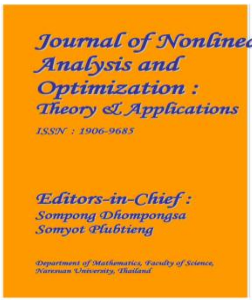
Journal of Nonlinear Analysis and Optimization
Vol. 17, Issue. 2 : 2026
ISSN : **1906-9685**



CHILD MORTALITY PREDICTION USING MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE TECHNIQUES

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

NRI INSTITUTE OF TECHNOLOGY, VISADALA

Meghana nuthalapati **B Srinivas**

ABSTRACT

Child mortality remains one of the major public health challenges across the world, especially in developing and underdeveloped countries. It refers to the death of children under the age of five due to factors such as malnutrition, poor healthcare facilities, infectious diseases, lack of vaccination, unsafe drinking water, poverty, and inadequate maternal care. Predicting child mortality at an early stage can help healthcare organizations and governments take preventive actions to reduce death rates and improve child health conditions. With the rapid growth of healthcare data and computational technologies, Machine Learning (ML) techniques have become highly effective tools for analyzing medical data and predicting health-related outcomes accurately.

This project presents a Machine Learning-based Child Mortality Prediction System designed to identify the probability of child mortality using various healthcare and socio-economic factors. The system uses datasets containing information such as age, birth weight, nutritional status, immunization records, maternal health, access to clean water, household income, and medical history. Data preprocessing techniques including data cleaning, handling missing values, normalization, and feature selection are applied to improve the quality of the dataset and increase prediction performance.

Several Machine Learning algorithms such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Naive Bayes are implemented to classify and predict child mortality risks. Among these models, ensemble methods like Random Forest

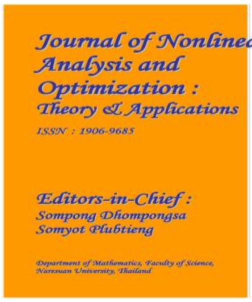
provide better accuracy and reliability due to their ability to handle complex healthcare data efficiently. The system evaluates model performance using metrics such as accuracy, precision, recall, and F1-score.

The proposed system helps healthcare professionals and government agencies identify high-risk cases and take timely preventive measures. It supports early diagnosis, effective healthcare planning, and improved decision-making in child healthcare management. The implementation of Machine Learning in healthcare prediction systems can significantly reduce child mortality rates and enhance public health services. This project demonstrates how advanced ML techniques can contribute to building intelligent healthcare systems that support child survival and overall community well-being.

INTRODUCTION

Child mortality is one of the most critical public health issues affecting many countries around the world, particularly developing and underdeveloped nations. It refers to the death of children below the age of five due to various health, environmental, and socio-economic factors. According to global health organizations such as World Health Organization and United Nations Children's Fund, millions of children die every year from preventable causes including malnutrition, infectious diseases, premature birth complications, lack of immunization, poor sanitation, unsafe drinking water, and inadequate healthcare facilities. Reducing child mortality is considered one of the major goals of public health and sustainable development programs worldwide.

Traditional methods for predicting child mortality mainly rely on statistical analysis and manual examination of healthcare records. Although these approaches provide useful insights, they are often time-consuming and less effective in handling large and complex healthcare datasets. In recent years, the rapid growth of digital healthcare systems and medical databases has created opportunities for using advanced technologies to improve healthcare prediction and decision-making processes.



Machine Learning (ML), a branch of Artificial Intelligence (AI), has emerged as a powerful technology for analyzing medical and healthcare data. ML algorithms can identify hidden patterns, relationships, and trends within large datasets, enabling accurate prediction of health outcomes. By learning from historical healthcare records, ML models can predict the probability of child mortality and identify high-risk cases at an early stage. This helps healthcare professionals and government agencies take preventive actions and provide timely medical support to vulnerable children.

This project focuses on developing a Machine Learning-based Child Mortality Prediction System using healthcare and socio-economic data. The system considers several important factors such as birth weight, nutritional status, maternal health, immunization records, household income, access to clean water, sanitation facilities, and medical history. These features are processed and analyzed using various ML algorithms including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Naive Bayes.

The proposed system aims to improve the accuracy and efficiency of child mortality prediction while reducing manual effort in healthcare analysis. It can assist doctors, healthcare workers, policymakers, and public health organizations in identifying high-risk children and planning effective healthcare interventions. The integration of Machine Learning into healthcare systems can support early diagnosis, better resource allocation, and improved healthcare management.

