학습방법을 제안하였다[2]. Park 등은 신용카드 사기 검출 문제에서 불균형적인 클래스 분포로 인한 오검출을 효과적으로 줄이기 위해 신경망과 결정 트리(decision tree)의 하이브리드 기법을 제안하였으며 [3], Chan 등은 각기 다른 학습 데이터를 이용해서 생성된 복수개의 선형 분류기들을 meta 학습 기법을 이용하여 결합하는 방법을 사용하였다[4,5]. 그들은 이 방법을 확장하여 신용카드 사기 검출의 비용 모델을 도입하고, 분류에 따른 비용을 최소화하는 학습 데이터내의 클래스 분포에 대해 고찰하였다[6].

살펴보았듯이, 대부분의 연구에서는 결정 트리 형태의 선형 분류기를 사용하여 사기 검출에 접근하였다. 하지만, 보다 높은 성능(높은 사기 검출율(true positive rate)과 낮은 오검출율(false positive rate))을 위해서 대부분의 상용 시스템에서는 신경망 분류기를 사용하고 있다[2,3]. 이는 겹쳐져 있는 클래스 분포 상에서는 비선형적인 분류 경계선이 훨씬 효과적인 분류를 가능하게 하며 또한 신경망의 출력값에 문턱값(threshold)을 기준으로 분류를 결정함으로써 검출율의 정도를 연속적으로 조절할 수 있다는 장점이 있기 때문이다.

일반적으로 신경망을 학습할 때에 각 입력에 대한 목표 출력 (여기서는 -1/1)을 정하고 신경망의 출력과 목표 출력값의 차이, 즉 오차(mean square error, MSE)가 최소가 되도록 한다. 하지만, 사기 검출 문제에서는 클래스가 상당히 겹쳐있으므로 MSE가 0이 될 수 없으며, 대부분이 (-1,1)사이의 출력값을 가지게 된다. 따라서, 실제 사용 시에는 적절한 성능을 만족하도록, 즉, 오검출율과 검출율의 적절한 범위 내에서 문턱값을 결정하여 사용하게 된다. 문턱값에 따라서 사기 검출 분류기의 성능이 달라지게 되고, 이는 receiver operating characteristics(ROC) 곡선으로 명확히 표현되어 진다[8]. 실제로 신용카드 사기 검출 문제에서는 허용 가능한 오검출율 범위가 존재하므로, 이 범위 내에서의 분류기의 성능이 중요하며, 그 범위 밖의 성능은 관심 밖이라 할 수 있다. 이런 측면에서 MSE를 기준으로 신경망 분류기를 학습하는 경우는, 관심 범위 내에서의 분류기의 성능을 최대화하기 위한 직접적인 방법은 아니라고 할 수 있다. 다시 말해서 모든 입력 패턴의 오차를 최소화하는 것이 특정 문턱값에서의 성능 최대화와 정확히 일대일 대응하지 않는다. Dorronsoro 등이 사용한, Fisher의 선형 discriminant[9]를 기준으로 하여 정상과 사기거래의 구분정도를 최대화하는 학습 방법도 목표 출력값을 사용하지 않으나, 특정 영역에서의 성능을 선택적으로 최대화하기에는 적절하지 않다. 실제로 신용카드 사기 검출 시스템에 사용될 최적의 신경망 분류기를 특정 문턱값에서의 성능을 기준으로 학습/선택하기 위한 효과적이며 직접적인 방법이 필요하다 하겠다.

필드에서 운용시에 발생하는 실제적인 문제중의 하나로서, 입력되는 거래의 사기와 정상 거래의 비율이 기간에 따라서 변하거나 학습 시와는 다른 거래 패턴이 입력되는 경우가 발생하게 된다. 이런 경우에는 학습시 예상한 성능과 차이가 나게된다. 이에 대해서 신경망 분류기의 재학습이 좋은 대책이지만, 실제 운용되고 있는 상황에서는 한시적이지만 그에 따른 성능을 조정하기 위하여 기설정된 문턱값을 변경하여야 하는 경우가 발생한다. 그러므로, 신경망 분류기의 학습 시에 이러한 입력 환경의 변화를 감안하여 평가하여야 할 필요가 있다. 따라서, 신경망 분류기를 학습할 때에 일정 동작점(특정 문턱값에 대응)에서 학습 결과를 평가하는 것이 아니라, 여러 동작점 혹은 일정 동작 영역에서 분류기의 성능을 함께

고려하여야 한다.

