

基于LSTM-AR模型的光伏指数保险欺诈检测方法研究

贾伟, 杨扬

广西大学中国-东盟经济学院, 广西 南宁

收稿日期: 2024年7月4日; 录用日期: 2024年7月30日; 发布日期: 2024年9月20日

摘要

光伏指数保险是一种基于太阳辐射总量的天气指数保险产品, 旨在为光伏电站提供发电量不足的风险保障。然而, 光伏指数保险中存在着因缺乏客观数据与评价标准而导致的保险欺诈问题。本文针对这一问题, 提出了一种基于深度学习的模型, 结合LSTM、注意力机制和残差连接, 该模型使用太阳辐射量对光电企业发电量进行预测。为验证模型的有效性, 本文使用2019年新疆光电光伏数据集进行测试并与传统的最小二乘法的回归预测进行比较。测试结果表明, 本文所提出的模型可以有效使用太阳辐射量对光电企业发电量进行预测, 进而帮助保险企业针对光伏企业的发电量有关的保险欺诈行为进行识别, 为光伏保险业务的规范化和发展提供了一种有效的技术手段。

关键词

光伏指数保险, 保险欺诈, 神经网络

Research on Photovoltaic Index Insurance Fraud Detection Method Based on LSTM-AR Model

Wei Jia, Yang Yang

China-ASEAN School of Economics, Guangxi University, Nanning Guangxi

Received: Jul. 4th, 2024; accepted: Jul. 30th, 2024; published: Sep. 20th, 2024

Abstract

Photovoltaic (PV) index insurance is a weather index insurance product based on the total amount of solar radiation, designed to provide risk protection against insufficient power generation for PV

文章引用: 贾伟, 杨扬. 基于 LSTM-AR 模型的光伏指数保险欺诈检测方法研究[J]. 金融, 2024, 14(5): 1737-1745.

DOI: 10.12677/fin.2024.145177

power plants. However, PV index insurance faces issues of insurance fraud due to the lack of objective data and evaluation standards. To address this problem, this paper proposes a deep learning-based model that combines LSTM, attention mechanism, and residual connections. This model uses solar radiation data to predict the power generation of photovoltaic enterprises. To validate the effectiveness of the model, the 2019 Xinjiang photovoltaic dataset was used for testing and compared with traditional least squares regression predictions. The test results show that the proposed model can effectively use solar radiation data to predict the power generation of photovoltaic enterprises, thereby helping insurance companies identify fraudulent activities related to power generation in PV enterprises. This provides an effective technical means for the standardization and development of PV insurance business.

Keywords

Photovoltaic Index Insurance, Insurance Fraud, Neural Network

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着“双碳”目标的提出,我国各地区各有关部门围绕能源绿色低碳发展制定了一系列政策措施,推动了太阳能、风能、水能、生物质能、地热能等清洁能源的开发利用。

光伏发电是一种利用太阳能产生清洁电力的方式,作为主要的清洁能源之一,它有利于实现我国的“双碳”目标。然而,现今光伏电站建设运营以及光伏产品生产过程中仍存在较大风险。例如:光伏发电是利用半导体的光电效应,将太阳能直接转化为电能的一种技术,虽然这项技术有很高的应用前景,但这项技术很大程度上取决于天气状况,所以存在天气原因导致发电不足的风险[1]。并且虽然光伏发电不需要燃料,但是其初始投资较高,其光伏单位千瓦成本约为6500元,且其运行和投资回收期更长[2]。这就导致了光伏企业需要大量融资,但银行等金融机构对光伏项目的贷款意愿不高的问题[3]。

天气指数保险作为一种新兴的绿色保险产品,在国内具有重要意义。它能够帮助风光水电企业摆脱“靠天吃饭”的困境,平滑其年度经营成果。同时,通过稳定企业的年度收入,天气指数保险为企业融资提供了增信支持。这不仅有助于绿色能源企业的长期可持续发展,还通过保险的形式为实现“双碳”目标贡献了金融力量[4]。

