International Journal of
*Environmental Research
and Public Health*

MDPI

*Article*

# Integrating Individual Factors to Construct Recognition Models of Consumer Fraud Victimization

**Liuchang Xu [1], Jie Wang [2], Dayu Xu [1] and Liang Xu [2,*]**

[1] College of Mathematics and Computer Science, Zhejiang A&F University, Hangzhou 311300, China; xuliuchang@zafu.edu.cn (L.X.); xdysie@zafu.edu.cn (D.X.)

[2] Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310058, China; wj_psy@zju.edu.cn

* Correspondence: xuliang_psy@zju.edu.cn

**Abstract:** Consumer financial fraud has become a serious problem because it often causes victims to suffer economic, physical, mental, social, and legal harm. Identifying which individuals are more likely to be scammed may mitigate the threat posed by consumer financial fraud. Based on a two-stage conceptual framework, this study integrated various individual factors in a nationwide survey (36,202 participants) to construct fraud exposure recognition (FER) and fraud victimhood recognition (FVR) models by utilizing a machine learning method. The FER model performed well (f1 = 0.727), and model interpretation indicated that migration status, financial status, urbanicity, and age have good predictive effects on fraud exposure in the Chinese context, whereas the FVR model shows a low predictive effect (f1 = 0.565), reminding us to consider more psychological factors in future work. This research provides an important reference for the analysis of individual differences among people vulnerable to consumer fraud.

**Keywords:** consumer financial fraud; individual factors; machine learning; fraud exposure; fraud victimization

check for updates

**Citation:** Xu, L.; Wang, J.; Xu, D.; Xu, L. Integrating Individual Factors to Construct Recognition Models of Consumer Fraud Victimization. *Int. J. Environ. Res. Public Health* **2022**, *19*, 461. https://doi.org/10.3390/ijerph19010461

Academic Editor: Jimmy Efird

Received: 22 November 2021
Accepted: 21 December 2021
Published: 1 January 2022

## 1. Introduction

The rapid development of mobile technology has facilitated social communication crossing geographical divides, but it has also increased the chances of fraudsters defrauding money [1]. In recent years, consumer financial fraud, defined as "crimes against consumers in which deceptive or false acts are committed for personal financial gain" [2,3], has attracted increasing attention from academia and government. Consumer financial fraud not only poses a threat to the consumer economy [4–7] but also causes victims to suffer physical, mental, social, and legal harm [8–15]. The great threat to public health makes anti-fraud an urgent issue.

Consumer financial fraud is essentially a process of deception used to defraud consumers of their money (and perhaps personal information) in a consumer context (e.g., when buying something). From the perspective of information transmission, Shadel and Pak indicated that fraudsters often portray themselves as a positive role model to convey fraudulent information, and some recipients of the information may believe it for specific reasons, leading to being deceived [16]. Grazioli used an information-processing model of deception detection to study the underlying reasons of consumers' failure to detect intentional deception and found that the deceived usually rely on "trust" cues and heavily discount "assurance" cues [17]. After noting the importance of cue processing in the fraud process, Wright et al. proposed that individual factors also need to be considered [18]. The Process Model of Deception Detection improved by Wright et al. indicated that, in a fraud scenario, different individuals have different cue processing, evaluation, and decision-making methods [18]. Therefore, investigating who is more likely to be deceived has become a critical topic in fraud research.

Demographic factors, such as sex, age, education experience, and financial status, were the most frequently considered individual factors in previous fraud studies. For age, researchers found that older adults are less exposed to fraud because they are less likely to go online [18–21]. However, once older adults encounter fraud, they are more likely to become victims of fraud, that is, once they are targeted (exposed to fraud), they are more likely to be victimized. Older adults may face aging-related cognitive decline and retirement-related social isolation [4,22,23]. For financial status, Kerley and Copes found that people with incomes between USD 15,000 and 24,000 are more likely to be victims of fraud [19]. Sex and education experience may affect the individual's probability of victimization in certain fraud scenarios. For example, although sex does not affect the victimization of fraud in general [19,24], males are more likely to be victims of investment fraud, lottery fraud, and advertising fraud [16,25]. For education experience, groups with low education levels are more likely to encounter loan and lottery fraud [26] but less likely to encounter investment fraud [27].

The above studies presented various pieces of evidence of the correlation between demographic factors and fraud, while psychologists are dedicated to finding the psychological factors that affect the victimization of fraud [28]. Previous psychological studies have shown that various psychological traits, such as personality, self control, aloneness, impulsion, cognitive ability, and emotional status, may affect the individual's probability of being deceived [29–36]. For example, Gottfredson and Hirschi indicated that individuals with low self-control are more likely to be impulsive, focus on the present, and pursue immediate pleasure, so they are more likely to be involved in fraud [37]; van de Weijer and Leukfeldt found that individuals with low conscientiousness, low neuroticism, and high openness are more likely to be deceived [38]; inducing individual fear emotions can increase the individual's trust in fraudulent information [39,40]; and Fischer et al. found that fraud victimization is related to high motivation, trust and excessive self-confidence [41]. In addition, according to the depth interviews with fraud victims, Button et al. summarized a series of reasons for the victims of online fraud, including embarrassing fraud, visceral appeals, pressure, and other psychological factors [42]. These psychological studies have given us a deeper understanding of the individual factors that affect the victimization of fraud.

Whether studying demographic or psychological factors, these studies help us understand why some individuals appear to be more susceptible to being deceived [43]. However, most of the studies have focused on a single or small number of factors and rarely considered different types of individual factors together. With the development of machine learning (ML) technology, it is possible to compare the effects of different features from the perspective of computational modeling [44]. Thus, can we use ML technology to compare the importance of individual factors in fraud scenarios? Can we build a predictive model to identify which individuals are more susceptible to fraud? These are the questions that this research intends to explore.

