基于深度卷积网络的可疑交易识别

陈 靖　丁启禄

摘　要：可疑交易识别是反洗钱工作的重要内容，以算法模型为工具分析和识别可疑交易已成为新的趋势。深度卷积神经网络可有效自动提取数据中的分类特征，在众多分类任务中表现出较好的识别效果，已被广泛应用于各领域研究。本文首先基于深度学习理论，选取一维卷积神经网络，并设计了包含7层的模型框架应用于可疑交易识别分析。其次，将Elliptic数据集以7:3的比例划分为训练集和测试集，采用划分的数据对模型进行训练和测试，以GCN模型、Skip-GCN模型、EvolveGCN模型等深度神经网络模型为对照组，验证本文所提出模型的有效性。最后，通过将输入数据中各元素的排序方式随机打乱，探讨模型对数据输入的稳健性。研究结果表明，一维卷积神经网络对可疑交易识别具有较好的适用性，Elliptic数据集总体分类精度可达98%，F_1值达到80%，具有较好的分类效果。对比GCN模型、Skip-GCN模型、EvolveGCN模型等在Elliptic数据集上的识别效果，本文所提模型总体上具有较好的识别精度，总体正确率和F_1值均达到较高水平。

关键词：反洗钱　可疑交易识别　算法模型　深度卷积网络

洗钱是指借助各种手段将非法所得合法化的行为。由于洗钱可为毒品犯罪、黑社会性质的组织犯罪、恐怖活动犯罪等上游犯罪提供资金来源，扰乱社会稳定和金融市场秩序，具有较大的危害，各国均采取不同程度的反洗钱措施防范和遏制洗钱行为。我国针对洗钱行为也制定了许多法律法规，如《中华人民共和国反洗钱法》《金融机构反洗钱规定》等，并通过多种监管方式督促金融机构认真履行反洗钱义务。2012年，FATF发布最新的国际反洗钱标准《四十项建议》，突出了反洗钱有效性指标在反洗钱国际评估中的重要地位。2019年2月，FATF通过《中国反洗钱和反恐怖融资互评估报告》，标志着我国反洗钱工作进入了一个新的阶段。虽然我国反洗钱工作取得了阶段性成效，但比对国际标准和发达国家，仍存在较大差距。面对趋严的国际反洗钱标准、FATF互评估强化后续审查以及新一轮互评估，提升我国反洗钱工作有效性的任务仍紧迫和艰巨。金融机构作为资金流动的枢纽，反洗钱体系中的重要防线，其反洗钱工作有效性将对我国反洗钱工作的实效产生较大影响。对金融机构反洗钱有效性进行评估，一直以来也是FATF和反洗钱监管部门的一项重

陈靖、丁启禄，中国人民银行福州中心支行。

要任务。但从日常监管实践发现，金融机构反洗钱有效性与监管要求仍存在较大差距，部分机构反洗钱工作仅停留在表面，未深入理解和落实“以风险为本”的监管理念，阻碍了我国反洗钱工作的开展。

可疑交易识别是金融机构发现和防范洗钱行为的重要手段，其主要通过系统预设的监测标准对客户存在的异常行为进行实时监测，并及时预警提示金融机构业务人员关注异常客户情况，做到洗钱行为的精准预防。我国现行的反洗钱相关法律法规仅从原则性层面规定了金融机构异常交易监测标准设计的要求，金融机构对异常交易监测标准设计具有较大的自主决定权，造成异常交易监测的有效性无法得到保证，如有的金融机构全年未预警任何异常交易或金融机构预警了异常交易但核实后发现大部分均是客户正常交易，造成金融机构反洗钱资源的浪费，进而导致金融机构反洗钱有效性的降低，故提升可疑交易识别的有效性对反洗钱工作具有重要作用。

