Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, December : 2023

A DEEP LEARNING MODEL FOR INTERNET LOAN USING ANTI-FRAUD MODEL

[1]ALAM VIDYA SANTHI, M.Tech , [2]Mr. U.DHANUNJAYA,M.Tech(PhD)

[1]alamvidya282@gmail.com, [2]dhananjaycse@skucet.org

[1,2]Sri Krishnadevaraya University College Of Engineering And Technology, Sri Krishnadevaraya University, Anantapur,515003

Abstract

Online banking has been growing in popularity as of late. Internet banking institutions, however, now face a grave danger from bad debt. Traditional financial institutions often use logistic regression as a fraud detection methodology. Logistic regression still has room for improvement in terms of accuracy, even if it is interpretable. This exploration examines the practicality of involving a profound brain network to recognize extortion utilizing a major public credit dataset, like that of Loaning Club. An irregular woodland is utilized to at first fill in the missing factors. The best qualities for separation are then picked utilizing the XGBoost calculation. We then, at that point, recommend tending to the example irregularity utilizing a manufactured minority oversampling procedure.We build a deep neural network to identify online loan fraud using the preprocessed data. A large body of experimental evidence shows that deep neural networks outperform the most popular models. Financial engineers at small and medium Internet financial organisations might profit from such a straightforward approach, which may improve the use of deep learning for anti-fraud purposes related to online loans.

KEYWORDS : xbboost, antifraud, logistic function, internet banking

1. INTRODUCTION

1.1 About Deep Learning

A subfield of machine learning, deep learning (also referred to as deep structured learning) makes use of ANNs and RL to learn representations. Supervised, semi-supervised, and unsupervised learning are the three main categories.Neural networks that learn from data, such as recurrent neural networks, deep belief networks, and convolutional neural networks, have been used in many different domains to achieve results that are on par with or even better than those of human experts. These domains include computer vision, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection, and speech recognition, among others.The dispersed communication nodes and information processing in biological systems provided the inspiration for artificial neural networks (ANNs). ANNs vary from biological brains in many ways. To be more precise, most biological brains are dynamic (plastic) and analogue, while artificial neural networks are often static and symbolic.

One meaning of the term "deep" in the context of deep learning is the use of numerous network layers. Although linear perceptrons are not capable of universal classifiers, earlier research shown that networks with non-polynomial activation functions also, one unbounded-width stowed away layer may. Profound learning is a later variation that core interestson bounded-size layers with an unlimited number of them. This allows for optimised implementation and practical application while as yet keeping up with hypothetical all inclusiveness under moderate conditions. The "organized" angle comes from the way that profound learning permits layers to be different and withdraw altogether from physiologically roused connectionist models toachieve efficiency, trainability, and understandability.

Artificial neural networks

Computer systems that mimic the structure and function of the neural networks found in the brains of living things are known as artificial neural networks (ANNs) or connectionist systems. Without task-specific programming, these systems learn (become better at doing tasks) by looking at examples. As an illustration, in the realm of image recognition, they may train to detect cat photographs by studying previously annotated examples with the labels "cat" or "no cat" applied by hand, and then applying the same analysis to new images. Applications that defy expression by conventional rule-based programming algorithms have seen the most success with these.

A network of interconnected units known as artificial neurons (like the neurons in a real brain) is the foundation of an ANN. A neuron may communicate with another neuron via any of its synapses. Postsynaptic neurons may receive signals, analyse them, and then send signals to neurons downstream of them. The states of neurons may be represented by real numbers, usually ranging from 0 to 1. The intensity of the signal sent downstream by neurons and synapses may be affected by their weight, which changes as learning progresses.

Neurons are usually arranged in layers. The inputs to the various levels could undergo a variety of modifications. Possibly after making many trips through the levels, signals make their way from the first (input) to the final (output) layer.

