

深度学习在金融领域的应用研究综述

付钰菲, 汪明艳

(上海工程技术大学管理学院, 上海 201620)

✉291732841@qq.com; wmy61610@126.com



摘要: 随着人工智能和金融科技的快速发展, 机器学习尤其是深度学习在金融领域的应用引起了人们浓厚的研究兴趣。为了探索金融深度学习的应用领域, 对近十年金融深度学习的文献进行了总结, 并分别从模型介绍和应用领域两个方面进行了归纳。结果发现: 金融深度学习常用的模型包括卷积神经网络、循环神经网络和长短期记忆神经网络, 并且它们在金融文本分析、金融风险评估和异常检测以及投资组合管理方面均有着广泛的应用。未来可以将新的文本挖掘和自然语言处理技术应用到行为金融学领域进行更深入的研究, 同时还可以探索将深度学习应用到加密货币和区块链等新兴金融领域的更多可能性。

关键词: 深度学习; 神经网络; 文本分析; 风险评估; 投资组合管理

中图分类号: TP399 **文献标识码:** A

A Research Overview of the Application of Deep Learning in the Financial Field

FU Yufei, WANG Mingyan

(School of Management, Shanghai University of Engineering Science, Shanghai 201620, China)

✉291732841@qq.com; wmy61610@126.com

Abstract: With the rapid development of artificial intelligence and financial technology, the application of machine learning, especially deep learning in the financial field, has aroused strong research interest. In order to explore the application fields of financial deep learning, this paper proposes to summarize the literature of financial deep learning over the past ten years, from two separate aspects: model introduction and application field. Results show that the commonly used models of financial deep learning include convolutional neural network, recurrent neural network and long short-term memory neural network. They have a wide range of applications in financial text analysis, financial risk assessment and anomaly detection, and portfolio management. In future, new text mining and natural language processing techniques can be applied to the field of behavioral finance for more in-depth research. At the same time, more possibilities for applying deep learning to emerging financial fields such as cryptocurrency and blockchain can also be explored.

Keywords: deep learning; neural network; text analysis; risk assessment; portfolio management

1 引言(Introduction)

早在40年前的国外, 金融就一直是机器学习研究最多的应用领域之一。我国关于机器学习在金融领域的应用研究虽然起步较晚, 但是随着我国人工智能的发展, 文献数量也急剧增长, 股票市场预测、金融风险评估、投资组合管理等是其中最需要解决的几个问题。而在机器学习领域中, 深度

学习又是一个新兴领域。相对于传统的机器学习算法在训练之前需要预处理和特征提取, 深度神经网络可以使用非线性处理单元的级联来进行自动特征提取和转换, 这些层可以使神经网络发现数据之间的非线性关系, 并提高从原始数据学习的能力(LECUN, 2015)^[1]。基于此, 本文首先介绍了金融领域最常用的三个深度学习模型, 包括卷积神经网络

(Convolutional Neural Network, CNN)、循环神经网络(Recurrent Neural Network, RNN)和长短期记忆神经网络(Long Short-Term Memory, LSTM);然后通过收集并整理相关文献对深度学习在金融领域的几个应用做了详细介绍;最后进行了总结分析。

2 金融领域常用的深度学习模型介绍(Introduction to deep learning models commonly used in the financial field)

深度学习模型是一种具有多层神经网络的复杂体系结构,可以逐步从输入中提取高级特征。金融领域出现最多的深度学习模型主要有三种:卷积神经网络(CNN)、循环神经网络(RNN)、长短期记忆神经网络(LSTM)。

2.1 卷积神经网络(CNN)

CNN是一种受人类视觉机制启发的深度学习模型,主要用于图像分类、图像识别问题,后来也被用于文本分类。在大多数CNN架构中,有不同类型的层:输入层、卷积层、池化层(平均值或最大值)和全连接层。具有不同层的广义CNN架构如图1所示。

