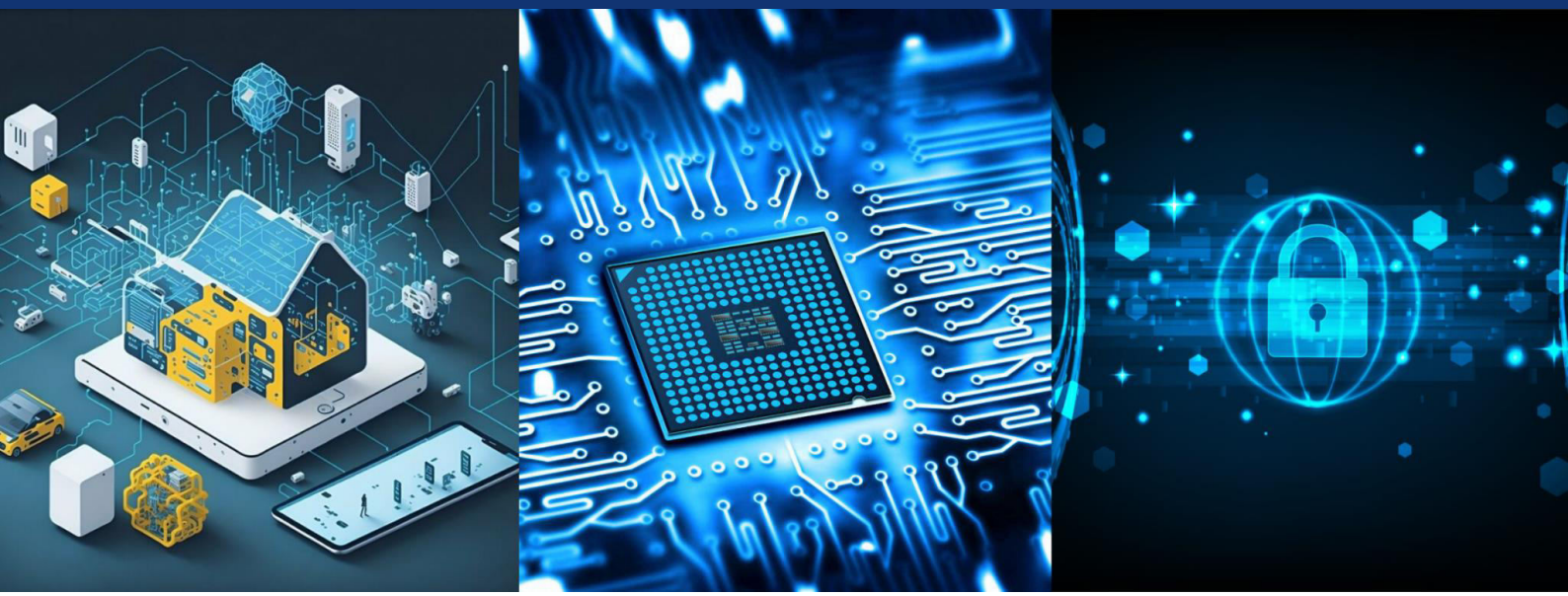
IJIRCCE

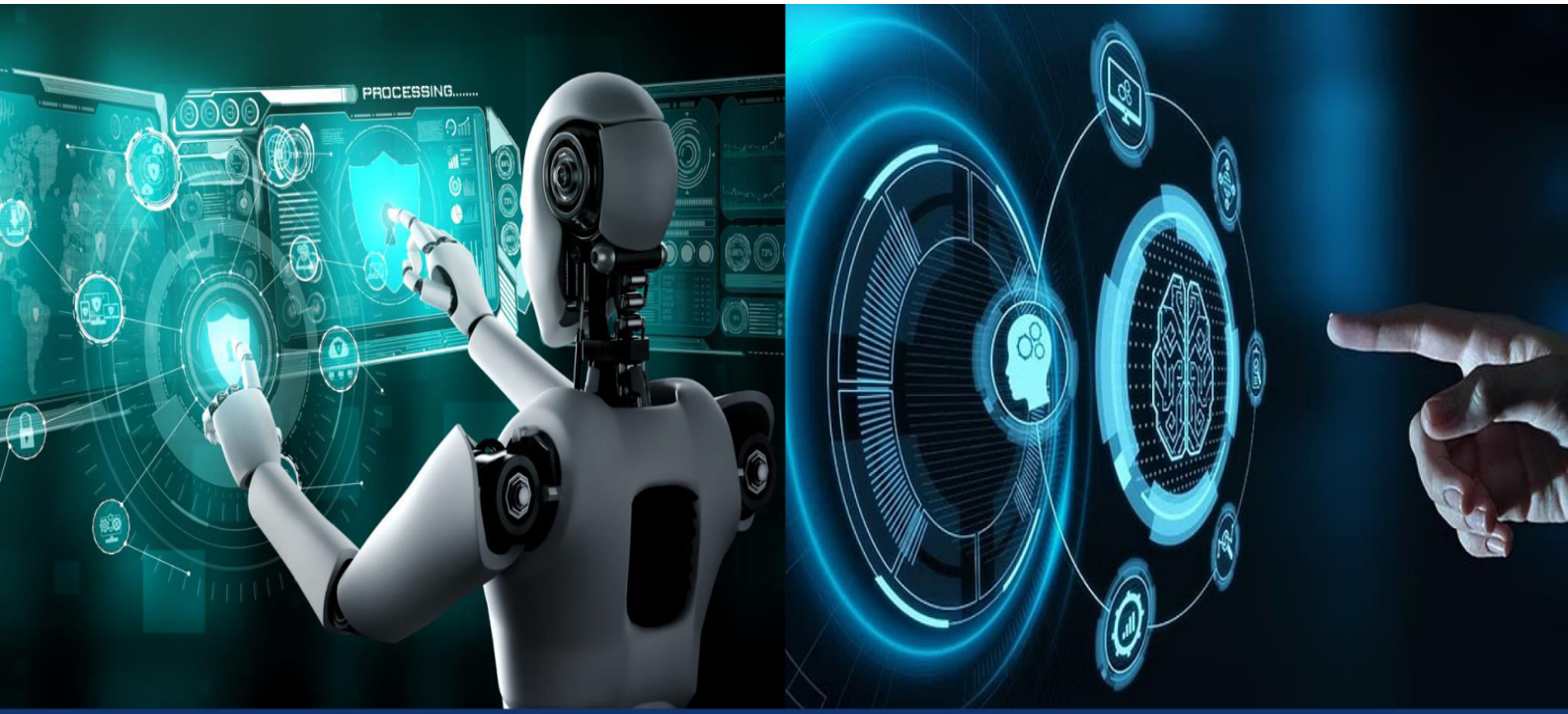
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 13, Issue 3, March 2025

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303116

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sepsis Detection Using Machine Learning

Vishnu Sowbaghya C, Dr.V.Sangeetha

III-B.Sc., Department of Computer Science with Data Analytics, Dr.N.G.P.Arts and Science College, Coimbatore, India

Assistant Professor, Department of Computer Science with Data Analytics, Dr.N.G.P.Arts and Science College, Coimbatore, India

ABSTRACT: Sepsis is a life-threatening medical condition caused by the body's extreme response to an infection, leading to organ failure and potentially death if not detected early. Traditional diagnostic methods often result in delayed identification, increasing mortality risks. This project leverages Machine Learning (ML) techniques, specifically the Random Forest algorithm, to develop an early detection system for sepsis. The model is trained on patient health data, including vital signs and laboratory test results, to predict sepsis onset accurately.

Key steps include data preprocessing, feature selection, and model training to ensure optimal performance. The trained model is evaluated using accuracy, precision, recall, F1-score, and AUC-ROC to measure its effectiveness. The system is then deployed using Python and Flask, enabling real-time predictions in a healthcare setting. By implementing this ML-based approach, the project aims to assist medical professionals in making faster and more accurate sepsis diagnoses, ultimately improving patient outcomes and reducing fatalities.

KEYWORDS: Sepsis Detection, Machine Learning (ML), Random Forest Algorithm Early Diagnosis, Medical Data Analysis

I. INTRODUCTION

Sepsis is a critical and life-threatening medical emergency that arises when the body's response to an infection triggers widespread inflammation, leading to tissue damage, organ failure, and potentially death. It is a major global health concern, with millions of cases reported annually. The key to improving patient survival rates is **early** detection and timely intervention. However, traditional diagnostic methods rely on clinical observations and laboratory tests, which can be time-consuming, leading to delays in treatment and increased fatality risks.

