ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 5| May 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

EARLY PREDICTION OF SEPSIS USING MACHINE LEARNING ALGORITHM: A BRIEF CLINICAL PERSPECTIVE

Dr. A. A. Bardekar[1], Rajshree Datey[2], Rohini Mankar[3], Saukhya Jainwar[4], Sampada Gole[5], Isha Upadhye[6]

[1]*Computer Science and Engineering, Sipna College of Engineering and Technology*

[2]*Computer Science and Engineering, Sipna College of Engineering and Technology*

[3]*Computer Science and Engineering, Sipna College of Engineering and Technology*

[4]*Computer Science and Engineering, Sipna College of Engineering and Technology*

[5]*Computer Science and Engineering, Sipna College of Engineering and Technology*

[6]*Computer Science and Engineering, Sipna College of Engineering and Technology*

Article DOI: https://doi.org/10.36713/epra10142
DOI No: 10.36713/epra10142

ABSTRACT

Sepsis is a worldwide cause of death owing to infection and associated immune system response. In situations of septic shock, mortality rates are highest in both developed and underdeveloped countries. Sepsis is a medical illness that requires immediate medical attention but can be avoided with advance warning. Sepsis affects an estimated 30 million people worldwide, with more than 6 million people dying each year. Among them More than 4.2 million new born and children are at risk of contracting the disease. To treat Sepsis, hospitals spend $24 billion (about 13% of all healthcare spending in the United States). The importance of early identification of sepsis in improving sepsis outcomes cannot be overstated. Each hour of delay in treatment increases the risk of death by 4 to 8%. As a result, developing a good model that may be able to tackle this problem becomes urgent and critical. We want to analyse, assess, and design an algorithm that will help us address some of the problems we've had when attempting the project.

KEYWORDS: ***Sepsis, Machine learning, early prediction, Supervised learning, KNN.***

I. INTRODUCTION

Sepsis is a leading cause of death worldwide. It is described as a life-threatening organ failure induced by a dysregulated host response to infection. Septic shock is the leading cause of death in both developing and industrialized countries.Sepsis is the leading cause of death in hospitals, with survivors facing an increased risk of physical organ failure, neurological impairment, and chronic illness. Early and customized therapy has been demonstrated to improve sepsis outcomes significantly. Predictive models were utilized to improve treatment and could be used in critical care settings like the Intensive Care Unit (ICU) to detect septic patients early [1]. Machine learning has emerged as a viable technique to reduce diagnostic uncertainty, identify patients with severe sepsis, and select appropriate medications. The current state of reliable and early detection of sepsis is highly challenging, which might cause therapeutic delays. Despite this, there has been a lot of interest in the topic, which has been highlighted in recent articles. According to the research, persons who get Sepsis and receive delayed antibiotic treatment had a higher adjusted death rate. The effect is more noticeable in patients suffering from septic shock, a more advanced form of sepsis. In this situation, hourly treatment delays result in an increase in mortality of 3.6 to 9.9% each hour. Clinical criteria for diagnosing and treating sepsis have been proposed by some professional care societies.

However, the critical necessity for early and accurate Sepsis detection has yet to be attained. The use of a computational technique has the potential to improve the early diagnosis of Sepsis. It entails integrating machine learning algorithms to clinical data in order to provide time-specific predictions up to a day before Sepsis is clinically determined. Sepsis is an emergency-oriented clinical illness that can be averted with early notice. It's a SIRS, a normal autoimmune response to infections, and most people don't need medication or antibiotics. Because other infections without acute

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 5| May 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

hypothermia, digestive, inflammatory, and lung problems, and long-term immunological reactions (e.g. infections of the urinary system) make it difficult to diagnose early, the response to postoperative sepsis diagnosing inflammation is unknown. Early differentiation of inflammation based on sepsis is required, and then successful treatment modalities must be integrated into the critical care plan [2].

Artificial intelligence technology has emerged as an effective method in medical assistance, including early sepsis diagnosis, through the integration of electronic medical records, medical data, and other data. These methods have been developed to model and predict the health of the human body, as well as obtain accurate prescription information to assist clinicians in making quick and effective decisions. Clinical characteristics were employed in real time to construct a significant predictive model that can accurately anticipate the onset of sepsis in an intensive care unit (ICU) before clinical understanding.

