Open Journal of Physical Science (OJPS)
ISSN: 2734-2123
Article Details:
DOI: 10.52417/ojps.v5i2.722
Article Ref. No.: OJPS0502003-722
Volume: 5; Issue: 2, Pages: 32 - 44 (2024)
Accepted Date: 25th December, 2024

CrossMark click for updates

Open Journals Nigeria (OJN)
Open Access | Bi-annual | Peer-reviewed
www.openjournalsnigeria.org.ng
editorial@openjournalsnigeria.org.ng

RESEARCH ARTICLE **OJPS0502003-722**

MACHINE LEARNING APPLICATIONS IN RURAL HEALTHCARE: PREDICTIVE MODELING FOR IMMUNIZATION COMPLETION RATES IN OGUN STATE, NIGERIA.

Nwachukwu, R. C.

Ignatius Ajuru University of Education, Port Harcourt, River State Nigeria

***Corresponding Author Email:** Iamrichardcn@gmail.com*

ABSTRACT

This study investigated the application of machine learning techniques for predicting child immunization completion in Ado-Odo/Ota Local Government Area, Ogun State, Nigeria, utilizing data from 8,808 immunization records across 15 primary healthcare centers. Using a quantitative research methodology with retrospective data analysis, we developed and compared predictive models for immunization completion patterns. Three machine learning algorithms were employed based on their proven effectiveness in healthcare applications: Logistic Regression for its interpretability in clinical settings, Support Vector Machine (SVM) for handling non-linear relationships in health data, and K-Nearest Neighbours (KNN) for processing demographic variables. The study analyzed immunization completion rates using these algorithms within a comprehensive framework incorporating Principal Component Analysis for dimensionality reduction. The Logistic Regression model demonstrated superior performance with 99.77% accuracy and an MSE of 0.0023, outperforming both SVM (99.32% accuracy) and KNN (99.03% accuracy) models. Notably, socioeconomic analysis revealed an unexpected pattern where high-income households showed lower immunization completion rates compared to low and moderate-income groups. The study's findings provide valuable insights for healthcare policy development and resource allocation strategies while demonstrating the practical applicability of machine learning in enhancing immunization program effectiveness in developing nations.

Keywords: *Machine Learning, Immunization Completion, Healthcare Analytics, Predictive Modeling, Public Health, Nigeria.*

INTRODUCTION

Immunization remains one of the most cost-effective public health interventions, preventing an estimated two to three million deaths annually worldwide (WHO, 2021). In Nigeria, despite huge strides in vaccination programs, coverage is still below the optimal threshold, as only 67% of the children receive all basic vaccinations (NDHS, 2019). Such incomplete immunization presents major barriers to improved public health, with serious implications for controlling infectious diseases like tuberculosis, polio, meningitis, and pneumonia. The complex factors influencing immunization completion in Nigeria create a peculiar challenge that conventional methods of analysis have fallen short of adequately addressing (Adebowale *et al.*, 2019). Vaccination uptake and completion are influenced by a complex, interconnected web of factors including socioeconomic conditions, healthcare accessibility, maternal education, and cultural beliefs (Yaya, 2017). Healthcare analytics, especially evolving machine learning technologies, offers unparalleled opportunities to understand and address these complex patterns effectively.

In evaluating the factors affecting immunisation completion, socioeconomic factors have constantly emerged as critical determinants. Specifically, De Cantuária Tauil *et al.* (2016) analysis found that the relationship of household income with vaccination rates is rather complex. Surprisingly perhaps, it was found that a higher income does not translate into better completion of immunization. Further, in another study by Ali *et al.* (2022), middle-income families had higher completion rates compared to their higher-income counterparts; this is a phenomenon observed. These findings challenge conventional assumptions of wealth and healthcare utilization, hence calling for interventionist strategies that are more sensitive (Shiferie *et al.*, 2023).