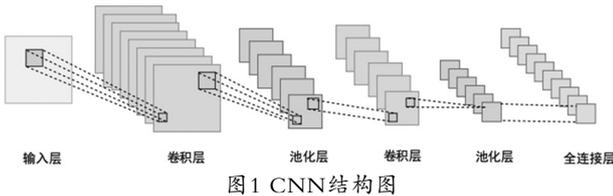


Fig.1 Convolutional neural network architecture

输入层获取原始输入,接下来特征提取层(包括卷积层和池化层)学习相关的特征。卷积层运用被称为特征检测器的滤波器来学习特征并生成特征图,一般使用 1×1 、 3×3 和 5×5 的滤波器尺寸。池化层也被称为降维方法,用来提取相关的特征,留下不必要的特征,这样就减少了输入的空间维度。最大池化是运用最多的池化类型,它从每个窗口取最大值。经过卷积运算获得的输出层和池化层的降维可以加快计算速度。最后,将特征提取层产生的特征传递到全连接层,其中全连接层中的每个神经元都连接到前一层的神经元。

CNN的优点是可以通过卷积自动对需要的特征进行提取,并且它还共享卷积核,可以用来处理高维数据。CNN的缺点是在池化层采样的过程中丢失很多有用的信息,并且由于特征提取的封装,为改进网络性能罩了一层黑盒。

2.2 循环神经网络(RNN)

RNN主要用于处理时间序列数据,包括音频、语音和语言等序列数据,由连续结构化的RNN单元组成,结合之前状态的RNN结构,如图2所示。用于传统预测分析的人工神经网络(Artificial Neural Network, ANN)模型不适用于序列数据,因为它们将每个输入视为一个独立的实体,而序列数据中的观察结果不是相互独立的。与其他前馈网络不同的是,RNN使用内部存储器来处理传入的输入,在其操作过程中,逐个处理序列数据。值得注意的是,它考虑了处理序列中元

素的时间因子,使用隐藏状态保留先前处理的观察结果,并在下一个将要处理的观察结果中使用。因此,RNN中的输出不仅取决于当前的输入,而且还取决于从网络先前隐藏状态计算出的输出。

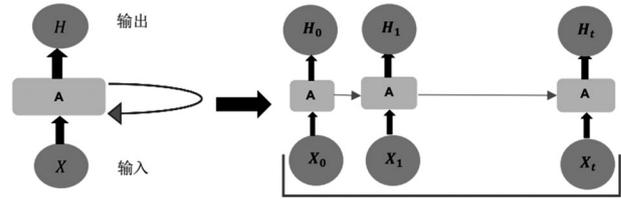


图2 结合之前状态的RNN结构图

Fig.2 RNN architecture combined with previous states

RNN的特殊结构使其具有如下两个优点:首先,它可以针对任意长度的序列输入向量进行建模;其次,在处理序列数据时,可以考虑到每个时间步骤的前后信息。RNN中的信息通过循环传播,这使得模型可以使用相同的参数,从而降低参数的复杂度。RNN的缺点是不支持长期记忆并且面临梯度消失问题。

2.3 长短期记忆神经网络(LSTM)

LSTM是RNN的一种变体,LSTM网络由LSTM单元组成。LSTM单元由输入门、输出门和遗忘门组成,这三个门控制着信息的流动。遗忘门决定应该保留或丢弃单元状态中的哪些信息,而输入门负责在单元状态中应该存储哪些新信息,输出门接收当前输入、先前隐藏状态输出和新计算的单元状态,以便为当前输入观察依次生成新的隐藏状态和输出。有了这些功能,每个单元格都会记住任意时间间隔内的所需值。基本的LSTM单元如图3所示(σ : sigmoid函数; \tanh : 双曲正切函数; \times : 乘法; $+$: 加法)。

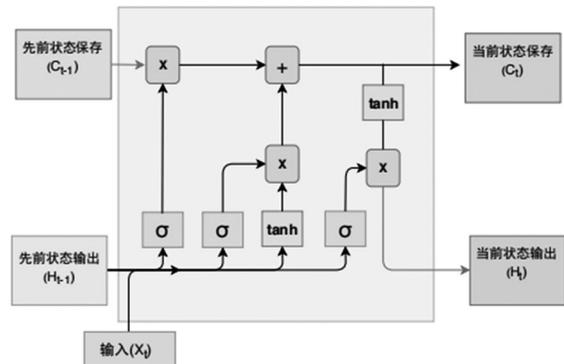


图3 基本的LSTM单元

Fig.3 Basic LSTM unit

LSTM通过引入门控机制解决了RNN的缺点,如梯度消失和不支持长期记忆问题。LSTM网络的优势在于可以记住网络中的短期和长期值,因此被广泛应用于自动语音识别、语言翻译、手写字符识别、时间序列数据预测等序列数据分析。

