e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

DESIGN AND IMPLEMENTATION OF EARLY SEPSISDETECTION SYSTEM USING MACHINE LEARNING

Ms. Prathibha BS[*1], Girija Patil[*2], Bharat P[*3], Manish MP[*4], Dayanand S[*5]

[*1]Assistant Professor Information Science And Engineering The Ntaional Institute Of Engineering Mysuru, Karnataka, India.

[*2,3,4,5]Information Science And Engineering The National Institute Of Engineering Mysuru, Karnataka, India.

ABSTRACT

Sepsis remains a critical challenge in healthcare, necessitating timely detection and intervention to mitigate itslife- threatening consequences. This project aims to revolutionize sepsis detection by use of real-time machine learning algorithms. Through the utilization of sophisticated techniques including Random Forest, Naive Bayes, and Decision Trees, the study aims to develop a dynamic and accurate system for early sepsis pre- diction. By analyzing diverse patient data sources encompassing vital signs and clinical history, the project prioritizes the early identification of sepsis, aiming to facilitate prompt interventions and improve patient outcomes.

The research aligns with the Third International Consensus on sepsis-3 criteria, emphasizing the importance of evolving diagnos- tic accuracy. By leveraging comprehensive datasets and advanced algorithms, the project seeks to overcome the limitations of existing sepsis detection methodologies, such as limited sensitivity and delayed interventions. The system's objectives encompass the creation of a user-friendly interface for healthcare professionals, enabling seamless integration into existing clinical workflows.

Furthermore, the project's impact extends beyond immediate cost-effectiveness, aiming to enhance overall healthcare efficiency and patient care. Through optimized resource allocation and improved diagnostic speed, the system aims to streamline health- care processes and contribute to long-term cost savings. This research holds promise for transforming healthcare practices, improving patient outcomes, and establishing a robust frameworkfor efficient sepsis detection in clinical settings.

Keywords: Sepsis, Detection, Machine Learning.

I. INTRODUCTION

Sepsis remains a significant challenge in contemporary healthcare, demanding prompt intervention to mitigate its severe outcomes. Despite medical progress, early sepsis de- tection continues to present hurdles due to its intricate clinical presentation and the absence of highly sensitive diagnostic tools. The integration of machine learning (ML) presents a promising avenue for enhancing sepsis detection by employing algorithms capable of analyzing extensive patient data to identify nuanced patterns indicative of the condition. This research endeavors to exploit the potential of ML to develop a real-time predictive model for early sepsis detection, with the overarching goal of enhancing clinical outcomes andhealthcare efficiency.

Machine learning algorithms, including Random Forest, Naive Bayes, and Decision Trees, help in sepsis detection by providing a dynamic and data-driven approach. By amalga- mating diverse patient data, such as vital signs, laboratory results, and clinical history, these algorithms have the po- tential to identify subtle indicators of sepsis that may evade traditional diagnostic methods. Through iterative learning and optimization, the predictive model aims to prioritize the early identification of sepsis, thereby enabling timely interventions and improving patient outcomes and mitigating risk of fatality. Aligned with the paradigm shift outlined in the Third International Consensus on sepsis-3 criteria, this research underscores the imperative of advancing diagnostic accuracy through ML techniques. By leveraging comprehensive datasets and sophisticated algorithms, the project endeavors to sur- mount the limitations of existing detection methodologies, such as limited sensitivity and delayed interventions. More- over, the development of a user-friendly interface tailored for healthcare professionals and users will facilitate seamless integration into clinical workflows, empowering practitioners with actionable insights derived from ML-driven analytics.