In conclusion, the application of Machine Learning techniques in child mortality prediction has the potential to save lives by enabling early risk detection and preventive healthcare strategies. The proposed system contributes to the development of intelligent healthcare solutions that promote child survival, public health improvement, and overall community well-being.

LITERATURE SURVEY

Child mortality prediction has become an important research area in the healthcare domain due to the increasing need for early diagnosis and preventive healthcare management. Researchers across the world have developed various statistical, data mining, and Machine Learning-based approaches to analyze healthcare data and predict child mortality risks. This literature survey discusses existing studies, techniques, and methodologies used for child mortality prediction and highlights their advantages and limitations.

Earlier studies on child mortality mainly used traditional statistical methods such as regression analysis and demographic analysis. These methods focused on identifying relationships between factors such as maternal health, nutrition, vaccination, sanitation, and mortality rates. Although statistical approaches provided valuable insights, they were limited in handling large and complex healthcare datasets. They also required manual interpretation and were less efficient in predicting individual risk cases accurately.

With the advancement of Artificial Intelligence and data analytics, researchers started applying Machine Learning techniques to healthcare prediction systems. Logistic Regression was one of the most commonly used algorithms for predicting child mortality because of its simplicity and effectiveness in binary classification problems. Several studies reported that Logistic Regression performs well when the dataset contains structured and balanced healthcare records. However, the algorithm struggles with highly complex and non-linear relationships among features.

Decision Tree algorithms were later introduced to improve prediction accuracy and interpretability. Decision Trees can easily represent healthcare decision-making processes and identify important risk factors affecting child mortality. However, single Decision Trees may suffer from overfitting problems when handling large datasets. To overcome this issue, researchers implemented ensemble learning methods such as Random Forest and Gradient Boosting. Random Forest combines multiple decision trees to improve prediction accuracy, reduce overfitting, and handle missing healthcare data efficiently. Many research studies showed that Random Forest provides better performance for healthcare prediction tasks compared to traditional ML models.

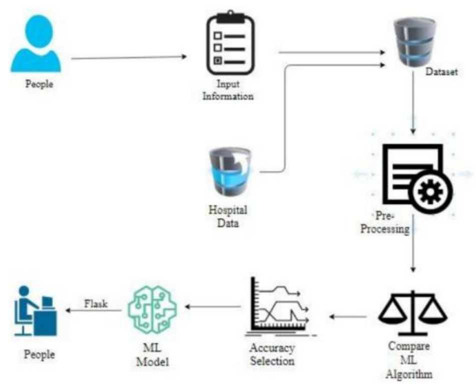
Journal of Nonlinear Analysis and Optimization
Vol. 17, Issue. 2 : 2026
ISSN : **1906-9685**

Support Vector Machine (SVM) has also been widely used in child mortality prediction because of its capability to classify complex healthcare data accurately. SVM performs effectively for high-dimensional datasets but requires higher computational resources for large-scale data processing. Naive Bayes classifiers were applied in some studies because of their fast processing speed and simplicity, especially for probabilistic healthcare analysis.

Recent research has focused on Deep Learning techniques such as Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). These models can automatically learn hidden patterns from large healthcare datasets and improve prediction accuracy significantly. However, Deep Learning methods require large training datasets, high computational power, and longer training time.

Several studies also emphasized the importance of socio-economic and environmental factors such as poverty, education, sanitation, and access to healthcare services in child mortality prediction. Researchers integrated healthcare records with demographic and environmental data to improve prediction reliability and support public health decision-making.

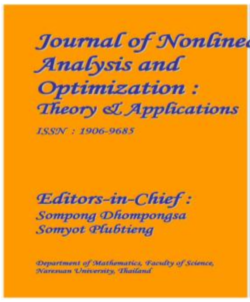
SYSTEM ARCHITECTURE



IMPLEMENTATION

The implementation of the Child Mortality Prediction System using Machine Learning involves several stages such as data collection, preprocessing, feature extraction, model training, prediction, and performance evaluation. The main objective of the implementation is to develop an intelligent healthcare system capable of predicting child mortality risks accurately using medical and socio-economic data.

1. Data Collection



The first step in implementation is collecting healthcare-related datasets from hospitals, healthcare organizations, government health departments, and public healthcare databases. The dataset contains important information related to child health and environmental conditions. Common attributes include age, birth weight, nutritional status, vaccination records, maternal health conditions, access to healthcare services, sanitation facilities, household income, education level of parents, and medical history. The dataset also contains target labels indicating whether the child survived or faced mortality risk.

