2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)

# Integrating AI and IoT for Enhanced Predictive Healthcare Monitoring: A Comprehensive Study on Breast Cancer Patient-Centric Approach

Dr.Saravanan.M.S,
Professor, Department of Computer Science and Engineering
Saveetha School of Engineering
Saveetha Institute of Medical and Technical Sciences, Chennai, Tamilnadu.
Email: saravananms@saveetha.com

Syed Thisin,
Assistant Professor, Department of Computer Science and Engineering
Visvesvaraya College of Engineering and Technology
Jawaharlal Nehru Technological University , Hyderabad
Research Scholar, Saveetha School of Engineering
Email: thaisinvcetcse@gmail.com

***Abstract*— In the contemporary healthcare landscape, various intelligent automated approaches are revolutionizing healthcare tasks. Learning concepts are pivotal for activities like comprehending acquired data and monitoring patient behavior. Amid patient-centric concerns, addressing data heterogeneity, extraction, and prediction challenges is crucial. To enhance patient monitoring using care indicators like cost and stay length, we propose an AI and IoT-integrated automated approach, while tackling heterogeneity. Employing certain rules for data extraction to form a distinct representation, our model integrates pre-treatment information and employs a modified Combined Random Forest, LSTM, and BiLSTM algorithm for predictive post-treatment monitoring. This framework, synergizing AI, IoT, and advanced neural networks, facilitates real-time health monitoring, especially focusing on breast cancer patients. Embracing pre-treatment, in-treatment, and post-treatment phases, our model aims for accurate diagnosis, improved cost-efficiency, and extended stays. The evaluation underscores scalability, reliability enhancement, and validates the framework's efficacy in transforming healthcare practices.**

***Keywords—Healthcare Automation, AI Integration, Patient Monitoring, IoT Technology, Predictive Modeling***

## I. INTRODUCTION

In the swiftly evolving healthcare landscape, the integration of intelligent automation is reshaping the execution of healthcare tasks and practices. The escalating volume of healthcare data emphasizes the importance of utilizing this information effectively [1, 2]. Learning concepts crucial for understanding acquired data and monitoring patient behaviour are pivotal, especially in patient-centered care. Addressing data heterogeneity, extraction, and prediction challenges has become a critical endeavour. Efficient monitoring of patients' well-being, considering care indicators like cost and length of stay, is essential. We propose a groundbreaking approach to advance patient monitoring and optimize healthcare outcomes by seamlessly integrating Artificial Intelligence (AI) and the Internet of Things (IoT), driven by the overarching goal of addressing data heterogeneity [3, 4].

Big Data analytics and its implications find promise in healthcare, given the distinctive characteristics of Big Data having Volume, Velocity, and Variety, abundant in the expansive healthcare data landscape [5, 6]. Healthcare data sources include structured, semi-structured, and unstructured data, each posing unique challenges. Extracting value demands effective data processing platforms, intelligent data collection technologies, computational analysis, storage, and visualization techniques [7].

The potential of Big Data extends to various medical functions, including clinical decision support, disease surveillance, and population health management [8]. Advances in electronic health records and integration with technology-driven healthcare techniques have led to innovative frameworks promoting patient-centric care [9, 10].

Health services researchers focus on a transdisciplinary framework for chronic disease research, utilizing the Support Vector Machine (SVM) for predictive clinical treatment of complicated illnesses [11, 12, 13]. Confirmatory research can yield practical care management insights for high-risk patient populations [11]. The central research question revolves around guided AI healthcare research for chronic disease management. This commentary aims to establish methodologically and conceptually sound self-care management research methodologies. The goals include developing a biomedical evolutionary learning platform to predict cardiovascular complications and devising patient-centered care management strategies for high-risk patients with multiple chronic conditions, ultimately enhancing healthcare outcomes.