In fact, in different stages of fraud, the role of individual factors may also be different. As mentioned before, older adults are less exposed to fraud but more likely to become victims of fraud once exposed [16,22]. This reminds us that fraud has two important stages: fraud exposure and fraud victimization. Fan and Yu recently developed a two-stage conceptual framework to investigate age-related differences in fraud exposure and fraud victimization [4]. The first stage of fraud exposure was defined as whether consumers experienced fraud regardless of whether they were victimized, and the second stage was whether consumers became victims (lost money) of fraud that they were exposed to [4]. The above work found that some factors (e.g., sex, assets, and region) have a significant impact on fraud exposure but have no influence on whether they are victimized. In fact, previous research on fraud mechanisms has mainly focused on investigating the second stage of fraud, that is, why some people are more easily deceived [31,43]; however, there is less research on the first stage. Therefore, to better investigate fraud mechanisms, this study built two predictive models of fraud exposure and fraud victimization.

In summary, the present study uses individual factors as inputs to construct fraud exposure recognition (FER) and fraud victimhood recognition (FVR) models. From a practical perspective, we built a predictive model to automatically identify which individuals are more likely to be exposed to fraud and which individuals are more susceptible to fraud. Identifying individuals susceptible to fraud in advance may help us reduce the number of victims. From a theoretical perspective, we used a computational modeling method to compare the impact of various individual factors on fraud exposure and fraud victimhood. The results of factor comparison can provide an important reference for subsequent research on individual differences in fraud.

## 2. Materials and Methods

### *2.1. Data and Participants*

The present work used the China Household Finance Survey (CHFS) microdata, a public-use database collected by the Survey and Research Center for China Household Finance [45], to achieve our objectives. The CHFS is a nationwide biennial survey of Chinese household finances containing information about household demographics, geographic location, assets and liabilities, income and expenditure, employment, consumer fraud, and so forth [4,45]. The 2015 CHFS database (Accessible at https://chfser.swufe.edu.cn/datas/Products/Datas/DataList, accessed date 10 December 2020) [45], including 37,289 households in 351 counties of 29 Chinese provinces, was applied in this study. After removing samples with severe missing data, the data of 36,202 households were considered in subsequent modeling and analysis. In addition, since the person who is the most knowledgeable about their household finance was interviewed in CHFS [4], he or she was regarded as the household reference person here. The average age of the final 36,202 participants (47.29% females) was 52.83 years (SD = 14.95).

### *2.2. Measurements and Feature Processing*

This section describes how the variables were measured in the 2015 CHFS and how they were processed for subsequent modeling. The main features, including fraud exposure, fraud victimhood, demographic features, and financial-related features, are introduced in detail. However, considering the length of the article, the complete feature list (a total of 150 features) is presented in Supplemental Materials Table S1.

#### 2.2.1. Fraud Exposure and Victimhood

Consumer fraud exposure was investigated before fraud victimhood. First, the survey asked "whether the household encountered the following consumer fraud over the past year", including telephone fraud, SMS fraud, social application (such as QQ and WeChat) fraud, phishing, fraud from acquaintances, and other methods. Multiple options were allowed for this question. If the respondent encountered at least one of the fraud methods, the household was considered exposed to consumer fraud, and then he or she was asked "whether the household has suffered monetary losses as a result of fraud". If the respondent answered "yes", the household was regarded as a consumer fraud victim. Consumer fraud exposure and victimhood were treated as ground truth in our predictive models. As binary features, they were labeled as 1 (exposed to consumer fraud; fraud victim) or 0 (not exposed to consumer fraud; fraud survivor).

#### 2.2.2. Demographic Features

Demographic information included age, sex (male of female), education (from 1 no schooling at all to 9 doctorate degree), marital status, employment (having a job or not), self-evaluation of physical condition, migration status (having which type of registered residence), region (eastern, western, or central China), resident type (rural or urban), financial knowledge, and political status (whether a Chinese Communist Party member). As model inputs, all single-choice multi-categorical features were labeled as different numbers from 1 to $n$ ($n$ is the number of options), all multiple selection features were transformed into dummy variables, and all continuous inputs were scaled to a value between 0 and 1. The above feature-processing methods were also used to process other input features.

#### 2.2.3. Financial-Related Features

Defrauding money is the main purpose of customer financial fraud [3], so the individual's financial-related features are the key considerations of this research. The 2015 CHFS collected information about assets, liabilities, income, and expenditures, and each feature had multiple molecular dimensions. For instance, income information included not only the total income but also sources of income, such as lottery winnings, the sale of a house, sale of intellectual property, and so forth; expenditures included expenditures on different items, such as food, transportation, communication, luxury, education, and so forth. Most of the molecular dimensions were transformed into independent features as model inputs.

#### 2.2.4. Other Features

In addition to the above main features, other potentially relevant features, such as risk appetite and subjective well-being, were also treated as model inputs. For example, participants' risk appetite was considered an input in this study because risk appetite may influence an individual's financial behavior [46,47]. The risk appetite was investigated by two questions (e.g., "Which of the choice below do you want to invest most if you have adequate money?" from 1 "project with high-risk and high-return" to 5 "unwilling to carry any risk"), and the answer to each question was used as an input feature. Subjective well-being was measured by asking "how happy does the respondent feel", and the respondent was asked to respond on a scale from 1 "extremely happy" to 5 "extremely unhappy". In sum, the name and description of all features are presented in Supplemental Materials Table S1.

### *2.3. Model Construction and Evaluation*

The present work constructed two recognition models of fraud exposure and fraud victimhood. Both fraud exposure and victimhood were binary features, so we formulated fraud exposure recognition and fraud victimhood recognition as classification problems. We used fraud exposure and fraud victimhood as ground truth, used the other features introduced above as inputs, and applied a machine learning algorithm to construct the predictive models.

For the machine learning algorithm, this work applied the random forest classification (RFC) algorithm. As RFC has shown good performance in classification tasks [48,49], we can easily interpret the constructed RFC models by calculating the feature importance [50–54]. The predictive effect of each model was evaluated by the tenfold cross-validation technique. The tenfold cross-validation technique uses 90% of the data as training data to train the models and the remaining instances as testing data, and this procedure is repeated ten times [50]. Finally, the prediction accuracy of each classifier was measured using precision, recall, and F1 values as follows [55]:

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{1}$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{2}$$

$$\text{F1} = 2 \times \text{Precision} \times \text{Recall}/(\text{Precision} + \text{Recall}) \tag{3}$$

where TP (true positive) is the number of positive samples predicted by the classifier as positive; FP (false positive) is the number of negative samples predicted by the classifier as positive; and FN (false negative) is the number of positive samples predicted by the classifier as negative.

