International Journal of Medical Toxicology & Forensic Medicine
2026, Volume 16, E52602

Review Article

A Personalized Context-Aware Recommendation System using Natural Language Processing for Drug Supply Chain Management in Smart City Environments

Gali Nageswara Rao[1*], Madhavarao Divya shree[2], Arunachalam Elaiyaraja[3], Bhanupriya Periaswamy[4], Mattaparthi Bala Prabhakar[5], Sandeep Gupta[6]

1. *Department of Information Technology, Aditya Institute of Technology and Management, Tekkali, Srikakulam, India.*
2. *Department of ECE, St. Joseph's College of Engineering, Chennai, Tamil Nadu, India*
3. *Department of Master of Business Administration, Panimalar Engineering College, Chennai, Tamil Nadu, India*
4. *Department of ECE, Saveetha Engineering College, Thandalam, India*
5. *Department of Mathematics, Aditya University, Surampalem, Kakinada, Andhra Pradesh, India*
6. *Department of Electrical Engineering, Graphic Era Deemed to be University, Dehradun (Uttarakhand), India.*

Citation Rao GN, Shree MD, Elaiyaraja A, Priya PB, Prabhakar MB, Gupta S. A Personalized Context-Aware Recommendation System using Natural Language Processing for Drug Supply Chain Management in Smart City Environments. International Journal of Medical Toxicology and Forensic Medicine. 2026; 16:E52602.

https://doi.org/ 10.22037/ijmtfm.v16.52602

Article info:

Received: 26 June, 2026

Accepted: 03 July, 2026

Published: 08 August, 2026

Keywords:

Drug recommendation system, Natural language processing, Deep learning, Gannet optimization algorithm

ABSTRACT

Background: Personalized Context-Aware Drug Recommendation Systems leverage machine learning (ML) techniques and patient data to provide tailored medication suggestions based on medical history, demographic information, and individual health profiles. By analyzing electronic health records, real-time physiological data, and genomic information, the system provides timely recommendations. This manuscript presents a personalized context-aware drug recommendation system using the optimal deep learning (PCADRS-ODL) method. The proposed technique aims to assist medical professionals in making effective treatment decisions.

Methods: Initially, the input data undergoes preprocessing. Then, a BERT-based word embedding approach is employed for feature extraction. Next, the PCADRS-ODL system utilizes a Double Attention-Convolutional Neural Network with a Bidirectional Gated Recurrent Unit framework. Furthermore, the parameters of the DAC-BiGRU model are optimized using the Gannet Optimization Algorithm.

Results: Experimental results on a benchmark dataset demonstrate that the proposed method achieves superior performance over existing methods.

Conclusion: Overall, the proposed PCADRS-ODL method provides an accurate and effective personalized drug recommendation framework, demonstrating its potential to support clinical decision-making and improve patient-specific treatment outcomes.

*** Corresponding Authors:**
Gali Nageswara Rao, PhD
Address: *Department of Information Technology, Aditya Institute of Technology and Management, Tekkali, Srikakulam, India.*
E-mail: *gnraoaitam@gmail.com*

Introduction

Health-related recommender systems can create a primary analysis, forecast disease progression, and generate suitable suggestions as per the health condition of patients. A recommender system is a method that delivers a list of proposals by forecasting selections so that consumers can rapidly discover what they need in a huge quantity of data [1]. In recent years, machine learning (ML) has significantly enhanced the excellence of medical recommender systems through delivering recommendations, which are built on patient desires and responses [2]. Owing to the exponential growth in the volume of data and rapid progress in PC hardware, deep learning (DL) has been used in numerous areas with outstanding outcomes. A DL-based recommender system has effectively enhanced the present performance of the recommender system and accuracy by removing significant features from information like images and review texts [3].

Although existing drug recommendation methods have shown promising performance, they still face challenges in providing personalized recommendations by effectively utilizing patient context and medical information. To address this issue, this manuscript presents the Personalized Context-Aware Drug Recommendation System using the Optimal Deep Learning (PCADRS-ODL) technique. The major contributions of this work are as follows:

- Utilization of the BERT-based word embedding approach to generate meaningful feature representations from the input data.
- Application of the Double Attention-Convolutional Neural Network with a Bidirectional Gated Recurrent Unit framework (DAC-BiGRU) for learning contextual information and generating accurate drug recommendations.
- Adoption of the Gannet Optimization Algorithm (GOA) for optimal tuning of the DAC-BiGRU model parameters.

Performance evaluation on a benchmark dataset to demonstrate the effectiveness of the proposed model compared to existing methods.