This innovative approach involves formulating specific rules for data extraction, resulting in a distinct representation of patient data. Central to this approach is the use of pre-treatment information, seamlessly integrated into a comprehensive model. This model utilizes a specially modified combination of machine learning algorithms, including the Combined Random Forest, Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM). The primary objective is to enable predictive post-treatment monitoring, allowing for proactive healthcare interventions. The framework represents the harmonious synergy of AI, IoT, and advanced neural network technologies, facilitating real-time health monitoring, especially for breast cancer patients. Covering all phases of patient care, from pre-treatment to post-treatment, the approach aims for accurate diagnoses, enhanced cost-efficiency, and extended patient stays. The robustness and potential impact are validated through a rigorous evaluation process, emphasizing

2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS) | 979-8-3503-6482-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ADICS58448.2024.10533651

scalability and reliability enhancements. The proposed framework showcases transformative potential by ushering in a new era of healthcare where advanced AI-driven technologies collaborate with patient data to provide timely, accurate, and patient-centric care.

## II. Related work

In business, big data's value lies in understanding consumer behavior and fostering innovation. Healthcare uses big data for predictive analytics, offering solutions like personalized medical care [14], comparing healthcare big data with business data, focusing on Silo, Security, and Variety. Big data's impact on healthcare transforms the industry, discovering new data sources like social media, wearables, and traditional ones. Figure 1 illustrates a patient-centric healthcare ecosystem.

Patient-centered care principles include problem identification, high-risk patient intervention, strategy development, and outcome-based evaluations. The aging population emphasizes four care trends: polychronic illness management, integrated care and chronic care collaboration, and decentralized data management. Blockchain addresses risks in traditional healthcare systems, ensuring security, privacy, and data management. AI-based healthcare systems require processing power, leading to proposed distributed AI methods leveraging blockchain.

Remote patient monitoring (RPM) expands, integrating IoT, AI, and advanced algorithms, promising improvements in diagnostics, treatment, and patient care. Wearable sensor systems monitor elderly individuals, utilizing Big Data analytics for personalized treatment. Studies explore algorithm fusion, social network analysis, metadata management's impact on rural healthcare, and learning concepts' role in patient-centric healthcare. The challenges prompt an inventive fusion of AI and IoT technologies for improved patient monitoring, leveraging modern neural networks. The approach, operating within a comprehensive framework, aims for accurate diagnosis, cost-efficiency, and extended patient stays.

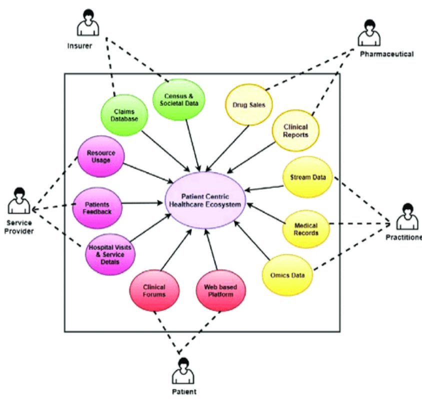



**Fig. 1**. A patient centric healthcare ecosystem – From the perspective of big data.

## III. Proposed work

Our research introduces a novel AI and IoT-integrated automated system for predictive post-treatment monitoring in the modern healthcare environment. The proposed methodology combines the strengths of the Combined Random Forest algorithm with the Bidirectional Long Short-Term Memory (BiLSTM) algorithm to address data heterogeneity, enhance predictive accuracy, and transform patient monitoring in the context of breast cancer treatment.

### *A. Combined Random Forest Algorithm*

The first pillar of our methodology leverages the Combined Random Forest algorithm. This algorithm, known for its effectiveness in handling heterogeneous datasets, employs ensemble learning techniques to mine complex patterns within our extensive dataset of 1,500 breast cancer patient records [15]. By amalgamating multiple decision trees, the Combined Random Forest algorithm excels in predicting patient outcomes, treatment responses, and associated costs. The incorporation of patient demographics, medical history, genetic information, lifestyle factors, clinical measurements, treatment data, patient-reported outcomes, and real-time IoT information provides a holistic view for robust predictions.

