A COMPREHENSIVE STUDY OF MACHINE LEARNING AND DEEP LEARNING ALGORITHMS FOR PRODUCT CATEGORIZATION

Meenal N. Pande
Department of Computer Science, Degree College of Physical Education, HVPM, Amravati, Maharashtra, India*meenal.pande@gmail.com*

Dr. S. E. Yedey
P.G. Department of Computer Science and Technology Degree College of Physical Education, HVPM, Amaravati, Maharashtra, India
sanjayeyedey@gmail.com

Abstract
Product categorization is an important task in e-commerce, which helps to organize and retrieve products efficiently from large inventories. With the growth of e-commerce platforms, the complexity and scale of product catalogues have increased, creating challenges for traditional classification methods. This review considers the model of machine learning and deep learning algorithms for automating and improving the product categorization process. Traditional methods may struggle with the extreme diversity of products, but AI-based approaches, such as CNNs, RNNs, and transformers, tend to provide much better accuracy and efficiency in terms of categorization. The review discusses the advantages and disadvantages of different algorithms and identifies some major challenges, such as computational resource demands, quality of data, and model interpretability. Advances notwithstanding, deep learning models face other issues related to their "black-box" nature, limiting transparency and trust in real-world applications. The paper presents a discussion on the large size and quality of high-quality datasets and provides countermeasures against scarcity and imbalances. Additionally, it discusses the integration of explainable AI (XAI) to face such concerns and improve the transparency of ML and DL models. Finally, the study points out where future researches may focus such as algorithm's scalability improvement; ethical concerns handling in AI-Driven systems and developing robust model for practical deployment in real word applications. Improvement of these is likely to impact significantly the efficacy and usability efficiency of product classification in the future e-commerce ventures.
Keywords: *Artificial Intelligence, Classification, Deep Neural Networks, E-commerce Categorization, Natural Language Processing.*

I. Introduction

In the digital age today, e-commerce platforms as well as online marketplaces are growing at an unbelievable rate, which leads to massive product inventories as well as diverse customer requirements[1]. The proper categorization of products is essential to enhance searchability, boost user experience, then perform efficient inventory management[2]. This challenge has led to the adoption of automation using ML and deep learning DLtechniques, providing scalable and intelligent solutions for product categorization tasks[3].

For ages, machine learning has been a staple of classification problems[4]. Decision trees, random forests, then SVMs stand some of the algorithms which have proven to be efficient in processing structured data as well as text-based product information[5]. Despite this, these models are effective for small-to-medium-sized datasets and produce intelligible results [6],[7].

Advanced extensions of ML, deep learning introduces the neural networks through which automatic extraction of the hierarchical features from raw input data takes place[8]. Deep learning algorithms have revolutionized product categorization with such architectures as CNNs (for images) and RNNs (for any sequential input, such as product description)[9],[10]. Techniques such as transformers and vision transformers have further advanced capabilities in DL[11].

The integration of ML and DL algorithms with virtualization technologies has further enhanced applicability[12]. Virtualized environments further make such models scalable for deployment on cloud infrastructures, where businesses can run large datasets in real time[13]. This means that tools like Kubernetes or Docker make resource management very efficient so that ML and DL algorithms can easily be executed for categorizing products[14]. These pre-trained models and APIs also become available through cloud-based platforms like AWS, Azure, and Google Cloud, making it easier for small and medium enterprises to introduce automated categorization systems[15].

The present paper reviews the development and use of ML and DL in the broader area of product classification, discussing the strengths and weaknesses of different approaches as well as their robustness and flexibility to different types of data and deployment frameworks.

The need to categorize has grown exponentially with the widespread growth in e-commerce websites and the digital marketplace in recent years[16]. Categorization boosts product discoverability and subsequently optimizes your inventory, personalization algorithm, and recommendation system; however, for a long period, these were manual rules-based, relying on existing taxonomies and human inference[17]. These methods were effective only for small-scale datasets. They were labor-intensive and time-consuming, and full of inconsistencies, especially for large, dynamic inventories[18].