## 3. Results

### 3.1. Basic Statistics

As the first step of data exploration, several basic statistics of the dataset were examined. As shown in Table 1, among all the participants, 58.62% ($n$ = 21,221) reported being exposed to fraud, but among those who were scammed, only 6.05% ($n$ = 1284) became fraud victims (losing money). For age, we observed that younger people were more likely to be exposed to fraud (F(1, 36,201) = 395.051, $p < 0.001$, $\eta^2 = 0.011$). However, once exposed to fraud, the mean age of the victims was older than that of the survivors (F(1, 21,220) = 32.002, $p < 0.001$, $\eta^2 = 0.002$). The financial status of an individual affects the probability of an individual encountering fraud. Compared with individuals not exposed to fraud, individuals exposed to fraud have more assets (F(1, 36,201) = 899.874, $p < 0.001$, $\eta^2 = 0.024$), debts (F(1, 36,201) = 89.349, $p < 0.001$, $\eta^2 = 0.002$), income (F(1, 36,201) = 291.647, $p < 0.001$, $\eta^2 = 0.008$), and consumption (F(1, 36,201) = 666.952, $p < 0.001$, $\eta^2 = 0.018$). Once exposed to fraud, the financial status does not affect whether individuals are victimized. These results describe the linear relationship between several individual factors and fraud. The individual factors of finer granularity and the nonlinear relationship between variables were analyzed through subsequent computational modeling analysis.

**Table 1.** The basic statistics for the dataset.

| Individual Factors | | Fraud Exposure ($n$ = 36,202) | | Fraud Victimhood ($n$ = 21,221) | |
|---|---|---|---|---|---|
| | | Exposed ($n$ = 21,221) | Not Exposed ($n$ = 14,981) | Fraud Victim ($n$ = 1284) | Fraud Survivor ($n$ = 19,937) |
| Age (M ± SD) | | 51.52 ± 14.88 | 54.68 ± 14.85 | 53.80 ± 16.44 | 51.37 ± 14.77 |
| Sex (%) | Male (Female) | 52.20 (47.80) | 53.44 (46.56) | 52.80 (47.20) | 52.15 (47.85) |
| Asset (¥) | M (SD) | $1.15\times10^6$ $(2.03\times10^6)$ | $5.80\times10^5$ $(1.34\times10^6)$ | $1.04\times10^6$ $(1.83\times10^6)$ | $1.16\times10^6$ $(2.05\times10^6)$ |
| Debt (¥) | M (SD) | $9.17\times10^4$ $(2.16\times10^5)$ | $5.70\times10^4$ $(1.47\times10^5)$ | $8.18\times10^4$ $(2.15\times10^5)$ | $9.23\times10^4$ $(2.16\times10^5)$ |
| Income (¥/year) | M (SD) | $5.67\times10^4$ $(2.49\times10^5)$ | $3.42\times10^4$ $(1.81\times10^5)$ | $6.39\times10^4$ $(2.17\times10^5)$ | $5.62\times10^4$ $(2.50\times10^5)$ |
| Consumption (¥/year) | M (SD) | $6.57\times10^4$ $(7.51\times10^4)$ | $4.62\times10^4$ $(6.36\times10^4)$ | $6.54\times10^4$ $(7.25\times10^4)$ | $6.57\times10^4$ $(7.53\times10^4)$ |

$n$ indicates the number of participants, M indicates the mean values, and SD indicates standard deviation.

### 3.2. Model Prediction Results

This study used fraud exposure and fraud victimhood as ground truth, used various individual factors as inputs, and applied the RFC algorithm to construct predictive models. In addition, only 6.05% of the participants were victimized after being exposed to fraud. Therefore, for the FVR model, random undersampling was used to balance the number of positive and negative samples [28]. For both the FER and FVR models, a grid parameter search was applied to select the best parameters, and the best parameters for each model are presented in Table 2.

**Table 2.** The best parameters for RFC models.

| Parameters | Models | |
|---|---|---|
| | FER | FVR |
| n_estimators | 127 | 47 |
| max_depth | 20 | 10 |
| min_samples_leaf | 5 | 10 |
| min_samples_split | 35 | 45 |
| max_features | 0.3 | 0.9 |

FER indicates fraud exposure recognition and FVR indicates fraud victimhood recognition.

The performance of the FER model is shown in Figure 1. We observed that the FER model achieved a mean precision value of 0.733, a mean recall value of 0.721, and a mean f1 value of 0.727 (see Figure 1a). The receiver operating characteristic (ROC) curve of our FER model is shown in Figure 1b, and we observed that the mean area under the curve (AUC) was 0.675. The FVR model reached a mean precision value of 0.580, a mean recall value of 0.553, and a mean f1 value of 0.565 (see Figure 2a). The ROC curve of our FVR model is shown in Figure 2b, and the mean AUC was 0.577. In general, the performance of the FER model was better than that of the FVR model (f1: t(9) = 10.405, $p < 0.001$, $d = 3.290$).

