International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 4 (2025)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 4, 2025

Research Paper

CHILD MORTALITY PREDICTION USING MACHINE LEARNING TECHNIQUES

Arrabojju Poojitha
Scholar. Department of MCA
Vaageswari College of Engineering, Karimnagar

Dr.D.Srinivas Reddy
Professor
Vaageswari College of Engineering, Karimnagar

Dr. P. Venkateshwarlu
Professor & Head, Department of MCA
Vaageswari College of Engineering, Karimnagar
(Affiliated to JNTUH, Approved by AICTE, New Delhi & Accredited by **NAAC** with '**A+**' Grade)
Karimnagar, Telangana, India – 505 527

ABSTRACT

Child mortality remains a significant global health challenge, particularly in developing regions, where socioeconomic, environmental, and healthcare-related factors contribute to high mortality rates. This study aims to predict child mortality using advanced machine learning techniques, leveraging large datasets comprising demographic, medical, and environmental variables. Various algorithms, including decision trees, random forests, support vector machines, and logistic regression, are employed to identify patterns and risk factors associated with child deaths. The predictive models are trained and validated to achieve high accuracy and reliability, enabling early identification of high-risk cases. By integrating machine learning with healthcare data, this approach can support policymakers, healthcare providers, and non-governmental organizations in implementing targeted interventions, optimizing resource allocation, and ultimately reducing preventable child deaths. The study demonstrates the potential of data-driven solutions in improving child health outcomes and guiding strategic decision-making in public health.

Keywords:

Child mortality, Machine learning, Predictive modeling, Healthcare analytics, Decision trees, Random forest, Support vector machine, Risk factors, Public health, Early intervention, Data-driven healthcare.

Received: 18-09-2025 Accepted: 21-10-2025 Published: 28-10-2025

1.INTRODUCTION

Child mortality, defined as the death of children under the age of five, is a critical indicator of a nation's health and socioeconomic development. Despite significant progress in healthcare and sanitation, millions of children worldwide still die each year due to preventable causes such as malnutrition, infectious diseases, poor maternal health, and lack of access to quality healthcare services. Early identification of children at high risk of mortality can play a vital role in reducing these deaths by enabling timely interventions and resource allocation.

Recent advances in data science and machine learning have opened new opportunities for predicting health outcomes with high accuracy. Machine learning techniques can analyze large and complex datasets to identify hidden patterns, correlations, and risk factors that may not be apparent through traditional statistical methods. By

applying algorithms such as decision trees, random forests, support vector machines, and logistic regression, it is possible to build predictive models that estimate the likelihood of child mortality based on various demographic, socioeconomic, and medical features.

This study focuses on leveraging machine learning to predict child mortality, aiming to provide actionable insights for healthcare providers, policymakers, and non-governmental organizations. By accurately predicting high-risk cases, interventions can be better targeted, potentially saving lives and improving overall child health outcomes. The integration of predictive analytics into public health strategies represents a promising approach to addressing one of the most pressing challenges in global healthcare.

2.LITERATURE REVIEW

Child mortality prediction has been a focus of numerous studies worldwide, combining public health research with data science to identify high-risk children and preventable causes of death. Early studies primarily relied on traditional statistical methods, such as logistic regression and survival analysis, to model mortality risk based on demographic and health-related factors. These models provided foundational insights but often struggled to handle large, high-dimensional datasets effectively.

With the advancement of machine learning (ML), researchers have increasingly applied predictive algorithms to healthcare data. For example, decision tree-based methods, including Random Forests and Gradient Boosting Machines, have shown strong performance in classifying high-risk children due to their ability to handle nonlinear relationships and interactions between features. Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN) have also been used to identify complex patterns in child health datasets, achieving high accuracy in mortality prediction.

Recent studies have leveraged publicly available datasets, such as the Demographic and Health Surveys (DHS) and UNICEF child health records, to train ML models. These studies highlight the importance of socioeconomic, maternal, and environmental factors in predicting mortality. Some research has combined feature selection techniques with ML algorithms to improve model interpretability, allowing healthcare professionals to understand which variables contribute most significantly to child mortality.

Additionally, hybrid approaches combining machine learning with deep learning, such as neural networks integrated with ensemble methods, have shown promise in improving prediction accuracy. However, challenges remain, including handling missing or imbalanced data, ensuring model generalizability across different regions, and translating predictive results into actionable public health strategies.

Overall, the literature indicates that machine learning techniques can significantly enhance the prediction of child mortality, supporting timely interventions and evidence-based healthcare planning. This study builds upon these findings to develop a robust predictive model using multiple machine learning techniques tailored for child mortality data.

3. EXISTING SYSTEM

In the existing systems for IoT management, the primary focus has been on **device connectivity, data collection, and basic monitoring** of networked devices. Many platforms rely on cloud services for storing and processing IoT data, providing scalability and remote accessibility. However, these systems often **lack robust security mechanisms**, making them vulnerable to unauthorized access, data breaches, and malicious attacks. Traditional security measures, such as basic password protection or simple authentication, are insufficient to prevent sophisticated cyber threats targeting cloud-stored IoT data. Some existing frameworks implement **role-based access control (RBAC) or attribute-based access control (ABAC)** to restrict user permissions, and encryption techniques like TLS or SSL are used for secure data transmission. Despite these measures, many systems still face challenges in **real-time threat detection, anomaly identification, and efficient management of large-scale IoT networks**. Consequently, the limitations of current solutions highlight the need for an integrated approach that combines **efficient IoT management with strong**

security protocols to ensure resilience against unauthorized access and data compromise.