As datasets became complex, DL appeared as a revolutionary advance in this field, able to handle multimodal data, like images, text, and metadata[19]. More recently, other innovations such as transformers and vision transformers (ViTs) have further pushed the envelope in automated categorization to handle scalability, understanding of context, and even real-time performance[22].

II. Research Gap

It is true that many studies have brought great progress in applying machine learning and deep learning algorithms to product categorization in e-commerce. Nevertheless, there are challenges toward that end-to-end approach having models generalizing effectively across various datasets and domains and without the need for much retraining. Most algorithms highly depend on a quality labeled dataset and often are incapable of doing effectively in a noisy real world environment. Another factor is the lack of interpretability in DL models referred to as "black-box" problems, creating a challenge to implement them in dynamic business settings. All these challenges require new approaches that improve scalability and interpretability, with reduced computational demand.

III. Research Questions

RQ1. How do AI models improve the accuracy and scalability of product categorization?

RQ2. Which techniques can make the requirement of large labeled datasets for training in AI minimal?

RQ3. How can deep learning models be made more transparent and usable in practice?

IV. Inclusion And Exclusion Criteria

Fig. 1 shows, from the database searches, 150 records were identified as shown in the figure. Duplicates were removed and the remaining 130 records were screened and gave 120 screened records. Out of the records, 20 were excluded; 10 of them because they were not relevant to load balancing, 5 were not relevant to cloud computing, and 5 were review papers. The remaining articles were 100 that were further assessed for eligibility. In the screening process, 12 papers with full-text articles were excluded because of inadequate reporting of data. Thus, 50 papers were included in the detailed survey.