II. RELATED WORK

[1] Ruiqing Ding; Yu Zhou; Jie Xu; Yan Xie; Qiqiang Liang; He Ren; Yixuan Wang; Yanlin Chen Cross-Hospital Sepsis Early Detection via Semi-Supervised Optimal Transport With Self-Paced Ensemble IEEE Journal of Biomedical and Health Informatics DOI: 10.1109/JBHI.2023.3253208 : Leveraging machine learning techniques for Sepsis early detection and di- agnosis has attracted increasing interest in recent years. How- ever, most existing methods require a large amount of labeled training data, which may not be available for a target hospital that deploys a new Sepsis detection system. More seriously, as treated patients are diversified between hospitals, directly applying a model trained on other hospitals may not achieve good performance for the target hospital. Which can efficiently transfer knowledge from the source hospital (with rich la- beled data) to the target hospital (with scarce labeled data). Specifically, SPSSOT incorporates a new optimal transport- based semi-supervised domain adaptation component that can effectively exploit all the unlabeled data in the target hospital. Moreover, self-paced ensemble is adapted in SPSSOT to alleviate the class imbalance issue during transfer learning. In a nutshell, SPSSOT is an end-to-end transfer learning method that automatically selects suitable samples from two domains (hospitals) respectively and aligns their feature spaces. Exten- sive experiments on two open clinical datasets, MIMIC-III and Challenge, demonstrate that SPSSOT outperforms state-of-the- art transfer learning methods by improving 1–3% of AUC. [2] Mengsha Fu; Jiabin Yuan; Menglin Lu; Pengfei Hong; Mei ZengAn Ensemble Machine Learning Model For the Early Detection of Sepsis From Clinical Data 2019 Computing in Cardiology (CinC) DOI: 10.22489/CinC.2019.317 : Sepsis is a life-threatening disease with high mortality and expensive cost of treatment. In order to improve the outcomes of patients, it is important to detect at-risk patients with sepsis at an early stage. An ensemble model, which combined boosting and bagging tree models (lightgbm, xgboost and random forest) were designed to predict sepsis based on the records of the patient's hourly data. We compared the ensemble model and each single model of evaluation metrics results on selected inner test data Offline, the best performance was achieved AUC of 0.792, ACC of 0.727. Finally, the proposed model was evaluated on the full test sets received an official utility score, defined by the organizers, was 0.087 ranked 75/105. While the single model of lightgbm only received a utility score of -0.036. The ensemble model utilized the preprocessing data and achieved better performance than a single tree-based model [3]Teh Xuan Ying; Asma Abu-Samah Early Prediction of Sepsis for ICU Patients using Gradient Boosted Tree 2022 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) DOI: 10.1109/I2CACIS54679.2022.9815467 In-tensive care unit patients, especially those who have undergone surgeries or have severe health issues, tend to have a higher risk of developing sepsis due to a weaker immune system. Due to late detection of sepsis, no preventive actions can be taken to treat sepsis patients. Therefore, this research aims to identify, validate, and test suitable machine learning algorithms for the early prediction of sepsis using pre-processed data produced from the Medical Information Mart for Intensive Care III, MIMIC-III database. This research will be designing prediction models for 15 hours before sepsis onset using pre- processed data obtained from MIMIC-III database using Decision Tree, Random Forest, AdaBoost, Gradient Boosted Tree, and Multilayer Perceptron. A 10 cross-validation is used in validating the models. The performance of prediction models is evaluated mainly using ROC-AUC score. In model comparison, an extra set of prediction models using the same algorithms is developed for 10 hours before sepsis onset to compare its performance with the earlier prediction model developed. The result of model comparison shows that for the prediction model of 15 and 10 hours before sepsis onset, ROC- AUC score for Gradient Boosted Tree is the best with 0.777 for 15 hours and 0.769 respectively from 10 hours prediction model. The results can be optimized further using more data and using derived Boosted Trees algoritms.

III. PROPOSED SYSTEM

The proposed system constitutes a real-time application designed to serve hospitals, catering to the needs of administra- tors, doctors, and patients alike. Its primary objective revolves around the early detection of sepsis utilizing machine learning (ML) algorithms. Leveraging efficient algorithms such as Decision Trees, Random Forest, and Na¨ıve Bayes, the system is adept at handling numerical data, ensuring swift execution speeds, and harnessing the power of classification effectively. By enabling timely intervention, reducing manual analysis costs, and providing a valuable tool for healthcare profession- als, the proposed solution aims to significantly impact sepsis management in hospital settings.

The integration of machine learning algorithms within the proposed system marks a notable advancement, offering a dynamic and data-driven approach to enhance sensitivity and specificity in identifying subtle indicators of sepsis. These algorithms operate on extensive datasets, enabling them to discern intricate patterns and anomalies that might evade traditional diagnostic methods. Moreover, the incorporation of adaptive algorithms such as Random Forest and K-Nearest Neighbors (KNN) introduces ongoing optimization to the system. Continuously learning from fresh data and adapting to diverse patient profiles, these algorithms ensure consistent high accuracy and reliability in early sepsis detection. This adaptability and optimization significantly enhance the sys- tem's robustness, ensuring its effectiveness across a wide range of patient scenarios and cases.

