Advances in Applied Mathematics 应用数学进展, 2021, 10(2), 386-395
Published Online February 2021 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2021.102044

Hans汉斯

基于CNN-SVM的信用卡诈骗检测方法

丁　冲，常磊雅，景英川*

太原理工大学，数学学院，山西　晋中
Email: dingchong_1123@163.com, 2892841748@qq.com, *shyjyc1970@163.com

收稿日期：2021年1月7日；录用日期：2021年2月11日；发布日期：2021年2月18日

摘　要

随着经济的发展，信用卡的普及，越来越多的信用卡交易出现了违规欺诈等行为，给国家和个人带来了巨大的经济损失。针对信用卡交易数据量大、特征数多和高度不平衡性(正常样本数量远高于诈骗样本数量)等特性，使得欺诈检测系统需进一步改进和完善。为减少银行和持卡人的损失，提出了一种基于卷积神经网络(CNN)和支持向量机(SVM)相结合的方法，即CNN-SVM法。该模型首先用SMOTE算法对原始数据中小样本进行处理以达到平衡数据的效果，再利用CNN对数据进行隐式特征提取，最后用SVM对提取后的特征数据进行检测。结合实例分析并比较得出：基于CNN-SVM的欺诈检测模型与传统的分类模型相比，有更加精准优良的效果。

关键词

信用卡诈骗，不平衡数据，SMOTE，卷积神经网络，支持向量机

A Credit Card Fraud Detection Method Based on CNN-SVM

Chong Ding, Leiya Chang, Yingchuan Jing*

School of Mathematics, Taiyuan University of Technology, Jinzhong Shanxi
Email: dingchong_1123@163.com, 2892841748@qq.com, *shyjyc1970@163.com

Received: Jan. 7th, 2021; accepted: Feb. 11th, 2021; published: Feb. 18th, 2021

Abstract

With the development of economy and the popularity of credit card, more and more credit card transactions have been illegal and fraudulent, which has brought huge economic losses to the country and individuals. Due to the large amount of credit card transaction data, the large number of features and the high imbalance (the number of normal samples is much higher than the num-

文章引用: 丁冲, 常磊雅, 景英川. 基于 CNN-SVM 的信用卡诈骗检测方法[J]. 应用数学进展, 2021, 10(2): 386-395.
DOI: 10.12677/aam.2021.102044

ber of fraud samples), the fraud detection system needs to be further improved and perfected. In order to reduce the losses of banks and cardholders, a method based on the combination of convolutional neural network (CNN) and support vector machine (SVM), namely the CNN-SVM method, is proposed. This model firstly uses SMOTE algorithm to treat the small sample of the original data to achieve the effect of balance data, then uses CNN to extract the implicit feature of the data, and finally uses SVM to detect the extracted characteristic data. Based on the example analysis and comparison, the fraud detection model based on CNN-SVM is more accurate and better than the traditional classification model.

Keywords

Credit Card Fraud, Skewed Data, SMOTE, Convolution Neural Network, Support Vector Machine

1. 引言

信用卡交易伴随着现代信息技术和全球化的发展变得越来越多。同时，信用卡诈骗问题也在急剧地增加。根据中国银行业协会的报告显示：信用卡诈骗的违法案例在逐年增加，并且每年由于信用卡犯罪导致的经济损失已达百亿元[1]。信用卡诈骗检测已经成为一个亟待解决的问题。信用卡诈骗形式千奇百怪，主要分为两种：一种是线上诈骗，一种是线下诈骗[2]，例如恶意透支，伪造信用卡等[3]。为了解决这种信用诈骗问题主要有两种方法：一是预防欺诈，二是欺诈检测。许多国内外学者针对信用卡诈骗检测做出了很多研究，并取得了巨大的成功。E. Aleskerov (1997) [4]讨论了一种可以基于神经网络的用于信用卡诈骗检测模型，并可以在自动加入接口的基础上提出了三种不同的神经网络结构；Chan (1999) [5]讨论了一种用于信用卡欺诈检测的分布式数据挖掘方法，使用可伸缩的黑箱方法来构建有效的欺诈探测器，可以显著减少由于非法诈骗行为造成的损失；Fiore (2019) [6]将生成对抗网络应用到信用卡诈骗检测系统，实验表明在增强集上训练的分类器上训练的分类器性能更好，特别是在灵敏度方面，形成了一种有效的欺诈检测机制；Yang (2020) [7]提出了联邦学习框架下卷积神经网络对信用卡检测，有效地利用了联邦学习特性达到了不仅可以保证算法准确性，也极大地避免了客户隐私地泄露的目的。