2. Data Preprocessing

Healthcare datasets often contain missing values, duplicate entries, inconsistent formats, and noisy data. Therefore, preprocessing is necessary to improve data quality and model performance. Missing values are handled using techniques such as mean replacement or data removal. Categorical data like gender, vaccination status, and disease type are converted into numerical values using encoding methods. Data normalization and scaling are applied to ensure that all features are within a similar range. Preprocessing helps improve the efficiency and accuracy of Machine Learning algorithms.

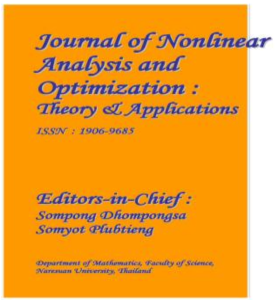
3. Feature Selection and Extraction

Feature selection is an important step in identifying the most relevant factors affecting child mortality. Features such as birth weight, malnutrition level, maternal age, immunization status, and healthcare access are extracted and analyzed. Unnecessary or less important features are removed to reduce complexity and improve prediction accuracy. Proper feature extraction helps the system identify meaningful healthcare patterns effectively.

4. Machine Learning Model Training

After preprocessing, the dataset is divided into training and testing datasets. Various Machine Learning algorithms such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Naive Bayes are implemented and trained using historical healthcare records. The models learn relationships between healthcare factors and child mortality outcomes. Among these

Journal of Nonlinear Analysis and Optimization
Vol. 17, Issue. 2 : 2026
ISSN : **1906-9685**



algorithms, Random Forest generally provides better accuracy and reliability because it combines multiple decision trees and reduces overfitting problems.

5. Prediction and Risk Analysis

Once the models are trained, the system predicts the probability of child mortality for new healthcare records. The trained model analyzes the input features and classifies cases into high-risk or low-risk categories. If the system identifies a child as high-risk, healthcare professionals can take preventive measures such as medical treatment, nutritional support, and regular monitoring.

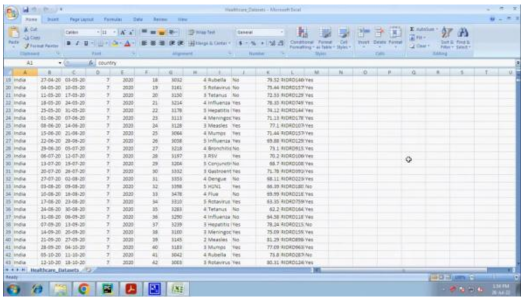
6. Performance Evaluation

The performance of the implemented models is evaluated using metrics such as accuracy, precision, recall, F1-score, and confusion matrix. These evaluation methods help determine the effectiveness of the prediction system in identifying mortality risks correctly while minimizing prediction errors.

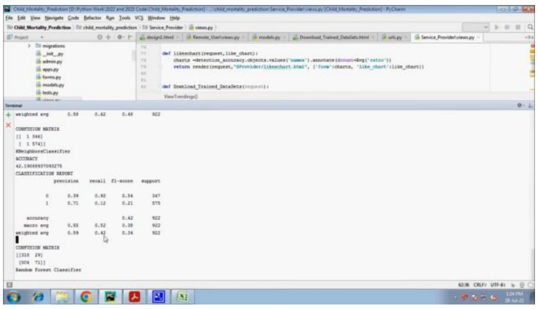
SCREEN SHOTS

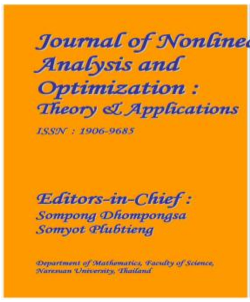
Fig No: 1 Login Page

DATASET

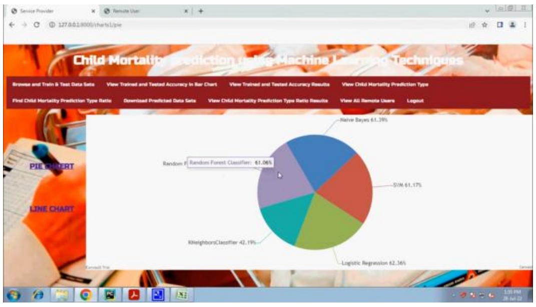


ALGORITHM :