3 深度学习在金融领域的应用(The application of deep learning in the financial field)

3.1 金融文本挖掘

随着社交媒体和实时新闻媒体的快速传播,基于即时文

本的信息检索也被应用于金融模型开发。通过对新闻、财务报表、公司信息披露等内容的上下文进行分析可以得到很多有用的信息，因此金融文本挖掘研究在最近几年变得非常流行。例如，HAN等人^[2]首先基于一种新的事件类型模式分类算法，对中国企业的不同事件类型进行了分类。此外，还使用其他的输入因素对股票价格进行了预测。KRAUS等人^[3]基于金融新闻和股票市场数据，使用文本挖掘技术实现了具有迁移学习的LSTM模型。DANG等人^[4]使用Stock2Vec和双流GRU(TGRU)模型从金融新闻和股票价格中生成输入数据，用于股票价格分类。张梦吉等人^[5]提出了一种基于CNN-RNN的新闻事件分类模型，并运用LSTM模型结合资金流向和公司财务来对个股趋势进行预测。

3.2 金融文本情感分析

行为金融学最重要的组成部分之一是情绪或投资者情绪。文本挖掘技术的进步为通过网络媒体或社交媒体成功提取情感提供了可能性。人们对金融情感分析越来越感兴趣，尤其是将使用深度学习的情感分析模型应用于金融市场预测方面。ZHUGE等人^[6]利用指数数据的价格和文本帖子中的情绪数据预测第二天的股票开盘价。DAS等人^[7]利用Twitter情绪数据和股票价格数据预测谷歌、微软和苹果三个公司股票的价格。耿立校等人^[8]提出了一种基于CNN的情感分析模型来构建投资者情绪特征，并运用LSTM模型进行股票趋势的预测。温玉莲等人^[9]采用Word2vec和BiLSTM相结合的方法进行情感分析，然后利用股票时序数据和情感指标进行股票预测。

3.3 异常检测

金融欺诈检测也被称为异常检测，是各国政府都非常重视的一个问题，亟待研究出一个良好的解决方法。异常检测可以说是机器学习研究最广泛的领域之一，一般是分类问题。

信用卡异常检测是此研究领域研究最多的一个问题。ANDER等人^[10]运用MLP网络对信用卡交易是否存在异常进行分类。JURGOVSKY等人^[11]使用LSTM从信用卡交易序列中检测信用卡欺诈问题，并将结果与RF进行了比较。杨磊等人^[12]应用深度学习中的Transformer编码器挖掘用户的数据和信息，然后将其作为分类模型的输入来判断用户是否违约。刘颖等人^[13]首先构建了利用boosting算法增强的支持向量机(SVM)和随机森林(RF)结合的基分类器簇，然后运用深度信念网络(DBN)和基分类器簇共同进行信用卡欺诈的分类。

对于其他异常检测问题，WANG等人^[14]使用文本挖掘和DNN模型来检测汽车保险欺诈。GOUMAGIAS等人^[15]使用深度Q学习预测风险厌恶型企业的逃税行为，并提供了使税收收入最大化的相关建议。黄良瑜等人^[16]提出一种基于网络嵌入和LSTM模型的异常交易行为检测方法，能有效检测出市场上的异常交易行为。

3.4 金融风险评

风险评估指的是识别任何给定资产、公司、个人、产品、银行等的“风险”，包括但不限于破产预测、财务困境

预测、企业信用评级、债券评级、贷款申请、消费者信用评分等。能够准确地识别这些风险是至关重要的，因为资产定价高度依赖于这些风险评估措施。其中一个最广为人知的例子就是2008年全球金融危机，由于金融机构间信用违约互换(CDS)风险评估不当的抵押贷款危机导致房地产泡沫破灭，进而导致全球经济大萧条，而将深度学习应用于风险评估领域可以获得更高的准确性。

