Computer Science and Application 计算机科学与应用, 2021, 11(12), 3019-3027
Published Online December 2021 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2021.1112305

Hans汉斯

一种基于特征敏感的信用卡欺诈检测模型

黄家元

同济大学，上海

收稿日期：2021年11月25日；录用日期：2021年12月21日；发布日期：2021年12月28日

摘　要

金融欺诈交易检测一直是金融领域关注的重点问题。然而，现有的欺诈检测方法，在模型训练过程中忽略了对数据的特征敏感性做处理，导致对减少实际经济损失没有明显作用；因此提出了一种基于特征敏感的stacking集成学习方法FSBS (feature-sensitive based stacking)，首先选择几种不同质的基分类器当作第一层基准模型，然后通过交叉验证的方式得到概率形式的输出，最后通过一层特殊的stacking集成方法使模型对大金额交易样本有所偏置。最终实验证明，FSBS模型可以有效减少欺诈交易带来的经济损失。

关键词

欺诈检测，特征敏感，集成学习，结合策略，对比实验

A Credit Card Fraud Detection Model Based on Feature Sensitivity

Jiayuan Huang

Tongji University, Shanghai

Received: Nov. 25th, 2021; accepted: Dec. 21st, 2021; published: Dec. 28th, 2021

Abstract

Financial fraud transaction detection has always been a key issue in the financial field. However, the existing fraud detection methods ignore the feature sensitivity of the data in the model training process, which has no obvious effect on reducing the actual economic loss; therefore, a feature-sensitive stacking integrated learning method FSBS (feature-sensitive based stacking), firstly

文章引用: 黄家元. 一种基于特征敏感的信用卡欺诈检测模型[J]. 计算机科学与应用, 2021, 11(12): 3019-3027.
DOI: 10.12677/csa.2021.1112305

selects several different qualitative base classifiers as the first-level benchmark model, and then obtains the probabilistic output through cross-validation, and finally uses a special layer of stacking integration method to make the model for large amounts. The trading sample is biased. The final experiment proved that the FSBS model can effectively reduce the economic loss caused by fraudulent transactions.

Keywords

Fraud Detection, Feature Sensitivity, Ensemble Learning, Combination Strategy, Contrast Experiment

1. 引言

随着互联网普及率的提高，电子交易比重日渐提升，然而在技术进步带来电商交易繁荣的背后，欺诈交易也以惊人的速度在增长[1]。近年来数据挖掘和机器学习的逐步发展，许多智能机器学习模型被应用在欺诈检测领域。在现有的研究中，一些基于神经网络的算法具有识别率高、鲁棒性好、施展性强的特点得到了广泛应用[2] [3] [4] [5] [6]。以决策树模型为基础的随机森林模型也以其性能好、泛化性强、实现简单、可以应对数据不均衡等优点被给予极大关注[7] [8] [9] [10] [11]。除此之外，许多利用技术组合或安全策略等方法进行欺诈检测的研究也给予启发[12] [13] [14]。

在算法落地过程中，对交易金额越大的交易样本的误判往往会造成更大的经济损失，也就是说，模型对“交易金额”这一特征更加敏感。例如，一笔流水为 1 万元的交易与流水为 100 元的交易，理应受到不同程度的重视。那些大金额的交易数据才应该是模型关注的重点。然而在现有的研究中，往往忽略了交易数据的特征敏感问题，导致模型可能取得了不错的性能检测指标，但是错判的都是交易金额相对较大的样本，使得经济损失反而上升，这给欺诈检测算法的落地带来了重大隐患。

本文针对上述问题提出了一种基于特征敏感的 stacking 学习方法 FSBS (feature-sensitive based stacking)。具体来说，首先选择五种不同质的基础分类模型，分别对训练样本进行预测，同时采用 5 折交叉验证的形式来避免过拟合；之后将上一步的五个结果，拼接上各个样本的真实 label 带入新的基于特征敏感的逻辑斯蒂回归进行训练，最终再预测的结果就是 FSBS 的最终结果。本篇文章的主要工作可以总结为：1) 文章提出了一种基于特征敏感学习的 stacking 集成方法 FSBS；2) 在数据集上进行对比实验，实验结果表明 FSBS 相比其他方法，其错判样本的平均交易金额更小，在降低欺诈经济损失这一目的上效果更加显著。