The initial motivation behind the neural network method was to model problem-solving after the human brain. Backpropagation, in which data is sent in the other way and the network is adjusted to reflect that, is one example of how the emphasis shifted from biology to matching certain mental talents.

Numerous applications have found usage for neural networks, such as medical diagnosis, social network filtering, machine translation, medical imaging, and computer vision.

The usual number of units and connections in a neural network ranges from a few thousand to several million as of 2017. These networks are capable of performing a wide range of activities at a level exceeding human capability, even though the number of neurons on a human brain is many orders of magnitude more.

Deep neural networks

An ANN with a few layers between its feedback and result layers is known as a profound brain organization (DNN). The structure blocks of each and every brain organization — neurons, neurotransmitters, loads, inclinations, and capabilities — stay steady no matter what the sort of organization. These parts can be shown very much like some other ML calculation and work like human minds. As an example, a DNN with breed recognition training may examine a provided picture furthermore, decide the probability that the canine being referred to has a place with a specific variety. Before returning the suggested label, the user may examine the findings and tell the network which probabilities to show (above a specific threshold, for example). There is a layer for every mathematical operation in and of itself, [citation required]. "Deep" networks come from the fact that sophisticated DNNs contain several layers.

DNNs are able to represent intricate non-linear connections. In DNN designs, the item is addressed as a layered structure of natives, and the models that outcome are compositional. The extra layers consider the gathering of qualities from lower levels, which could bring about the displaying of muddled information utilizing less units contrasted with a shallow organization that performs in much the same way.One example is the demonstration that DNNs outperform shallow networks when it comes to approximating sparse multivariate polynomials.

Deep architectures include several iterations of a handful of fundamental methods. There are several areas where each architecture has been successful. Without evaluating them on the same data sets, it might be difficult to compare the performance of different designs.

Most DNNs are feedforward networks, meaning that information goes directly from the contribution to the result withoutany kind of feedback loop. As a first step, the DNN builds a network of virtual neurons and gives every association between them an irregular mathematical worth, or "loads".A value between zero and one is returned as an output by multiplying the weights and inputs. In the event that the network failed to correctly identify a certain pattern, a weight adjustment technique would be used. Until it finds the right mathematical manipulation to analyse the input completely, the algorithm might make certain parameters more important.

Applications like language modelling make use of repetitive brain organizations (RNNs), which permit contribution to stream in one way or the other. This is an applicationwhere long short-term memory really shines.

The field of computer vision makes use of convolutional deep neural networks (CNNs). Automatic speech recognition (ASR) acoustic modelling is another area that has seen CNN use.

1.2 ABOUT THE PROJECT

Internet financial models have grown rapidly in recent years, and conventional financial institutions used to handle most of the Internet business. This has led to a huge increase in the employment of fraudulent tactics on the Internet. This is why the danger of online fraud is so high for loan firms operating on the Internet. Financial risk management and big data analysis have been blessed with new prospects brought about by the convergence of rapidly improving computer technology, ever-increasing data sets, and novel methods for analysing this data. Several fraud prevention methods and anti-fraud procedures have been created by researchers over the years. For the purpose of detecting fraud, Leonard [1] suggested an expert system that is rule-based. Bank fraud specialists built the rules of this model by hand. To aid risk analysts in extracting additional fraud rules, Sanchez et al. [2] proposed utilizing affiliation rules to recognize misrepresentation. Monetary misrepresentation displaying language (FFML) was proposed by Edge and Sampaio [3]to aid in fraud analysis by better specifying and integrating fraud rule sets. On the other hand, rule-based models can't be updated quickly enough to account for new frauds since they rely on accurate and extensive expert knowledge.

Consequently, fraud detection using machine learning models has been implemented. Credit card fraud may be detected using neural networks, according to Ghosh and Reilly [4].