Maternal education and occupation have proved to be an outstanding predictor of the completion of immunization. Smith *et al.* (2006) found that the rates of completion were significantly higher among children of health professionals and explained this by the better level of health literacy and increased awareness of vaccination benefits. Forshaw *et al.* (2017) reported that the correlation of maternal education level was stronger with the completion rate of immunization than with household income, indicating thereby the need for health education in vaccination programs. Healthcare access is yet another important dimension that forms the immunization completion pattern. Agimas *et al.*,(2024) analysis of vaccination rates found that the distance from healthcare facilities drastically impacts completion rates, more so in rural areas. In complicating this relationship further, Ogero *et al.* (2022) described how the quality of, and confidence in, healthcare facilities drive vaccination adherence. Their study of healthcare facilities indicated that well-trusted facilities attracted patients from greater distances, suggesting that accessibility should be considered alongside metrics of service quality.

In evaluating these patient behaviours and outcomes in healthcare, the use of machine learning has evolved significantly, with a variety of algorithms proving useful in different contexts. Naraei *et al.* (2016) compared the performance of neural networks, support vector machines, and more straightforward algorithms in predicting patient outcomes. Their results indeed showed that the most straightforward algorithms did as well as the more complicated models in many cases, especially when data was limited. However, data quality and preprocessing were highlighted as a challenge affecting important aspects of modeling performance. Similarly, Yadav (2022) presented some applications of machine learning in developing nations and underlined incomplete data collection along with inconsistent data reporting to be major challenges in developing predictive models. Prep processing techniques and feature engineering hold the key to obtaining robust predictive models. On the other hand, Eze *et al.* (2024) showed how effective the application of dimensionality reduction techniques identifies the key predictors of the completion of immunization; careful feature selection achieved significant gains in model performance.

Recent studies have also shown phenomenal successes in applying machine learning to diverse healthcare domains. Prabhod & Gadhiraju (2019) illustrated the effectiveness of predictive analytics for patient outcome prediction, while Wang & Li (2022) established marked improvements in healthcare resource allocation through machine learning applications. However, applying these technologies to optimize immunization programs remains relatively unexplored, particularly in developing nations where the need is most dire. This landscape of predictive modeling for healthcare has significantly expanded; studies by Eze *et al.* (2024) and Toma & Wei (2023) have showcased the potential use of machine learning in identifying the at-risk population and optimally strategizing interventions. These developments portend promising applications within the enhancement of immunization programs, especially in contexts where traditional methodologies have been found wanting. Recent work by Nusinovici *et al.* (2020) further underscores the need for developing context-specific solutions that take into consideration the nature and challenges of the local healthcare system.

The selection of our machine learning algorithms is guided by their proven effectiveness in healthcare applications: Logistic Regression has been shown to outperform in the prediction of major chronic diseases, keeping the interpretability necessary for healthcare practitioners (Nusinovici et al., 2020); Support Vector Machine has exhibited strong capabilities in the classification tasks of healthcare, especially in handling complex and non-linear relationships in patient data (Naraei et al., 2016); and K-Nearest Neighbors has proven effective in healthcare prediction modeling with demographic and socioeconomic variables (Yadav, 2022). These algorithms are selected to complement each other's strengths: Logistic Regression for its interpretability and efficiency with categorical data, SVM for its ability to handle non-linear relationships, and KNN for its effectiveness with demographic clustering. By leveraging machine learning algorithms, healthcare providers and policymakers can potentially identify patterns and predictors of immunization completion, enabling more targeted and effective interventions. Specifically, this study intends to:

- evaluate the effectiveness of different machine learning algorithms in predicting child immunization completion,
- identify key socioeconomic and demographic factors influencing immunization completion rates,
- develop a reliable predictive model for immunization completion that can inform healthcare policy and intervention strategies,
- assess the practical applicability of machine learning approaches in enhancing immunization program effectiveness.