文献综述

现阶段金融机构可疑交易识别标准的设置主要大体存在两种模式，一是根据监管部门发布的风险提示或金融机构自身总结的风险点，结合客户的身份信息、交易频度、关联客户交易行为等设计监测预警指标，并通过设置不同的阈值实现可疑交易的预警，具有操作简单方便的优点。但现阶段洗钱形式多样复杂，洗钱手法不断变化，以及国际反洗钱内容的不断外延，基于单指标或者多指标的可疑交易识别标准已无法满足现阶段可疑交易预警的需求。二是基于机器学习算法，将已总结的洗钱犯罪行为，如地下钱庄、电信诈骗、跨境赌博等进行模型化，利用算法模型可不断更新迭代的优势，持续优化模型，提高可疑交易预警的有效性。

将机器学习算法应用于可疑交易识别大致可分为无监督的学习算法模型和有监督的学习算法模型。无监督的学习算法即依据客户的身份特性或者交易特性，采用聚类算法模型，对未知类别的样本进行聚类分群，进而识别出存在可疑交易的客户。如李欣月等提出基于CURE聚类算法的交易离群点识别模型；陈好孟将K-Means聚类算法应用于地下钱庄可疑交易监测分析中，并以已侦办的某地下钱庄案为例，验证了聚类算法在可疑交易监测中具有较好的识别效果；丁晓基于交易的时序特征，利用基于密度的聚类算法对交易资金序列进行自动分群成组，进而结合其他风险特征对可疑交易进行准确识别。聚类算法无须样本类别信息，对数据要求较低，具有较强的适用性。但由于其未利用样本类别信息，导致其无法保证聚类结果的准确性。

有监督的学习算法，又称分类算法，其与聚类算法最大区别在于训练模型时，分类算法的训练数据必须带有分类标签，以便指导模型进行更新迭代，达到可接受的分类精度。如刘云翔提出基于Fisher判断准则的BalanceCascade分类模式，并采用Adaboost分类算法，应用于银行卡异常

交易监测；肖琨采用多种分类算法，如支持向量机、多层感知机以及逻辑回归等应用于货币异常交易识别，并采用Paysim模拟器创建的货币交易数据验证模型的有效性；卢睿等采用随机森林算法进行特征选择，设计了可疑交易监测模型，并应用于信用卡交易数据可疑交易识别；Hassan H M将分类算法应用于检测欺诈性信用卡交易，并采用欧洲信用卡欺诈数据集分别验证了多层感知器、支持向量机、决策树等分类模型在信用卡异常交易监测的有效性。分类学习算法因训练过程采用样本的类别信息，使其具有较好的分类识别效果。但传统的分类算法，如随机森林、支持向量机等，为保证模型的分类精度，需耗费大量的人工成本对模型的特征进行选择，以便确定较优的特征集合，且对于不同的数据可能需要重新设计不同的分类特征变量，存在模型特征设计成本过高的缺点。

深度神经网络可在一定程度解决特征选择的问题，其可通过构造多层神经网络模型提取样本中较为稳健的特征集合，并在众多分类任务中表现优异。如陈泽瀛将商户异常交易监测应用于深度神经网络模型，提出基于LSTM和LightGBM组合集成分类模型，并以银联商务某地区的餐饮行业部分收单商户数据验证了模型的识别精度；黄良瑜采用LSTM模型对银行间债券市场异常交易行为进行检测，并通过实验数据证明了模型可有效提升异常交易检测的准确性；Weber M、Bellei C、Alarab I等基于交易拓扑图信息，采用图卷积神经网络模型或其改进模型对比特币异常交易数据进行识别，取得较好的识别效果。考虑到整合并提取交易拓扑图信息将降低计算效率，且一维卷积神经网络在其他分类任务取得较好效果，故本研究结合深度神经网络理论，选取一维卷积神经网络，并设计了包含7层的模型框架应用于可疑交易识别分析，以期为可疑交易识别提供一种新的方式。

研究方法

结合深度卷积神经网络理论，本文选取一维卷积神经网络应用于可疑交易识别分析，该模型框架包含7层神经网络，其中4层卷积层，3层全连接层，具体见图1。对于模型中的卷积层其步长均设置为1，Padding方式采用Valid，卷积核大小为3×1，激活函数采用Relu函数。第一层和第二层卷积层均采用128个卷积核对输入数据进行卷积操作，且在第二层卷积后使

