UTILIZING MACHINE LEARNING PREDICTIVE ANALYTICS TO ENHANCE EARLY SEPSIS DIAGNOSIS IN CRITICAL CARE SETTING

PARON DAOTHONG

Graduate School Srinakharinwirot University

2024

การใช้การวิเคราะห์เชิงคาดการณ์ของการเรียนรู้ของเครื่องเพื่อปรับปรุงการวินิจฉัยภาวะการติดเชื้อในกระแสเลือดระยะแรกเริ่มในผู้ป่วยวิกฤต

ปารณ ดาวทอง

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมชีวการแพทย์
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2567

UTILIZING MACHINE LEARNING PREDICTIVE ANALYTICS TO ENHANCE EARLY SEPSIS DIAGNOSIS IN CRITICAL CARE SETTING

PARON DAOTHONG

An Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of MASTER OF ENGINEERING

(M.Eng. (Biomedical Engineering))

Faculty of Engineering, Srinakharinwirot University

2024

THE THESIS TITLED

UTILIZING MACHINE LEARNING PREDICTIVE ANALYTICS TO ENHANCE EARLY SEPSIS DIAGNOSIS IN CRITICAL CARE SETTING

BY

PARON DAOTHONG

HAS BEEN APPROVED BY THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE MASTER OF ENGINEERING
IN M.ENG. (BIOMEDICAL ENGINEERING) AT SRINAKHARINWIROT UNIVERSITY

(Assoc. Prof. Dr. Chatchai Ekpanyaskul, MD.)
Dean of Graduate School

ORAL DEFENSE COMMITTEE

..... Major-advisor
(Dr.Wongwit Senavongse, Ph.D.)

..... Chair
(Asst. Prof.Watcharin Tungsuksan)

..... Committee
(Dr.Theerasak Chanwimalueng, Ph.D.)

Title	UTILIZING MACHINE LEARNING PREDICTIVE ANALYTICS TO ENHANCE EARLY SEPSIS DIAGNOSIS IN CRITICAL CARE SETTING
Author	PARON DAOTHONG
Degree	MASTER OF ENGINEERING
Academic Year	2024
Thesis Advisor	Dr. Wongwit Senavongse , Ph.D.

This study develops machine learning and deep learning models to enhance early sepsis diagnosis in critical care using the eICU Collaborative Research Database, which includes over 200,000 ICU admissions from hospitals across the United States. Sepsis remains a global health challenge, responsible for an estimated 49 million cases and 11 million deaths annually. Early detection is crucial but difficult due to the condition's rapid progression and variable presentation.The research implements and compares several algorithms—Support Vector Machine, Logistic Regression, Random Forest, XGBoost, and Deep Neural Networks—using clinical features such as vital signs, bedside scores, and hemodynamic indicators. Both a core and a comprehensive feature set were used to assess the effect of data richness on performance. Models were evaluated on accuracy, AUROC, F1-score, recall, and specificity, with an emphasis on minimizing false negatives.XGBoost trained on the comprehensive dataset achieved the highest overall performance (AUROC: 0.88, F1-score: 0.74), offering strong sensitivity and specificity. Meanwhile, a dual-input deep learning model achieved the highest recall (0.70), highlighting its suitability for early-warning systems where identifying all potential sepsis cases is critical.This research confirms the value of machine learning in leveraging electronic health records for predictive diagnostics. Practical considerations for clinical integration are also discussed, including model interpretability, deployment within ICU workflows, and risk-based alerting strategies.

Keyword : Sepsis prediction, machine learning, deep learning, critical care, ICU, eICU database, XGBoost, clinical decision support, early diagnosis

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Wongwit Senavongse, my thesis advisor, for his invaluable guidance, thoughtful advice, and unwavering support throughout the course of this study. His expertise and encouragement have been instrumental in shaping both the direction and the quality of this work.

I am also sincerely grateful to the team behind the eICU Collaborative Research Database for providing access to a high-quality, large-scale dataset. Their contribution was essential to the development and evaluation of the predictive models used in this research.

PARON DAOTHONG

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1
INTRODUCTION

Background

Sepsis emerges as a severe medical condition when the immunity's reaction to inflection leads to harm to ordinary tissues and organs. It is considered one of the most serious and life-threatening medical emergencies.[1] Recent epidemiological studies estimate that sepsis affects approximately 49 million individuals annually and is responsible for around 11 million deaths worldwide. A substantial number of these cases—nearly half—occur in children, particularly in the neonatal period, with an estimated 2.8 million deaths in the first month of life. Many of these deaths could potentially be avoided with timely diagnosis and appropriate clinical intervention. In many low-resource settings, common causes of sepsis-related mortality include diarrheal and respiratory infections.[2,3]

According to the definition from The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)[4], Sepsis is defined as a life-threatening condition arising from a dysregulated immune response to infection, which may result in organ dysfunction. Therefore, prompt recognition and management are essential. Current clinical guidelines recommend using at least two screening criteria—such as qSOFA, SIRS, NEWS, or MEWS—in combination to enhance early detection of sepsis and septic shock.[5] In order to confirm sepsis, Despite the availability of laboratory diagnostics, delays in test results and interpretation can impact clinical decision-making. Even though laboratory tests, along with analyzed data from the hospital staff, can provide insightful information, most of them take time to diagnose affect the physician's correctness in perceiving and prescribing.

The purpose of this study is to develop and evaluate predictive models using machine learning techniques to support early sepsis detection in intensive care units by leveraging electronic medical record (EMR) data. The predictive machine learning model utilized in this research could assist physicians in diagnosing sepsis before its onset in

critical care patients, improve survival outcomes, and reduce both the incidence of septic shock and overall treatment costs.

Objectives of the Study

1. Develop the predictive machine learning model to early diagnosis sepsis in critical care patients.
2. Analyze electronic health records in the dataset to discover variables potentially associated with sepsis occurrence.

Significance of the Study

Presently, the concept of using prediction model on sepsis is widespread among the researchers in medical data science field as it is a life-threatening disease which may be avoided through timely recognition and effective medical intervention. There are some studies the use of sepsis clinical prediction model as it could reduce patient mortality, the result show that the model performs well in some hospitals but not performs well in some hospitals, even when hospitals are using the same underlying system.[6,7] In 2019, PhysioNet, organized the annual George B. Moody PhysioNet Challenge in collaboration with Computing in Cardiology, create a challenge for participant to build early sepsis prediction based on patient clinical information. The result of the challenge shows that various computational methods forecast sepsis onset hours before clinical symptoms become evident, but transferring these models across various hospital infrastructures remains a significant hurdle as most of the model performs well in hospitals system from training set but not in the different hospital system in hidden test set.[8,9]

In pursuit of advancing healthcare outcomes, this research embarked on a groundbreaking endeavor: utilizing electronic medical records (EMR) sourced from diverse hospital systems and countries. The objective was to harness this extensive data pool to construct an advanced machine learning model focused on identifying sepsis at an early stage in ICU settings. By aggregating information from various healthcare settings and geographical locations, this study sought to improve model consistency and predictive precision. The significance of this research lies in its potential to revolutionize

sepsis diagnosis, a condition often characterized by rapid deterioration and requiring prompt intervention. By leveraging the power of predictive analytics and machine learning, this innovative approach holds the promise of early detection, enabling timely and targeted medical interventions, ultimately saving lives, and improving the quality of care for critically ill patients across the globe.

Research Scope

The eICU-CRD (eICU Collaborative Research Database)[10,11] serves as a large-scale ICU dataset containing anonymized health information covering more than 200,000 intensive care admissions throughout U.S. hospitals during 2014–2015. It offers a wide array of patient information, such as patient demographics, physiological measurements, lab values, and therapeutic interventions. an overview of the primary variables available in the dataset is presented below:

Table 1 Variable contain on the dataset

No.	Name	Description
1.	Arterial Line MAP (mmHg)	Mean arterial pressure via arterial line
2.	Bedside Glucose	Glucose level measured at the bedside
3.	Best Eye Response	Eye response score from the Glasgow Coma Scale (GCS)
4.	Best Motor Response	Motor response score from the GCS
5.	Best Verbal Response	Verbal response score from the GCS
6.	CI	Cardiac index (cardiac output normalized by body surface area)
7.	CO	Cardiac output (volume of blood pumped by the heart per minute)
8.	CPP	Cerebral perfusion pressure
9.	CV/ PV Assessment	Cardiovascular and peripheral vascular assessment
10.	CVP	Central venous pressure
11.	CVP (mmHg)	Central venous pressure value in mmHg
12.	Delirium Scale/Score	Clinical scoring of delirium severity
13.	ECG (secs)	Electrocardiogram duration or intervals

Table 1 (Continued)

No.	Name	Description
14.	ECMO	Extracorporeal membrane oxygenation in use
15.	Electrolyte Replacement	Administration of electrolytes (e.g., potassium, magnesium)
16.	End Tidal CO2	CO_2 level measured at the end of exhalation
17.	Eye Opening	Eye opening component of GCS
18.	Eye, Ear, Nose, Throat Assessment	EENT physical assessment
19.	Fall Risk	Assessment of patient risk for falling
20.	Gastrointestinal Assessment	Assessment of GI system
21.	Genitourinary Assessment	Assessment of GU system
22.	Glasgow coma score	Assessment of consciousness using GCS scale (3-15)
23.	Heart Rate	Heartbeats per minute
24.	IAP	Intra-abdominal pressure
25.	ICP	Intracranial pressure
26.	Impella	Impella heart pump support device present
27.	Integumentary Assessment	Skin and tissue integrity assessment
28.	Invasive BP	Blood pressure measured directly via arterial line
29.	LVAD	Left ventricular assist device present
30..	Level of Assistance	Required assistance level for mobility or care
31.	MAP (mmHg)	Mean arterial pressure in mmHg
32.	Mental Status Assessment	Assessment of alertness and cognition
33.	Motor Response	Observed motor response (part of GCS)
34.	Musculoskeletal Assessment	Assessment of muscles and skeletal function
35.	Neurological Assessment	Overall neurologic system assessment
36.	Non-Invasive BP	Blood pressure measured using a cuff (non-invasive method)
37.	O2 Admin Device	Device used for oxygen administration (e.g., nasal cannula, mask)

Table 1 (Continued)

No.	Name	Description
38.	O2 L/%	Oxygen flow rate or percentage concentration
39.	O2 Saturation	Oxygen saturation level in peripheral blood
40.	P.O.	Ingestion of food or medication by mouth
41.	PA	Pressure in the pulmonary artery
42.	PAOP	Pressure measured during pulmonary artery occlusion (wedge pressure)
43.	PVR	Resistance in the lung vasculature
44	PVRI	Indexed measure of pulmonary vascular resistance
45	Pain Assessment	Nursing assessment of pain level
46	Pain Present	Boolean indicator of pain presence
47	Pain Score/Goal	Patient pain score and/or targeted pain goal
48	Patient s Comfort/Function (Pain) GOAL At Rest	Comfort and functional pain goals at rest
49	Pulse	Heart rate measured manually or by pulse oximeter
50	Pulse Ox Mode	Mode of pulse oximetry monitoring
51	RASS	Richmond Agitation-Sedation Scale
52	Respiratory Assessment	Clinical assessment of respiratory system
53	Respiratory Rate	Number of breaths per minute
54	SEDATION SCORE	Quantitative sedation level
55	SV	Stroke volume (amount of blood pumped per heartbeat)
56	SVO2	Oxygen saturation in mixed venous blood
57	SVR	Resistance in systemic circulation
58	SVRI	Systemic vascular resistance index
59	Score (Glasgow Coma Scale)	Total Glasgow Coma Scale score
60	Sedation Scale/Score/Goal	Sedation level or goal as assessed by clinical staff
61.	SpO2	Peripheral oxygen saturation

Table 1 (Continued)

No.	Name	Description
62.	Symptoms of Delirium Present	Observation of delirium-related behaviors
63.	Temperature	Patient's core or peripheral body temperature
64.	Verbal Response	Observed verbal response (part of GCS)

Project plan

The research is planned to start the process on 1st August 2023 and finish on 15 January 2024 at the Biomedical Engineering Department at Srinakharinwirot University.

Table 2 Grant Graph

No.	Task	Month							
		AUG-DEC	JAN-APR	MAY-JUL	AUG-DEC	JAN	FEB	MAR	APR
1.	Literature review related work								
2.	Plan and edit the research topic								
3.	Collect the dataset								
4.	Preprocessing dataset								
5.	Working on thesis chapter 1-3 for 1st Presentation								
6.	Develop the machine learning model								

Table 2 (Continued)

No.	Task	Month							
		AUG-DEC	JAN-APR	MAY-JUL	AUG-DEC	JAN	FEB	MAR	APR
7.	Analyze and conclude the research result								
8.	Prepare the paper for the international conference								
9.	Working on thesis chapter 4-5 for 2nd Presentation								
10.	Present the research Project in the conference								
11.	Edit and Summary the final thesis submission								

Chapter 2
LITERATURE REVIEW

Immune System

Definition of Immunity

Immunity refers to the body's intricate defense system that protects against harmful pathogens, infections, and diseases. It serves as a fundamental biological function, enabling organisms to recognize and neutralize foreign invaders while preserving the body's ability to tolerate its own cells and molecules.[12] The immune system is broadly divided into innate immunity, offering rapid general defense, and adaptive immunity, which evolves over time to target specific pathogens for lasting immunity.[13]

Mechanisms of Immunity

The immune system functions as a sophisticated and highly coordinated network of cells and signaling molecules that detect, neutralize, and eliminate pathogenic threats.[14] It is broadly classified into innate and adaptive immunity, both of which are critical for sustaining health and preventing infections.