信用卡诈骗检测问题的本质可以被看成是一个二分类问题，即通过数据集特征区别被诈骗和未被诈骗两类。许多的统计学方法都可以用于分类，其中主流的分类方法可以分为两类，传统的机器学习算法和神经网络方法。传统的机器学习方法包括随机森林[8] [9] [10]，支持向量机[11]和 boosting [12]算法等。虽然以上这些机器学习算法针对小样本数据集可以达到很好的预测或分类效果，但当遇到大批量高维数据时无法达到理想的效果，然而深度学习算法可以更好地解决高维度复杂的数据，基于深度学习的算法可以更加准确地从大数据中提取有效的特征，从而构建更完美的模型。深度学习算法针对分类常用算法有：卷积神经网络[13] [14]、深度神经网络等。其中 CNN 在面对高维数据特征选取能力突出，所以广泛地应用在各个领域的检测方面。

本文首先利用 SMOTE [15]算法对原始数据中较少类样本进行随机采样处理达到平衡数据的效果，然后利用卷积神经网络对高维数据特征提取的能力将数据特征提取出来，最后应用支持向量机对提取出来的特征进行检测和分类，进而提出一种基于 CNN-SVM 方法用于信用卡诈骗检测。

2. 卷积神经网络和支持向量机

2.1. 卷积神经网络

当今人工智能正处于高速发展的阶段，神经网络在“深度学习”的名义下快速崛起。其中卷积神经网络作为神经网络中最具有良好性能的网络之一。它的主要架构包括：输入层、卷积层、池化层、全连接层和 Softmax 层等核心层次。如图 1 所示。

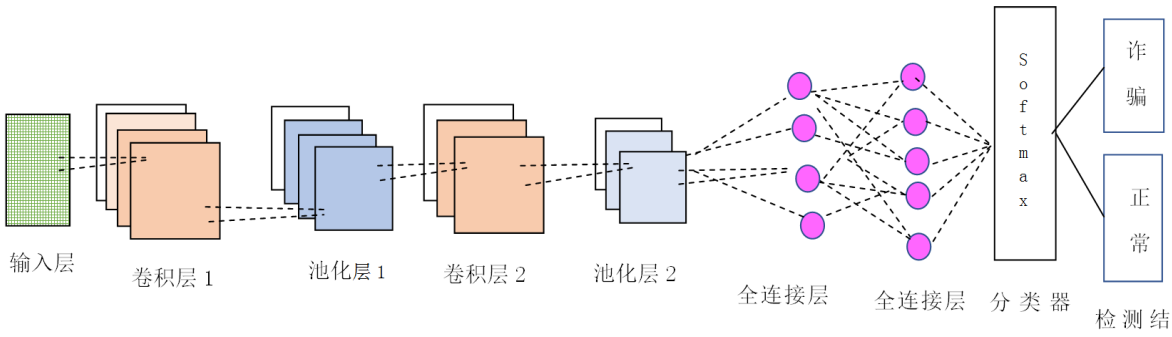

Figure 1. Structure of Convolutional neural network
图 1. 卷积神经网络结构

2.1.1. 卷积层

卷积层是通过卷积核(Convolution kernel)矩阵对输入层得到的数据进行卷积计算，即对输入数据的特征进行信息上特征选取或降维，并产生特征图(feature map)。卷积核类似一个窗口大小的扫描器，通过多次的扫描输入数据，来提取数据中重要的特征。

卷积的数学一般表达式如：

$$x_j^l = f\left(\sum_{i\in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{1}$$

其中，l 表示第 l 层卷积层；x_j^l 为第 l 层输出；x_j^{l-1} 为第 l 层输入；k_{ij}^l 表示权重矩阵；b_j^l 表示偏置；M_j 表示 $l-1$ 层特征图的第 j 个卷积区域；$f(\cdot)$ 表示激活函数。