To address this challenge, this project aims to develop a Machine Learning (ML)- based sepsis detection system that can analyze patient health data to predict sepsis onset at an early stage. By leveraging advanced ML techniques, specifically the Random Forest algorithm, the model will classify patients at risk with high accuracy. The system will be trained on a dataset containing vital signs, laboratory test results, and other clinical parameters to identify significant patterns and risk factors associated with sepsis.

The core objectives of this project include data preprocessing, feature selection, model training, evaluation, and deployment to ensure optimal performance. The final model will be integrated into a Flask-based web application, enabling real-time predictions to assist healthcare professionals in making faster and more reliable clinical decisions. This research aims to demonstrate how machine learning can revolutionize early disease detection, ultimately improving patient outcomes and reducing sepsis-related mortality.

II. RELATED WORK

Various studies have investigated the use of machine learning techniques for early sepsis detection to improve diagnosis and reduce mortality rates. Traditional diagnostic methods rely on clinical observations and laboratory tests, which can lead to delays in treatment. To overcome these challenges, machine learning models such as Random Forest, Support Vector Machines (SVM), and XGBoost have been employed to predict sepsis based on patient data.

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303116

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Deep learning models, including Long Short-Term Memory (LSTM) networks, have also been explored for analyzing time-series medical data, enabling continuous patient monitoring. Additionally, feature selection techniques have been applied to identify key biomarkers such as heart rate, blood pressure, and white blood cell count, improving model accuracy. However, challenges such as data imbalance, lack of interpretability, and real-time implementation remain significant barriers.

This project builds upon previous research by utilizing Random Forest for better accuracy and explainability while integrating a Flask-based web application for real- time sepsis prediction. By addressing existing limitations, the proposed system aims to assist medical professionals in making faster and more accurate diagnoses.

III. METHODOLOGY

3.1 Dataset

For this study, we use a publicly available sepsis dataset, which includes patient health records containing various physiological parameters. Key features in the dataset include:

- **Vital signs:** Heart rate, blood pressure, respiratory rate, and temperature.
- **Laboratory test results:** White blood cell count, lactate levels, and other biomarkers.
- **Sepsis label:** Indicates whether a patient has sepsis (1) or not (0).

3.2 Data Preprocessing

Data preprocessing is crucial to improve model accuracy and reliability. The following steps were applied:

- **Handling missing values:** Imputing or removing records with incomplete critical features.
- **Feature scaling and normalization:** Standardizing numerical values to ensure uniformity.
- **Feature selection:** Identifying the most relevant parameters for sepsis detection.

3.3 Machine Learning Model

The Random Forest algorithm is used for sepsis prediction due to its robustness and ability to handle high-dimensional medical data. The model is trained using patient health features to distinguish between sepsis and non-sepsis cases.

3.4 Evaluation Metrics

Model performance is assessed using the following metrics:

- **Accuracy:** Measures overall correctness of predictions.
- **Precision:** The proportion of correctly predicted sepsis cases.
- **Recall:** The ability of the model to identify sepsis-positive cases.
- **F1-Score:** A balance between precision and recall.
- **AUC-ROC Curve:** Evaluates the model's effectiveness in distinguishing between sepsis and non-sepsis cases.

IV. EXPERIMENTAL SETUP

4.1 Tools and Libraries

The following tools were used for model implementation:

- **Python:** Programming language for machine learning model development.
- **Libraries:** Pandas, NumPy, Scikit-Learn, Matplotlib (for data analysis and visualization).
- **Flask:** For deploying the model as a web-based application.
- **Google Colab:** Development environments for training and testing.

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303116

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4.2 Model Training & Testing