Innate Immunity

- Innate immunity serves as the body's immediate barrier against infection, acting quickly through generalized defense processes. This built-in defense system exists from birth and functions without needing earlier contact with pathogens.[15] It primarily relies on physical barriers, chemical secretions, and immune cells to identify and neutralize pathogenic invaders.

- Physical and Anatomical Barriers: The skin and mucous membranes serve as effective barriers between the internal and external environments, preventing pathogen entry. The epidermis's outermost layer is primarily composed of keratinocytes, connected by desmosomes and surrounded by extracellular matrix components.[16]
- Effector Cells and Soluble Mediators: Macrophages, dendritic cells, neutrophils, and natural killer cells are key immune cells that help identify and neutralize threats. They recognize pathogens using PRRs that bind to molecular structures like PAMPs, triggering an instant immune response.[17]

Adaptive Immunity

- Adaptive immunity is a specialized and highly efficient defense mechanism that provides long-term protection through immunological memory. Unlike innate immunity, which responds immediately in a broad manner, adaptive immunity targets specific antigens and depends on immune cell activation and proliferation for a robust response during repeated exposures.[18]
- Lymphocytes such as B cells and T cells mediate adaptive immunity. B cells generate antibodies to neutralize pathogens, while T cells help destroy infected cells and coordinate immune defense mechanisms.[19]

- Immunological Memory: One of adaptive immunity's defining traits is its capacity to recall previous pathogens. When re-encountered, it reacts more rapidly and strongly. This principle underpins the effectiveness of vaccines in providing long-term immunity.[20]

Enhancement of Immunity

Immune activity is influenced by several elements such as diet, vaccines, personal habits, and the gut microbiome's makeup. These factors significantly influence both innate and adaptive immune responses, thereby contributing to overall health and disease resistance.[21]

Nutrition and Dietary Interventions

- Optimal nutrition is essential for maintaining a robust immune system. A well-balanced diet provides vital micronutrients and bioactive compounds that support immune cell function and mitigate susceptibility to infections.[22]
- Gut Microbiota: The gut's microbial community significantly influences immune regulation. Early exposure to diverse microbes helps condition the immune system and may prevent chronic autoimmune conditions and allergies.[23]

Vaccination and Immune Priming

- Vaccines continue to be a highly effective method for strengthening immunity and lowering infection risks. By introducing harmless components of pathogens, vaccines stimulate the development of memory cells, Preparing the immune system to respond quickly and powerfully when exposed again.[19]

- Mechanisms of Vaccines help train the immune system to detect and respond more effectively to harmful microbes, lowering the risk of serious illness.[24]

Microbiome and Immune Regulation

- The gut microbiota is essential for immune balance, helping to regulate inflammation and communicate with immune cells. A well-balanced gut microbiome supports immune stability and helps protect the body from harmful microbes. [16]
- Immune Modulation: Beneficial gut microbiota produces metabolites that regulate immune responses and mitigate excessive inflammation.[23]

Sepsis and Septic shock

Definition and Overview

Sepsis is a severe condition that can lead to death from infection when it is not identified and treated in a timely manner. Its identification demands urgent attention. The progression of sepsis is influenced by various host and microbial elements, including age, genetics, health status, and environmental conditions. Sepsis is identified by an abnormal immune response to infection that leads to organ failure, setting it apart from ordinary infections. It's crucial to note that sepsis-induced organ dysfunction might be hidden, making its consideration imperative in any patient with an infection. Conversely, a new-onset organ dysfunction could be the result of an unrecognized infection. Therefore, organ failure without a clear cause should prompt consideration of a possible hidden infection. Furthermore, Sepsis symptoms may vary depending on underlying conditions, recent treatments, or coexisting diseases. It's essential to recognize that specific infections can lead to localized organ dysfunction without triggering a dysregulated systemic host response.[4]

Sepsis is formally defined as a critical condition marked by a severe condition caused by abnormal immune responses leading to organ failure. The identification of organ dysfunction involves an acute change in the total Sequential Organ Failure Assessment (SOFA) score by at least 2 points attributable to the infection. In patients without known pre-existing organ dysfunction, the baseline SOFA score is assumed to be zero. Notably, a SOFA score of 2 or more reflects an overall mortality risk of approximately 10% in a general hospital population with suspected infection. It is crucial to recognize that even patients with modest dysfunction can experience further deterioration, underscoring how serious this condition is and the urgency of immediate treatment when required.[25]

In simpler terms, sepsis occurs when the body's immune reaction to infection harms its own tissues and organs, putting life at risk. Swift identification of patients who may require extended ICU care or are at higher risk of death can be achieved at the bedside using the quick SOFA (qSOFA) criteria, which include alterations in mental status, systolic blood pressure less than 100 mm Hg, or a respiratory rate exceeding 22/min. Additionally, septic shock, a subset of sepsis, is characterized by profound circulatory and problems in how cells and metabolism function. Identification of patients with septic shock involves a clinical construct of sepsis with persistent hypotension necessitating vasopressors to maintain a mean arterial pressure of at least 65 mm Hg and a serum lactate level surpassing 2 mmol/L (18 mg/dL) despite adequate volume resuscitation. It is crucial to note that meeting these criteria results in a hospital mortality rate exceeding 40%, underscoring the severity of septic shock and emphasizing the critical need for timely intervention.[26]

Epidemiology and Global Impact

Sepsis and septic shock are significant global health concerns, impacting nearly 49 million individuals and causing about 11 million deaths globally in 2017, representing close to one-fifth of all deaths worldwide.[27] In 2017, the World Health Organization (WHO) identified sepsis as a major global health issue and called for improved strategies in prevention, early diagnosis, and effective management.

Sepsis and Septic Shock Mortality and Morbidity

- Sepsis-related deaths account for 20% of global mortality.
- Septic shock mortality rates range from 40% to 60%, making it the deadliest stage of sepsis.[4]
- Neonatal sepsis contributes to 2.9 million deaths globally per year, primarily in low-income settings.[27]

Regional and Demographic Disparities

- Sepsis and septic shock pose a greater challenge in low- and middle-income countries (LMICs) where healthcare infrastructure is often limited, high rates of infectious diseases, and inadequate sanitation.[28] Neonatal and maternal sepsis remain major contributors to mortality, particularly in regions with limited access to antibiotics and proper medical care.[29]
- High-Income Countries (HICs): Older adults, immunocompromised individuals, and people with chronic diseases are at greater risk of developing sepsis. Hospital-acquired infections are a leading cause.[30]
- Low- and Middle-Income Countries (LMICs): Higher incidence of neonatal sepsis, malaria-related sepsis, and infections due to poor sanitation.[27]
- Neonates and Infants: Approximately 3 million neonatal sepsis cases occur annually, with mortality rates exceeding 30% in resource-limited settings.[29]

Pathophysiology and Disease Progression

Sepsis and septic shock arise from an abnormal immune response to infection, which triggers widespread inflammation, endothelial injury, blood clotting irregularities, and metabolic disturbances.[31] Sepsis pathophysiology reflects an unstable interplay between inflammatory and anti-inflammatory processes, where excessive immune activation can lead to multi-organ dysfunction syndrome (MODS).

Mechanisms of Sepsis Progression

1. **Immune System Activation and Cytokine Storm**
   - Pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) trigger Toll-like receptors (TLRs), activating the innate immune system.[32]
   - Overproduction of inflammatory mediators like TNF-α, IL-1β, and IL-6 can trigger a cytokine storm, which contributes to extensive tissue injury and organ dysfunction.[33]
2. **Endothelial Dysfunction and Vascular Leakage**
   - Sepsis induces endothelial barrier disruption, causing increased vascular permeability, capillary leak syndrome, and edema.[34]
   - Abnormal blood clotting mechanisms can cause disseminated intravascular coagulation (DIC), worsening small vessel blockages and limiting oxygen delivery to tissues.[35]
3. **Mitochondrial Dysfunction and Metabolic Failure**
   - Impaired mitochondrial respiration results in decreased ATP production, shifting metabolism towards anaerobic glycolysis and lactic acidosis.[4]
   - Elevated levels of reactive oxygen species (ROS) contribute to cellular damage through oxidative stress mechanisms.[36]

4. **Multi-Organ Dysfunction Syndrome (MODS) and Septic Shock**
    - Persistent hypoxia, tissue damage, and inflammatory dysregulation lead to MODS, involving failure of the lungs, kidneys, heart, and brain, increasing mortality.[34]
    - In septic shock, excessive nitric oxide (NO) production causes profound vasodilation and refractory hypotension.[37]

Risk Factors and Causes

Sepsis and septic shock arise from diverse infections and patient-specific risk factors. Recognizing these contributors enables early intervention and targeted therapies.

Common Causes of Sepsis and Septic Shock

- Bacterial Infections: The leading cause, particularly involving gram-negative bacteria such as Escherichia coli and Pseudomonas aeruginosa, and gram-positive strains like Staphylococcus aureus and Streptococcus pneumoniae.[38]
- Fungal Infections: Higher risk in immunocompromised individuals, particularly Candida species.[39]
- Viral Infections: Sepsis can result from influenza, COVID-19, and other viral infections[5]

Key Risk Factors for Sepsis and Septic Shock

1. **Advanced Age**
    - Individuals older than 65 are at significantly elevated risk due to age-related decline in immune function and existing health conditions.[40]
2. **Chronic Diseases**
    - Underlying conditions like diabetes, liver cirrhosis, renal insufficiency, and heart disease compromise immune defenses and heighten vulnerability to infections.[41]

3. **Immunosuppression**
    - Cancer, HIV/AIDS, prolonged corticosteroid use, and post-transplant immunosuppression significantly elevate sepsis risk.[33]
4. **Nosocomial Infections and Invasive Procedures**
    - Infections acquired during hospitalization, including those linked to ventilators and central lines, are major contributors to sepsis incidence.[30]
    - Surgical procedures, central venous catheters, and mechanical ventilation introduce pathogens into the bloodstream, raising sepsis risk.[42]

Signs and Symptoms of Sepsis and Septic Shock

sepsis presents with a broad spectrum of symptoms due to its systemic nature, making early identification crucial for improving patient outcomes.[4] Symptoms range from mild to severe and progress rapidly, leading to multi-organ dysfunction syndrome (MODS) if left untreated.[27]

Early Signs of Sepsis

1. Fever or Hypothermia: An elevated body temperature (>38.3°C) or hypothermia (<36°C) is common due to immune system dysregulation.[4]
2. Tachycardia, characterized by a heart rate over 90 beats per minute, often signals systemic inflammatory response.[43]
3. Tachypnea, or a respiratory rate above 22 breaths per minute, suggests compensatory effort in response to metabolic imbalance.[5]
4. Altered Mental Status: Confusion, disorientation, or lethargy suggest cerebral hypoperfusion and inflammation.[32]

Severe Symptoms Indicative of Septic Shock

1. Persistent low blood pressure, defined as MAP under 65 mmHg even after fluid therapy, points to circulatory system collapse.[44]
2. Lactic Acidosis: Elevated lactate levels above 2 mmol/L indicate inefficient oxygen use and tissue hypoperfusion.[4]
3. Oliguria or Anuria: Decreased urine output (<0.5 mL/kg/hour) is indicative of acute kidney injury (AKI) due to hypoperfusion.[45]
4. Cold, Clammy Skin: Poor peripheral perfusion results in cyanosis and mottling, particularly in the extremities.[41]

Multi-Organ Dysfunction and Late-Stage Symptoms

- Respiratory Failure: Progression to ARDS may necessitate mechanical ventilatory support due to severe lung impairment.[46]
- Liver Dysfunction: Elevated bilirubin and transaminases, indicating hepatocellular injury.[35]
- Coagulopathy and Disseminated Intravascular Coagulation (DIC): Elevated prothrombin time (PT), activated partial thromboplastin time (aPTT), and d-dimer indicate sepsis-induced clotting dysfunction.[47]

Sepsis Criteria

According to the recommendation from sepsis and septic shock 2021 guideline[5] ,the screening method on sepsis should be use more than one criterion to confirm the sepsis on patient. The guideline recommends by meeting a minimum of two indicators from the criteria below:

qSOFA

- qSOFA serves as a primary screening approach for sepsis, often applied in conjunction with other criteria. To suspect sepsis with qSOFA is using the following list, a score above 2 indicates a strong suspicion of sepsis in the patient.

