
저자 (Authors)	전한나, 김종우
출처 (Source)	한국지능정보시스템학회 학술대회논문집 , 2021.6, 17-18 (2 pages)
발행처 (Publisher)	한국지능정보시스템학회 Korea Intelligent Information Systems Society
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE10588070
APA Style	전한나, 김종우 (2021). 거래 정보와 순차 정보를 활용한 딥러닝 기반의 신용카드 거래 이상탐지 기법. 한국지능정보시스템학회 학술대회논문집, 17-18.
이용정보 (Accessed)	한양대학교 166.***.182.218 2022/02/28 22:24 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관 소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

C1.3 거래 정보와 순차 정보를 활용한 딥러닝 기반의 신용카드 거래 이상탐지 기법

전한나
한양대학교 일반대학원
경영학과
junhannah@hanyang.ac.kr

김종우
한양대학교 경영대학
경영학부
kjuw@hanyang.ac.kr

Abstract - 본 연구에서는 신용카드 이상탐지를 위한 기법으로 거래 정보와 순차 정보를 모두 활용한 딥러닝 기반의 앙상블 모델을 제시하였다. 이를 위해, 신용 카드 거래 데이터에서 각 거래 정보와 순차 정보 특징을 추출하고, 앙상블 모델에 학습하였다. 이를 통해 기존 연구에서 단편적인 정보만을 사용한 모델의 성능과 비교하여 성능 향상이 이뤄졌음을 보였다.

Key Terms - 이상탐지, 앙상블 기법, Credit card fraud detection, Deep learning

사사표기

이 논문 또는 저서는 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2020S1A3A2A02093277)

I. 서론

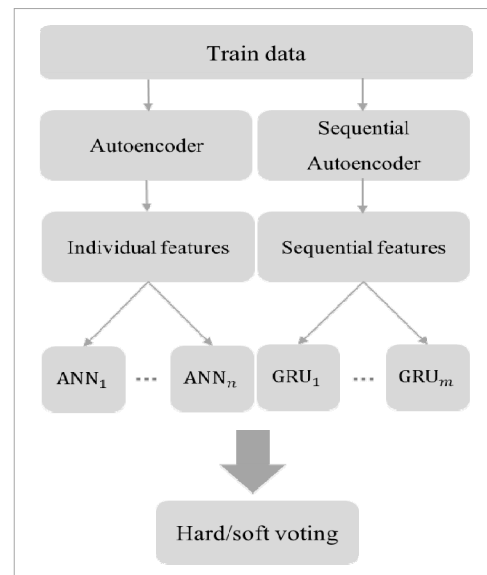
오늘날 금융 거래에 있어 금융 사기는 막대한 비용 손실을 가져올 수 있어 정확한 이상 거래의 탐지가 필요하다. 이를 위해 기존 많은 연구에서는 머신 러닝 기법(Bahnsen et al, 2016), 딥러닝 기법 (Jurgovsky et al, 2018)을 활용하여 신용카드 거래 이상탐지를 수행하였다.

그러나, 기존 연구에서는 각 거래의 정보만을 활용하여 시간적 정보를 고려하지 않거나, 시간적 정보만을 활용하여 단편적인 정보만 사용한다는 한계점이 존재한다. 이에 본 연구에서는 신용 카드 거래에서 거래 정보와 시간 정보를 모두 활용한 딥러닝 기반의 앙상블 기법을 제시하여 앞서 언급한 한계점을 개선하고자 하였다.

II. 데이터 및 제안 모델

본 연구에서 사용한 데이터는 캐글 웹사이트 의 신용카드 이상 거래 데이터를 사용하였다. 이 데이터는 2013년 유럽 카드 소유자들의 이틀 간의 거래 내역으로 총 284,807개의 거래로 구성된다. 거래는 정상 거래와 이상거래로 구분되며, 이 중 492개의 거래만이 이상거래로, 불균형한 데이터 분포를 이룬다. 모델 학습을 위한 학습 데이터는 전체 데이터의 70%, 테스트 데이터는 20%, 검증용 데이터는 10%를 사용하였다.