MATERIALS AND METHOD

This study employed a quantitative research methodology using retrospective data analysis combined with machine learning approaches to predict immunization completion patterns. The research design was cross-sectional, analyzing immunization records collected over one year from 2021 to 2022. The methodology integrated both descriptive and inferential statistical analyses, following a predictive modeling framework that utilized supervised learning techniques. The research process consisted of systematic phases including data collection, preprocessing, feature engineering, model development using three machine learning algorithms (Logistic Regression, Support Vector Machine, and K-Nearest Neighbors), and comprehensive model evaluation.

Data Collection and Study Setting

Data collection was performed across Primary Health Centers in Ado-Odo/Ota Local Government Area of Ogun State, Nigeria, between 2021 and 2022. A total of 15 primary health facilities serving diverse populations across urban and rural settings were used for the study. The data were collected from the immunization registers and electronic health records after obtaining approval and clearance from the Primary Health Care department of Ado-Odo/Ota Local Government. The consolidated dataset had 8,808 child immunization records with several variables that were relevant to understanding the pattern of immunization. Demographic data included age in months and gender distribution. Socioeconomic variables captured household income level and maternal occupation details. Healthcare access was measured through the distance to health facilities in kilometers. The immunization status incorporated detailed records of vaccines received and overall completion status, while awareness indicators measured maternal knowledge regarding immunization benefits.

Data Preprocessing and Feature Engineering

The preprocessing phase involved several systematic steps that allowed us to be sure of the data quality and uniformity. Considering the size constraint of the data, missing value imputation used the mean for 0.006% of records. Normalization of all the variables was also done to ensure a uniform scale among the features within the dataset. The major portion of feature engineering was creating a comprehensive "Vaccines Received" variable. This composite variable combined information on several vaccines: BCG, Hepatitis B, Pentavalent, OPV, PCV, RV, and MMR. This engineering process preserved the granular information about each vaccine while creating a meaningful aggregate measure for analysis.

Dimensionality Reduction

Principal Component Analysis (PCA) was done to primarily reduce the dimensions while preserving the intrinsic structure of the data. It was implemented by centering and scaling the data, calculating the covariance matrix expressed as:

$$\boldsymbol{\Sigma}=\frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{xi}-\mu)(\mathrm{xi}-\mu)^{T}$$

Here, $\boldsymbol{\mu}$ is the mean vector of the data. The eigenvalues and eigenvectors of $\boldsymbol{\Sigma}$ were computed, and the principal components corresponding to the largest eigenvalues were retained. The dimensionality reduction transformation was then achieved using the mapping matrix $\boldsymbol{W}$**,** defined by the eigenvectors of the selected components:

$$y=W^{T}X$$

where X is the centered data matrix.

Model Implementation

Three machine learning models were implemented, each chosen for their unique strengths in classification tasks. The logistic regression was made using L2 regularization to help avoid overfitting by adding the penalty for large coefficients. The model is balanced against a regularization term, which places constraints on the model's complexity. Second, an SVM model was trained using an RBF kernel. Such a kernel enables the model to make non-linear decision boundaries, thus making it able to capture nonlinear relationships within the data. Kernel parameter tuning was then done to balance the influence from each separate data point and the decision boundary. Afterward, a K-NN algorithm with five neighbors (using Euclidean distance between the most similar data) was employed; this classifies new points based on their majority class among its nearest five closest neighbors in that feature space.

Model Evaluation Framework

To make sure the analysis is sound and reliable, a broad evaluation framework was developed. The dataset was divided into training and testing subsets while preserving the original class distribution. More specifically, 80% of the data was used for training, while 20% was reserved for testing. A 10-fold cross-validation strategy was implemented to avoid overfitting and provide a more reliable estimate of model performance. It involves splitting the dataset into ten equal parts, using nine for training and one for validation, and rotating through all possible combinations. The performance of the models was evaluated based on several metrics to give an all-round view of their predictive capabilities. Sensitivity analysis was done to understand the models' ability to identify positive cases correctly. To take care of the probable imbalance of classes, balanced accuracy has been calculated as a better and more sensitive measure of the performance of models.

