

# 基于关联规则——LSTM神经网络模型的“能源双碳数智”平台满意度分析

黄聘博, 王滢琦, 胡靖辉, 钱佳瑜

绍兴文理学院数理信息学院, 浙江 绍兴

收稿日期: 2023年1月4日; 录用日期: 2023年2月1日; 发布日期: 2023年2月8日

## 摘要

“能源双碳数智”平台是能源新形势政策背景下, 为政府机构、电网公司、企业、能源服务商等市场主要客户群体提供实现全方位碳监测与计量、多维度碳分析与评估、综合性碳资产管理、个性化碳账户与减排等服务的一体化智慧管控数字平台。该平台助力碳达峰、碳中和进程。在此背景下, 本文首先基于ACSI模型建立平台满意度评价指标体系, 借助关联规则——LSTM神经网络模型得出评价数据, 对评价结果进行全面、详尽、客观的分析, 为政府部门及企业提供客观、科学的数据分析资料, 就如何推进“能源双碳数智”平台运行, 提出一些具有推广价值的可行性经验。

## 关键词

“能源双碳数智”平台, ACSI模型, 关联规则, LSTM神经网络

## Satisfaction Analysis of “Energy Dual Carbon Digital Intelligence” Platform Based on Association Rule—LSTM Neural Network Model

Pinbo Huang, Yanqi Wang, Jinhui Hu, Jiayu Qian

School of Mathematical Information, Shaoxing University, Shaoxing Zhejiang

Received: Jan. 4<sup>th</sup>, 2023; accepted: Feb. 1<sup>st</sup>, 2023; published: Feb. 8<sup>th</sup>, 2023

## Abstract

Under the background of the new energy situation and policies, the “Energy Dual Carbon Digital

文章引用: 黄聘博, 王滢琦, 胡靖辉, 钱佳瑜. 基于关联规则——LSTM神经网络模型的“能源双碳数智”平台满意度分析[J]. 计算机科学与应用, 2023, 13(2): 172-179. DOI: 10.12677/csa.2023.132018

**Intelligence” platform provides government agencies, power grid companies, enterprises, energy service providers and other major customer groups in the market with an integrated intelligent management and control digital platform to achieve all-round carbon monitoring and measurement, multi-dimensional carbon analysis and assessment, comprehensive carbon asset management, personalized carbon account and emission reduction and other services. The platform facilitates the process of carbon peaking and carbon neutrality. In this context, this paper first establishes a platform satisfaction evaluation index system based on the ACSI model, obtains the evaluation data with the help of the association rule—LSTM neural network model, conducts a comprehensive, detailed and objective analysis of the evaluation results, provides objective and scientific data analysis data for government departments and enterprises, and puts forward some feasible experience with promotion value on how to promote the operation of the “energy dual carbon digital intelligence” platform.**

## Keywords

**“Energy Dual Carbon Digital Intelligence” Platform, ACSI Model, Association Rules, LSTM Neural Network**

Copyright © 2023 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

“能源双碳数智”平台是双碳目标下的全新产物，是助力碳达峰、碳中和目标的有利保障，但是一切新事物的诞生必定需要经历实践的证明，其可行性和实施效果需要得到考证。因此，本文以“能源双碳数智”平台为中心点，调研杭州市重点耗能企业工作人员对“能源双碳数智”平台的满意度，通过关联规则——LSTM神经网络挖掘调研数据背后隐藏的信息(关联规则具有较好的数据挖掘能力，使输入神经网络中的数据更加合理，而LSTM神经网络利用高精度的算法得到最终训练结果)，以此找到目前“能源双碳数智”平台发展出现的困境和存在的挑战，最终得到推动“能源双碳数智”平台运转且稳健发展的突破点。