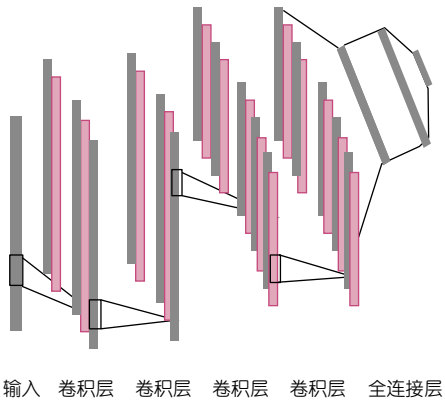

图1 模型框架图

用最大池化和Batch Normalization层。第三层和第四层卷积均采用256个卷积核抽取数据特征，并于第四层卷积后加入最大池化层和Batch Normalization层。第五层和第六层为全连接层，均设置了512个神经元与上一层神经元全连接，每层全连接层后均采用Dropout层防止模型过拟合。第七层为分类层，通过模型抽取的特征，采用Softmax函数对数据进行分类，得到每一类的分类概率。

数据来源与实验设置

Elliptic数据集

为了验证本文所提分类模型的有效性，选取加密货币合规公司Elliptic发布的比特币交易数据集作为实验数据训练一维卷积神经网络模型，主要理由如下：一是由于金融机构交易数据的敏感性以及客户数据的保密要求，目前暂无官方发布的金融机构可疑交易相关数据集；二是Elliptic数据集预先划分了训练和测试集，且存在较多研究者对其研究并发布相关模型精度指标，便于模型间精度的公平对比。

Elliptic数据集[①]为比特币交易信息数据集，其发布目的是为了提供真实的虚拟货币数据，方便研究者测试及验证可疑交易模型分类效果，推进可疑交易识别模型的发展。该数据集包括了20万笔交易记录，共203 769个节点和234 355条边，总价值达到60亿美元，其对交易数据标注分为三种类别，一是合法类别，即交易所、钱包提供商以及矿工等发起的正常合法交易数据，共有4 545条交易记录，占总量的2%；二是非法交易，即诈骗、恶意软件、恐怖分子组织等发起的非法交易数据，共计42 019条交易记录，占总量的21%；三是未知数据，即尚未标注的交易数据，占总量的77%。而且对于每条交易记录数据，Elliptic数据集提供了166个交易特征，其中前94个特征主要描述数据的基本情况，如时间步长、输入/输出数量、交易费用等基本信息，后72个特征主要描述关联主体间的交易特征，如相邻交易的最大值、最小值等。对于时间步长特征，其主要依据与交易记录关联的时间戳将所有交易数据划分为49个独立的时间步长，每个时间步长平均间隔两周左右，且不同的时间步长间不存在相关联的交易。综上所述，鉴于本文主要研究有监督的可疑交易识别模型，故仅采用Elliptic数据集存在标签的交易数据训练和测试模型，即合法类别4 545条交易记录，非法交易42 019条交易记录，数据总量共计46 564条交易记录。

实验设置和模型评价指标

本研究基于Elliptic数据集，结合深度神经网络理论，选取一维卷积神经网络，并设计了包含7层的模型框架应用于可疑交易识别分析。实验相关环境和参数设置如下：（1）实验环境。采用Windows10系统，硬件规格为：显卡GTX1650（4G），

① Elliptic, www.elliptic.co.