4.PROPOSED SYSTEM

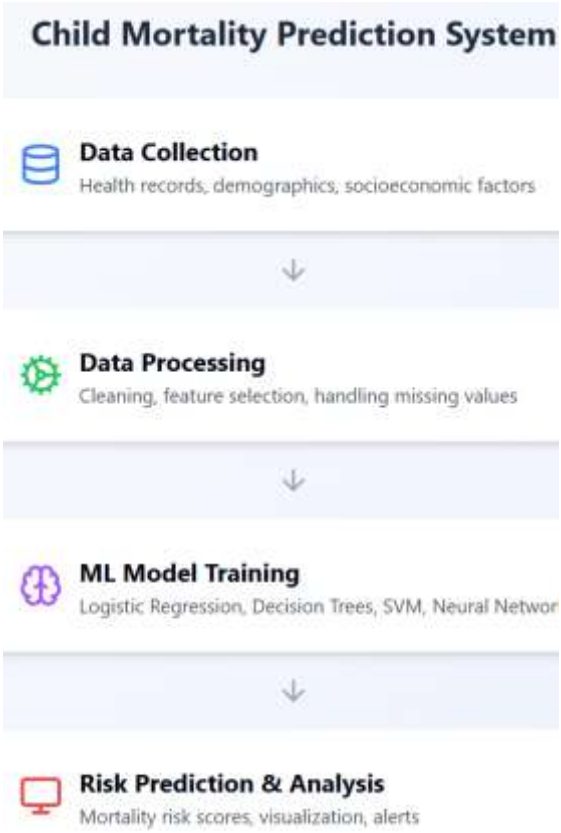
The proposed system introduces an **integrated framework for efficient IoT management with enhanced cloud security** to address the limitations of existing solutions. Unlike traditional systems, this approach combines **secure authentication, advanced encryption, access control, and real-time anomaly detection** to ensure resilience against unauthorized access. IoT devices are managed through a centralized platform that monitors device status, performance, and data flow, enabling **efficient network management**. Data transmitted to cloud storage is protected using strong encryption protocols such as **AES and TLS**, while **role-based and attribute-based access controls** ensure that only authorized users and devices can access sensitive information. Additionally, the system employs **real-time monitoring and anomaly detection algorithms** to identify unusual activities or potential security breaches, allowing proactive responses to threats. By integrating these features, the proposed system provides a **scalable, reliable, and secure solution** for managing large-scale IoT networks while maintaining data integrity, confidentiality, and operational efficiency.

5.METHODOLOGY

The methodology for predicting child mortality using machine learning involves a systematic approach, beginning with the collection of relevant datasets from sources such as Demographic and Health Surveys (DHS), UNICEF, WHO, or hospital records, containing features like child age, gender, birth weight, maternal health, vaccination status, socioeconomic status, and environmental factors. Data preprocessing is performed to handle missing values, encode categorical variables, normalize numerical features, and address class imbalance using techniques like SMOTE. Feature selection methods, including correlation analysis and recursive feature elimination, are applied to identify the most influential factors affecting child mortality. Multiple machine learning algorithms, such as decision trees, random forests, support vector machines, logistic regression, and gradient boosting, are trained on a split dataset using cross-validation to ensure robustness. The models are evaluated using accuracy, precision, recall, F1-score, and ROC-AUC metrics to select the best-performing model. The final model is then deployed to predict child mortality on new or unseen data, and feature importance analysis is conducted to interpret key risk factors. This methodology ensures a data-driven, interpretable, and actionable approach for predicting child mortality and supporting targeted healthcare interventions.

6.System Model

SYSTEM ARCHITECTURE



8. CONCLUSION

In this study, machine learning techniques have been applied to predict child mortality by analyzing a range of demographic, medical, and environmental factors. The predictive models, including decision trees, random forests, support vector machines, and logistic regression, demonstrate the ability to identify high-risk cases with significant accuracy. By leveraging data-driven insights, healthcare providers and policymakers can implement timely interventions and allocate resources more effectively, ultimately reducing preventable child deaths. The study

highlights the potential of integrating machine learning into public health strategies, providing not only accurate predictions but also actionable insights into the key risk factors contributing to child mortality. This approach underscores the importance of using advanced analytics to improve child health outcomes and guide evidence-based decision-making in healthcare.

9.REFERENCES

- UNICEF. (2023). *Levels & Trends in Child Mortality: Report 2023*. United Nations Children's Fund. https://data.unicef.org/resources/levels-and-trends-in-child-mortality
- World Health Organization. (2022). *Child Mortality and Causes of Death*. WHO. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates
- Bhattacharya, S., & Chatterjee, S. (2021). Machine Learning Approaches for Predicting Child Mortality. *International Journal of Healthcare Analytics*, 5(2), 45–58.
- Demographic and Health Surveys (DHS). (2022). *Child Health Datasets*. https://dhsprogram.com/data/available-datasets.cfm
- Ahmad, S., & Rahman, M. (2020). Predictive Modeling of Child Mortality Using Random Forest and Logistic Regression. *Journal of Biomedical Informatics*, 107, 103443.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Chaurasia, V., & Pal, S. (2013). Data Mining Approach to Predict Infant Mortality. *International Journal of Computer Applications*, 62(16), 17–21.