위와 같이 사기 검출을 위한 신경망 분류기의 학습시의 현실적인 문제들을 함께 해결하는 방법으로서, 본 논문에서는 ROC 곡선을 이용하여 선택적 영역에서 신경망 분류기의 성능을 최대화하는 직접적 방식을 제시한다. 제안하는 방식에서의 분류기 평가 함수는 신경망의 출력에 의한 기울기(gradient)를 유도하기가 용이하지 않으므로, 전역 최적화 기법(global optimization)인 진화 연산 기법을 사용하여 최적의 신경망 분류기를 학습한다. 또한 본 논문에서는, 기존의 대부분 연구에서처럼 부분적으로 수집된(일정범위에서 “왜곡된” 분포의) 거래 데이터에서의 성능 분석이 아니라, 일정 기간동안의 실제 신용카드 거래 데이터(실제의 정상/사기 분포와 거의 동일)에 적용하여, 결과를 분석 고찰하고 제안한 방식의 유용성과 타당성을 보이고자 한다.

본 논문의 구성은 다음과 같다. 2장에서 일반적인 사기 검출 문제에서의 분류기의 평가를 위한 ROC 그래프와, 제안하는 방식에서 사용하는 ROC 기반 분류기의 평가 함수에 대하여 상세히 살펴본다. 3장에서는 실제 신용카드 거래 데이터를 사용하여 사기 검출 분류기를 구성하여 실험한 결과를 보이고, 제시한 방법의 효과 및 장단점에 대하여 논의하며, 4장에서 결론을 맺는다.

2. 신용카드 사기 검출 분류기의 학습

2.1 ROC 그래프 기반 분류기의 성능 평가

사기 검출을 위한 분류기 학습 시에는 일반적으로 사용하는 정확도(accuracy) 기준은 적절하지 않다. 예를 들어, 사기 정상의 클래스 비율이 1:1999인 경우에 모두 정상으로 분류하는 분류기는 실제로 많은 경우에 의미가 없음에도 불구하고, 정확도가 99.95%가 된다. 따라서 정확도를 기준으로 하지 않고, 다음과 같은 검출율(true positive rate, TPR)과 오검출율(false positive rate, FPR)을 분류기의 성능 평가 기준으로 사용하는 것이 적합하다.

$$TPR = \frac{\text{검출된 사기거래 수}}{\text{총 사기거래 수}}$$

$$FPR = \frac{\text{오검출된 정상거래 수}}{\text{총 정상거래 수}}$$

FPR과 TPR를 각각 X축과 Y축으로 하고, 분류기의 출력 문턱값의 변화에 따른 분류기의 성능 즉, (FPR, TPR) 값들을 도시한 ROC 그래프를 이용하면, 보다 명확히 사기 검출 분류기의 성능을 나타낼 수 있다[8]. 그림 1에 3개의 사기 검출 분류기의 성능을 예시하였다. ROC 그래프 공간에서는 상단에 위치할수록 좋은 성능을 나타내는 것이며, (0,1)점은 완벽한 분류 성능을 가진 분류기를 나타내며, 실제로 사기 검출 문제는 겹쳐진 클래스 분포를 가지므로 (0,0)-(0,1) 라인과 (0,1)-(1,1)라인에 근접할수록 좋은 성능을 가진 분류기로 판단할 수 있다. 그림 1에서 보듯이, 분류기 A는 분류기 B보다 상단에 위치하므로, 분류기 A는 분류기 B보다 항상 우수한 성능을 나타낸다. 즉, 같은 오검출율에서 상대적으로 검출율이 항상 높으며, 이에 반해 분류기 C는 분류기 A보다 높은 오검출율 구간에서는 낮은 검출율을 보이나, 낮은 구간에서는 오히려 높은 검출율을 나타낸다. 본 논문에서는 분류기 A와 C의 경우에서처럼, ROC 그래프 공간의 전 영역에서 분류기의 성능을 최대화하지 않고, 관심있는 특정 영역에서의 성능을 최대화하도록 분류기를 학습하고자 한다. 다음절에서

신용카드 사기 검출 문제에 사용되는 평가 함수에 대해서 상세히 설명한다.