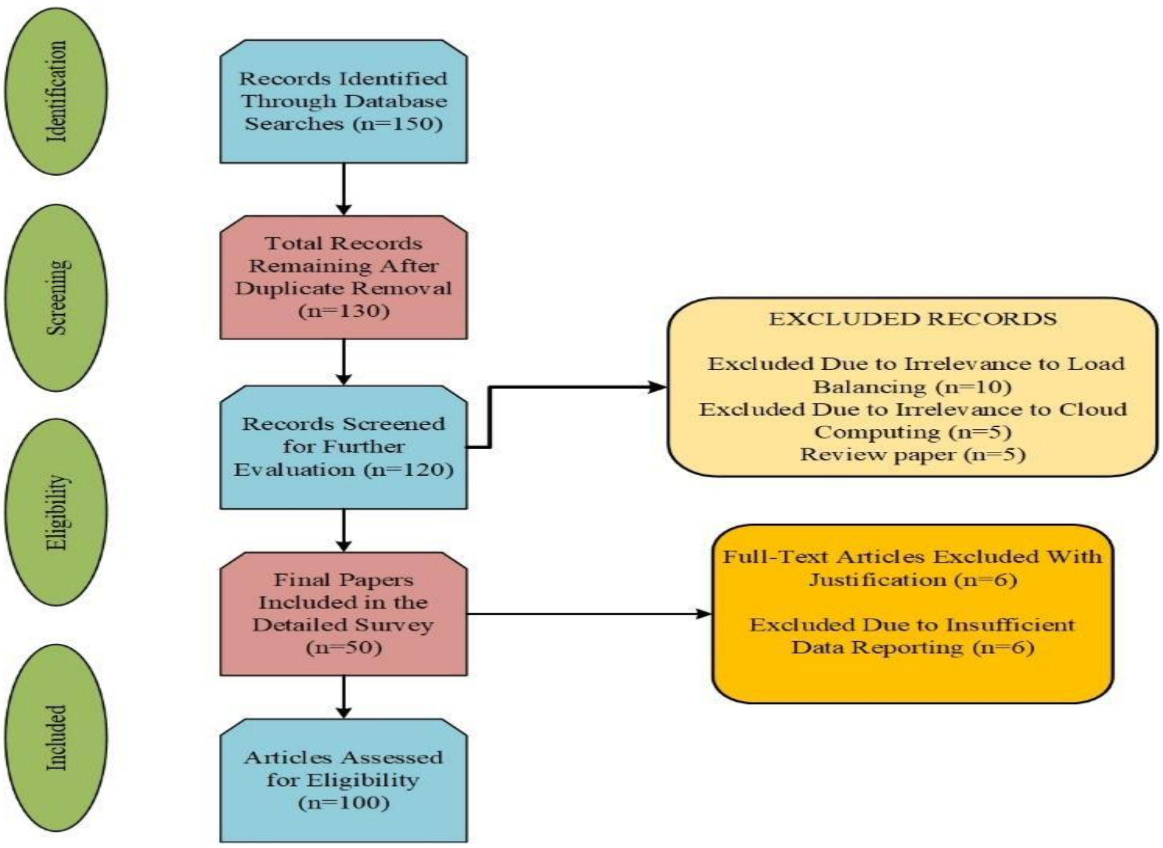

Fig. 1. Systematic Review Process

Ml And Dl Models In Product Categorization

Table I summarizes major research on the product categorization of e-commerce, integrating the use of techniques in ML, DL, and NLP, to progress accuracy in classification

TABLE I. Comparative Overview Of Ml And Dl Models In Product Categorization

Research Title	**Proposed Method**	**Advantages**	**Limitations**	**Proposed Solutions**
Robust Product Classification with Instance-Dependent Noise[23]	Data de-noising algorithm for noisy product titles	Improves classification accuracy by removing noise	Sensitive to complex or ambiguous noise patterns; effectiveness	Further studies on model adaptability to datasets with varying data quality
Hybrid Deep Learning Approach for Product Categorization in E-commerce, 2022[24].	Integration of LSTM with CNN for predicting product categories	Achieves 97% F1 score; combines LSTM's sequential learning with CNN's feature extraction	May not generalize across product categories; requires significant computational resources	Optimize model to handle heterogeneous data and improve scalability
Improving Product Categorization in E-commerce Using NLP and Machine Learning[25].	Combines NLP techniques with traditional ML algorithms	Improves classification precision by integrating word embeddings and text preprocessing	Struggles with unclear or inconsistent product descriptions; scalability issues with large datasets	Develop more robust algorithms and infrastructure to manage dynamic e-commerce data
lti-level Product Category Prediction through Text Classification, 2022[26].	Uses LSTM and BERT for text classification	Improves categorization precision, particularly with a Brazilian dataset	Ambiguity in product descriptions; scalability issues	Explore scalability solutions and improve handling of ambiguous descriptions
Comparative Analysis of Machine and Deep Learning Techniques for Text Classification[27].	BiLSTM for text classification	Achieves 98.5% accuracy, outperforming Naive Bayes, SVM, and Gradient Boosting	High-quality data preprocessing is critical; poor data standardization reduces performance	Focus on high-quality preprocessing and more robust feature extraction methods
A Machine Learning Approach for Product Matching and Categorization[28].	Combines ML algorithms with deep learning for betterproduct categorization and match	Improved categorization and matching accuracy	Dependent on quality and consistency of product data; errors due to noisy or inconsistent data	Refine preprocessing techniques and optimize models for diverse product categories
Applying Machine Learning for Automatic Product Categorization[29].	Comparison of LR, MNB, and SVM for product categorization	SVM performs best on high-dimensional data	SVM is susceptible to misclassification due to variations in input data; requires significant feature engineering	Improve preprocessing algorithms and automated feature selection methods
Deep Categorization Network for E-commerce Product Classification [30].	Introduces DeepCN, a deep learning-based model for product classification	High accuracy on large e-commerce platforms	Performance affected by training data quality and diversity; requires significant computational resources	Explore more efficient architectures and training schemes to overcome resource constraints
Hybrid Deep Learning Approach for Low-Quality Data in Product Categorization [31]	Combines deep learning with data augmentation techniques	Improves classification accuracy for noisy or sparse data	Dependent on the quality of the augmentation process, which can lead to overfitting	Develop more complex data augmentation strategies and combine them with other machine learning techniques

The reviewed studies emphasize some of the significant advancements in product categorization through learning techniques. Ravi et al. [23] have proposed a data de-noising algorithm that can improve the accuracy of classification but struggles with ambiguous noise. Sharma et al. [24].We have achieved a 97% F1 score by integrating LSTM with CNN but the scalability remains a challenge. Almeida et al. [26] deployed LSTM and BERT to enhance the categorization of Brazilian datasets, encountering problems with ambiguous descriptions and scalability. It is more important to enhance preprocessing for optimized models in several ecommerce applications.

