

Retraction

Retracted: Deep Learning-Based Consumer Behavior Analysis and Application Research

Wireless Communications and Mobile Computing

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] Y. Zhang, A. Wang, and W. Hu, "Deep Learning-Based Consumer Behavior Analysis and Application Research," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 4268982, 7 pages, 2022.

Research Article

Deep Learning-Based Consumer Behavior Analysis and Application Research

Yuan Zhang , Aiqiang Wang, and Wenxin Hu

School of Economics & Management, Beijing Institute of Petrochemical Technology, Daxing, 102617 Beijing, China

Correspondence should be addressed to Yuan Zhang; zhangyuan@bipt.edu.cn

Received 7 March 2022; Accepted 28 March 2022; Published 26 April 2022

Academic Editor: Mohammad Farukh Hashmi

Copyright © 2022 Yuan Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The core of consumer purchase behavior analysis lies in building prediction models. This paper combines consumer behavior prediction with deep learning models, proposes rDNN models and KmDNN models, uses AUC and F value as evaluation indicators, implements algorithms using Python as experimental tools, derives prediction results, and conducts comparative analysis. Focusing on deep neural network models in deep learning, the performance of deep neural network models for consumer purchase behavior analysis is explored from three aspects: the underlying theory of deep neural network models, the construction and implementation of the models, and the improvement of the models, and empirical analysis is conducted through experimental results.

1. Introduction

In recent years, the rapid development of Internet technology has promoted the development of e-commerce [1]. The scale and forecast of e-commerce transactions show that e-commerce has a broad development prospect. The development of e-commerce has led to the emergence of online retailing, with the volume of online transactions showing a sharp rise. The emergence of e-commerce has changed the traditional trading model, and the timely, convenient, and even visual communication between businesses and consumers makes it easier for businesses to understand the market and the real, personalized needs of their customers, solving the problem of time-consuming and inefficient traditional transactions [2]. In order to better meet consumers' needs, e-commerce platforms both at home and abroad need to tap into consumers' buying habits and preferences in their growing shopping logs, so as to refine product demand, identify the range of consumers for each product, and carry out precise marketing and recommendation services [3].

It is a challenge for e-commerce platforms to analyze the intrinsic patterns, consumption levels, personality preferences, and other implicit characteristics of consumers' purchasing behavior, to estimate the likelihood of consumers

purchasing products and to make accurate recommendations based on the predicted results. The success of e-commerce companies lies in the fact that they have access to the personal information, shopping data, and consumption habits of many consumers [4]. Bringing the latest scientific technology into consumer behavior analysis, improving the accuracy of analysis and prediction, enabling one-to-one accurate recommendations, and developing effective marketing strategies are the key issues that need to be addressed today. For e-commerce companies, such as Tmall, Jingdong, and other large e-commerce websites, the number of products is huge and showing a rapid growth trend. How to develop marketing strategies, accurate advertising, and to show consumers the right goods is the dilemma of e-commerce enterprises to face; for ordinary consumers, having the same background consumers in the purchase of goods, commodity prices, sales volume, the amount of good and bad reviews, collection volume, and other related factors are their concerns; for the social economy, these are the following factors: the development of social consumption trends, the development of regional consumption trends, and industry. The development of consumption trends is one of the most important concerns of social entrepreneurs. In order to achieve the goal of benefiting all three parties, there is an urgent need to face

the problem of how to effectively analyze consumer behavior and accurately grasp consumer demand trends [5].

In order to solve these problems, the thesis will be based on the consumer behavior data of consumers interacting with goods in the past month from the perspective of big data, extracting the implicit features affecting consumer behavior from different perspectives, learning the internal patterns of the data with different models, respectively, constructing a predictive model for consumer behavior analysis, realizing the possibility of predicting consumers' future purchases of goods, and providing theoretical and technical references for e-commerce platforms and their marketing strategy formulation [6].

2. Related Work

The main objective of consumer behavior analysis is to identify products that are valuable to consumers and that consumers are more likely to buy. Consumer behavior analysis is a key part of accurate recommendations, and researchers in different fields have provided different ideas to solve the problem from the perspective of consumer behavior processes [7].