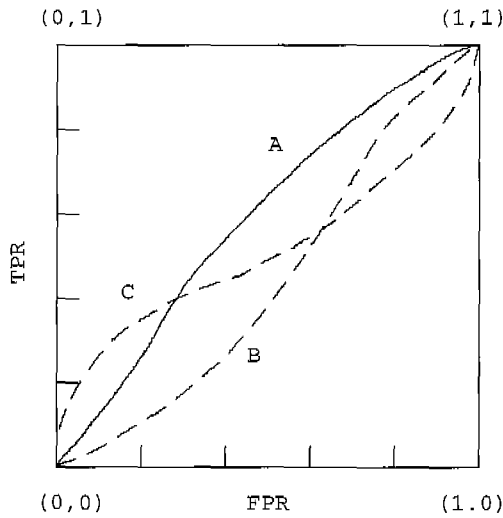


그림 1. ROC 그래프의 예시 (3개의 분류기).

Fig. 1 An example of ROC graphs (three classifiers).

2.2 신용카드 사기 검출 분류기의 진화 학습

실제 필드에서 사용하는 신용카드 사기 검출기의 성능 평가 방법에는 여러 기준이 있다. 거래 단위나 회원단위로 검출/오검출로 평가할 수도 있고, 사기거래를 검출함으로써 발생 가능하였던 손실 금액을 고려하는 등의 금액기준을 사용하기도 한다. 하지만, 본 논문에서는 가장 기본적이고 일반적인 평가 대상인, 2.1절에서 설명한 검출/오검출 거래 건수를 기준으로 한 ROC 그래프 공간을 이용한다.

사기 검출 시스템을 사용하는 사용자측에서는, 여러 현실적인 문제로 인해 거절율(rejection rate, 검출과 오검출 포함)에 제약을 가지고 있다. 많은 사기 거래를 검출하려하여도 사기 검출율의 증가에 따라서 정상 오검출율도 함께 증가하기 때문이다. 따라서, 전체 거래에 대한 거절율을, 실제로 사용자측의 처리 가능한 일정 수준이하로 유지하면서 사기 검출 분류기를 사용하게 된다. 그러므로 신경망 분류기의 학습시의 평가는 실제 사용자가 사용하게 될 거절을 범위 내에서 사기 검출 거래 수를 최대화하도록 설정한다. 또한 그 범위 내에서, 시간에 따른 입력 데이터 분포의 변화에 따른 분류기 성능의 변화를 고려하여서 해당 거절을 구간내의 평균적인 분류기 성능, 즉 국소 구간내의 총 검출 사기 거래 수(local integrated true detection, LITD)를 최대화하도록 분류기를 학습한다. 그림 2는 학습시 사용하는 평가 함수의 개념을 나타내며, 식 (1)은 LITD의 정량적인 표현이다.

$$LITD = \frac{1}{Z} \int_{r_1}^{r_2} N_F(R) dR \quad (1)$$

여기서 R 은 전체 거래에 대한 거절 거래의 퍼센트(%)이며, $N_F(R)$ 은 R 에 따른 신경망 분류기에 의해서 검출되는 사기 거래 수를 나타내고, Z 는 정규화 상수이다. 식 (1)의 LITD가 최대가 되도록 신경망 분류기의 가중치(weight)를 학습함으로써, 분류기의 사용 영역 범위 내에서 성능이 최적인 신경망 분류기를 선택적으로 취할 수 있는 것이다. 이에 반해 사기 검출 신경망 분류기의 MSE 최소화 학습은 전체

학습 데이터에 대해 목표 출력값에 근접하도록 함으로써, 전체적으로 ROC 그래프를 상단으로 움직이는 방법이나, 비슷한 MSE를 가지지만 ROC 그래프 공간상에서의 성능 차이가 나는 분류기를 구별하여 학습하지 못한다. 즉, 지정된 거절을 범위 ($r_1 \sim r_2$)에 대해서 직접적인 성능을 최적화하지 못한다는 것이다. 이에 비해 LITD 최대화 방법은 선택 구간내의 성능을 직접적이면서 보다 용이하게 분류기의 최적화를 수행할 수 있다.