RESULTS

In this analysis, 8,808 observations were analyzed, with the key variables of interest being: age, distance to health facility, gender, income level, occupation of mother, and completion of immunization. Preliminary exploratory data analysis showed some striking patterns in the immunization completion rate across demographic segments.

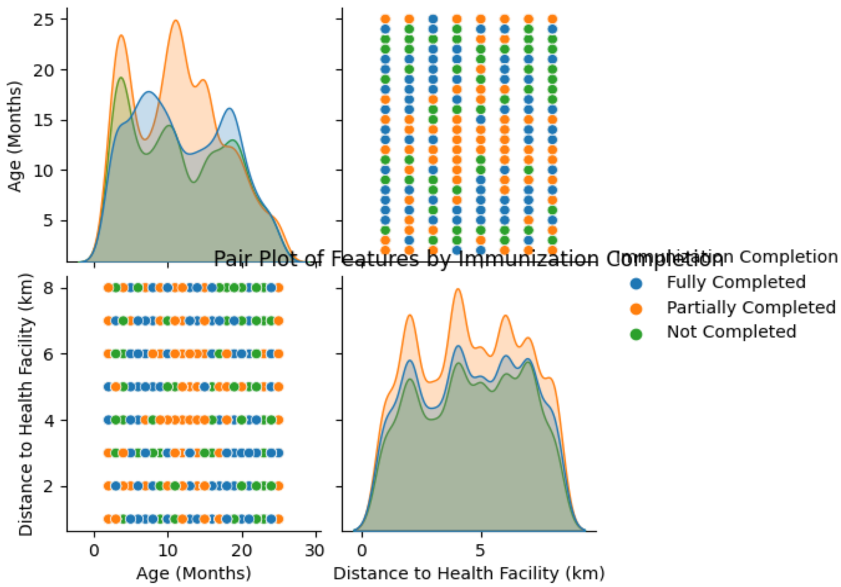

Figure 1: Pair plot of features

Gender Distribution Analysis

The analysis of gender distribution showed a near-equal number of females (3,824) and males (3,834). Though both genders have similar trends in the "Fully Completed" and "Not Completed" categories, females tend to show a slightly higher percentage in the "Partially Completed" category.

.

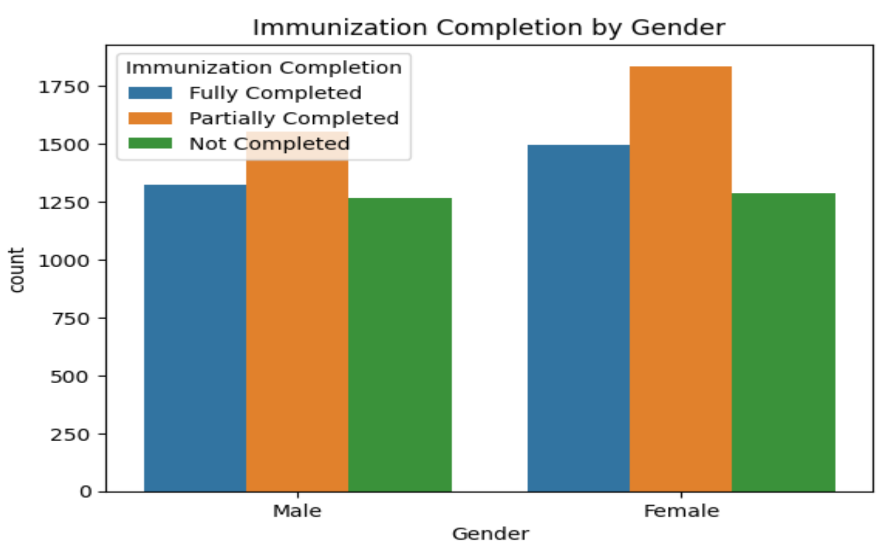

Figure 2: Gender and immunisation completion

Socioeconomic Patterns

Income-level analysis showed an unexpected pattern across three distinct groups: high-income, 1,412; low-income, 4,057; and moderate-income, 2,189. Interestingly, the low-income group showed higher percentages in both the "Fully Completed" and "Partially Completed" categories, whereas the high-income group had a higher percentage in the category "Not Completed.