Literature Review

This section offers a brief review of literature works on context-aware drug recommendation. Long et al. [4] aim to utilize AI methods to generate intellectual decisions for healthcare supply chain mode assortment. In [5], Sentence Extractor with Keywords is a significant and explainable categorization technique presented for removing data from medical notes in text identification. Yazdinejad et al. [6] present a safe drug supply chain management structure that must attain enhanced safety and intellect utilizing ML techniques. Ali et al. [7] propose an NLP-specific website to mechanize evidence-driven MDT choices for skin lesions with BCC as the use case. In [8], a novel technique was presented in order to measure and achieve dangers by accepting NLP. Shaik and Athithan [9] project a new technique that combines BC technology and ML methods for quality assessment in farming supply chains. Rathor et al. [10] projected an Archimedes Optimizer with improved DL-driven recommendation systems. The authors [11] executed a method like an LSTM neural network with a hyperparameter process and fusion technique.

Proposed System

In this manuscript, the PCADRS-ODL method is presented. The proposed technique contains different kinds of procedures involved as data pre-processing, BERT-based word embedding, DAC-BiGRU-based categorization, and GOA-based parameter tuning. The complete structure of the PCADRS-ODL method is displayed in Figure 1.

BERT-driven Word Embedding

The BERT-based word embedding approach has been used to generate features. A word embedding is a lower-dimension vector representation of NLP that computers can understand. Fine-tuning and feature-based tasks are two traditional techniques to apply pre-trained language representation.

$$r_n = BERT(s_N) s_N \in S, n \in [0, N] \qquad (1)$$

Here, r_n indicates the sentence vector, N refers to the sentence counts, the index of a sentence is represented as n, s_n indicates the text of n^{th} sentence and the term S shows the sentence.

Classification using DAC-BiGRU Model

In this work, a DAC-BiGRU hybrid DL algorithm is introduced for drug recommendation. The DAC-BiGRU utilizes the capability of CNN for feature extraction and local patterns and Bi-GRU for modelling sequential data, but internally nests 2 attention modules for improving the model's capability to focus on the essential parts [12]. BiGRU is a bi-directional RNN that fuses backwards and forward GRUs for sequential data processing. The present HLf Bi-GRU is defined by the