后续文章的组织结构如下：第 2 节详细介绍了 FSBS 算法模型。第 3 节介绍对比实验及结果分析。第 4 节为结论部分。

2. 算法模型

该模型运用了集成学习模型融合 stacking 结合策略，它利用原始数据集训练第一层的五个初级学习器(也称为基学习器)，以获得五个不同的预测结果；然后第一层得到的预测结果被最后一层基于特征敏感的次级学习器作为输入特征进行训练，得到最终预测结果，从而提高模型的准确度，减少泛化误差，并

使得模型对交易金额有一定敏感度。FSBS 的结构如图 1 所示。

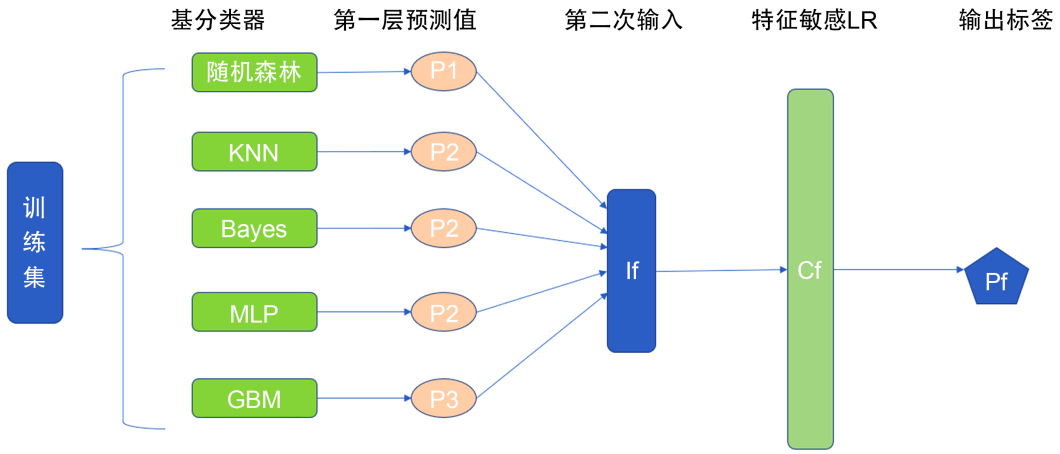

Figure 1. FSBS model structure
图 1. FSBS 模型结构

2.1. 第一层模型

对于输入进模型的原始数据来说，可能是杂乱无规律的。在 stacking 中，通过第一层学习器后，有效的特征被学习出来了。从这个角度看，stacking 第一层就是特征抽取的过程。在[15]的研究中，图 2 上排是未经 stacking 的数据，下排是经过 stacking (多个无监督学习算法)处理后的，我们可以显著的发现红色和蓝色的数据在下排的分界更为明显。*数据经过了压缩处理，这说明有效的 stacking 可以对原始数据中的特征有效的抽取。

