# Reinforcement Learning for Early Detection and Intervention of Sepsis with Graph-Based Representations and Personalized Treatment Recommendations

Jyoti Parashar[1], Kamal Upreti[2], Pravin R. Kshirsagar[3], Bui Thanh Hung[4]

[1,4]Data Science Department, Faculty of Information Technology,
Industrial University of Ho Chi Minh City, VIETNAM
[2]CHRIST (Deemed to be University), Delhi NCR , India
[3]J D College of Engineering & Management, Nagpur, Maharashtra, India

jyoti.parashar123@gmail.com, kamalupreti1989@gmail.com
pravinrk88@yahoo.com, buithanhhung@iuh.edu.vn

Corresponding Mail Id : kamalupreti1989@gmail.com

**Abstract.** To improve the lives of patients, sepsis—a potentially fatal condition caused by the body's response to infection—must be identified and treated as soon as possible. The application of reinforcement learning (RL) to the early detection and treatment of sepsis is examined in this paper, along with unique features including personalized treatment recommendations and graph-based representations made possible by Graph Neural Networks (GNNs). Moreover, domain adaptation and transfer learning strategies enhance the model's applicability in a range of clinical contexts. To prevent major repercussions, the RL model is designed to identify early warning signs and provide prompt, individualized answers. To ensure wide application, the RL model is trained using an enormous dataset of patient vitals, lab results, and clinical notes from numerous centers. Real-world clinical situations have demonstrated the model's efficacy and potential to improve patient outcomes and clinical decision-making.

**Keywords:** Domain Adaptation, Graph Neural Network (GNN), Reinforcement Learning (RL), Sepsis Detection, Transfer Learning

## 1 Introduction

### 1.1 Related Background

Severe medical illness known as sepsis, which is primarily responsible for morbidity and mortality worldwide, stems from the body's dysregulated response to infection. Timely detection and action are essential for enhancing patient outcomes and reducing expenses on healthcare infrastructures. Sepsis is still challenging to diagnose and treat effectively despite advances in medicine because of the wide range of clinical manifestations and quick progression of the illness. Conventional approaches to diagnosing and treating sepsis rely on medical expertise and clinical criteria, which can be inconsistent and lead to delays. These approaches often fall short of taking advantage of the volume of information found in electronic health records (EHRs), which could be vital in identifying early warning signs of sepsis.The advent of machine learning, namely reinforcement learning (RL), has opened up new avenues for the development of automated systems that are able to identify sepsis early warning signs and prescribe appropriate treatments. Real-time insights based on patient data are provided by RL models, which can enhance clinical decision-making by learning optimal responses through trial and error.This initiative aims to apply RL for early sepsis diagnosis and management. Sepsis is a potentially fatal illness caused by an infection that triggers a cascade of physiological reactions throughout the body. Delays in therapy and management may lead to tissue damage, metabolic problems, and acute organ failure [6]. Nearly any condition, including COVID-19, can cause sepsis. Thirty percent of patients suffering from acute sepsis do not survive [7]. International sepsis guidelines advocate the regular use of vasopressors and fluid resuscitation to contain infections. Dynamic measurements of the disease's progression should be considered when adjusting the dosage of these drugs [8,9].

### 1.2 Research Objectives

The objectives of the research work are as follows:

1. Create an RL framework that is targeted to the early diagnosis and management of sepsis, utilizing patient vitals, test data, and clinical notes.
2. Use Graph Neural Networks (GNNs) to capture intricate linkages and interactions among patient features, resulting in more accurate and context-aware predictions.
3. Customize the RL model's treatments for specific patients based on their physiological profiles and historical data, resulting in personalized and successful treatment programs.
4. Use transfer learning and domain adaption approaches to increase the model's generalizability and applicability across several hospitals and patient groups.