In response to the increasingly complex influencing factors, some researchers have started from the data itself, cleaned, transformed and generalized the data, and used machine learning models for mining. In [8], we took the purchase information data of automobile customers as the research object, used SQL server to store the automobile marketing analysis data as the input data for the improved ID3 behavioral tree model and association rule model in the paper, and mined the data to compare the prediction results of the two models, and the results showed that the improved ID3 behavioral tree had higher prediction accuracy [9]. The improved behavioral tree model was used to analyze and predict consumer behavior. The results of the experiments showed that the improved model was more effective in predicting consumer behavior, which in turn proved the value and potential of the improved decision tree model. Based on the data mining research background, [10] uses multiple multidimensional cube clustering operators to classify and cluster mobile communication customers, solving the problem of customer's consumption behavior analysis, providing new ideas and methods for consumer behavior analysis methods. [11] takes the shopping behavior of consumers on the Tmall website within a certain time period as the research context and proposes a machine learning method based on model combination to predict consumer purchase behavior. Using the user product interaction data provided by Ali as the background, [12] gradually carried out data preprocessing, selecting sample data, constructing features, and building prediction models and evaluation models to make predictions on users' consumption behavior. In the paper, logistic regression and iterative decision tree are used to construct the prediction models, respectively. The iterative decision tree algorithm has better prediction effect through the validation comparison of the test set. The customer mining model built by [13] for customer analysis is used throughout the paper to firstly preprocess

the customer data and extract important user features using neural networks, followed by associative classification of customers in order to obtain the consumption of each category of customers, the purpose of classification is to make Bayesian inference for each category of customers, and finally, the Bayesian approach is used to predict customer consumption. Using users' bank card payments as the research context, [14] proposed a behavioral prediction method based on secondary clustering and Hidden Markov Chain (HMC) theory, using penalty factors to cluster users' consumption behavior after clustering on the basis of secondary clustering and then using HMC theory to estimate the shift of consumption hierarchy states in the sequence to predict consumers' future consumption behavior [15]. The experimental results show that the prediction accuracy of the plain Bayesian method reaches more than 70%, while the prediction accuracy of the neural network model can reach more than 85%, indicating that the artificial neural network is more capable of learning the relationship between the consumer behavior and the consumer behavior than the plain Bayesian. The relationship between consumer behavior characteristics is obtained [16]. The authors of [17] proposed four key problems in modeling consumer behavior and predicting consumer behavior, namely, preparing large amounts of data, labeling data, the existence of conceptual drift, and computational complexity, based on these four key problems the authors proposed to use machine learning methods to model consumers to predict consumer behavior [18].

3. Statistical Analysis of Data

3.1. Consumption Category Statistics. The collected consumption items were divided into 9 categories of 159 fields, and the specific consumption shares of the different categories of consumption types were counted. The two more extreme cases are obtained in this paper, those with a monthly consumption of over 10,000 and the two cases with less than 2,000, differed considerably in terms of their spending power, but had similar consumption shares, reflecting similarities in their consumption perceptions. It is worth noting that the higher proportion of high consumers spend on entertainment is related to consumption habits and the difference in value between entertainment items and basic living expenses [19].

As shown in Table 1, the focus of this paper is on the characteristics that characterize "irrational consumption" in terms of consumption behavior, so that the low spending power and the high spending on entertainment or beauty are more important characteristics. This is in line with the common perception of irrational consumption.

3.2. Consumption Time Statistics. The assessment of the timing of consumption provides more feedback on one aspect of the characteristics of consumption habits. Generally speaking, groups that concentrate their spending on holidays and weekends tend to have a stable work situation, which facilitates their ability to meet their financial obligations and repay their loans on time. As shown in Table 2

TABLE 1: Statistics on consumption categories.

Feature category	Category description	Characteristic field	Number of fields (PCs.)
Amuse_info	Entertainment consumption	Consumption of K IV V, bars, bath centers, etc.	21
Dailyshop_info	Daily commodity consumption	Food, furniture, toiletries, etc.	27
Largegoods_info	Large commodity consumption	Automobile, housing, decoration, study abroad, etc.	25
Generalgoods_info	Consumption of medium value goods	Mobile phones, computers, etc.	27
Life_info	Living consumption	Water, electricity, internet, insurance, etc.	12
Medical_info	Health care consumption	Medicine, treatment, fitness and other expenses	15
Educate_info	Educational consumption	Educational institutions, books, etc.	8
Decorate_info	Beauty and clothing consumption	Consumption of clothing, bags and cosmetics	6
Travel_info	Travel consumption	Transportation, hotels, catering, etc.	18
Total			159

TABLE 2: Consumption time statistics.