## 2. 研究背景

当前，全球气候的变化对人类的生产、生活带来了极大的不利影响，包括中国在内的很多国家将“碳达峰、碳中和”上升为国家战略。2020年，作为世界上最大的发展中国家和最大的煤炭消费国，中国基于推动实现可持续发展的内在要求和构建人类命运共同体的责任担当，宣布了“3060”碳达峰、碳中和的目标愿景[1]。提出双碳目标能够有效减少碳排放，减轻温室效应的影响，是保护环境的重要途径，也是可持续发展的必然要求。但是碳达峰和碳中和的任务重大而艰巨，且受到我国现状的制约，我们国家经济体量巨大，发展速度飞快，能源需求量大，发展起步较晚，且相关新能源结构体系尚未构建，对煤炭的需求量较大。我们国家正处于城镇化和工业化的非常时期，产业和能源结构对煤炭的依赖度高，短期内难以实现双碳目标。

中央财经委员会第九次会议，习近平主席特别强调，要把碳达峰和碳中和列为我国生态文明建设规划布局，建立健全绿色低碳技术评价、交易体系和技术创新综合服务平台[2]。与其他西方发达国家相比较，中国实现“双碳”目标正陷入碳排放总量大、碳减排时间紧、阻碍因素多等不容忽视的问题。从某

种角度来看,这对于整个能源行业来说,产业结构、安全运营、新能源消纳能力等都提出了非常大的挑战,但也为以电力为核心的能源信息行业带来了前所未有的机遇。为顺应国家能源革命和“双碳”战略构想,主动响应建立健全绿色低碳技术评价、交易体系和技术创新服务平台的总体安排,在新形势新政策背景下,“能源双碳数智”平台的研究与开发启动。

“能源双碳数智”平台以区域内智慧能源综合服务平台为基础,以能源大数据应用为依托,从政府部门、企业公司、交易市场的角度,着眼于碳资产的全流程管理,开展“能源双碳数智”平台建设工作。平台借助汇聚海量能源信息,结合大数据、人工智能等先进技术,研究碳核算方法、计算方法和预测方法,逐步形成标准算法模型,构建起碳监测体系、碳评估体系和碳排放控制体系,构建能源“双碳”数字智能产品,为政府部门提供“碳监测”辅助能源和碳管理,为企业提供“碳分析”指导节能减碳,为交易市场提供“碳认证”优化资源配置,促进社会经济高质量发展绿色高位[3]。

在此背景下,本文通过基于关联规则——LSTM 神经网络模型对“能源双碳数智”平台在杭州市的建设及发展现状做出评价,为政府机构和企业提供客观、科学的数据分析资料 and 实践经验,推动“能源双碳数智”平台的运行发展。

### 3. 关联规则——神经网络组合模型

#### 3.1. 关联规则介绍

关联规则最早由 R. Agrawal 教授于 1993 年提出,能够衡量特定事物之间的相关性。能够借助冲突法来发现大量数据库数据中项目集相互间的关系。如果两个或更多个属性之间存在关联,则能够参照结合其他属性值预测另一个属性。可以用形象的语言将数据库比作超市,不同类别的项目集是各种商品,寻找关联项目集就是挖掘客户购物篮中较常同时出现的商品。关联法则挖掘问题,能够分为两个子问题:

- 1) 找到事物数据库文件中所有大于等于用户选定的最小支持度的数据项集;
- 2) 借助频繁项集生成所需要的关联规则,根据用户设置的最小置信度进行取舍,最终得到强关联规则。

最有名的关联规则挖掘算法是 Apriori 算法,它主要借助于向下封闭属性:假如一个项集是频繁项目集,那么它的非空子集必然是频繁项目集。即频繁集的子集也肯定是频繁集。例如,如果{A, B, C}是 3 项频繁集,则其子集{A, B}, {B, C}, {A, C}也肯定是 2 项的频繁集。它先生成 1-频繁项目集,再借助 1-频繁项目集生成 2-频繁项目集,然后再根据 2-频繁项目集生成 3-频繁项目集……以此类推,直至最后生成所有最频繁项比较集,然后继续从最频繁项比较集中发现规则符合条件的关联项。在代码操作中,Apriori 算法主要采用逐层迭代探索的步骤,简单、直接、容易实现,没有复杂的理论计算。