CPU（Core I7）、内存（8G），并采用Google开源的TensorFlow1.14深度神经网络库实现可疑交易一维卷积神经网络模型的构建和训练。（2）Elliptic数据训练和测试集划分。为了保证实验结果的可对比性，本文参考Mark Weber划分规则，依据交易的时间步长将训练数据和测试数据分为7:3，其中前34个时间步长所包含交易数据作为训练数据，后15个时间步长所包含的交易数据作为测试数据。（3）模型训练阶段参数设置。由于数据中正负样本比例差异较大，为了保证每一次迭代均存在正负样本，故本研究将Mini-batch设置为2 048个样本；Epoch设置为500；LR设置为0.01；损失函数采用交叉熵损失函数，并采用L2范数对模型权重进行正则化。（4）模型输入特征的选择。Elliptic数据集每条记录包含166个特征，但由于时间步长仅作为划分不同时间步长的标识，故将该特征剔除，仅利用其余的165个特征作为模型的输入。

本文采用四个分类评价指标对模型的分类效果进行评估，包括总体正确率（Overall Accuracy）、召回率（Recall）、精确率（Precision）以及F_1值，其中总体正确率为测试样本分类正确的总数量占所有样本的比重；召回率为测试数据中被正确预测的正样本占总正样本数的比重；精确率为测试数据中被预测为正样本中实际为正样本的比重；F_1值为召回率和精确率的加权平均值。上述四个指标数值均越大表明效果越好。

结果与分析

本文主要通过两部分实验探讨一维卷积神经网络应用于可疑交易识别任务的分类效果：（1）对比其他模型的分类表现，验证本文所提可疑交易一维卷积神经网络模型的有效性。（2）将输入数据随机打乱，探讨输入层特征的排列顺序对模型分类的影响，进而验证模型的稳健性。

模型对比实验

为了验证本文所提可疑交易一维卷积神经网络模型的有效性，且保证对比的合理性和有效性，参照前期相关研究成果，选择GCN模型、Skip-GCN模型、EvolveGCN模型、Label-GCN模型以及GCN-based模型等深度神经网络模型作为对比模型，其中GCN、Skip-GCN以及EvolveGCN模型LR设置为0.001，Epoch设置为1 000；Label-GCN模型LR设置为0.01，Epoch采用Early Stopping的方式，即当训练集上的损失函数值不再降低的时候停止继续训练；GCN-based模型LR设置为0.001，Epoch设置为50。各模型分类结果见表1，对比其他模型的分类效果，本文所提模型分类效果总体上较好，其中总体正确率与对比模型中表现最差的模型（GCN）相比高于2%，与对比模型中表现最好的模型相比高于1%；F_1值与对比模型中表现最差的模型（GCN）相比高于17%，与对比模型中表现最好的模型（GCN-based）相比高于3%；精确率与对比模型中表现最差的模型（GCN）相比高于13%，与对比模型中表现最好的模型

表 1 模型各评价指标精度对比

模型	精确率	召回率	F_1值	总体正确率
GCN	81%	51%	63%	96%
Skip-GCN	81%	62%	71%	97%
EvolveGCN	85%	62%	72%	—①
Label-GCN	86%	74%	76%	97%
GCN-based	90%	68%	77%	97%
本文模型（平均值）	94%	70%	80%	98%

（GCN-based）相比高于4%。虽召回率低于Label-GCN模型4%，但其他三项指标均高于Label-GCN模型，如精确度高于8%，F_1值高于4%，总体正确率高于1%。综上所述，对比其他深度神经网络模型的分类表现，本文所提可疑交易一维卷积神经网络模型分类效果明显较优，一定程度验证了该模型对可疑交易识别的有效性。

模型稳健性实验

由于模型卷积层的计算是通过滑动窗口对输入数据抽取其空间特征或数据间的关联特征，输入数据中各元素的排列方式对模型将产生一定影响，进而影响模型的分类效果。为了探讨输入数据中各元素的排序方式对模型产生的影响，进而验证模型是否稳健，本文采用无放回的简单随机抽样方式，将输入数据的165个特征随机排列，生成五组特征排列顺序不同的输入数据。并在相同的参数下分别训练模型，且为了直观反映其数据波动情况，将5组数据的测试分类评价指标绘制成箱线图（见图2）。由图2可知，五组数据所训练模型的分类效果总体趋于一致，标准差均小于0.05，未出现明显波动。波动最小的为总体正确率，平均值达到97.66%，标准差为0.001，其次是F_1值，平均值为79.64%，标准差为0.007。波动最大的为精确率，平均值为91.65%，标准差为0.024，其中五组数据中精确率最大值为93.52%，最小值为87.46%，极差达6.06%。由上述可知，本文所提模型具有一定的稳健性，对于不同排列方式的输入数据分类效果差异较小，总体未出现明显波动，可有效抽取可疑交易数据中潜在的代表性特征。