Table 3 qSOFA (Quick SOFA) Criteria

No.	Criteria
1.	Respiratory rate >= 22/min
2.	Altered mentation
3.	Systolic blood pressure <=100 mm Hg

SIRS (Systemic Inflammatory Response Syndrome)

- SIRS provides an alternative method for identifying sepsis through the checklist below, sepsis is likely if the total score exceeds 2.

Table 4 SIRS (Systemic Inflammatory Response Syndrome) Criteria

No.	Criteria
1.	Temperature >38°C or <36°C
2.	Heart rate >90/min
3.	Respiratory rate >20/min or Paco2 <32 mm Hg (4.3 kPa)
4.	White blood cell count >12 000/mm3 or <4000/mm3 or >10% immature bands

National Early Warning Score (NEWS)

- NEWS is another criterion that is applied to help detect potential sepsis cases in clinical settings. By using NEWS check list as following, a NEWS score of 5 or higher indicates a possible sepsis-related infection in the patient.

Chart 1: The NEWS scoring system

Physiological parameter	Score 3	2	1	0	1	2	3
Respiration rate (per minute)	≤8		9–11	12–20		21–24	≥25
SpO_2 Scale 1 (%)	≤91	92–93	94–95	≥96			
SpO_2 Scale 2 (%)	≤83	84–85	86–87	88–92 ≥93 on air	93–94 on oxygen	95–96 on oxygen	≥97 on oxygen
Air or oxygen?		Oxygen		Air			
Systolic blood pressure (mmHg)	≤90	91–100	101–110	111–219			≥220
Pulse (per minute)	≤40		41–50	51–90	91–110	111–130	≥131
Consciousness				Alert			CVPU
Temperature (°C)	≤35.0		35.1–36.0	36.1–38.0	38.1–39.0	≥39.1	

Figure 1 NEWS score

Modified Early Warning Score (MEWS)

- MEWS is one of the clinical scoring tools used to identify patients potentially at risk for sepsis, with MEWS scores of 5 or more may indicate a suspected sepsis infection.

Score	3	2	1	0	1	2	3
Respiratory rate		<9		9-14	15-20	21-29	>30
Heart rate		<40	41-50	51-100	101-110	111-129	>130
Systolic blood pressure	<70	71-80	81-100	101-199		>200	
Temperature		<35		35-38,4		>38,5	
Level of consciousness				Alert	Voice	Pain	Unresponsive
Contact physician when MEWS score > 4, if oxygen saturation drops to <90 % with oxygen treatment, or if you are concerned about the patient's condition.							

Color-code	MEWS score	Follow up/new measurements
Blue	0	24 hours
Yellow	1	8-12 hours
Orange	2	4-8 hours
Red	3-4	1-4 hours
	>4	Contact physician

Figure 2 MEWS score

Machine learning

Machine learning in healthcare has gained increasing attention due to its transformative potential. Researchers explore whether machine learning enhances patient care, with factors like hospital settings and data quality influencing model performance.[48]

A primary application of machine learning in healthcare is predictive modeling, used to assess patient health conditions and improve treatment plans. Studies evaluating machine learning's real-world performance in critical care settings in the United States indicate that these models can effectively handle clinical situations. However, performance varies based on hospital size and resource availability. Some models perform well in larger hospitals with more structured data, while in smaller hospitals, performance may degrade due to factors such as inconsistent data collection and resource limitations.[49]

Machine learning is widely applied in healthcare for image processing, disease diagnosis, and patient monitoring. Algorithms like deep convolutional networks, support vector machines (SVMs) and random forest algorithms are commonly employed in medical diagnosis and clinical outcome prediction.[50–52]

However, implementation challenges persist across diverse healthcare systems. Hospital-related factors, such as the availability of structured data, clinician acceptance of AI-based recommendations, and regulatory constraints, contribute to variations in model performance.[53]

In this study, I propose using well-established machine learning algorithms that are frequently applied in healthcare. These algorithms have demonstrated strong performance in areas such as disease prediction, patient monitoring, and medical imaging. The table below summarizes their definitions and common applications in medical research.

Table 5 Definition of algorithms and models in this study

No.	Name of algorithm and model	Definition
1	Deep Learning	Deep learning, a branch of machine learning, utilizes layered neural networks to model intricate patterns and relationships in data. n healthcare, deep learning is extensively applied in medical imaging, diagnostic support, and the development of individualized treatment strategies. It excels in tasks like tumor detection, medical image segmentation, and electronic health record analysis.[54]

Table 5 (Continued)

No.	Name of algorithm and model	Definition
2	Linear Support Vector Machine (LinearSVM)	Linear Support Vector Machine (SVM) is designed for datasets where classes can be separated by a straight hyperplane. It is particularly effective for binary classification tasks with clear class boundaries.[55]).
3	Logistic Regression (LR)	Logistic regression is a classification model that uses several independent parameters to predict a binary-dependent outcome. It is a highly effective technique for identifying the relationship between data or cues or a particular occurrence.[56]
4	Random Forest (RF)	Random Forest is an ensemble learning method that aggregates predictions from multiple decision trees to enhance model accuracy and reduce overfitting. It is widely used for both classification and regression tasks.[57]
5	eXtreme Gradient Boost (XGBoost)	XGBoost is an optimized gradient-boosting framework that uses decision trees to improve predictive accuracy. Known for its speed and performance, it supports regression, classification, and ranking tasks, and includes regularization to reduce overfitting.[58]

Related work

Several previous studies have pursued similar objectives to this research. A summary of related literature, including datasets and machine learning models used, is presented in the following table.

Table 6 Related work

Study	Dataset	Algorithms and Model
Sudarsan Sadasivuni et al.	Research dataset	ANN, LinearSVM, LR, RF,
Sadik Aref et al.	The Physionet challenge 2019	XGBoost
S. Babu et al	The Physionet challenge 2019 and MIMIC-III	XGBoost, MLP, GB, LDA
Xiao Lu et al.	MIMIC-III and MIMIC-IV	Gradient Boosting Decision Tree (GBDT), XGBoost, RF, LightGB, Support Vector Machine (SVM)
Divya Bhaskaracharya et al.	The Physionet Challenge 2019	LR, Naïve Bayes classifier, KNN, XGBoost, RF
Benjamin Roussel et al.	The Physionet Challenge 2019	RNN

CHAPTER 3
RESEARCH METHODOLOGY

Data Source

This study uses the eICU Collaborative Research Database (eICU-CRD) [10,11], a large-scale, deidentified critical care dataset comprising more than 200,000 ICU admissions from multiple centers across the United States between 2014 and 2015. It contains rich clinical data including vital signs, care plans, illness severity scores, diagnoses, and treatments. The data was collected through the Philips eICU telehealth program, which supports real-time remote monitoring of ICU patients.

Data Preparation

This section describes the preprocessing steps applied to the dataset to ensure consistency, manage missing values, and prepare the data for analysis. The key steps include patient identification, sepsis labeling, handling of missing data, feature aggregation, and dataset construction.

Patient Identification and Cohort Segmentation

Each patient in the dataset is uniquely identified using PatientUnitStayID. This identifier ensures that data remains patient-specific and prevents redundancy or duplication. By structuring the dataset in this manner, patient records are maintained as distinct units, thereby preserving data integrity and preventing data leakage during model development.

Sepsis labeling

The dataset includes predefined labels indicating whether a patient has been diagnosed with sepsis. Based on this labeling, patients are classified into two distinct groups:

- Sepsis group: Patients who meet the criteria for sepsis based on the dataset's predefined labels.

- Non-sepsis group: Patients who do not meet the criteria for sepsis.

This approach ensures consistency in patient classification and eliminates the need for additional manual labeling based on external criteria.

Handling missing data

Missing data is managed using a structured two-step approach to ensure the reliability of the dataset while retaining informative features:

1. Feature Elimination: Features exhibiting excessive missingness, specifically over 60–80%, were excluded from the dataset to maintain data quality. This threshold is chosen to balance data retention with feature completeness, ensuring that excessively sparse features do not negatively impact analysis.

2. Missing Value Imputation: For the remaining features, missing values are systematically replaced with zero (0). This approach standardizes the dataset while minimizing potential biases introduced by other imputation methods.

This strategy ensures that the dataset remains comprehensive while mitigating the influence of incomplete data.

Feature Engineering and Aggregation

To facilitate analysis and enhance model performance, key statistical metrics are computed for each feature at the patient level. These statistical summaries transform raw time-series data into structured representations, improving interpretability and reducing dimensionality. The computed metrics include:

- Mean: The average value over the recorded period.
- Mode: The value that appears most often in the patient's data.
- Median: A central value that better represents skewed data distributions.
- Maximum: The highest observed value for each patient.

By aggregating data in this manner, patient-level characteristics can be more effectively analyzed in subsequent modeling steps.

Dataset Construction

Following preprocessing and feature engineering, two datasets are created to support different aspects of the analysis:

1. Comprehensive Feature Dataset: This dataset includes all retained features following missing data handling and feature aggregation. It provides a complete representation of patient data.
2. Core Feature Dataset: This dataset consists of the six features with the highest data availability, ensuring a streamlined and reliable subset for focused predictive modeling.

The construction of these datasets allows for both broad exploratory analysis and targeted model development, ensuring flexibility in the research methodology.

Experiment

This section describes the experimental setup used to develop and evaluate predictive models for sepsis detection. The process consists of dataset partitioning, model training, hyperparameter tuning, validation, and final performance assessment. The workflow for model development is illustrated in Figure 3.

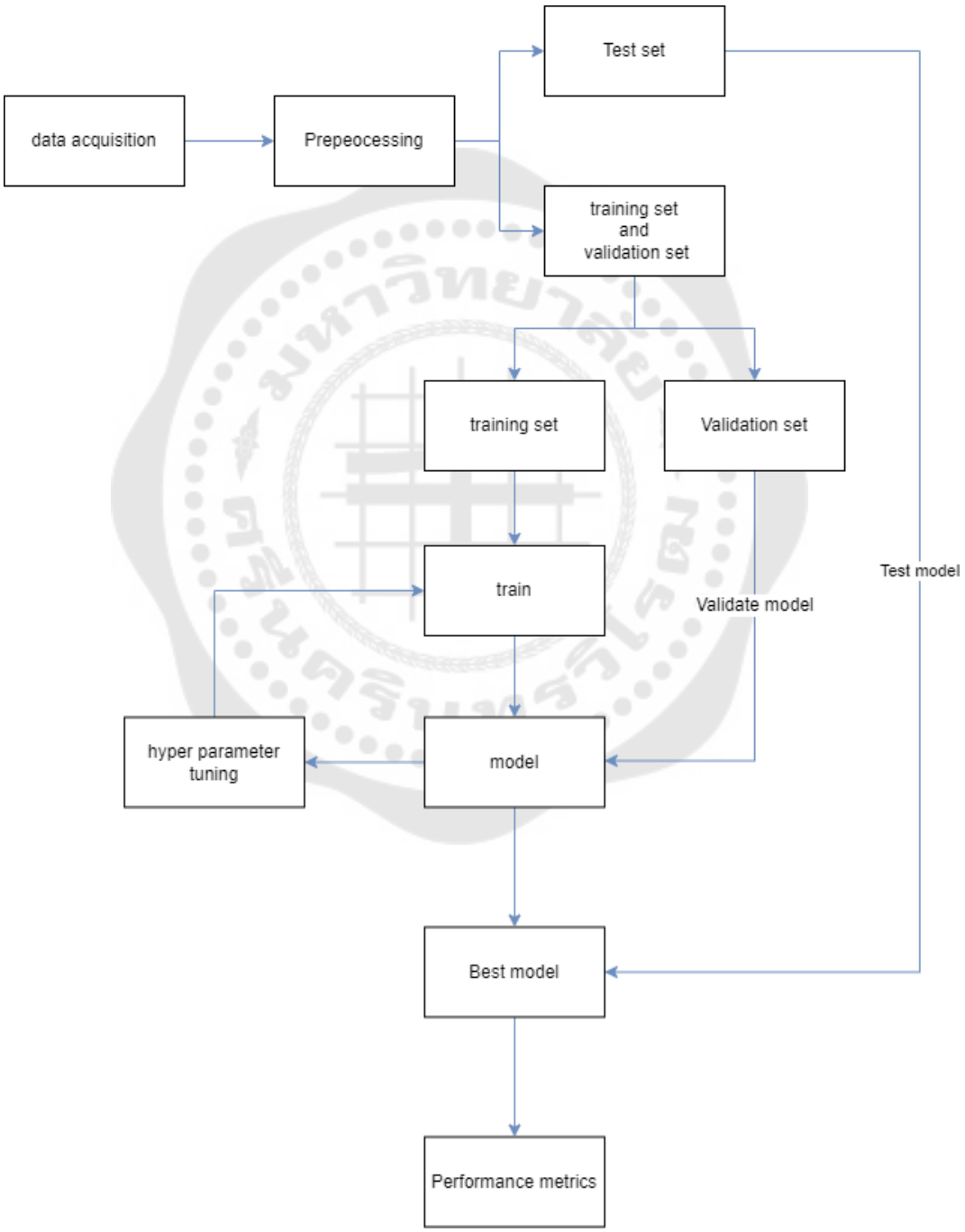


Figure 3 Model development

Dataset Splitting

To ensure a robust evaluation process, the dataset is randomly partitioned into three subsets:

- Test Set (20%) – Reserved for final model evaluation.
- Development Set (80%) – Used for training and validation.

