Development of an Artificial Intelligence-Based Model for Patient's Vital Signs Deterioration Prediction

Aluko Oluwadare Abiodun[1]; Olawuni Adeolu[2];
Babalola Donald Abayomi[3]; Oyebiyi Adewale Julius[4]

[1,3,4]Department of Computer Engineering, School of Engineering,
Federal Polytechnic, Ile-Oluji, Ondo State
[2]Department of Electrical/Electronic Engineering, School of Engineering,
Federal Polytechnic, Ile-Oluji, Ondo State

Publication Date: 2025/10/13

Abstract: This research focuses on establishing reliable machine learning models for clinical decision support by highlighting the crucial roles of data preprocessing and quality. Using the MIMIC-IV database, we developed and validated algorithms based on vital physiological indicators, including blood pressure, temperature, heart rate, respiratory rate, and blood oxygen saturation (SpO2).. The study reveals that the quality of the data provided to machine learning models significantly impacts their performance and reliability in clinical environments. To preserve data accuracy and quality, we have enforced rigorous data preprocessing and quality control guidelines, involving univariate and multivariate analyses. The refined data was utilized to educate an artificial neural network (ANN), which formulated an Early Warning Score (EWS) system. Remarkable performance was displayed by the model, with perfect classification scores (precision, recall, F1 score, and accuracy all equaling 1.0) for individual vital sign predictions. Additionally, the model's MSE and MAE were close to zero, indicating negligible error in the regression metrics. The AUC curve's area was consistently high across all parameters (ranging from 0.992 to 1.000), while the validation accuracy ranged from 94.6% to 100%. Such results are achievable when using high-quality data. Conversely, they also illustrate the negative effects of compromised data quality on performance. In conclusion, the successful development and trustworthy deployment of machine learning systems in healthcare settings rely on robust data preprocessing and quality control, as this research illustrates.

Keywords*: Early Warning Systems (EWS), Machine Learning (ML), Area Under the ROC Curve (AUC), Artificial Neural Network (ANN).*

How to Cite: Aluko Oluwadare Abiodun; Olawuni Adeolu; Babalola Donald Abayomi; Oyebiyi Adewale Julius (2025) Development of an Artificial Intelligence-Based Model for Patient's Vital Signs Deterioration Prediction.
International Journal of Innovative Science and Research Technology, 10(9), 2950-2963.
https://doi.org/10.38124/ijisrt/25sep1438

I. INTRODUCTION

Timely identification of life degradation in clinical settings is the basis of effective health care. Traditionally, early warning values (EWS) were used to monitor important physiological parameters to predict adverse events. However, these evaluation systems are often static and have limited prediction accuracy, according to study by [1]. Further developments in machine learning have emerged new opportunities for developing dynamic data- driven models that can more accurately recognize degradation and provide early intervention. The potential of machine learning models, particularly deep learning-based algorithms like the Artificial Neural Network algorithm, to use structured data and identify significant patient patterns makes them extremely promising for the healthcare industry [1]. Despite its power, these models do not contradict the quality of the data being trained and used. Missing values, erroneous entries, inconsistent formats, and outdated information can severely interfere with the output of the model and lead to insecure clinical recommendations. This study looks at how data quality affects ANN-based machine learning's ability to predict outcomes. This was taught to recognize deterioration in life. Physiological signals such as temperature, systolic blood pressure, respiratory rate, heart rate, and SPO_2 were used to construct predictive models and analyze their performance under high quality conditions. With excellent power metrics, this study lays the basis for

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

further investigation of how performance deteriorates under the conditions of endangered data. Telemedic platforms can combine patients with health service providers, allowing remote deployment and monitoring. Using technology can improve access to healthcare and reduce stress in the health system. Therefore, timely identification of the patient's important functions for effective emergency response and patient degradation for improved health outcomes is extremely important. Delays in detection of critical conditions can lead to increased morbidity, mortality, and health costs. Traditional monitoring methods for critical functions are based on manual observations that are susceptible to human error and inconsistent reporting, [2]. Recent advances in machine learning (ML) and data analysis have made it possible to develop predictive models for patient degradation. However, existing solutions often face challenges in prioritizing critical patients and optimizing emergency response. This model can analyze large data records, identify patterns, and provide early warnings for potential emergencies based on machine learning. Machine learning offers an innovative approach to patient monitoring. Actual applications are hampered by a critical issue: data quality. Patient data collected at hospitals is often scattered with missing, inconsistent, or inaccurate entries due to human error, sensor errors, or integration challenges through medical devices. Such low-quality data can lead to misleading predictions, clinical risk malfunctions, and ultimately negative patient outcomes. Most current studies of patient degradation for ML-based patients focus on algorithmic accuracy, thereby limiting the role of data quality as a determinant of model reliability. There is a gap in systematic assessment of how well these models are cut as the quality of data input changes. The goal of this research is to evaluate performance under ideal circumstances and create degradation predictions using the Artificial Neural Network (ANN) algorithm. In other words, the scale is designed to evaluate possible dismantling when the data quality is altered or deteriorated.

II. LITERATURE REVIEW

As reviewed by [3], the prevention, management, and treatment of disease as well as the enhancement of mental and physical health are all included in healthcare. It includes a variety of services rendered by allied health, nursing, and medical specialists. A health system is a complex network of organizations and individuals working together to improve population health, [3]. According to [4] a health system as a network of organizations and individuals whose primary purpose is to promote, restore, and maintain health. This includes a range of interventions, from prevention and treatment to palliative care, [5]. Among its many components are healthcare facilities, providers, and the community itself. Home healthcare (HHC) provides medical and non-medical services to patients in their homes. Medical services may include diagnosis, treatment, wound care, and vital sign monitoring, and are delivered by licensed healthcare professionals.

- *Clinical Deterioration*

Early definitions concentrated on the outcome (e.g., sepsis or cardiac arrest), whereas current models highlight the presence of abnormalities in vital signs and other clinical observations. The idea of clinical deterioration is intricate and has changed over time [5]. These models are constrained, though, by their dependence on discrete vital sign data and their disregard for other significant patient characteristics that may affect results, such as frailty, co-morbidities, pre-morbid function, and treatment delays. According to [18], clinical deterioration is a major physiological disturbance that shows up as a patient's condition worsens. Furthermore, according to study by [7], this demonstrated that it was a dynamic state that jeopardizes hemodynamic stability and is accompanied by physiological decompensating and subjective or objective findings. Inpatient mortality is determined by the identification of clinical deterioration, which can happen at any time during hospitalization and is impacted by environmental factors.

- *Vital Signs*

In order to guarantee patient safety and direct treatment, vital signs monitoring entails the intermittent or continuous observation of a patient's vital signs. According to [8], these medical indicators are essential for detecting clinical deterioration and offer insights into the body's essential processes. The five main vital indicators are blood pressure (BP), body temperature (BTemp), heart rate (HR), respiration rate (RR), and (peripheral) oxygen saturation (SpO2). Nurses can manually monitor these or use automated devices Vital signs are typically taken in general wards every four to six hours, although wearable technology may allow for continuous monitoring. This might cause wearable technology to become the norm for tracking vital signs. Each vital sign has normal values that vary according to age, sex, weight, level of activity capacity, and general health. At rest, healthy individuals typically have the following normal values: The heart rate (pulse) ranges from 60 to 100 bpm. 12–18 breaths per minute is the breathing rate. Between 90/60 mmHg and 120/80 mmHg, the blood pressure reading [9] Average temperature: 98.6°F (37°C), SpO2: 95%€"100%, and temperature spectrum: 97.8°F (36.5°C) to 99.1°F (37.2°C).

- *Importance of Vital Signs in Machine Learning Models Examine the Function of Vital Signs in Such Models, as well as the Impact of Data Quality on Predicted Accuracy*

The majority of ML-based early warning systems rely on vital indicators. Features including heart rate, respiratory rate, temperature, and oxygen saturation are among the most reliable markers of clinical deterioration, according to[2]. The risk of inaccurate or missing vital sign records, which can lower model reliability, was also brought to light by the study. Using these characteristics, [8] created a machine learning system that, with stringent data quality constraints, attained an accuracy of over 93%. Although ML algorithms can deal with noise to a certain degree, bias is frequently introduced or the prediction results are skewed by poor data quality. The work reviewed by [26] claimed that even the

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

greatest algorithms lose precision and generalizability when training data is inconsistent or lacking. In their comparison of logistic regression models with artificial neural networks, they discovered that the addition of missing values and simulated noise significantly reduced performance. In comparison to models trained without data pretreatment, [8]found a 12% increase in model accuracy using MEDLENS, a framework with integrated data quality checks. In medical AI applications, this emphasizes the need for high-quality data pipelines.