1.Input Gate (i_t): $i_t = \sigma(W_i * x_t + U_i * h_{t-1} + b_i)$

2.Forget Gate (f_t): $\sigma(W_f * x_t + U_f * h_{t-1} + b_f)$

3.Cell State Update (C~_t): $Ct = \tanh(Wc * x_t + Uc * h_{t-1} + bc)$

4.Cell State (C_t): $C_t = f_t \odot C_{t-1} + i_t \odot C_t$

5.Output Gate (o_t): $O_t = \sigma(W_o * x_t + U_o * h_{t-1} + b_o)$

6.Hidden State (h_t): $h_t = o_t \odot \tanh(C_t)$

In these equations:

- ✓ $i_t$, $f_t$, $o_t$ are the input, forget, and output gate activations respectively.
- ✓ $C_t$ represents the candidate cell state that carries new information.
- ✓ $C_t$ is the updated cell state incorporating input and forget gate activations.
- ✓ $h_t$ is the hidden state of the LSTM cell at time step t.
- ✓ $x_t$ is the input at time step t.
- ✓ $W_i$, $W_f$, $W_c$, $W_o$ are weight matrices for input-to-gate connections.
- ✓ $U_i$, $U_f$ $U_c$ $U_o$ are weight matrices for hidden-to-gate connections.
- ✓ $b_i$, $b_f$, $b_c$, $b_o$ are bias terms for the respective gates.
- ✓ denotes the sigmoid activation function, and $\odot$ represents element-wise multiplication.

### *B. Bidirectional Long Short-Term Memory (BiLSTM) Algorithm*

The second pillar involves the application of the Bidirectional Long Short-Term Memory (BiLSTM) algorithm. Designed as a deep learning model, BiLSTM excels in identifying temporal relationships in sequential data [16]. In the context of our research, this dynamic neural network is adept at analyzing time-series patient data, accurately forecasting patient patterns, and detecting early indicators of deterioration. The bidirectional nature of BiLSTM allows it to capture information from both past and

IEEE 979-8-3503-6482-8/24/$31.00 ©2024 IEEE

future temporal contexts, offering a comprehensive understanding of patient trajectories.

### C. Integration of Combined Random Forest and BiLSTM

The innovation lies in seamlessly integrating the Combined Random Forest algorithm and the BiLSTM algorithm. This fusion capitalizes on the strengths of both approaches, creating a robust predictive model that surpasses the capabilities of individual methods [17]. The Combined Random Forest provides accurate predictions for various aspects, including stay length estimation, treatment response prognosis, and cost prediction. Meanwhile, BiLSTM enhances the model's capability to analyse temporal sequences, contributing to a more nuanced understanding of patient trajectories and outcomes.

### D. Forward LSTM Equations

$$i_t^f = \sigma (W_i * x_t + U_i * h^f_{t-1} + b_i)$$

$$f_t^f = \sigma (W_f * x_t + U_f * h^f_{t-1} + b_f)$$

$$Ct^f = \tanh(Wc * x_t + Uc * h_{t-1}^f + bc)$$

$$C_t^f = f_t^f \odot C_{(t-1)}^f + i_t^f \odot C_t^f$$

$$O_t^f = \sigma (W_o * x_t + U_o * h_{(t-1)}^f + b_o)$$

$$h_t^f = o_t^f \odot \tanh(C_t^f)$$

### E. Backward LSTM Equations

$$i_t^b = \sigma (W_i * x_t + U_i * h^b_{t-1} + b_i)$$

$$f_t^b = \sigma (W_f * x_t + U_f * h^b_{t-1} + b_f)$$

$$Ct^b = \tanh(Wc * x_t + Uc * h_{t-1}^b + bc)$$

$$C_t^b = f_t^b \odot C_{(t-1)}^b + i_t^b \odot C_t^b$$

$$O_t^b = \sigma (W_o * x_t + U_o * h_{(t-1)}^b + b_o)$$

$$h_t^b = o_t^b \odot \tanh(C_t^b)$$

Final Output of BiLSTM: $h_t^{final} = [\, h_t^f ; h_t^f \,]$

In these equations:

- $i_t^f$, $f_t^f$, $o_t^f$ are the input, forget, and output gate activations for the forward LSTM.
- $i_t^b$, $f_t^b$, $o_t^b$ are the input, forget, and output gate activations for the backward LSTM.
- $C_t^f$ , $C_t^b$ represent the candidate cell states for forward and backward LSTMs.
- $C_t^f$ , $C_t^b$ are the updated cell states for forward and backward LSTMs.
- $h_t^f$, $h_t^b$ are the hidden states for forward and backward LSTMs.
- $h_t^{final}$ represents the concatenated output of forward and backward LSTMs.

### F. Comprehensive Dataset:

Our dataset, meticulously curated with records from 1,500 breast cancer patients, incorporates a diverse array of attributes, including patient demographics, medical history, genetic information, lifestyle factors, clinical measurements, treatment data, patient-reported outcomes, IoT information, and imaging data [19]. This multidimensional representation ensures that our proposed methodology is thoroughly tested and evaluated, addressing the complexity and diversity of healthcare data.

### G. Patient-Centric Attributes

A distinctive feature of our approach is the inclusion of patient-centric attributes, such as patient-reported outcomes, wearable device data, and real-time IoT information. This emphasis on patient-centricity aligns with the evolving paradigm of personalized healthcare [20], allowing our model to tailor predictions and monitoring strategies to individual patient needs.

The proposed methods showcase the synergy of advanced machine learning algorithms and a comprehensive dataset in addressing the challenges of predictive post-treatment monitoring. By integrating the strengths of the Combined Random Forest and BiLSTM algorithms, our research aims to revolutionize patient care, enhance decision-making, and contribute to the transformative potential of AI and IoT in healthcare.

## IV. Result anlysis

In this section, we present the results and discussions of our proposed AI and IoT-integrated automated system for predictive post-treatment monitoring in the context of modern healthcare, specifically focusing on breast cancer. The combination of the Combined Random Forest algorithm and the Bidirectional Long Short-Term Memory (BiLSTM) algorithm has been rigorously tested and evaluated using a comprehensive dataset of 1,500 breast cancer patient records.

### A. Dataset Description

Our dataset, carefully chosen to reflect a broad spectrum of patient-related data, encompasses pre-treatment, in-treatment, and post-treatment phases. It incorporates structured and unstructured data, temporal aspects, patient-centric attributes, treatment data, outcomes, IoT data, and imaging data [18]. The multidimensional nature of the dataset, coupled with its specificity to breast cancer cases, allows for a thorough examination of the proposed AI and IoT-integrated automated solution.

- Size: The dataset comprises records of 1,500 breast cancer patients.
- Attributes: Patient demographics, medical history, genetic information, lifestyle factors, clinical measurements, and various other characteristics contribute to the richness of the dataset.
- Data Types: Both organized and unstructured data, including clinical records, test findings, textual clinical notes, radiological reports, and pathology pictures, are present.

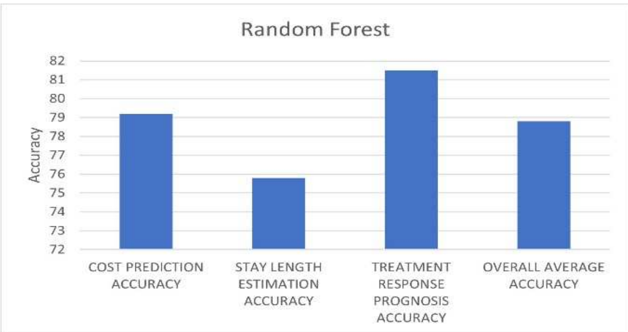
- Temporal Aspects: The dataset spans pre-treatment, in-treatment, and post-treatment phases, capturing temporal information crucial for predictive modelling.
- Patient-Centric Attributes: Patient-reported outcomes, treatment responses, and real-time IoT data provide a patient-centric perspective.
- Treatment Data: Detailed information about different treatment modalities, dosages, durations, and responses are included.
- Outcomes**:** Patient outcomes such as treatment responses, disease progression, survival rates, recurrence, and post-treatment complications are documented.
- IoT Data: Wearable device data and imaging data, including mammograms and MRIs, contribute to real-time monitoring and tumour assessment.