In order to identify fraudulent transactions, Kokkinaki [5] suggested using choice trees and Boolean rationale capabilities to depict normal exchange designs. To recognize extortion, Peng et al. [6] thought about nine ML models. These findings show that linear logistic and Edition of 2021, #9 An Attribution 4.0 Licence from Creative Commons is in effect for this work. The full text may be found at this URL: https://creativecommons.org/licenses/by/4.0/. The Future of Online Loan Anti-Fraud Models Using Deep Learning, by W. Fang et al. The effectiveness of Bayesian networks is higher. A novel clustering approach, the better serious learning organization (ICLN) and the directed a superior cutthroat learning organization (SICLN),were suggested by Lei and Ghorbani [7]. A cost-sensitive decision tree was developed by Sahin et al. [8]. To combat fraud, Halvaiee and Akbari [9] suggested using an enhanced AIRS algorithm. Unfortunately, model risk is readily introduced by these conventional machine learning approaches, which depend substantially on subjective rules that are manually applied. Because the training dataset is imbalanced and heavily contaminated with noise, these approaches also have a tendency to overfit. In order to incorporate many models for complex fraud detection, ensemble learning approaches have also been created. A sacking outfit model including k-reliance probabilistic organizations wassuggested by Louzada and Ara [10]. According to the findings, the suggested ensemble model is more capable of accurate modelling. Utilizing a histogram-based exception score (HBOS) calculation to address the client's earlier way of behaving, Carminati et al. [11] fostered a half breed of semi-managed and unaided methodologies for extortion and inconsistency identification.

1.3 OBJECTIVE OF PROJECT

This project's overarching goal is to investigate the feasibility of using a deep neural network to identify fraudulent activities. To begin, we use an irregular timberland to fill in the spaces. From that point forward, the most separating highlights are picked utilizing a XGBoost calculation.Next, in order to address the imbalance in the sample, we suggest using a synthetic minority oversampling strategy.

2. LITERATURE SURVEY

2.1 The creation of a consumer credit fraud warning rule-based expert system model

Insurance firms have a significant hurdle when trying to identify fraudulent claims. Fraudsters aren't hesitant to use newer and more complex ways, even when detection chances are improving. It is not enough to just set up new fraud detection systems; the current ones also need to be upgraded and made as good as they can be. A rule-based expert system is a typical kind of detection system; it examines established rules and issues warnings when certain circumstances are satisfied. In most cases, the rules are examined independently, and relationships between the regulations are not given enough weight. In this study, we show how rule-based systems may benefit from association rule mining categorization by including relationships between rule pairs. A genetic optimizer is used to establish the rule weights.

2.2 The use of association principles to the detection of credit card fraud

When it comes to data mining models, association rules have been the most extensively researched. Here, we suggest using them to glean information about fraudulent exchanges from value-based charge card data sets, with the goal of preventing and detecting fraud. Credit card fraud data from some of Chile's most prominent retailers has been subjected to the suggested technique.

2.3 Framework for Financial Data Modelling Language (FFML) Development: A Language for Rule-Based Policy Modelling for Financial Data Streams Security Management

Financial organisations must have the competence to develop fraud detection systems and fraud management policies in order to diminish the adverse consequence of extortion on client care, monetary misfortunes, and the association's standing and brand. The limit of fraudsters to effectively re-engineer their techniques because of specially appointed security convention organizations is exhibited by the quickly changing assaults continuously monetary help stages. This features the unmistakable hole between the speed of exchange execution inside streaming monetary information and the relating misrepresentation innovation structures that shield the stage. To make it more straightforward to depict and apply proactive misrepresentation controls inside multi-channel monetary help stages, this paper subtleties the formation of FFML, a standard based strategy displaying language and a sweeping engineering. This article tells the best way to limit strategy displaying intricacy and sending latencies by utilizing a space explicit language to digest the monetary stage into an information stream based data model. The language is novel and can be involved by both master and non-master clients for strategy planning. As part of a larger suite of knowledge-based systems and assistive tools, FFML helps fraud analysts with their day-to-day tasks, such as creating new high-level fraud management policies, translating the underlying API into executable code, and implementing active monitoring and compliance features in the financial platform.