➢ *Software Programming Language (Anaconda for Python).*

The finding from [16} showed that a free and open-source tool for creating and running Python code is called Anaconda Python. It was made by the Python programming company continuum.io. The most well-liked method for learning and using Python in data science, machine learning, and scientific computing is the Anaconda platform. More than 30 million people use it worldwide, and it works with Linux, macOS, and Windows. According to [16], Anaconda Python is widely used because it makes package deployment and maintenance easier. Additionally, it contains a large number of libraries and packages that you can utilize for your projects. Because Anaconda Python is open-source and free, anyone can help with its development. Python is a flexible language that can be utilized on a web application's front end, back end, or entire stack. Even though the standard Python library has a lot of functionality, there are numerous circumstances in which we need modules and libraries that are not included in the standard library. You can create an environment for multiple Python versions and package versions using Anaconda software. Packages in project environments can also be installed, uninstalled, and upgraded using Anaconda. Our main tool for building ML models was Jupyter Notebook. Key packages like pandas, seaborn, matplotlib, and scikit-learn, along with Python 3.11.1, were installed. These packages all come with built-in libraries and functions to aid in data manipulation, analysis, visualization, and model building._Our preferred tool was Jupyter Notebook due to its easy-to-use interface, sophisticated data cleaning capabilities, and ability to quickly develop modeling procedures using the Python programming language, as illustrated in Figure 1.

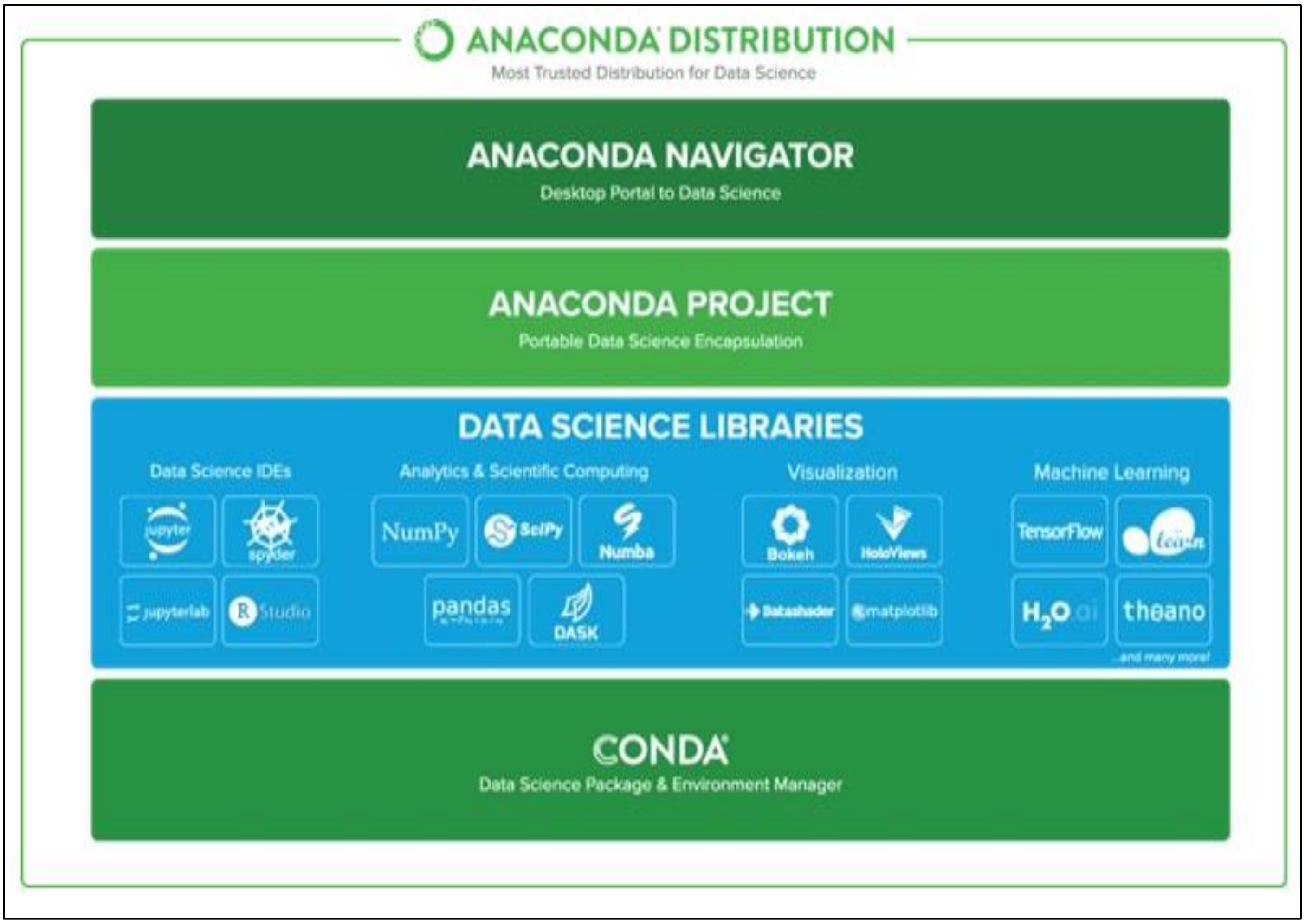


Fig 1 Shows the Interface of Anacoda Software Applications

➢ *Related Work*

Data-driven and AI-enabled technologies have been used more and more in recent healthcare breakthroughs to improve shared decision-making (SDM), clinical decision-making, and patient management. For instance, [11] showed how real-time predictive algorithms that assist dynamic

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

patient care decisions are made possible by the collection of high-frequency intensive care unit (ICU) data stored in cloud infrastructures. Proactive interventions are made possible by the scalable, continuous data analytics made possible by this cloud computing interface. However, there are still major obstacles to general adoption, including worries about data privacy, security, and the environmental impact of massive data centers, which are brought up by this shift to cloud-based systems. In order to encourage person-centered shared decision-making, the integration of patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) into electronic health records (EHRs) was emphasized in the study by [12]. Clinicians can more effectively customize care plans by integrating these digital tools, which provide them access to more contextual information about patients' experiences and outcomes. The full use of such systems is nevertheless hampered by implementation hurdles such clinician time limits, workflow interruption, and opposition to digital adoption, despite the obvious advantages. Tools for clinical prioritization are also essential for increasing the effectiveness of care. Using EHR data, [13] created risk-based informatics models that effectively triage diabetic patients, giving priority to those who are most at risk for negative consequences. The model's effectiveness may be impacted by patient heterogeneity, such as socioeconomic status and demographic variety, which have an impact on both disease presentation and healthcare access, even if this strategy shows promise in maximizing clinical resources. Promising answers to privacy issues in multicenter health data analysis can be found in distributed machine learning systems. Federated learning approaches, in which data stays within local jurisdictions while models are trained collaboratively, are made possible by systems like the Personal Health Train and the Australian Cancer Trials (AusCAT) network [14], [15]. Technical challenges include providing interoperable data formats, harmonizing heterogeneous data sources, and overcoming model generalizability limits across various healthcare contexts, but these approaches also protect patient confidentiality. [16] Discovered that infection risk prediction models employing ensemble machine learning algorithms were very successful in forecasting patient infections based on retrospective hospital data. Despite their efficacy, these models' applicability in other contexts is limited by biases associated with hospital-specific practices and patient groups. Furthermore, [17] emphasized the necessity of strong data governance frameworks and the significant impact that completeness and quality of ICU data have on the dependability of decision-making support systems. Healthcare prediction powers are further improved by creative use of unstructured and multimodal data. The study reviewed by [18] improved hospital readmission predictions by applying natural language processing (NLP) to home healthcare nursing notes. Although the study's results are limited by problems with data quality and its single-center methodology, this method makes use of rich textual data that typical structured records sometimes overlook. In order to create individualized relapse prevention methods for child mental health, [19] integrated patient participation to promote engagement and used multimodal clinical and self-report data. Although these early trials show promise, they need to be validated on a broader scale. Electronic patient-reported outcome (ePRO) applications are one example of a technology that supports goal-oriented care and offers important insights into complex care environments. In their investigation of the adoption and usage of ePRO apps, [26] uncovered anthropological viewpoints that underscore the difficulties associated with hiring, digital literacy, and technology acceptance in settings with multiple stakeholders._Numerous clinical contexts have also looked into the adoption of digital health and shared decision-making. Digital interventions in kidney transplant care and cardiac rehabilitation were examined in recent studies by [21] and [22], respectively. Both stressed how crucial it is to use longitudinal qualitative research to capture changing needs and perceptions by involving a variety of stakeholders, including patients, physicians, and caregivers. However, the generalizability of these research' findings was constrained by their frequent use of single-center designs or small sample numbers. The views of nursing professionals about AI-enabled clinical decision support systems have been studied by [7] and [23]Their polls showed a generally favorable attitude regarding AI's potential to enhance healthcare, but they also raised issues with usability, trust, and the requirement for proper training. These insights' broader relevance is limited by self-report biases and geographical focus, indicating the need for larger, more varied investigations. Digital interventions that are educational and community-focused support therapeutic initiatives. In university settings, [24] showed how game-based learning strategies improve nurse decision-making abilities, encouraging engagement and information retention. [2] Developed an ICT-based person-centered community care platform that improved shared decision-making and the quality of life for community members, despite user requests for more advanced features. A computerized decision support tool for managing chronic pain in underprivileged groups was assessed by [25]. The challenges of integrating such tools into standard care workflows are highlighted by the poor physician uptake and slight increases in clinical outcomes, despite the feasibility having been proven. Predictive modeling has been used to maximize operational efficiency in primary care settings. In order to predict patient no-shows, [26] used machine learning approaches, which improved scheduling and resource allocation while simplifying feature sets. However, the model's forecast accuracy is limited by its failure to incorporate external contextual variables, such as weather and transportation circumstances. An integrated e-health system for diabetes self-management was assessed by [9], who found that it increased quality of life and glycemic control. Notwithstanding these encouraging results, effectiveness differed depending on the type of diabetes, and potential biases are introduced by the lack of participant and physician blinding.