V. Ml And Dl Models With Data Virtulization In Product Categorization

Table II represents a comparative overview of different studies on product categorization research studies, highlighting proposed methods, main findings, and limitations, in addition to their suggestions for improvement.

TABLE II. Comparative Overview Of Ml And Dl Models With Data Virtulization In Product Categorization

Research Title	Proposed Method	Advantages	Limitations	Proposed Solutions
Product Categorization Using Deep and Machine Learning Techniques[32].	CNNs and LSTMs for product categorization, comparing with SVM and decision trees.	High accuracy (95%) in product classification, especially with deep learning models like CNNs.	Model training is difficult due to the need for large labeled datasets and high computational power. Inconsistencies in product descriptions and noisy data affect model performance.	Improve data preprocessing techniques and explore model architectures that are less computationally expensive.
Understanding the Product Classification Methodologies for Cloud Computing[33].	Regression analysis to assess factors influencing investment decisions.	Highlights important factors like financial literacy and risk tolerance that influence investment decisions.	Small sample size, self-reported biases.	Expand the study to a more heterogeneous group of investors and use qualitative methods for more in-depth analysis.
Data Management Challenges in Deep Learning Systems: A Multiple Case Study[34]	By DL systems analyzes data management challenges in industry environments.	Provides insights into the limitations of current data management solutions for DL systems.	Lack of cost-effective and scalable solutions for real-world deployments, and lack of necessary infrastructure or skills.	Develop cost-effective and scalable data management solutions, and provide infrastructure support to bridge the gap in real-world deployments.
IoT-Based Efficient Data Visualization Framework for Business Intelligence in Corporate Finance[35]	IoT-EDVF for data visualization and business intelligence management.	Improves data quality and business intelligence decision-making with significant improvements in revenue analysis and performance metrics.	Relies on IoT infrastructure, which may not be available everywhere, and requires advanced technical knowledge for implementation.	Ensure scalability of the IoT infrastructure for different business environments and streamline the implementation process to make it more accessible.
AI-Assisted Prediction of Environmental Impacts to Support Life Cycle Assessment Practitioners[36].	AI using NLP-processed data and Random Forest for predicting environmental impacts.	High prediction accuracy (R^2: 68%-81%) for various environmental impact categories.	Accuracy heavily depends on the quality and volume of evaluating data. The model cannot replace a comprehensive life cycle assessment (LCA).	Develop methods to improve the quality and quantity of training data, and integrate AI predictions with detailed LCA methods for a more comprehensive approach.

Immersive Virtual Shopping Experiences in the Retail Metaverse: Insights on Consumer-Driven E-Commerce and Blockchain-Based Assets[37].	Integration of DL algorithms, NN, and data-driven decision-making in virtual retail.	Enhances customization and user engagement in the retail metaverse, improving operational efficiency and customer intelligence.	Results may not generalize across all retail environments due to diversity, and reliance on existing literature limits generalization.	Conduct empirical validation in broader retail settings and investigate how the findings apply in various contexts.

The reviewed works present significant advancement in all areas. Khan et al. [32]claimed the accuracy of CNNs and LSTMs in product classification, but these are noisy and computationally intensive. Ahmed et al. [34] further mentioned that a scalable solution should be used for the management of data for the deep learning systems. Smith et al. [36] achieved accuracy of high degree in environmental impact prediction, but relied more on data quality. Roberts et al. [37]improved user engagement into virtual retail, but this was not generalizable across contexts of different retails, and such kinds of results so obtained require improvements in scalability, better preprocessing, and wider validation.