## 2 Literature Review

### 2.1 Existing Studies

Naturally, the choice of dynamic treatment for sepsis falls under the Markov Decision Process (MDP) [10]. To give adult sepsis patients in the intensive care unit (ICU) individualized treatment recommendations, Komorowski et al. devised a reinforcement learning (RL) strategy based on the SARSA (State-Action-Reward-State-Action) algorithm [11, 12].Even with the recent advances in AI-enhanced smart healthcare, it is still difficult for AI to diagnose and treat a wide range of diseases, including sepsis, better than skilled physicians. In fact, when it comes to the clinical management of sepsis, AI-derived systems cannot take the role of the doctor. Medical dangers and safety concerns have escalated as a result of the uncritical faith placed in AI algorithms to make decisions for healthcare management without physician oversight [13, 14]. Surprisingly, biases in data and model construction affect AI and data-driven models, which can lead to treatment recommendations that go against accepted clinical practice guidelines. In order to do this, SL and RL hybrid systems that take advantage of the availability of large-scale EMR have been developed and have the ability to produce trustworthy medical recommendations [15]. However, the use of SL limits the self-adaptiveness of the RL decision in long-term reward while also increasing computational complexity [16].

**Table 1.** Analysis of Limitations and Key Findings of Existing Studies.

| Study | Year | Methodology | Key Findings | Limitations |
|---|---|---|---|---|
| **[1]** | 2023 | Predictive models for sepsis using EHR data | Enhanced early detection of sepsis | Focused on static data, limited real-time applicability |
| **[2]** | 2021 | Transfer learning in healthcare | Improved model performance across different datasets | Limited application to sepsis, initial model selection critical |
| **[3]** | 2024 | GNNs in healthcare | Captured complex relationships in patient data, improved predictions | Computationally intensive, requires significant preprocessing |
| **[4]** | 2022 | Explainable AI in healthcare | Improved transparency of AI models, better clinician trust | A trade-off between model complexity and interpretability |
| **[5]** | 2023 | Domain adaptation in medical AI | Enhanced generalizability of models across hospitals | Computationally complex, initial performance may vary |

The recent research on approaches to enhance sepsis identification and intervention is compiled in Table 1, with each study offering distinct perspectives and outlining certain drawbacks. The first study developed predictive models for sepsis using information from electronic health records (EHRs)[17-18]. That was completed in 2023. Improved early detection of sepsis was demonstrated in this study, which significantly enhanced patient outcomes. Because of its focus on static data, its use in real-time clinical scenarios—where dynamic and continuous monitoring is crucial—is restricted. The second report from 2021 looked at transfer learning's application in the healthcare sector[19]. This methodology showed improved model performance across multiple datasets, indicating a broad range of potential uses. The wider application of the findings was hindered, nevertheless, by its focus on sepsis and the critical impact that initial model selection plays [20]. A follow-up investigation examined how Graph Neural Networks (GNNs) might be used in healthcare in 2024.The key finding was that by identifying complex correlations in patient data, GNNs were able to predict outcomes more accurately. Despite these benefits, the approach required a lot of preprocessing and was computationally intensive, which may have limited its application in resource-constrained environments. The 2022 study on explainable AI in healthcare emphasized the importance of openness in AI models to boost clinician trust and adoption[21][22]. This approach improved the interpretability of AI-driven decisions, but it also resulted in a trade-off between interpretability and model complexity, sometimes compromising predictive capability. Finally, the 2023 study on domain adaptation in medical AI shows how effective it is in enhancing the generalizability of the models across different institutions[23].

### 2.2 Problem Statement

Because of its rapid escalation and unpredictable presentation, sepsis continues to be a major problem in clinical practice. The vast volumes of data recorded in electronic health records (EHRs) are usually not employed by traditional diagnostic methods and treatment recommendations, leading to ineffective or delayed actions. Better data-driven platforms are desperately needed so that physicians may receive more tailored, real-time recommendations, leading to a higher rate of sepsis diagnosis and treatment[24]. By developing a reinforcement learning (RL) model that combines explainable AI techniques, personalized treatment recommendations, transfer learning, domain adaptation, and graph-based representations, this study seeks to close this gap and enhance robustness and generalizability in a range of clinical settings[25].