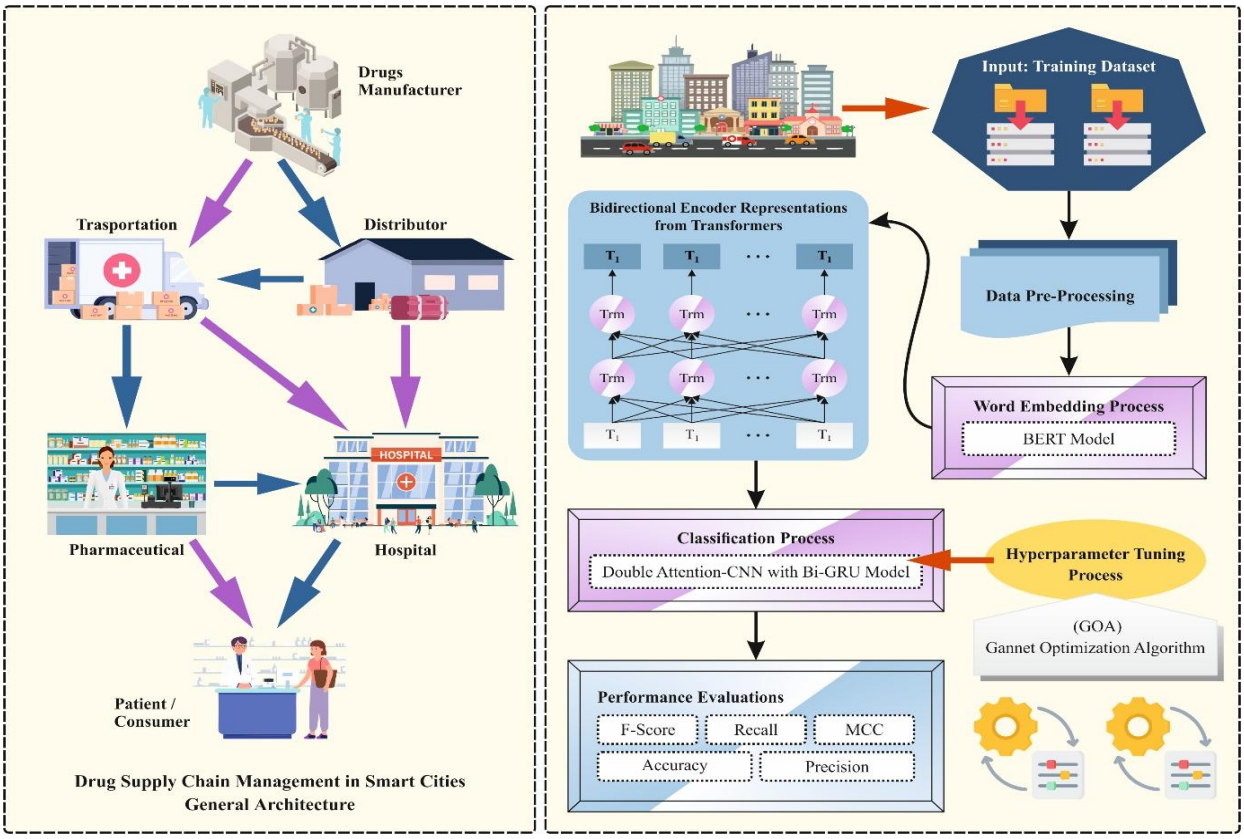

Figure 1. Workflow of PCADRS-ODL technique.

xisting inp a_t, the input of HL at the time $(H_{t-1})^{\rightarrow}$, and the output of the reverse HL $(H_{t-1})^{\leftarrow}$.

$$\overrightarrow{H_t} = I\left(a_t, \overrightarrow{H_{t-1}}\right) \quad (2)$$

$$\overleftarrow{H_t} = I\left(a_t, \overleftarrow{H_{t-1}}\right) \quad (3)$$

$$H_t = u_t\overrightarrow{H_t} + v_t\overleftarrow{H_t} + b_t \quad (4)$$

GRU is a non-linear transformation of the input vectors and encodes the vector, u_t, and v_t Are the weights of the forward-backwards HL? The double-attention mechanism operates on the contextual representations generated by the BiGRU layer. The channel attention mechanism assigns higher weights to the most informative feature channels while suppressing less relevant ones. Subsequently, the temporal attention mechanism identifies the most significant sequence positions within the BiGRU hidden representations by assigning attention weights to important contextual information. The combined effect of channel and temporal attention enables the proposed framework to focus on the most discriminative features and contextual patterns, thereby enhancing the overall text classification performance.

The attention module stimulates human consideration by enabling the model to concentrate on various input measures, thereby automatically changing the weight to enhance critical information opinion.

$$y = \sum_{i=1}^{n} \omega_i H_i \quad (5)$$

$$\omega_i = \frac{\exp(e_i)}{\Sigma_{j=1}^{n} e_j} \quad (6)$$

$$e_i = \alpha_i Att\ (\beta_i H_i + c_i) \quad (7)$$

Where e_i and ω_i Are the corresponding weights to feature with the time , α_i, and β_i denotes the shared weighted, b represents the biased, and c_i refers to the resultant at t time. Fig. 2 shows the infrastructure of the DAC-BiGRU system (Figure 2).

Model Optimization Strategy

The GOA module is utilized to optimize the hyperparameters of the DAC-BiGRU architecture [13]. All the individuals from the gannet population initialization.

$$a_i = rand \times\ (ub - lb) + lb \quad (8)$$

The parameters lb and ub illustrate minimal and

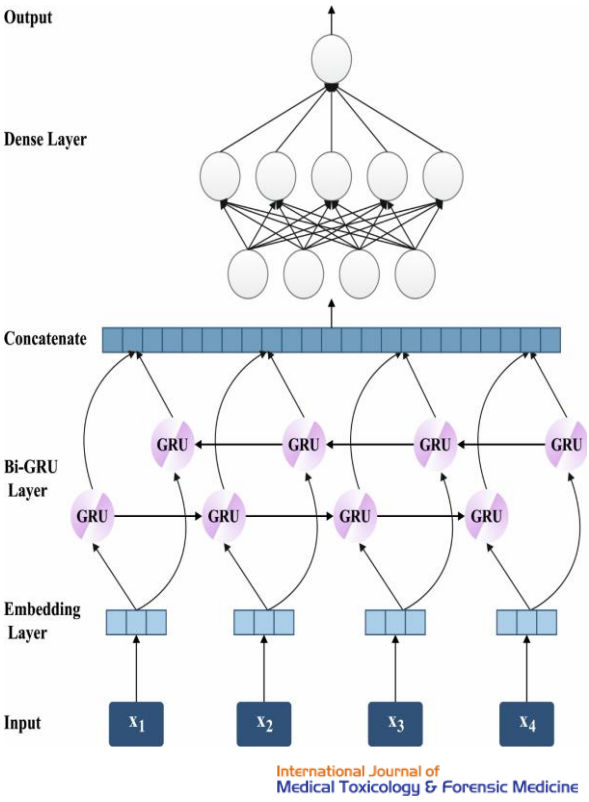

Figure 2. Architecture of DAC-BiGRU.