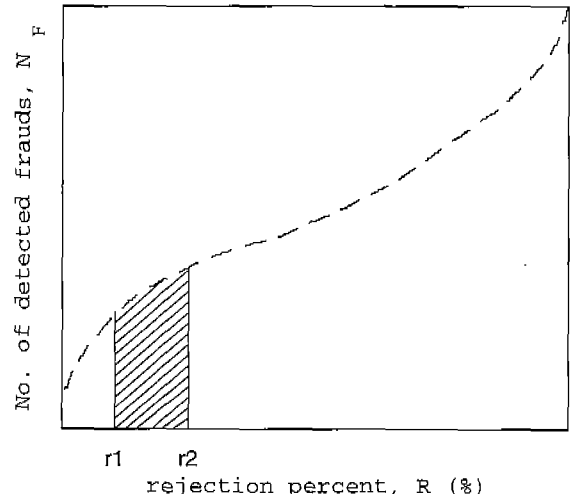


그림 2. LITD 성능 평가 함수의 예시.

Fig. 2 LITD-based evaluation function.

LITD를 최대화하기 위해서는 일반적으로 신경망 학습에 사용하는 오역전파(backpropagation) 방법과 같은 gradient 기반 학습 방법을 이용하는 것은 용이하지 않다. gradient를 구하기 위해서 신경망 출력과 식 (1)과의 관계를 수식적 형태의 비용함수로 표현하는 것이 용이하지 않기 때문이다. 따라서 본 논문에서는 gradient에 기반을 두지 않고 LITD를 최대화하는 방법으로서 진화 연산 기법[10]을 사용한다. 여러 진화 연산 기법 중에서 신경망 학습에 많이 사용되고 있는 진화 프로그래밍(evolutionary programming, EP) 기법을 사용한다. 진화 연산 기법은 신경망의 학습시 국소점 문제 해결뿐만 아니라, 본 논문에서처럼 비용 함수에 대한 gradient를 구하기 어려운 경우에 효과적인 학습 방법이며 학습 성능도 효과적인 것으로 알려져 있다[11].

3. 실험 결과 및 논의

3.1 학습 및 테스트 데이터

실험에 사용한 신용 카드 거래 데이터는 1년 간의 거래 중에서 선별적으로 수집된 약 5만 건의 거래(학습 데이터로 사용)와 이후 3개월 동안의 약 700만 건의 거래(테스트 데이터로 사용)이다. 사기 대 정상 비율은 학습 데이터는 1:4정도이며, 테스트 데이터는 약 1:1700정도로서, 실제 신용카드 거래가 상당히 불균형적인 분류 문제임을 알 수 있다. 높은 성능의 사기 검출 시스템을 위해서는 거래 데이터로부터 효과적인 특징(feature)을 추출하여야 하는데, 거래 횟수와 빈도, 요청 금액과 요청 시간 등의 정보를 분석하여 6개의 특징을 실험에 사용하였다.

3.2 신경망 기반 사기 검출 분류기

사기 검출 시스템에는, 개인의 규칙적인 사용 패턴과 평균적인 사용 행태에 대한 특징을 추출하고, 거래 발생시 이와의 차이를 기준으로 사기를 검출하는 방법(user profiling)과 사기 거래들의 공통적인 사용 패턴을 추출하고, 이와 유사한 거래 발생 시에 사기로 판단하는 두 가지 접근 방식이 가능하다. 본 논문에서는 거래 데이터에 일정 기간이상의 연속적인 사용 거래 데이터를 포함하고 있지 않으므로 후자의 방식의 사기 검출 분류기를 사용한다.