Characteristic field	Field description	Number of fields (PCs.)
Month info	Month of consumption	1
Day info	Date of consumption (day of week)	1
Time info	Period of consumption (every 2 h)	1
Festival long info	Whether it is a holiday (long holiday > 3 d)	1
Festival short info	Whether it is a holiday (short holiday < 3 d)	1
Total		5

groups that concentrate their spending on long holidays tend to have a habit of travelling for holidays, this group tends to have a better financial background and is usually less likely to be late in repaying their loans. Conversely, large purchases that occur at particular times (late at night) and on particular days (working days) are likely to be “special purchases” in an emergency situation or sporadic purchases by the “unemployed.” Whether it is “special consumption” in an emergency situation or sporadic consumption by the “unemployed,” when the consumption accounts for a large proportion of the total, it is reasonable to doubt their financial ability to support the repayment on time; so, this aspect is logically related to financial credit and can be used as input layer for deep learning.

3.3. *Statistics on Consumption Habits.* The user’s spending habits are a processed feature to better establish the target mapping as shown in Figure 1, where the impulse consumption index is

$$Im_con = con_con/month, \quad (1)$$

where Im_con is the impulse consumption index, con_con refers to the number of consecutive consumptions, and month is the month, where con_con is defined as consumption across 5 categories in a 1d period. In general, a large amount of consumption across categories in a short period of time is often indicative of “impulsive consumption” in a certain consumer environment and stimulus. It is easy to assume that people who spend impulsively are more likely to be late with their payments or have poor financial literacy

and behavior. The Consumer Concentration Index is calculated as

$$Fo_con = Lar_amon/month, \quad (2)$$

where Fo_con is the consumption concentration index and $Lar_amon/month$ is the number of large consumptions in a month, where large consumption is defined as consumption exceeding 20% of the user’s average monthly total consumption (monthly average consumption is the monthly average over a year).

The consumption distribution index, as shown in Table 3, refers to the concentration of consumption behavior, i.e., a user who spends more than 60% of their total consumption in any 2 categories in the month is considered to have spent a significant amount of money 1 time.

4. Building RDNN Models

From the data statistics and discussion in Section 3, it is clear that the categories of the sample data are unbalanced; so, this subsection designs an improved DNN model, the RDNN model, to address the characteristics of unbalanced categories of consumer behavior data.

The RDNN model is an extension of the DNN model, which differs from the DNN model in that the negative samples of data are randomly sampled before using the DNN model, this model will be more effective for unbalanced data, and the proportion of unbalanced data is experimentally explored.

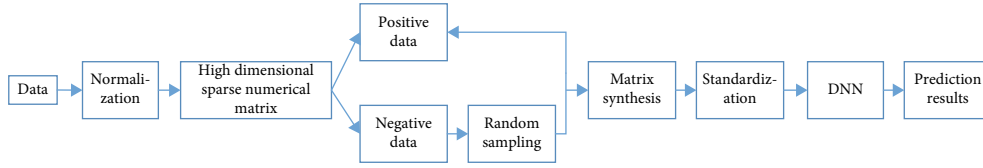


FIGURE 1: rDNN model.

TABLE 3: Statistics on consumption habits.

Characteristic field	Field description	Number of fields (PCs.)
Impulse consumption index *	Number of consecutive consumption/month	1
Consumption concentration index *	Number of consecutive consumption/month	1
Consumption distribution index *	Index representing consumption preference	1
Total		3

The idea of building an improved model, rDNN, is to retain the ability of the DNN model to dig deeper into the data and reduce the impact of data category imbalances on the effectiveness of the model. rDNN models can reduce the training cost of the model, eliminate the redundancy of negative data samples, and automatically learn features at the bottom of the network to uncover more valuable information about the data, providing a new technical tool for consumer behavior analysis and prediction.

4.1. Selecting the Right N/P Ratio. To reduce the impact of data category imbalance on model performance, the ratio of negative to positive samples is also known as the negative positive sample ratio (N/P) [20]. The paper balances the negative sample with the positive sample by drawing a random subsample from the negative sample, and the number of subsamples drawn forms a certain ratio with the number of positive samples (as shown in Figure 2).

In Figure 2, the negative samples are in black and the positive samples in white, the negative samples are randomly sampled to achieve a balanced ratio with the positive samples, and the N/P ratio in the figure is 1. The study of the N/P ratio can select the right amount of positive and negative samples, avoid the homogeneity of data characteristics to a certain extent, and enhance the generalization ability of the model.