maximal bounds vectors of the searching space, correspondingly. The term $a_i$defines the vector position of i^{th}gannet. The initialization of GOA is done; this technique has entered the exploration step.

$$U_a = 2\cos(2\pi \times r1) \times \frac{C_{\max} - C}{T_{\max}} \quad (9)$$

$$\begin{cases} V_b = 2T \times (2\pi \times r_2) \times \dfrac{C_{\max} - C}{C_{\max}} \\ C = \begin{cases} 1 - \dfrac{S}{\pi}, s \in (0, \pi) \\ \dfrac{S}{\pi} - 1, s \in (\pi, 2\pi) \end{cases} \end{cases} \quad (10)$$

$$x_i^{C+1} = \begin{cases} x_i^C + U_{a1} + U_{a2}, & rand \le p(a) \\ x_i^C + V_{b1} + V_{b2}, & rand > p(b) \end{cases} \quad (11)$$

$$\begin{cases} U_{a2} = A \times (x_i^C - x_r^C) \\ V_{b2} = B \times (x_i^C - \bar{x}^C) \end{cases} \quad (12)$$

$$\begin{cases} A = (2_{r3} - 1) \times U_a \\ B = (2_{r4} - 1) \times V_b \end{cases} \quad (13)$$

$$\bar{x}^C = \frac{\sum_{i=1}^{N} x_i^C}{N} \quad (14)$$

$$x_i^{C+1} = \begin{cases} x_i^C + t \times Del \times (x_i^C - x_{best}^C), Cap > c(a) \\ x_{best}^C - (x_i^C - x_{best}^C) \times t \times lv, Cap \le c(b) \end{cases} \quad (15)$$

When the gannet loses its prey beyond the hunting range, it switches to a Lévy flight (LF) strategy.

$$\begin{cases} Del = cap \times |x_i^C - x_{best}^C| \\ lv = Levy(D) \end{cases} \quad (16)$$

$$\begin{cases} Levy(D) = 0.01 \times \dfrac{\mu\sigma}{|v|1/\beta} \\ \mu = \left(\dfrac{\sin\left(\frac{\pi\beta}{2}\right)\Gamma(1+\beta)}{\Gamma\left(\frac{(1+\beta)}{2}\right)\beta \times 2^{(\beta-1)/2}} \right) 1/\beta \end{cases} \quad (17)$$

GOA considers accuracy as the condition to define the fitness function (FF) as:

$$Fitness = \max(P) \quad (18)$$

$$P = \frac{TP}{TP + FP} \quad (19)$$

Whereas, FP and TP imply false positives and true positives.

Results and Discussions

The empirical assessment of the PCADRS-ODL method is examined using the benchmark dataset [14]. This dataset contains 20000 samples with two classes, as illustrated in Table 1. Table 2 represents the parameter settings of the proposed methodology.

Figure 3 depicts confusion matrices of the PCADRS-ODL model with 80:20 and 70:30 splits, showing high classification accuracy with minimal false predictions across training and testing phases.

Figure 4 presents the classification outcomes of the PCADRS-ODL method with 80:20 TRANP/TESTP. With 80% TRANP, the PCADRS-ODL system obtains an average accuracy of 97.34%, precision of 97.35%, recall of 97.34%, F-score of 97.34%, and MCC of 94.69%. With 20% TESTP, the PCADRS-ODL algorithm got an average accuracy of 97.25%, precision

Table 1. Details of the dataset.

Class (Ratings)	No. of Instances
Negative (1-5)	10000
Positive (6-10)	10000
Total Instances	20000

International Journal of Medical Toxicology & Forensic Medicine

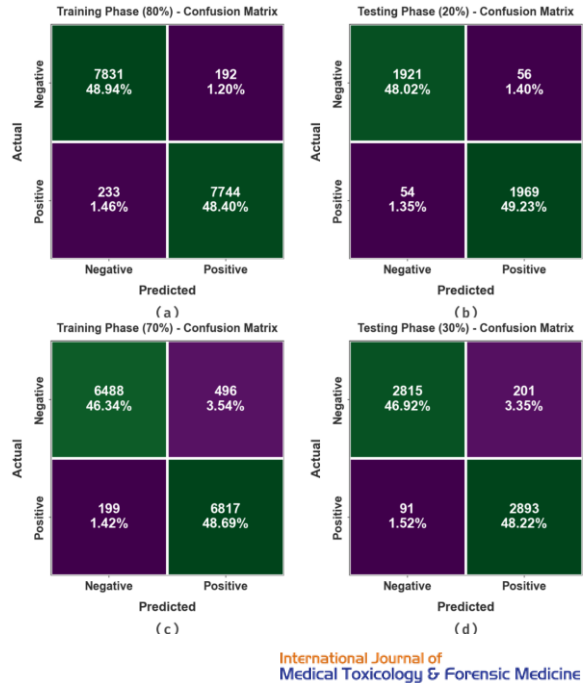

Figure 3. Confusion matrices of PCADRS-ODL with (a-b) 80:20 and (c-d) 70:30 split.

