IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

GANGRENE PREDICTION USING MACHINE LEARNING

[1]Sanskar Randive

[1]Student, Dept. of Information Technology, G.H.Raisoni College of Engineering, Nagpur, India.

Abstract: Gangrene is a severe medical condition characterized by the death of body tissues due to a lack of blood supply or infection. Timely and accurate prediction of gangrene can significantly impact patient outcomes and guide appropriate medical interventions.

This paper provides an overview and analysis of machine learning-based prediction models for gangrene. We begin by discussing the clinical significance of gangrene, its causes, and the challenges associated with its prediction. We emphasize the need for early detection and intervention to prevent further tissue damage and potential complications.Various machine learning algorithms, including logistic regression, K-Nearest Neighbors, support vector machine, naive bayes, decision trees and random forests, are examined for their application in gangrene prediction. We explore the features and data sources commonly used in these models, such as patient Diabetes, blood pressure, heart rate, respiratory rate, body temperature, skin condition and line of demarcation. Additionally, we discuss the importance of feature selection and extraction techniques to enhance the predictive performance of the models.

Data preprocessing steps, including missing data imputation, outlier detection, and normalization, are addressed to ensure data quality and model robustness. We delve into the evaluation metrics employed to assess the performance of gangrene prediction models, such as accuracy.

Keywords: Logistic regression, K-Nearest Neighbors, SVM, Naive Bayes, Decision trees, Random forests..

I. INTRODUCTION

Gangrene is a severe medical condition characterized by the death and decay of body tissues, typically caused by a lack of blood supply or infection. It is a critical condition that requires immediate medical attention to prevent further complications and potentially save lives. Gangrene can affect various parts of the body, including the limbs, internal organs, and skin.

The development of gangrene can occur due to several factors, such as traumatic injuries, diabetes, peripheral artery disease, frostbite, or severe infections. It is classified into different types based on its underlying causes, including dry gangrene, wet gangrene, gas gangrene, and internal gangrene. Regardless of the type, gangrene poses a significant threat to the affected individual's health and well-being.

Machine learning, a subfield of artificial intelligence, provides computational methods that can analyze large amounts of data, identify patterns, and make predictions. By leveraging diverse data sources, including patient characteristics, medical histories and laboratory results, machine learning algorithms can learn from historical cases and identify key risk factors associated with gangrene development.

II. Proposed Methodology

A. Collection of data

Collecting data is the foremost and one of the most important steps in this module. For the perfect prediction of the collection the dataset should be proper.The dataset taken into account consists of parameters like diabetes, blood pressure, heart rate, respiratory rate, body temperature, skin condition and line of demarcation.

B. Cleaning and pre-processing

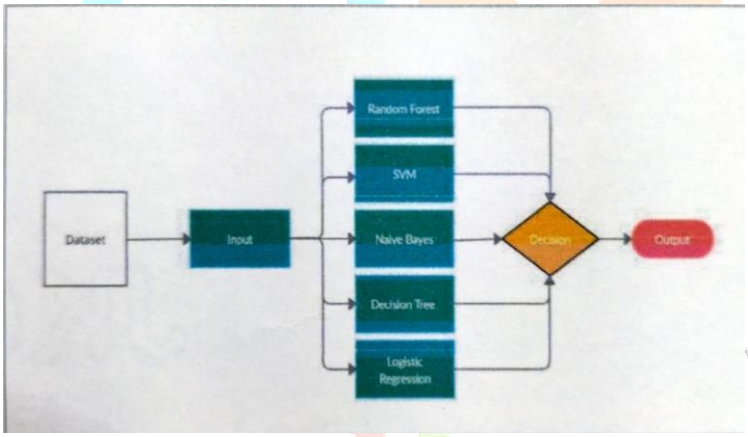
In this phase the obtained data will be carefully assessed for any missing values and it will be cleaned or filled. This is done to transform and improve the dataset before using it, to make it compatible to further processing

C. Feature Selection

There can be number of elements in the dataset of Gangrene however not all of these are required. For example, in our model the body temperature are dropped as it has no relation to Gangrene. This is done to reduce the dimensionality of the dataset to make it easier to evaluate.