频繁项集的常见评估标准包括支持度、置信度和提升度。

支持度是数据集中若干关联数据出现的次数占总数据集的比例,或者几个数据关联出现的概率。如果要分析两个数据 X 和 Y 的相关性,则对应的支持度是:

$$Support(X,Y) = P(XY) = \frac{number(XY)}{num(AllSamples)}$$

以此类推,如果我们要分析三个数据 X、Y、Z 的相关性,则对应的支持度为:

$$Support(X,Y,Z) = P(XYZ) = \frac{number(XYZ)}{num(AllSamples)}$$

一般来说,支持度高的数据不一定构成频繁项集,但是支持度太低的数据一定不构成频繁项集。

从置信度中可以看出一个数据出现后另一个数据出现的概率,也就是数据的条件概率。如果要分析两个数据 X 和 Y 的相关性,X 对 Y 的置信度为:

$$Confidence(X \leftarrow Y) = P(X|Y) = \frac{P(XY)}{P(Y)}$$

同样，依此类推，可以得到多个数据的关联置信度，例如，对三个数据  $X$ ， $Y$ ， $Z$ ，则  $X$  对于  $Y$  和  $Z$  的置信度为：

$$Confidence(X \leftarrow YZ) = P(X|YZ) = \frac{P(XYZ)}{P(YZ)}$$

提升度表示在包含  $Y$  的条件下，同时含有  $X$  的概率与  $X$  整体发生的概率的比值，即：

$$Lift(X \leftarrow Y) = \frac{P(X|Y)}{P(X)} = \frac{Confidence(X \leftarrow Y)}{P(X)}$$

提升度反映了  $X$  和  $Y$  间的关联关系，如果提升度大于 1，则  $X \leftarrow Y$  是一个有效的强关联规则。若提升度小于或等于 1，则  $X \leftarrow Y$  为无效强关联规则。在特殊情况下，如果  $X$  和  $Y$  是独立的，则  $Lift(X \leftarrow Y) = 1$ ，因为  $P(X|Y) = P(X)$ 。

通常，要在数据集中选择频繁的数据集，需要自定义评估标准。最常用的评估标准是自定义的支持度，或自定义支持度和置信度的组合。本文主要利用关联规则建立数据与数据的关联关系，优化输入神经网络中的数据，提升结果精度。

### 3.2. LSTM 神经网络基本原理

神经网络是模仿动物神经网络的行为特征并进行分布式并行信息处理的算法数学模型。网络依靠系统的复杂性，通过调节大量内部节点之间的关系来处理信息[4]。它可以对所有类型的数据进行训练，并产生符合用户需求的训练结果。与传统的数据分析和拟合方法不同，它可以学习各种类型的非线性数据，从而产生尽可能接近实际的结果。LSTM 神经网络是神经网络的加强分支之一，其网络特点为能够利用记忆细胞状态将梯度消失和梯度爆炸等神经网络训练中的常见问题做出较好地控制[5]。它具有图 1 所示的类似链式结构，它由四个神经网络层构成，并且以一种特殊的方式进行交互。