The development set is further divided into:

- Training Set (80%) – Used for model training.
- Validation Set (20%) – Used for hyperparameter tuning and performance validation.

This splitting strategy ensures that the test data remains completely independent from the training and validation process, preventing data leakage and overfitting.

Machine Learning Model Training

For traditional machine learning models, a **10-fold cross-validation** approach is employed to optimize hyperparameters and improve model performance. The process is as follows:

1. Hyperparameter Tuning
   - Training data is split evenly into ten parts..
   - Each part takes a turn as validation data, with the others used for training.
   - The model is trained repeatedly with varying data splits, and tuning is guided by performance outcomes.
2. Final Model Training
   - The best-performing hyperparameters are selected.
   - The final model is retrained on the complete training data with the selected parameters.

The following machine learning algorithms are each trained separately using both datasets (Comprehensive and Core Feature datasets):

- Linear Support Vector Machine (Linear SVM) – Trained with both datasets
- Logistic Regression (LR) – Trained with both datasets
- Random Forest (RF) – Trained with both datasets
- eXtreme Gradient Boosting (XGBoost) – Trained with both datasets

Each model's performance is evaluated based on the validation set, and the best-performing model proceeds to final testing.

Deep Learning Model Training

Unlike traditional models, deep learning networks are trained directly without k-fold cross-validation. To assess how different different feature sets, three deep learning models are developed:

1. Model A – Trained using the Comprehensive Feature Dataset.
2. Model B – Trained using the Core Feature Dataset.
3. Model C – Trained using a combination of both datasets, incorporating a broader feature set.

Each deep learning model undergoes training and fine-tuning with optimization techniques such as learning rate adjustments and dropout regularization. The final model is selected based on validation performance.

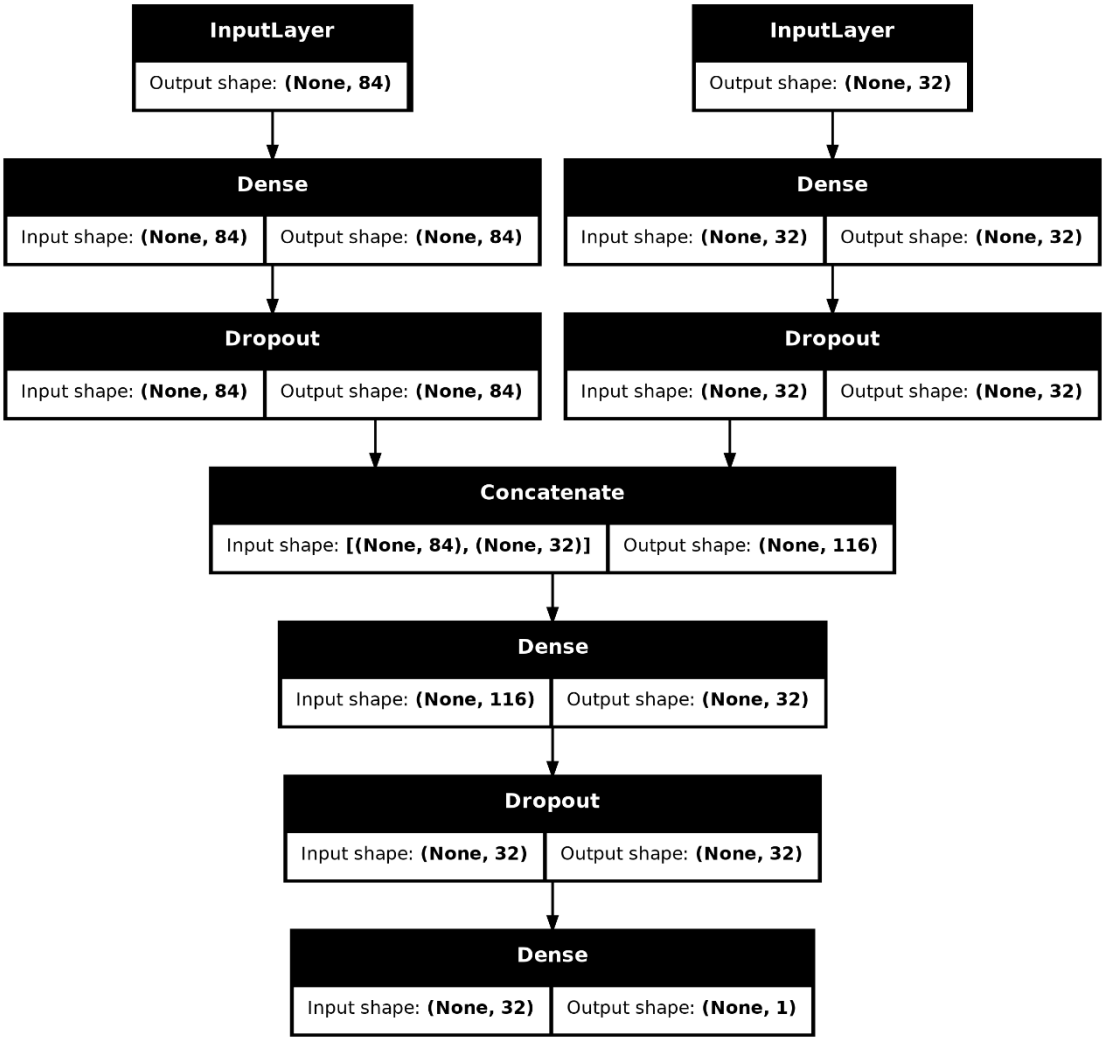


Figure 4 Deep Learning Model C

Model Development Workflow

Figure 3 illustrates the full model development pipeline, from preprocessing through training, validation, and testing.

Steps in Model Development: Acquiring and Preprocessing Data:

Data Acquisition and Preprocessing

- The data undergoes collection, cleaning, and preprocessing, as outlined earlier.

Dataset Splitting

- The dataset was split into three subsets: training, validation, and testing.

Model Training

- Machine learning models were tuned using 10-fold cross-validation to determine the best-performing hyperparameters.
- Each deep learning model was trained independently on distinct dataset configurations.

Hyperparameter Tuning

To optimize model performance, all algorithms underwent structured hyperparameter tuning using grid search and validation AUC as the primary evaluation metric. This approach ensured robust model selection, particularly in the context of class imbalance and clinical sensitivity.

The following hyperparameters and value ranges were explored:

- Support Vector Machine (SVM):
  - kernel: (linear, rbf)
  - C: (0.1, 1, 10)
  - gamma: (scale, auto)
  - class weight: (None, balanced)
- Random Forest (RF):
  - n_estimators: (100, 200)
  - max_depth: (10, 20, None)
  - min_samples_split: (2, 5)
- XGBoost:
  - max_depth: (3, 5, 7, 9)
  - learning_rate: (0.01, 0.05, 0.1, 0.2)
  - subsample: (0.6, 0.8, 1.0)

- colsample_bytree: (0.6, 0.8, 1.0)

- Logistic Regression (LR):
  - penalty: (l2, elasticnet)
  - C: (0.001, 0.01, 0.1, 1, 10, 100)
  - solver: (liblinear, saga)
  - l1_ratio: (0.1, 0.5, 0.7, 1.0)
- Deep Learning Models:
  - dropout: (0.3–0.5)
  - learning rate: (0.001, 0.01, 0.0001)
  - hidden layers: (2–4)
  - neurons per layer: (32–256)
  - Optimizer: Adam

Model Evaluation and Selection

- Validation results were used to select the best-performing model.
- Final evaluation was conducted using the holdout test set to ensure objective performance measurement.

Performance Assessment

- The final model is evaluated using key performance metrics, including accuracy, precision, recall, F1-score, and AUC-ROC.

Performance metrics

Model performance was evaluated using standard metrics such as accuracy, precision, recall (sensitivity), F1-score, specificity, and area under the ROC curve (AUROC).

Confusion matrix

- The confusion matrix summarizes classification results by showing correct and incorrect predictions per class, helping identify where the model confuses one class for another.[59]

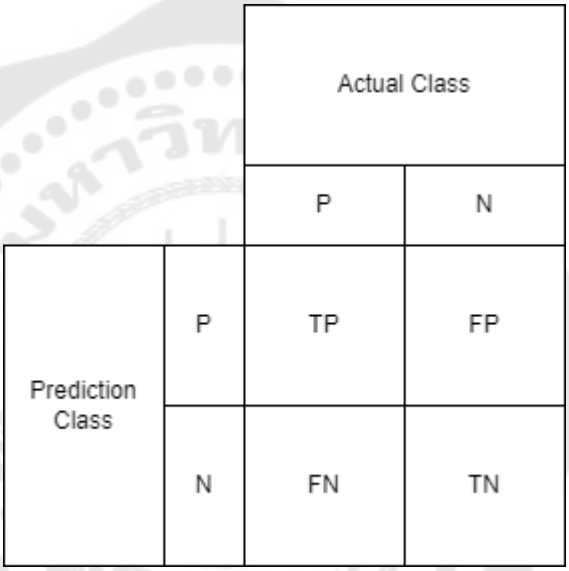


Figure 5 Confusion matrix

- P: Positive, N: Negative
- TP: True Positive, FP: False Positive
- FN: False Negative, TN: True Negative

Accuracy

- Accuracy measures the proportion of correct predictions out of all predictions made by the model. Commonly, Accuracy is used in balanced dataset. For the unbalanced dataset accuracy should not be use[59]. From Figure 2, accuracy can be calculated as:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)}$$

Precision

- Precision is suited for the cases where the study required confidence in positive outcomes. As Precision assesses the proportion of true positive predictions among all positive predictions made. From Figure 2, precision can calculate as:

$$Precision = \frac{TP}{TP + FP}$$

Sensitivity or Recall

- Recall is especially important for imbalanced datasets and measures the proportion of actual positives correctly identified. Sensitivity or Recall also known as True Positive Rate (TPR). From Figure 2, Sensitivity or Recall can calculate as:

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

- F1-Score is the combination of precision and recall, used to maintain a balance between precision and recall. Frequently use in imbalance dataset. F1-Score can be calculated as:

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Specificity

- Specificity measures the proportion of actual negatives that are correctly identified by the model. Specificity is also known as True Negative Rate (TNR). From Figure 2 Sensitivity can calculate as:
- $Specificity = \frac{TN}{TN+FP}$

Area Under the Receiver Operating Characteristic (AUROC)

- AUROC is calculated as the area under the ROC curve. The ROC curve illustrates how sensitivity and false positive rate change across thresholds, helping visualize classifier performance.
- False Positive rate (FNR) can be calculated as:
- $FNR = \frac{FP}{FP+TN} = 1 - Specificity$

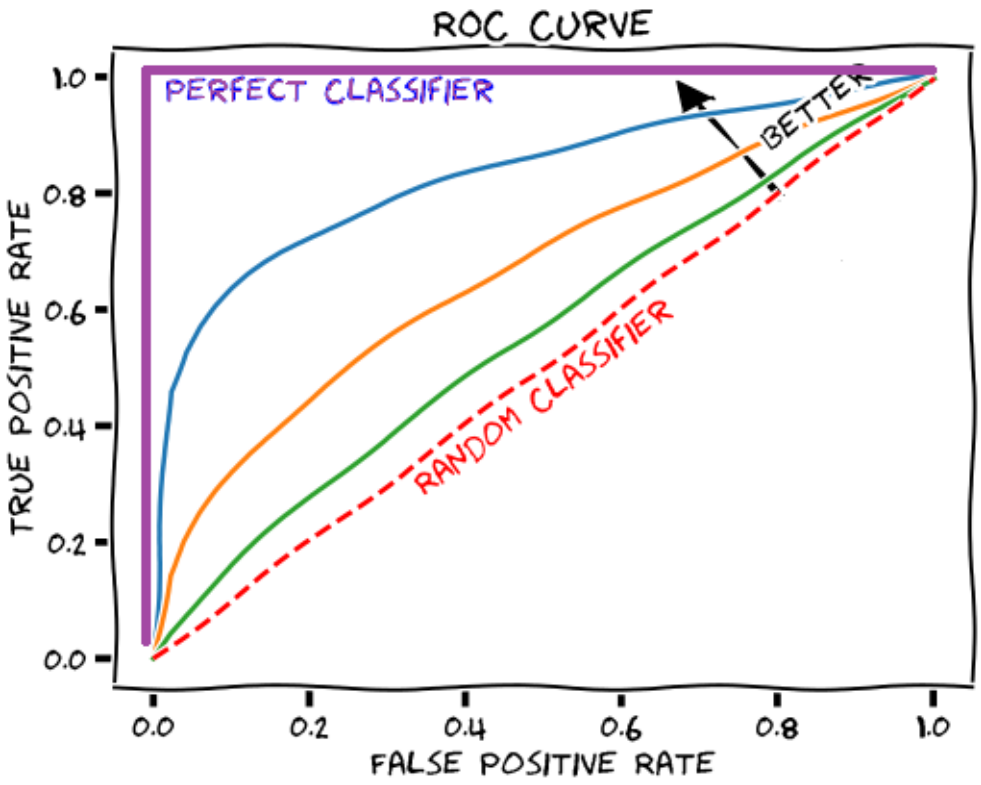


Figure 6 Receiver Operating Characteristic Curve (ROC Curve)

CHAPTER 4

RESULT

Feature Selection

The original dataset included 64 clinical features collected from ICU patients, ranging from vital signs and laboratory values to neurological and sedation scores. To ensure model readiness and clinical relevance, a structured feature selection process was conducted based on data type, completeness, and modeling utility.