IV. SYSTEM DESIGN

The system design of this project revolves around a three- tier architecture, comprising the data layer, business layer, and presentation layer. The data layer is responsible for storing and retrieving data from the database, while the business layer contains the core logic of the application, including data processing and prediction algorithms. Finally, the presenta- tion layer provides the user interface elements, facilitating interaction between users and the system. By adopting this architecture, the project ensures modularity, scalability, and separation of concerns, facilitating easier maintenance andfuture enhancements.

In addition to the core functionality, the system design also encompasses features aimed at enhancing usability and scalability. Utilizing a modular approach, the system allows for easy integration of future enhancements and updates, ensuring adaptability to evolving healthcare needs and technological ad- vancements. Furthermore, the system's architecture prioritizes security and data integrity, implementing robust authentica- tion mechanisms and data encryption protocols to safeguard sensitive patient information. By adhering to best practices in system design and development, the project not only addresses the immediate challenge of sepsis prediction but also lays the foundation for a versatile and resilient healthcare solution with potential for widespread impact in clinical settings.

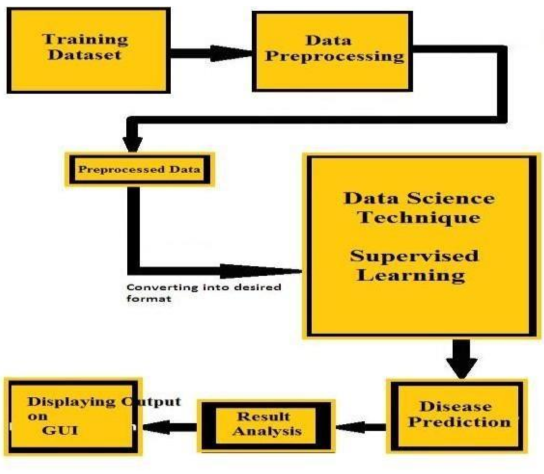

Fig. 1. System Architecture

Fig. 2. WorkFlow

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

The system's workflow is delineated through a comprehensive flow diagram, illustrating the interaction between users and modules. Admin modules encompass functions such as user management, discussion forum moderation, and dataset management, while member/doctor modules include user registration, disease prediction, and dataset management. This modular approach facilitates ease of use and navigation for users, ensuring seamless interaction with the system and fostering collaboration among healthcare professionals.

V. SYSTEM IMPLEMENTATION

The implementation of the proposed system harnesses the capabilities of C .NET and SQL to create a robust and efficient platform for sepsis identification. Leveraging C .NET, the system's business logic layer is meticulously designed, encap- sulating data processing and algorithmic functionalities. SQL serves as the backbone of the system, providing a reliable and scalable database infrastructure for seamless data storage and retrieval. Additionally, the system integrates machine learning (ML) techniques, including Naive Bayes, Random Forest, and Decision Tree algorithms, to develop predictive models for early sepsis detection. Through the synergistic utilization of C.NET, SQL, and ML algorithms, the system aims to empower healthcare professionals with a comprehensive toolset for timely intervention and improved patient care.

This project adopts a three-tier architecture, comprising the presentation layer, business logic layer, and data layer, to ensure modularity and scalability. The presentation layer, built using ASP.NET, provides the user interface elements, facilitating user interaction and navigation. Meanwhile, the business logic layer, implemented with C classes, contains the core functionalities of the system, including data processing and prediction algorithms. The data layer, powered by MS SQL Server 2005, serves as the backend database, enabling efficient data storage and retrieval. The integration of these layers enables seamless communication and interaction be- tween different components of the system, ensuring smooth operation and robust performance.

The system implementation follows a structured approach, encompassing various functional prototypes and implementation steps. The presentation layer invokes the business logic layer through user interactions, initiating processes such as data retrieval and prediction. The business logic layer, in turn, interacts with the database using table adapters and SQL data sources, executing operations such as data insertion and retrieval. Through comprehensive testing and validation procedures, the system ensures accuracy, reliability, and adherence to functional requirements. Overall, the implementation of the proposed system represents a culmination of advanced tech- nologies and methodologies, aimed at delivering an effective and user-friendly solution for sepsis identification in healthcaresettings.