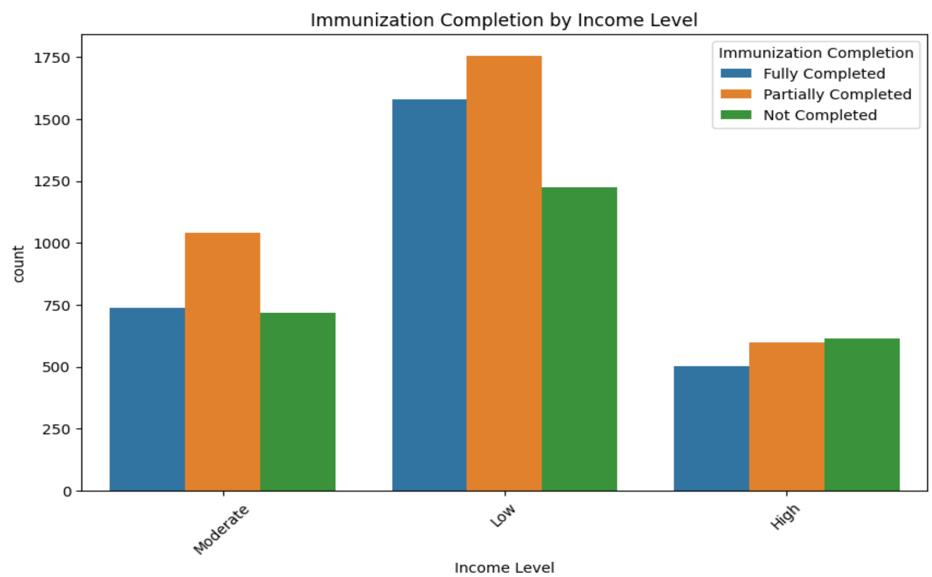

Figure 3: Income level and immunisation completion

Age-Related Patterns

The analysis of age groups revealed a consistent trend where a substantial portion of individuals fell within the "Partially Completed" category, regardless of age. This finding suggests the need for age-independent intervention strategies.