KPI-based health information systems have been shown by [5] to improve performance monitoring and decision-making in ophthalmology departments. However, this study's single-center design limits the applicability of its findings to other medical specialties or systems. Emerging

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

studies and privacy-preserving analytical techniques broaden the scope of digital health in emergency care. Although study findings are still pending, [28] reported a large-scale randomized trial strategy evaluating computerized diagnostic decision support systems in emergency rooms. In order to balance data security and computational resource demands, the study by [27] demonstrated homomorphic encryption for privacy-preserving postoperative mortality prediction across institutions. By addressing conceptual issues with identifying and operationalizing decision-making constructs, [29] provided a rigorously validated scale to measure shared problem-solving in mental healthcare, facilitating improved evaluation of collaborative therapeutic processes. Predictive models with high accuracy are still developing quickly. A LightGBM-based machine learning framework was reported by [8] to predict 30-day mortality among sepsis patients in intensive care units. The framework achieved exceptional accuracy (AUC 0.983) and identified unique predictive factors, including hospital length of stay. However, its generalizability needs to be examined in prospective, real-world cohorts. An AI-powered community-based blood pressure control program targeted at underprivileged groups was piloted by [30]. With 69% of participants reaching blood pressure goals within 12 weeks, the intervention demonstrated therapeutic benefit and feasibility; however, the lack of control groups and poor completion rates weaken the conclusions. Compared to conventional methods, machine learning algorithms have demonstrated promise in identifying patients who are deteriorating earlier. ML models like random forests, XGBoost, and neural networks routinely beat conventional EWS like NEWS and MEWS, according to a thorough review by [33]. The same was true for [32], who used Long Short-Term Memory (LSTM) networks to predict COVID-19 patients' deterioration with over 90% accuracy. However, dependable and regular data inputs are crucial for these models to function. When taken as a whole, these studies show how digital health has advanced in many ways, including shared decision-making, clinical decision support, predictive analytics, and patient participation in a variety of contexts and demographics. Predictive monitoring has been dominated by traditional Early Warning Scores (EWS) such as MEWS and NEWS [32], [33]. While simple and interpretable, these rule-based systems depend on static thresholds that cannot capture the complex, non-linear interactions of physiological variables, often resulting in high false-alarm rates or missed detections. Despite encouraging technical advancements, enduring difficulties still exist. These include improving model generalizability, resolving privacy and ethical issues, eliminating physician adoption barriers, guaranteeing data quality and interoperability, and carrying out larger, more extensive, and more varied evaluations. To fully utilize AI and informatics to enhance patient outcomes, healthcare delivery, and system efficiency globally, these problems must be resolved. The present study addresses these gaps by developing a machine learning–based Early Warning System trained on data from MIMIC III database. It emphasizes data quality assessment and preprocessing while integrating the model into mobile app for real-time monitoring. This approach enhances both methodological innovation and contextual relevance, aiming to improve reliability, clinical utility, and emergency responsiveness.

III. METHODOLOGY

A. Introduction

In order to train and evaluate a predictive model, this study used patient physiological data obtained from the MIMIC III database, a licensed and de-identified vital sign database. Dataset collection, input vital sign variable identification, data processing and analysis steps, predictive model development, and a machine learning model that identifies the most critically ill patient in predicting for emergency levels interpretation, evaluation, and validation using both classification and regression evaluation metrics were all part of the algorithm exploration process for building various models for the patients' vital sign deterioration prediction. To eliminate noise and missing values, the data underwent preprocessing and anonymization. The following tasks were part of the algorithm exploration process in order to create different models for the prediction of patients' vital sign deterioration: i. Real-time data collecting and identification of input vital sign variables. Predictive model development using the chosen machine learning model, ii. Early warning score computation as a reliable variable for identifying critically ill patients, iii. Data preprocessing steps for predictive model development, iv. Integration of machine learning algorithms for prioritization and optimization for emergency responses, evaluation and validation using both classification and logistic evaluation metrics, and clinical guidelines as illustrated in the block diagram in Figure 2.

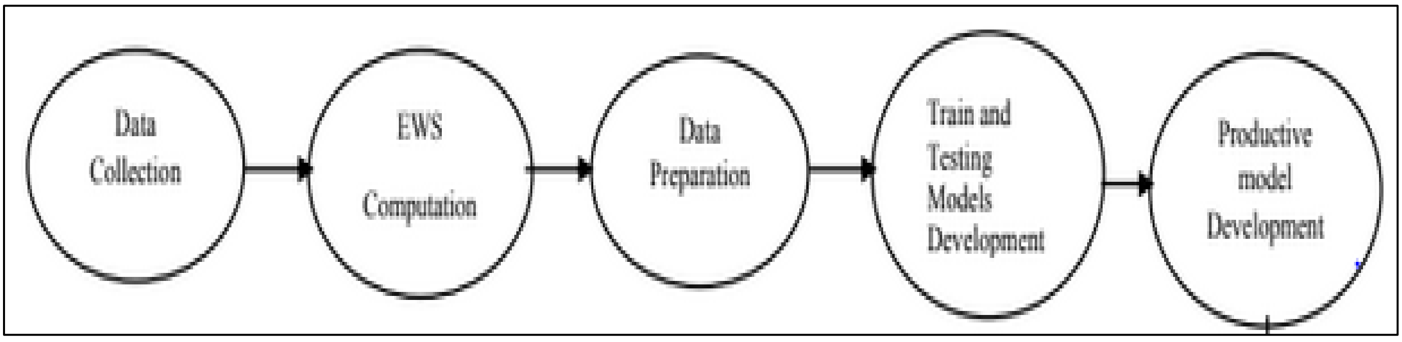


Fig 2 Block Diagram for Proposed Model: Impact of Data Quality on Machine Learning for Patient Deterioration Prediction

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

➢ *Exploratory Data Analysis*

Two (2) techniques were utilized to check user input, data quality, and pre-processing, which are essential to creating an effective prediction model. This ensured that the dataset is well-structured, suited for a variety of analyses, and understood its variability. A clear summary of one variable at a time is provided by univariate analysis, which is used to understand the distribution and characteristics of individual features or parameters. It also aids in the identification of patterns, trends, anomalies, outliers, and data entry errors that may compromise the predictive model's accuracy. ii. Multivariate analysis is used to examine the relationship between variables, including multicol-linearity, numerical variables into categorical variables, and how the target variable changes when categorical features are removed or repeated rows are consolidated. In order to view high-dimensional data, identify anomalies, and comprehend trend seasonality and noise in the dataset, a sophisticated technique for dimensionality reduction and time series decomposition was observed.

➢ *MLEWS Computation*

The Early Warning Score (EWS) is a commonly used instrument for detecting individuals who are at risk of clinical deterioration. EWS is calculated by giving points to physiological indicators that fall outside of normal ranges, with higher scores suggesting a greater likelihood of deterioration. The process of computing the EWS, assigning points to five physiological indicators, and categorizing patients into various risk frames based on their scores. Our ML early warning score (ML EWS), which assigns points to a number of physiological indicators, illustrates the scoring methodology used to gauge an individual's level of disease. These symptoms include: Temperature: 0–2 points; Oxygen Saturation (SpO2): 0–2 points; Heart Rate (HR): 0–3 points; Respiratory Rate (RR): 0–3 points; Systolic Blood Pressure (SBP): 0–3 points Refer to Table 3.1. A score is given to each parameter according to how much it deviates from the norm. The overall MLEWS score, which ranges from 0 to 13, is then calculated by adding the scores.

Table 1 Shows the Standard Values for Allocating Early Warning Score (MLEWS)

Physiological Parameter	**3**	**2**	**1**	**0**	**1**	**2**	**3**
Temperature (◦C)		<35.1	35.1–36.5	36.5–37.3	37.4-38.5	>38.5	
Heart Rate (BPM		<40	40 –50	51–100	101–110	111–130	>130
Respiration Rate (BPM)		<9	9 -11	12–20	21-24	25–30	>30
Oxygen Saturation (SPO2) (%)		<91	91–94	95-100			
Systolic Blood Pressure (mmHg)	<70	70–80	81–89	90–120	121–140	141-180	>180

➢ *Data Preprocessing*

The dataset underwent preprocessing, which included data transformation by encoding categories variables in vital signs; data generalization and regularization; data cleaning; handling missing data; normalization; feature scaling; and pertinent features. To guarantee relevance and quality in our prediction, the data relevance features were also extracted and filtered. Both the numerical and category features were scaled and encoded. Both transformations were used in the preprocessing pipeline to guarantee that missing values were handled correctly and that the data was prepared for model training. The structure of the processed feature set is (50, 42), meaning that after encoding and scaling, we get 50 samples and features.