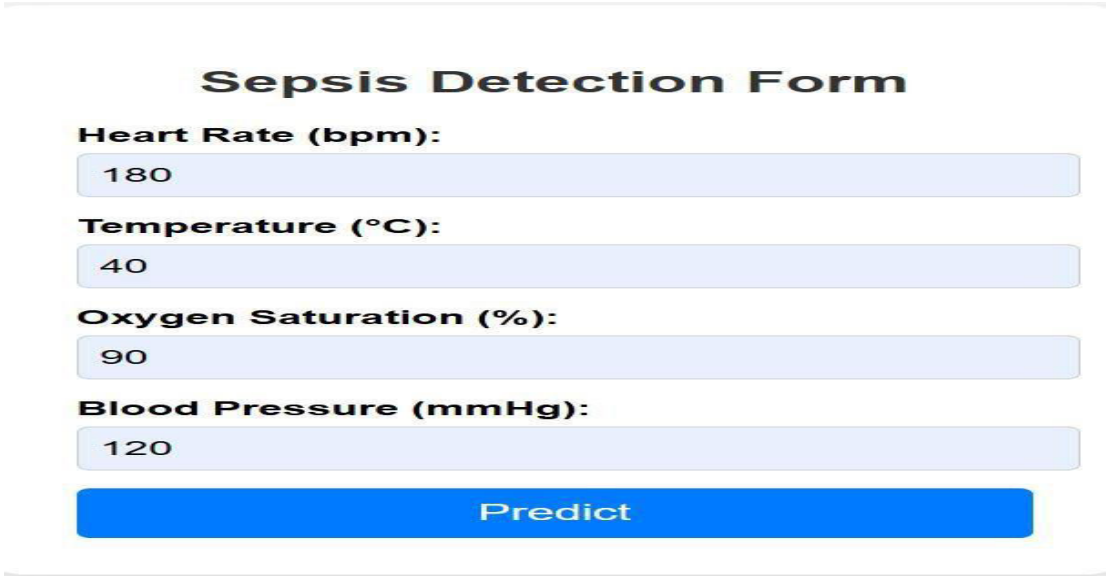
The dataset was split into 80% training and 20% testing to evaluate model performance. Hyperparameter tuning was performed to optimize the accuracy and reliability of the model. The final trained model is deployed using Flask, enabling real-time sepsis predictions in a clinical setting.

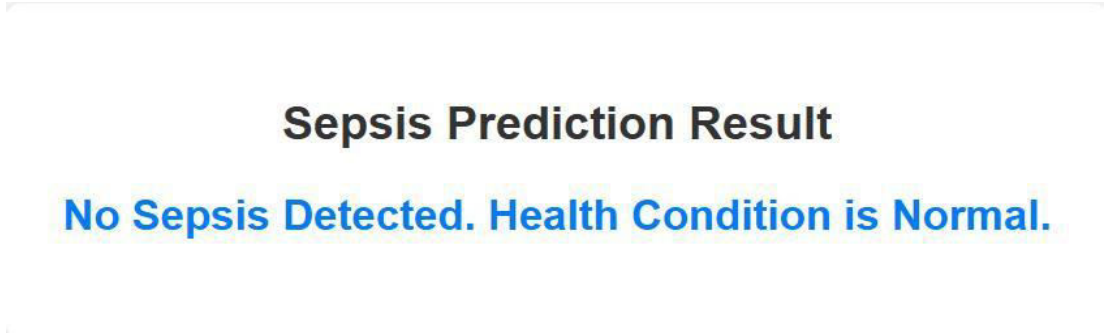
4.3 Implementation

For model training, Google Colab was used to train the machine learning model, while vs code was utilized for testing and deployment. The trained Random Forest model was saved and integrated into a Flask-based web application.

For front-end development, Sublime Text Editor was used, and the user interface was designed using HTML, CSS, and JavaScript to ensure an interactive and user-friendly experience. The Flask framework was used to connect the front-end with the trained machine learning model, enabling real-time sepsis predictions. The system allows users to input patient data, and the model processes it to predict the likelihood of sepsis, assisting healthcare professionals in early detection.