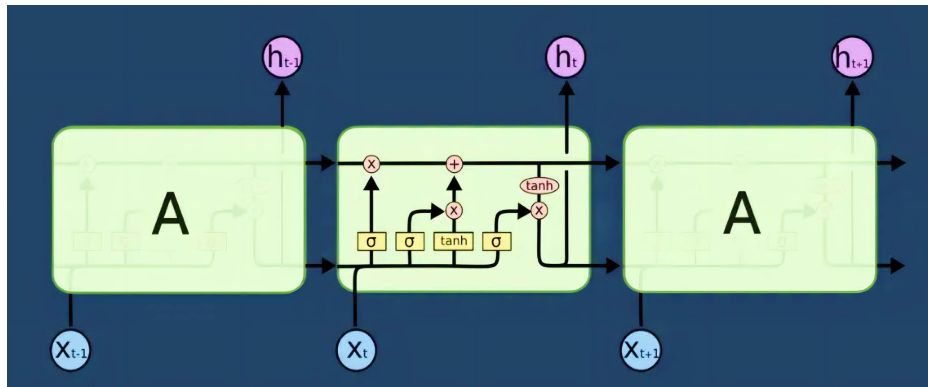


Figure 1. Chain structure  
图 1. 链式结构

在确定输入层和输出层神经元节点数后，使用梯度下降法在进行反向传播的同时，实时对其模型参数进行更新，训练需要预测的模型，直到使得预测模型的精度需求达到满足给定条件为止，训练完成。LSTM 神经网络结构如图 2 所示。

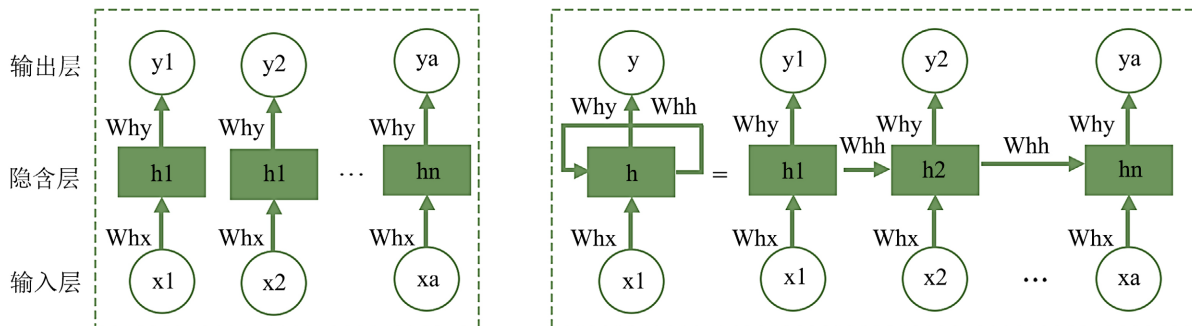


Figure 2. LSTM neural network structure  
图 2. LSTM 神经网络结构

### 3.3. 关联规则——神经网络组合模型

#### 3.3.1. 合理性

评价模型通常会选择多个影响因子作为评价的指标，但是各因子与评价结果的关联性大小各异。在设计上述平台满意度评价模型时，因各影响因子与预测结果无法较好地进行了区分，并且很难判定各影响因子的选择恰当与否。所以通过关联规则对影响因素进行分类和选择，分类的结果越好，此种评价模式的准确性越高，通常为得到更加准确的评价结果，所选择的每个影响因子都应是有利于最终分类结果准确性的，也就是与分类结果存在较强的关联性。

#### 3.3.2. 模型构成

模型训练样本的质量一定程度上影响着预测精度，提高样本质量对提升模型精度有积极影响。利用神经网络模型搭建平台满意度评价模型时，训练样本即平台满意度影响因子的实测数据，应用关联规则对其进行挖掘，分析各因子的重要性，并将选取重要性较高的因子作为输入值传递到神经网络模型中进行学习训练。不仅可以在提高模型精度，还可以增强学习器的泛化能力，减少训练耗时。

组合模型是在神经网络模型的基础上，增加一个关联规则模型，借助关联规则来提高神经网络模型输入数据的质量，进而减小误差、提高预测精度和效率等。模型结构如图 3 所示。