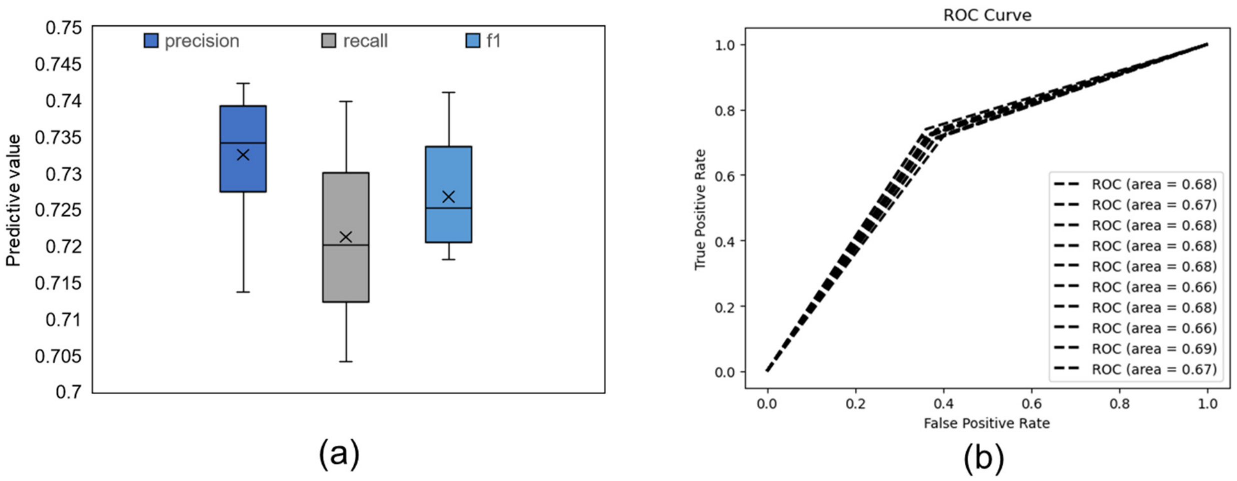

**Figure 1.** Predictive results of the FER model. (**a**) Each box shows the distribution of the predictive results of ten test sets. The black line in the middle of each box indicates the median; "×" indicates the mean value. (**b**) ROC curve of ten test sets.

### *3.3. Model Interpretability*

We then explained the constructed models by examining the information gain of features (calculating feature importance) [51]. The feature importance of the FER and FVR models is presented in Figures 3 and 4, respectively. Since tenfold cross-validation technology was used to evaluate our models, the feature importance was different when predicting different test sets. Therefore, the distribution of feature importance was arranged in descending order of the mean value, and only the top 20 features were included for visibility. We observed that for the FER model, the most important feature is a registered residence, accounting for 6.98% of the model. Total assets are the second most important feature (4.93%), followed by funds (4.07%), total consumption (3.66%), rural areas (3.50%), total income (3.05%), stock accounts (2.75%), and so on. For the FVR model, the total income is the most important feature (7.34%), followed by age (6.33%), clothing expenses (4.82%), total assets (4.46%), cash (4.39%), and total consumption (3.67%). The above results compare the potential impact of different features on fraud exposure and fraud victimization, and the complete importance scores are presented in supplementary material Tables S2 and S3.

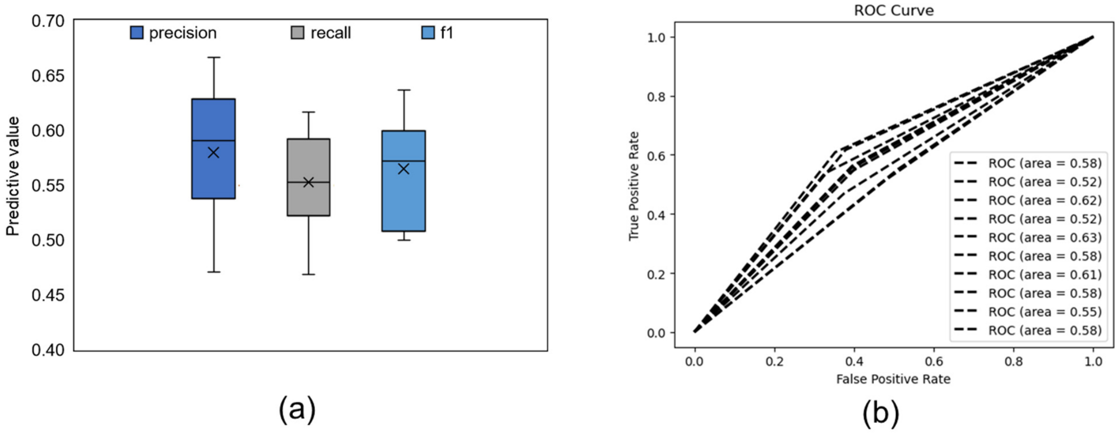

**Figure 2.** Predictive results of the FVR model. (**a**) Each box shows the distribution of the predictive results of ten test sets. The black line in the middle of each box indicates the median; "×" indicates the mean value. (**b**) ROC curve of ten test sets.

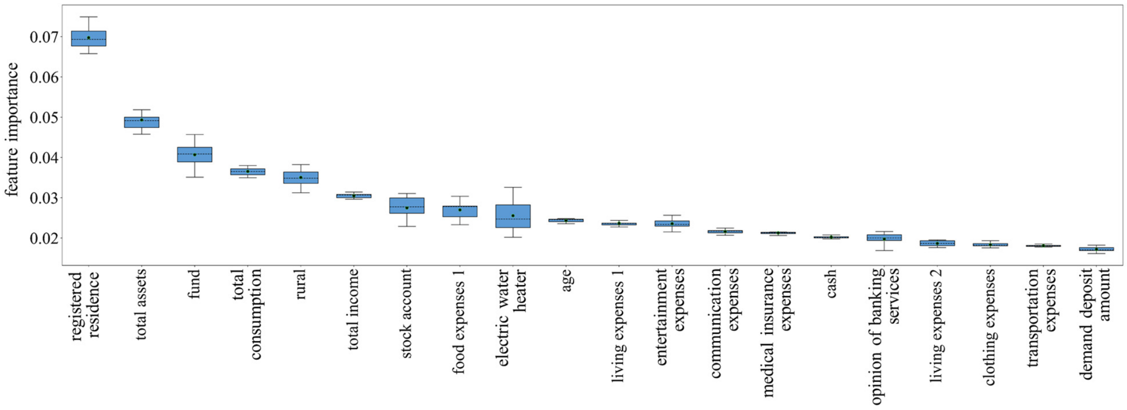

**Figure 3.** Distribution of feature importance for FER model. Arranged in descending order of the mean value, the top 20 features were included for visibility, and the trend of the remaining features was approximately the same. Error bars indicate the standard deviations, the black dots indicate the mean values, and the black dotted lines indicate the median values.