E. Classification Methods

Various types of algorithm were considered for this module, to compare and find out the best algorithm that would provide the maximum accuracy in forecasting the result. The Algorithm used are:



1) KNN algorithm:

In order to forecast the label of an input data point based on the majority label of its K neighbors, KNN first locates the K data points that are the closest to an input data point. The Euclidean distance formula is used to determine the distance between data points.

2) Logistic Regression:

Logistic regression is a statistical method for examining the relationship between a dependent variable and one or more independent factors. It is commonly utilized when there is a binary dependent variable and the problem involves binary classification. Logistic regression determines the probability that the dependent variable will be in one of two possible states based on the values of the independent variables.

3) Naïve Bayes:

The name "naive" comes from the fact that Naive Bayes makes the assumption that the features or attributes of the input data are independent of one another, which is frequently true in reality. Despite this oversimplifying presumption, Naive Bayes can frequently produce classification results with high precision, especially when the dataset is sizable and the features are largely independent.

4) Decision tree:

Decision trees work by recursively splitting the dataset into subsets based on the most informative features, in order to create a tree-like structure that can be used to make
predictions on new data. The decision is based on the value of a certain feature at each internal node of the tree, and the algorithm divides the data until a stopping requirement is satisfied.

5) Random Forest algorithm:

Random forest algorithm is an ensemble learning technique that merges multiple decision trees and creates a forest of trees. The final forecast is based on the consensus of all the trees in the forest, and each tree is constructed using a random subset of features and data points.

F. Training and Testing:

Firstly, the data set and the object of the data set were preprocessed and label encoded by using the module preprocessing of sklearn library. After that, the dataset was divided in the proportion of 75:25 for the training and testing phase with the help of train_test_split function. It was done using the sklearn.model_selection