结论与建议

准确有效的识别客户交易记录中的可疑交易对防范洗钱风险和提高反洗钱履职能力具有重要意义。基于传统机器学习算法的可疑交易识别需耗费大量的人工成本对模型的特征进行选择，以便确定较优的特征集合，但其仍无法保证获得有效的分类特征。因此，本文利用深度神经网络自动抽取特征的优势，基于深度卷积神经网络理论基础，选取一维卷积神经网络，并设计了包含7层的模型框架，通过将其

① 参考文献未公开该模型的总体正确率指标数值。

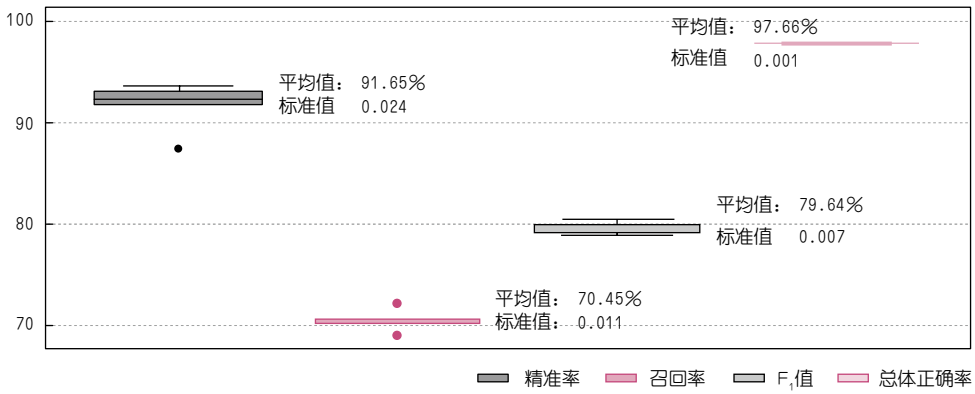

图 2 五组数据评价指标箱线图

应用于Elliptic比特币交易数据集以及对比不同分类模型，探讨模型的适用性和有效性。研究结论如下：（1）通过对比GCN模型、Skip-GCN模型、EvolveGCN模型、Label-GCN模型以及GCN-based模型等深度神经网络模型，本文所提可疑交易一维卷积神经网络模型识别效果均高于上述对比模型，验证了模型具有较好的有效性。（2）模型具有较好的稳健性，对模型输入数据元素的排列分布具有一定的容忍度，即将模型输入数据随机打乱，对模型分类效果的影响较小，未产生明显的分类精度波动。综上所述，一维卷积神经网络模型对可疑交易有较好的分类表现，且模型分类结果具有较好的稳健性和有效性，可为可疑交易识别提供一种反洗钱“科技+履职”新方案。

同时，对提高我国可疑交易识别有效性提出如下建议：一是积极组织开展我国各行业洗钱类型分析研究，由行业监管部门牵头，调动金融机构反洗钱人员积极投入相关研究，深度挖掘各行业存在高风险点和洗钱风险薄弱环节，并有针对性地研究防范措施，提升我国金融业反洗钱、反恐怖融资、反扩散融资的有效性。二是牵头行业监管部门或协会制定并发布各行业可疑交易特征的参考指引，总结行业可疑交易监测的履职共性问题和典型案例。金融机构在监管指引和典型案例的基础上自主开展自查自纠工作，并不断完善可疑交易监测自定义指标，以解决当前金融业可疑交易监测指标参差不齐、有效性不足等问题。三是推动金融机构充分利用洗钱风险自评估、可疑交易监测指标评估等机制。收集相关数据及资料，识别易被利用于洗钱的客户群、业务种类或交易渠道等，深入分析可疑交易报告，梳理出适用于本机构的洗钱类型及风险防控措施机制。N