光伏指数保险是一种基于太阳辐射总量的天气指数保险产品,旨在为光伏电站提供发电量不足的风险保障。

我国光伏指数保险虽然起步较晚,但发展迅速。2014年,安邦保险公司向保监会提交了新保险备案,主要承保因极端天气导致的电力输出不足所带来的损失。2016年,中国人寿财险与汉能发电公司合作,首次推出了“光伏+保险”新模式。尚德益家新能源公司与中华财险合作,推出了第一个综合性保险方案,涵盖了发电量损失险、机器设备损坏险、营业中断险、财产损失险以及公众责任险。2022年1月,中国保险行业发布首个光伏领域保险评估标准——《晶体硅光组件承保验厂评估规范》,完善了我国保险业在光伏组件方面的风险评估标准[5]。

然而,我国光伏保险产品仍面临着许多的问题,如在理赔过程中,企业需要提供相关的损失数据。然而保险公司往往缺乏客观的参考数据和第三方损失评估标准。一方面,光伏发电项目的运行数据往往

由项目方自行监测和报告, 保险公司难以获得第一手真实信息, 也难以区分正常损耗和异常损耗。这就给保险欺诈者提供可乘之机, 他们可以通过篡改数据、伪造证据、夸大损失等手段, 向保险公司索赔, 一定程度上导致保险公司的赔付额度与赔付率的上升。另一方面, 由于光伏保险市场还不够成熟, 缺乏统一的行业规范、专业的评估机构、完整的数据平台等, 导致保险公司在理赔过程中难以判断真假、定责定损、追究法律责任等, 也给保险欺诈者留下空间。

在其他保险领域, 保险欺诈已经有了深入的研究, 如车险欺诈一直是保险欺诈的主要领域。为了有效识别车险欺诈, 国内外许多学者将机器学习模型应用于车险欺诈检测技术, 并取得了显著效果。例如: Viaene 等人[6]、Hanafizadeh 等人[7]、Li 等人[8]分别使用贝叶斯模型、聚类模型、随机森林对车险欺诈进行了研究; He 等人[9]进一步研究了深度学习在此领域是否能有很好的应用; Majhi 等人[10]对混合模型进行了研究。赵桂芹等人[11]、汤俊等人[12]、王海巍等人[13]使用传统的机器学习模型对国内车险欺诈进行了研究; 闫春等人[14]、喻炜等人[15]、徐徐等人[16]从深度学习网络、混合模型的角度出发, 在车险欺诈检测任务上得到了较好的效果。

光伏指数保险欺诈不同于车险欺诈的地方在于光伏指数保险的数据是时间序列数据, 它具有周期性、趋势性和随机性等特征。而车险的数据既包括结构化数据, 也包括非结构化数据, 具有多样性、复杂性和稀疏性等特征。因此, 车险欺诈识别方法主要依赖于机器学习对这些结构化和非结构化的数据进行分析。

所以本文选择使用有别于常规车险欺诈所用到的深度学习模型, 对 2019 年新疆光电光伏数据集进行训练以识别在光伏指数保险中因谎报发电量而引起的保险欺诈问题。

2. 文献综述

光伏保险根据面向客户的不同, 有不同的险种, 包括面向组件制造商的巩固率保障险、面向电站建造商的职业责任险、面向已建成电站的财产险、面向家庭光伏的家庭太阳能光伏设备综合保险等[17]。例如, 财产一切险(BRI 保险)承保财产因自然灾害或意外事故造成的直接物质损失。其中, 德国 Conergy 公司提出用财产一切险减少施工及施工后造成的光伏电站损失; 在“华龙一号”项目运营期, 由中再产险担任执行机构的中国核保险共同体为项目提供约 130 亿元的核风险保障; 人保财险在南平、宁德和莆田等地推广家庭太阳能光伏设备综合保险, 对因火灾、暴雨、洪水、台风、泥石流、地面突然下降下沉等造成的设备损伤进行保险。