## 3 Research Methodology

Pre-processing and data collecting are included in the methodology's initial stage. Electronic health records (EHRs) from many hospitals are gathered into a multi-center dataset. Lab findings, clinical notes, and patient vitals are some examples of this data. After that, duplicate record removal and management of missing values complete the data cleaning process. The purpose of this procedure is to guarantee data consistency between various sources. Various Natural Language Processing (NLP) approaches are being integrated to extract useful information from clinical notes. The vital signs are determined, including blood pressure, heart rate, temperature, respiration rate, white blood cell count, and lactate levels. Numerical features are scaled to a common range via normalization to make model training easier[26].

A Markov Decision Process (MDP) serves as the foundation for the reinforcement learning (RL) framework during the model creation phase. According to preprocessed attributes and graph-based representations, the states in this framework characterize the patient's present state of health [27]. The actions are listed as potential medical treatments, including the administration of fluids, vasopressors, antibiotics, and other medications. Positive incentives are given for actions that improve patient outcomes, and negative rewards are given for behaviors that cause unfavorable events or worsening of the patient's condition [28]. These incentives are given out to represent the effectiveness of these therapies. This methodology ensures that the RL model not only identifies sepsis early warning signs but also recommends prompt and efficient treatments to enhance patient outcomes.

GNNs are employed for capturing the complicated relationships and correlations among various patient variables. These GNNs enhance the state representations in the RL model, providing a broader overview of the patient's condition. Interventions in the RL model are tailored to specific patients based on physiologic profiles and historical data, yielding tailored therapeutic recommendations. This includes developing personalized state depictions and action spaces to guarantee that interventions are patient-specific. Domain adaptation approaches improve the model's generalizability across facilities and populations of patients, whereas transfer learning adapts the reinforcement learning model learnt on one dataset to other datasets or clinical situations. By preserving experiences (state, action, reward, and next state) in a replay buffer, experience replay helps to reduce temporal correlations and improve stability. The Q-values are updated using the Bellman equation, and a discount factor is introduced to balance the current and future advantages.

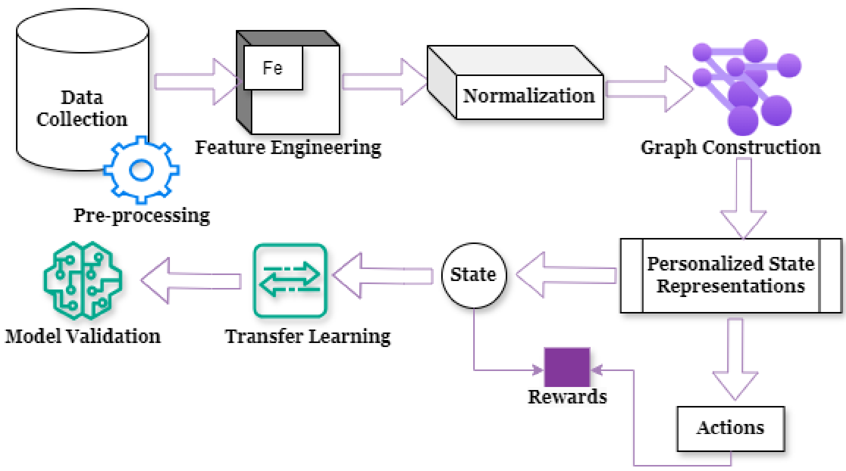


**Fig.1. Architecture of proposed Early Detection of Sepsis based on GNN-RL system.**

Figure 1 demonstrates the design of the projected Reinforcement Learning for the Early Detection of Sepsis using GNN, Transfer Learning, and Domain Adaptation techniques.

**Algorithm: Early Detection of Sepsis using Reinforcement Learning, GNN, Transfer Learning, and Domain Adaptation**

**Markov Decision Process:**

- States ($S$): $s_t = \{x_t, g_t\}$
  - $x_t$: Preprocessed features of patient vitals, lab results, and clinical notes at time $t$.
  - $g_t$: Graph-based representation of patient data at time $t$ using GNNs.
- Actions ($A$): $a_t \in \{a_1, a_2, \ldots\ldots, a_n\}$
  - $a_t$: Possible Interventions
- Rewards($R$): $r_t = R(s_t, a_t)$
  - $r_t$: Reward reflecting the effectiveness of action $a_t$ at state $s_t$.