4.4 EXPERIMENTAL RESULT:



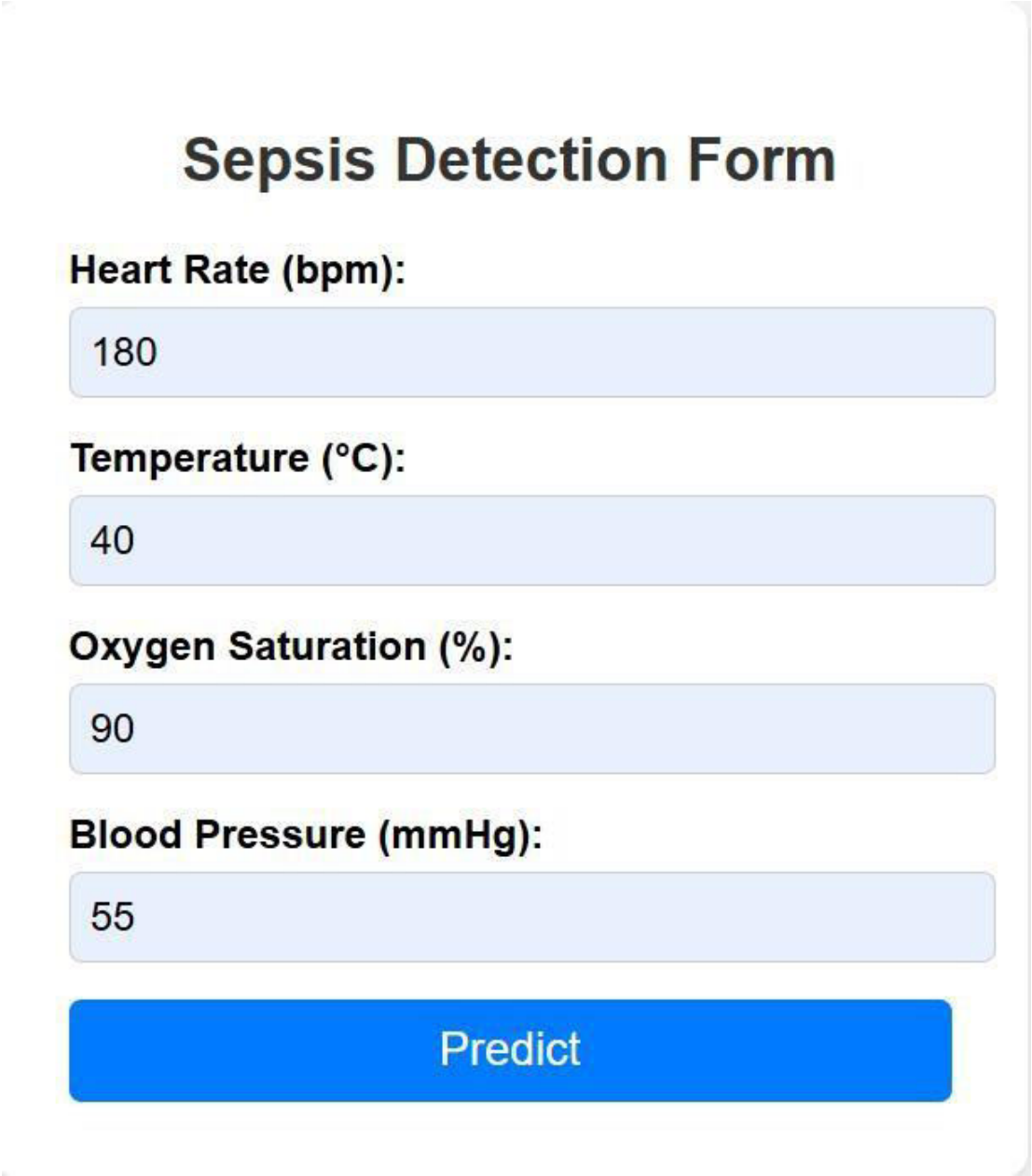


© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303116

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



V. CONCLUSION

This study developed a Random Forest-based machine learning model for early Sepsis Detection, improving diagnosis accuracy. The model was trained on patient data, deployed using Flask, and integrated into a web application for real-time predictions. This system can assist healthcare professionals in early detection, potentially reducing sepsis-related fatalities. Future enhancements may include deep learning integration and real-time patient

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

monitoring to further improve accuracy and usability.

REFERENCES

[1] Singer, M., Deutschman, C. S., et al. (2016). "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)." *JAMA, 315(8), 801-810.*
https://jamanetwork.com/journals/jama/fullarticle/2492881

[2] Vincent, J. L., et al. (2013). "The SOFA (Sepsis-Related Organ Failure Assessment) Score to Describe Organ Dysfunction/Failure." *Intensive Care Medicine, 22(7), 707-710.*
https://link.springer.com/article/10.1007/BF01709751

[3] Johnson, A. E. W., et al. (2016). "MIMIC-III, a Freely Accessible Critical Care Database." *Scientific Data, 3, 160035.*
https://www.nature.com/articles/sdata201635

[4] Mohanty, S., et al. (2021). "Early Detection of Sepsis Using Machine Learning: A Review." *Biomedical Signal Processing and Control, 68, 102684.*
https://www.sciencedirect.com/science/article/abs/pii/S174680942100517X

[5] Abdeltawab, H., et al. (2019). "A Random Forest Model for Sepsis Prediction in ICU Patients." *IEEE Access, 7, 112483-112495.*
https://ieeexplore.ieee.org/document/8787497

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details