2.1.2. 池化层

卷积神经网络中的池化层作为一种对输入特征提取的核心方式，它一般位于卷积层后面，它通过池化核对输入特征向量进行降采样，然后同时可以进一步对数据的特征进行提取，所以其具备以下特性：一是它实现了对输入原始数据的压缩，二是可以很大程度上减少关于参数的计算量，在一定意义上减小了模型的复杂度，提升了计算效率。

现在最常用的池化层可以分为两种：平均池化层(meaning pooling)和最大池化层(max pooling)。池化的数学一般表达式如下：

$$x_{i+1} = f\left(\beta * down\left(x_i\right) + b\right) \tag{2}$$

式中，x_i 表示输入；$down(\cdot)$ 表示为池化函数；β 表示乘性偏置；b 表示为加性偏置；$f(\cdot)$ 表示激活函数。

2.1.3. 全连接层和 Softmax 层

全连接层在整个神经网络中的主要作用是相当于“分类器”的作用，它将经过卷积和池化操作后的输入数据映射到样本标记空间。

原始数据经过卷积层和池化层处理后得到的特征矩阵通过全连接层得到的一维向量经常使用Softmax 函数进行分类。此函数常用于处理多分类问题，对于特例的二分类问题它可以使用广泛的逻辑函数。Softmax 函数是将一个 n 维的输入向量映射为 n 维的向量，得到由取值范围从 0 到 1 之间的元素组成的输出向量，且其所有组成元素和为 1，即所得的向量可以作为事件发生的概率。概率最大的数值所属类别概率最大，作为预测类别。

全连接层和 Softmax 层的数学表达式如下：

$$y^k = softmax\left(\omega^k * x^{k-1} + b^k\right) \tag{3}$$

其中：x^{k-1} 表示为全连接层的输入；b^k 表示为全连接层的输出；ω^k 表示为权重系数；b^k 表示为加性偏置；k 表示为第 k 层网络。

2.1.4. 激活函数

由于线性复合函数存在拟合能力有限的问题，所以引入激活函数使模型简单的映射转化为非线性映射，这样有助于提高模型的表达能力。常见的激活函数包括：sigmoid，ReLU 等。

Nair 和 Hinton 于 2010 年将修正线性单元(Rectified Linear Unit，简称 ReLU)应用于神经网络，函数图像如图 2 和定义如下：

$$f(z) = \max(0, z) \tag{4}$$

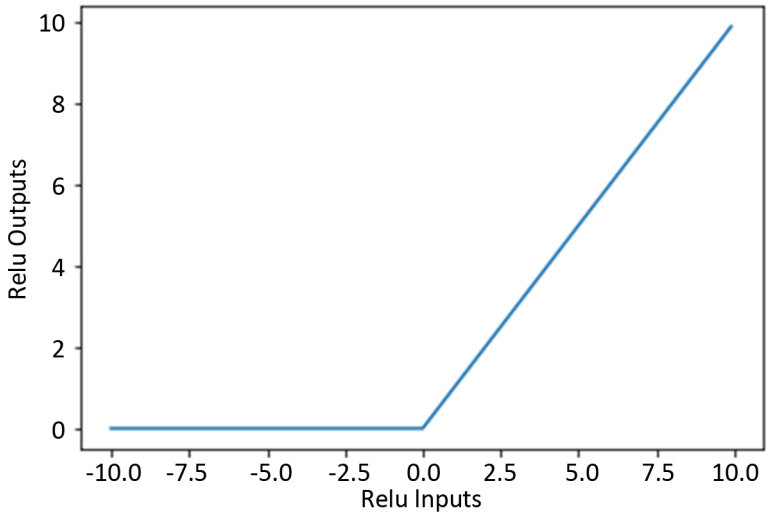

Figure 2. ReLU functional image
图 2. ReLU 函数图像

通过表达式可以看出 ReLU 函数目的是求最大值函数，它具备以下优点，当大于 0 时候的正区间上消除了梯度弥散的问题，并且它只要求判读是否大于 0，所以极大得提高了计算速度和收敛速度，所以一般在选取激活函数时，选取 ReLU 函数作为激活函数。