II. LITERATURE REVIEW

In several medical domains, a diagnosis system based on artificial intelligence (AI) has proven to be useful. Machine learning methods utilized in the diagnosis, prognosis, and therapy of sepsis include supervised learning and reinforcement learning. By reviewing breast tissue imaging, Beck et al. [3] construct the C-Path (Computational Pathologist) method to automatically diagnose breast cancer and forecast whether patients would survive or not.

The use of many physiological signs and modelling efficient machine learning algorithms for the diagnosis, prediction, and therapy of sepsis are the two primary problems in the current research. Similarly, in order to predict sepsis in advance, it's critical to select the appropriate variables and create useful algorithms in the clinic.In this perspective, we provide a brief, clinician-oriented vision on the following relevant aspects concerning the use of machine learning predictive models for the early detection of sepsis in daily practise: (i) the controversy over sepsis definition and its impact on the development of prediction models; (ii) the choice and availability of input features; and (iii) the measure of model performance, the output, and their utility in medical practice.

Hemodynamic sepsis administration in the emergency department relies primarily on endotracheal intubation and vasopressor care to maintain a sufficient heart rate and total organ blood flow, according to Prasad et al. [4]. Although frequent pastimes are associated with low blood pressure (65 mmHg for venous blood pressure and 90 mmHg for systolic blood pressure [SBP]), little attention is paid to the conceptual nature of heart rate. He employed unsupervised re-examination techniques within two hours of the onset of hypertensive moments (SBP: 90 mmHg) or immediately before to the onset of vasopressor medication. The data revealed that certain hypotensive patients appeared to be affected fast (within 40 minutes). Patients who had hypotension in the previous hour as a result of a significant and severe decline in daily SBP had a higher prevalence of subsequent vasopressor administration than those who had a more progressive decrease in hypertension.

III. OBJECTIVE

The purpose of this study was to create an algorithm for predicting Sepsis in the early stages utilizing commonly available clinical data. Early prediction, in particular, can save lives, whereas late or missed predictions can be fatal, and misleading algorithms can waste hospital resources and destroy trust in the algorithm itself.

The algorithm will be created at several levels and will use clinical data, automatically recognizing a patient's risk of developing Sepsis and making a positive or negative forecast of at least six hours and no more than twelve hours before Sepsis start. We utilize KNN, which is based on Supervised Learning, since it is most beneficial when labelled data is too expensive or difficult to gather, and it can achieve high accuracy in a wide range of prediction situations.

IV. PROPOSED SYSTEM

This Project has following Modules:

1. Admin panel

- Login
- Register hospital sub-admins
- Register doctors
- View doctors
- View hospital admins
- View patients
- View sepsis predicted patients details

2. Hospital Sub admin

- Login
- Admit patient
- View admitted patients
- Add admitted patient clinical data
- View patients data
- View sepsis prediction report