2.4 Using a Neural Network to Identify Credit Card Fraud

Credit card fraud has grown in frequency as the majority of us use them more and more for online purchases. Reason being, with the rise of internet shopping and other forms of electronic payment, frauds involving large sums of money have been more popular. To reduce the loss as much as possible, effective methods are required. To further deceive people into giving over their credit card information, con artists utilise methods such as phishing, impersonation, and bogus SMS and phone calls. Using a variety of machine learning techniques, such as support vector machine (SVM), k-nearest neighbour (KNN), and artificial neural network (ANN), this article aims to predict the probability of fraud. By combining deep learning with supervised machine learning, we can also identify valid transactions from fraudulent ones.

2.5 On atypical database transactions: Identification of probable frauds using machine learning for user profiling

In order to identify suspicious or fraudulent transactions, as well as changes in user conduct, this article presents a framework for generating profiles of usual user activity. We lay out the anomaly detection issue and talk about some of the earlier efforts to fix it. An algorithm to generate user profiles and another to recognise unusual transactions are both provided by the suggested method, which demonstrates the feasibility of building individual user profiles. There are also upper and lower limits given for the number of misclassifications. An analysis of this method is presented, along with a few areas that need more investigation.

2.6 Evaluation of categorization algorithms for financial risk prediction: an empirical study

For the purpose of early risk identification in the financial sector, a variety of categorization algorithms have been used recently. Predicting financial risks requires careful consideration of which classifiers, if any, will be most effective given a certain dataset. Depending on the performance metric and the specifics of the situation, prior research suggests that classifiers' abilities to forecast financial risk may differ. The primary objective of this research is to provide a two-stage procedure for assessing categorization algorithms in the context of financial risk prediction. The article ranks classifiers using a performance score that is constructed using three different various measures navigation (MCDM) calculations: TOPSIS, PROMETHEE, and

VIKOR. In this exact exploration, we utilize seven certifiable credit chance and misrepresentation risk datasets fromsix different nations to evaluate several categorization algorithms. The findings demonstrate that among the main three classifiers as indicated by TOPSIS, PROMETHEE, and VIKOR are straight calculated, Bayesian Organization, and troupe draws near. Besides, this study dives into the subject of building an information rich monetary gamble the board technique to improve the utility ofcategorization findings when detecting financial risks.

2.7 Enhanced neural networks trained via competition for the purpose of detecting fraud and network intrusion

Our study introduces two novel bunching calculations — the directed better serious learning organization (SICLN) and the superior cutthroat learning organization (ICLN) —for use in detecting network intrusions and fraud.

A new set of rules is applied to the traditional competitive learning neural network (SCLN) via the unsupervised clustering technique known as the ICLN. The ICLN's network neurons learn a new rule for updating rewards and punishments so that they can accurately portray the data's centre. In light of this new update rule, the SCLN is presently not temperamental. A directed variation of the ICLN is the SICLN. The SICLN's new directedupdating strategy improves clustering performance by training using data labels as input. Because of its extreme sensitivity to missing or delayed labels, the SICLN is applicable to data that is either labelled or unlabeled. Plus, the SICLN can fix its own mistakes and is hence totally unconcerned about the starting point for cluster size.We have conducted experimental comparisons using both real-world and research information in extortion recognition and organization interruption identification to assess the calculations thathave been suggested. The findings show that the SICLN achieves better performance than conventional unsupervised clustering methods, and that the ICLN also achieves good performance.

2.8 A decision tree method to fraud detection that takes cost into consideration

There has been a dramatic increase in global financial losses due to fraud as a consequence of advancements in information technology. When it comes to online credit card fraud, the most typical forms include fraudulent use of credit cards via virtual POS (Point of Sale) terminals or postal orders, even though fraud control methods like CHIP&PIN were created for credit card systems. The most effective and crucial measure to prevent this kind of fraud is, hence, fraud detection. This research compares the efficiency of two popular conventional characterization models utilizing a certifiable charge card dataset and afterward presents a clever expense delicate choice tree technique that limits the complete misclassification costs while picking the parting quality at each non-terminal hub. In this method, the costs of misclassification are assumed to be variable.