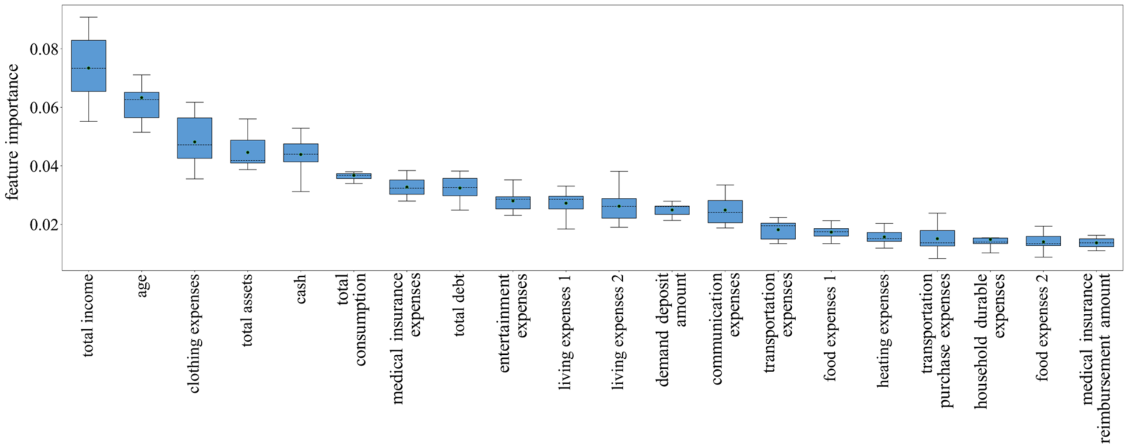

**Figure 4.** Distribution of feature importance for FVR model. Arranged in descending order of the mean value, the top 20 features were included for visibility, and the trend of the remaining features was approximately the same. Error bars indicate the standard deviations, the black dots indicate the mean values, and the black dotted lines indicate the median values.

## 4. Discussion

The present work used a computational modeling method to examine the impacts of various individual factors on consumer financial fraud. Based on the two-stage conceptual framework of fraud [4], we used a nationwide dataset (CHFS) [45] to construct the fraud exposure recognition and fraud victimhood recognition models. The importance of the model features indicates which individuals are more likely to be exposed to fraud and which individuals are more likely to be victimized by fraud. The model we build also has practical value because early identification of vulnerable individuals may be able to reduce the harm caused by fraud by intervening early. Our models may also be useful in practices in finance area, such as in the mobile banking applications.

The basic statistics of the CHFS dataset were first conducted to explore the linear relationship among several individual factors, fraud exposure, and fraud victimhood. The results show that people exposed to fraud were younger than those not exposed, whereas the victims were older than the survivors once exposed to fraud. This finding supports the previous opinion that older adults are less exposed to fraud but more likely to become victims of fraud [16,22]. We also found that individuals with more assets, debts, incomes, and consumption were more likely to be exposed to fraud. This result reflects that more economic activity or possession of more assets increases the likelihood of being the target of consumer financial fraud because defrauding money is the main purpose of customer financial fraud [3]. However, similar to previous findings [31], financial factors do not have any significant relationship with whether the individual will become a victim. More factors were then explored in subsequent modeling.

The FER model we constructed has a predictive effect, reaching a mean f1 value of 0.727. Identifying individuals who are more vulnerable to fraud may be able to directly prevent individuals from being exposed to fraud. Therefore, we believe that using the current prediction model may indirectly mitigate the threat posed by consumer financial fraud. The FER model was then interpreted by calculating feature importance. The results show that the migration status (i.e., which type of registered residence) is the most important factor in fraud exposure. Although the socioeconomic and health disadvantages of Chinese migrants have been reviewed in previous works [56,57], the fact that being migrants put consumers at a substantially higher risk of being targeted by perpetrators was only recently discovered [4]. Our results further lay the foundation for the decisive impact of immigration on fraud exposure. Similar to the basic statistical results, many financial-related features (e.g., assets, total consumption, funds, income, food expenses, and living expenses) play an important role in predicting fraud exposure. This result once again supports the opinion that defrauding money is the main purpose of customer financial fraud [3]. We also observed that, in the Chinese context, urbanicity (whether living in rural areas) and age have important influences on whether fraud will be encountered. Further statistical analysis shows that living in rural areas and being older decreased the chance of being a fraud target, similar to previous survey results [19,20]. Finally, although other factors, such as the opinion of banking services (accounting for 1.98% of the model), attitudes toward online financial products (1.26%), and risk appetite (0.79%), also have predictive effects, their effects are much smaller than the features mentioned above.

The FVR model shows a low predictive effect on fraud victimhood, only reaching a mean f1 value of 0.565. Model interpretation shows that a large number of financial-related features, such as income, clothing expenses, consumption, and debt, played important roles in the FVR model, although statistical tests were not significant. Both the model recognition results and the feature importance results remind us that the individual factors considered in this study are not sufficient to effectively predict fraud victimization. Predicting whether someone will be victimized may require more consideration of the influence of individual cognition and psychological factors [28,58]. For FER, its essence is to investigate the criteria for fraudsters to choose fraud targets. Individuals' demographic, economic and geographic factors considered in this study are also decisive factors for fraudsters. However, once exposed to fraud (i.e., for FVR), an individual's psychological factors may determine whether he or she will be a victim of fraud. Previous works have shown that victims often fail to recognize deception cues due to psychological factors, such as personality, low self-control, and impulsivity [29,32,34,58]. Vishwanath et al. found that individuals holding the psychological characteristics associated with victimhood are more likely to use heuristics to make quick (usually erroneous) decisions [35]. Unfortunately, few psychological factors were investigated in the CHFS survey, and the considered factors, including risk appetite and well-being, have little effect on fraud victimhood. Therefore, measuring and integrating various psychological factors to construct FVR models is an important direction for future research.