PIE CHART:



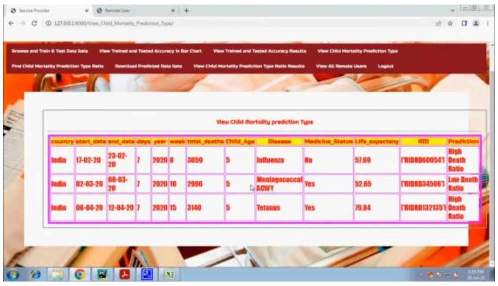


Fig No: 3 Prediction Page

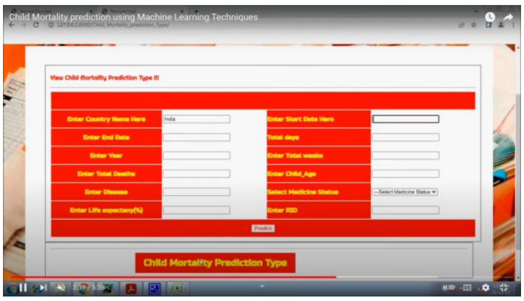
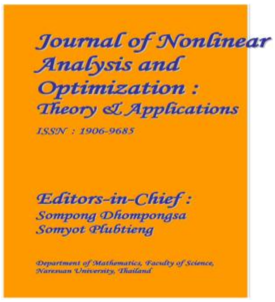


Fig No: 4 Details

CONCLUSION

Child mortality remains a major global healthcare challenge, especially in developing countries where access to quality healthcare, nutrition, sanitation, and medical facilities is limited. Early identification of children who are at high risk of mortality is essential for providing timely treatment and preventive healthcare services. Traditional prediction methods based on manual analysis and statistical approaches are often time-consuming and less effective in handling large healthcare datasets. Therefore, the use of advanced technologies such as Machine Learning has become highly important in modern healthcare systems.

This project presented a Machine Learning-based Child Mortality Prediction System designed to analyze healthcare and socio-economic data to predict mortality risks among children. The system used various features such as birth weight, nutritional status, immunization records, maternal health, sanitation facilities, household income, and medical history for prediction analysis. Several Machine Learning algorithms including Logistic Regression, Decision Tree, Random Forest, Support



Vector Machine (SVM), and Naive Bayes were implemented and evaluated.

Among the implemented algorithms, ensemble learning methods such as Random Forest demonstrated better accuracy, reliability, and prediction performance compared to traditional methods. The system successfully identified high-risk cases and supported early healthcare intervention. By using preprocessing, feature extraction, and classification techniques, the proposed model improved prediction accuracy and reduced manual effort in healthcare analysis.

The project also highlighted the significant role of Machine Learning in transforming healthcare systems into intelligent and data-driven environments. The developed system can assist doctors, healthcare professionals, government agencies, and policymakers in improving healthcare planning, resource allocation, and child welfare programs. Early prediction of mortality risks can help reduce child death rates and improve overall public health conditions.

Although the proposed system provides effective prediction results, certain challenges still exist, such as incomplete healthcare data, imbalanced datasets, and varying healthcare conditions across regions. Continuous dataset updates and advanced learning models are required to improve prediction accuracy further. Future integration of Deep Learning, cloud computing, and real-time healthcare monitoring systems can enhance the efficiency and scalability of the system.

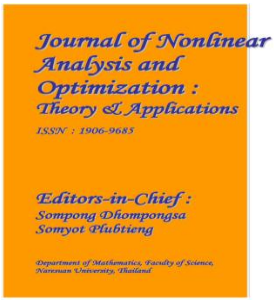
REFERENCE

1. World Health Organization. Child Mortality and Global Health Reports.
   WHO Official Website

2. United Nations Children's Fund. Child Survival and Health Statistics Reports.
   UNICEF Official Website

3. Machine Learning Applications in Healthcare Prediction Systems, *International Journal of Computer Science and Information Technologies*, 2021.

4. Predicting Child Mortality Using Machine Learning Algorithms,



International Journal of Advanced Research in Computer Science, Vol. 12, No. 4, 2022.

5. Healthcare Data Analysis and Mortality Prediction Using Artificial Intelligence Techniques, *IEEE International Conference on Healthcare Informatics*, 2020.

6. Random Forest-Based Prediction Model for Child Healthcare Monitoring, *International Journal of Engineering Research & Technology (IJERT)*, 2023.