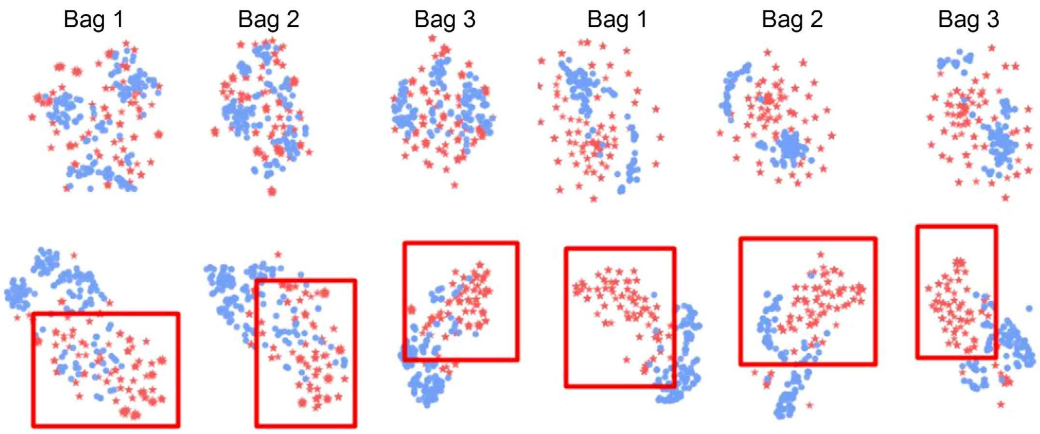

Figure 2. Data distribution before and after stacking
图 2. Stacking 前后数据分布

Stacking 中的第一层可以等价于神经网络中的前 $n-1$ 层，而 stacking 中的最终分类层可以类比于神经网络中最后的输出层。不同点在于，stacking 中不同的分类器通过异质来体现对于不同特征的表示，神经网络是从同质到异质的过程且有分布式表示的特点(distributed representation)。Stacking 中应该也有分布式的特点，主要表现在多个分类器的结果并非完全不同，而有很大程度的相同之处。多个分类器应该尽量在保证效果好的同时尽量不同，stacking 集成学习框架的对于基分类器的两个要求：① 差异化(diversity)

要大；② 准确性(accuracy)要高。

因此从理论上讲，基学习器应该尽量做到“好而不同”，stacking 中各个初级学习器的学习能力越强、关联程度越低，模型预测效果就越好。由于随机森林、支持向量机和梯度提升树算法在之前的研究中虽然在某些条件下有一定的缺陷，但是仍是具有较好的性能，比如随机森林和支持向量机能较好地解决小样本情况下分类器过拟合数据的问题；人工神经网络具有强大的学习能力，善于挖掘控制因素和结果之间的非线性关系等。根据理论分析和实验结果，最终的模型结构的第一层基学习器包含以下五种模型：

① 随机森林(random forest)

② K 近邻(KNN)

③ 贝叶斯分类器(Bayes)

④ 多层感知机(MLP)

⑤ 梯度提升算法(Gradient Boosting Machine)

2.2. 第二层模型

在 stacking 分类模型中，我们在最后一层分类器使用了本文创新性提出的基于特征敏感的损失函数，它是一种基于逻辑回归(Logistic Regression)交叉熵损失函数的变种，逻辑回归是一种非常经典的分类模型，被广泛的应用于二分类的机器学习问题上。本文提出的最终损失函数 $\mathcal{L}_{main}$ 其具体形式定义如下：

$$\mathcal{L}_{main} = \mathcal{L}_{cls} + \alpha \cdot \mathcal{L}_{extra} + \gamma \|\theta\|^2$$

其中，$\mathcal{L}_{main}$ 为机器学习模型总的函数表达，$\mathcal{L}_{cls}$ 为在分类问题中的损失函数，常见的比如负对数似然损失、交叉熵损失、或者指数损失，$\mathcal{L}_{extra}$ 为在标准分类问题的损失函数上、按照实际情况添加的某种约束或偏重，使得最终的代价朝向一个特定的方向偏置，而这个偏置就是具体业务场景更加关注的部分；在本文 FSBS 模型中，$\mathcal{L}_{cls}$ 为标准的交叉熵损失函数，$\mathcal{L}_{extra}$ 为根据交易金额对样本额外添加的权重项。α 为超参数。为惩罚项，γ 为正则化参数，用于对参数进行约束；θ 表示所有可学习的模型参数，用于 L1/L2 惩罚来防止过拟合。$\mathcal{L}_{cls}$ 和 $\mathcal{L}_{extra}$ 表达式如下所示：

$$\mathcal{L}_{cls} = -\frac{1}{m}\sum_{i=1}^{m}\left[y_i * \log h_\theta(x_i) + (1-y_i) * \log\left(1-h_\theta(x_i)\right)\right]$$

$$\mathcal{L}_{extra} = \log\left(\frac{x_i^t}{K}+1\right) * \left(-\frac{1}{m}\sum_{i=1}^{m}\left[y_i * \log h_\theta(x_i) + (1-y_i) * \log\left(1-h_\theta(x_i)\right)\right]\right)$$