3. Doctors

- Login
- View sepsis patients
- View patient data

4. Patients

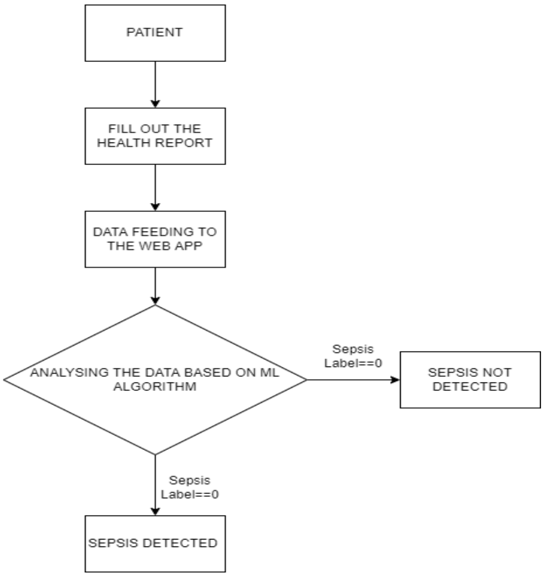
- Login
- View own/patient data

- View sepsis report

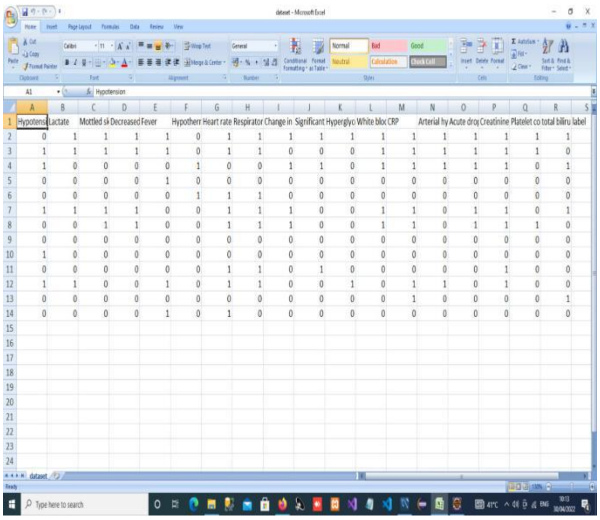
V. METHODOLOGY

Dataset consist of clinical data of patients. For predicting the output label based on the qualities, the K-Nearest Neighbors (KNN) method is utilized. Machine learning algorithms build a mathematical model based on "training data" so that forecasts or choices can be made without the need for specialized task programming.

A. Model Diagram



B. Dataset



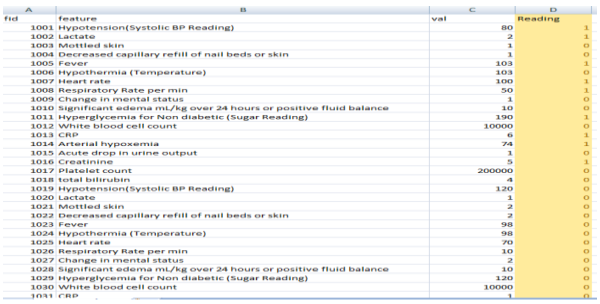
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Hypotens	Lactate	Mottled sk	Decreased	Fever	Hypotherr	Heart rate	Respirator	Change in	Significant	Hyperglyo	White bloc	CRP	Arterial hy	Acute dro	Creatinine	Platelet co	total biliru	label
2	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	
3	1	1	1	1	1	0	1	1	0	0	0	1	1	1	1	1	1	0	
4	1	0	0	0	0	1	0	0	1	1	0	1	1	1	1	1	0	1	
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	
7	1	1	1	1	0	0	1	1	1	0	0	1	1	0	1	1	0	1	
8	0	0	1	1	0	0	1	1	1	0	0	1	1	0	1	1	1	0	
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	
12	1	1	0	0	1	0	1	1	0	0	1	0	1	1	0	1	0	0	
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	
14	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	

The description for true value of sepsis label is given below:

1. Hypotension (systolic blood pressure) is a condition in which the blood pressure is less than 90 mm Hg.
2. Lactate concentrations more than 1 mmol/L
3. Skin that is mottled
4. Nail bed or skin capillary refill is reduced.
5. Fever of more than 38.3 degrees Celsius (101 degrees Fahrenheit)
6. Hypothermia is defined as a core temperature of 36 degrees Celsius (96.8 degrees Fahrenheit).
7. More than 90 beats per minute
8. Respiratory rate per minute greater than 20
9. Alteration in mental state
10. Positive fluid balance (>20 mL/kg over 24 hours) or significant edema
11. Hyperglycemia (>140 mg/dL) in a non-diabetic person
12. A white blood cell count of more than 12,000
13. C-reactive protein (CRP) levels in the blood are higher than 3 (according to your lab's cutoffs).
14. Hypoxemia of the arteries is less than 75
15. Urine output drops abruptly
16. Creatinine level rises by more than 0.5 mg/dL
17. a platelet count of less than 100,000
18. Bilirubin levels are high (total bilirubin > 4 mg/dL).