4.2. Model Construction and Algorithm Implementation. The data set is divided into a training set and a test set, and a deep learning model is constructed on this basis (shown in Figure 3). In the deep learning model construction, the advantages and disadvantages of DNN, rDNN, and KmDNN models will be compared first, and the results will be analyzed. The KmDNN model is another improvement of the DNN model, which is a deep neural network model with random sampling based on clustering of negative class data samples, and this model will only be used for comparative analysis in this section.

In Figure 3, three deep learning models with DNN as the core are constructed, and the rDNN model and KmDNN

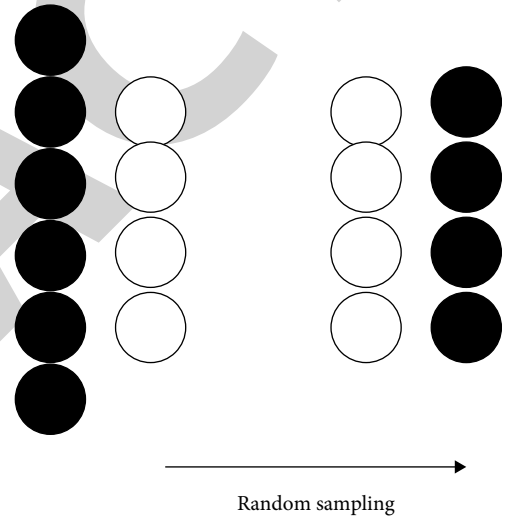


FIGURE 2: Schematic of N/P ratio of 1.

model are introduced according to the imbalance of the category data.

5. Experimental Results and Analysis

Before using the DNN model, the negative samples were randomly sampled, resulting in different results for each experiment; so, the experiment was repeated 30 times, the mean value was taken as the final result, and Table 4 shows the experimental results of various methods:

Table 3 shows that the DNN improvement algorithm has a significant effect on improving the AUC. The rDNN model gets a significant improvement in the prediction effect on the basis of the DNN model for negative sample random sampling; therefore, reducing the proportional disparity between positive and negative samples is important for the prediction of the model; in other words, the balance of the category data has a significant impact on the model to perform well. Due to the sheer volume of data in the thesis and the limited experimental environment and platform,

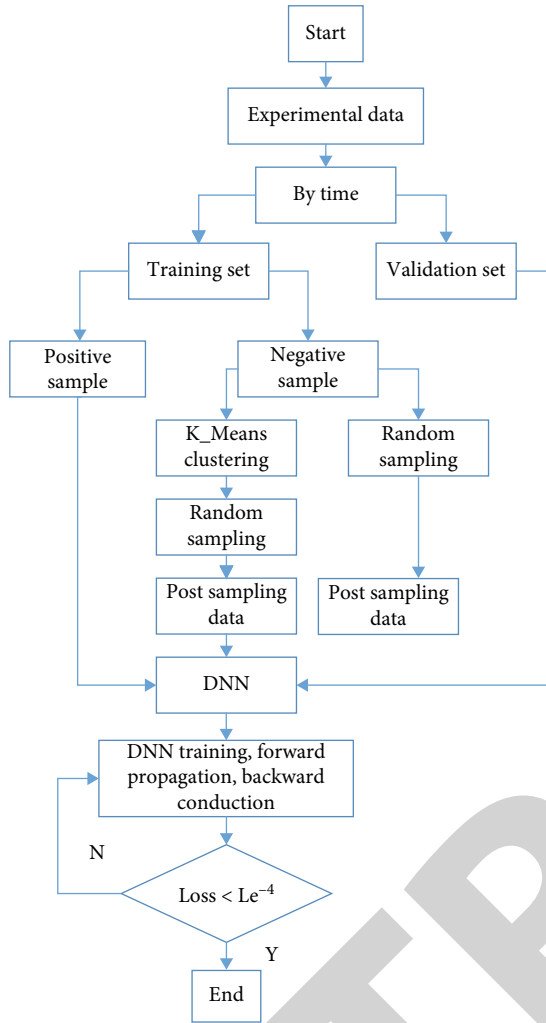


FIGURE 3: Deep learning model construction flowchart.

TABLE 4: Comparison of experimental results.

Deep learning model	Average AUC
DNN	0.7893
rDNN	0.8322
KmDNN	0.8064

only the treatment of random sampling of negative samples was done. Further comparisons of methods for dealing with unbalanced data, such as experimenting with methods such as oversampling, can be conducted on a case-by-case basis in subsequent experiments. In real life, oversampling methods have greater limitations. The data in this experiment has 7507 positive samples and 2106772 negative samples, and using oversampling methods increases the sample size, requires greater memory consumption and computation time, and is not very practical. Therefore, random sampling can be more suitable for realistic needs [21, 22].