In addition, the different types of scams are not distinguished in this study. In fact, the victims of different scams may have different characteristics. For example, according to data from Tencent 110, men are more likely to be scammed in pornography and dating scams, while women are less likely to be scammed in these two types of scams [59]. At the same time, the incidence of different types of scams may vary in different countries. In China, transaction scams and loan scams are more common, followed by identity impersonation scams, financial management scams, and online dating scams [60]. For future research in the Chinese context, separate modeling for more high-incidence scams can be considered to improve the accuracy of recognition models.

From a practical perspective, our recognition models can be applied to combat fraud. The FER model can be used to identify who are more likely to be the targets of fraud and the FVR model can be used to predict who are more likely to be defrauded and suffer losses (although our results showed that this model need more psychological factors). For government departments, after identifying susceptible groups, they can carry out more anti-fraud training and publicity for these people to improve their anti-fraud consciousness and avoid victimization. For banks, they can identify vulnerable users by using our recognition models, so that certain security measures can be taken for these users (such as blocking possible fraudulent transactions). In addition, banks can also consider the recognition results of our models in the fraudulent financial transaction detection model. Referring to our recognition models, the accuracy of fraud detection may be improved. Notably, similar to fraud detection, our recognition models also rely on consumers' demographic data and financial data, so privacy intrusiveness should be considered when collecting information [61].

## 5. Conclusions

In conclusion, the present study integrated a large number of individual factors to predict customer fraud victimization. Based on the two-stage conceptual framework, we constructed FER and FVR models, respectively. The FER model performed well, and model interpretation indicated that migration status, financial status, urbanity, and age have good predictive effects on fraud exposure in the Chinese context, whereas the FVR model shows a low predictive effect, reminding us to consider more psychological factors in future work.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/ijerph19010461/s1, Table S1: Feature list, Table S2: Feature importance of fraud exposure recognition (FER) model, Table S3: Feature importance of fraud victimhood recognition (FVR) model.

**Author Contributions:** Conceptualization, L.X. (Liang Xu); methodology, L.X. (Liang Xu); validation, L.X. (Liang Xu); formal analysis, L.X. (Liuchang Xu); data curation, D.X.; writing—original draft preparation, L.X. (Liuchang Xu); writing—review and editing, J.W. and L.X. (Liang Xu); visualization, L.X. (Liuchang Xu); supervision, L.X. (Liang Xu). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Zhejiang Provincial Natural Science Foundation of China, grant number LGG22D010001, the National Natural Science Foundation of China, grant number 72001190, Scientific Research Fund of Zhejiang Provincial Education Department, grant number Y202147221, and Ministry of Education of Humanities and Social Science Project, grant number 20YJC630173.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw data are available at https://chfser.swufe.edu.cn/datas/Products/Datas/DataList (accessed date 10 December 2020).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Williams, E.J.; Beardmore, A.; Joinson, A.N. Individual Differences in Susceptibility to Online Influence: A Theoretical Review. *Comput. Hum. Behav.* **2017**, *72*, 412–421. [CrossRef]
2. Deevy, M.; Lucich, S.; Beals, M. Scams, Schemes & Swindles. Financial Fraud Research Center, Stanford University. Available online: http://longevity.stanford.edu/wp-content/uploads/2017/01/Scams-Schemes-Swindles-FINAL-On-Website.pdf (accessed on 10 December 2021).
3. Irvin-Erickson, Y.; Ricks, A. Identity Theft and Fraud Victimization: What We Know about Identity Theft and Fraud Victims from Research-and Practice-Based Evidence. Available online: https://ncvc.dspacedirect.org/handle/20.500.11990/1544 (accessed on 10 December 2021).
4. Fan, J.X.; Yu, Z. Understanding Aging and Consumer Fraud Victimization in the Chinese Context: A Two-Stage Conceptual Approach. *J. Elder Abus. Negl.* **2021**, *33*, 230–247. [CrossRef] [PubMed]
5. Anderson, K.B. Consumer fraud in the United States: An FTC Survey. Federal Trade Commission. Available online: https://www.ftc.gov/sites/default/files/documents/reports/consumer-fraud-united-states-ftc-survey/040805confraudrpt.pdf (accessed on 10 December 2021).
6. Saunders, L.; Pizor, A.; Twomey, T. Desperate Homeowners: Loan mod Scammers Step in When Loan Services Refuse to Provide Relief. National Consumer Law Center. Available online: https://www.nclc.org/images/pdf/pr-reports/report-loan-mod-scams-2009.pdf (accessed on 10 December 2021).
7. Lee, C.S. How Online Fraud Victims are Targeted in China: A Crime Script Analysis of Baidu Tieba C2C Fraud. *Crime Delinq.* 2021; advance online publication. [CrossRef]
8. Ganzini, L.; McFarland, B.H.; Cutler, D. Prevalence of Mental Disorders after Catastrophic Financial Loss. *J. Nerv. Ment. Dis.* **1990**, *178*, 680–685. [CrossRef]
9. Spalek, B. Exploring the Impact of Financial Crime: A Study Looking into the Effects of the Maxwell Scandal upon the Maxwell Pensioners. *Int. Rev. Vict.* **1999**, *6*, 213–230. [CrossRef]
10. Deem, D.L. Notes from the Field: Observations in Working with the Forgotten Victims of Personal Financial Crimes. *J. Elder Abus. Negl.* **2000**, *12*, 33–48. [CrossRef]
11. Identity Theft Resource Center. The Aftermath: The Non-Economic Impacts of Identitytheft. Available online: https://www.idtheftcenter.org/wp-content/uploads/2018/09/ITRC_Aftermath-2018_Web_FINAL.pdf (accessed on 10 December 2021).
12. Federal Trade Commission. FTC Identity Theft: Planning for the Future Conference. Available online: https://www.ftc.gov/system/files/documents/videos/identity-theft-planning-future-part-1/ftc_identity_theft_planning_for_the_future_transcript_segment_1.pdf (accessed on 25 May 2017).
13. Button, M.; Lewis, C.; Tapley, J. Not a victimless crime: The impact of fraud on individual victims and their families. *Secur. J.* **2014**, *27*, 36–54. [CrossRef]
14. Cross, C. 'They're very lonely': Understanding the fraud victimisation of seniors. *Int. J. Crime Justice Soc. Democr.* **2016**, *5*, 60. [CrossRef]
15. Kadoya, Y.; Khan, M.S.R.; Narumoto, J.; Watanabe, S. Who is next? A study on victims of financial fraud in Japan. *Front. Psychol.* **2021**, *12*, 649565.
16. Shadel, D.; Pak, K.B.S. The Psychology of Consumer Fraud. Ph.D. Thesis, Universiteit van Tilburg, Tilburg, The Netherland, 2007.
17. Grazioli, S. Where Did They Go Wrong? An Analysis of the Failure of Knowledgeable Internet Consumers to Detect Deception over the Internet. *Group Decis. Negot.* **2004**, *13*, 149–172. [CrossRef]
18. Wright, R.; Chakraborty, S.; Basoglu, A.; Marett, K. Where Did They Go Right? Understanding the Deception in Phishing Communications. *Group Decis. Negot.* **2010**, *19*, 391–416. [CrossRef]
19. Kerley, K.R.; Copes, H. Personal Fraud Victims and Their Official Responses to Victimization. *J. Police Crim. Psych.* **2002**, *17*, 19–35. [CrossRef]
20. Schoepfer, A.; Piquero, N.L. Studying the Correlates of Fraud Victimization and Reporting. *J. Crim. Justice* **2009**, *37*, 209–215. [CrossRef]
21. Pratt, T.C.; Holtfreter, K.; Reisig, M.D. Routine online activity and internet fraud targeting: Extending the generality of routine activity theory. *J. Res. Crime Delinq.* **2010**, *47*, 267–296. [CrossRef]
22. Shi, J.; Wu, C.; Qian, X. The Effects of Multiple Factors on Elderly Pedestrians' Speed Perception and Stopping Distance Estimation of Approaching Vehicles. *Sustainability* **2020**, *12*, 5308. [CrossRef]
23. Shao, J.; Zhang, Q.; Ren, Y.; Li, X.; Lin, T. Why Are Older Adults Victims of Fraud? Current Knowledge and Prospects Regarding Older Adults' Vulnerability to Fraud. *J. Elder Abus. Negl.* **2019**, *31*, 225–243. [CrossRef]