学术编辑：卢超群

参考文献：
[1] 陈好孟.基于聚类算法的地下钱庄监测分析研究[J].金融发展研究,2020(06):9-16.
[2] 陈泽瀛,陶森林,蔡朝辉.基于LSTM和LightGBM组合模型的商户异常交易行为检测模型构建[J].数字技术与应用,2020,38(12):113-117.
[3] 丁晓.基于时间特征的可疑资金交易识别研究[J].现代计算机,2021(19):62-67.
[4] 黄良瑜,王蕙婷,詹杭龙,金健.基于深度学习的银行间债券市场异常交易行为检测[J].计算机应用与软件,2021,38(09):78-85.
[5] 刘云翔,唐泽苇,徐齐.基于级联平衡算法的银行卡异常交易检测[J].计算机仿真,2019,36(12):370-373.
[6] 李欣月,张高煜,彭兰舒,袁顺捷,常宇星,刘梦冬.基于聚类算法的金融交易离群点识别[J].电子技术,2016,45(01):24-28.
[7] 卢睿,李林瑛,孙永义.一种基于随机森林的可疑交易检测方法[J].辽宁工程技术大学学报(自然科学版),2021,40(01):82-89.
[8] 肖琨,王云,张桂刚.基于识别和多重分类的反洗钱系统[J].小型微型计算机系统,2019,40(10):2046-2051.
[9] Alarab I,Prakoonwit S,Nacer M I.Competence of Graph Convolutional Networks for Anti-Money Laundering in Bitcoin Blockchain[C].Proceedings of the 2020 5th International Conference on Machine Learning Technologies.2020: 23-27.
[10] Bellei C,Alattas H,Kaaniche N.Label-GCN: An Effective Method for Adding Label Propagation to Graph Convolutional Networks[J].arXiv preprint arXiv:2104.02153, 2021.
[11] Hassan H M,Ahmed A H A.The effectiveness of the Random Forest algorithm in monitoring abnormal withdrawals to detect credit cards frauds[J]. AL-BUTANA JOURNAL Of APPLIED SCIENCE,2022, 44-64.
[12] Weber M,Domeniconi G,Chen J,et al.Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics[J]. arXiv preprint arXiv:1908.02591, 2019.

Suspicious Transaction Recognition Based on Deep Convolutional Networks

CHEN Jing　DING Qilu
(Fuzhou Central Sub-branch of People's Bank of China)

Abstract　The identification of suspicious transactions is an important part of anti-money laundering work. The use of algorithms as tools to analyze and identify suspicious transactions has become a new trend. The use of deep convolutional neural networks could effectively extract classification features in data automatically. It has shown a good recognition effect in many classification tasks and has been widely used in various fields of research. The research steps of this article were as follows: Firstly, based on deep learning theory, this paper selected a one-dimensional convolutional neural network, and designed a model framework of seven layers for suspicious transaction identification and analysis. Secondly, the Elliptic data set was divided into a training set and test set at a ratio of 7:3. The model was trained and tested with the divided data. The GCN model, Skip-GCN model, EvolveGCN model and other deep neural network models were used as control groups to verify the effectiveness of the model proposed in this paper. Finally, the robustness of the model to the data input was discussed by randomly scrambling the order of each element in the input data. The research results were as follows: The one-dimensional convolutional neural network had good applicability for suspicious transaction recognition. The overall classification accuracy of the Elliptic data set could reach 98%, and the F_1 value could reach 80%, which had a good classification effect. Comparing the recognition effects of the GCN, Skip-GCN, EvolveGCN and other models on the Elliptic data set, the model proposed in this study had generally good recognition accuracy. The overall accuracy and F_1 value reached a high level.

Keywords　Anti-Money Laundering, Identification Of Suspicious Transactions, Algorithm Model, Deep Convolutional Neural Network
JEL Classification　C45　C69　G29