In comparison to the established, well-known methods, this cost-sensitive decision tree algorithm achieves better results on the provided problem set. This holds true not only for the commonly used metrics like precision and genuine positive rate, yet additionally for a recently characterized cost-delicate metric that is profoundly pertinent to the space of Mastercard extortion identification.

3. PROBLEM STATEMENT

The police force's service is often confined to crime response rather than prevention, despite the fact that crime prevention is a significant focus of the force, due to the short human resource capacity of the force compared to the population. The development of some wearable technology and smartphone apps has progressed over the years with the goal of protecting women. On the other hand, the majority of these apps and wearables use visual or auditory signals to alert contacts (guardians) or authorities. These methods are useless if a lady leaves the city or her guardians' care. These mechanisms fail miserably at immediately protecting women from harm because they rely on insufficient social action. There are frequently gaps in assuring women's safety due to the lack of integration between crime response, crime analysis, and crime prevention strategy.

3.1 LIMITATIONS

Rule-based models can't be updated quickly enough to account for new frauds since they rely on accurate and comprehensive expert knowledge.

4. ANTI FRAUD MODEL

A public loan dataset containing 200,000 records may be mined for fraud using a deep learning approach. Customers' real conditions may be better understood via our analysis of their credit rating. It seems to reason that a customer's propensity to engage in fraudulent activity increases in direct correlation with their credit rating. A lower score, such an E, suggests this. In order to construct anti-fraud regulations based on their customers' actual information, small lending organisations that operate on the internet put varying thresholds on their customers' credit rating data. The goal of this project is to help small financial credit organisations enhance their risk management and anti-fraud measures by providing them with a straightforward model.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, December : 2023

4.1 FEATURES OF ANTI FRAUD MODEL

Improved performance of the deep neural network bodes well for its potential use in detecting online fraud in the financial sector. Inadequate risk management capabilities coupled with a lack of resources for data engineering, optimisation, and modelling.

5. ARCHITECTURE

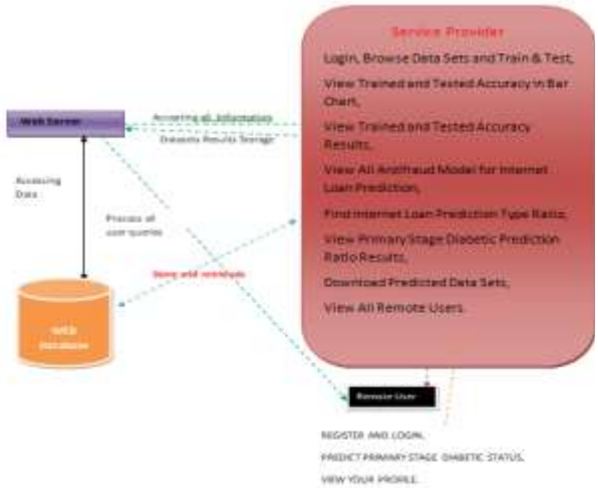

Fig. 5: System Architecture

6. IMPLEMENTATION

6.1 Service Provider

Access to the system by service providers is handled in this section. Once he logs in, he can do a lot of cool stuff, like check the trained and tested accuracies, see the predicted anti-fraud model type, compare the predicted anti-fraud model types, download the dataset with the predictions, and see who else is using the system remotely.

6.2 Remote User

Users must sign up for an account and then log in to access this module. Upon logging in, he is able to do things like examine profiles and anticipate the sort of anti-fraud model.