B. Model Development

The approach of Artificial neural network involves computation which attempts to perform simulations based the behaviour of nerve cells (neuron), The Architecture of Neural network is mainly composed of connecting neuron with a connection which exists between the neurons. In medical procedures, it has been brought to limelight that neural network is very applicable. Preliminary data analysis involves sorting of the attributes into rows and columns which yields a matrix. The matrix is represented as the input; the data label which was part of the data set was used for deriving the target data which is the output. Definition of input, hidden and output layers: Cleveland dataset consists of 8 attributes, therefore the input layer in the model consists of 8 neurons, also the hidden layer consists of 20 neurons while the output layer consist of 1 neuron. The network is ready for training process, the samples (patient's vital signs) will be divided automatically into training, testing and validation sets. The essence of training set is to teach the network and it also continues to improve the measure of the needed network accuracy. Figure 3 shows the training interface for neural network.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

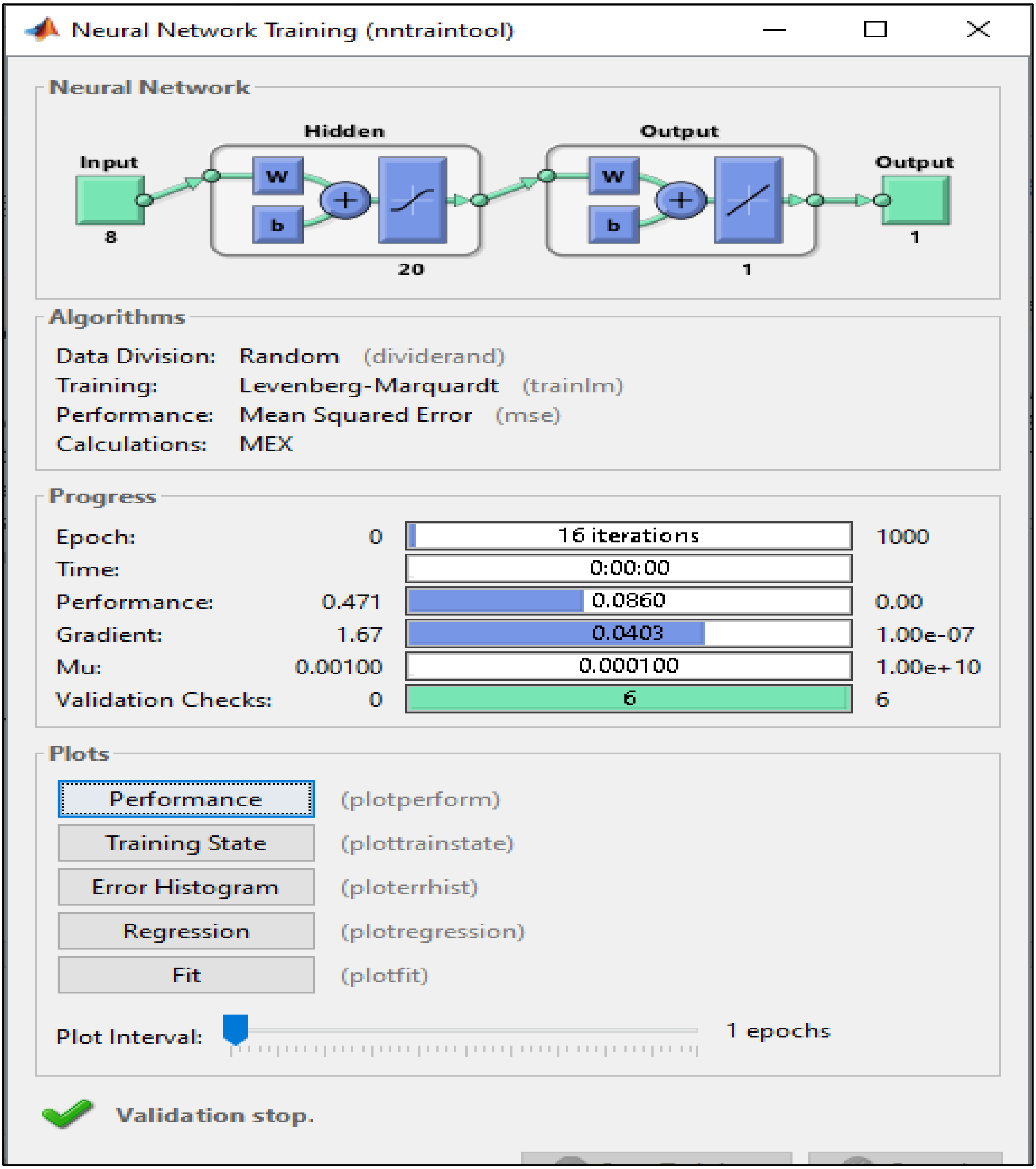


Fig 3 Training Interface for Neural Network

The ANN algorithm was chosen for its efficiency and high performance on structured data The model was evaluated by both classification performance metrics, include: (Recall, Precision F1-score, Sensitivity, and Specificity) and regression performance metrics include: (Mean Absolute Error, Mean Squared Error, and Root Mean Squared Error (RMSE).

- *Model Train-Test Split:*

Our medical data is divided as part of the following procedure. An 80-20 split was used to separate the dataset into training and testing sets. This indicates that 80% of the data was utilized to train the model, with the remaining 20% set aside for cross-validation testing, specifically 5-fold cross-validation. By using the scikit-learn Python library,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

the stratify argument makes sure that the target variable's distribution is maintained in both groups. Training and testing sets of the dataset were being separated. The shapes of each set are as follows:

C. Individual Vital Sign Prediction

The trained model's capacity to predict the individual vital sign deterioration of each patient allowed for the early detection of potential health conditions. The projections were based on the vital signs that were considered for this investigation. Installing Anaconda software with Python 3.11.1 and required packages like pandas, seaborn, matplotlib, and scikit-learn made it easier to manipulate, visualize, analyze, and build machine learning models. These packages come with built-in functions and libraries. Because of its user-friendly interface, advanced data cleaning capabilities, and speedy execution of modeling processes utilizing the Python programming language, Jupyter Notebook was our preferred tool.

D. Moral Aspects to Take into Account

Before analysis, all patient data was anonymised. The dataset was utilized in compliance with relevant data protection laws and ethical guidelines.

IV. RESULTS

Exploratory data analysis was carried out to guarantee data quality before model training, and visualization techniques were used to find patterns, outliers, missing values, and correlations in the datasets. To improve the dataset's appropriateness for predictive modeling, further data pretreatment procedures included transformation, feature engineering, data cleansing, and standardization. Our propose paper also presents the performance evaluation of the proposed machine learning model (ANN) for predicting the deterioration of patient vital signs and triggering emergency responses. The results indicate that prioritizing important patients at high risk of deterioration and improving clinical decision-making can be achieved by combining machine learning approaches with Early Warning Scores (MLEWS). A univariate analysis was performed in order to understand the distributions of individual features and parameters. Table 1 and Figures 4 (a), 4 (b), and 4 (c) present the findings, respectively.

A. Univariate Analysis Result for the Distributions of Key Numerical Variables

This demonstrated the dataset's numerical variables' distribution pattern, revealing that all of the variables are symmetrically distributed and indicating a generally normal pattern.