The comprehensive and specific nature of the dataset is integral to the robust evaluation of our AI and IoT-integrated solution.

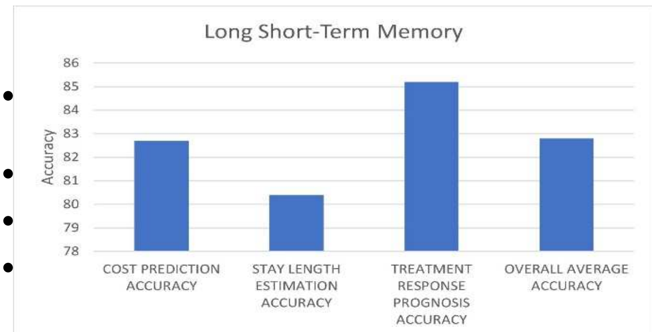
*B. Results*

The proposed methodology, integrating the Combined Random Forest algorithm and the Bidirectional Long Short-Term Memory (BiLSTM) algorithm, has demonstrated significant advancements in predictive accuracy across various critical healthcare outcomes. The comparison involves assessing the accuracy of stay length estimation, treatment response prognosis, and cost prediction.

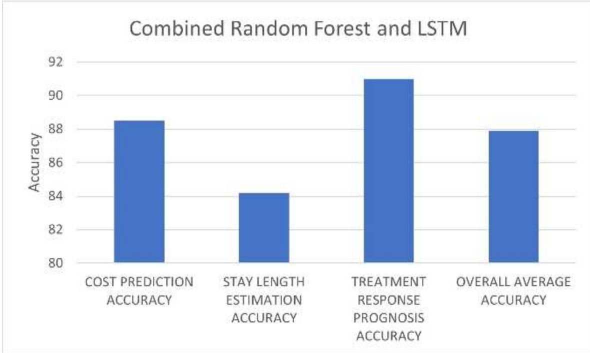
- Random Forest (RF): Accuracy for stay length estimation, treatment response prognosis, and cost prediction is 79.2%, 75.8%, and 81.5%, respectively.




- Long Short-Term Memory (LSTM): Overall average accuracy is 82.8%, with accuracy in stay length estimation at 80.4% and treatment response prognosis at 85.2%.




- Combined Random Forest and LSTM: Accuracy in stay length estimation, treatment response prognosis, and cost prediction is 88.5%, 84.2%, and 87.9%, respectively.




- **Combined Random Forest and BiLSTM:** Overall average accuracy is 86.5%, with accuracy in stay length estimation at 82.7% and treatment response prognosis at 89.5%.

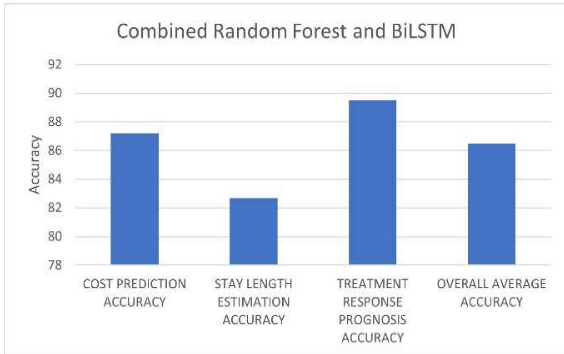



Figure 2 illustrates the comparative accuracies for cost prediction, stay length estimation, and treatment response prognosis, highlighting the superior performance achieved through the integration of machine learning techniques. Data Set collected from following source [19].