The KmDNN model with the introduction of K_mens clustering also improves the results, but when clustering negative samples, the number of clusters needs to be set in

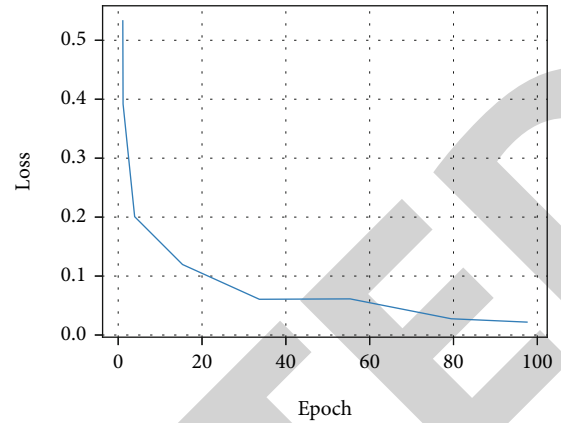


FIGURE 4: rDNN training set iterative process.

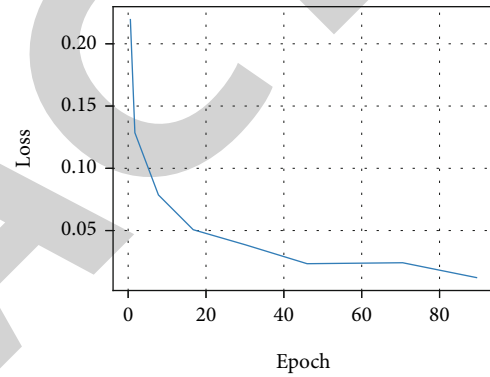


FIGURE 5: rDNN validation set iterative process.

advance, the compatibility between algorithms is different, the theoretical interpretability is not strong, and each cluster in the clusters does not have equal negative samples; so, the sampling cannot be guaranteed to be drawn randomly and with equal probability, thus leading to biased results, and this method needs to be improved. The DNN-based model produced better predictions than the traditional prediction model analyzed, thus further confirming that the deep learning model has a stronger learning capability and more powerful feature representation, as shown in Figures 4 and 5.

According to the image test of the iterative process of rDNN training and validation sets, when the number of iterations reaches 160, the drop is slow, and the model is well trained. When the number of iterations reaches 160, the decline is slow, and the model is well trained.

The different consumption effects are shown in Figure 6, by comparing the experimental results based on DNN models and improved models, the improved models KmDNN and rDNN models are better than DNN models, reducing the proportion of data imbalance before building the model, which has a greater improvement on the prediction effect of the model. The deeper learning models are more effective than the DNN and rDNN models, and the deeper feature learning helps the models to produce better results. In the consumer behavior analysis and prediction problem, the proposed rDNN model retains the ability of

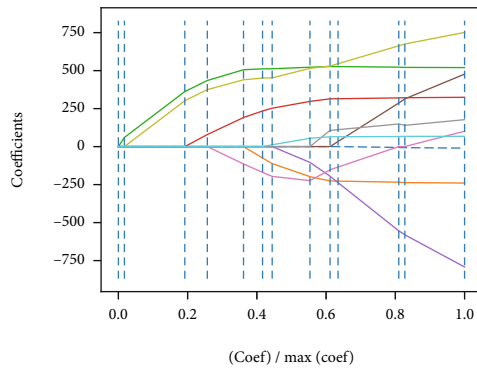


FIGURE 6: Different consumption factors.

the original DNN to automatically learn deeper features and abstract higher-level features and incorporates methods to solve the category data imbalance problem on top of this, reducing the amount of data and easing the DNN training burden, which has better practicality.

6. Conclusions

On the basis of consumer behavior data processing and feature engineering, deep learning models are studied in depth, and a deep learning model framework is designed and constructed, differing from traditional analysis and prediction models. In the deep learning models for consumer behavior analysis and prediction, we focus on exploring DNN models and improving on DNN models, proposing rDNN models and KmDNN models, using AUC and F values as evaluation indicators, using Python as an experimental tool, implementing algorithms, deriving prediction results, and conducting comparative analysis.