6.3 Algorithms used

6.3.1 Support Vector Machine

One of the most notable directed learning calculations, Backing Vector Machine (SVM) is utilized for both order and relapse situations. By the by, its essential use is in AI arrangement errands. The help vector machine (SVM) strategy looks to lay out the ideal choice limit that can parcel n-layered space into classes, permitting us to easily allocate new information focuses to the legitimate classification going ahead. In this case, a hyperplane is the optimal choice boundary. Support vector machines choose the hyperplane's extreme points and vectors.

6.3.2 Naïve Bayes

One managed learning approach that utilizes the Gullible Bayes calculation to address order issues is the one that depends on Bayes hypothesis. With a high-layered preparing dataset, it is for the most part utilized for text order. With regards to developing fast AI models that can create speedy forecasts, one of the best and simple arrangement calculations

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, December : 2023

is the Innocent Bayes Classifier. It makes forecasts in light of the probability of a thing happening, since it is a probabilistic classifier.

6.3.3 Logistic Regression

A logistic regression model may forecast the value of a dependant variable that is categorical. There can be no other consequence but a discrete or downright worth. It gives probabilistic qualities that fall somewhere in the range of 0 and 1, as opposed to the exactvalues of yes or no, true or false, etc.

7. RESULTS

Fig.7.1: Comparison graph of machine learning models.

Fig. 7.2: Line Graph of model accuracies

Fig. 7.3: prediction of loan approval using different models

Fig. 7.4: ratio of loan approval vs Not Approval

8. CONCLUSIONS

In this research, we sample from the lending club company's public loan data set using actual customer information. We proceed to construct an online fraud detection algorithm that relies on deep learning. In order to discover the best possible combination of model parameters, we first provide the model's primary parameters and then optimise them. Lastly, the suggested model is tested against the most common logistic regression in the financial sector and other comparable models to determine its performance. The findings show that the deep neural network outperforms the other options, which bodes well for its potential use in detecting online fraud in the banking sector.

9. Future Enhancement

We want to work with established Chinese banks and online financial tech firms on whitelists and blacklists in the near future. It is believed that the fraud detection capabilities may be enhanced by combining deep neural networks with expert anti-fraud rules, blacklists, and whitelists.

10. REFERENCES

[1] K. J. Leonard, ‘‘The development of a rule based expert system modelfor fraud alert in consumer credit,’’ Eur. J. Oper. Res., vol. 80, no. 2,pp. 350–356, Jan. 1995.

[2] D. Sánchez, M. A. Vila, L. Cerda, and J. M. Serrano, ‘‘Association rulesapplied to credit card fraud detection,’’ Expert Syst. Appl., vol. 36, no. 2,pp. 3630–3640, Mar. 2009.

[3] M. E. Edge and P. R. F. Sampaio, ‘‘The design of FFML: A rule-based policy modelling language for proactive fraud management in financial datastreams,’’ Expert Syst. Appl., vol. 39, no. 11, pp. 9966–9985, Sep. 2012.

[4] S. Ghosh and D. L. Reilly, Credit Card Fraud Detection With a NeuralNetwork. Wailea, HI, USA: IEEE, 1994.

[5] A. I. Kokkinaki, ‘‘On atypical database transactions: Identification ofprobable frauds using machine learning for user profiling,’’ in Proc. IEEEKnowl. Data Eng. Exchange Workshop, 1997, pp. 229–238.

[6] Y. Peng, G. Wang, G. Kou, and Y. Shi, ‘‘An empirical study of classificationalgorithm evaluation for financial risk prediction,’’ Appl. Soft Comput.,vol. 11, no. 2, pp. 2906–2915, Mar. 2011.

[7] J. Z. Lei and A. A. Ghorbani, ‘‘Improved competitive learning neuralnetworks for network intrusion and fraud detection,’’ Neurocomputing,vol. 75, no. 1, pp. 135–145, Jan. 2012.

[8] Y. Sahin, S. Bulkan, and E. Duman, ‘‘A cost-sensitive decision treeapproach for fraud detection,’’ Expert Syst. Appl., vol. 40, no. 15,pp. 5916–5923, Nov. 2013.