其中 m 为样本数量，x_i 为训练数据集的特征，y_i 为训练数据集的标签，α，K 为超参数，x_i^t 为训练数据集第 i 条交易数据的交易金额 trade_amount；α 越高，则损失函数中金额较大的交易数据权重就越大；K 越小，就越大，模型就对金额大的数据更加敏感，最终结果就越趋向于将这些大金额样本预测正确。$h_\theta(x)$ 为 sigmoid 函数：

$$h_\theta(x) = g\left(\theta^{\mathrm{T}} * x\right) = \frac{1}{1+\mathrm{e}^{-\theta^{\mathrm{T}} * x}}$$

最终，FSBS 的算法流程如下：

算法 3.1：FSBS 算法流程

输入：数据集 $D = \{d_1, d_2, \cdots, d_n\}$，$d_i = (x_i, y_i)$，$x_i$ 是一个 d 维的特征向量，y_i 是一个一维的标签；

过程：

第一步：训练初级分类器

1：**for** $t=1,2,3,\cdots,T$ **do**

2: $h_t=\mathcal{L}_t(D)$

3: **end for**

第二步：使用训练出来的初级分类器来得到预测结果，并作为数据集训练基于特征敏感的逻辑斯蒂分类器 $\mathcal{L}$

4: $D'=\varnothing$

5: **for** $i=1,2,3,\cdots,m$ **do**

6: **for** $t=1,2,3,\cdots,T$ **do**

7: $z_{it}=h_t(x_i)$;

8: **end for**

9: $D'=D'\cup\left(\left(z_{i1},z_{i2},z_{i3},\cdots,z_{iT}\right),y_i\right)$;

10: **end for**

11: 以梯度 $\theta_j:=\theta_j+\mu\cdot\alpha\cdot\log\left(\frac{x_i^t}{K}+1\right)\left(\mathcal{L}_\theta(x_i)-y_i\right)x_i^j$ 更新 $\mathcal{L}$ 的参数

12:**return** $\mathcal{L}$

输出：一组初级学习算法 $\mathcal{L}_1,\mathcal{L}_2,\mathcal{L}_3,\cdots,\mathcal{L}_T$

基于特征敏感的逻辑斯蒂回归(次级学习算法) $\mathcal{L}$

3. 实验

3.1. 数据集

为了验证上述算法的有效性，本节在已有数据集上进行了对比实验验证，其中数据集的具体信息如下：本实验采用的数据集大小为 5,125,107 条，其中欺诈样本的有 147,829 条，正常样本有 4,977,278 条，数据维度为 10 维。样本的标签分布比例和交易金额分布比例如图 3 和图 4。