2.2. 支持向量机

在统计学习方法中，支持向量机(SVM)作为机器学习算法中最常用算法之一，构建它的条件就是训练数据必须是线性可分，其学习策略是最大间隔法。目标是利用结构风险最小化原则构造最优决策函数来解决二分类问题定义：分类的超平面数学表达示如下：

$$\omega^{\mathrm{T}} * x + b = 0 \tag{5}$$

其中：ω 为超平面的法向量，b 为相对偏移量。

Vanpik 证明了对于模式识别情况下训练一个支持向量机等价于二次最优化问题，如下所示。

$$\min:\ W(\alpha)=-\sum_{i=1}^{l}\alpha_i+\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}y_i y_j\alpha_i\alpha_j k(x_i,x_j) \tag{6}$$

$$\text{subject to:}\ \sum_{i=1}^{l}y_i\alpha_i=0 \tag{7}$$

$$\forall i:0\le\alpha_i\le C \tag{8}$$

其中 l 表示训练集中样本总数，α 表示 l 变量中的一个向量，α_i 对应于一个训练样本(x_i,y_i)。最优化问题的解法是在(7)，(8)式限制下通过最小化向量(6)式中的α 实现的。

定义矩阵 Q，Q 上式可以等价表示为：

$$\min:\ W(\alpha)=-\alpha^{\mathrm{T}}1+\frac{1}{2}\alpha^{\mathrm{T}}Q\alpha \tag{9}$$

$$\text{subject to:}\ \alpha^{\mathrm{T}}y=0 \tag{10}$$

$$0\le\alpha_i\le C1 \tag{11}$$

3. 基于 CNN-SVM 信用卡诈骗检测模型

卷积神经网络具有优秀的特征学习的能力，一个训练好的卷积神经网络模型可以通过卷积核池化操作从原始数据中学习和创建有用的特征，这些特征能最大化地表示所需任务的原始数据。因此本文构建了一个基于卷积神经网络自适应特征提取的信用卡诈骗检测系统，模型结构如图 3 所示。

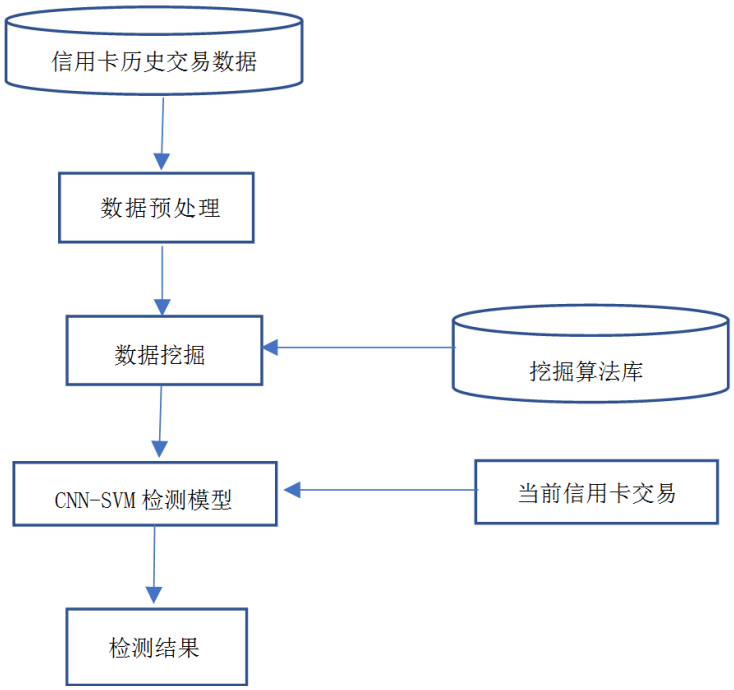

Figure 3. Fraud detection model based CNN-SVM
图 3. CNN-SVM 诈骗检测模型

该模型分为以下几个步骤：

1) 数据预处理：首先利用归一化将原始数据进行预处理，由于当前信用卡交易数据都属于高度不平

衡数据，由于随机过采样采取简单复制样本的策略来增加少数类样本，这样容易产生模型过拟合的问题，所以本文将采用 SMOTE 算法对少数类样本进行分析并根据少数类样本人工合成新样本添加到数据集中，算法如下所示。