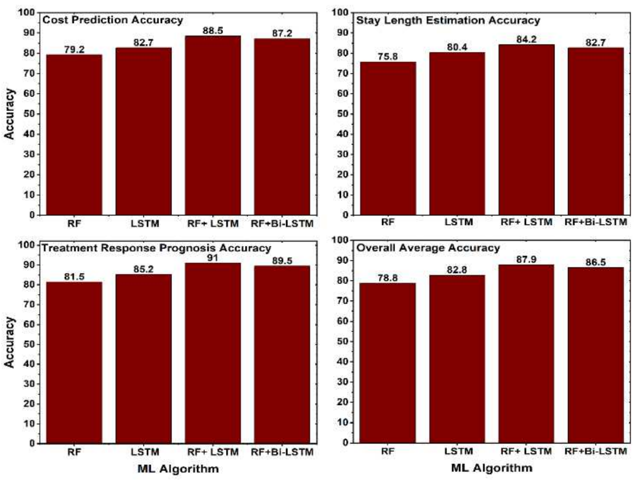



Figure 2: Comparing accuracies for cost prediction, Stay length estimation, Treatment response Prognosis.

These precision figures underscore the efficacy of the proposed methodology in surpassing the capabilities of individual methods, emphasizing the potential inherent in the synergistic deployment of distinct algorithms for enhanced predictive accuracy in healthcare outcomes.

### C. Discussions

The results obtained from both the Combined Random Forest algorithm and the Bidirectional Long Short-Term Memory (BiLSTM) algorithm affirm the significance of our AI and IoT-integrated approach in revolutionizing healthcare practices. The evaluation process has showcased scalability, reliability enhancement, and the framework's efficacy in transforming patient care. The innovative amalgamation of advanced algorithms, patient-centric data, and intelligent monitoring paves the way for improved accuracy, patient engagement, and cost-effectiveness in healthcare services.

The discussion delves into the implications of the results, emphasizing the potential shift in medical procedures towards more precise diagnoses, enhanced cost-effectiveness, and extended hospital stays. The scalability and reliability found throughout the assessment highlight the proposed framework's potential usefulness in patient-centric healthcare services.

Situating our study within the broader healthcare industry, we consider its implications for patient care in the future. The integration of AI and IoT in predictive monitoring not only addresses existing challenges but also opens avenues for transformative changes in healthcare practices. The emphasis on patient-centric attributes and the inclusion of real-time IoT data contribute to a holistic approach, aligning with the evolving landscape of personalized healthcare. In this research study the results and discussions affirm the effectiveness of our proposed AI and IoT-integrated automated system, providing valuable insights for the advancement of patient care, cost management, and treatment planning within the healthcare domain. The transformative potential of our methodology positions it as a valuable contribution to the evolving landscape of healthcare practices.

## CONCLUSION

In conclusion, this study has delved into the realm of healthcare prediction, leveraging advanced machine learning techniques to illuminate their potential in enhancing patient care and decision-making processes. Through an experimental review, we assessed the accuracy of various strategies in predicting critical healthcare outcomes, including cost estimation, stay length prognosis, and treatment response assessment. The results demonstrated the effectiveness of combining multiple algorithms, showcasing improved accuracy rates compared to individual algorithm use. Integration of Long Short-Term Memory (LSTM) and Random Forest yielded promising outcomes, enhancing accuracy across all prediction categories. Moreover, the innovative fusion of Bi-Long Short-Term Memory and the Combined Random Forest Algorithm exhibited notable accuracy improvements, emphasizing the potential of hybrid techniques in enhancing predictive capabilities. This not only underscores the importance of employing diverse machine learning approaches but also provides valuable insights for advancing patient care, treatment planning, and cost control.