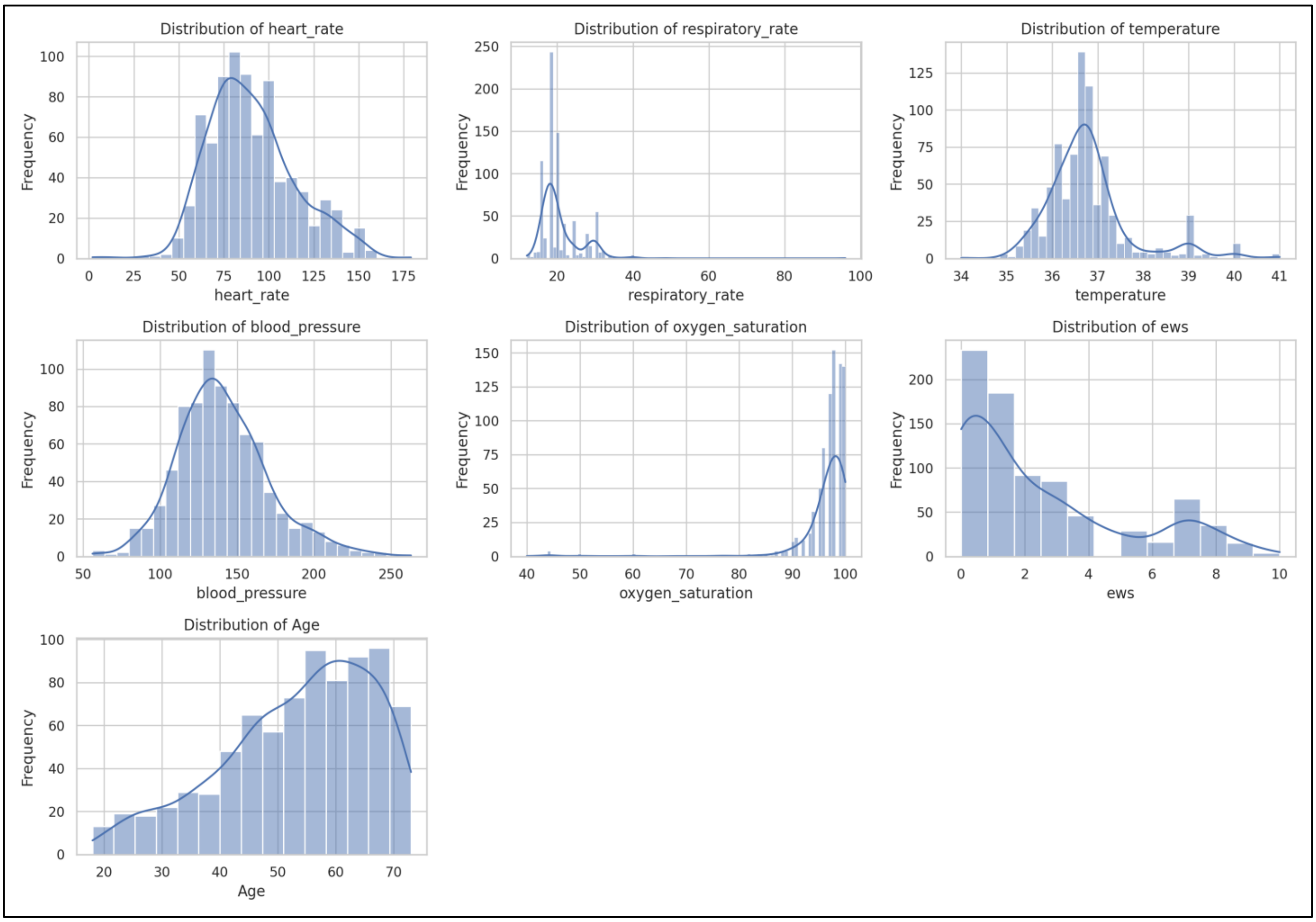


Fig 4 (a) Shows the Distributions of Key Numerical Variables

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

➢ *Univariate Analysis Result for the Outliers*

This demonstrates how the obvious extreme value that deviates from the true range was found. Because outlier data can result in randomized and uncertain outcomes at late stages, as illustrated in Figure 4.1(b), they were removed. By determining the interquartile range (IQR) and using violin plots to visualize the results, the research found outliers in a number of numerical columns.

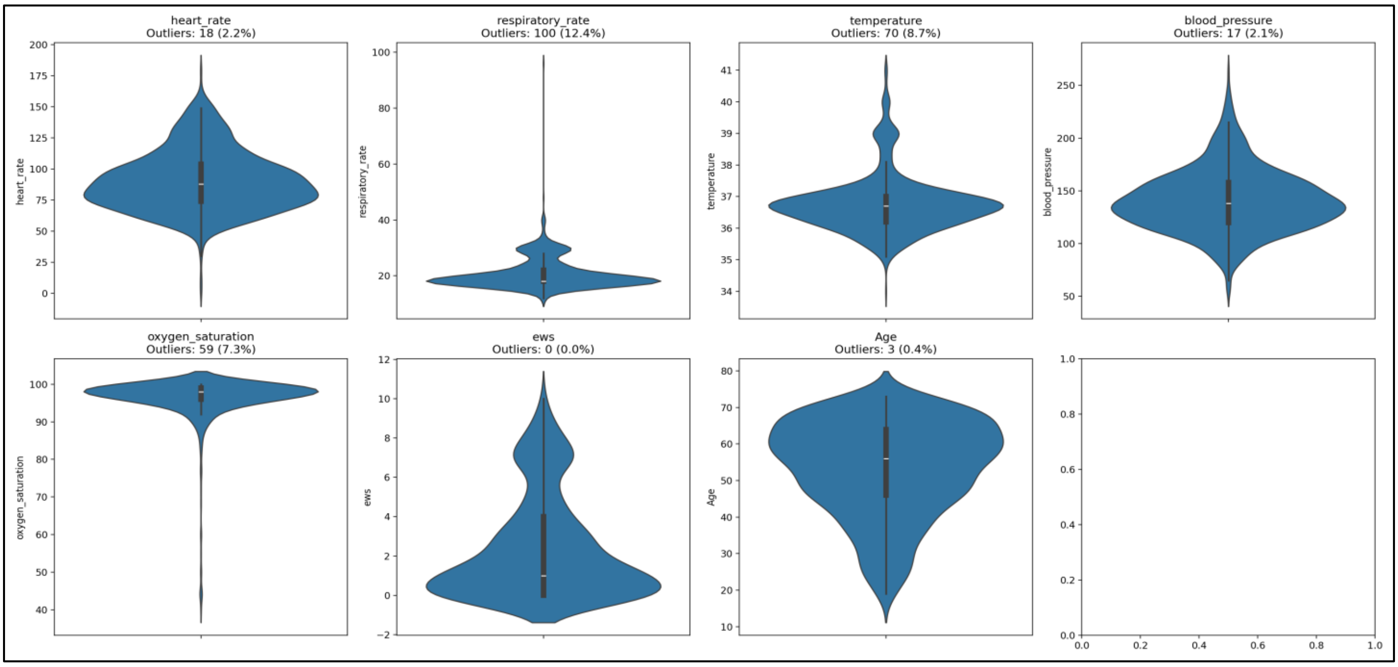


Fig 4 (b) Shows the Visualization of Outlier of Individual Feature in Distributions

➢ *Univariate Analysis Result for the Time Trends for Analysis for the Numeric Parameters*

The way each parameter varies throughout the time index we established is shown visually by these line charts. This offers information on how the patient metric changed over time and demonstrates the discernible trends in the numerical parameters. Trends: By displaying the variations in the corresponding parameter over time, each plot enables us to spot any trends, patterns, or abnormalities, and by observing these patterns, one can gain insight into how each parameter behaves and how it may be related to other parameters or outside influences.

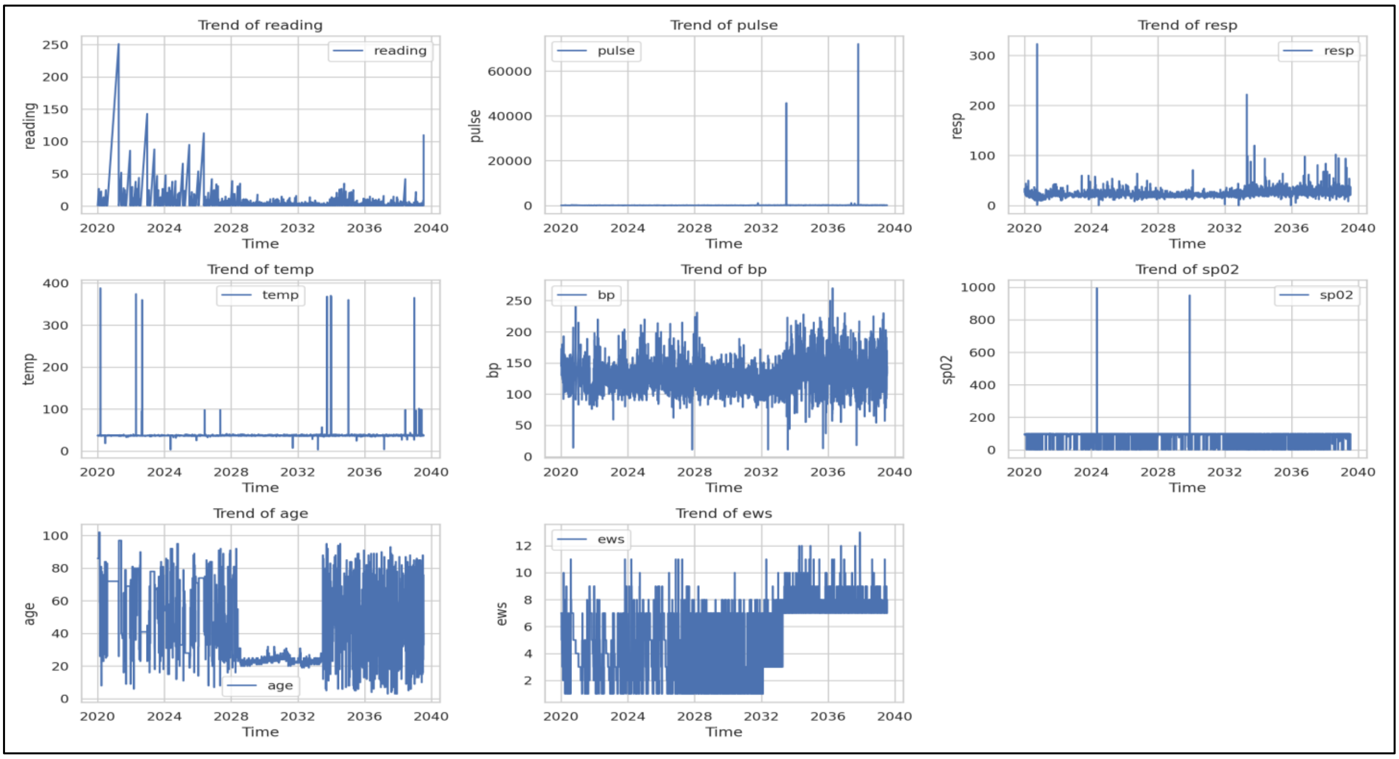


Fig 4 (c) Shows the Time Trends for Analysis for the Numeric Parameters

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

➢ *Multivariate Analysis;*

Key numerical variable distributions that were examined for correlations are shown in Figures 5 (a), and 5 (b), respectively. It assists us in recognizing important patterns and comprehending the connections among the data. As seen in Figure 5 (a), this heatmap offers information about whether factors are positively or negatively associated, which can be helpful for additional research or predictive modeling.