算法 1：SMOTE 算法

输入 少数类样本 x_i，以及少数类样本集 S_{min}

1.对于少数类中的每一个样本，以欧式距离为标准计算它到少数类样本集 S_{min} 中所有样本的距离，得到其 k 近邻。

2.根据样本不平衡比例设置一个采样比例以确定采样倍率 N，对于每一个少数类样本 x_i，从其 k 近邻中随机选取若干个样本，假设选取的近邻为 x_i'。

3.对于每一个随机选出的近邻 x_i'，分别与原样本按照如下的公式构建新的样本

$$x_{new} = x_i + rand(0,1) \times (x_i' - x_i)$$

输出 新的样本 x_{new}

2) 特征提取：首先搭建好一个完整的卷积神经网络模型，然后将训练集用于训练卷积神经网络。当卷积神经网络训练完成后，移除卷积神经网络的输出层就得到了基于卷积神经网络的特征提取器。

3) 诈骗检测：将经过卷积层和池化层提取到的特征输入到支持向量机进行是否被诈骗的检测与分类。

4. 实验结果及分析

4.1. 实验数据

本文将通过实验得到的一系列评价指标证明这个模型的优越性。此实验数据集是由 ULB ML Group [16] 提供的欧洲信用卡交易数据集。这个数据集一共有 284,807 条交易记录，但是仅有 492 条被欺诈记录，即未被诈骗与被诈骗数量比例为 1:578，所以这个数据集表现出高度不平衡。由于信用问题和为保护用户客人隐私，数据中客户具备的原始特征没有被提供。这是一个非常典型的关于信用卡诈骗数据如表 1 所示：

Table 1. Data set of credit card fraud
表 1. 银行信用卡诈骗数据集

未被诈骗	诈骗	特征数	数据总量
284315	492	29	284807

从中可以看出，将原始数据进行重新平衡是非常必要的，否则将会针对正常交易检测出现过拟合问题，针对一些被诈骗的形式也会被忽略。

4.2. 实验环境及评价指标

实验计算机环境配置为 16 G 内存，i7-8750H 2.2 GHZ 处理器，操作系统为 Windows 10，实验语言为 Python，其中的框架使用 TensorFlow，开发工具为 Anaconda3。

评价一个机器学习算法的性能是非常重要的，它可以帮助我们寻找信用卡诈骗检测系统的最优参数。在机器学习领域中，混淆矩阵(confusion matrix)，又称为可能性表格或是错误矩阵。它是一种特定的 2 × 2 矩阵用来呈现算法性能的可视化效果，通常是监督学习(非监督学习，通常用匹配矩阵：matching matrix)。以分类模型中的最简单的二分类模型为例。对于这种问题，模型最终需要判断样本的结果是 0 还是 1，或者说是 positive 还是 negative。其中：

- True Positive = TP：真实值是 positive，模型认为是 positive 的数量

- False Negative = FN：真实值是 positive，模型认为是 negative 的数量：这就是统计学上的第二类错误(Type II Error)
- False Positive = FP：真实值是 negative，模型认为是 positive 的数量：这就是统计学上的第一类错误(Type I Error)
- True Negative = TN：真实值是 negative，模型认为是 negative 的数量

模型评估的矩阵如表 2 所示：

Table 2. Model evolution matrix
表 2. 模型评估矩阵

真实值 预测值	Positive (欺诈样本)	Negative (正常样本)
Positive (欺诈样本)	TP	FN
Negative (正常样本)	FP	TN

另外，本文中的对比实验中使用了准确率、召回率、精准率等性能指标，这些性能指标数学公式如下：

$$\text{Accr} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \tag{12}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{13}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14}$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recal}} \tag{15}$$

其中 Accr (准确率)表明所有实验记录中被正确分类的结果占总观测值的比重；Recall (召回率)表明在模型预测是 positive 的所有结果中，模型预测为正确的比重；Precision (精准率)表明在真实值是 Positive 的所有结果中，模型预测结果为正确的比重；F1 值表明 precision 与 recall 调和平均值。

当数据集为极不平衡数据集，针对诈骗检测模型表现 accuracy 作为测量指标是不足够准确的。因此，我们将考虑 precision，recall，F1 和 AUC 作为评价性能指标；受试者工作特征曲线(Receiver operating characteristic cure，简称 ROC 曲线)，又称感受性曲线(Sensitivity cure)，ROC 可以更好的描述针对不平衡数据分类器的表现。AUC (Area under cure)是统计学习中普遍使用的对于二分类评价指标，数值含义是 ROC 曲线所包含的面积值。