然而以上保险产品所承包的风险主要为自然灾害或意外导致的设备损坏、设备质量缺陷、光伏组件损害等, 虽然也有涉及因非意外与非自然灾害的光照不足引起的发电损失的保险产品。但由于缺乏光伏保险数据的积累以及有效的欺诈行为识别手段, 在理赔方面, 企业索赔时需要提供相关的损失数据, 但保险公司缺乏客观的参考数据与评估标准, 一定程度上提高了保险公司的赔付额度与赔付率。

太阳辐射量反映在一段时间内特定地区的客观光照强度, 光伏数据作为典型的时间序列数据, 记录了光伏发电量或者太阳辐射量等随时间变化的变量。因此构建光伏发电量或者太阳辐射量的时间维度上的相关模型可以在一定程度上为光伏保险欺诈识别提供数据支持。

分析时间序列数据的模型可以分为传统的统计模型和基于深度学习的模型。传统的统计模型主要包括自回归模型(AR)、移动平均模型(MA)、自回归移动平均模型(ARMA)、自回归整合移动平均模型(ARIMA)等。这些模型通常基于线性假设, 能够有效地捕捉时间序列数据中的趋势和季节性, 但是对于非线性和复杂的时间序列数据, 它们的表现可能不够理想。基于深度学习的模型主要分为以下三类:

卷积神经网络(CNN)是一种利用卷积和池化操作来提取图像特征的深度学习模型, 最初是为了解决计算机视觉中的图像识别问题而设计的[18]。Li 等人[19]在 2017 年提出了一种将时间序列数据转换为图

像数据的方法, 然后用 CNN 模型对图像数据进行特征提取和聚类, 再结合天气等外部因素, 来预测电力负荷的变化。

循环神经网络(RNN)是一种深度学习模型, 它能够学习时间序列数据的特征, 由 Jordan 于 1990 年首次提出[20]。Schuster 等人[21]在 1997 年将 RNN 扩展为双向循环神经网络(Bi-RNN), 使其能够同时利用前后两个方向的信息。长短期记忆网络(LSTM)是 Hochreiter 于 1997 年设计的一种改进的 RNN 模型, 它通过引入门控机制, 解决了 RNN 在训练过程中容易出现的梯度消失和梯度爆炸的问题[22]。门控循环单元(GRU)是 Cho 等人[23]在 2014 年基于 LSTM 模型的简化版本, 它减少了参数的数量, 降低了训练时间, 同时保持了与 LSTM 相当的准确性。

Transformer 是一种新型的深度学习框架, 它与传统的 CNN 或 RNN 结构有所不同, 由 Vaswani 等人[24]于 2017 年提出。Transformer 的核心机制是自注意力模块, 它可以动态地生成权重矩阵, 根据输入模式之间的相似性来计算输出。自注意力模块的参数数量较少, 计算量也较低, 因此它能够有效地建模长距离的依赖关系[25]。

这些模型通常能够处理非线性和高维的时间序列数据, 并且具有更强的泛化能力和灵活性。然而, 这些模型也存在一些问题。如 CNN 的预测精度与循环神经网络相比, 对处理步长较长的时序预测问题略显不足。尽管循环神经网络(RNN)能够捕捉并利用时间依赖关系进行预测, 但在长时间序列预测任务中表现不佳。此外, RNN 通常采用串行计算, 这一计算方法会导致训练过程中内存消耗巨大, 并且梯度消失和梯度爆炸问题始终未能彻底解决。Transformer 类模型需要对输入序列的每个元素进行两两比较, 计算它们之间的相关性。这样的操作需要大量的内存和时间, 尤其是当序列很长时。

2022 年 3 月 Google 研究团队和瑞士 AI 实验室 IDSIA 提出了一种新的架构, 称为 Block Recurrent Transformer (DeLesley Hutchins *et al.* Block Recurrent Transformers (March 2022))。Block Recurrent Transformer 模型通过引入循环单元和滑动自注意力机制解决了 Transformer 模型在处理序列很长时非常耗时和消耗内存的缺点。该模型需要对输入序列进行分块处理, 这可能会导致一些信息丢失或者边界效应。