사용하는 신경망은 1개의 은닉층을 가지고 시그모이드(sigmoid) 함수를 사용하는 다층 퍼셉트론(MLP) 모델을 사용하였다. Radial basis function처럼 국소적인 성질을 가지거나 복잡도(complexity)가 높은 신경망 모델을 사용함으로써 약간의 검출 성능을 높일 수 있을 것으로 기대되나 본 논문에서는 이에 대한 실험은 생략하였다. 또한, 분류 판단단계에서 정상 대 사기 비율을 다소 낮추기 위한 조건 필터[2,3]는 사용하지 않았다.

3.3 실험 및 결과 논의

우선 MSE 최소화화 및 LITD 최대화의 학습 평가 지수의 차이를 비교 분석한다. 실험 조건은 다음과 같다. 신경망의 은닉층은 10, 20으로 바꾸어 가면서 두 학습 평가 지수를 최적화하도록 각각 10개의 신경망 분류기를 생성하였다. 비교시의 공정성을 위하여 MSE 기준으로 신경망을 학습 할 때에도 EP를 사용하였다. EP의 탐색 연산자는 가우시안 변이(Gaussian mutation)를 사용하였으며, 파라미터들은 수행 시간을 감안하여 두 경우 모두 동일하게 적용하였다(표 1). 식(1)에서의 LITD를 계산하는 구간은 (2% ~ 6%)으로 하였다.

그림 3은 학습 데이터에 대한, 각각 MSE와 LITD 기준으로 학습한 10개씩의 신경망 분류기(은닉 노드의 개수가 20)의 성능을 두 평가 지수 공간(X축: MSE, Y축: LITD)에 도시한 그림이다. MSE 기준으로 학습한 10개의 분류기를 (+로 표시)을 살펴보면, 상대적으로 높은 MSE 값을 가지는 분류기들이 낮은 LITD 값을 가지는 것을 알 수 있다. 즉, MSE 기준으로 학습한 분류기들은 전반적으로는 MSE 감소와 LITD 증가간에 약간의 상관관계를 가진다. 하지만, MSE 값이 일정 수준이하로 학습된 분류기들에서는 반드시 MSE 감소와 LITD 증가가 일치하지 않음을 알 수 있다. 즉, 비슷한 MSE 값을 가지는 신경망 분류기가 LITD 측면에서는 상당히 다른 성능을 보인다. LITD를 기준으로 학습한 분류기(x로 표시)는 MSE를 기준으로 학습한 경우에 비해 상당히 큰 MSE 값을 가지나, 대부분 MSE를 기준으로 학습한 경우보다 높은 LITD를 나타냄을 알 수 있다. 이는 예상하였던 것처럼 지정된 거절을 구간내의 성능을 최대화하기 위해서 구간 밖의 데이터에 대한 MSE가 높아지는 방향으로 분류 경계선이 형성되기 때문이다. 이러한 현상은 은닉 노드 개수가 10인 경우에서도 관찰할 수 있었다.

표 1. EP의 파라미터들.

Table 1. Parameter values of EP run.

수행(generation) 횟수	500
집단(population) 크기	10
토너먼트 경쟁 크기	8
가우시안 변이의 σ	1.0

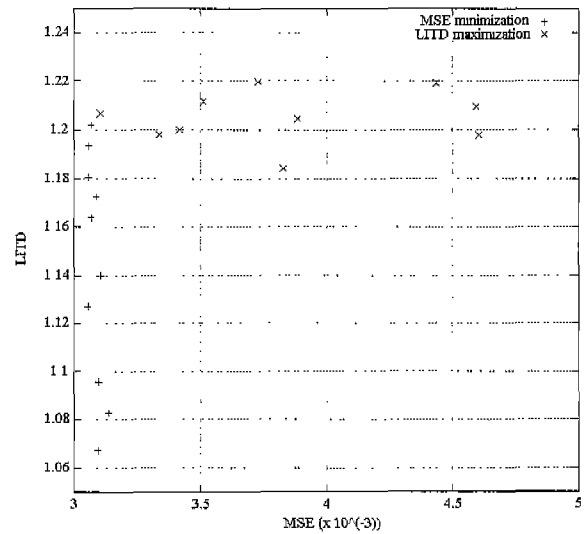
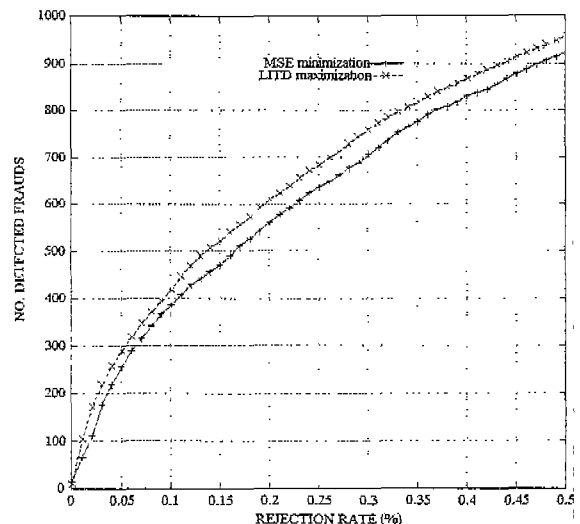