Figure 3. Proportion of sample label distribution
图 3. 样本标签分布比例

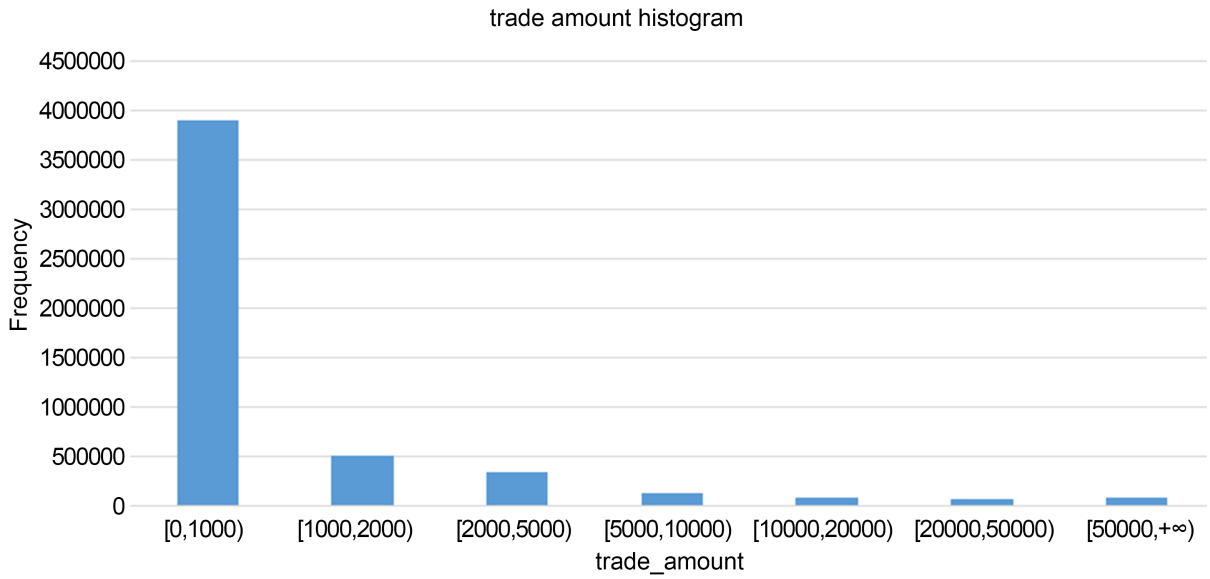

Figure 4. Proportion of sample amount distribution
图 4. 样本金额分布比例

3.2. 数据指标

在本实验中，主要依据基于混淆矩阵的精准率、召回率、f1 值和 loss 值这四个指标来衡量算法性能。对于二分类问题来说，将少数类的欺诈样本看作正样本，多数类看作负样本。混淆矩阵如表 1 所示：

Table 1. Confusion matrix
表 1. 混淆矩阵

	Predicted Positive	Predicted Negative
Actual Positive	TP (True Positive)	FN (False Negative)
Actual Negative	FP (False Positive)	TN (True Negative)

Precision 为精准率，表示实际为正例且预测为正例的样本数量在所有预测为正例的样本中的比例。Precision = TP/(TP + FP)。

Recall 为召回率，表示实际为正例且预测为正例的样本数量在所有正例样本中的比例。Recall = TP/(TP + FN)。

f1 值主要应用于二分类，其计算方式为：$f1 = (1 + \beta^2) * Precision * Recall/\beta^2 * (Precision + Recall)$。在本实验中 β 取 1，表示精准率和召回率的调和平均。

Loss 值表示模型错判样本的平均交易金额(单位：元)。

3.3. 实验过程

在实验过程中，模型选择和训练过程均采用第二节所述的基于特征敏感的 stacking 模型 FSBS。在训练阶段，由于次级的训练集是利用初级学习器产生的，如果直接将初级学习器的训练结果作为次级学习器的训练集，则过拟合风险较大；因此，在每一个基础模型中应用了 5 倍的交叉验证。在训练过程中，从全部数据集中抽样采用 70%的数据作为训练数据，30%的数据作为测试数据以调整超参数大小。在训练数据集 dataset 中，将其中的 50%用做训练第一层分类器，其余的 50%用做训练第二份分类器。具体过程如下：

① 把训练集随机分为 5 个相同数量的子数据集，每个子数据集不能重叠。

② 对于单个模型，4 个子数据集依次作为训练集，其余的子数据集作为预测集，每个模型都可以输出自己预测集的预测结果。

③ 在上述建立的 5 个模型过程中，每个模型分别对测试数据集进行预测，并最终保留 5 列结果，然后对这 5 列取平均值，作为第 1 个基模型对训练数据的一个 Stacking 转换。