➢ *Results on Multivariate Analysis with Heatmap*

The relationship between the numerical parameters is displayed in the heatmap representation. This illustrates the relationships and correlations between a number of metrics. Values close to 1 indicate strong positive correlations, while values close to -1 indicate strong negative correlations.

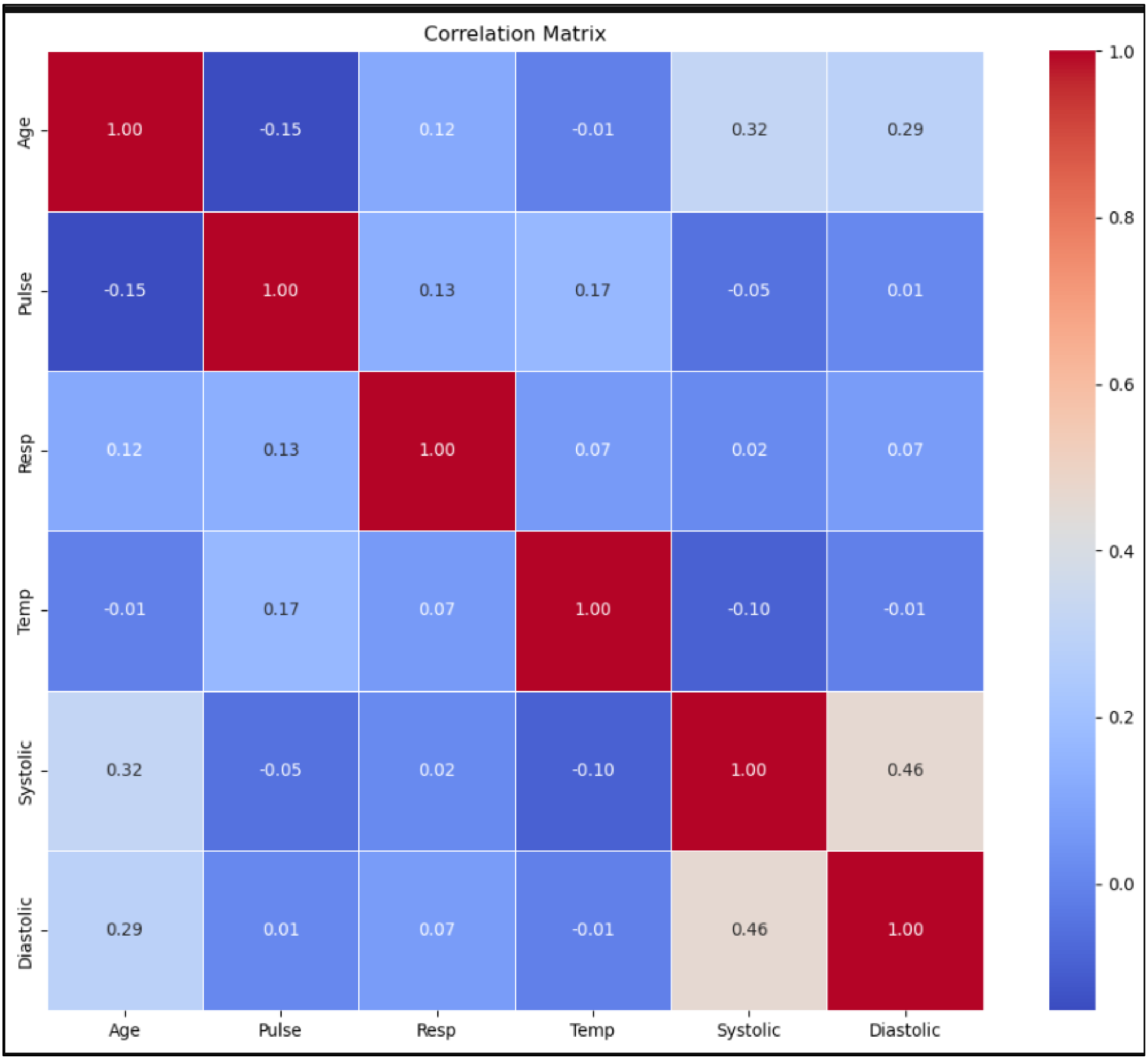


Fig 5 (a) Shows the Visualization of Heatmap for Multi-Variate Analysis

➢ *Results on Multivariate Analysis with Time Series Decomposition*

The "reading" parameter's time series decomposition aids in the analysis of the data's underlying patterns, seasonality, and residuals. Understanding the evolution of the reading data over time is possible thanks to the decomposition graphic.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

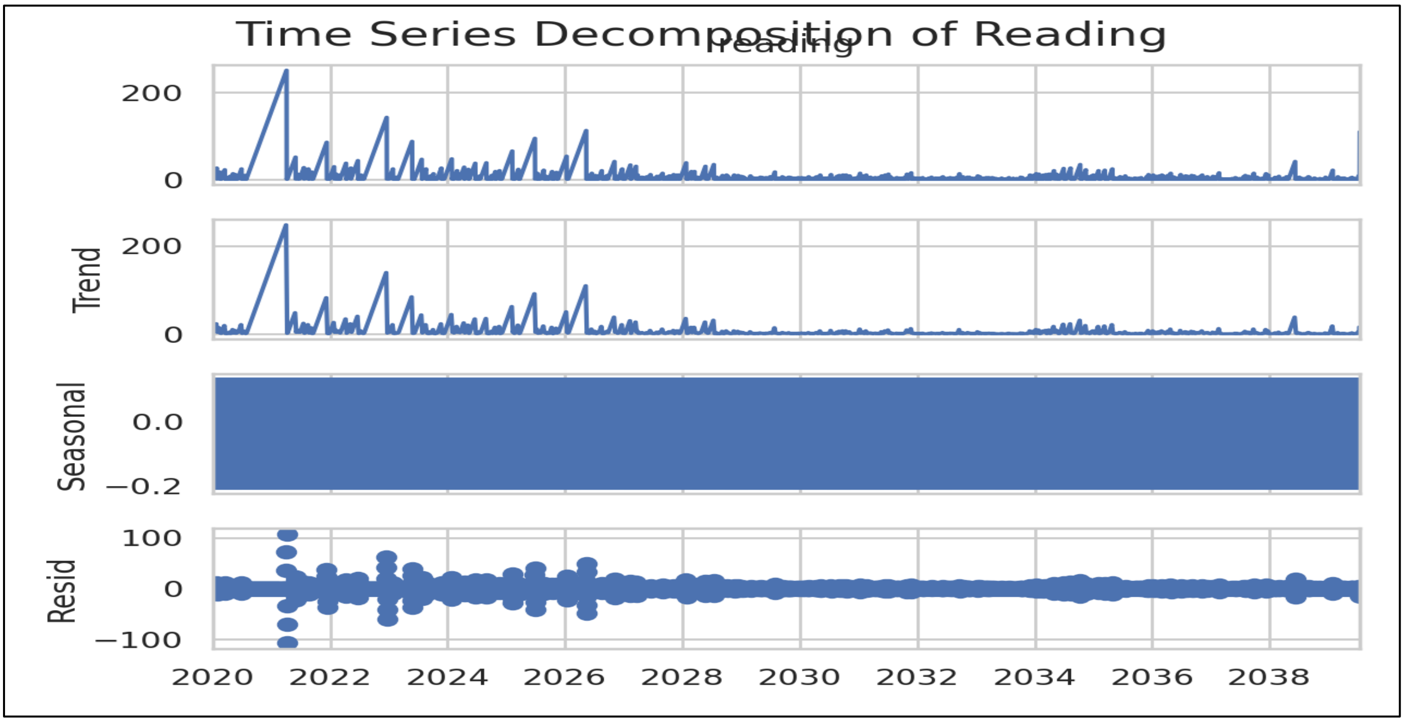


Fig 5 (b) Shows the Visualization of Time Series Decomposition for Multivariate Analysis

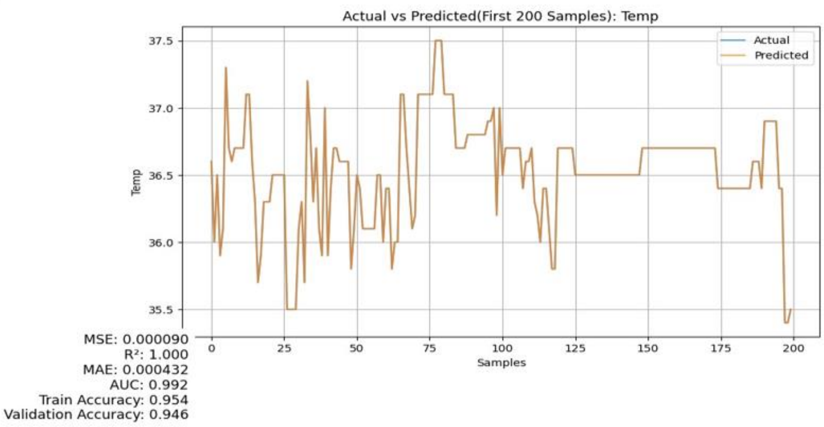
B. Predictions Model with ANN

The findings, which are displayed in Table 2 (a) shows Predictions using ANN with confusion matrix, comprise classification performance metrics reports, and Figure 6 (a) shows regression performance metrics reports, and cross-validation scores respectively.