그림 3. MSE와 LITD의 상관 관계.

Fig. 3. A relationship between MSE and LITD performance indices.

그림 4는 MSE와 LITD를 기준으로 학습한 신경망 분류기의 일반화 성능을 나타낸다. 각 평가 지수를 기준으로 상위 50% (5개)의 신경망 분류기의 테스트 데이터에서의 평균적인 성능을 도시하였다. 실제 사기 검출을 위해서 사용할 경우에 허용 가능한 거절율의 범위는, 높은 정상/사기 비율(약 1700/1 정도)을 고려하면 건수 기준으로 약 0.1% ~ 0.4% 정도로 낮아지고, 이 거절을 구간에서 신경망 분류기의 일반화 성능을 평가한다. 일반화 성능 평가시의 LITD 계산 구간이 학습시의 LITD 계산 구간 (2% ~ 6%)보다 낮은 이유는 테스트 데이터의 정상 거래수가 학습 데이터의 정상 거래수보다 훨씬 많으면서도 하루당 일정 수준 이하의 거절 건수를 유지하기 위해서는 분류기가 학습시 보다 낮은 거절을 구간에서 운용되어야 하기 때문이다. 그림에서 보듯이, LITD를 기준으로 학습한 경우의 신경망 분류기가, 비록 MSE 기준으로는 성능이 낮았지만, 은닉층의 노드 개수에 상관없이 0.1 ~ 0.3% 사이의 거절을 구간에서 더 많은 사기 거래를 검출함을 알 수 있다. 이는 학습한 분류기 10개 모두의 평균 성능에서도 동일한 양상을 보였다.



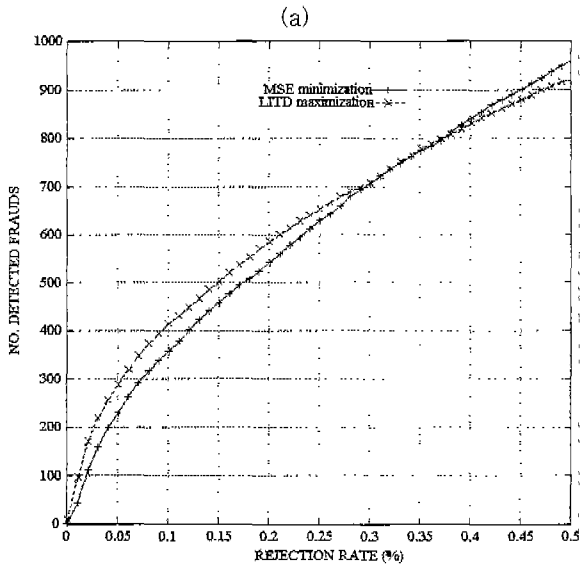


그림 4. MSE와 LITD를 기준으로 학습한 상위 5개의 신경망 분류기의 일반화 성능. (a) 은닉 노드 개수 10, (b) 은닉 노드 개수 = 20.

Fig. 4. Performance of the top five neural networks classifiers in terms of MSE and LITD, respectively. (a) No. of hidden nodes = 10, (b) No. of hidden nodes = 20.