Related work on categorization of the product discusses several AI techniques (ML, DL) algorithms. Most of the papers are published in 2023, as shown in Fig. 1.

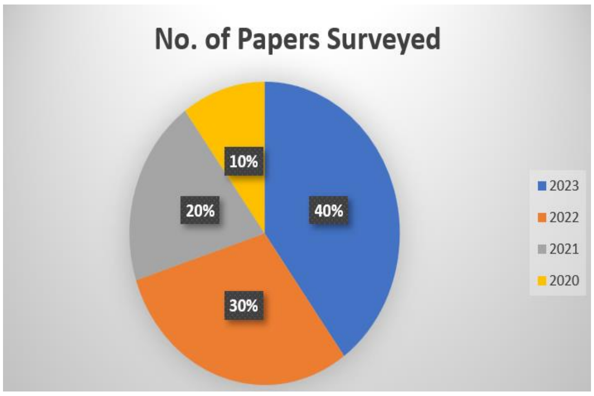

Fig. 2. Papers collected for the Review

VI. Discussion And Conclusion

A. Key Benefits of AI-Driven Product Categorization

1) Improved Accuracy: The promise shown by AI and machine learning algorithms, particularly in the area of deep learning models, is particularly convincing, as this encompasses both CNNs and transformers [42]. Through such models, which can deeply analyze complex datasets to find patterns not readily found through other systems, greater accuracy is expected in classifications, furthering better customer experiences [43].

2) Automation and Operational Efficiency: Using AI to automate the process of categorizing products can help businesses minimize manual efforts while speeding up tasks such as sorting and recommending products [44]. This process not only saves time but also reduces operational costs, thus streamlining and scaling product management.

3) Scalability in Large-Scale E-commerce: Since ML and DL models are naturally scalable, they come in handy with huge data sets of e-commerce sites[45]. Handling huge product catalogues spread over various regions means the scalability without sacrificing performance can be maintained effectively by businesses.

*4) Personalization opportunities:*It combines AI with customer behavior data, and personalized product recommendations can drive engagement and increase sales [46]. The customized shopping experience can result in a higher conversion rate through tailored suggestions based on user preferences.

5) Innovative Hybrid Approaches: Businesses will be able to develop stronger models for product categorization busing deep learning and domain-specific features.

B. Challenges of AI in Product Categorization

1) Computational Requirements: Deep learning models are computationally intensive and can significantly hinder small to mid-sized companies. Due to this, such models do require access to high performance infrastructure and may demand even increased operational costs with a requirement for high technical expertise [47].

2) Dependence on High Quality Data: This has led to reliance on the existence of large, high-quality datasets to enhance the effectiveness of AI models [48]. These datasets, however, often cost much, time, and human resources to obtain and annotate. Besides, biases within the data might hinder generalization in novel scenarios or product categories.

References

1. S. Mascarenhas and M. Agarwal, “A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image

Classification," in *2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON)*, IEEE, 2021, pp. 96–99. Accessed: Dec. 31, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9687944/

