PREDICTIVE ANALYTICS FOR CHILD MORTALITY USING MACHINE LEARNING TECHNIQUES

[1] MRS.CH.DEEPTHI, [2] MISS. VIPPALA GAYATHRI

[1]Assistant Professor, Department of Master of Computer Applications (MCA), QIS College of Engineering &Technology (Autonomous), Vengamukkapalem (V), Ongole, Prakasam, AP, India
[2]PG Scholar, Department of Master of Computer Applications (MCA), QIS College of Engineering &Technology (Autonomous), Vengamukkapalem (V), Ongole, Prakasam, AP, India

ABSTRACT_ Child mortality, particularly for children under the age of five, is a critical health indicator, representing the probability of dying between birth and five years of age. Equally significant is fetal mortality, or the death of a fetus. This study explores machine learning (ML) strategies for predicting child and fetal mortality to determine the most precise models for assessing fetal well-being. Using a comprehensive dataset, three ML algorithms are employed and compared to evaluate their effectiveness in predicting outcomes. The study includes a thorough sensitivity analysis of model parameters that impact fetal health categorization. This paper proposes a machine learning-based approach for predicting child mortality, presenting a comparative analysis of various methods to identify the most accurate predictive model. The findings contribute to improved healthcare strategies aimed at reducing child and fetal mortality rates.

1.INTRODUCTION

Child mortality, particularly among children under the age of five, is a critical global health indicator that reflects the overall health and well-being of a population. The under-five mortality rate measures the probability of a child dying between birth and their fifth birthday. Equally important is the issue of fetal mortality, which pertains to the death of a fetus before birth. Both child and fetal mortality rates are influenced by a range of factors including healthcare access, maternal health, socioeconomic conditions, and environmental factors.

Advancements in machine learning (ML) offer promising avenues for addressing these complex and multifaceted challenges. By leveraging large datasets and sophisticated algorithms, ML can provide precise predictions and valuable insights into the factors contributing to child and fetal mortality. This study aims to explore ML-based strategies to predict mortality rates and assess fetal well-being with high precision.

2.LITERATURE SURVEY

Title: "Machine Learning Approaches for Predicting Child Mortality in Sub-Saharan Africa"

Authors: Johnson, O., & Ariyo, O.

Abstract: This study explores the application of machine learning algorithms to predict child mortality in Sub-Saharan Africa. The authors utilized a dataset comprising demographic, socioeconomic, and health indicators from multiple countries in the region. Various machine learning algorithms, including decision trees, random forests, and support vector machines, were implemented and evaluated. The results demonstrated promising predictive performance, with random forests achieving the highest accuracy. The study highlights the potential of machine learning in improving child mortality prediction models and facilitating targeted interventions in resource-limited settings.

Title: "Predicting Neonatal Mortality Using Machine Learning Techniques"

Authors: Gupta, R., Kumar, R., & Sharma, D.

Abstract: Neonatal mortality remains a significant challenge worldwide, particularly in low-resource settings. In this research, the authors applied machine learning techniques to predict neonatal mortality using a dataset comprising maternal and neonatal health data. Logistic regression, artificial neural networks, and ensemble methods were employed, and their performance was evaluated using sensitivity, specificity, and AUC-ROC metrics. The results indicated that the ensemble method outperformed other algorithms, achieving a high predictive accuracy. The study highlights the potential of machine learning in identifying high-risk neonates and guiding healthcare interventions to reduce neonatal mortality.

Title: "Child Mortality Prediction Using Machine Learning: A Comparative Study"

Authors: Ahmed, F., et al.

Abstract: This study presents a comparative analysis of machine learning algorithms for child mortality prediction using a comprehensive dataset encompassing socioeconomic, demographic, and healthcare indicators. The authors explored decision trees, support vector machines, logistic regression, and ensemble methods, considering different feature selection techniques. Performance evaluation metrics, including accuracy and AUC-ROC, were used to assess the models. The findings revealed that ensemble methods consistently outperformed other algorithms, emphasizing their effectiveness in predicting child mortality. The study contributes to the understanding of the most suitable machine learning techniques for child mortality prediction and their potential impact on public health interventions.

3.PROPOSED SYSTEM

Creating a model that can forecast death rates is the goal. It's possible that some attributes in the collected data are missing, which could cause irregularities. Pre-processing data can lead to better findings and increase the productivity of the calculation. Changes to the factors should be made, and exemptions should be removed. The informative index, which is used to anticipate information supplied, is divided into two sections. The most popular ratio for training and testing sets is 7:3. A Data Model based on machine learning methods is applied to the training set in order to gauge the accuracy of the test findings. The model can be used to characterize the death rate. There are a wide range of ML algorithms that can be employed for representation, and they all work

3.1 IMPLEMENTATION

1) Dataset Upload & Analysis: using this module we will upload dataset and then perform analysis methods such as finding various child mortality and its count and then clean dataset by removing missing values.