Data Availability

The dataset used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

Acknowledgments

Project of Beijing Municipal Education Commission, project name: comparative study on tourism consumption patterns of Chinese and foreign college students in Beijing, Tianjin, and Hebei, No. SM201610017002.

References

- [1] G. Yolcu, I. Oztel, S. Kazan, C. Oz, and F. Bunyak, "Deep learning-based face analysis system for monitoring customer interest," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 1, pp. 237–248, 2020.
- [2] V. H. Lee, J. J. Hew, L. Y. Leong, G. W. H. Tan, and K. B. Ooi, "Wearable payment: a deep learning-based dual-stage SEM-ANN analysis," *Expert Systems with Applications*, vol. 157, article 113477, 2020.
- [3] G. A. Abbasi, L. Y. Tiew, J. Tang, Y. N. Goh, and R. Thurasamy, "The adoption of cryptocurrency as a disruptive force: deep learning-based dual stage structural equation modelling and artificial neural network analysis," *PLoS One*, vol. 16, no. 3, article e0247582, 2021.
- [4] A. Generosi, S. Ceccacci, and M. Mengoni, "A deep learning-based system to track and analyze customer behavior in retail store," *2018 IEEE 8th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*, 2018, pp. 1–6, Berlin, Germany, 2018.
- [5] A. Pumsirirat and L. Yan, "Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 1, pp. 18–25, 2018.
- [6] Y. Wang, Q. Chen, D. Gan, J. Yang, D. S. Kirschen, and C. Kang, "Deep learning-based socio-demographic information identification from smart meter data," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2593–2602, 2019.
- [7] L. F. Maimó, Á. L. P. Gómez, F. J. G. Clemente, M. G. Pérez, and G. M. Pérez, "A self-adaptive deep learning-based system for anomaly detection in 5G networks," *Ieee Access*, vol. 6, pp. 7700–7712, 2018.
- [8] Y. Mehta, N. Majumder, A. Gelbukh, and E. Cambria, "Recent trends in deep learning based personality detection," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2313–2339, 2020.
- [9] D. Chen, P. Wawrzynski, and Z. Lv, "Cyber security in smart cities: a review of deep learning-based applications and case studies," *Sustainable Cities and Society*, vol. 66, article 102655, 2021.
- [10] K. Yu, L. Lin, M. Alazab, L. Tan, and B. Gu, "Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4337–4347, 2021.
- [11] P. Sundaravadivel, K. Kesavan, L. Kesavan, S. P. Mohanty, and E. Kougianos, "Smart-log: a deep-learning based automated nutrition monitoring system in the iot," *IEEE Transactions on Consumer Electronics*, vol. 64, no. 3, pp. 390–398, 2018.
- [12] H. He, C. K. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, 2018.
- [13] X. Tao, Z. Chunjiang, and X. Yongjian, "Collaborative parameter update based on average variance reduction of historical gradients," *Journal of Electronics and Information Technology*, vol. 43, no. 4, pp. 956–964, 2021.
- [14] X. Liu, D. Lee, and K. Srinivasan, "Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning," *Journal of Marketing Research*, vol. 56, no. 6, pp. 918–943, 2019.
- [15] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: a deep learning-based intelligent mechanism," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.
- [16] L. Zhong, L. Hu, and H. Zhou, "Deep learning based multi-temporal crop classification," *Remote Sensing of Environment*, vol. 221, pp. 430–443, 2019.
- [17] C. Liu, Y. Cao, Y. Luo et al., "A new deep learning-based food recognition system for dietary assessment on an edge

- computing service infrastructure,” *IEEE Transactions on Services Computing*, vol. 11, no. 2, pp. 249–261, 2018.
- [18] P. An, Z. Wang, and C. Zhang, “Ensemble unsupervised auto-encoders and Gaussian mixture model for cyberattack detection,” *Information Processing & Management*, vol. 59, no. 2, article 102844, 2022.
- [19] S. Zhang, L. Yao, A. Sun, and Y. Tay, “Facial expression analysis under partial occlusion,” *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, pp. 1–49, 2019.
- [20] S. Tuli, N. Basumatary, S. S. Gill et al., “HealthFog: an ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments,” *Future Generation Computer Systems*, vol. 104, pp. 187–200, 2020.
- [21] E. Zinovyeva, W. K. Härdle, and S. Lessmann, “Antisocial online behavior detection using deep learning,” *Decision Support Systems*, vol. 138, article 113362, 2020.
- [22] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, “Deep learning-based text classification,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 3, pp. 1–40, 2021.