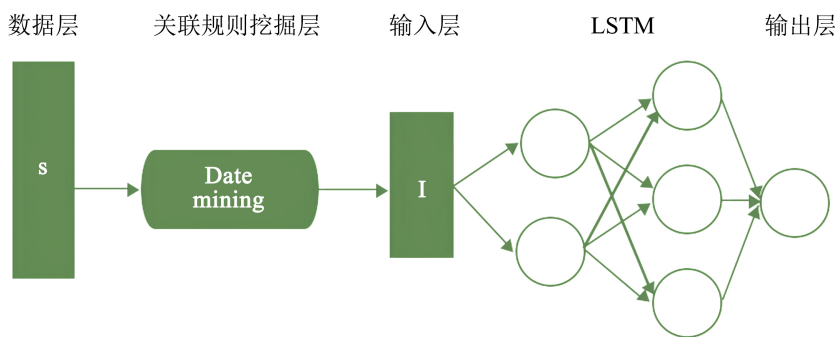


Figure 3. Structure of the combined model  
图 3. 组合模型结构

分为五个部分：

第一层是数据层(Data)。将平台满意度影响因子历史数据整合在一起的原始数据集。

第二层是关联规则挖掘层(Association rules)。在大量调研数据的基础上，进行数据挖掘找出这些看似不相关的指标间的关联性，从而确定平台满意度的关键因子，建立研究平台满意度评价指标体系，为研

究平台满意度评价模型的输入参数选择提供有力支持。

第三层是输入层(Input)。将第二层中确定的平台满意度评价指标作为此网络的输入数据,此数据便于开展下一层的训练学习,在实际的训练过程中输入数据往往需要先进行归一化处理。

第四层是隐含层(Hidden)。对输入数据进行学习,训练网络。这一层采用 LSTM 神经网络。

第五层是输出层(Output)。主要负责将第四层内学习到的详细信息转成为输出。然后将预测结果与真实样本数据进行对比,同时将误差信息进行反向传递到隐含层中。接着在隐含层内通过不断矫正模型参数,减少模型误差,然后将更新的参数用于下一次的学习训练。此过程通过往返训练迭代,使预测值与真实值之间的训练误差不断减小,最终达到理想范围[6]。

## 4. 基于关联规则——LSTM 神经网络模型的“能源双碳数智”平台满意度评价

### 4.1. 基于 ACSI 模型的平台满意度评价指标体系

ACSI 模型是一种以顾客为基础,用作评价与改良组织绩效的测评体系,是目前具备最完整体系,且应用效果最好的一个顾客满意度理论模型。可操作性指标体系是对满意度进行科学度量、分析、评价的基础,结合专家访谈的意见,本文对此模型进行适当改造,从 ACSI 模型的 6 个结构变量细分出的各变量中选取合适的评价指标,确定平台满意度评价指标体系。如表 1 所示。

**Table 1.** Platform satisfaction evaluation index system

**表 1.** 平台满意度评价指标体系

一级指标	二级指标	三级指标
平台满意度指数	顾客预期	能效评价较好
		能耗预算降低
		能耗预警减少
	感知质量	能源数据动态跟踪
		平台服务贴心
		数据核查全面
	感知价值	管理效率高
		平台流程简化
		数据核查及时
	顾客满意	政府政策支持
		平台使用便利
		及时了解用能状况
顾客抱怨	碳检测数据精度	
	平台收费价格	
	运转流畅度	
顾客忠诚	愿意使用“能源双碳数智”平台	
	愿意支持“能源双碳数智”平台	
	愿意推荐“能源双碳数智”平台	

## 4.2. 评价过程

### 1) 数据处理

本文总共选择了 21 个影响因素作为评价因子，从实际问题中提取的输入数据，一般都需要预先进行归一化处理，以此来降低神经网络目标的误差。经赋值归一化等处理之后可以直接作为输入进行运算。处理后的数据一般均被转换为[0, 1]、[0.1, 0.9]或[-1, 1]之间的数。数据归一化能够消除维度不同的数据之间的数量级差，消除奇异值，从而保障网络运行时的收敛性，并减小网络输出误差。数据归一化处理方法很多，例如最大最小法、函数处理法、使用 Matlab 软件自行编程等。