Initial Filtering

- Non-numeric fields (e.g., free-text clinical notes or descriptions) were excluded, as they were incompatible with the numerical inputs required for machine learning.
- The identifier column ("Unnamed: 0"), which served only as a patient reference ID, was removed from analysis.

Reduction Based on Data Availability

- Many features with a high percentage of missing values were removed to improve data reliability and reduce imputation bias.
- However, some features with substantial missingness—such as Score (Glasgow Coma Scale)—were still included. This decision was made to improve the model's generalizability across hospitals that may differ in how they collect and record clinical data. Including features with variable availability helps the model adapt to different data environments and supports its use in more diverse clinical settings.

Final Feature Set

- A total of 21 numeric features were selected for model development. These include vital signs, blood pressure components, respiratory

metrics, and ICU-specific scores. Some variables—such as invasive blood pressure and non-invasive blood pressure—are composed of three distinct subcomponents: systolic, diastolic, and mean arterial pressure. Each of these components was treated as a separate input feature. While this increased the number of inputs and allowed for more granular modeling of cardiovascular dynamics, it also contributed to higher missingness for those variables when any individual subcomponent was absent.

- The final list of features, presented in the exact order used during modeling, includes:
  - Temperature (°C)
  - Heart Rate
  - Non-Invasive BP Systolic
  - Non-Invasive BP Mean
  - Non-Invasive BP Diastolic
  - Respiratory Rate
  - O_2 Saturation
  - GCS Total
  - O_2 L/%
  - Pain Goal
  - Pain Score
  - Bedside Glucose
  - Sedation Goal
  - Sedation Score
  - Invasive BP Systolic
  - Invasive BP Mean
  - Invasive BP Diastolic
  - CVP

- P.O. Value
- Delirium Score
- Score (Glasgow Coma Scale)

Figure 7 shows the percentage of missing data for each candidate feature in both sepsis and non-sepsis groups. While many features with excessive missingness were removed, some—such as Score (Glasgow Coma Scale) and Delirium Score—were retained despite missing values exceeding 90%. This decision was made to support the model's generalizability across hospitals, as different institutions may collect different subsets of data. Including these sparse but clinically relevant features allows the model to operate in environments with incomplete data, increasing its real-world applicability. Conversely, features such as urine output, FiO_2, and arterial pH, which had missingness levels above 80% and low predictive contribution, were excluded.

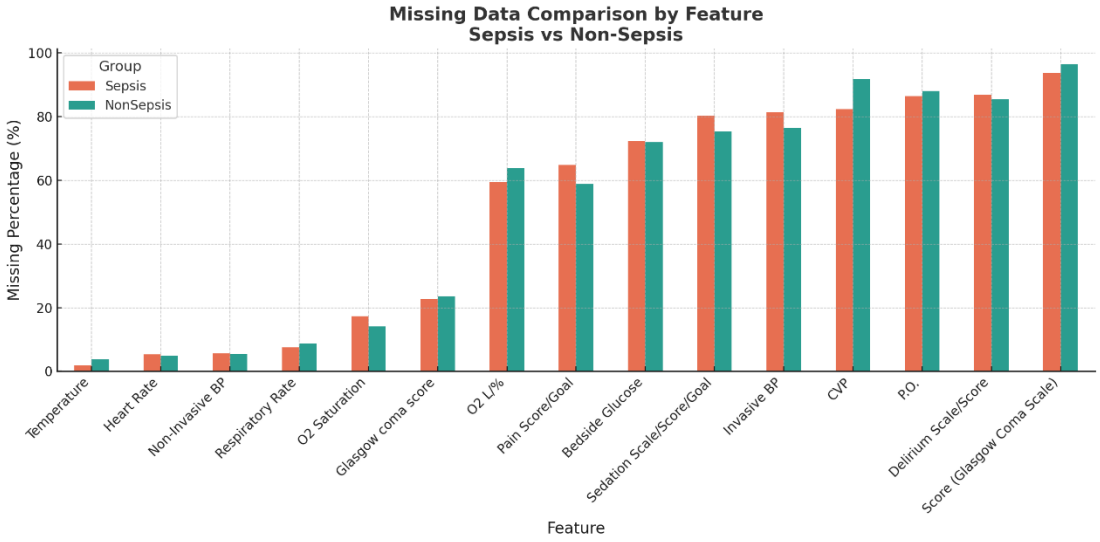


Figure 7 Missing data percentages by feature

Dataset Limitations

While the dataset provided a rich source of ICU patient data for model development, several limitations were identified that could impact the model's performance and generalizability.

High Rate of Missing Data

A number of features had substantial missing values, with some exceeding 90%. Although many of these were excluded, a few were retained to support broader applicability across hospitals with varying data availability. The use of imputation, particularly zero-filling—may introduce noise and limit clinical precision.

Lack of Key Laboratory Markers

Important clinical indicators for early sepsis detection, such as white blood cell count and lactate, were not present in the dataset. Their absence likely reflects variability in measurement frequency or availability across ICU systems. Although related vital signs and scoring metrics were included, the lack of standard lab values limits the model's alignment with clinical diagnostic criteria.

Single Dataset Source

All data were derived from the eICU Collaborative Research Database. Despite being multicenter, this dataset follows a standardized format, which may not fully represent the variability in documentation and workflows seen across all clinical institutions.

Class Imbalance

The dataset may exhibit an imbalance between sepsis and non-sepsis cases. If not fully addressed during training, this could skew the model toward the majority class, potentially lowering sensitivity for detecting sepsis cases.

Class Imbalance Handling

Sepsis was underrepresented compared to non-sepsis cases in the dataset, introducing a class imbalance challenge. To address this, several mitigation strategies were applied across different model types:

- Stratified sampling was used during train-test splits and cross-validation to maintain class proportions.
- The class_weight='balanced' parameter was enabled in models such as Support Vector Machine and Logistic Regression,

automatically adjusting the loss function to penalize misclassification of the minority class.

- While tree-based models like XGBoost did not use scale_pos_weight in this implementation, model evaluation relied on AUC rather than accuracy, reducing the impact of imbalance.
- In deep learning models, class imbalance was indirectly addressed by using AUC as the primary evaluation metric and a custom early stopping strategy that prioritized high validation AUC performance.

These techniques collectively helped reduce false negatives and improve recall, which are critical priorities for early sepsis detection in clinical settings.

Static Features Over Temporal Data

The model used summary statistics (e.g., mean, max) rather than time-series inputs. As a result, it may miss dynamic patterns and temporal trends critical to early sepsis recognition.

Excluded Non-Numeric Features

Clinical text data and categorical information—such as medication names or physician notes—were excluded due to their non-numeric format. This limits the model's ability to incorporate qualitative insights that may hold diagnostic value

Performance Evaluation

Confusion Matrix Analysis

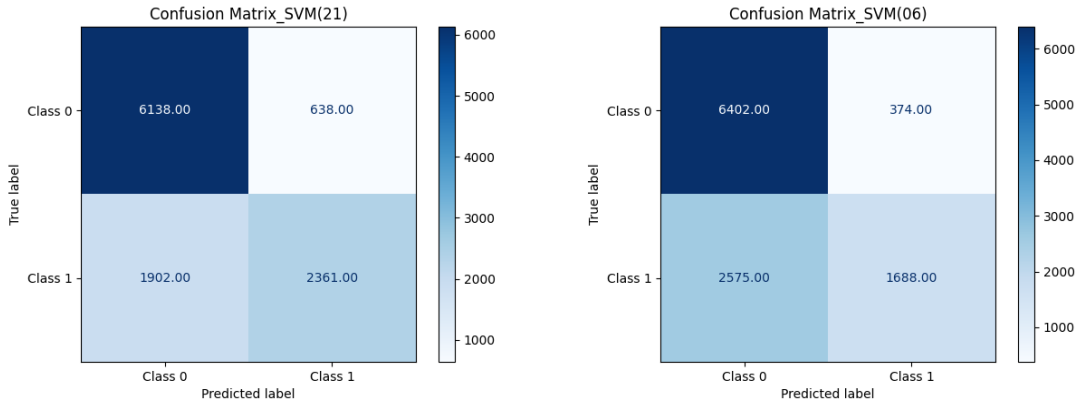


Figure 8 Combined Confusion Matrix for SVM Model

Support Vector Machine (SVM)

- Comprehensive Dataset (SVM_21)
  - TN = 6138, FP = 638
  - FN = 1902, TP = 2361

The use of a comprehensive dataset improves recall by reducing FN and increasing TP, although it slightly increases FP. This version shows better balance.

- Core Dataset (SVM_06)
  - TN = 6402, FP = 374
  - FN = 2575, TP = 1688

This model demonstrates high specificity, accurately identifying non-sepsis cases. However, it suffers from a high false negative rate, indicating that a substantial number of sepsis cases were missing.

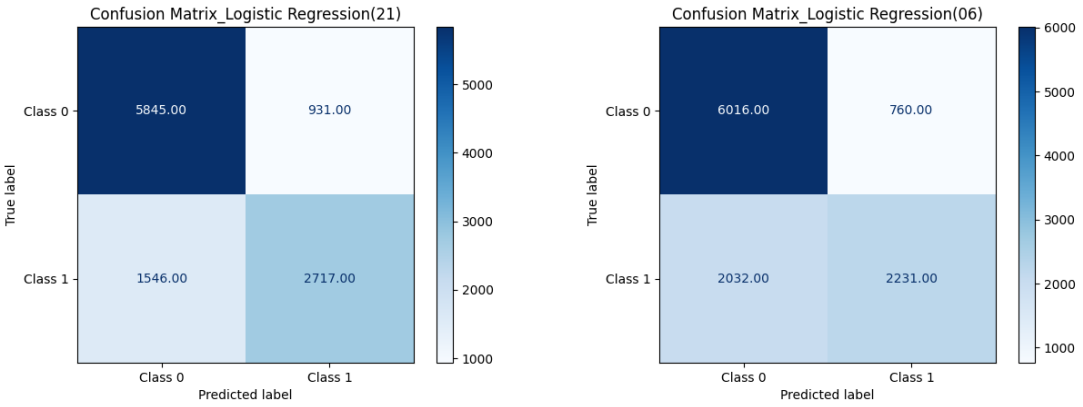


Figure 9 Combined Confusion Matrix for Logistic Regression Model

Logistic Regression (LR)

- Comprehensive Dataset (LR_21)
  - TN = 5845, FP = 931
  - FN = 1546, TP = 2717

Logistic regression benefits significantly from the comprehensive dataset, with a clear increase in TP and reduction in FN. However, this comes at the cost of more false alarms (higher FP).

- Core Dataset (LR_06)
  - TN = 6016, FP = 760
  - FN = 2032, TP = 2231

The model shows modest performance, with a fairly even trade-off between false positives and false negatives. The high FN count still makes it risky in critical settings.

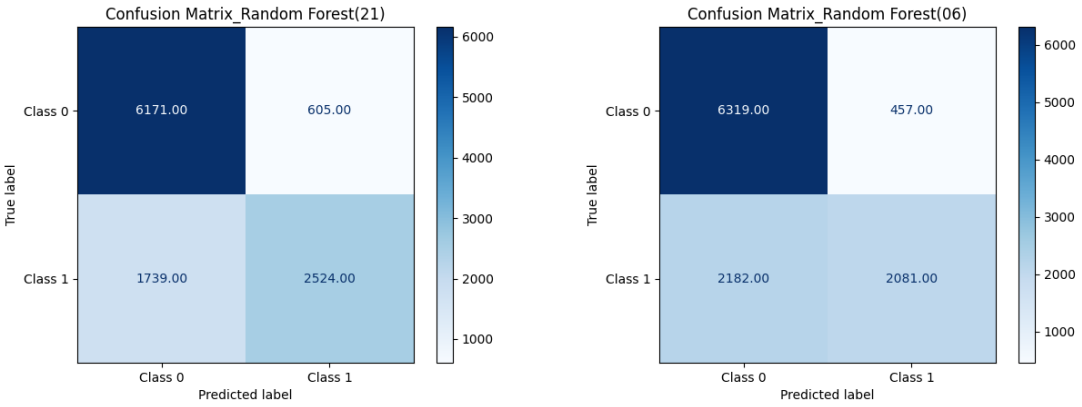


Figure 10 Combined Confusion Matrix for Random Forest Model

Random Forest (RF)

- Comprehensive Dataset (RF_21)
  - TN = 6171, FP = 605
  - FN = 1739, TP = 2524

Performance improves with the comprehensive dataset, particularly in reducing FN and increasing TP, though slightly more non-sepsis cases are misclassified as sepsis.