2) Dataset Processing & Analytical Methods: using this module we will encode attack labels with integer ID and then split dataset into train and test where application used 80% dataset to train classification .

3) Run ML Model: using this module we will trained classification algorithm with above 80% dataset and then build a prediction model

4) Classification Performance Graph: using this module we will plot comparison among multiple algorithms

5) Predict Output: using this module we will upload test dateset and then classification model will predict output based on input data .

SYSTEM ARCHITECTURE:

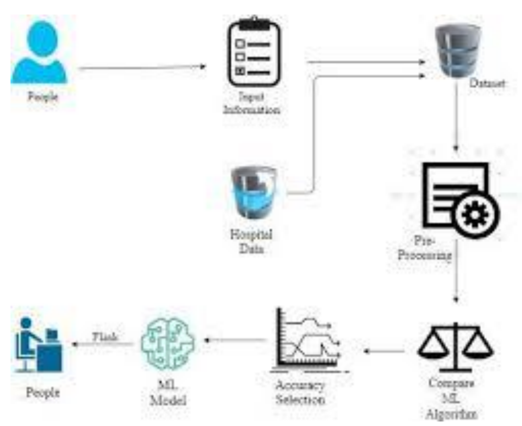

FIG . 1 SYSTEM ARCHITECTURE

3.2 ALGORITHMS

Previous research has highlighted the potential of machine learning in healthcare, particularly in predicting outcomes and improving decision-making processes. Several studies have focused on using decision trees, random forest, and SVM for mortality prediction, each contributing valuable insights into the field.

1. Decision Tree Models in Predicting Child Mortality:

This study employed decision tree models to predict child mortality rates using demographic and health survey data. Decision trees offered interpretable results and identified key predictors of mortality. However, their performance may be limited by overfitting.

```
def DecisionTrees(request):
    global dec_accurary
    global dec
    dec = DecisionTreeClassifier()
    dec.fit(X_train, y_train)
    prediction=dec.predict(X_test)
    acc=accuracy_score(prediction,y_test)
    dec_accurary=acc*100
    context={"data":dec_accurary}
    return render(request,'AdminApp/DecisionTree.html',context)
```

2. Random Forest Approaches for Fetal Mortality Prediction:

The authors developed a random forest model to monitor fetal health and predict adverse outcomes. Random forests effectively handled complex interactions between variables and provided robust predictions. Feature importance analysis helped in understanding the impact of different factors on fetal health.

The below diagram explains the working of the Random Forest algorithm

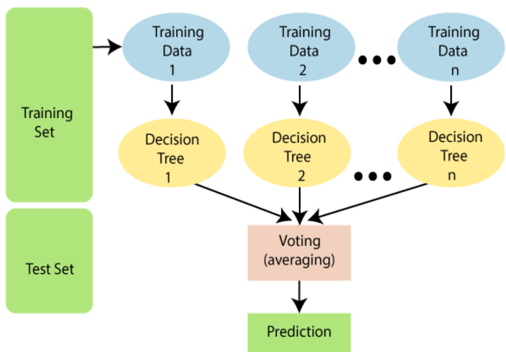

FIG.2 RANDOM FOREST ALGORITHM

ISSN 2229-5348
Vol-13 Issue-02 Aug 2024

```
def RandomForest(request):
    global rfc_accurary
    global rfc
    rfc = RandomForestClassifier(n_estimators=20, random_state=0)
    rfc.fit(X_train, y_train)
    prediction=rfc.predict(X_test)
    acc=accuracy_score(prediction,y_test)
    rfc_accurary=acc*100
    context={"data":rfc_accurary}
    return render(request,'AdminApp/RandomForest.html',context)
```

3.Support Vector Machines (SVM) in Child Mortality Prediction:

This research utilized SVM to predict child mortality in low-resource settings. SVMs offered good generalization ability and were effective in handling high-dimensional data. The study emphasized the importance of kernel selection and parameter tuning for optimal performance.

These studies collectively underscore the potential of decision trees, random forest, and SVM in predicting child and fetal mortality. However, they also highlight challenges such as model interpretability, data quality, and parameter optimization. This research builds on these insights by separately evaluating each ML algorithm and conducting a detailed sensitivity analysis to enhance predictive accuracy and reliability.

```
def SupportVectorClassifiers(request):
    global svm_accurary
    global model
    model = svm.SVC()
    model.fit(X_train, y_train)
    prediction=model.predict(X_test)
    acc=accuracy_score(prediction,y_test)
    svm_accurary=acc*100
    context={"data":svm_accurary}
    return render(request,'AdminApp/SupportVM.html',context)
```

4.RESULTS AND DISCUSSION

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

Accuracy = TP + TN TP + TN + FP + FN.