G. Performance Evaluation and Analysis:

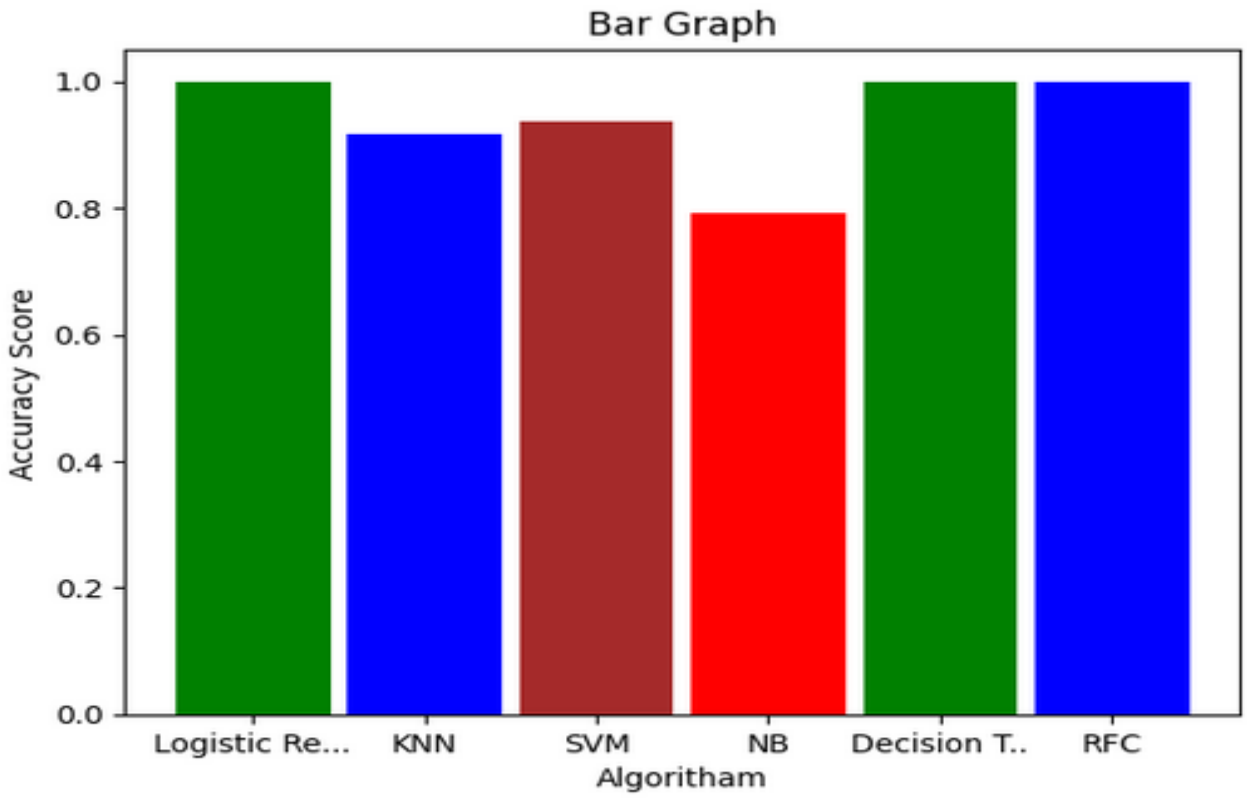


Fig. 1: Accuracy rate of algorithm

The K-Nearest Neighbors [KNN] algorithm, Decision tree, Logistic regression, Naïve Bayes, and Random Forest are the five distinct machine learning classification techniques that were used and later accuracy was calculated of every algorithm using a confusion matrix. After comparing it was found that decision tree , logistic regression and random forest has nearly about the same accuracy, however, decision tree and logistic regression was facing a problem in overfitting the training data thus, making it too complex and losing its ability to generalize new unseen data. Finally, random forest algorithm was chosen over decision tree to predict the result