- Core Dataset (RF_06)
  - TN = 6319, FP = 457
  - FN = 2182, TP = 2081

RF trained on core features shows strong specificity but still misses a considerable number of sepsis cases.

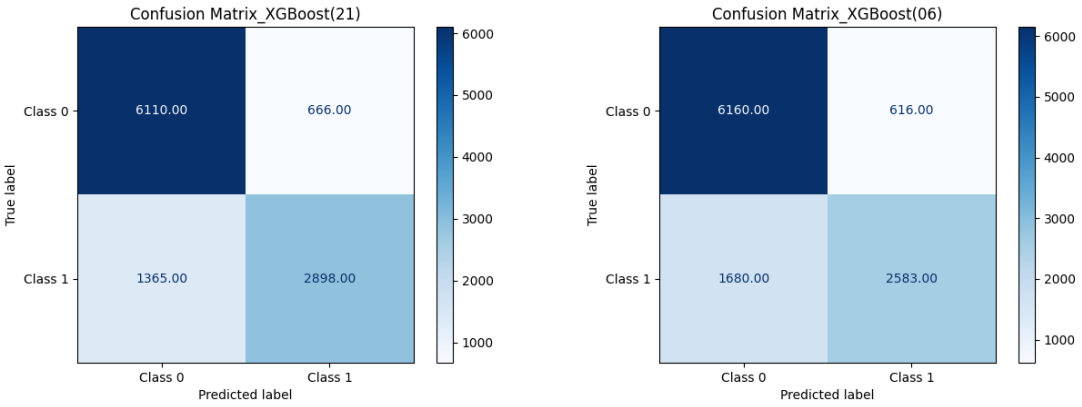


Figure 11 Combined Confusion Matrix for XGBoost Model

XGBoost (XGB)

- Comprehensive Dataset (XGB_21)
  - TN = 6110, FP = 666
  - FN = 1365, TP = 2898

This model achieved the best performance among traditional algorithms. With the lowest FN and highest TP, it presents a strong case for use in real-time clinical sepsis screening.

- Core Dataset (XGB_06)
  - TN = 6160, FP = 616
  - FN = 1680, TP = 2583

Among traditional models, XGBoost shows the most balanced performance, even with fewer features. It maintains low FN and relatively high TP.

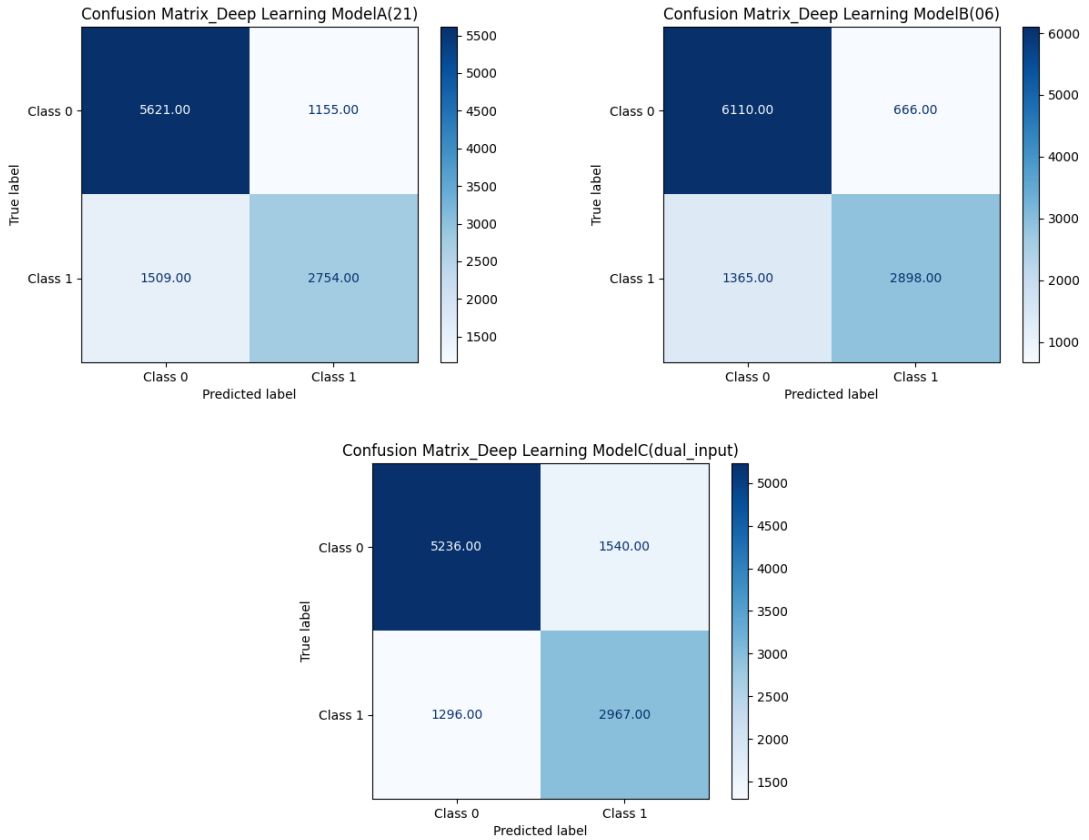


Figure 12 Combined Confusion Matrix for Deep Learning Model

Deep Learning Models

- Model A (Comprehensive Dataset)
  - TN = 5621, FP = 1155
  - FN = 1509, TP = 2754

Deep learning Model A captures more sepsis cases than most traditional models but at the cost of increased false positives. The FN is moderate and significantly better than some of the machine learning counterparts.

- Model B (Core Dataset)
  - TN = 6110, FP = 666
  - FN = 1365, TP = 2898

Surprisingly, Model B mirrors the performance of XGBoost with the same dataset. It shows high effectiveness despite fewer features.

- Model C (Combined Dataset)
  - TN = 5236, FP = 1540
  - FN = 1296, TP = 2967

Model C, which uses both datasets together (dual input), achieves the lowest false negative count (FN = 1296) and the highest number of true positives (TP = 2967) among all models. This suggests that combining datasets significantly enhances recall, which is essential for sepsis prediction.

Deep Learning Model C achieved the lowest number of false negatives among all models, making it the most effective for sepsis detection where recall is critical.

XGBoost trained on the comprehensive dataset (Model A) showed the best overall balance between false positives and false negatives, making it the most reliable model across both classes.

Support Vector Machine trained on the core dataset (Model B) had the lowest false positives but the highest false negatives, showing a conservative classification style that may miss sepsis cases.

Across all models, performance improved when trained on the comprehensive dataset. Deep learning models performed best when both datasets were combined, as seen in the strong recall and true positive rate of Model C.

Model Performance Metrics

Table 7 Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	Specificity	AUROC
Support Vector Machine (Model A)	0.77	0.79	0.55	0.65	0.91	0.73
Support Vector Machine (Model B)	0.73	0.82	0.40	0.53	0.95	0.67
Logistic Regression (Model A)	0.78	0.75	0.64	0.69	0.86	0.75
Logistic Regression (Model B)	0.75	0.75	0.52	0.62	0.89	0.71
Random Forest (Model A)	0.79	0.81	0.59	0.68	0.91	0.75
Random Forest (Model B)	0.76	0.82	0.49	0.61	0.93	0.71
XGBoost (Model A)	0.82	0.81	0.68	0.74	0.90	0.88
XGBoost (Model B)	0.79	0.81	0.61	0.69	0.91	0.85
Deep Learning (Model A)	0.76	0.71	0.65	0.67	0.83	0.83
Deep Learning (Model B)	0.67	0.56	0.70	0.63	0.65	0.75
Deep Learning (Model C)	0.74	0.66	0.70	0.68	0.77	0.82

Performance Analysis

Among all models, XGBoost trained on the comprehensive dataset (Model A) achieved the highest performance, with:

- Accuracy: 0.82
- F1-score: 0.74
- AUROC: 0.88

This indicates that XGBoost provides both high precision and recall, making it a strong candidate for sepsis prediction in practice.

Deep Learning Model C, trained with combined datasets, achieved the highest recall (0.70) among all models, along with a strong AUROC of 0.82. This suggests

that while deep learning may produce more false positives, it is less likely to miss sepsis cases—a crucial consideration in clinical applications.

The Support Vector Machine (Model B) and Random Forest (Model B) models trained on the core dataset had high specificity (0.95 and 0.93, respectively) but significantly lower recall (0.40 and 0.49), making them less suitable for detecting sepsis reliably.

Logistic Regression showed moderate and consistent performance, with improvements seen in Model A over Model B, aligning with other models' trends: the comprehensive dataset generally yields better overall results.

CHAPTER 5

SUMMARY DISCUSSION AND SUGGESTION

Summary

This study shows that both traditional machine learning and deep learning models can be effective for sepsis prediction, particularly when trained on rich, comprehensive datasets. Among all models evaluated, XGBoost and Deep Learning Model C emerged as the most promising. However, performance varies depending on the metric of interest—highlighting the need for careful selection based on the intended clinical use case. With further refinement and validation, these models have the potential to contribute meaningfully to early sepsis detection and improved patient outcomes in critical care settings.

Interpretation of Results

The experimental results demonstrate that predictive model performance varies significantly depending on both the algorithm used and the dataset on which it was trained. Overall, models trained on the comprehensive dataset consistently outperformed those trained on the core dataset, particularly in terms of recall and AUROC.

Among traditional machine learning methods, XGBoost trained on the comprehensive dataset (Model A) achieved the best performance across nearly all metrics. With an accuracy of 0.82, F1-score of 0.74, and AUROC of 0.88, it offered a strong balance between sensitivity and specificity.

In contrast, while Support Vector Machine (Model B) achieved the highest specificity (0.95), its recall was relatively low (0.40)—indicating a higher risk of missing sepsis cases. This underscores the importance of emphasizing recall and false negative rate (FNR) in clinical applications, where missing a sepsis diagnosis can have life-threatening consequences.

Deep learning models showed particularly strong performance in terms of recall. Notably, Model C, which utilized both datasets in a dual-input format, achieved the highest

recall (0.70) while maintaining a solid AUROC (0.82). This model also produced the lowest number of false negatives, which is a critical success indicator for early sepsis detection in intensive care settings.

Dataset Influence on Model Performance

The findings highlight the clear impact of dataset complexity on model accuracy and reliability. The comprehensive dataset, which includes a greater number of features and richer clinical information, enabled models to detect sepsis more accurately. All models trained on the comprehensive dataset showed consistent improvements in both recall and F1-score compared to their counterparts trained on the core dataset.

Interestingly, XGBoost and deep learning models were able to effectively leverage the added complexity without overfitting. In contrast, simpler models like Support Vector Machine showed limited benefit or even performance degradation when exposed to the expanded feature set.

The strong performance of Model C supports the hypothesis that combining feature-rich and high-coverage data inputs enhances a model's ability to generalize across varied ICU patient profiles and hospital systems.

Clinical Implications

In a clinical context, predictive models for sepsis must prioritize early detection (high recall) without overwhelming clinicians with false alarms (high precision and specificity). From this perspective:

Deep Learning Model C is ideal for deployment in early-warning systems, where detecting every possible sepsis case is critical—even at the cost of more false positives.

XGBoost Model A may be more appropriate in settings that require more balanced decision-making, offering a strong trade-off between identifying sepsis and maintaining diagnostic accuracy.

These findings suggest that hybrid deployment strategies could be beneficial—such as using deep learning to trigger initial alerts and XGBoost as a secondary filter to

confirm and prioritize alerts. This approach may reduce alarm fatigue while maintaining clinical sensitivity.

For successful deployment, models must be integrated into existing clinical infrastructure:

Workflow Integration:

The system should operate within electronic health records (EHRs) or ICU dashboards, running continuously in the background and generating risk scores in real time. Alerts should be triggered automatically when predefined risk thresholds are exceeded.

Outcome Impact:

Earlier detection enabled by these models could facilitate more timely interventions, reduce ICU length of stay, and potentially lower mortality rates. These outcomes should be evaluated in prospective studies or simulated clinical workflows.

Regulatory Considerations:

Clinical AI tools must meet regulatory standards such as those outlined by the FDA for Software as a Medical Device (SaMD). Requirements include rigorous external validation, explainability (e.g., SHAP or LIME support), and reproducibility before implementation in critical care settings.

With thoughtful integration and validation, these predictive models have the potential to significantly enhance ICU triage and early sepsis recognition in real-world clinical environments.

Model Deployment Considerations

To translate these predictive models into real-world clinical practice, several deployment strategies can be considered. The goal is to integrate early sepsis detection into existing hospital workflows without overburdening clinical staff or generating excessive false alarms.