2. D. Tchoffa, N. Figay, P. Ghodous, H. Panetto, and A. El Mhamedi, "Alignment of the product lifecycle management federated interoperability framework with internet of things and virtual manufacturing," *Computers in Industry*, vol. 130, p. 103466, 2021.
3. G. Taranto-Vera, P. Galindo-Villardón, J. Merchán-Sánchez-Jara, J. Salazar-Pozo, A. Moreno-Salazar, and V. Salazar-Villalva, "Algorithms and software for data mining and machine learning: a critical comparative view from a systematic review of the literature," *J Supercomput*, vol. 77, no. 10, pp. 11481–11513, Oct. 2021, doi: 10.1007/s11227-021-03708-5.
4. D. Soni and N. Kumar, "Machine learning techniques in emerging cloud computing integrated paradigms: A survey and taxonomy," *Journal of Network and Computer Applications*, vol. 205, p. 103419, 2022.
5. S. A. Shetty, T. Padmashree, B. M. Sagar, and N. K. Cauvery, "Performance Analysis on Machine Learning Algorithms with Deep Learning Model for Crop Yield Prediction," in *Data Intelligence and Cognitive Informatics*, I. Jeena Jacob, S. Kolandapalayam Shanmugam, S. Piramuthu, and P. Falkowski-Gilski, Eds., in Algorithms for Intelligent Systems. , Singapore: Springer Singapore, 2021, pp. 739–750. doi: 10.1007/978-981-15-8530-2_58.
6. V. Pham, D. C. Weindorf, and T. Dang, "Soil profile analysis using interactive visualizations, machine learning, and deep learning," *Computers and Electronics in Agriculture*, vol. 191, p. 106539, 2021.
7. R. R. Petrus, P. J. do Amaral Sobral, C. C. Tadini, and C. B. Gonçalves, "The NOVA classification system: A critical perspective in food science," *Trends in Food Science & Technology*, vol. 116, pp. 603–608, 2021.
8. S. V. Mahadevkar *et al.*, "A review on machine learning styles in computer vision—techniques and future directions," *Ieee Access*, vol. 10, pp. 107293–107329, 2022.
9. "GitHub - sohansai/product-categorization: E-commerce Product Categorization Model Using Deep Learning," GitHub. Accessed: Dec. 31, 2024. [Online]. Available: https://github.com/sohansai/product-categorization
10. S. Ertekin, "UNDERSTANDING THE PRODUCT CLASSIFICATION METHODOLOGIES FOR CLOUD COMPUTING," *Copyright 2023 by Institute for Global Business Research, Nashville, TN, USA*, p. 24, 2023.
11. M. Bertolini, D. Mezzogori, M. Neroni, and F. Zammori, "Machine Learning for industrial applications: A comprehensive literature review," *Expert Systems with Applications*, vol. 175, p. 114820, 2021.
12. F. A. Alenizi, S. Abbasi, A. H. Mohammed, and A. M. Rahmani, "The artificial intelligence technologies in Industry 4.0: A taxonomy, approaches, and future directions," *Computers & Industrial Engineering*, p. 109662, 2023.
13. [G. W. Scurati, S. Y. Kwok, F. Ferrise, and M. Bertoni, "A STUDY ON THE POTENTIAL OF GAME BASED LEARNING FOR SUSTAINABILITY EDUCATION," *Proc. Des. Soc.*, vol. 3, pp. 415–424, Jul. 2023, doi: 10.1017/pds.2023.42.
14. M. Nagy, G. Lăzăroiu, and K. Valaskova, "Machine Intelligence and Autonomous Robotic Technologies in the Corporate Context of SMEs: Deep Learning and Virtual Simulation Algorithms, Cyber-Physical Production Networks, and Industry 4.0-Based Manufacturing Systems," *Applied Sciences*, vol. 13, no. 3, p. 1681, Jan. 2023, doi: 10.3390/app13031681.
15. D. Saha and A. Manickavasagan, "Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review," *Current Research in Food Science*, vol. 4, pp. 28–44, 2021.
16. S. Rawat, A. Rawat, D. Kumar, and A. S. Sabitha, "Application of machine learning and data visualization techniques for decision support in the insurance sector," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100012, 2021.
17. P. A. Rauschnabel, "Augmented reality is eating the real-world! The substitution of physical products by holograms," *International Journal of Information Management*, vol. 57, p. 102279, 2021.
18. A. Tyvelä, "Version control and automated filtering of simulation data to support design for manufacturability in virtual new product introduction," Master's Thesis, A. Tyvelä, 2021. Accessed: Dec. 31, 2024. [Online]. Available: https://oulurepo.oulu.fi/handle/10024/19145