大多数风险评估研究集中在信用评分和财务困境预测上。关于信用评分分类问题，LUO等人^[17]使用信用违约互换数据进行企业信用评级，并将其信用等级划分为A、B、C三类。在测试的模型中，带有受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)的深度信念网络(Deep Belief Network, DBN)表现最好。类似地，YU等人^[18]采用DBN、反向传播(Back Propagation, BP)和SVM的级联混合模型进行信用分类，取得了良好的性能结果(准确率为80%—90%)。王重仁等人^[19]提出了一种融合注意力机制的LSTM和CNN相结合的神经网络模型来对个人信用进行评分，以提高信用风险评估的准确性。陈学彬等人^[20]运用擅长处理长时间依赖关系的LSTM神经网络方法构建了中国信用债违约风险预测模型，并且达到了较高的预测准确率。

关于银行和企业的财务困境预测问题，RONNQVIST等人^[21]运用文本挖掘识别银行是否处于困境，具体做法为从金融新闻中提取数据，然后使用深度前馈网络(Deep Feedforward Network, DFFN)对从词嵌入中提取的语义句子向量进行分类。KVAMME等人^[22]使用CNN和RF模型来预测客户是否会违约其抵押贷款。

3.5 投资组合管理

投资组合管理是在一个给定时期内选择投资组合中的各种资产的过程。一般来说，投资组合管理包括以下相关领域：投资组合优化、投资组合选择、投资组合分配。投资组合管理实际上是一个优化问题，确定在给定时期内选择表现最好的资产的最佳可能行动路线。

针对投资组合选择问题，LEE等人^[23]比较了用于股票价格预测的三个RNN模型(S-RNN、LSTM、GRU)，然后根据预测选择股票来构建基于阈值的投资组合。张虎等人^[24]运用自注意力神经网络模型结合过去60个交易日的因子数据预测沪深300成分股未来一个月的价格变动趋势，从而选出前50个上涨概率最大的股票构建投资组合。闫洪举^[25]提出了一种基于深度自动编码器的方法，以确定构建指数跟踪组合所需的股票，并利用深度神经网络模型对个股的权重进行测算。

4 结论(Conclusion)

金融领域常用的深度学习模型主要有三种：CNN、RNN、LSTM。CNN使用卷积滤波器来识别数据中的模式，广泛应用于图像识别领域和自然语言处理领域。RNN旨在识别序列模式，它在上下文至关重要的情况下尤其强大，因此常用于情感分析。LSTM网络是一种特殊的RNN，能够学习

长期上下文和依赖关系,弥补了RNN的梯度消失和不支持长期记忆的问题。金融深度学习主要的应用领域有金融文本分析、异常检测、金融风险评估和投资组合问题。其中,金融文本分析为研究最多的领域,通常运用文本挖掘和情感分类对财经新闻、社交媒体等文本信息进行分析,从而实现了对股票价格、股票趋势等的预测。

虽然文本挖掘和情感分析已经受到了很多的关注,但行为金融学是一个广泛的研究领域,还有很多值得深入研究的东西,尤其是在我国利用深度学习进行研究的文献还比较少。而文本挖掘技术、自然语言处理技术以及基于代理的计算金融和语义相结合的技术的发展为此研究领域提供了新的研究机会,未来可以进一步对这一领域进行研究。此外,随着一些新的金融领域的兴起和流行,它们对金融市场的影响也越来越大,例如加密货币和区块链技术,将深度学习创新地应用于这些新兴的金融领域可以为金融市场探索更多的可能性。

参考文献(References)