4.3. 实验结果

此欺诈模型中的神经网络结构由输入层，三层卷积层，两层池化层及全连接层组成。第一层输入层是由信用卡交易原始数据预处理后的数据集提取的原始数据特征组成，而后的第一层卷积层由 64 个尺寸为 5 × 5 的卷积核组成，第二层卷积层由 32 个尺寸为 3 × 3 的卷积核组成，第三层卷积层由 16 个尺寸为 3 × 3 的卷积核组成，其中前两层卷积层后面都连接一个最大池化层(Maxpooling)，最后一层是带有 ReLU 激活函数的全连接层；再经过卷积神经网络得到的提取特征输入到支持向量机对整个模型进行输出，对模型输出进行分类，检测交易数据是正常交易还是诈骗交易。

为了证明本文提出的模型的优越性，在实验中首先对信用卡数据通过SMOTE算法进行过采样处理，然后对数据进行随即分割，将处理后的数据的70%作为训练集，将原始数据的30%按照分层抽样的方法诈骗样本与正常交易样本比例进行抽样作为测试集。

在这部分，本文将利用 Python 语言分别对信用卡数据应用门控递归单元(GRU)，GRU-SVM，卷积神经网络，支持向量机及高斯朴素贝叶斯[17]作为参照模型，其中对于文章提出的诈骗检测系统中的卷积神经网络中的目标函数的优化操作上采取 Adam [18]优化算法，此优化算法是随机梯度下降算法的扩展式，通过计算梯度的一阶矩估计和二阶矩估计而为不同的参数设计独立的自适应性学习率，可以有效地减少过拟合问题并提高模型的泛化能力，能够获得更优秀的 CNN 模型。其中各个模型参数如表 3 所示。

Table 3. Parameters of each model
表 3. 各个模型参数

超参数	CNN-SVM	GRU	GRU-SVM	SVM	CNN
Batch_size	256	256	256	-	-
epochs	30	30	30	-	-
学习率	10^{-3}	10^{-3}	10^{-3}	-	10^{-3}
分类数量	2	2	2	2	2
Dropout	0.5	-	-	-	0.5

本实验中，分别应用本文提出的诈骗检测系统(CNN-SVM)与其他机器学习方法和经典的诈骗检测方法对欧洲信用卡实例数据进行训练，然后进行测试，得到 Precision，F1-score，Recall，AUC 的数值。以 Precision，F1-score，Ave_Precision，Auc 作为指标与 CNN-SVM 进行分类和检测效果对比，各个实验指标表格如表 4 所示和柱形图如图 4 所示。

Table 4. Experiment index of each model
表 4. 各个模型实验指标

评价指标	CNN-SVM	GRU	GRU-SVM	SVM	CNN	SMOTE + GNB
Precision	0.91	0.18	0.13	0.07	0.88	0.18
Avg Precision	0.96	0.59	0.57	0.54	0.94	0.15
F1 score	0.90	0.30	0.23	0.14	0.89	0.30
AUC	0.96	0.96	0.95	0.95	0.96	0.96

通过以上指标图表明，本文提出的 CNN-SVM 方法对正样本(诈骗样本)检测的准确率即 precision 达到 92%，相比于 CNN 算法大约有 3%的提升，并且远高于其他统计学习算法；F1 作为准确率(precision)和召回率(recall)的 Harmonic 平均值更适合区分诈骗样本和正常交易，作为本文提出的检测系统方法的 F1 值高达 90%，略高于 CNN 算法，远高于其他统计学习算法；CNN-SVM 针对信用卡交易样本的正负样本的平均值高达 96%，相比于 CNN 算法有 2%的提升，并且远高于其他统计学习算法；此检测系统的 AUC 达到 96.5%，对比其他检测算法略高于 1%，相比传统欺诈检测系统(使用 SMOTE(borderline2)平衡算法的传统欺诈检测系统[17]在相同数据集上的 AUC 达到 88%))性能大约提高了 10%。