Table 2. Parameter setting.

Method	Parameter	Value
Word Embedding	Pre-trained Model	BERT-Base
Word Embedding	Embedding Dimension	768
Tokenization	Maximum Sequence Length	128
CNN	Kernel Size	3
CNN	Number of Filters	128
BiGRU	Hidden Units	256
Attention	Mechanism	Double Attention
Dense Layer	Units	128
Optimizer	Hyperparameter Tuning	GOA
Learning Rate	Initial LR	0.001
Batch Size	Batch Size	32
Training	Epochs	100

of 97.25%, recall of 97.25%, F-score of 97.25%, and MCC of 94.50%.

Figure 5 reveals the overall classifier analysis of the PCADRS-ODL approach at 70:30 TRANP/TESTP. With 70%TRANP, the PCADRS-ODL algorithm achieves an average. $accur_y$ of 95.03%, $preci_n$ of 95.12%, $recal_l$ of 95.03%, F_{score} of 95.03%, MCC and of 90.15%. Moreover, with 30%TESTP, the PCADRS-ODL system realises an average. $accur_y$ of 95.14%, $preci_n$ of 95.19%, $recal_l$ of 95.14%, F_{score} of 95.13%, and MCC of 90.33%.

Figure 6, a detailed comparison assessment of the PCADRS-ODL technique is provided with existing

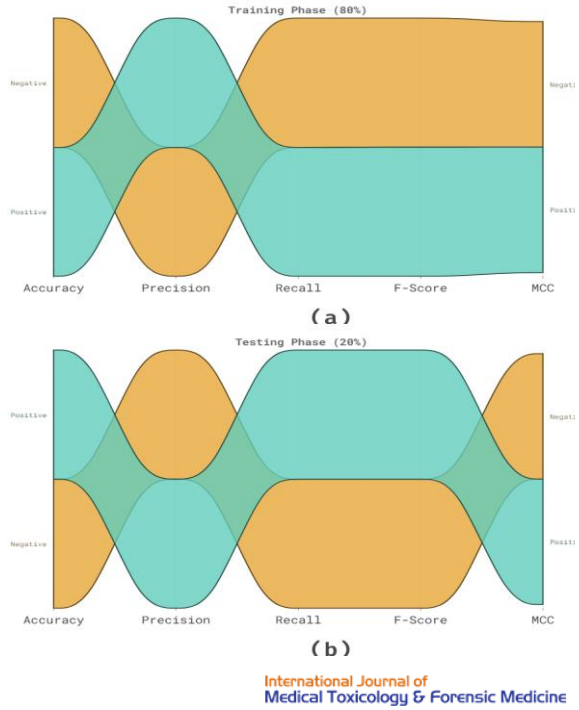

Figure 4. Classifier outcome of PCADRS-ODL technique with 80:20 TRANP/TESTP.

systems [15]. The experimental outcome depicted that the PCADRS-ODL system gains optimum performance based onattains a maximumatainsa maximum$accur_y$ of 97.34% while the AOAEDL-RS, DT, LR, RF, SVM, and NB methodologies cquire minimums$accur_y$ values of 96.94%, 90.90%, 94.92%, 90.42%, 93.84%, and 95.27%. esides based on$resuperior$ins a superior reriorrec *values* al_l of 97.34% while the AOAEDL-RS, DT, LR, RF, SVM, and NB algoriths achive lesser$recal_l$ values of 96.94%, 93.03%, 92.83%, 90.44%, 93.49%, and 94.37%.

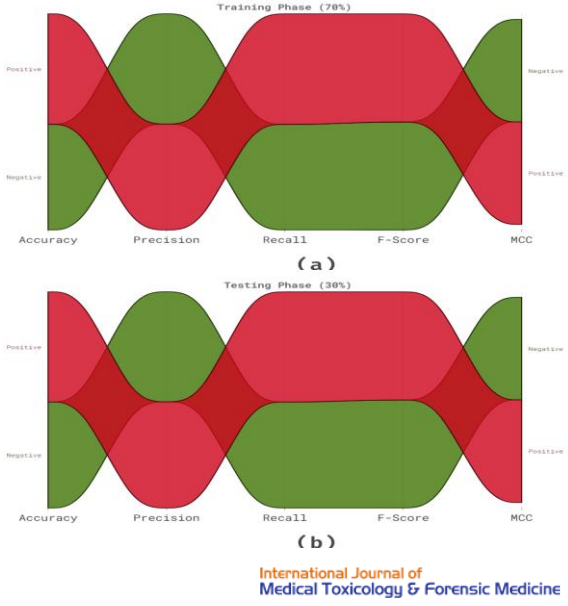

Figure 5. Classifier outcome of PCADRS-ODL technique with 70:30 split.