24. Titus, R.M.; Gover, A.R. Personal fraud: The victims and the scams. *Crime Prev. Stud.* **2001**, *12*, 133–152.
25. Anderson, B.B.; Vance, A.; Kirwan, C.B.; Jenkins, J.L.; Eargle, D. From Warning to Wallpaper: Why the Brain Habituates to Security Warnings and What Can Be Done About It. *J. Manag. Inf. Syst.* **2016**, *33*, 713–743. [CrossRef]
26. AARP. Off the Hook: Reducing Participation in Telemarketing Fraud. Available online: http://www.aarp.org/research/frauds-scams/telemarketing/aresearch-import-179-D17812.html (accessed on 10 December 2021).
27. Pak, K.; Shadel, D. AARP Foundation national fraud victim study. Available online: https://assets.aarp.org/rgcenter/general/fraud-victims-11.pdf (accessed on 10 December 2021).
28. Norris, G.; Brookes, A.; Dowell, D. The Psychology of Internet Fraud Victimisation: A Systematic Review. *J. Police Crim. Psych.* **2019**, *34*, 231–245. [CrossRef]
29. Modic, D.; Anderson, R.; Palomäki, J. We Will Make You like Our Research: The Development of a Susceptibility-to-Persuasion Scale. *PLoS ONE* **2018**, *13*, e0194119. [CrossRef]
30. Modic, D.; Lea, S.E.G. How Neurotic Are Scam Victims, Really? The Big Five and Internet Scams. *SSRN J.* **2012**. [CrossRef]
31. Purkait, S.; Kumar De, S.; Suar, D. An Empirical Investigation of the Factors That Influence Internet User's Ability to Correctly Identify a Phishing Website. *Inf. Manag. Comput. Secur.* **2014**, *22*, 194–234. [CrossRef]
32. Holtfreter, K.; Reisig, M.D.; Pratt, T.C. Low Self-Control, Routine Activities, and Fraud Victimization. *Criminology* **2008**, *46*, 189–220. [CrossRef]
33. Whitty, M.T.; Buchanan, T. The Online Dating Romance Scam: The Psychological Impact on Victims—Both Financial and Non-Financial. *Criminol. Crim. Justice* **2016**, *16*, 176–194. [CrossRef]
34. Pattinson, M.R.; Jerram, C.; Parsons, K.; McCormac, A.; Butavicius, M.A. Managing Phishing Emails: A Scenario-Based Experiment. In *Human Aspects of Information Security & Assurance, Proceedings of the Fifth International Symposium on Human Aspects of Information Security & Assurance (HAISA), London, UK, 7–8 July 2011*; Furnell, S.M., Clarke, N.L., Eds.; University of Plymouth: London, UK, 2011; pp. 74–85.
35. Vishwanath, A.; Herath, T.; Chen, R.; Wang, J.; Rao, H.R. Why Do People Get Phished? Testing Individual Differences in Phishing Vulnerability within an Integrated, Information Processing Model. *Decis. Support Syst.* **2011**, *51*, 576–586. [CrossRef]
36. Vasilopoulos, N.L.; Cucina, J.M.; McElreath, J.M. Do Warnings of Response Verification Moderate the Relationship Between Personality and Cognitive Ability? *J. Appl. Psychol.* **2005**, *90*, 306–322. [CrossRef]
37. Gottfredson, M.R.; Hirschi, T. *A General Theory of Crime*; Stanford University Press: Palo Alto, CA, USA, 1990.
38. van de Weijer, S.G.A.; Leukfeldt, E.R. Big Five Personality Traits of Cybercrime Victims. *Cyberpsychology Behav. Soc. Netw.* **2017**, *20*, 407–412. [CrossRef] [PubMed]
39. Kim, D.; Hyun Kim, J. Understanding Persuasive Elements in Phishing E-Mails: A Categorical Content and Semantic Network Analysis. *Online Inf. Rev.* **2013**, *37*, 835–850. [CrossRef]
40. Petty, R.E.; Briñol, P. Emotion and Persuasion: Cognitive and Meta-Cognitive Processes Impact Attitudes. *Cogn. Emot.* **2015**, *29*, 1–26. [CrossRef]
41. Fischer, P.; Lea, S.E.; Evans, K.M. Why do individuals respond to fraudulent scam communications and lose money? The psychological determinants of scam compliance. *J. Appl. Soc. Psychol.* **2013**, *43*, 2060–2072. [CrossRef]
42. Button, M.; Nicholls, C.M.; Kerr, J.; Owen, R. Online frauds: Learning from victims why they fall for these scams. *Aust. N. Z. J. Criminol.* **2014**, *47*, 391–408. [CrossRef]
43. Norris, G.; Brookes, A. Personality, Emotion and Individual Differences in Response to Online Fraud. *Personal. Individ. Differ.* **2021**, *169*, 109847. [CrossRef]
44. Vempala, N.N.; Russo, F.A. Modeling Music Emotion Judgments Using Machine Learning Methods. *Front. Psychol.* **2018**, *8*, 2239. [CrossRef]
45. Gan, L.; Yin, Z.; Jia, N.; Xu, S.; Ma, S.; Zheng, L. *Data You Need to Know about China*; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]
46. Brink, A.G.; Gouldman, A.; Victoravich, L.M. The Effects of Organizational Risk Appetite and Social Pressure on Aggressive Financial Reporting Behavior. *Behav. Res. Account.* **2018**, *30*, 23–36. [CrossRef]
47. Wood, S.; Liu, P.-J.; Hanoch, Y.; Xi, P.M.; Klapatch, L. Call to Claim Your Prize: Perceived Benefits and Risk Drive Intention to Comply in a Mass Marketing Scam. *J. Exp. Psychol. Appl.* **2018**, *24*, 196–206. [CrossRef] [PubMed]
48. Alam, M.S.; Vuong, S.T. Random Forest Classification for Detecting Android Malware. In Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, Beijing, China, 20–23 August 2013; pp. 663–669.
49. Arora, N.; Kaur, P.D. A Bolasso Based Consistent Feature Selection Enabled Random Forest Classification Algorithm: An Application to Credit Risk Assessment. *Appl. Soft Comput.* **2020**, *86*, 105936. [CrossRef]
50. Xu, L.; Zheng, Y.; Xu, D.; Xu, L. Predicting the Preference for Sad Music: The Role of Gender, Personality, and Audio Features. *IEEE Access* **2021**, *9*, 92952–92963. [CrossRef]
51. Quiroz, J.C.; Geangu, E.; Yong, M.H. Emotion Recognition Using Smart Watch Sensor Data: Mixed-Design Study. *JMIR Ment. Health* **2018**, *5*, e10153. [CrossRef]
52. Xu, L.; Wen, X.; Shi, J.; Li, S.; Xiao, Y.; Wan, Q.; Qian, X. Effects of Individual Factors on Perceived Emotion and Felt Emotion of Music: Based on Machine Learning Methods. *Psychol. Music* **2021**, *49*, 1069–1087. [CrossRef]