Data Collection

In this project we have collected data from the test readings submitted by hospital sub admin. The readings are in the raw format so we have preprocess it to convert into numeric format. We have stored the processed data in database as given below:



fid	feature	val	Reading
1001	Hypotension(Systolic BP Reading)	80	1
1002	Lactate	2	1
1003	Mottled skin	1	0
1004	Decreased capillary refill of nail beds or skin	1	0
1005	Fever	103	1
1006	Hypothermia (Temperature)	103	0
1007	Heart rate	100	1
1008	Respiratory Rate per min	50	1
1009	Change in mental status	1	0
1010	Significant edema mL/kg over 24 hours or positive fluid balance	10	0
1011	Hyperglycemia for Non diabetic (Sugar Reading)	190	1
1012	White blood cell count	10000	0
1013	CRP	6	1
1014	Arterial hypoxemia	74	1
1015	Acute drop in urine output	1	0
1016	Creatinine	5	1
1017	Platelet count	200000	0
1018	total bilirubin	4	0
1019	Hypotension(Systolic BP Reading)	120	0
1020	Lactate	1	0
1021	Mottled skin	2	0
1022	Decreased capillary refill of nail beds or skin	2	0
1023	Fever	98	0
1024	Hypothermia (Temperature)	98	0
1025	Heart rate	70	0
1026	Respiratory Rate per min	10	0
1027	Change in mental status	2	0
1028	Significant edema mL/kg over 24 hours or positive fluid balance	10	0
1029	Hyperglycemia for Non diabetic (Sugar Reading)	120	0
1030	White blood cell count	10000	0
1031	CRP	1	0

Data processing

It is the process of transforming raw data into understandable format. Data processing is a critical step in preparing data for machine learning. For the most frequent types of ML, a substantial amount of data is usually required. The data format must be correct in order to get better results from the model used in Machine Learning projects [5]. Machine Learning is heavily reliant on test outcomes. Data preprocessing is a crucial step in the data mining process.

As a result, the interpretation and correctness of the results are primarily determined prior to the execution of an experiment. Preprocessing data is always the most important phase in any deep learning activity, especially in computational modelling. It is more difficult to discover knowledge during the

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 5| May 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

training process if there is a lot of unneeded and redundant information, or noisy and erroneous data. The stages of data preparation and retrieval will take a long time to process.

```
# Load libraries
import pandas as pd
from sklearn.neighbors import KNeighborsClassifier
from sklearn.tree import DecisionTreeClassifier # Import Decision Tree Classifier
from sklearn.model_selection import train_test_split # Import train_test_split function
from sklearn import metrics #Import scikit-learn metrics module for accuracy calculation

def knnalgo(lst):
col_names = ['Hypotension','Lactate','Mottled skin','Decreased capillary refill of nail beds or skin','Fever','Hypothermia','Heart rate','Respiratory Rate per min','Change in mental status','Significant edema','Hyperglycemia','White blood cell count','CRP','Arterial hypoxemia','Acute drop in urine output','Creatinine','Platelet count','total bilirubin','label']
# load dataset
pima = pd.read_csv("dataset.csv", header=0, names=col_names)
#split dataset in features and target variable
feature_cols = ['Hypotension','Lactate','Mottled skin','Decreased capillary refill of nail beds or skin','Fever','Hypothermia','Heart rate','Respiratory Rate per min','Change in mental status','Significant edema','Hyperglycemia','White blood cell count','CRP','Arterial hypoxemia','Acute drop in urine output','Creatinine','Platelet count','total bilirubin']
X = pima[feature_cols] # Features
y = pima.label # Target variable
# Split dataset into training set and test set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=1) # 70% training and 30% test
# Create Decision Tree classifer object
clf = KNeighborsClassifier()

# Train Decision Tree Classifer
clf = clf.fit(X_train,y_train)
#Z_test=[[0,0,0,1,1,0,1,1,1,0,0,0,0,0,0,0,1,1]]
Z_test= [lst]
#Predict the response for test dataset
print(Z_test)

y_pred = clf.predict(Z_test)
print(y_pred)
return y_pred
# Model Accuracy, how often is the classifier correct?
#print("Accuracy:",metrics.accuracy_score(y_test, y_pred))
```