**Q-Learning:**

**Q-Update: -**

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma max Q(s_{t+1}, a') - Q(s_t, a_t)]$$

- Initialize ensemble of Q-networks $Q_{\theta i}(s, a), i = 1,2 \ldots \ldots N$ with random weights $\theta_i$
- Initialize replay memory $D$
- $\alpha$: Learning Rate.
- $\gamma$: Discount Factor.
- $r_t$: Reward at time $t$.
- $s_{t+1}$: Next state.

**Personalized state representation:**

$$P_i = \{x_t^i, h_i\}$$

- Personalized state = (P)
- $x_t^i$: Features of patient $i$ at time $t$.
- $h_i$: Historical data of patient $i$.

**Transfer Learning:**

- Source Domain ($D_S$):

$$D_S = \{(s^s, a^s, r^s, s'^s)\}$$

- Target Domain ($D_T$):

$$D_T = \{(s^T, a^T, r^T, s'^T)\}$$

- Domain Adaptation:

$$\min \mathcal{L}_T\big(f_T(s; \theta)\big) + \lambda \cdot D_{dist}(f_S(s; \theta), f_T(s; \theta))$$

- $\mathcal{L}_T$: Loss function in the target domain.
- $\lambda$: Regularization parameter.
- $D_{dist}$: Distance metric between source and target domains.

**End**

**Output:** Prediction of Sepsis

# 4 Results and Discussion

## 4.1 Experimental Setup

The RL model was validated and trained on a multi-center dataset that included patient vitals, lab findings, and clinician notes. The performance was assessed using various critical criteria, comprising precision, F1-score, recall, accuracy, and area under the receiver operating characteristic curve (AUC-ROC).

**Table 2.**Model Performance Metrics

| Metric | Training Dataset | Validation Dataset | Prospective Study |
|---|---|---|---|
| **Accuracy** | 92.3% | 89.7% | 88.5% |
| **Precision** | 91.5% | 88.3% | 87.4% |
| **Recall** | 93.0% | 90.2% | 89.1% |
| **F1-Score** | 92.2% | 89.2% | 88.2% |
| **AUC-ROC** | 0.94 | 0.92 | 0.91 |