또한, 실제 사기 검출 분류기의 성능 평가는 검출율과 오검출율도 중요한 평가중의 하나이지만, 비용 기준 평가도 중요한 평가 요소 중의 하나이다. 사기거래의 검출로 인한 경제적 이익(profit)은 사기 검출 시스템을 채택하는 입장에서 가장 중요한 관심사이기 때문이다. 하지만 정확하게서 절대적인 이익의 계산은 불가능하다. 거래 데이터 자체가 이미 기존의 검출기에 의해서 승인/거절되었던 거래이기 때문이다. 다시 말해서, 각 거래가 정상/사기 거래이건 간에 기존의 검출기에 의해서 이미 승인/거절되었다는 것은 기존의 검출기에 의해서 거래 패턴이 이미 영향을 받았다는 것을 의미하기 때문이다. 따라서, 정확히 기존의 사기 검출 분류기와 이익 비교는 불가능하지만, 정확하지는 않지만 대략의 상대적인 이익(relative profit)의 경향 비교는 가능하다. 표 2는 상대적인 이익을 계산하기 위한 비용 모델이다. [6]에서처럼 사기로 분류하는 경우에 발생하는 제비용을 고려한 overhead를 비용 모델에 채택하였다. Rate는 거래를 승인함으로써 발생하는 수수료를 나타낸다.

그림 5는 기존 검출 분류기에 대한 MSE와 LITD 기준으로 학습한 상위 50%의 신경망 분류기의 평균적인 상대적 이익 곡선이다(overhead와 rate는 1000원과 2.5%로 가정). 그림에서 보듯이 문턱값에 따라서 상대적 이익이 달라지는데, 거절율이 약 0.1% - 0.2% 사이에서 상대적 이익이 최대가 될 수 있다(거절율이 0%인 경우, 즉 전혀 검출을 하지 않는 경우의 이익은 실험에서 사용한 거래 데이터에 대해서 계산된 값일 뿐 실제 온라인 상으로 운영할 때의 이익과 전혀 다른 것이다). 따라서, 현재 사용하고 있는 분류기 대비 이익을 고려하여 사용할 경우에는 0.1% - 0.2%사이의 거절율 구간에서 사용하는 것이 바람직하다. LITD 최대화를 기준으로 학습하였지만, MSE를 최소화하도록 학습한 경우보다 실제 이익/비용측면에서도 유리함을 알 수 있다.

표 2. 기존의 분류기에 대한 상대적 이익을 계산하기 위한 이익 모델.

Table 2. A profit model for calculating relative profit with respect to the pre-installed classifier.

승인/거절 여부	해당거래의 정상/사기 여부	상대적 이익
기 "거절"된 거래를 "승인"	정상	overhead + 금액 * rate
	사기	overhead - 금액
기 "승인"된 거래를 "거절"	정상	-overhead - 금액 * rate
	사기	-overhead + 금액

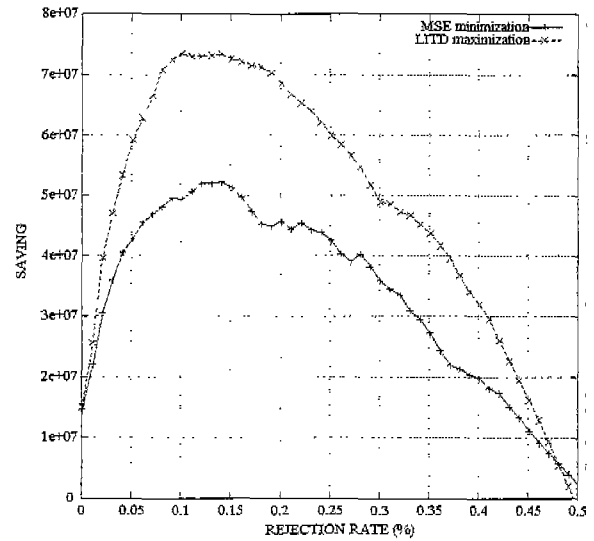


그림 5. HSE와 LITD를 기준으로 학습한 상위 5개의 신경망 분류기의 기존 검출기 대비 상대적 이익 곡선.