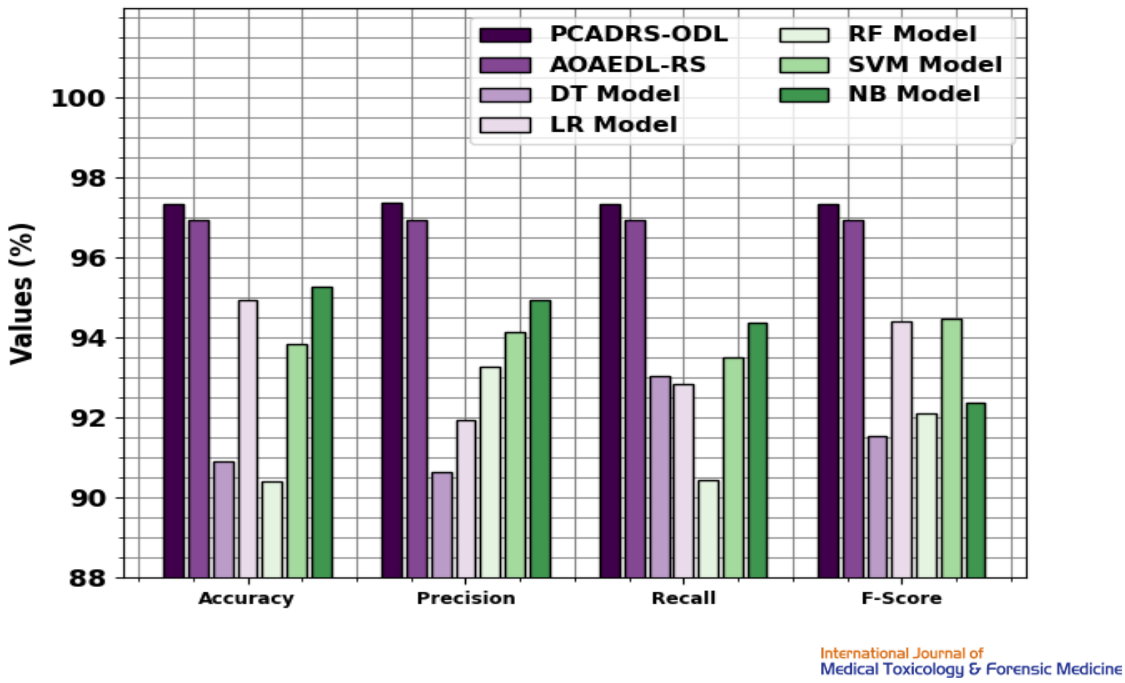

Figure 6. Comparative analysis of PCADRS-ODL with other approaches.

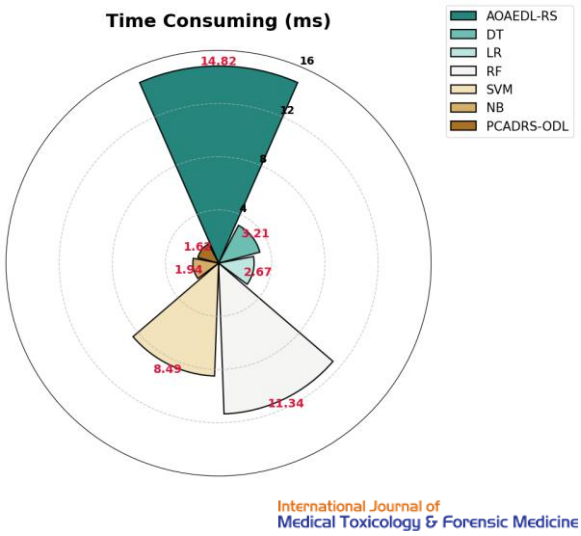

Figure 7. Comparative analysis of PCADRS-ODL technique with other approaches.

Finally wih respect to F_{score}, the PCADRS-ODL model accomplishes a greater F_{score} of 97.34% while the AOAEDL-RS, DT, LR, RF, SVM, and NB techniques attain minimal F_{score} values of 96.93%, 91.54%, 94.39%, 92.10%, 94.46%, and 92.38%. Thus, the proposed model is found to be better than existing models.