Classification Report: Temp

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1535
1	1.00	1.00	1.00	957
accuracy			1.00	2492
macro avg	1.00	1.00	1.00	2492
weighted avg	1.00	1.00	1.00	2492

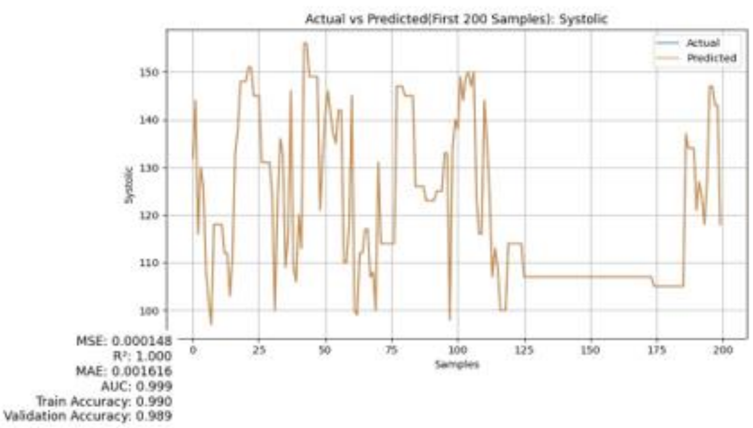
Temp Confusion Matrix:
[[1535 0]
[0 957]]



Classification Report: Systolic

	precision	recall	f1-score	support
0	1.00	0.67	0.80	39
1	0.99	1.00	1.00	2453
accuracy			0.99	2492
macro avg	1.00	0.83	0.90	2492
weighted avg	0.99	0.99	0.99	2492

Systolic Confusion Matrix:
[[26 13]
[0 2453]]



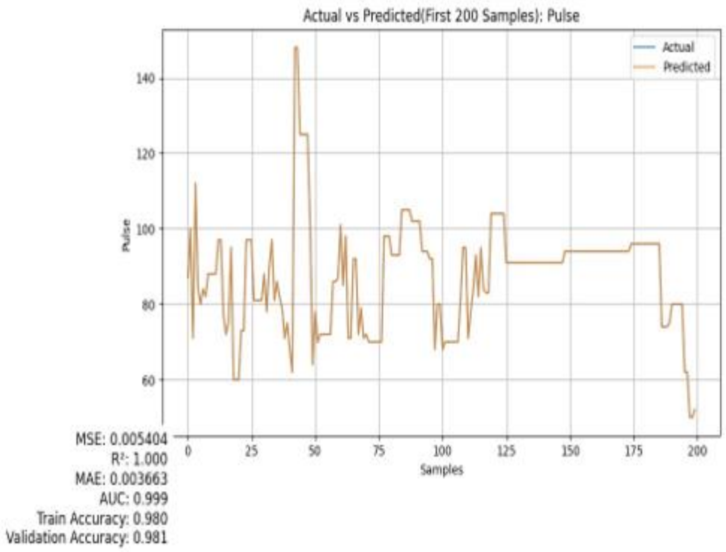
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

```
Classification Report: Pulse
              precision    recall  f1-score   support

           0       1.00      0.90      0.95       814
           1       0.95      1.00      0.98      1678

    accuracy                           0.97      2492
   macro avg       0.98      0.95      0.96      2492
weighted avg       0.97      0.97      0.97      2492

Pulse Confusion Matrix:
 [[ 730   84]
 [   0 1678]]
```

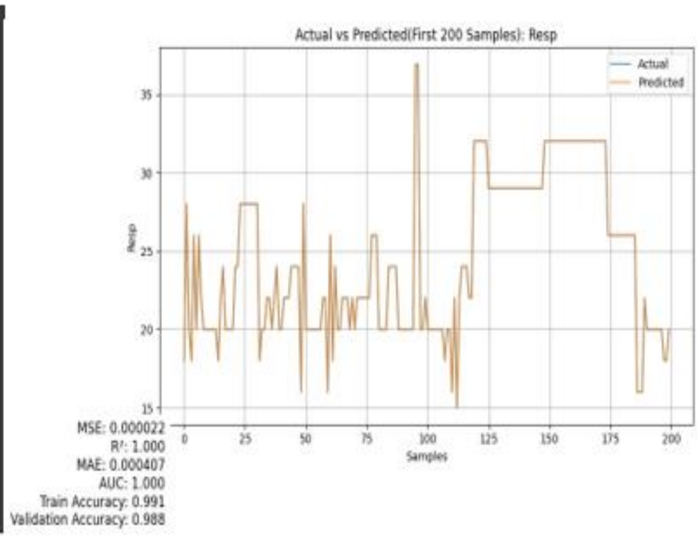



```
Classification Report: Resp
              precision    recall  f1-score   support

           0       1.00      0.16      0.28       140
           1       0.95      1.00      0.98      2352

    accuracy                           0.95      2492
   macro avg       0.98      0.58      0.63      2492
weighted avg       0.96      0.95      0.94      2492

Resp Confusion Matrix:
 [[  23  117]
 [   0 2352]]
```




```
Classification Report: SP02
              precision    recall  f1-score   support

           0       1.00      0.80      0.89         5
           1       1.00      1.00      1.00      2487

    accuracy                           1.00      2492
   macro avg       1.00      0.90      0.94      2492
weighted avg       1.00      1.00      1.00      2492

SP02 Confusion Matrix:
 [[   4    1]
 [   0 2487]]
```

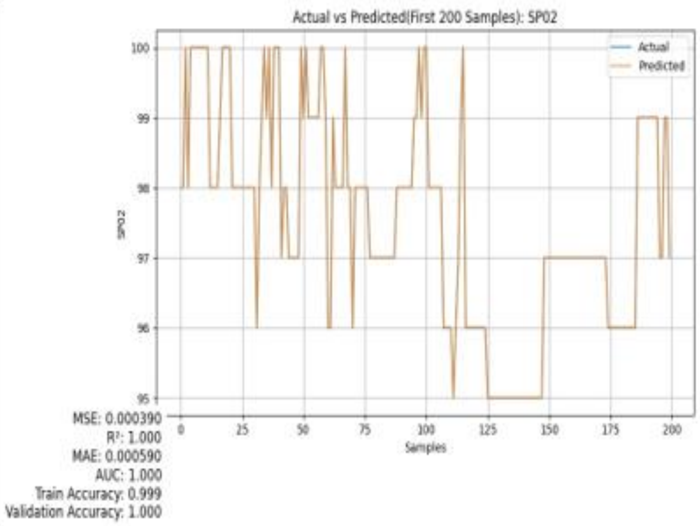


Fig 6 (a) Predictions using ANN and, Regression Performance Metrics Reports, and Cross-Validation Scores Table 2 (a), Confusion Matrix, Comprise Classification Performance Metrics Reports

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1438

V. DISCUSSION OF FINDINGS

These findings corroborate earlier studies by Gomez et al. (2024) and Ye et al. (2023), which discovered that high-quality data is optimal for machine learning models. The study's perfect classification results lend credence to the notion that data quality has a major impact on machine learning algorithms' capacity for prediction. On the other hand, earlier studies by Singh *et al*. (2023) showed that models' performance significantly declined when they were subjected to missing or noisy data. Even at the validation stage, the study's consistent performance highlights the dependability and suitability of ANN with the right data. The model proved that clean and well-preprocessed data greatly increases machine learning efficacy by predicting patient deterioration using vital signs with remarkable performance. These findings provide a starting point for potential future studies that mimic declining data quality to see how it affects prediction accuracy.

VI. CONCLUSION

The effect of data quality on machine learning models intended to forecast patient decline was examined in this study. The model achieved average perfect predictive accuracy performance using ANN and a dataset that included vital signs like temperature, systolic blood pressure, heart rate, respiratory rate, and SpO_2. All classification metrics scored 0.992, and the AUC was 0.998. These results confirm that a fundamental prerequisite for creating trustworthy and therapeutically useful machine learning-based early warning systems is data quality. The findings also imply that when medical facilities give priority on gathering and preparing high-quality data. ML has the potential to be a very effective tool for enhancing patient outcomes. It offered empirical support for the claim that high model accuracy in predicting patient deterioration is correlated with good data quality. ML can be an effective technique for improving patient outcomes. It presented empirical evidence that good data quality correlates with high model performance in patient deterioration prediction and established a performance benchmark to guide future research on data deterioration and resilience testing.

REFERENCES

[1]. Suh, J., Lee, G., Kim, J. W., Shin, J., Kim, Y.-J., Lee, S.-W., & Kim, S. (2024). Privacy-preserving prediction of postoperative mortality in multi-institutional data: Development and usability study. *JMIR Medical Informatics*, *12*, e56893.