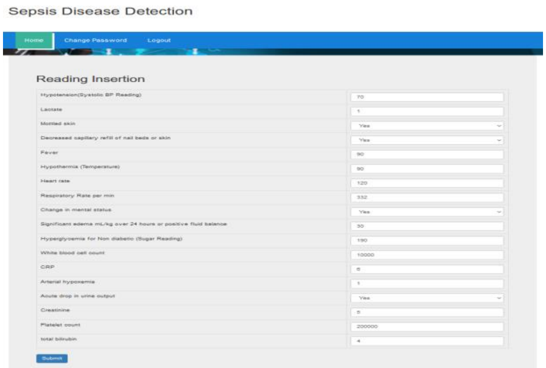
VI. CONCLUSION AND FUTURE SCOPE

We solely evaluated essential health measurements that are required to forecast sepsis disease in our experiment. We processed raw data during insertion in java just after getting raw readings from the hospital, reducing the time necessary to process raw data. The knn machine learning algorithm is given the processed data. To implement the knn algorithm, we used the sklearn python library. As a result, the execution time is lowered and the accuracy is increased. In order to establish a high level of confidence in the existing results, we need to conduct more analysis utilising techniques such as Random Forest, Principal Component Analysis, and so on. Next, we must employ the time component approach; this will necessitate the assistance of a domain specialist. We can consider employing SMOTE to deal with Imbalance. We will continue to develop on various machine learning techniques in the future.

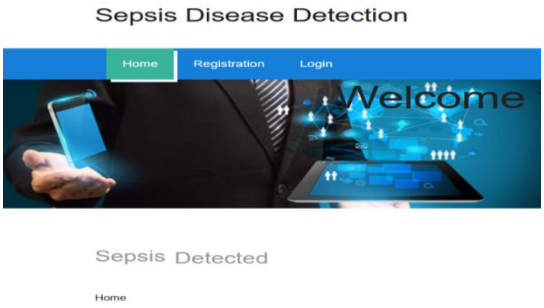
VII. RESULT

On the basis of clinical dataset, the corresponding greater than or less than values are computed to set the label to 1 otherwise 0.

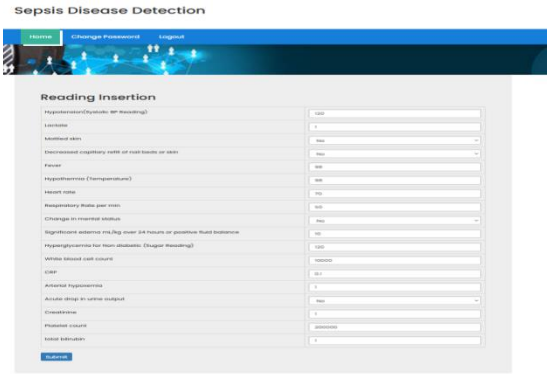
Input:



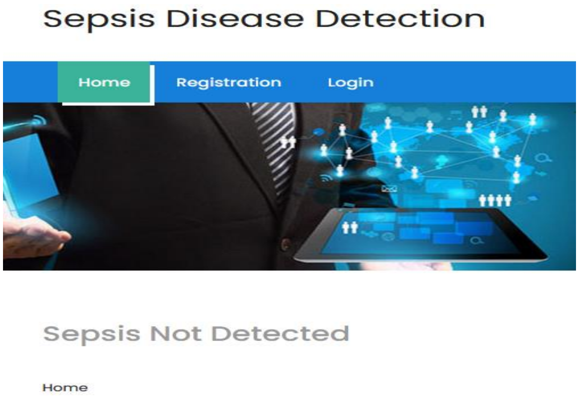
Output



Input



Output



VIII. REFERENCES

1. *R. M. Demirer and O. Demirer, "Early Prediction of Sepsis from Clinical Data Using Artificial Intelligence," 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 2019, pp. 1-4.*
2. *K. Gunnarsdottir, V. Sadashivaiah, M. Kerr, S. Santaniello and S. V. Sarma, "Using demographic and time series physiological features to classify sepsis in the intensive care unit," 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, 2016, pp. 778-782.*
3. *A. H. Beck, A. R. Sangoi, S. Leung et al., "Systematic analysis of breast cancer morphology uncovers stromal features associated with survival," Science Translational Medicine, vol. 3, no. 108, 108ra113 pages, 2011.*

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 5| May 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

4. *V. Prasad, J. C. Lynch, M. R. Filbin, A. T. Reisner and T. Heldt, "Clustering Blood Pressure Trajectories in Septic Shock in the Emergency Department," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019, pp. 494-497*
5. *J. Thakur, S. K. Pahuja and R. Pahuja, "Neonatal Sepsis Prediction Model for Resource-Poor Developing Countries," 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), Kolkata, 2018, pp. 1-5.*