

기계학습 기반의 이상거래 탐지 시스템 (사례연구)



2018년 8월



제 출 문

본 보고서를 “기계학습 기반의 이상거래 탐지 시스템”의
최종 결과보고서로 제출합니다.

2018년 8월

연 구 주 관 : 한양대학교 지식서비스연구소

연 구 기 간 : 2018. 04. ~ 2018. 08. (5개월)

연 구 자 : 정 미 연

지식서비스연구소장 귀하

* 본 보고서는 중소벤처기업부에서 시행한 컨설팅대학원 및 컨설팅 R&D 센터 사업지원으로 연구되었습니다.

목 차

I. 서론	4
II. 이론적 배경	5
2.1 CRM과 금융거래 이상탐지	5
2.2 이상탐지와 머신러닝	8
2.2.1 이상탐지와 머신러닝 사용현황	8
2.2.2 분류 알고리즘(로지스틱 회귀분석, 랜덤 포레스트, XGBoost, MLP)	9
III. 연구 설계 및 분석	13
3.1 연구 설계	13
3.2 연구 모형	13
3.3 사례 분석	14
3.3.1 탐색적 자료 분석(EDA)	14
3.3.2 변수 선택	16
3.3.3 로지스틱 회귀분석	18
3.3.4 랜덤 포레스트	19
3.3.5 XGBoost	20
3.3.6 Multi Layer Perceptron	22
3.4 분석 결과	22
IV. 결론 및 한계점	24
별첨. 참고문헌	

1. 서론

근래에 IT 기술의 발전에 따라 국내외로 금융 산업과 기술을 융합하여 전자금융거래 시장이 급속도로 성장하고 있다(정성훈 외 4명, 2015). 금융거래량이 늘면서 결제 수단이 개발되고 있지만 그만큼 사기탐지가 어렵도록 금융 사기는 치밀해져 가고 있다, 2017년 한국은행 지급 수단 이용행태 조사결과에 따르면 지급수단별 이용 경험은 현금 다음으로 신용카드가 79.1% 이고 모바일 기반 인터넷 쇼핑에서도 모바일 카드나 휴대폰 소액결제 등의 다양한 모바일 지급수단이 활용되고 있다. 그러나 조사 대상자의 11.0%는 1년 내로 지급 수단을 분실하거나 보이스피싱, 위/변조, 도난이나 유출된 개인정보를 악용하려는 경험을 했다고 응답하여 이를 보완해야 하는 노력이 필요할 것이다. (2017 지급수단 이용행태, 한국전화결제산업협회) 금융 당국은 이런 피해를 막고자 이상금융거래감지시스템(FDS, Fraud Detection System)을 필수로 구축할 것을 법으로 제정하였다.

FDS는 전자상 금융거래 시에 발생하는 정보나 내역들을 분석하여 금융산업 이상거래를 탐지하고 이를 차단하여 금융거래 보안성을 높인다. 현재 우리나라에서는 은행 20개사와 증권회사 26개사가 2014년 이후로 시스템을 개발하고 운영하고 있다. 이는 소비자의 피해를 막고 안전한 거래를 제공할 수 있는 서비스로, 일종의 CRM(고객관계관리)의 일종으로 볼 수 있다. 사기 탐지 시스템 혹은 이상거래 탐지 시스템이라고도 불리는 이 시스템은 실시간으로 사용자 혹은 결제 데이터를 받아 분석하고, 일반적인 패턴으로 정해 둔 이외의 다른 패턴의 이상 거래가 탐지될 경우에 즉시 사용자나 금융기관 탐지 사실을 알려 피해를 막는다. 피해를 막는다는 입장에서는 기업의 위기관리라고 볼 수 있을 것이다. 이를 위해서는 정확하면서도 빠른 탐지가 중요하고 알고리즘을 잘 사용하여 성능을 향상시키는 것이 중요하다.