以上实验表明，本文提出的基于 CNN-SVM 对信用卡交易诈骗检测具有良好的效果，能有效地从交易数据中检测出诈骗样本。在现实中，可以通过本文提出的信用卡欺诈检测模型检测当次交易结果是否存在欺诈行为，并且得到的结果需要在银行或者金融机构做进一步的检查和验证，所以本文提出的信用卡欺诈检测

系统对银行机构在对信用卡欺诈行为上起到了预警作用，并且极大地提高了检测的效率和结果的准确性。

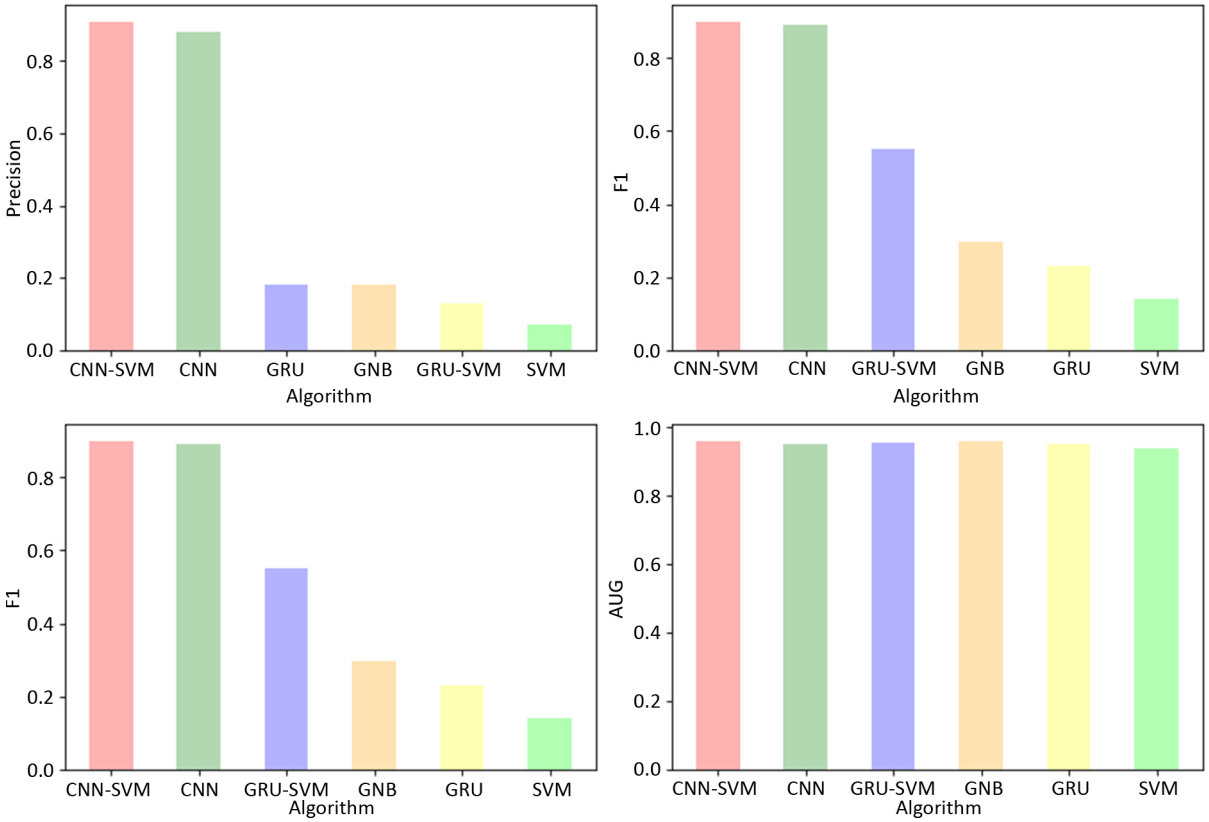

Figure 4. Comparison diagram of each algorithm index
图 4. 各个算法指标对比图

5. 结语

针对基于信用卡诈骗检测过程中存在由于卷积神经网络的全连接层的复杂而低效，由于银行交易的原始数据具有极度不平衡的特点，容易造成过拟合问题并且检测诈骗的准确率低的问题，本文提出并构建的一种将 CNN 和 SVM 结合方法来构建信用卡诈骗检测系统，不仅有效地解决了信用卡数据不平衡问题，而且此模型具有良好的泛化能力和模型检测的准确率，因为本文提出的检测模型综合了 CNN 和 SVM 的优势，所以 CNN-SVM 模型与 CNN、SVM 在很多情况下具有更好的鲁棒性和对正样本检测的准确性，通过卷积神经网络(CNN)在对目标函数的优化操作上采取 Adam 优化算法可以有效地减少过拟合问题并提高模型的泛化能力，能够获得更优秀的 CNN 模型。当前此诈骗检测模型只在 ULB 提供的数据集上运行，后期可以考虑使用其他的数据集运行以检测效果。同时，可以运用更复杂的卷积神经网络以进一步提高模型的准确率和运行效率。

基金项目

国家自然科学基金资助项目《高维数据变量间非线性交互作用的研究》(11571009)。

参考文献

[1] 中国银行业协会银行卡专业委员会. 中国银行卡产业发展蓝皮书 2018[M]. 北京: 中国金融出版社, 2018.

[2] Laleh, N. and Azgomi, M.A. (2009) A Taxonomy of Frauds and Fraud Detection Techniques. 2009 *International Con-*

ference on Information Systems, Technology and Management, Ghaziabad, 12-13 March 2009, 256-267. https://doi.org/10.1007/978-3-642-00405-6_28

[3] 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

[4] Chan, P.K., Fan, W., Prodromidis, A.L. and Stolfo, S.J. (1999) Distributed Data Mining in Credit Card Fraud Detection. *IEEE Intelligent Systems and their Applications*, **14**, 67-74. https://doi.org/10.1109/5254.809570

[5] Aleskerov, E., Freisleben, B. and Rao, B. (1997) CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection. *Computational Intelligence for Financial Engineering*, New York City, 24-25 March 1997, 220-226. https://doi.org/10.1109/CIFER.1997.618940

[6] Fiore, U., Santis, A.D., Perla, F., Zanetti, P. and Palmieri, F. (2017) Using Generative Adversarial Networks for Improving Classification Effectiveness in Credit Card Fraud Detection. *Information Sciences*, **479**, 448-455. https://doi.org/10.1016/j.ins.2017.12.030