④ 选择第 2 个初级学习器，重复以上①~③操作，再次得到训练整个数据集在第 2 个基模型的一个 Stacking 转换。

⑤ 以此类推，有几个初级学习器，就会对整个训练数据集生成几列新的特征表达，同样，也会对测试集有几列新的特征表达。

⑥ 使用本文提出的基于特征敏感的逻辑斯蒂回归作为第二层的模型进行建模，分类预测。

整个分类训练过程如图 5 所示。

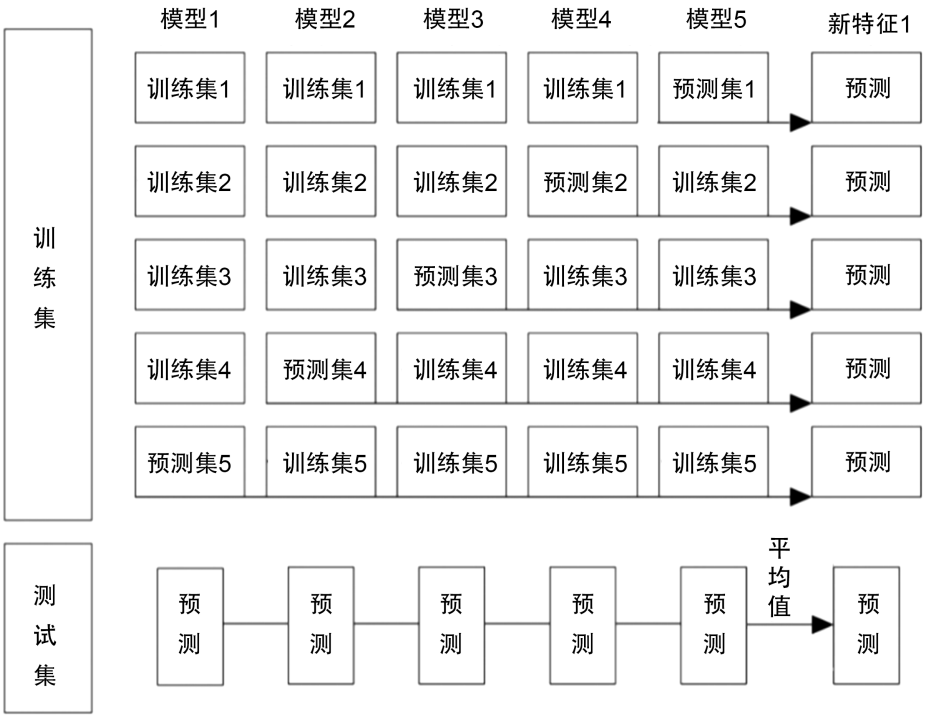

Figure 5. Training process
图 5. 训练过程

3.4. 实验结果

首先，选择几种常见的欺诈检测模型进行实验，其结果如表 2；之后，按照上文所述的 FSBS 模型进行训练，在使用 5 折交叉验证对第一层分类器进行处理后，根据经验以及训练结果不断对第二层分类器的 α 和 K 这两个超参数进行调整，最终实验结果如表 3。从实验结果的对比中可以看出：在泛化性上，本文提出的 FSBS 分类模型在 f1 值上与已被证实拥有良好性能的模型不分上下；在针对性上，FSBS 模型的平均 loss 值更低，说明 FSBS 对大金额的交易样本更加敏感，从而在减少金融领域欺诈检测经济损失的效果上更加明显。并且，调整 α 和 K 这两个超参数还可以改变模型对金额的敏感度，当 α 取 0 时，第二层分类器退化成普通的逻辑斯蒂回归。

Table 2. Results of common models
表 2. 常见模型实验结果

模型名称	精确率	召回率	F1 值	Loss 值
逻辑回归	76.0%	77.4%	76.7%	856.78
随机森林	92.7%	91.6%	92.2%	1551.72
Adaboost	80.8%	85.5%	83.3%	913.11
Original Stacking	92.7%	88.8%	90.7%	947.91

Table 3. Results of FSBS
表 3. FSBS 实验结果

参数设置	精确率	召回率	F1 值	Loss 值
α = 1，K = 1000	76.9%	79.6%	78.2%	468.32
α = 2，K = 1000	76.0%	75.6%	75.7%	362.03
α = 5，K = 1000	50.6%	71.3%	59.1%	220.66
α = 1，K = 2000	88.4%	85.6%	86.9%	517.63
α = 2，K = 2000	87.4%	81.5%	84.3%	363.20
α = 5，K = 2000	50.9%	80.0%	62.2%	272.47
α = 1，K = 5000	91.2%	88.7%	89.9%	591.05
α = 2，K = 5000	80.6%	90.4%	85.2%	397.44
α = 5，K = 5000	73.8%	87.9%	80.2%	303.56