[9] N. SoltaniHalvaiee and M. K. Akbari, ‘‘A novel model for credit card frauddetection using artificial immune systems,’’ Appl.

Soft Comput., vol. 24,pp. 40–49, Nov. 2014.

[10] F. Louzada and A. Ara, ‘‘Bagging k-dependence probabilistic networks:An alternative powerful fraud detection tool,’’ Expert Syst. Appl., vol. 39,no. 14, pp. 11583–11592, Oct. 2012.

[11] M. Carminati, R. Caron, F. Maggi, I. Epifani, and S. Zanero, ‘‘BankSealer:A decision support system for online banking fraud analysis and investigation,’’ Comput. Secur., vol. 53, pp. 175–186, Sep. 2015.

[12] K. Fu, D. Cheng, Y. Tu, and L. Zhang, ‘‘Credit card fraud detection usingconvolutional neural networks,’’ in Proc. Int. Conf. Neural Inf. Process.,Oct. 2016, pp. 483–490.

[13] B. Tu, D. He, Y. Shang, C. Zhou, and W. Li, ‘‘Deep feature representation for anti-fraud system,’’ J. Vis. Commun. Image Represent., vol. 59,pp. 253–256, Feb. 2019.

[14] M. E. Greiner and H. Wang, ‘‘Building consumer-to-consumer trust in EFinance marketplaces: An empirical analysis,’’ Int. J. Electron. Commerce,vol. 15, no. 2, pp. 105–136, Dec. 2010.

[15] D. G. Pope and J. R. Sydnor, ‘‘What’s in a picture? Evidence of discrimination from prosper.com,’’ J. Hum.Resour., vol. 46, no. 1, pp. 53–92, 2011

[16] S. Freedman and G. Z. Jin, ‘‘The information value of online socialnetworks: Lessons from peer-to-peer lending,’’ Int. J. Ind. Org., vol. 51,pp. 185–222, Mar. 2017.

[17] M. Herzenstein, U. M. Dholakia, and R. L. Andrews, ‘‘Strategic herdingBehavior in peer-to-peer loan auctions,’’ J. Interact. Marketing, vol. 25,no. 1, pp. 27–36, Feb. 2011.

[18] M. Herzenstein, R. L. Andrews, and U. M. Dholakia, ‘‘The democratization of personal consumerloans? Determinants of successin online peer-topeer lending communities,’’ J. Marketing Res., vol. 15, pp. 274–277, 2008.

[19] N. Mykhalchenko and J. Wiegratz, ‘‘Anti-fraud measures in SouthernAfrica,’’ Rev. Afr. Political Economy, vol. 46, no. 161, pp. 496–514,Jul. 2019.

[20] B. Riley, ‘‘Anti-fraud technologies: A business essential in the card industry,’’ Card Technol. Today, vol. 19, no. 10, pp. 10–11, Oct. 2007.

[21] L. Breiman, ‘‘Random forests,’’ Mach. Learn., vol. 45, no. 1, pp. 5–32,2001.

[22] T. Chen and C. Guestrin, ‘‘XGBoost: A scalable tree boosting system,’’in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining,Aug. 2016, pp. 785–794.

[23] T. Pan, J. Zhao, W. Wu, and J. Yang, ‘‘Learning imbalanced datasts basedon SMOTE and Gaussian distribution,’’ Inf. Sci., vol. 512, pp. 1214–1233,Feb. 2020.

[24] E. Micheu-Tzanakou, ‘‘Artificial neural networks: An overview,’’ Netw.,vol. 22, nos. 1–4, pp, pp. 208–230, 2011.

[25] T. Zhou, G. Han, X. Xu, Z. Lin, C. Han, Y. Huang, and J. Qin, ‘‘δ-agreeAdaBoost stacked autoencoder for short-term traffic flow forecasting,’’Neurocomputing, vol. 247, pp. 31–38, Jul. 2017.