VI. ALGORITHM IMPLEMENTATION

In our research focused on sepsis prediction, we have implemented a machine learning framework employing super- vised learning algorithms such as the Naive Bayes classifier, Decision Tree, and Random Forest. These algorithms were selected due to their adaptability and efficiency across datasets of varying sizes and complexities, which are essential for accurate predictions in the context of time-critical medical scenarios. The Naive Bayes algorithm, for instance, starts by scanning the dataset and computing probabilities for each at- tribute value, followed by classification based on these probabilities. Similarly, the Decision Tree algorithm involves dataset scanning, gain calculation for attributes, and outcome determination through recursive segmentation and classification. Conversely, the Random Forest algorithm utilizes parameter splitting, Gini index calculation, and outcome determination to classify instances effectively.

Assessing the performance of these algorithms is crucial for evaluating their efficacy in sepsis prediction. According to our evaluations, the Naive Bayes algorithm demonstrates an impressive accuracy of 96.5 showcasing its proficiency in accurately identifying sepsis occurrences. Additionally, the Deci- sion Tree algorithm exhibits a commendable accuracy of 89.7, providing reliable predictions based on dataset characteristics and feature selection. Despite a slightly lower accuracy rate of 76.1, the Random Forest algorithm still offers valuable insights into sepsis prediction. Despite variations in accuracy rates, each algorithm contributes to the overall effectiveness of the predictive model, catering to different dataset complexities and prediction requirements. Through comprehensive algorithm implementation and accuracy assessment, our research aims to advance sepsis prediction, offering healthcare professionals a robust tool for early detection and intervention.

VII. CONCLUSION

In conclusion, our research underscores the significance of machine learning algorithms in enhancing sepsis prediction, a critical aspect of early intervention and patient care in health- care. Through the adoption of supervised learning techniques like the Naive Bayes classifier, Decision Tree, and Random Forest, we have demonstrated the feasibility of accurate sepsis identification based on patient parameters. The impressive accuracies achieved by these algorithms validate their efficacy in handling diverse datasets and providing reliable predictions. Moreover, the comprehensive evaluation of algorithm performance highlights their distinct strengths and suitability for different prediction scenarios.

Moving forward, our research paves the way for further advancements in sepsis prediction systems, offering healthcare professionals valuable insights into patient risk assessment and intervention strategies. By leveraging the power of machine learning and data-driven approaches, we aim to contribute to improved patient outcomes and healthcare delivery practices. Additionally, our findings emphasize the importance of continuous evaluation and refinement of predictive models to ensure their effectiveness in real-world clinical settings. Ultimately, our research endeavors to bridge the gap between data science and clinical practice, fostering innovation and excellence in sepsis management.

VIII. REFERENCES

[1] Cross-Hospital Sepsis Early Detection via Semi-Supervised Optimal Transport With Self-Paced Ensemble IEEE Journal of Biomedical and Health Informatics DOI: 10.1109/JBHI.2023.3253208 Authors: Ruiqing Ding; Yu Zhou; Jie Xu; Yan Xie; Qiqiang Liang; He Ren; Yixuan Wang; Yanlin Chen;

[2] Early Prediction of Sepsis for ICU Patients using Gradient Boosted Tree 2022 IEEE International Conference on Automatic Control and Intelli- gent Systems (I2CACIS)
DOI: 10.1109/I2CACIS54679.2022.9815467 Authors: Teh Xuan Ying; Asma Abu-Samah

[3] S. K. Gadre, M. Shah, E. Mireles-Cabodevila, B. Patel, and A. Duggal, "Epidemiology and Predictors of 30-Day Readmission in Patients With Sepsis," Chest, 155(3), 2019, pp.483-490.

[4] A. Keeley, P. Hine and E. Nsutebu, "The recognition and management of sepsis and septic shock: a guide for non-intensivists," Postgraduate Medical Journal. 93, 2017, pp.626-634

[5] P. E. Marik and A. M. Taeb, "SIRS, qSOFA and new sepsis definition," Journal of thoracic disease, 9(4), 2017, p.943.

[6] Schonfeld, D. and R. Paul, Sepsis: an Update on Current Improvement Efforts. Current Treatment Options in Pediatrics, 2020. 6(4): p.

[7] Morrill, J.H., et al., Utilization of the Signature Method to Identify the Early Onset of Sepsis From Multivariate Physiological Time Series in Critical Care Monitoring. Crit Care Med, 2020. 48(10): p. e976- e981