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553):436-444.
- [2] HAN S Q, HAO X L, HUANG H L. An event-extraction approach for business analysis from online Chinese news[J]. Electronic Commerce Research and Applications, 2018, 28(3/4):244-260.
- [3] KRAUS M, FEUERRIEGEL S. Decision support from financial disclosures with deep neural networks and transfer learning[J]. Decision Support Systems, 2017, 104:38-48.
- [4] DANG L M, SADEGHI-NIARAKY A, HUYNH H D, et al. Deep learning approach for short-term stock trends prediction based on Two-Stream Gated Recurrent Unit Network[J]. IEEE Access, 2018, 6:55392-55404.
- [5] 张梦吉,杜婉钰,郑楠.引入新闻短文本的个股走势预测模型[J].数据分析与知识发现,2019,3(05):11-18.
- [6] ZHUGE Q, XU L Y, ZHANG G W. LSTM neural network with emotional analysis for prediction of stock price[J]. Engineering Letters, 2017, 25(2):167-175.
- [7] DAS S, BEHERA R K, KUMAR M, et al. Real-time sentiment analysis of twitter streaming data for stock prediction[J]. Procedia Computer Science, 2018, 132:956-964.
- [8] 耿立校,刘丽莎,李恒昱.多源异构数据融合驱动的股票指数预测研究[J].计算机工程与应用,2021,57(20):142-149.
- [9] 温玉莲,林培光.基于行业背景差异下的金融时间序列预测方法[J].南京大学学报(自然科学),2021,57(01):90-100.
- [10] ANDER G J, AREVALO J, PAREDES R, et al. End-to-end neural network architecture for fraud scoring in card payments[J]. Pattern Recognition Letters, 2017, 105(04):175-181.
- [11] JURGOVSKY J, GRANITZER M, ZIEGLER K, et al. Sequence classification for credit-card fraud detection[J]. Expert Systems with Applications, 2018, 100(06):234-245.
- [12] 杨磊,姚汝婧.基于Transformer的信用卡违约预测模型研究[J].计算机仿真,2021,38(08):440-444.
- [13] 刘颖,杨轲.基于深度集成学习的类极度不平衡数据信用卡欺诈检测算法[J].计算机研究与发展,2021,58(03):539-547.
- [14] WANG Y B, XU W. Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud[J]. Decision Support Systems, 2018, 105(01):87-95.
- [15] GOUMAGIAS N D, HRISTU-VARSAKELIS D, ASSAEL Y M. Using deep Q-learning to understand the tax evasion behavior of risk-averse firms[J]. Expert Systems with Application, 2018, 101(07):258-270.
- [16] 黄良瑜,王蕙婷,詹杭龙,等.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.
- [17] LUO C C, WU D S, WU D X. A deep learning approach for credit scoring using credit default swaps[J]. Engineering Applications of Artificial Intelligence, 2017, 65:465-470.
- [18] YU L A, ZHOU R T, TANG L, et al. A DBN-based resampling SVM ensemble learning paradigm for credit classification with imbalanced data[J]. Applied Soft Computing, 2018, 69:192-202.
- [19] 王重仁,王雯,余杰,等.融合深度神经网络的个人信用评估方法[J].计算机工程,2020,46(10):308-314.
- [20] 陈学彬,武靖,徐明东.我国信用债个体违约风险测度与防范——基于LSTM深度学习模型[J].复旦学报(社会科学版),2021,63(03):159-173.
- [21] RONNQVIST S, SARLIN P. Bank distress in the news: Describing events through deep learning[J]. Neurocomputing, 2017, 264:57-70.
- [22] KVAMME H, SELLEREITE N, AAS K, et al. Predicting mortgage default using convolutional neural networks[J]. Expert Systems with Application, 2018, 102(07):207-217.
- [23] LEE S I, YOO S J. Threshold-based portfolio: The role of the threshold and its applications[J]. The Journal of Supercomputing, 2018, 76(10):8040-8057.
- [24] 张虎,沈寒蕾,刘晔诚.基于自注意力神经网络的多因子量化选股问题研究[J].数理统计与管理,2020,39(03):556-570.
- [25] 闫洪举.基于深度学习的指数跟踪方法研究[J].统计与决策,2021,37(05):143-147.

作者简介:

付钰菲(1996-),女,硕士生.研究领域:信息管理。
汪明艳(1975-),女,博士,教授.研究领域:应急管理,舆论治理,电子商务。