53. Götz, F.M.; Stieger, S.; Gosling, S.D.; Potter, J.; Rentfrow, P.J. Physical Topography Is Associated with Human Personality. *Nat. Hum. Behav.* **2020**, *4*, 1135–1144. [CrossRef] [PubMed]
54. Xu, L.; Sun, Z.; Wen, X.; Huang, Z.; Chao, C.; Xu, L. Using Machine Learning Analysis to Interpret the Relationship between Music Emotion and Lyric Features. *PeerJ Comput. Sci.* **2021**, *7*, e785. [CrossRef]
55. Sun, Z.; Ji, Z.; Zhang, P.; Chen, C.; Qian, X.; Du, X.; Wan, Q. Automatic Labeling of Mobile Apps by the Type of Psychological Needs They Satisfy. *Telemat. Inform.* **2017**, *34*, 767–778. [CrossRef]
56. Li, J.; Rose, N. Urban Social Exclusion and Mental Health of China's Rural-Urban Migrants—A Review and Call for Research. *Health Place* **2017**, *48*, 20–30. [CrossRef] [PubMed]
57. Zhang, L.; Sharpe, R.V.; Li, S.; Darity, W.A. Wage Differentials between Urban and Rural-Urban Migrant Workers in China. *China Econ. Rev.* **2016**, *41*, 222–233. [CrossRef]
58. Harrison, B.; Vishwanath, A.; Ng, Y.J.; Rao, R. Examining the Impact of Presence on Individual Phishing Victimization. In Proceedings of the 2015 48th Hawaii International Conference on System Sciences, Kauai, HI, USA, 5–8 January 2015; pp. 3483–3489.
59. Tencent 110. Online Fraud Governance Report in 2020. Available online: https://download.mouse0232.cn/pdf/0226/%E3%80%90%E8%85%BE%E8%AE%AF110%E3%80%912020%E5%B9%B4%E7%BD%91%E7%BB%9C%E8%AF%88%E9%AA%97%E6%B2%BB%E7%90%86%E6%8A%A5%E5%91%8A.pdf (accessed on 10 December 2021).
60. China Academy of Information and Communications Technology. Research Report on Telecommunication Network Fraud Management under the New Situation. Available online: http://www.caict.ac.cn/kxyj/qwfb/ztbg/202012/P020201218393889946629-5.pdf (accessed on 10 December 2021).
61. Găbudeanu, L.; Brici, I.; Mare, C.; Mihai, I.C.; Șcheau, M.C. Privacy Intrusiveness in Financial-Banking Fraud Detection. *Risks* **2021**, *9*, 104. [CrossRef]