Fig. 5. Relative profit curves of the top five classifiers that were trained in terms of MSE and LITD performance indices, respectively with respect to the pre-installed classifier.

4. 결론

본 논문에서는 사기 검출을 위한 신경망 분류기의 학습에 적합한 방식을 제안하였다. 본 학습 방식은, 출력단의 오차를 최소화하는 방식이 아닌, 진화 연산 기법을 사용하여 사기 검출 분류기의 특정 동작 영역에서의 성능을 선택적으로 그리고 직접적으로 최적화하는 체계적인 학습 틀/framework)을 제공하며, 또한 실제 신용카드 사기 검출 시스템에서 발생하는 입력 거래의 분포 변화에도 덜 민감한 신경망 분류기를 생성할 수 있다. 실제 신용카드 거래 데이터에 대한 실험 결과를 통해서 제안한 방식이 사기 검출에서의 분류기 학습 시에 검출율/오검출을 뿐만 아니라 상대적인 이익측면에서도 좋은 성능을 나타냄을 보였다.

이외에도, 제안한 진화 연산 기반 학습 방식은 하나의 분류기의 최적화뿐만 아니라 상호 보완적인 복수개의 분류기 군을 학습할 수 있는 장점이 있다. 이와 함께 이익 모델에 기반하여 사기 검출 분류기를 학습하는 방법에 대하여 연구중

참 고 문 헌

[1] T. Fawcett and F. Provost, "Adaptive fraud detection," *Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 291-316, 1997.

[2] J. R. Dorronsoro, F. Ginel, C. Sanchez, and C. S. Cruz, "Neural fraud detection in credit card operations," *IEEE Trans. on Neural Networks*, vol. 8, no. 4, pp. 827-834, 1997.

[3] L.-J. Park, S.-A. Kim, H.-J. Cho, T.-S. Kim, and B.-H. Wang, "A credit card fraud detection system based on hybrid of neural networks and decision tree," *Proc. of the Korea-U.S. Science and Technology Symposium*, 1998.

[4] P. K. Chan and S. J. Stolfo, "Toward scalable learning with non-uniform class and cost distributions : A case study in credit card fraud detection," *Proc. of Int'l Conf. Knowledge Discovery and Data Mining*, pp. 164-168, 1998.

[5] P. K. Chan and S. J. Stolfo, "Learning with non-uniform class and cost distributions : Effects and a distributed multi-classifier approach," *KDD-98 Workshop on Distributed Data Mining*, pp. 1-9, 1998.

[6] P. K. Chan, W. Fan, A. L. Prodromidis, and S. J. Stolfo, "Distributed data mining in credit card fraud detection," *IEEE Intelligent Systems*, vol. 14, no. 6, pp. 67-74, 1999.

[7] F. Provost, "Learning with imbalanced data sets 101," *AAAI'2000 Workshop on Imbalanced Data Sets*, 2000.

[8] F. Provost and T. Fawcett, "Analysis and visualization of classifier performance : Comparison under imprecise class and cost distributions," *Proc. of Int'l Conf. Knowledge Discovery and Data Mining*, pp. 43-48, 1997.

[9] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, New York: John Wiley, 1973.

[10] T. Bäck, D. B. Fogel, and Z. Michalewicz, *Handbook of Evolutionary Computation*, IOP Publishing and Oxford University Press, 1997.

[11] X. Yao, "Evolving artificial neural networks," *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1423-1447, 1999.

저 자 소개



박래정(Lae-Jeong Park)

1991년 : 서울대학교 전기공학과 공학사
 1993년 : 한국과학기술원 전기 및 전자공학과 공학석사
 1997년 : 한국과학기술원 전기 및 전자공학과 공학박사
 1997년~1999년 : LG종합기술원 정보기술연구소 선임연구원
 2000년~현재 : 강릉대학교 정보전자공학부 전임강사

관심분야: 기계 학습, 데이터 마이닝, 진화 연산
 Phone : +82-33-640-2389
 Fax : +82-33-640-2244
 E-mail : ljpark@kangnung.ac.kr