[7] Yang, W., Zhang, Y., Ye, K., Li, L. and Xu, C.-Z. (2019) FFD: A Federated Learning Based Method for Credit Card Fraud Detection. *International Conference on Big Data* 2019, San Diego, 25-30 June 2019, 18-32. https://doi.org/10.1007/978-3-030-23551-2_2

[8] Breiman, L. (2001). Random Forests. *Machine Learning*, **45**, 5-32. https://doi.org/10.1023/A:1010933404324

[9] 董师师, 黄哲学. 随机森林理论浅析[J]. 集成技术, 2013, 2(1): 1-7.

[10] 王奕森, 夏树涛. 集成学习之随机森林算法综述[J]. 信息通信技术, 2018, 12(1): 49-55. http://dx.chinadoi.cn/10.3969/j.issn.1674-1285.2018.01.009

[11] Cortes, C. and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, **20**, 273-297. https://doi.org/10.1007/BF00994018

[12] Freund, Y. and Schapire, R.E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, **55**, 119-139. https://doi.org/10.1006/jcss.1997.1504

[13] Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. https://doi.org/10.1109/5.726791

[14] Lecun, Y. Boser, B., Denker, J. and Henderson, D. (2014) Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, **1**, 541-551. https://doi.org/10.1162/neco.1989.1.4.541

[15] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002) SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, **16**, 321-357. https://doi.org/10.1613/jair.953

[16] Ccfraud Dataset. https://www.kaggle.com/mlg-ulb/creditcardfraud

[17] Awoyemi, J.O., Adetunmbi, A.O. and Oluwadare, S.A. (2017) Credit Card Fraud Detection Using Machine Learning Techniques: A Comparative Analysis. 2017 *International Conference on Computing Networking and Informatics*, Lagos, 29-31 October 2017, 1-9. https://doi.org/10.1109/ICCNI.2017.8123782

[18] Kingma, D. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. Computer Science. arXiv:1412.6980. https://arxiv.org/abs/1412.6980