Figure 7 depicts the computation efficiency of the PCADRS-ODL model with existing methods. The PCADRS-ODL method provides a lesser inference time of 1.62ms, compared with the existing methods, such as AOAEDL-RS, DT, LR, RF, SVM, and NB models, which attain a greater time of 14.82ms, 3.21ms, 2.67ms, 11.34ms, 8.49ms, and 1.94ms.

Table 3 represents the ablation study of the PCADRS-ODL model. The BiGRU attains an $accu_y$ of 93.62%, $prec_n$ of 93.58%, $reca_l$ of 93.55%, and $F1_{score}$ of 93.56%. and DAC-BiGRU offering an $accu_y$ of 94.84%, $prec_n$ of 94.79%, $reca_l$ of 94.76%, and $F1_{score}$ of 94.77%. Besides, the BERT + BiGRU

Table 3. Ablation study of the PCADRS-ODL methodology.

Methods	Accuracy	Precision	Recall	F1-Score
BiGRU	93.62	93.58	93.55	93.56
DAC-BiGRU	94.84	94.79	94.76	94.77
BERT + BiGRU	95.68	95.65	95.62	95.63
BERT + DAC-BiGRU	96.63	96.61	96.59	96.60
PCADRS-ODL (BERT + DAC-BiGRU + GOA)	97.34	97.35	97.34	97.34

International Journal of Medical Toxicology & Forensic Medicine

obtained a higher. $accu_y$ to 95.68%, $prec_n$ to 95.65%, $reca_l$ to 95.62%, and $F1_{score}$ to 95.63%. Moreover, the BERT + DAC-BiGRU technique is additionally enhanced. $accu_y$ to 96.63%, $prec_n$ to 96.61%, $reca_l$ to 96.59%, and $F1_{score}$ to 96.60%. At last, the PCADRS-ODL (BERT + DAC-BiGRU + GOA) model achieved the highest outputs with an $accu_y$ of 97.34%, $prec_n$ of 97.35%, $reca_l$ of 97.34%, and $F1_{score}$ of

97.34%, emphasising the efficiency of the components.

Table 4 depicts the K-fold cross-validation of the PCADRS-ODL approach with various metrics. The PCADRS-ODL method achieved a constant value through each fold, with values ranging from 94.91% to 96.83%. The other metrics also remained steady across folds, implying consistent classifier solutions. The maximal performance is noticed in K-3, whereas the minimal value is observed in K-2. This confirms the steadiness and effectiveness of the proposed PCADRS-ODL system.

Table 4. K-fold cross-validation of the proposed PCADRS-ODL techniques.

Fold	Accuracy	Precision	Recall	F1-Score
K-1	95.42	94.76	94.88	94.82
K-2	94.91	94.35	94.17	94.26
K-3	96.83	96.10	96.15	96.19
K-4	96.12	95.63	95.71	95.67
K-5	95.36	94.92	95.03	94.97
K-6	96.38	95.84	95.92	95.88
K-7	95.79	95.31	95.26	95.28
K-8	96.51	95.96	96.02	95.99
K-9	96.27	95.71	95.83	95.77
K-10	95.87	95.21	95.48	95.34

International Journal of
Medical Toxicology & Forensic Medicine

The experimental results demonstrate that the proposed framework achieves consistent and reliable classification performance across multiple evaluation metrics, including Accuracy, Precision, Recall, F-Score, and Matthews Correlation Coefficient (MCC). The inclusion of k-fold cross-validation further confirms the robustness and consistency of the proposed model. Furthermore, the GOA effectively optimizes the model hyperparameters, contributing to improved classification performance. The comparative inference time analysis also demonstrates the practical suitability of the proposed framework for intelligent drug supply chain monitoring. Although the proposed framework demonstrates promising performance, several aspects remain for future investigation. These include uncertainty estimation, confidence calibration, and robustness assessment under noisy textual data and evolving pharmaceutical information.

Conclusion

In this study, the MST-AMIC model has been designed for automatic medical image captioning using CXR images. The input images and radiology reports are first preprocessed. EfficientNet-B4 is used to extract multi-scale features that capture low-, mid-, and high-level visual details. These features are then fused to obtain a single representation. Finally, captions are generated in an autoregressive manner using greedy decoding. The model is examined on the MIMIC-CXR dataset and shows better performance compared to existing methods.