Integration with Electronic Health Records (EHRs):

The model can be embedded within hospital EHR systems to run continuously in the background, analyzing incoming patient data in real-time. When predefined risk thresholds are exceeded, the system can automatically trigger alerts to notify physicians or nurses for further assessment.

Tiered Alerting Systems:

A tiered system could be implemented to manage alarm fatigue. For example, a deep learning model (e.g., Model C) can be used as a broad screening tool to maximize sensitivity, while XGBoost (Model A) can serve as a secondary filter to confirm and prioritize alerts, improving precision.

Visualization and Explanation:

To build clinician trust, visual explanations (e.g., SHAP plots, feature contributions) should accompany predictions, highlighting which features contributed most to the risk score.

Threshold Customization by Unit:

Risk score thresholds can be adjusted depending on the ICU type (e.g., surgical vs. medical) or patient population to align with department-specific practices and acceptable risk tolerances.

Real-Time Simulation Testing:

Before full implementation, the model should be tested in a live clinical simulation environment to evaluate real-time performance, response workflows, and any unintended consequences.

By deploying the model with thoughtful integration and layered validation, hospitals can use it to enhance early sepsis detection, improve triage, and potentially reduce ICU mortality.

Limitations

While the study provides promising results, several limitations must be acknowledged:

Generalizability:

All models were trained on data from the eICU Collaborative Research Database. Although this dataset is multicenter, it follows a standardized structure that may not reflect the variability found in real-world hospital systems. Differences in clinical workflows, documentation practices, and variable definitions could affect model performance when applied elsewhere. Without external validation, generalizability remains uncertain.

Class Imbalance:

Sepsis is relatively rare in the dataset, resulting in a class imbalance that can bias models toward the majority (non-sepsis) class. Although metrics like F1-score and AUROC were used and class weighting was applied during training, the imbalance was not fully mitigated and may impact sensitivity.

Temporal Features:

This study relied on statistical summaries (e.g., mean, median, max) rather than time-series data. This limits the model's ability to capture the dynamic progression of clinical signs, which is often critical for early sepsis detection. Time-aware models could offer improved performance.

Limited Interpretability:

Especially in deep learning models, interpretability remains a challenge. Without model explanation tools such as SHAP or LIME, it can be difficult for clinicians to trust or understand how predictions are made, posing a barrier to clinical adoption.

Lack of External Validation:

The models have not been evaluated on independent external datasets (e.g., MIMIC-IV, HiRID). This limits confidence in their real-world applicability and regulatory readiness. External benchmarking is essential for confirming model robustness across diverse healthcare settings.

Recommendations for Future Work

Based on the findings and limitations, several directions for future research are recommended:

External Validation:

Validate the trained models using independent ICU datasets such as MIMIC-IV or HiRID. This step is critical for confirming model generalizability, identifying overfitting, and supporting regulatory and clinical acceptance across diverse healthcare environments.

Interpretability and Explanation Tools:

Apply model-agnostic interpretability frameworks such as SHAP (SHapley Additive exPlanations) or LIME to visualize how input features influence predictions. These tools are essential for clinician trust and explainable AI deployment in safety-critical settings.

Model Ensembling:

Investigate ensemble approaches that combine high-recall deep learning models with high-precision models like XGBoost. Such hybrid strategies can improve the overall diagnostic balance by reducing false negatives while maintaining specificity.

Incorporation of Temporal Models:

Explore sequence-based architectures like LSTM, GRU, or transformer models to better capture dynamic patient condition trajectories. This could enhance model sensitivity to subtle changes over time, particularly useful in early-stage sepsis detection.

Real-Time Simulation and Clinical Integration:

Deploy and test models in simulated clinical workflows to evaluate integration feasibility, response timing, and system latency. These simulations can uncover implementation challenges and refine alert strategies before real-world deployment.

Statistical Significance Testing:

While comparative model metrics were reported (e.g., AUROC, F1-score), no formal statistical testing was performed to assess whether these differences were significant. Future studies should use paired statistical tests (e.g., Wilcoxon signed-rank or DeLong test) on cross-validation folds to confirm whether observed performance gaps are statistically meaningful.

REFERENCES

1. Sepsis [Internet]. [cited 2023 Aug 10]. Available from: https://www.who.int/health-topics/sepsis

2. WHO calls for global action on sepsis - cause of 1 in 5 deaths worldwide [Internet]. [cited 2023 Aug 10]. Available from: https://www.who.int/news/item/08-09-2020-who-calls-for-global-action-on-sepsis---cause-of-1-in-5-deaths-worldwide

3. Guideline Development Group meeting on updating WHO recommendations on management of serious bacterial infection in young infants aged 0-59 days [Internet]. [cited 2023 Oct 24]. Available from: https://www.who.int/news-room/articles-detail/guideline-development-group-meeting-on-updating-who-recommendations-on-management-of-serious-bacterial-infection-in-young-infants-aged-0-59-days

4. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016 Feb 23;315(8):801–10.

5. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med. 2021 Nov;47(11):1181–247.

6. Mogri M, Grant RW, Liu VX. Use of Sepsis Clinical Prediction Models to Improve Patient Care. JAMA Internal Medicine. 2023 Jun 1;183(6):612–5.

7. Lyons PG, Hofford MR, Yu SC, Michelson AP, Payne PRO, Hough CL, et al. Factors Associated With Variability in the Performance of a Proprietary Sepsis Prediction Model Across 9 Networked Hospitals in the US. JAMA Internal Medicine. 2023 Jun 1;183(6):611–2.

8. Reyna M, Josef C, Jeter R, Shashikumar S, Moody B, Westover MB, et al. Early Prediction of Sepsis from Clinical Data: The PhysioNet/Computing in Cardiology Challenge 2019 [Internet]. PhysioNet; [cited 2023 Aug 11]. Available from: https://physionet.org/content/challenge-2019/1.0.0/

9. Reyna MA, Josef CS, Jeter R, Shashikumar SP, Westover MB, Nemati S, et al. Early Prediction of Sepsis From Clinical Data: The PhysioNet/Computing in Cardiology Challenge 2019. Critical Care Medicine. 2020 Feb;48(2):210.

10. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data. 2018 Sep 11;5(1):180178.

11. Pollard TJ, Johnson AEW, Raffa J, Badawi O. The eICU Collaborative Research Database [Internet]. physionet.org; 2017 [cited 2023 Nov 21]. Available from: https://physionet.org/content/eicu-crd/

12. Merck Manual Professional Edition [Internet]. [cited 2025 Mar 17]. Overview of the Immune System - Immunology; Allergic Disorders. Available from: https://www.merckmanuals.com/professional/immunology-allergic-disorders/biology-of-the-immune-system/overview-of-the-immune-system

13. Kumar T, Sharma A, Dutta S, J S, Dutta G, Sharma RP. A Concise Review of Immune System and Natural Immune Modulators. IJPSRR [Internet]. 2021 Jun 15 [cited 2025 Mar 17];68(2). Available from: https://globalresearchonline.net/journalcontents/v68-2/12.pdf

14. Keselowsky BG, Acharya A, Lewis JS. 2.2.3 - Innate and Adaptive Immunity: The Immune Response to Foreign Materials. In: Wagner WR, Sakiyama-Elbert SE, Zhang G, Yaszemski MJ, editors. Biomaterials Science (Fourth Edition) [Internet]. Academic Press; 2020 [cited 2025 Mar 17]. p. 747–75. Available from: https://www.sciencedirect.com/science/article/pii/B9780128161371000490

15. Singh A, Kaur H, Gupta G, Naranje K, Verma A, Roy A, et al. Enhancement of Immunity and Health in Neonates and Infants. Journal of Neonatology. 2021 Sep 1;35(3):138–54.

16. Belkaid Y, Hand TW. Role of the Microbiota in Immunity and Inflammation. Cell. 2014 Mar 27;157(1):121–41.

17. Riksen NP, Netea MG. Immunometabolic control of trained immunity. Molecular Aspects of Medicine. 2021 Feb 1;77:100897.

18. Pal A, Chakravarty AK. Chapter 6 - Genetics for adaptive immunity. In: Pal A, Chakravarty AK, editors. Genetics and Breeding for Disease Resistance of Livestock [Internet]. Academic Press; 2020 [cited 2025 Mar 19]. p. 119–25. Available from: https://www.sciencedirect.com/science/article/pii/B9780128164068000061

19. Pulendran B, Ahmed R. Immunological mechanisms of vaccination. Nat Immunol. 2011 Jun;12(6):509–17.

20. Sviridov D, Miller YI, Bukrinsky MI. Trained Immunity and HIV Infection. Front Immunol [Internet]. 2022 Jul 8 [cited 2025 Mar 21];13. Available from: https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2022.903884/full

21. Gasmi A, Shanaida M, Oleshchuk O, Semenova Y, Mujawdiya PK, Ivankiv Y, et al. Natural Ingredients to Improve Immunity. Pharmaceuticals. 2023 Apr;16(4):528.

22. Nutrition, immunity and COVID-19 | BMJ Nutrition, Prevention & Health [Internet]. [cited 2025 Mar 21]. Available from: https://nutrition.bmj.com/content/3/1/74

23. Markowiak P, Śliżewska K. Effects of Probiotics, Prebiotics, and Synbiotics on Human Health. Nutrients. 2017 Sep;9(9):1021.

24. Lee WS, Wheatley AK, Kent SJ, DeKosky BJ. Antibody-dependent enhancement and SARS-CoV-2 vaccines and therapies. Nat Microbiol. 2020 Oct;5(10):1185–91.

25. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016 Feb 23;315(8):762–74.

26. Shankar-Hari M, Phillips GS, Levy ML, Seymour CW, Liu VX, Deutschman CS, et al. Developing a New Definition and Assessing New Clinical Criteria for Septic Shock: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016 Feb 23;315(8):775–87.

27. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. The Lancet. 2020 Jan 18;395(10219):200–11.

28. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. Lancet. 2018 Jul 7;392(10141):75–87.

29. Plunkett A, Tong J. Sepsis in children. BMJ. 2015 Jun 9;350:h3017.

30. Fleischmann-Struzek C, Mellhammar L, Rose N, Cassini A, Rudd KE, Schlattmann P, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. Intensive Care Med. 2020 Aug;46(8):1552–62.

31. Hotchkiss RS, Moldawer LL, Opal SM, Reinhart K, Turnbull IR, Vincent JL. Sepsis and septic shock. Nat Rev Dis Primers. 2016 Jun 30;2(1):1–21.

32. van der Poll T, van de Veerdonk FL, Scicluna BP, Netea MG. The immunopathology of sepsis and potential therapeutic targets. Nat Rev Immunol. 2017 Jul;17(7):407–20.

33. Delano MJ, Ward PA. The immune system's role in sepsis progression, resolution, and long-term outcome. Immunol Rev. 2016 Nov;274(1):330–53.

34. Levy MM, Evans LE, Rhodes A. The Surviving Sepsis Campaign Bundle: 2018 update. Intensive Care Med. 2018 Jun;44(6):925–8.

35. Iba T, Levy JH, Warkentin TE, Thachil J, van der Poll T, Levi M, et al. Diagnosis and management of sepsis-induced coagulopathy and disseminated intravascular coagulation. J Thromb Haemost. 2019 Nov;17(11):1989–94.

36. Sanfilippo F, Corredor C, Fletcher N, Landesberg G, Benedetto U, Foex P, et al. Diastolic dysfunction and mortality in septic patients: a systematic review and meta-analysis. Intensive Care Med. 2015 Jun;41(6):1004–13.

37. Landry DW, Oliver JA. The pathogenesis of vasodilatory shock. N Engl J Med. 2001 Aug 23;345(8):588–95.

38. Vincent JL, Rello J, Marshall J, Silva E, Anzueto A, Martin CD, et al. International study of the prevalence and outcomes of infection in intensive care units. JAMA. 2009 Dec 2;302(21):2323–9.

39. Micek ST, Wunderink RG, Kollef MH, Chen C, Rello J, Chastre J, et al. An international multicenter retrospective study of Pseudomonas aeruginosa nosocomial pneumonia: impact of multidrug resistance. Crit Care. 2015 May 6;19(1):219.

40. Angus DC, Poll T van der. Severe Sepsis and Septic Shock. New England Journal of Medicine. 2013 Aug 29;369(9):840–51.

41. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Crit Care Med. 2013 Feb;41(2):580–637.

42. Sartelli M, Catena F, Abu-Zidan FM, Ansaloni L, Biffl WL, Boermeester MA, et al. Management of intra-abdominal infections: recommendations by the WSES 2016 consensus conference. World J Emerg Surg. 2017;12:22.

43. Miller RR, Lopansri BK, Burke JP, Levy M, Opal S, Rothman RE, et al. Validation of a Host Response Assay, SeptiCyte LAB, for Discriminating Sepsis from Systemic Inflammatory Response Syndrome in the ICU. Am J Respir Crit Care Med. 2018 Oct;198(7):903–13.