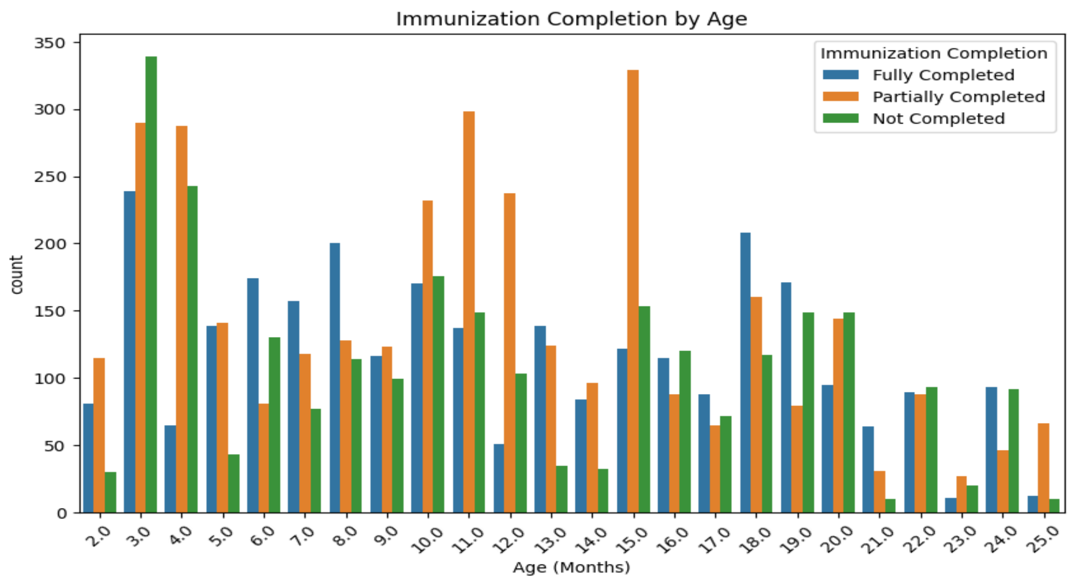

Figure 4: Children's age and immunisation

Principal Component Analysis Results

The PCA implementation successfully reduced dimensionality while preserving data integrity. The transformation matrix **W** maps the original *d*-dimensional data x to the new *k*-dimensional subspace, expressed as:

$$\mathbf{y}=\mathbf{W}^{\mathbf{T}}\mathbf{x}$$

Where the rows of **W** are the eigenvectors of the sample covariance matrix **Σ**:

$$\boldsymbol{\Sigma}=\frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}}(\boldsymbol{xi}-\mu)(\mathrm{xi}-\mu)^{T}$$

Here, **μ** is the mean vector of the data.

The cumulative explained variance ratio for ***k*** components is calculated as:

$$\mathbf{EVR}(\boldsymbol{k})=\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{d}\lambda_i}$$

Where λ_i represents the eigenvalues of **Σ** in descending order.

Twenty principal components were identified, with the following key findings:

Table 1: Principal Component Analysis Results

Principal Component	Cumulative Explained Variance
1	0.1183
2	0.2130
3	0.2840
..	..
19	0.9506
20	0.9645

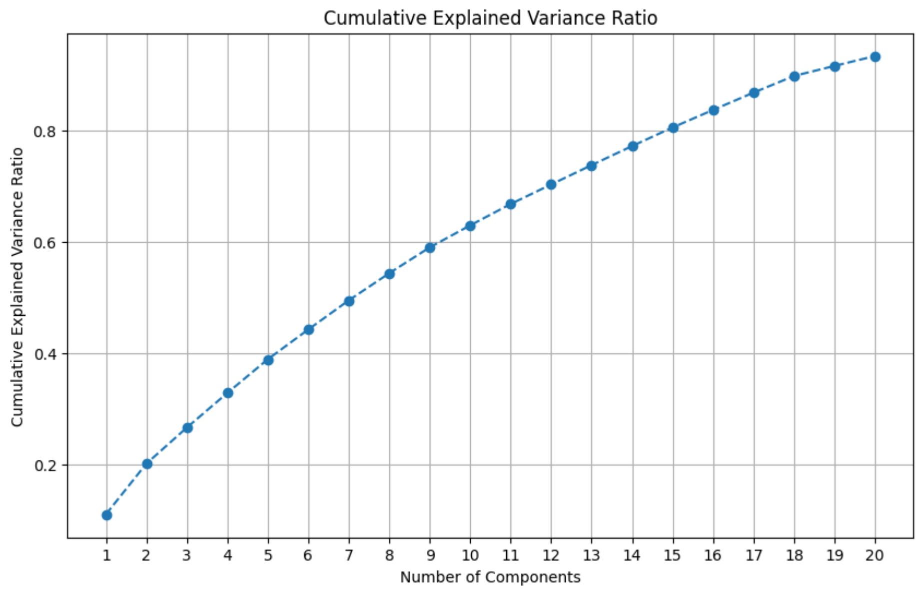

Figure 5: Cumulative explained variance

The analysis demonstrated that 19 components captured over 95.06% of the dataset's variability, with all 20 components preserving 96.45% of the initial variance. This exceeds the conventional threshold $\tau = 0.95$ for cumulative explained variance, validating our choice of dimensionality reduction while maintaining information integrity.