실제로 국내 금융기업들이 FDS를 구축하여 연간 446억원 규모의 사고를 무사히 막을 수 있었고 예방률이 95.4%에 달한다(한국금융연구원, 2018). 미탐율과 탐지 후 사고율은 2.3% 수준으로 효과적으로 차단하고 있다고 하는데 이는 정확도가 높은 모델을 구성하게 된다고 해도 잘못 탐지하는 데에 있어서 손실을 고려해야 함을 알 수 있다. 오탐의 경우에는 사용자 신뢰성을 잃고, 미탐의 경우에는 회사 입장에서의 손실을 발생시킨다. 그러나 여기서 발생하는 재무적인 가치 이상으로 소비자 신뢰성의 비재무적 가치가 존재할 것이므로 꾸준한 오탐 방지가 중요할 것이다. 본 연구에서는 국내 금융감독원이 제시한 프레임워크 일부인 정보 수집 및 가공 - 분석/탐지 순으로 진행한다. 또한 로지스틱 회귀분석, 랜덤 포레스트, XGBoost 알고리즘을 사용하여 각 알고리즘에 따른 비용을 계산하여 줄어드는 것을 확인하고 오탐율과 미탐율을 함께 줄이도록 성능이 높은 알고리즘을 적용하여 이상 탐지 시스템의 중요성을 시사한다.

II. 이론적 배경

2.1 CRM과 금융거래 이상탐지

고객관계관리(CRM, Customer relationship management)란 기업 입장에서 소비자들을 고객으로 유치 및 유지하고자 하는 경영의 일종이다. 기업들이 고객과 관련된 정보를 기반으로 숨겨진 고객 패턴을 분석하고 관리하는데 있어서 모든 분야를 의미한다(김승현, 진승헌).

가치전달의 의미로는 기업에 대한 고객충성도 향상을 위해 우량고객 중심의 차별 서비스 제공을 하며 고객 평생가치 극대화 혹은 이탈을 막는 전략이라고 볼 수도 있다(한상린, 박경숙, 2001). CRM을 일반적으로 크게 나누면 지속적인 관계를 유지하고자 하는 마케팅 행위, 이용자의 특성에 따른 관리 및 서비스, 고객의 데이터를 기반으로 한 정보기술관리, 전사적 차원의 관리로 구분해볼 수 있다(김형수, 김영걸, 박찬욱. CRM 고객관계관리 전략 원리와 응용. 2009). CRM을 통해서 수익성 증대, 고객의 정보에 따른 전략수립, 이탈 탐지 및 방지, 고객 활성화, 고객가치 증대 등의 효과를 기대할 수 있다. 본 논문에서는 CRM 관점에서 금융거래 이상탐지 연구에 접근해 보려고 한다.

금융거래 이상탐지란, 실시간으로 행동을 모니터링하여 이상 금융거래를 찾는 것이다. 특히 최근에 금융 거래량 증가와 수단의 다양화로 인하여 금융 사기가 다양해지고 있다. 최근에 두각되고 있는 핀테크의 경우에는 금융, 기술을 합쳐 말하는 것으로 모바일로 은행업무를 처리하거나 클라우드 펀딩 등의 정보기술 기반의 최신 금융기술이다. 금융관리, 금융데이터 분석, 금융 소프트웨어, 플랫폼으로 나뉘어 있고 핀테크의 목표는 사용자들이 편리하게 금융서비스에 접근하여 안전하게 지불할 수 있도록 하는 것이다. 핀테크와 간편결제의 활성화로 고수준의 보안시스템이 요구되고 있으며 금융당국도 2014년 이후 제 1금융권과 제 2금융권을 중심으로 FDS 구축을 법으로 의무화 하였다. 이런 머신러닝 기반의 탐지 시스템을 통해 시스템의 성능 향상, 신속한 대응, 자동화된 결과를 잘 반영하는 것을 기대하고 있다(CyberSource. 2016)

이상탐지를 하는 경우 발생하는 문제점은 미탐과 오탐이다. 이상 탐지를 못 했을 경우, 즉 미탐 시에는 신고 전에는 금융회사가 알 수 있는 방법이 없기 때문에 후에 회사 손실비용 발생 및 경제적 신뢰를 떨어뜨릴 수 있다. 오탐 시 소비자의 입장에서는 정상 행위를 했는데도 불구하고 오인받은 것이기 때문에 보상비용 발생과 회사 시스템에 대한 신뢰성 저하로 이어질 수 있다. 아래 [그림 1]의 도표를 살펴보면 사기 탐지(이상 탐지)는 시각화, 분류, 군집화로 이루어져 있으며 은행, 보험, 기타금융관련에서는 경제적 신뢰성을, 거래와 통신에서는 사회적 신뢰성을 이룬다고 정의하였다(전병국 외 3인.2014).