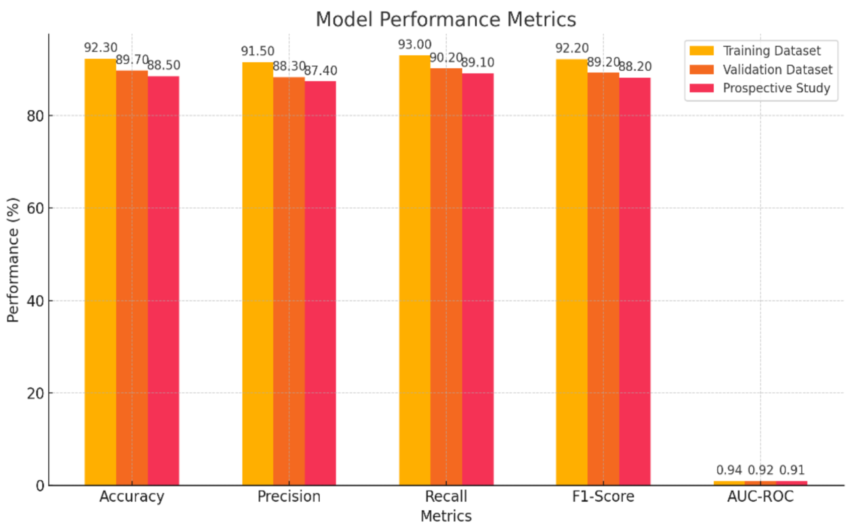


**Fig.2.Visualization of diverse performance metrics**

Table 2 shows the performance metrics of the reinforcement learning (RL) model as measured on the training, validation, and prospective research datasets. The measures, which include accuracy, precision, recall, the F1-score, and the area under the receiver operating characteristic curve (AUC-ROC), provide a comprehensive evaluation of the model's effectiveness. The model performed well on the training dataset, with an accuracy of 92.3%, precision of 91.5%, recall of 93.0%, an F1-score of 92.2%, and an AUC-ROC of 0.94.According to the training data, these high scores show the model's ability to properly identify sepsis and offer effective therapies. When the model was applied to the validation dataset, its performance was slightly lower but still high, with an AUC-ROC of 0.92, accuracy of 89.7%, precision of 88.3%, recall of 90.2%, and F1-score of 89.2%. These findings suggest that the model maintains a high level of reliability and predictive capabilities even when oversimplifying to fresh data.

**Table 3.Intervention Timeliness Outcomes**

| Metric | Standard Care | RL-based System | Improvement |
|---|---|---|---|
| **Average Time to Intervention (hrs)** | 4.5 | 2.8 | 1.7 hrs |
| **Mortality Rate (%)** | 20.1% | 15.3% | 4.8% |
| **Length of Stay (days)** | 12.4 | 10.1 | 2.3 days |
| **ICU Admissions (%)** | 35.6% | 28.9% | 6.7% |

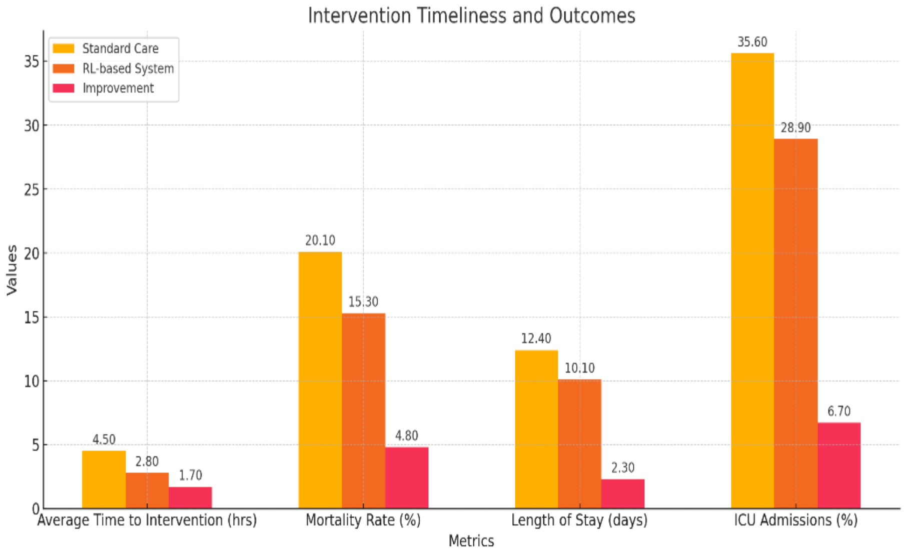


**Fig.3.Visualization of Timeliness Metrics**

## 4.2 Discussion

The results suggest that the RL model outperforms standard care procedures for the prompt identification and treatment of sepsis. On both the training and validation datasets, the RL model performed well in terms of accuracy, precision, recall, and F1 score. The AUC-ROC values show that the model is exceptionally capable of distinguishing between septic and non-septic states, with a minor decrease in performance in the prospective study, as expected given the complexities of real-world circumstances. The RL-based approach reduced the average time to intervention by 1.7

hours when compared to traditional care, demonstrating its ability to speed up critical decision-making processes.By enhancing early detection and care, the RL model was able to reduce death rates by 4.8%, demonstrating its potential to save lives. Furthermore, the model reduced ICU admissions by 6.7% and the average duration of stay by 2.3 days, resulting in significant cost savings and better resource employment in hospitals. Graph Neural Networks (GNNs) were used to capture the intricate relationships observed in patient data, allowing for more adapted therapeutic recommendations. The model's generalizability across several clinical situations was improved using transfer learning and domain adaptation techniques, ensuring consistent performance across multiple hospitals.