[2]. Park, M., Bui, L. K., Jeong, M., Choi, E. J., Lee, N., Kwak, M., Kim, J., Jung, J., Giap, T.-T.-T., Guk, H., & Na, J. (2021). ICT-based person-centered community care platform (IPC3P) to enhance shared decision-making for integrated health and social care services. *International Journal of Medical Informatics*, *156*, 104590.

[3]. Lai, H.-J., Tan, T.-H., Lin, C.-S., Chen, Y.-F., & Lin, H.-H. (2020). Designing a clinical decision support system to predict readmissions for patients admitted with all-cause conditions. *Journal of Ambient Intelligence and Humanized Computing*, *11*(3), 1251–1264.

[4]. World Health Organization (2021): guide for action. Geneva: Rehabilitation in health system; Licence: CC BY-NC-SA 3.0 IGO.

[5]. Abdelgawad, Y. H., Abd El Razik, M. S., Saleh, D. A., Abuelela, M. H., & Salem, M. R. (2024). Promoting health information system in guiding decisions for improving performance: An intervention study at the Research Institute of Ophthalmology, Giza, Egypt. *Frontiers in Digital Health*, *6*, Article 1288776.

[6]. Alruwaili, M. M., & Abuadas, F. H. (2024). Exploring nurses' awareness and attitudes toward artificial intelligence: Implications for nursing practice. *Digital Health*. Advance online publication. 23(1):185. doi: 10.1186/s12411-023-02259-0

[7]. Yu, Z., Ashrafi, N., Li, H., Alaei, K., & Pishgar, M. (2024). Prediction of 30-day mortality for ICU patients with Sepsis-3. *medRxiv*. 31(11), pp. 2679-2688.

[8]. Ruissen, M. M., Torres-Peña, J. D., Uitbeijerse, B. S., Arenas de Larriva, A. P., Huisman, S. D., Namli, T., Salzsieder, E., Vogt, L., Ploessnig, M., van der Putte, B., Merle, A., Serra, G., Rodríguez, G., de Graaf, A. A., de Koning, E. J. P., Delgado-Lista, J., Sont, J. K., & POWER2DM Consortium. (2023). Clinical impact of an integrated e-health system for diabetes self-management support and shared decision making (POWER2DM): A randomised controlled trial. *Diabetologia*, *66*, 2213–2225.

[9]. Ye-wen Chen, Yu-jie Li, Peng Deng, Zhi-yong Yang, Kun-hua Zhong, Li-ge Zhang, Yang Chen, Hong-yu Zhi, Xiao-yan Hu, Jian-teng Gu, Jiao-lin Ning, Kai-zhi Lu, Ju Zhang, Zheng-yuan Xia, Xiao-lin Qin, Bin Yi(2022) Learning to Predict In-Hospital Mortality Risk in the Intensive Care Unit with Attention-Based Temporal Convolution Network *BMC Anesthesiol*. 22: 119. doi: 10.1186/s12871-022-01625-5

[10]. Amsterdam UMC. (2023). A data pipeline for cloud-based ICU patient data processing and prediction. University of Amsterdam Research Priority Agenda Program AI for Health Decision-Making. doi: 10.2359/fcvm.2023.1359699

[11]. Rodríguez, M. O., Martinez, R. M., Martínez, A. P., et al. (2024). Closing the value-based circle in shared decision-making: A digital framework for informing the shared decision-making process through patient reported outcome and experience measures. *Frontiers in Public Health*. Advance online publication.

[12]. Karalliedde, J., French, O., Burnhill, G., et al. (2023). A pragmatic digital health informatics-based approach for aiding clinical prioritisation and reducing backlog of care: A study in cohort of 4022 people with diabetes. *Diabetes Research and Clinical Practice*, *203*, 110834.

[13]. Field, M., Thwaites, D. I., Carolan, M., Delaney, G. P., Lehmann, J., Sykes, J., Vinod, S., & Holloway, L. (2022). Infrastructure platform for privacy-preserving

distributed machine learning development of computer-assisted theragnostics in cancer. *Journal of Biomedical Informatics, 134*, 104181.

[14]. Deist, T. M., Dankers, F. J. W. M., Ojha, P., Price, G., Lambin, P., Dekker, A., (2020). Distributed learning on 20,000+ lung cancer patients – The Personal Health Train. *Radiotherapy and Oncology, 144*, 189–200.

[15]. Feng, T., Noren, D. P., Kulkarni, C., Mariani, S., Zhao, C., Ghosh, E., Swearingen, D., Frassica, J., McFarlane, D., & Conroy, B. (2023). Machine learning-based clinical decision support for infection risk prediction. *Frontiers in Medicine, doi:10*1213411.

[16]. Kramer, O., Even, A., Matot, I., Steinberg, Y., & Bitan, Y. (2021). The impact of data quality defects on clinical decision-making in the intensive care unit. *Computer Methods and Programs in Biomedicine*, 209, doi: 10.2136/67253

[17]. Gan, S., Kim, C., Chang, J., Lee, D. Y., & Park, R. W. (2024). Enhancing readmission prediction models by integrating insights from home healthcare notes: Retrospective cohort study. *International Journal of Nursing Studies*, 158, 104850.

[18]. Kooiman, B. E. A. M., Robberegt, S. J., Albers, C. J., Bockting, C. L. H., Stikkelbroek, Y. A. J., & Nauta, M. H. (2024). Congruency of multimodal data-driven personalization with shared decision-making for StayFine: Individualized app-based relapse prevention for anxiety and depression in young people. *Journal Name*. Advance online publication. doi: 10.1246/s40001-055-02422-5

[19]. Sassi, Z., Eickmann, S., Roller, R., Osmanodja, B., Burchardt, A., Samhammer, D., Dabrock, P., Möller, S., Budde, K., & Herrmann, A. (2024). Prospectively investigating the impact of AI on shared decision-making in post kidney transplant care (PRIMA-AI): Protocol for a longitudinal qualitative study among patients, their support persons and treating physicians at a tertiary care centre. *BMJ Open, 14*(10), e073422.

[20]. Wurhofer, D., Neunteufel, J., Strumegger, E.-M., Höppchen, I., Mayr, B., Egger, A., Sareban, M., Reich, B., Neudorfer, M., Niebauer, J., Smeddinck, J. D., & Kulnik, S. T. (2024). Investigating shared decision-making during the use of a digital health tool for physical activity planning in cardiac rehabilitation. *Frontiers in Digital Health, 5,* Article 1324488.

[21]. Almagharbeh, W. T. (2024). The impact of AI-based decision support systems on nursing workflows in critical care units. *International Nursing Review*. Advance online publication, doi: 10.2196/66293

[22]. Vázquez-Calatayud, M., García-García, R., Regaira-Martínez, E., & Gómez-Urquiza, J. (2024). Real-world and game-based learning to enhance decision-making. *Nurse Education Today, 140,* 106276.

[23]. Dhingra, L., Schiller, R., Teets, R., Shuman, S., DiFiglia, S., & Portenoy, R. (2021). Pain management in primary care: A randomized controlled trial of a computerized decision support tool. *The American Journal of Medicine, 134*(12), 1546–1554.

[24]. Steele Gray, C., Nie, J., Loganathan, M., & Tahsin, F. (2022). Supporting goal-oriented care using mobile technology: Findings from the mixed-methods trial of the ePRO tool. *Annals of Family Medicine, 20*(Supplement 1), 2763.

[25]. Al-Khatib, H., Zahran, H., & Mahmud, F. (2024). Telemedicine adoption during COVID-19 in rural areas: Barriers and facilitators. *Telemedicine and e-Health*, 30(1), 15-25.

[26]. Marcin, T., Hautz, S. C., Singh, H., Zwaan, L., Schwappach, D., Krummrey, G., Schauber, S. K., Nendaz, M., Exadaktylos, A. K., Müller, M., Lambrigger, C., Sauter, T. C., Lindner, G., Bosbach, S., Griesshammer, I., & Hautz, W. E. (2024). Effects of a computerised diagnostic decision support tool on diagnostic quality in emergency departments: Study protocol of the DDx-BRO multicentre cluster randomised cross-over trial. *BMJ Open*. 32(1), pp.2152-2164.

[27]. Shoesmith, W. D., Abdullah, A. C., Tan, B. Y., Kamu, A., Ho, C. M., Giridharan, B., Forman, D., & Fyfe, S. (2022). Development of a scale to measure shared problem-solving and decision-making in mental healthcare. *Patient Education and Counseling, 105*(7), 2480–2488.

[28]. Deo, R. C., Smith, R., & Patel, R. (2024). Software-driven chronic disease management: Algorithm design and implementation in a community-based blood pressure control pilot. *SAGE Open Medicine, 12*, 20503121241284025

[29]. Kim, H., & Park, S. (2023). Virtual reality for pain management in pediatric patients: A systematic review. *Pediatric Pain Management*, 18(2), 45-56

[30]. Koch, G., Wilbaux, M., Kasser, S., Schumacher, K., Steffens, B., Wellmann, S., & Pfister, M. (2022). Leveraging predictive pharmacometrics-based algorithms to enhance perinatal care Application to neonatal jaundice. *Frontiers in Pharmacology*, 14(5), pp. 1053-1065.

[31]. Singh, R., Patel, A., & Kumar, S. (2023). Machine learning for early detection of diabetic retinopathy using retinal images. *Journal of Medical Imaging*, 10(3), 230-239.