本文采用最大、最小法进行归一化处理。连续属性值可以直接进行归一化，而离散属性值则需要用数字代表各个属性的不同类型，即将离散属性值数值化后，再进行归一化处理，作为网络输入。函数形式如下所示：

$$x_k = (x_k - x_{\min}) / (x_{\max} - x_{\min})$$

将所有数据归一化处理构成评价属性集，作为输入训练网络。

由于关联规则是针对项与项之间关系开展挖掘分析的，所以明确挖掘中涉及到的要素后，需要将各因子实测数据转化为适合关联规则挖掘的数据形式。对问卷结果统一数据标准时，选择一位数作为标识符，分别标号为 1, 2, 3, 4, 5。统一影响因子数据标准时，分别设定 12 种影响因子的标号为 1、2、3……20、21。

### 2) 关联规则——LSTM 神经网络模型进行满意度评价

使用关联规则确定输入指标的训练集，输入向量为 6 个指标数据，期望输出值为满意度评分。

比较使用关联规则前后训练结果，在关联规则优化下，迭代次数减少，训练用时减少，同时有了更小的误差，根据理论，结果也更符合实际，由此可见加入关联规则的合理性。下图 4 为使用关联规则前后 ROC 曲线及 AUC 值：

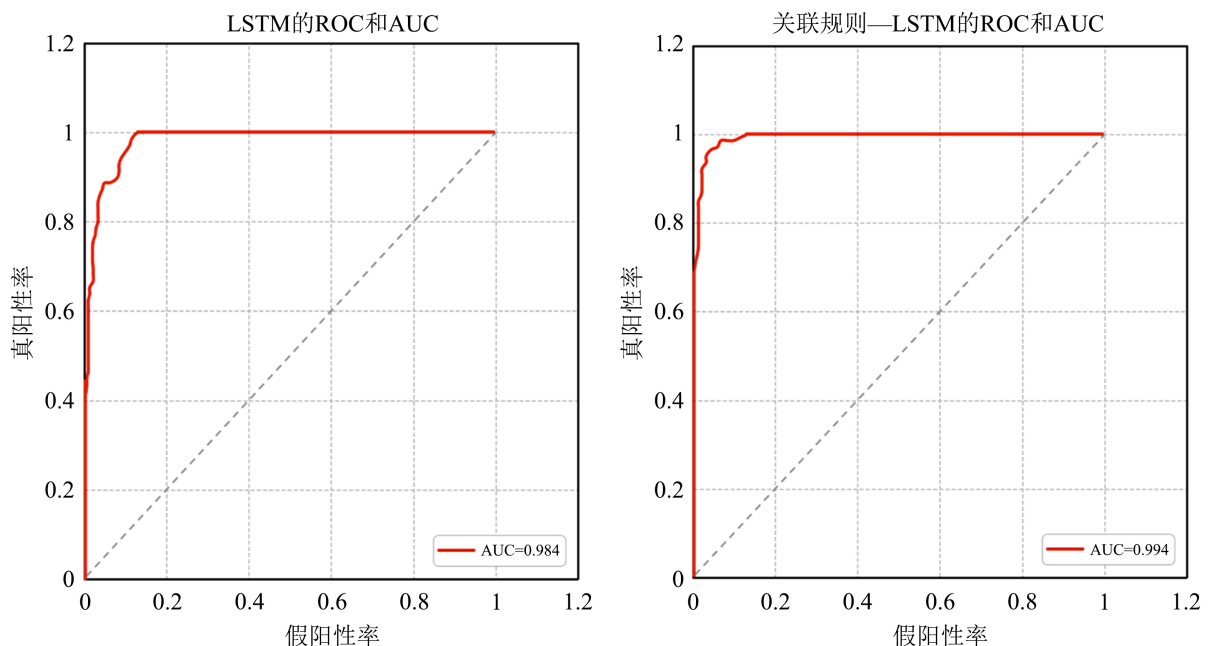


Figure 4. ROC curves and AUC values before and after using association rules

图 4. 使用关联规则前后 ROC 曲线及 AUC 值

从中可以看出加入关联规则后 AUC 值提升, 更加表明组合模型分类器性能表现突出。

使用关联规则——LSTM 神经网络模型, 将各指标数据带入网络得到评价结果, 评价结果如所示。

由下表 2 可以看出, 六个因子满意度总体平均指数均位于较高水平。其中顾客忠诚因子的满意度最高, 达到 4.058, 表明顾客对“能源双碳数智”平台的认同感较强。感知质量因子的满意度其次, 为 3.873, 说明顾客对现阶段“能源双碳数智”平台的服务项目设计较为满意, 这可能会成为顾客选择使用“能源双碳数智”平台的关键影响因素。得分最低的是顾客预期因子, 说明顾客对该平台的硬件设施配备认同感不强, 可能会成为顾客接受“能源双碳数智”平台的最大障碍。按降序排列依次为顾客忠诚因子、感知质量因子、顾客满意因子、感知价值因子、顾客抱怨因子和顾客预期因子。

**Table 2.** Satisfaction rating scale

**表 2.** 满意度评分表

分类	顾客预期	感知质量	感知价值	顾客满意	顾客抱怨	顾客忠诚	满意度评分