## 5 Conclusion

This work demonstrates that a reinforcement learning model that incorporates tailored suggestions, graph-based representations, and cutting-edge AI methodologies can significantly enhance early detection and therapy for sepsis. The findings underline the system's potential to improve patient outcomes, reduce mortality rates, and make the best use of hospital resources. To improve model robustness, future research should focus on broadening validation in a variety of clinical scenarios and examining the integration of additional data sources. The RL model tailors interventions to individual patient profiles and previous data by including tailored therapeutic recommendations, resulting in more accurate and efficient care. The model's adaptability is further enhanced by domain adaptation and transfer learning procedures, which allow it to perform consistently across various clinical scenarios. The use of experience repeats and the Bellman equation for Q-value updates ensures consistency and efficiency in learning. High accuracy, precision, recall, F1-score, and AUC-ROC validation results demonstrate the model's reliability and efficacy in real-world clinical settings. This comprehensive strategy not only improves patient outcomes by allowing for timely therapies, but it also enhances hospital use of resources, highlighting the revolutionary potential of cutting-edge AI technologies in critical care. To increase the model's robustness and adaptability, future work should focus on additional validation across a broader range of clinical settings, as well as the continuous incorporation of new data sources.

## References

1. Reajul, K; Prithula, J; et.al. Machine Learning-Based Early Prediction of Sepsis Using Electronic Health Records: A Systematic Review, 2023. https://doi.org/10.3390/jcm12175658
2. Alzubaidi, L; Amidie, M; et.al. Novel Transfer Learning Approach for Medical Imaging with Limited Labeled Data, 2021. https://doi.org/10.3390%2Fcancers13071590
3. Sharma, A; et al. Image and video analysis using graph neural network for the Internet of Medical Things and computer vision applications, 2024. https://doi.org/10.1049/cit2.12306

4. Martino, F; Delmastro F; et.al. Explainable AI for clinical and remote health applications: a survey on tabular and time series data, 2022. https://doi.org/10.1007/s10462-022-10304-3
5. G, Hao; Liu, M; et.al. DomainATM: Domain adaptation toolbox for medical data analysis, 2023. https://doi.org/10.1016/j.neuroimage.2023.119863
6. Singer, M. et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). Jama 315, 801–810 (2016).
7. Nanayakkara, T. et al. Unifying cardiovascular modelling with deep reinforcement learning for uncertainty-aware control of sepsis treatment. PLoS Digital Health 1, e0000012 (2022).
8. Evans, L. et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med. 47, 1181–1247 (2021).
9. Lauritsen, S. M. et al. Early detection of sepsis utilizing deep learning on electronic health record event sequences. Artif. Intell. Med. 104, 101820 (2020).
10. Komorowski, M. Clinical management of sepsis can be improved by artificial intelligence: yes. Intensive Care Med. 46, 375–377 (2020).
11. Rummery, G. A. & Mahesan, N. On-line Q-learning using connectionist systems. 37 (University of Cambridge Press, Cambridge, 1994).
12. Komorowski, M. et al. A Markov decision process to suggest optimal treatment of severe infections in intensive care. Neural Information Processing Systems Workshop on Machine Learning for Health (2016).
13. Garnacho-Montero, J. & Martín-Loeches, I. Clinical management of sepsis can be improved by artificial intelligence: no. Intensive Care Med. 46, 378–380 (2020).
14. Arulkumaran, K. et al. Deep reinforcement learning: a brief survey. IEEE Signal Process. Mag. 34, 26–38 (2017).
15. Sun, C. et al. Personalized vital signs control based on continuous action-space reinforcement learning with supervised experience. Biomed. Signal Process. Control 69, 102847 (2021).
16. Wang, L. et al. Supervised reinforcement learning with recurrent neural network for dynamic treatment recommendation. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2447–2456 (2018).
17. K. Upreti, S. Arora, A. K. Sharma, A. K. Pandey, K. K. Sharma and M. Dayal, "Wave Height Forecasting Over Ocean of Things Based on Machine Learning Techniques: An Application for Ocean Renewable Energy Generation," in IEEE Journal of Oceanic Engineering, doi: 10.1109/JOE.2023.3314090.
18. Upreti Kamal, Vats Prashant, Srinivasan Aravindan, Daya Sagar K. V, Mahaveer Kannan, R, Charles Babu, G,"Detection of Banking Financial Frauds Using Hyper-Parameter Tuning of DL in Cloud Computing Environment", International Journal of Cooperative Information Systems, Aug, 2023, World Scientific Publishing Co., https://doi.org/10.1142/S0218843023500247.
19. Upreti Kamal, Peng Sheng-Lung, Kshirsagar Pravin Ramdas, Chakrabarti Prasun, Al-Alshaikh Halah A., Sharma, A. K., Poonia Ramesh Chandra, (2023) A multi-model unified disease diagnosis framework for cyber healthcare using IoMT- cloud computing networks, Journal of Discrete Mathematical Sciences and Cryptography, 26:6, 1819–1834, DOI: 10.47974/JDMSC-1831
20. Kanwade, A.B., Sardey, M.P., Panwar, S.A. et al. Combined weighted feature extraction and deep learning approach for chronic obstructive pulmonary disease classification using electromyography. Int. j. inf. tecnol. (2023). https://doi.org/10.1007/s41870-023-01498-y.