19. K. Liu, X. Hu, H. Zhou, L. Tong, W. D. Widanage, and J. Marco, “Feature analyses and modeling of lithium-ion battery manufacturing based on random forest classification,” *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 6, pp. 2944–2955, 2021.
20. T. Kaluarachchi, A. Reis, and S. Nanayakkara, “A review of recent deep learning approaches in human-centered machine learning,” *Sensors*, vol. 21, no. 7, p. 2514, 2021.
21. G. Lazaroiu, A. Androniceanu, I. Grecu, G. Grecu, and O. Neguriță, “Artificial intelligence-based decision-making algorithms, Internet of Things sensing networks, and sustainable cyber-physical management systems in big data-driven cognitive manufacturing,” *Oeconomia Copernicana*, vol. 13, no. 4, pp. 1047–1080, 2022.
22. H. W. Kim *et al.*, “NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products,” *J. Nat. Prod.*, vol. 84, no. 11, pp. 2795–2807, Nov. 2021, doi: 10.1021/acs.jnatprod.1c00399.
23. F. Sun, D.-B. Luh, Y. Zhao, and Y. Sun, “Product Classification With the Motivation of Target Consumers by Deep Learning,” *IEEE Access*, vol. 10, pp. 62258–62267, 2022, doi: 10.1109/ACCESS.2022.3181624.
24. M. Gupta, R. Kumar, C. Ved, and S. Taneja, “Hybrid deep learning approach for product categorization in e-commerce,” presented at the INTERNATIONAL CONFERENCE ON INTELLIGENT AND SMART COMPUTATION (ICIASC-2023), Mohali, India, 2024, p. 040018. doi: 10.1063/5.0198666.
25. N. S. Satwik Reddy, V. V. A. Rohith, P. S. Abhiram, M. D. S. R. Saran, and S. Rebecca, “Enhancing Product Categorization in E-commerce using NLP and Machine Learning,” in *2024 International Conference on Inventive Computation Technologies (ICICT)*, Lalitpur, Nepal: IEEE, Apr. 2024, pp. 1–6. doi: 10.1109/ICICT60155.2024.10544665.
26. W. F. Maia *et al.*, “Multi-level Product Category Prediction through Text Classification,” Mar. 03, 2024, *arXiv*: arXiv:2403.01638. doi: 10.48550/arXiv.2403.01638.
27. D. S. Gochhait, “Comparative Analysis of Machine and Deep Learning Techniques for Text Classification with Emphasis on Data Preprocessing,” *Qeios*, May 2024, doi: 10.32388/XHC9J1.
28. P. Ristoski, P. Petrovski, P. Mika, and H. Paulheim, “A machine learning approach for product matching and categorization: Use case: Enriching product ads with semantic structured data,” *SW*, vol. 9, no. 5, pp. 707–728, Aug. 2018, doi: 10.3233/SW-180300.
29. A. Roberson, “Applying Machine Learning for Automatic Product Categorization,” *Journal of Official Statistics*, vol. 37, no. 2, pp. 395–410, Jun. 2021, doi: 10.2478/jos-2021-0017.
30. J.-W. Ha, H. Pyo, and J. Kim, “Large-Scale Item Categorization in e-Commerce Using Multiple Recurrent Neural Networks,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2022, pp. 107–115. doi: 10.1145/2939672.2939678.
31. P. Mavaie, L. Holder, and M. K. Skinner, “Hybrid deep learning approach to improve classification of low-volume high-dimensional data,” *BMC Bioinformatics*, vol. 24, no. 1, p. 419, Nov. 2023, doi: 10.1186/s12859-023-05557-w.
32. Md Abu Sufian Mozumder *et al.*, “Enhancing Customer Satisfaction Analysis Using Advanced Machine Learning Techniques in Fintech Industry,” *JCSTS*, vol. 6, no. 3, pp. 35–41, Aug. 2024, doi: 10.32996/jcsts.2024.6.3.4.
33. F. Ahmad and Z. Rahimi, “23 PUBLICATIONS 44 CITATIONS SEE PROFILE,” vol. 7, no. 1, 2023.
34. A. R. Munappy, J. Bosch, H. H. Olsson, A. Arpteg, and B. Brinne, “Data management for production quality deep learning models: Challenges and solutions,” *Journal of Systems and Software*, vol. 191, p. 111359, Sep. 2022, doi: 10.1016/j.jss.2022.111359.
35. C. Shao, Y. Yang, S. Juneja, and T. GSeetharam, “IoT data visualization for business intelligence in corporate finance,” *Information Processing & Management*, vol. 59, no. 1, p. 102736, Jan. 2022, doi: 10.1016/j.ipm.2021.102736.
36. A. Koyamparambath, N. Adibi, C. Szablewski, S. A. Adibi, and G. Sonnemann, “Implementing Artificial Intelligence Techniques to Predict Environmental Impacts: Case of Construction Products,” *Sustainability*, vol. 14, no. 6, p. 3699, Mar. 2022, doi: 10.3390/su14063699.
37. T. Jenkins, “Immersive Virtual Shopping Experiences in the Retail Metaverse: Consumer-driven E-Commerce, Blockchain-based Digital Assets, and Data Visualization Tools,” *Linguistic and Philosophical Investigations*, no. 21, pp. 154–169, 2022.