分数	3.350	3.873	3.548	3.575	3.417	4.058	3.733
评价指标	III	II	II	II	III	I	II

## 5. 结论

我国双碳目标提出以来, “能源双碳数智”平台试点工作逐步展开, 但有关平台建设情况、实施效果、发展现状的调查并不多, 本文尝试使用关联规则组合 LSTM 神经网络对“能源双碳数智”平台的满意度进行评价, 该方法结合关联规则的优势使输入数据间的联系更加精密, 同样利用神经网络的高精度、高准确性拟合数据, 快速得到满意度评分, 确保了评价的准确性、便捷性和信服力。满意度评价结果一方面可以反映“能源双碳数智”平台当前基本建设情况, 另一方面也是用户对平台使用意愿、使用满意情况的体现。

从分析结果看, 各因子满意度说明“能源双碳数智”平台的建设及发展现状良好, 但要想进一步推广落实, 还存在以下问题: 首先, “政府主导, 企业迎合”的模式仍不明确, 政策的作用不明显这也导致一部分企业不愿参与其中, 因此平台政策支持力度仍需加强; 其次, 安全可信的“双碳”数据体系需要进一步完善, 只有数据合理, 平台公信力才得以提高; 最后, 大数据技术、物联网技术、AI 技术的运用也是平台发展的重要动力。基于本文分析, “能源双碳数智”平台具备较好的前景和发展机遇, 可谓机遇与挑战并存, 因此, 只要将各个问题解决, “能源双碳数智”平台就能步入新的征程, 双碳目标的实现指日可待。

## 参考文献

- [1] 陈赞, 刘昌维, 潘智俊, 王晓慧, 李永庆, 李琳. 新形势下智慧“能源 + 双碳”服务平台的建设与应用[J]. 供用电, 2022, 39(2): 15-21.
- [2] 周吉, 许自豪, 刘熙, 李杰玲. 江西构建“数智控碳”平台体系的思考与建议[J]. 中国国情国力, 2022(6): 40-44.
- [3] 胡熠, 靳曙畅. 数字技术助力“双碳”目标实现: 理论机制与实践路径[J]. 财会月刊, 2022(6): 111-118.
- [4] 田珂. 基于聚类关联规则神经网络组合算法初速预测[J/OL]. 兵工学报, 1-12. <http://www.co-journal.com/CN/10.12382/bgxb.2021.0687>, 2022-12-29.
- [5] 肖岚. 基于 LSTM 网络的经营性租赁固定资产实物期权定价预测研究[J]. 武汉船舶职业技术学院学报, 2022, 21(4): 125-127+132.
- [6] 杨敬武. 基于关联规则与神经网络的滑坡危险性评价模型及应用研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2020.