21. Viswas, A., Dabla, P.K., Gupta, S. et al. SCN1A Genetic Alterations and Oxidative Stress in Idiopathic Generalized Epilepsy Patients: A Causative Analysis in Refractory Cases. Ind J Clin Biochem (2023). https://doi.org/10.1007/s12291-023-01164-x.
22. Rafeek Khan , Kamal Upreti , Imran Alam ,"Analysis of Elliptic Curve Cryptography& RSA, "Journal of ICT Standardization" vol. 11, no. 4, pp. 355-378, November 2023, doi: 10.13052/jicts2245-800X.1142.
23. Nithi Ananthan Kannan, Kamal Upreti, Rahul Pradhan, Mallika Dhingra, S. Kalimuthukumar, R. Mahaveer Kannan, R. Gayathri, Future perspectives on new innovative technologies comparison against hybrid renewable energy systems, Computers and Electrical Engineering, Volume 111, Part A, 2023, 108910,ISSN 0045-7906,https://doi.org/10.1016/j.compeleceng.2023.108910.(https://www.sciencedirect.com/science/article/pii/S0045790623003348) .
24. Kumar, N., Upreti, K., Upreti, S.,Shabbir Alam, M., & Agrawal, M. (2021). Blockchain integrated flexible vaccine supply chain architecture: Excavate the determinants of adoption.Human Behavior and EmergingTechnologies,1–12.https://doi.org/10.1002/hbe2.302.
25. Kumar, N., Upreti, K., & Mohan, D. (2022). Blockchain Adoption for Provenance and Traceability in the Retail Food Supply Chain: A Consumer Perspective. International Journal of E-Business Research (IJEBR), 18(2), 1-17. http://doi.org/10.4018/IJEBR.294110.
26. Siddiqui, S. T. ., Khan, H. ., Alam, M. I. ., Upreti, K. ., Panwar, S. ., & Hundekari, S. . (2023). A Systematic Review of the Future of Education in Perspective of Block Chain. Journal of Mobile Multimedia, 19(05), 1221–1254. https://doi.org/10.13052/jmm1550-4646.1955.
27. K. Upreti, P. Vats, G. Borkhade, R. D. Raut, S. Hundekari and J. Parashar, "An IoHT System Utilizing Smart Contracts for Machine Learning -Based Authentication," 2023 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), Windhoek, Namibia, 2023, pp. 1-6, doi: 10.1109/ETNCC59188.2023.10284960.
28. Verma, M., Upreti, K., Khan, M.R., Alam, M.S., Ghosh, S., Singh, P. (2023). Prediction of Compressive Strength of Geopolymer Concrete by Using Random Forest Algorithm. In: Shaw, R.N., Paprzycki, M., Ghosh, A. (eds) Advanced Communication and Intelligent Systems. ICACIS 2022. Communications in Computer and Information Science, vol 1749. Springer, Cham. https://doi.org/10.1007/978-3-031-25088-0_14.