4. 结论

在实际金融欺诈检测的过程中，对交易金额越大的样本的误判往往会带来更大的经济损失，但目前针对整体数据样本检测性能提升的方法，并不能保证这些样本的有效识别。因此本课题在现有欺诈检测方法的基础上，进一步关注特征敏感的数据样本的有效检测。通过改进集成学习 stacking 集成策略，修改模型第二层的损失函数，提出基于特征敏感的 stacking 分类方法 FSBS，以达到提升特殊样本判别准确率的目的。并最终通过实验证明，FSBS 可以在保证模型泛化性的基础上，有效降低信用卡欺诈交易的经济损失。

参考文献

[1] 360 企业安全. 2019 网络诈骗趋势研究报告[R]. 2019.

[2] Ghosh, S. and Reilly, D.L. (1994) Credit Card Fraud Detection with a Neural-Network. *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, Wailea, 4-7 January 1994, 621-630. https://doi.org/10.1109/HICSS.1994.323314

[3] Aleskerov, E., Freisleben, B. and Rao, B. (1997) CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection. *Proceedings of the IEEE/IAFE* 1997 *Computational Intelligence for Financial Engineering* (*CIFEr*), New York, 24-25 March 1997.

[4] Kolalikhormuji, M., Bazrafkan, M., Sharifian, M., *et al*. (2014) Credit Card Fraud Detection with a Cascade Artificial Neural Network and Imperialist Competitive Algorithm. *International Journal of Computer Applications*, **96**, 1-9. https://doi.org/10.5120/16947-6736

[5] 凌晨添. 进化神经网络在信用卡欺诈检测中的应用[J]. 微电子学与计算机, 2011, 28(10): 14-17.

[6] Maes, S., Tuyls, K., Vanschoenwinkel, B. and Manderick, B. (2002) Credit Card Fraud Detection Using Bayesian and Neural Networks. *First International NAISO Congress on Neuro Fuzzy Technologies*, Havana, 16-19 January 2002,

1-8.

[7] Sahin, Y., Bulkan, S. and Duman, E. (2013) A Cost-Sensitive Decision Tree Approach for Fraud Detection. *Expert Systems with Applications*, **40**, 5916-5923. https://doi.org/10.1016/j.eswa.2013.05.021

[8] Liu, C.W., *et al*. (2015) Financial Fraud Detection Model: Based on Random Forest. *International Journal of Economics & Finance*, **7**, 178-188. https://doi.org/10.5539/ijef.v7n7p178

[9] Lu, H., Yang, L., Yan, K., *et al*. (2016) A Cost-Sensitive Rotation Forest Algorithm for Gene Expression Data Classification. *Neurocomputing*, **228**, 270-276. https://doi.org/10.1016/j.neucom.2016.09.077

[10] Elkan, C. (2001) The Foundations of Cost-Sensitive Learning. In: Nebel, B., Ed., *Proceedings of the* 17*th International Joint Conference on Artificial Intelligence*, Morgan Kaufmann Publishers, Washington DC, 973-978.

[11] Zhang, Y. and Zhou, Z.H. (2010) Cost-Sensitive Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**, 1758-1769. https://doi.org/10.1109/TPAMI.2009.195

[12] Bnrescu, A. (2015) Detecting and Preventing Fraud with Data Analytics. *Procedia Economics and Finance*, **32**, 1827-1836. https://doi.org/10.1016/S2212-5671(15)01485-9

[13] Alkhateeb, Z.K. and Maolood, A.T. (2019) Machine Learning-Based Detection of Credit Card Fraud: A Comparative Study. *American Journal of Engineering and Applied Sciences*, **12**, 535-542. https://doi.org/10.3844/ajeassp.2019.535.542

[14] Webga, K. and Lu, A. (2015) Discovery of Rating Fraud with Real-Time Streaming Visual Analytics. 2015 *IEEE Symposium on Visualization for Cyber Security* (*VizSec*), Chicago, 25 October 2015. https://doi.org/10.1109/VIZSEC.2015.7312770

[15] Micenkova, B., McWilliams, B. and Assent, I. (2014) Learning Representations for Outlier Detection on a Budget.