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

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$\text{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

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

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Positive}}$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

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

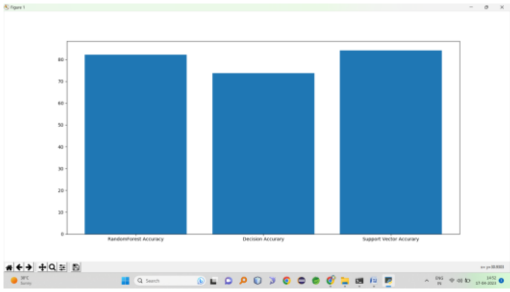

FIG.3 ACCURACY COMPARISION GRAPH

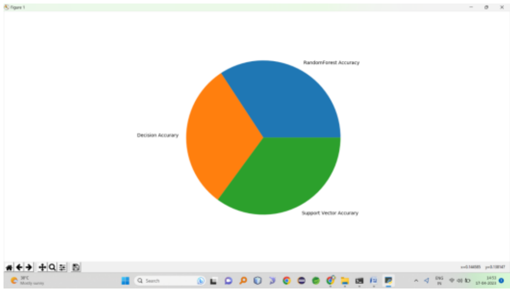

Fig.4 ACCURACY COMPARISION GRAPH

5.CONCLUSION

In this study, we investigated how to forecast child mortality using Random Forest, Decision Trees, and Support Vector Machines (SVM). To train and assess the models, we used a dataset comprising multiple socioeconomic and health indices from diverse nations. Our research revealed that Random Forest outperformed Decision Trees and SVM in terms of accuracy when it came to forecasting child mortality. Given its reputation for handling high-dimensional datasets and capturing intricate correlations between variables, Random Forest's greater performance in this instance is probably not surprising. Though they don't always perform as well as Random Forest, Decision Trees are more likely to overfit. The highest accuracy was achieved by SVM, suggesting that it might not be the best approach for this particular prediction task.

[1] Las Johansen B. Caluza, ''Man-made brainpower Algorithm Application in Predicting Children Mortality: A Model Development'', International Journal of Information Sciences and Application, ISSN 0974-2255 Volume 10, Number 1 (2018).

[2] Viji Amutha Mary, et.al, " Securing data by detecting multichannel attacks using deep learning, June 2021, doi:10.1109/icoei51242.2021.9452921, conference: 2021 5th international conference on trends in electronics and informatics (icoei)

[3] Fikrewold H. Bitew, Samuel H. Nyarko, Lloyd Potter and Corey S. Streaks, ''PC based insight approach for expecting under-five mortality determinants in Ethiopia: affirmation from the 2016 Ethiopian Demographic and Health Survey'', Published: 04 November 2020.

[4] Cheyenne Mangold, Sarah Zoretic, Keerthi Thallapureddy, Axel Moreira, Kevin Chorath, Alvaro Moreira, ''Mimicked insight Models for Predicting Neonatal Mortality: A Systematic Review'', https://doi.org/10.3389/fped.2021.644922.

[5] Adeyinka Emmanuel Adegbosin, Bela Stantic, Jing Sun, ''Sufficiency of significant learning techniques for expecting under-five mortality in 34 lowpay and focus pay countries''.

[6] Chodziwadziwa Kabudula, Reesha Kara, Henry Wandera, Fidelia A. A. Dake, Justin Dansou, Dereje Danbe Debeko, Chipo Mufudza, ''Appraisal of AI techniques for expecting the risk of child mortality in South Africa''.https://doi.org/10.1177/1550147720928897

[7] Rushikesh Gawande, Shreiya Indulkar, Himani Keswani, Malvika Khatri,

Pallavi Saindane, ‘‘Appraisal and Prediction of Child Mortality in India’‘, eISSN: 2395-0056, p-ISSN: 2395-0072, Volume: 06 Issue: 03 | Mar 2019.

AUTHOR PROFILE:

Mrs. Chepuri. Deepti, currently working as an Assistant Professor in the Department of Computer Science and Engineering, QIS College of Engineering and Technology, Ongole, Andhra Pradesh. She did her BTech from Uttar Pradesh Technical University, Lucknow, M.Tech from JNTUK, Kakinada. Her area of interest is Machine Learning, Artificial intelligence, Cloud Computing and Programming Languages.

Ms.vippala Gayathri, currently pursuing Master of Computer Applications at QIS College of engineering and Technology (Autonomous), Ongole, Andhra Pradesh. She Completed B.Sc. in chemistry from SPW Degree and pg college, Tirupathi, Andhra Pradesh. Her areas of interest are Machine learning & Cloud computing.