다음으로, 본 연구에서 제안한 모델의 알고리즘 구조는 <그림 1>과 같다. 순서는 다음과 같이 수행되었다. 학습 데이터를 각각 ANN(Artificial Neural Network)로 구성된 Autoencoder, GRU(Gated Recurrent Units)로 구성된 Sequential Autoencoder를 통해, 각각의 거래 정보의 특징(individual features)과 순차적 정보의 특징(Sequential features)에 대한 특징벡터를 추출한다. 이후 각 거래 정보는 n개의 ANN 앙상블 모델에 학습하고, 순차적 정보는 m개의 GRU 앙상블 모델에 학습한다. 마지막으로, 앙상블 모델이 예측한 벡터 값들을 majority voting인 hard voting과 soft voting을 통해 최종 예측 값을 구하였다.



<그림 1> 알고리즘 구조

III. 모델 성능 평가

본 연구의 제안 모델은 거래 정보만을 사용한 모델, 시간 정보만을 사용한 모델과 성능을 비교하여 두가지 정보를 활용한 효과 및 성능 향상을 보고자 하였다. 따라서, 비교군으로는 거래 정보를 사용한 ANN앙상블 모델, 시간 정보를 사용한

GRU 앙상블 모델, 선행 연구에서 제안한 거래 정보를 사용한 딥러닝 모델(Misra et al, 2020), 시간 정보를 활용한 딥러닝 모델 (Forough and Momtazi, 2020)을 택하였다.

모델의 성능 평가를 위해 사용한 성능지표는 Precision(정밀도), Recall(재현율), F1-score를 사용하여 평가하였다. 성능 비교 결과는 다음 <표 1>과 같다.

<표 1> 모델의 성능 비교 결과

Model	Precision	Recall	F1-score
Forough and Momtazi, 2020	93.51%	73.51%	82.29%
Misra et al, 2020	85.34%	80.15%	82.65%
ANN 앙상블 모델	100%	76.7%	86.8%
GRU 앙상블 모델	100%	76.7%	86.8%
제안 모델	100%	77.9%	87.89%

성능 비교 결과를 볼 때, 제안 모델은 Precision 100%, Recall 77.9%, F1-score 87.89%를 기록하였다. 이 수치는 개별, 시간 정보만을 활용한 모델보다 성능 향상이 이뤄졌음을 볼 수 있다.

IV. 결론

본 연구에서는 거래 정보와 순차 정보를 활용한 딥러닝 기반의 신용카드 거래 이상 탐지 모델을 제시하였다. 추후 연구 방향은 제안된 모델에서 Recall성능을 향상시키기 위한 방안을 모색해 나갈 것이다.

V. 참고문헌

Bahnsen, A. C., D. Aouada, A. Stojanovic, and B. Ottersten, "Feature engineering strategies for credit card fraud detection," Expert Systems with Applications, Vol.51, (2016), 134~142.

Forough, J., and S. Momtazi, "Ensemble of deep sequential models for credit card fraud detection," Applied Soft Computing, Vol.99, 2021.

Jurgovsky, J., M. Granitzer, K. Ziegler, S. Calabretto, P. E. Portier., L. He-Guelton, and O. Caelen, "Sequence classification for credit-card fraud detection," Expert Systems with Applications, Vol.100, (2018), 234~245.

Zamini, M., and G. Montazer, "Credit Card Fraud Detection using autoencoder based clustering," 2018 9th International Symposium on Telecommunications (IST), (2018), 486~491.

Misra, S., S. Thakur, M. Ghosh, and S. K. Saha, "An Autoencoder Based Model for Detecting Fraudulent Credit Card Transaction," Procedia Computer Science, Vol.167, (2020), 254~262.