44. Marik P, Bellomo R. A rational approach to fluid therapy in sepsis. Br J Anaesth. 2016 Mar;116(3):339–49.

45. Poston JT, Koyner JL. Sepsis associated acute kidney injury. BMJ. 2019 Jan 9;364:k4891.

46. Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, et al. Acute respiratory distress syndrome. Nat Rev Dis Primers. 2019 Mar 14;5(1):18.

47. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. The Lancet Haematology. 2020 Jun 1;7(6):e438–40.

48. Zhou X. A study of machine learning applications in healthcare. Applied and Computational Engineering. 2024 Nov 8;102:128–33.

49. Shailaja K, Seetharamulu B, Jabbar MA. Machine Learning in Healthcare: A Review. In: 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA) [Internet]. 2018 [cited 2025 Mar 24]. p. 910–4. Available from: https://ieeexplore.ieee.org/document/8474918

50. Bolón-Canedo V, Remeseiro B, Alonso-Betanzos A, Campilho A. Machine learning for medical applications. In 2016 [cited 2025 Mar 24]. Available from: https://www.semanticscholar.org/paper/Machine-learning-for-medical-applications-Bol%C3%B3n-Canedo-Remeseiro/10bb593cbff79fe1562304005474b7642d4ad8db?utm_source=consensus

51. Mitra D, Paul A, Chatterjee S, Mitra D, Paul A, Chatterjee S. https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-3092-4.ch002. IGI Global Scientific Publishing; 1AD [cited 2025 Mar 24]. Machine Learning in Healthcare. Available from: https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/271745

52. Anbarasan K. AI Innovation in Medical Imaging Diagnostics.

53. Machine Learning in Healthcare System – IJSREM [Internet]. [cited 2025 Mar 24]. Available from: https://ijsrem.com/download/machine-learning-in-healthcare-system/

54. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017 Dec;42:60–88.

55. Support Vector Machine (SVM) Algorithm - Javatpoint [Internet]. [cited 2023 Nov 23]. Available from: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm

56. Logistic regression: Definition, Use Cases, Implementation [Internet]. [cited 2023 Nov 23]. Available from: https://www.v7labs.com/blog/logistic-regression, https://www.v7labs.com/blog/logistic-regression

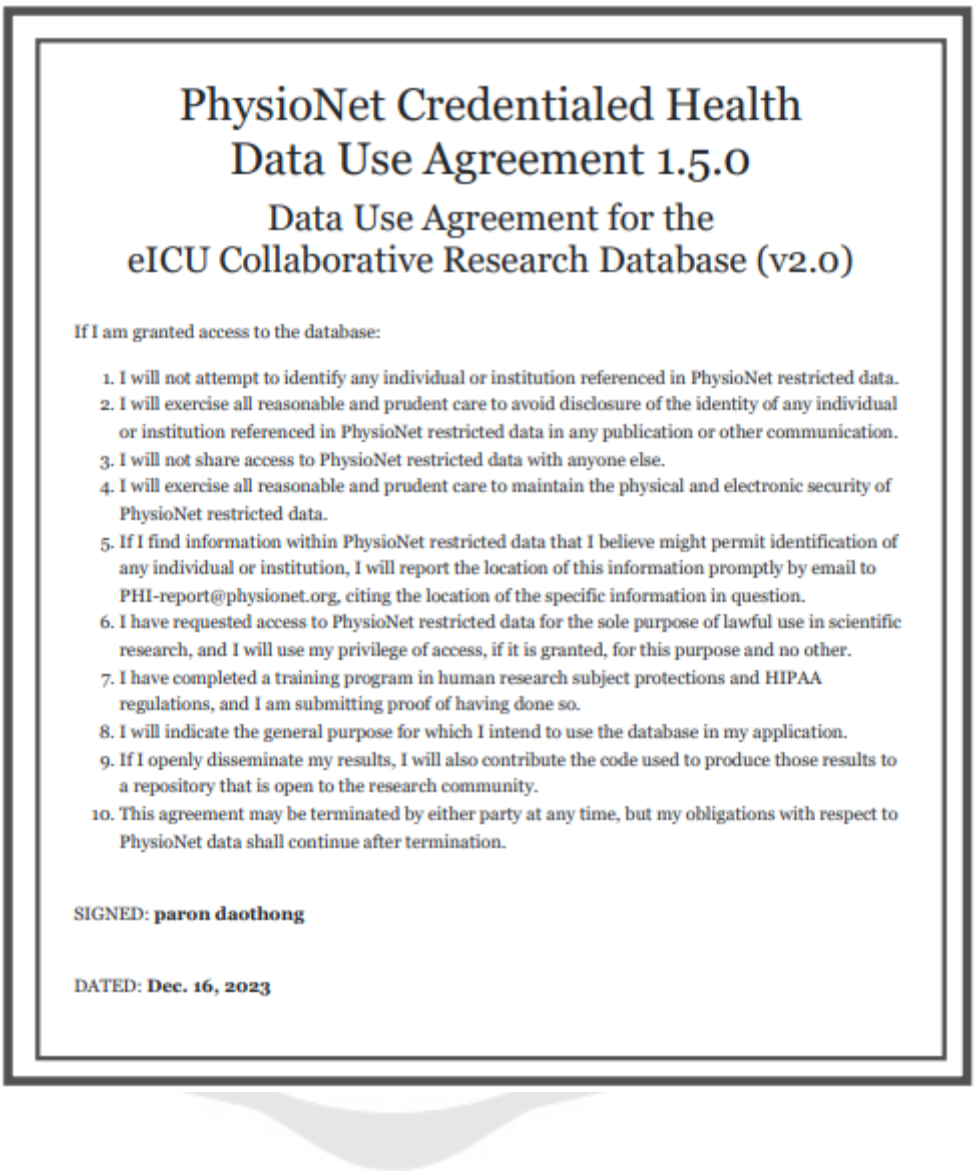
57. R SE. Understand Random Forest Algorithms With Examples (Updated 2023) [Internet]. Analytics Vidhya. 2021 [cited 2023 Nov 23]. Available from: https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/

58. NVIDIA Data Science Glossary [Internet]. [cited 2023 Nov 23]. What is XGBoost? Available from: https://www.nvidia.com/en-us/glossary/data-science/xgboost/

59. Tiwari A. Chapter 2 - Supervised learning: From theory to applications. In: Pandey R, Khatri SK, Singh N kumar, Verma P, editors. Artificial Intelligence and Machine Learning for EDGE Computing [Internet]. Academic Press; 2022 [cited 2023 Nov 22]. p. 23–32. Available from: https://www.sciencedirect.com/science/article/pii/B9780128240540000265

Appendix

Data Use Agreement



PhysioNet Credentialed Health Data Use Agreement 1.5.0

Data Use Agreement for the eICU Collaborative Research Database (v2.0)

If I am granted access to the database:

1. I will not attempt to identify any individual or institution referenced in PhysioNet restricted data.
2. I will exercise all reasonable and prudent care to avoid disclosure of the identity of any individual or institution referenced in PhysioNet restricted data in any publication or other communication.
3. I will not share access to PhysioNet restricted data with anyone else.
4. I will exercise all reasonable and prudent care to maintain the physical and electronic security of PhysioNet restricted data.
5. If I find information within PhysioNet restricted data that I believe might permit identification of any individual or institution, I will report the location of this information promptly by email to PHI-report@physionet.org, citing the location of the specific information in question.
6. I have requested access to PhysioNet restricted data for the sole purpose of lawful use in scientific research, and I will use my privilege of access, if it is granted, for this purpose and no other.
7. I have completed a training program in human research subject protections and HIPAA regulations, and I am submitting proof of having done so.
8. I will indicate the general purpose for which I intend to use the database in my application.
9. If I openly disseminate my results, I will also contribute the code used to produce those results to a repository that is open to the research community.
10. This agreement may be terminated by either party at any time, but my obligations with respect to PhysioNet data shall continue after termination.

SIGNED: **paron daothong**

DATED: **Dec. 16, 2023**

Figure 13 Signed Data Use Agreement

ETHICAL APPROVAL

AF20-03-03.1
August, 2023

Certificate of Exemption

This is to certify that

Protocol Title: UTILIZING MACHINE LEARNING PREDICTIVE ANALYTICS TO ENHANCE EARLY SEPSIS DIAGNOSIS IN CRITICAL CARE SETTING
Principal investigator: Mr. Paron Daothong
Institution: Faculty of Graduate School, Srinakharinwirot University
Protocol code: SWUEC-671035

The Human Research Ethics Committee of Srinakharinwirot University agreed that this research study has met the criteria of the Exemption Determination Regulations and considered exempt from the full review process. If the changes are made to the research protocol regarding research methodology and target population, a new research protocol must be submitted for approval from the ethical committee. Upon completion of the research study, please submit the protocol closing form and the complete research report, or a copy of a published journal article. The Ethics and Research Standards Division will archive your research documents for 3 years after the date of approval.

Date of approval: 24/04/2024
Date of expiration: 23/04/2027

(Dr. Sureeporn Patrasuwan M.D.)
Chairman, Human and Research Ethics Committee (Panel 1)
Srinakharinwirot University

Ethics and Research Standards Devision
Innovation Building Prof. Dr. Saroch Buasri, Floor 17
Srinakhanarinwirot University, Thailand
Tel.: +66-26-495000, 17503, 17506 Fax: +66-22-042590

Figure 14 Ethical Approval

Table 8 Missing Data Comparison

No.	Name	Missing Data (Percentage)		
		Sepsis	Non-Sepsis	Overall
1	Temperature	1.78	3.66	3.44
2	Heart Rate	5.33	4.85	4.9
3	Non-Invasive BP	5.56	5.47	5.48
4	Respiratory Rate	7.45	8.64	8.5
5	O2 Saturation	17.25	14.03	14.41
6	Glasgow coma score	22.66	23.62	23.51
7	O2 L/%	59.33	63.69	63.18
8	Pain Score/Goal	64.77	58.81	59.51
9	Bedside Glucose	72.22	71.93	71.97
10	O2 Admin Device	78.2	82.31	81.83
11	Sedation Scale/Score/Goal	80.21	75.27	75.85
12	Invasive BP	81.24	76.44	77
13	CVP	82.27	91.79	90.68
14	P.O.	86.36	88.02	87.83
15	Delirium Scale/Score	86.81	85.34	85.51
16	Best Motor Response	89.58	92.72	92.35
17	Best Verbal Response	89.59	92.72	92.36
18	Best Eye Response	91.24	95.12	94.67
19	Pain Assessment	92.97	90.69	90.96
20	Score (Glasgow Coma Scale)	93.67	96.5	96.17
21	Fall Risk	93.76	94.24	94.19
22	Respiratory Assessment	93.8	92.44	92.6
23	CV/ PV Assessment	93.8	92.45	92.6
24	Neurological Assessment	93.8	92.47	92.62

Table 8 (Continued)

No.	Name	Missing Data (Percentage)		
		Sepsis	Non-Sepsis	Overall
25	Gastrointestinal Assessment	93.82	92.51	92.66
26	Genitourinary Assessment	93.82	92.55	92.7
27	Musculoskeletal Assessment	93.83	92.59	92.73
28	Mental Status Assessment	93.84	92.55	92.7
29	Integumentary Assessment	93.84	92.57	92.72
30	Pulse Ox Mode	93.85	92.51	92.67
31	Eye, Ear, Nose, Throat Assessment	93.85	92.61	92.75
32	Level of Assistance	93.92	92.71	92.85
33	SEDATION SCORE	93.93	95.48	95.29
34	Eye Opening	93.95	93	93.12
35	Pulse	94.02	95.24	95.1
36	Electrolyte Replacement	94.39	96.58	96.33
37	Patient s Comfort/Function (Pain) GOAL At Rest	94.51	93.11	93.27
38	Symptoms of Delirium Present	94.77	94.81	94.8
39	Motor Response	95.09	94.15	94.26
40	Verbal Response	95.09	94.15	94.26
41	RASS	95.73	96.95	96.81
42	MAP (mmHg)	95.74	95.45	95.48
43	End Tidal CO2	96.13	96.61	96.56
44	Arterial Line MAP (mmHg)	97.54	96.81	96.89
45	ECG (secs)	98.17	97.96	97.98
46	SpO2	98.21	98.22	98.22
47	CVP (mmHg)	98.3	98.87	98.8
48	Pain Present	98.38	97.26	97.39
49	CI	98.87	96.32	96.61

Table 8 (Continued)

No.	Name	Missing Data (Percentage)		
		Sepsis	Non-Sepsis	Overall
50	SV	99.32	98.23	98.36
51	CO	99.35	97.54	97.75
52	SVR	99.7	97.03	97.34
53	SVO2	99.72	98.45	98.6
54	PA	99.78	96.2	96.62
55	SVRI	99.85	98.77	98.89
56	PVR	99.88	99.36	99.42
57	PAOP	99.91	99.62	99.66
58	ICP	99.93	99.38	99.44
59	PVRI	99.95	99.72	99.74
60	CPP	99.96	99.54	99.59
61	IAP	99.96	99.95	99.95
62	ECMO		100	
63	Impella		100	
64	LVAD		100	

VITA