综上, 本文选择在 LSTM 的基础上引入注意力机制并使用残差连接, 降低模型的复杂度和训练时间。

3. 模型构建

本文所用到的模型是在标准的 LSTM 模型基础上, 加入了注意力机制和残差连接。具体如下:

输入层后接一个双向 LSTM 层, 可以捕捉文本的前后上下文信息;

在 LSTM 层后接一个注意力层, 可以学习到不同特征的重要程度;

注意力输出后接一个残差连接, 将注意力输出与 LSTM 输出相加, 既包含了注意力信息, 也保留了 LSTM 层提取的特征;

最后接一个输出层做分类预测。结构图见图 1 所示。

LSTM 模型是一种改进的循环神经网络, 其由四个交互的子层分别是遗忘门、输入门、输出门和记忆单元组成, 基本原理见图 2 所示。

自注意力机制是一种用于捕捉输入向量之间的相关性的方法, 详细见图 3 所示。它可以让模型不仅关注当前的输入, 还能考虑其他位置的输入, 从而提高模型的表达能力和泛化能力。自注意力机制的基本思想是, 对于每个输入向量, 都要计算它与其他输入向量的相似度, 然后根据相似度对其他输入向量进行加权求和。具体步骤如下:

首先, 将序列中的每个元素转换为三个向量: 键(Key)、查询(Query)和值(Value)。这些向量通过不同的权重矩阵生成。