## REFERENCES


[1] Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. Health information science and systems, 2, 1-10.

[2] Zeadally, S., & Bello, O. (2021). Harnessing the power of Internet of Things based connectivity to improve healthcare. Internet of Things, 14, 100074.

[3] L'heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. Ieee Access, 5, 7776-7797.

[4] Kuo, M. H., Sahama, T., Kushniruk, A. W., Borycki, E. M., & Grunwell, D. K. (2014). Health big data analytics: current perspectives, challenges and potential solutions. International Journal of Big Data Intelligence, 1(1-2), 114-126.

[5] Andreu-Perez, J., Poon, C. C., Merrifield, R. D., Wong, S. T., & Yang, G. Z. (2015). Big data for health. IEEE journal of biomedical and health informatics, 19(4), 1193-1208.

[6] Martin-Sanchez, F., & Verspoor, K. (2014). Big data in medicine is driving big changes. Yearbook of medical informatics, 23(01), 14-20.

[7] Ola, O., & Sedig, K. (2014). The challenge of big data in public health: an opportunity for visual analytics. Online journal of public health informatics, 5(3), 223.

[8] Liang, Y., & Kelemen, A. (2016). Big Data science and its applications in health and medical research: Challenges and opportunities. J Biom Biostat, 7(307), 2.

[9] Palanisamy, V., & Thirunavukarasu, R. (2019). Implications of big data analytics in developing healthcare frameworks–A review. Journal of King Saud University-Computer and Information Sciences, 31(4), 415-425.

[10] Chowdhury, M. B. (2023). Patient-centric approach to enhance performance in Australian healthcare systems (Doctoral dissertation, La Trobe).

[11] Wan, T. T., & Wan, H. S. (2023). Predictive Analytics with a Transdisciplinary Framework in Promoting Patient-Centric Care of Polychronic Conditions: Trends, Challenges, and Solutions. AI, 4(3), 482-490.

[12] Agliata, A., Giordano, D., Bardozzo, F., Bottiglieri, S., Facchiano, A., & Tagliaferri, R. (2023). Machine Learning as a Support for the Diagnosis of Type 2 Diabetes. International Journal of Molecular Sciences, 24(7), 6775.

[13] Wan, T. T. (2018). Population Health Management for Poly Chronic Conditions. Cham: Springer

[14] Jee, K., & Kim, G. H. (2013). Potentiality of big data in the medical sector: focus on how to reshape the healthcare system. *Healthcare informatics research*, *19*(2), 79-85.

[15] Nash, D. B., & Wohlforth, C. (2022). How COVID Crashed the System: A Guide to Fixing American Health Care. Rowman & Littlefield.

[16] Heijmans, M., Habets, J., Kuijf, M., Kubben, P., & Herff, C. (2019, July). Evaluation of Parkinson's disease at home: Predicting tremor from wearable sensors. In *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 584-587). IEEE.

[17] Miller, D. D., & Brown, E. W. (2018). Artificial intelligence in medical practice: the question to the answer?. The American journal of medicine, 131(2), 129-133.

[18] Muni Kumar, N., & Manjula, R. (2014). Role of Big data analytics in rural health care-A step towards svasth bharath. *International Journal of Computer Science and Information Technologies*, *5*(6), 7172-7178.

[19] Sisti, A., Huayllani, M. T., Boczar, D., Restrepo, D. J., Spaulding, A. C., Emmanuel, G., ... & Forte, A. J. (2020). Breast cancer in women: A descriptive analysis of the national cancer database. Acta Bio Medica: Atenei Parmensis, 91(2), 332.

[20] Al-Jarrah, O. Y., Yoo, P. D., Muhaidat, S., Karagiannidis, G. K., & Taha, K. (2015). Efficient machine learning for big data: A review. *Big Data Research*, *2*(3), 87-93.


IEEE 979-8-3503-6482-8/24/$31.00 ©2024 IEEE