Although the proposed PCADRS-ODL framework demonstrates promising performance for drug supply chain text classification, several limitations remain. The current study primarily focuses on classification performance and does not include uncertainty estimation or confidence calibration. Furthermore, the robustness of the proposed framework under noisy textual inputs and previously unseen data distributions has not been explicitly evaluated. Future work will focus on incorporating uncertainty-aware prediction and confidence calibration techniques to improve the reliability of the proposed framework. Domain adaptation and continual learning strategies will also be explored to enhance the scalability and generalization capability of the proposed framework across diverse pharmaceutical supply chain environments.

Acknowledgment

None.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript. All authors contributed to the manuscript and approved the final version.

References

[1] Akram W, Joshi R, Haider T, Sharma P, Jain V, Garud N, et al. Blockchain technology: A potential tool for the management of pharma supply chain. Res Social Adm Pharm. 2024. [DOI: 10.1016/j.sapharm.2024.02.014]

[2] Babu CS, Akshara PM. Revolutionizing conversational AI: unleashing the power of ChatGPT-Based applications in generative AI and natural language processing. In: Advanced applications of generative AI and natural language processing models. IGI Global; 2024. p. 228-248. [DOI: 10.4018/979-8-3693-0502-7.ch011]

[3] Jackson I, Jesus Saenz M, Ivanov D. From

natural language to simulations: applying AI to automate simulation modelling of logistics systems. Int J Prod Res. 2024;62(4):1434-57. [DOI: 10.1080/00207543.2023.2276811]

[4] Long P, Lu L, Chen Q, Chen Y, Li C, Luo X. Intelligent selection of healthcare supply chain mode-an applied research based on artificial intelligence. Front Public Health. 2023;11:1310016. [DOI: 10.3389/fpubh.2023.1310016]

[5] Tavabi N, Pruneski J, Golchin S, Singh M, Sanborn R, Heyworth B, et al. Building large-scale registries from unstructured clinical notes using a low-resource natural language processing pipeline. Artif Med. 2024;102847. [DOI: 10.1016/j.artmed.2024.102847]

[6] Yazdinejad A, Rabieinejad E, Hasani T, Srivastava G. A BERT-based recommender system for secure blockchain-based cyber physical drug supply chain management. Cluster Comput. 2023;26(6):3389-403. [DOI: 10.1007/s10586-023-04088-6]

[7] Ali SR, Dobbs TD, Tarafdar A, Strafford H, Fonferko-Shadrach B, Lacey AS, et al. Natural language processing to automate a web-based model of care and modernize skin cancer multidisciplinary team meetings. Br J Surg. 2024;111(1):znad347. [DOI: 10.1093/bjs/znad347]

[8] Silvestri S, Islam S, Amelin D, Weiler G, Papastergiou S, Ciampi M. Cyber threat assessment and management for securing healthcare ecosystems using natural language processing. Int J Inf Secur. 2023:1-20. [DOI: 10.21203/rs.3.rs-3307666/v1]

[9] Shaik J, Athithan S. Utilizing Blockchain Technology and Machine Learning for Quality Evaluation in Agricultural Supply Chains. Int J Comput Digit Syst. 2024;16(1):1-14. [Link]

[10] Rathor K, Chandre S, Thillaivanan A, Raju MN, Sikka V, Singh K. Archimedes Optimization with Enhanced Deep Learning based Recommendation System for Drug Supply Chain Management. In: 2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN). IEEE; 2023. p. 1-6. [DOI: 10.1109/ICSTSN57873.2023.10151666]

[11] MK NP, Rastogi S, Alice K. Demand Forecasting in Supply Chain Management using CNN-LSTM Hybrid Model. In: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE; 2023. p. 1-5. [DOI: 10.1109/ICCCNT56998.2023.10307665]

[12] Wang Q, Liu P, Zhu Z, Yin H, Zhang Q, Zhang L. A text abstraction summary model based on BERT word embedding and reinforcement learning. Appl Sci. 2019;9(21):4701. [DOI: 10.3390/app9214701]

[13] Peng W, Jiang Z, Yao H, Zhang L, Yu J. A Hybrid Deep Learning Model for Short-Term Hydropower Generation Prediction Incorporating Hydrometeorological Factors. Research Square [Preprint]. 2023. [DOI: 10.21203/rs.3.rs-3483896/v1]

[14] UC Irvine Machine Learning Repository. Drug Review Dataset (Drugs.com) [Internet]. Irvine (CA): University of California, School of Information and Computer Science; 2018 [cited 2026 Aug 8]. [Link]

[15] Zhao ML, Ni SQ, Du ZG, Wang XY, Tian AQ, Ma XL. Multi-Objective Gannet Optimization Algorithm for Dynamic Passenger Flow Allocation in Train Operation Plan Optimization. IEEE Access. 2023. [DOI: 10.1109/ACCESS.2023.3318262]