III. Result

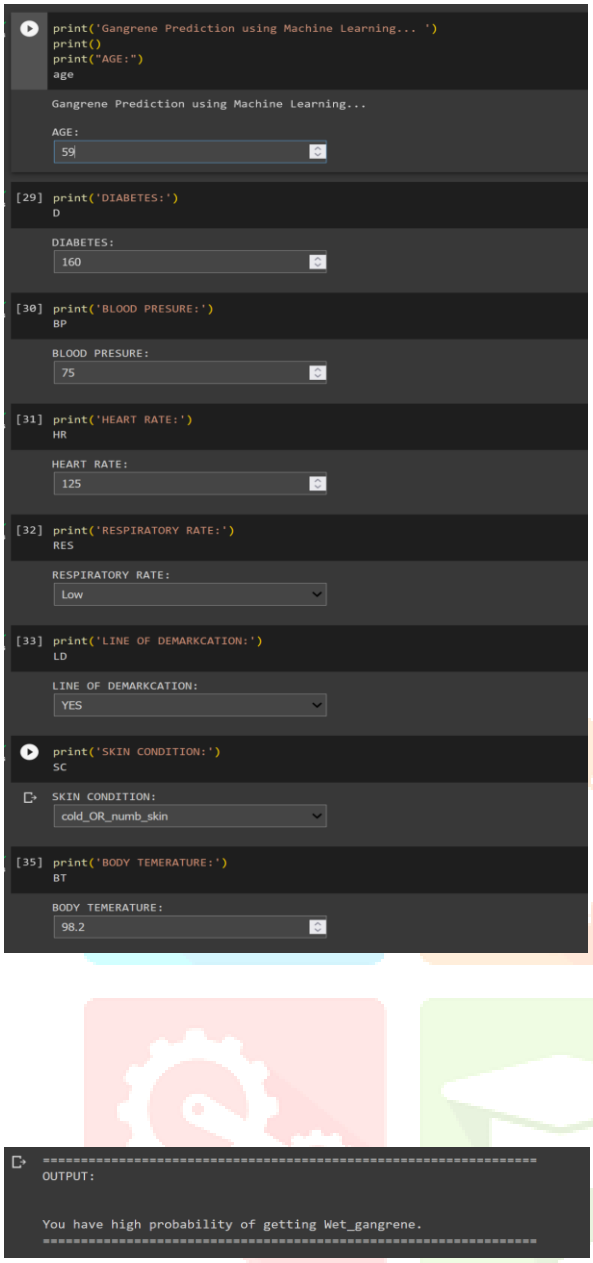
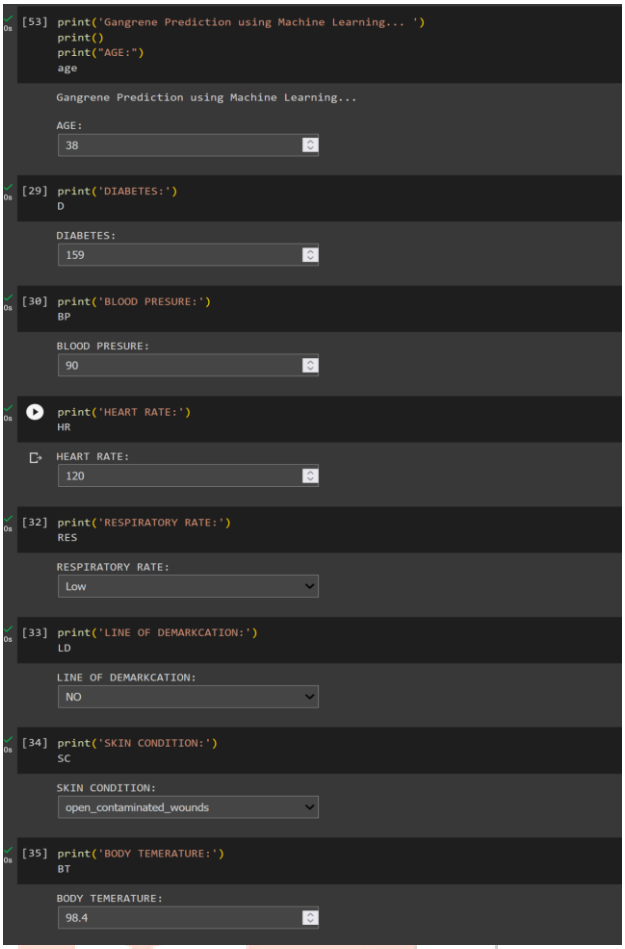


Fig 2: Prediction Output (1)



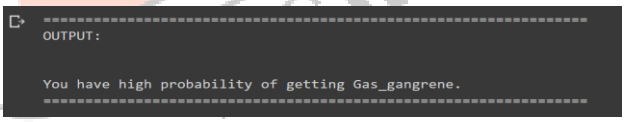


Fig 3: Prediction Output(2)

IV. Conclusion:

Predicting gangrene using machine learning can be a valuable tool in healthcare, aiding in early detection and timely intervention. By analyzing various factors and patterns in medical data, machine learning models can assist in identifying patients who are at a higher risk of developing gangrene. However, it is important to note that machine learning models should not be considered as a replacement for clinical expertise and medical diagnosis. They should be used as an additional tool to support healthcare professionals in making informed decisions.

V. Reference:

[1] Bhattacharya S, et al. "Machine Learning-based Automated Identification of Diabetic Foot Complications Using Natural Language Processing of Electronic Health Record Notes." Journal of Biomedical Informatics. 2019; 98: 103280.

[2] .Sharma A, et al. "Detection and Analysis of Diabetic Foot Ulcer Images Using Machine Learning Techniques." Journal of Medical Systems. 2019; 43(5): 132.

[3] Zhang Z, et al. "Detection of Diabetic Foot Ulcers by Using Ensemble Deep Learning Models." IEEE Access. 2020; 8: 46990-47002.