38. S. U. Hassan, J. Ahamed, and K. Ahmad, “Analytics of machine learning-based algorithms for text classification,” *Sustainable Operations and Computers*, vol. 3, pp. 238–248, 2022, doi: 10.1016/j.susoc.2022.03.001.
39. S. Jiang, J. Hu, C. L. Magee, and J. Luo, “Deep Learning for Technical Document Classification,” *IEEE Trans. Eng. Manage.*, vol. 71, pp. 1163–1179, 2024, doi: 10.1109/TEM.2022.3152216.
40. H. Chen, L. Wu, J. Chen, W. Lu, and J. Ding, “A comparative study of automated legal text classification using random forests and deep learning,” *Information Processing & Management*, vol. 59, no. 2, p. 102798, Mar. 2022, doi: 10.1016/j.ipm.2021.102798.
41. Omri Suissa and M. K. S. Uddin, “A REVIEW OF UTILIZING NATURAL LANGUAGE PROCESSING AND AI FOR ADVANCED DATA VISUALIZATION IN REAL-TIME ANALYTICS,” *GMJ*, vol. 1, no. 4, pp. 34–49, Jul. 2024, doi: 10.62304/ijmisds.v1i04.185.
42. T. Horak, P. Strelec, M. Kebisek, P. Tanuska, and A. Vaclavova, “Data Integration from Heterogeneous Control Levels for the Purposes of Analysis within Industry 4.0 Concept,” *Sensors*, vol. 22, no. 24, p. 9860, 2022.
43. M. Andronie *et al.*, “Big data management algorithms, deep learning-based object detection technologies, and geospatial simulation and sensor fusion tools in the internet of robotic things,” *ISPRS International Journal of Geo-Information*, vol. 12, no. 2, p. 35, 2023.
44. Q. Dai, “Designing an Accounting Information Management System Using Big Data and Cloud Technology,” *Scientific Programming*, vol. 2022, pp. 1–11, Aug. 2022, doi: 10.1155/2022/7931328.
45. H. Nguyen and D. Khatwani, “Robust Product Classification with Instance-Dependent Noise,” in *Proceedings of The Fifth Workshop on e-Commerce and NLP (ECNLP 5)*, Dublin, Ireland: Association for Computational Linguistics, 2022, pp. 171–180. doi: 10.18653/v1/2022.ecnlp-1.20.
46. L. Hua, “Smart Product Marketing Strategy in a Cloud Service Wireless Network Based on SWOT Analysis,” *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–11, Jun. 2022, doi: 10.1155/2022/7539860.
47. A. Bisquert, A. Hmimou, J. L. Berral, A. Gutierrez-Torre, and O. Romero, “HealthMesh: An Architectural Framework for Federated Healthcare Data Management,” 2022.
48. F. Khatami, P. De Bernardi, Š. Vilamová, E. Cagno, and F. Ricciardi, “Innovation ecosystem for smart product: empirical quantification of its key dimensions in SMEs of 21 European countries,” *Rev Manag Sci*, vol. 18, no. 9, pp. 2723–2748, Sep. 2024, doi: 10.1007/s11846-024-00763-z.