其次, 对于每个元素, 计算其与其他所有元素的键向量的点积, 得到注意力得分。

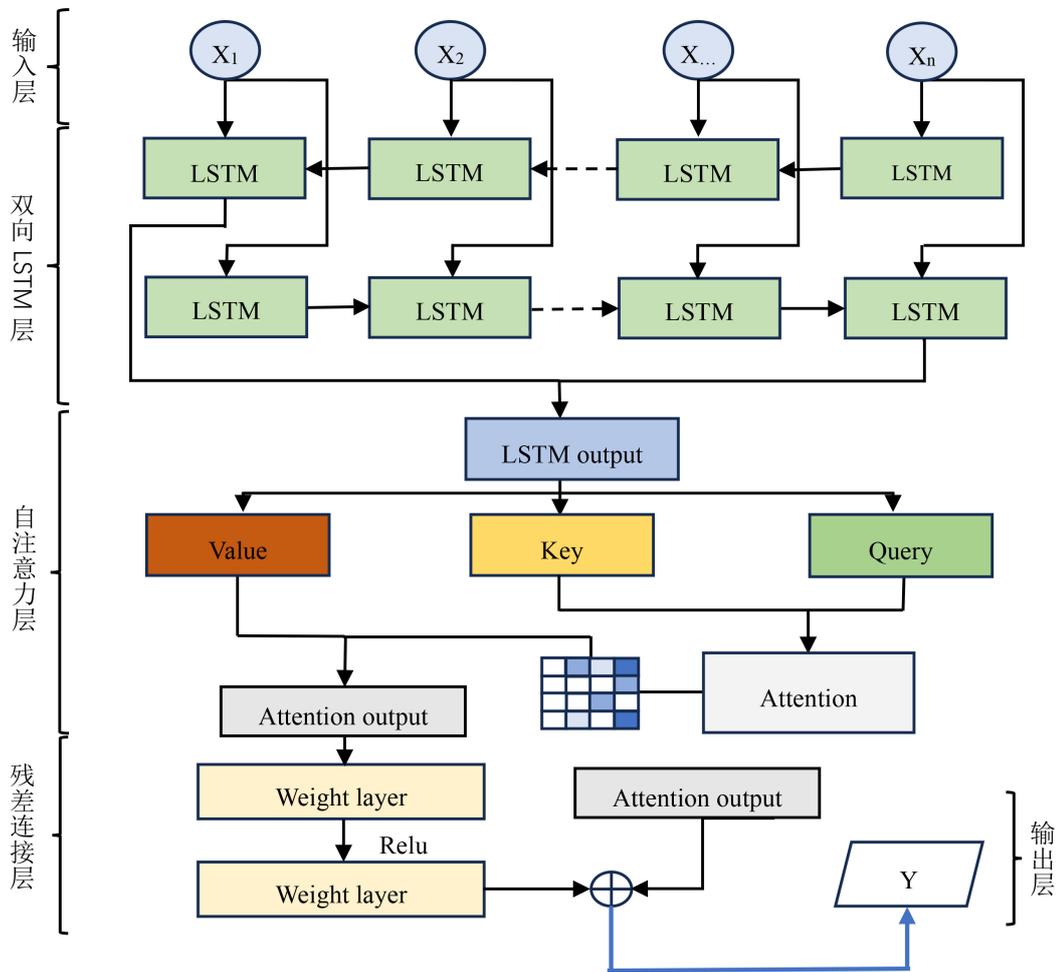


Figure 1. LSTM-AR flowchart

图 1. LSTM-AR 流程图

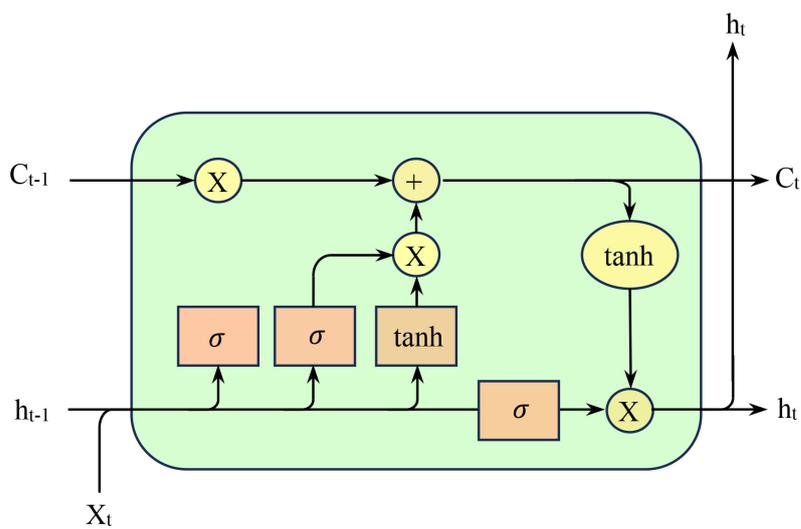


Figure 2. LSTM structure diagram

图 2. LSTM 结构图

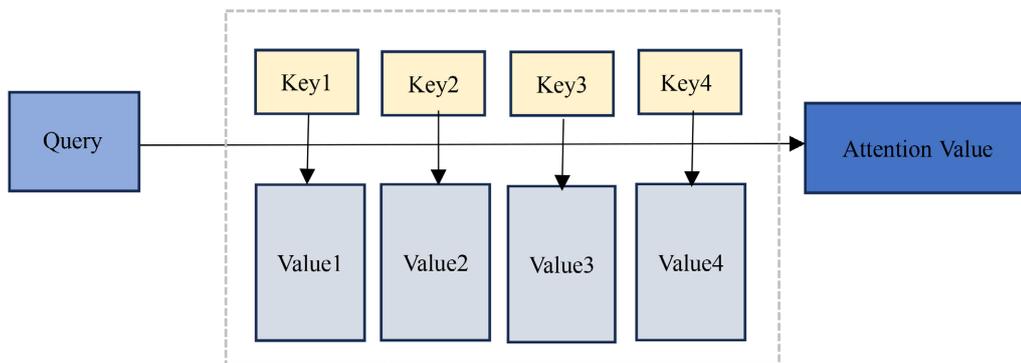


Figure 3. Self-attention mechanism diagram
图 3. 自注意力机制图

然后, 对于每个元素, 对其得分应用 softmax 函数, 使得它们的和为 1, 这样就可以将得分转换为权重。

最后, 将每个元素的值向量乘以其 softmax 权重, 并求和。这样就得到了一个对整个序列的加权表示, 其中更重要的元素会有更大的影响。

残差连接的基本思想是, 对于每个网络层或模块, 不仅保留它的原始输出, 还将其他的输入直接加到输出上。这样每个输出都包含了输入的信息, 但是不同层或模块的输入有不同的权重, 权重越大越重要。具体来说, 残差连接的计算公式如下:

$$y = x + F(x) \tag{1}$$

其中, x 是输入 y 是输出, $F(x)$ 是一个非线性变换函数, 可以包含卷积、激活、归一化等操作。

4. 实验设计

为验证所提方法的有效性, 本文使用 2019 年新疆风电的光伏数据, 对实际发电量进行学习预测。该数据集包括新疆风电 2019 年中每隔 15 分钟的光伏数据, 一共包含了 35,040 个观测时点和 8 个特征属性, 部分数据信息见表 1 所示:

Table 1. Photovoltaic data table of Xinjiang wind power in 2019
表 1. 2019 年新疆风电的光伏数据表

时间	2019-01-01 09:30:00	2019-01-01 09:45:00	2019-01-01 10:00:00	2019-01-01 10:15:00	2019-01-01 10:30:00
组件温度(°C)	-24.73	-23.5875	-19.5775	-15.115	-9.0225
温度(°)	-15.44	-15.064	-14.108	-14.491	-14.232
气压(hPa)	926.063	926.063	926.063	926.057	926.057
湿度(%)	62.197	60.738	57.275	57.702	57.983
总辐射量(W/m ²)	7.3	39.375	97.1	142.375	154.875
直接辐射(W/m ²)	6.57	35.4375	87.39	128.137	139.387
散射辐射(W/m ²)	9.46667	19.7067	44.6933	64.8533	75.1333
发电功率(mw)	0.94007	2.98933	5.21873	7.6602	9.50367

由于夜晚没有阳光的照射, 所以会存在大量的时间点出现总辐射量、直接辐射和散射辐射为 0 的情况, 为了避免因此对建模形成影响, 删除总辐射量、直接辐射和散射辐射为 0 的时间点的样本。最后光伏数据集包含了 17,044 个观测时间点。同时对数据进行划分, 将其中 80% 作为训练集, 剩余 20% 作为测试集。

本文采用了两种指标, 首先均方误差(MSE), 它是实际输出值和预测输出值之差的平方和的平均值, 它衡量了模型预测的准确性。

$$\text{MSE} = \frac{\text{SSE}}{n} = \frac{1}{n} \sum_{i=1}^m \omega_i (y_i - \hat{y}_i)^2 \quad (2)$$

$$\text{SSE} = \sum_{i=1}^m \omega_i (y_i - \hat{y}_i)^2 \quad (3)$$

其中, y_i 是真实数据, \hat{y}_i 是拟合的数据, n 是样本个数, $\omega_i > 0$ 。

其次是平均绝对误差(MAE), 它是实际输出值和预测输出值之差的绝对值的平均值。他衡量了模型预测的稳健性

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |h(x_i) - y_i| \quad (4)$$

其中, y_i 是真实数据, $h(x_i)$ 是拟合的数据, m 是样本个数。

5. 实验结果与分析

出于性能的比较, 用最小二乘法对实际发电率进行面板回归进行对比, 其中用最小二乘法所得到的 MSE 和 MAE 分别为 66.0903 和 5.9163; 用本文所引用的模型所得到的 MSE 和 MAE 分别是 0.0613 和 0.2151。从结果来看, 本文所用到的模型在预测实际发电率时, 明显优于最小二乘法的回归预测, 其 MSE 和 MAE 都比最小二乘法低了几个数量级。这说明本文所引用的模型可以更准确地拟合数据中的非线性和复杂特征, 而用最小二乘法则可能存在欠拟合和过拟合的问题。

本文所引用的模型之所以能优于最小二乘法, 可能有以下几个原因:

首先, 该模型使用了 LSTM 层来处理面板数据中的时间序列特征, 可以捕捉数据中长期依赖和动态变化, 而最小二乘法则忽略了时间维度的信息, 只考虑了截面维度的信息。

其次, 该模型使用了注意力机制来学习不同特征的重要程度, 可以给予更多的权重给对预测有贡献的特征, 从而减少对无关或噪声特征的影响, 而用最小二乘法做回归分析则假设所有特征都有相同的影响力, 可能会导致特征选择不合理或多重共线性的问题。

最后该模型使用了残差连接来将 LSTM 层的输出与输入相加, 可以保留 LSTM 层提取的特征, 同时也包含了注意力信息, 而用最小二乘法做回归分析只是使用了线性层来输出预测值, 可能会丢失一些有用信息或增加一些冗余的信息。

6. 结论

本文针对光伏指数保险欺诈的识别问题, 提出了一种基于深度学习的模型, 结合 LSTM、注意力机制和残差连接, 来处理时间序列数据, 并对 2019 年新疆光电光伏数据集进行训练和测试。实验结果表明, 本文所提出的模型在预测实际发电率方面, 明显优于最小二乘法的回归预测, 其 MSE 和 MAE 都比最小二乘法低了几个数量级。本文所提出的模型可以有效地识别光伏指数保险中因谎报发电量而引起的保险欺诈问题, 为光伏保险业务的规范化和发展提供了一种有效的技术手段。

本文的主要贡献有以下几点:

1) 本文综合考虑了时间序列数据的特征和复杂性, 设计了一种结合 LSTM、注意力机制和残差连接的深度学习模型, 能够有效地捕捉数据中的长期依赖、动态变化和非线性特征。

2) 本文在 2019 年新疆光电光伏数据集上进行了实验验证, 结果表明本文所提出的模型在预测实际发电率方面, 明显优于最小二乘法的回归预测, 其 MSE 和 MAE 都比最小二乘法低了几个数量级。

3) 本文所提出的模型可以有效地识别光伏指数保险中因谎报发电量而引起的保险欺诈问题, 为光伏保险业务的规范化和发展提供了一种有效的技术手段。

本文还存在以下不足之处:

1) 本文只使用了一个数据集进行实验验证, 未考虑不同地区、不同季节、不同天气条件下的数据特征差异。

2) 本文只使用了 MSE 和 MAE 作为评价指标, 未考虑其他更能反映欺诈检测性能的指标, 如准确率、召回率、F1 值等。

3) 本文未对比其他基于深度学习的模型, 如 CNN、RNN、Transformer 等, 未能充分展示本文所提出模型的优势和局限性。为了进一步提高本文的研究水平。

未来的工作可以从以下几个方面进行:

1) 使用更多的数据集进行实验验证, 考察本文所提出模型在不同场景下的泛化能力和稳定性。

2) 使用更多的评价指标进行性能分析, 考察本文所提出模型在欺诈检测任务上的准确性和效率。

3) 对比其他基于深度学习的模型, 分析本文所提出模型的优势和局限性, 探索更适合光伏指数保险欺诈识别问题的模型结构和参数设置。

参考文献

- [1] 李嘉. 论光伏发电天气指数保险的可行性[J]. 保险职业学院学报, 2017, 31(6): 30-33.
- [2] 余永清. 光伏发电企业面临的财务风险探讨[J]. 产业创新研究, 2022(2): 80-82.
- [3] 黄梦哲. 光伏发电企业融资现状及创新路径[J]. 商业会计, 2023(3): 114-116.
- [4] 李源. 浅谈国内绿色能源企业推广天气指数保险的现存问题和重要意义[J]. 科技与金融, 2023(5): 83-88.
- [5] 刘金霞, 谢美玲. 关于光伏保险发展的问题分析与对策建议[J]. 黑龙江金融, 2022(7): 70-72.
- [6] Viaene, S., Dedene, G. and Derrig, R. (2005) Auto Claim Fraud Detection Using Bayesian Learning Neural Networks. *Expert Systems with Applications*, **29**, 653-666. <https://doi.org/10.1016/j.eswa.2005.04.030>
- [7] Hanafizadeh, P. and Paydar, N.R. (2013) A Data Mining Model for Risk Assessment and Customer Segmentation in the Insurance Industry. *International Journal of Strategic Decision Sciences*, **4**, 52-78. <https://doi.org/10.4018/jsds.2013010104>
- [8] Li, Y., Yan, C., Liu, W. and Li, M. (2018) A Principle Component Analysis-Based Random Forest with the Potential Nearest Neighbor Method for Automobile Insurance Fraud Identification. *Applied Soft Computing*, **70**, 1000-1009. <https://doi.org/10.1016/j.asoc.2017.07.027>
- [9] He, X. and Chua, T. (2017) Neural Factorization Machines for Sparse Predictive Analytics. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tokyo, 7-11 August 2017, 355-364. <https://doi.org/10.1145/3077136.3080777>
- [10] Majhi, S.K. (2019) Fuzzy Clustering Algorithm Based on Modified Whale Optimization Algorithm for Automobile Insurance Fraud Detection. *Evolutionary Intelligence*, **14**, 35-46. <https://doi.org/10.1007/s12065-019-00260-3>
- [11] 赵桂芹, 吴洪. 汽车保险市场中存在道德风险吗?——来自动态续保数据的分析[J]. 金融研究, 2010(6): 175-188.
- [12] 汤俊, 莫依雯. 基于数据挖掘技术的车险反欺诈系统构建[J]. 上海保险, 2013(11): 39-42.
- [13] 王海巍. 我国险企运营中道德风险甄别问题研究——以大数据 Hadoop 聚类分析技术为视角[J]. 保险研究, 2016(2): 59-67.
- [14] 闫春, 李亚琪, 孙海棠. 基于蚁群算法优化随机森林模型的汽车保险欺诈识别研究[J]. 保险研究, 2017(6): 114-

127.

- [15] 喻炜, 冯根福, 张文珺. 机动车辆保险欺诈检测系统及团伙识别研究[J]. 保险研究, 2017(2): 63-73.
- [16] 徐徐, 王正祥, 王牧群. 基于深度学习技术的机动车辆保险欺诈识别模型与实证研究[J]. 上海保险, 2019(8): 53-58.
- [17] 沈燕鸿. 绿色保险发展的国内外实践[J]. 金融纵横, 2023(1): 53-59.
- [18] Goodfellow, I., Bengio, Y., Courville, A., *et al.* (2016) Deep Learning. MIT Press, 326-366.
- [19] Li, L., Ota, K. and Dong, M. (2017) Everything Is Image: CNN-Based Short-Term Electrical Load Forecasting for Smart Grid. 2017 14th International Symposium on Pervasive Systems, Algorithms and Networks & 2017 11th International Conference on Frontier of Computer Science and Technology & 2017 Third International Symposium of Creative Computing (ISPAN-FCST-ISCC), Exeter, 21-23 June 2017, 344-351. <https://doi.org/10.1109/ispan-fcst-iscc.2017.78>
- [20] Goodfellow, I., Bengio, Y., Courville, A., *et al.* (2016) Deep Learning. MIT Press, 363-405.
- [21] Schuster, M. and Paliwal, K.K. (1997) Bidirectional Recurrent Neural Networks. *IEEE Transactions on Signal Processing*, **45**, 2673-2681. <https://doi.org/10.1109/78.650093>
- [22] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [23] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., *et al.* (2014) Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, 25-29 October 2014, 1724-1734. <https://doi.org/10.3115/v1/d14-1179>
- [24] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [25] Li, S., Jin, X., Xuan, Y., *et al.* (2019) Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Vancouver, 8-14 December 2019, 5243-5253.