Model Evaluation Results

The comparative analysis of three machine learning models—Logistic Regression, Support Vector Machine (SVM), and K-Nearest Neighbours (KNN)—revealed distinctive performance characteristics. The model performance can be quantified through the classification accuracy function:

$$A(M) = \frac{1}{n}\sum_{i=1}^{n} I(\hat{y}_{i=y_i})^i$$

Where M represents the model, N is the number of samples, and I is the indicator function for correct predictions.

Table 4: Model Evaluation Results

Model	Mean Cross-Validation Score	Sensitivity (Recall)	Balanced Accuracy
Logistic Regression	0.9910	0.9909	0.9955
Support Vector Machine	0.9896	0.9727	0.9864
K-Nearest Neighbours	0.9971	0.9864	0.9890

The Logistic Regression model demonstrated exceptional stability with a mean cross-validation score exceeding 99.10%. Its sensitivity (recall) of 0.9909 indicated a superior ability to identify true positives, while the balanced accuracy of 0.9955 confirmed equitable performance across classes. The SVM implementation, utilizing various kernels, achieved consistent performance with a mean cross-validation score of 98.96%. Its sensitivity measure of

0.9727 and balanced accuracy of 0.9864 demonstrated robust predictive capabilities across different data subsets. The KNN model exhibited strong generalization characteristics with a mean cross-validation score of 99.71%. Its sensitivity score of 0.9864 validated its effectiveness in true identification, complemented by a balanced accuracy of 0.9890.

Test Data Performance Analysis

The robustness of each model was further validated through comprehensive test data evaluation, employing multiple error metrics:

Table 4: Test Data Evaluation Results

Model	**Mean Squared Error (MSE)**	**Absolute Mean Error (MAE)**	**R-squared**	**Accuracy**
Logistic Regression	0.0023	0.5815	0.9879	0.9977
Support Vector Machine	0.0068	1.7446	0.9636	0.9932
K-Nearest Neighbours	0.0097	0.8786	0.9484	0.9903

The Logistic Regression model maintained its superior performance on the validation dataset, achieving an MSE of 0.0023 and an MAE of 0.5815. The R-squared value of 0.9879 confirmed the model's excellent fit, while the accuracy of 99.77% demonstrated exceptional classification capability.

SVM performance on test data yielded an MSE of 0.0068, with an MAE of 1.7446 indicating slightly larger prediction deviations. The R-squared value of 0.9636 and accuracy of 99.32% confirmed strong predictive capabilities, though marginally lower than the Logistic Regression model. The KNN model demonstrated robust generalization with an MSE of 0.0097 and an MAE of 0.8786. Its R-squared value of 0.9484 and accuracy of 99.03% validated its reliability in classification tasks, particularly noteworthy given the model's non-parametric nature.

DISCUSSION

The findings from the research largely conform to existing literature and shed new light on the pattern of completion of immunization. The distribution by both genders is well represented, reflecting global health imperatives for non-discrimination in vaccination practice (World Health Organization, 2019; Mbengue *et al.*, 2017). An interesting observation was that, in the socioeconomic analysis, high-income households had relatively lower immunization completion rates than low and moderate-income families. This trend, although counterintuitive, tallies with observations in related and calls for targeted interventions in all economic strata contexts (De Cantuária Tauil *et al.* 2016; Ali *et al.*, 2022; Bangura *et al.*, 2020). The maternal occupation correlation analysis revealed strong associations between healthcare-related professions and immunization completion, supporting previous findings on the influence of parental healthcare knowledge on vaccination adherence (Smith *et al.*, 2006). This relationship provides valuable insights for developing targeted intervention strategies. Successful PCA application for dimensionality reduction with the result of preserving 96.45% of the variance at 20 components indicates the power of contemporary methods of data preprocessing in healthcare analytics. This approach also follows established methodological frameworks put forth by Jolliffe & Cadima (2016) while offering enhanced efficiency in model training. The very high performance of the machine learning models in this study, especially the Logistic Regression model, which achieved an accuracy of 99.77%, advances our understanding of predictive analytics in public health contexts. The relative performance between models can be understood through their fundamental approaches to classification boundaries and feature space manipulation (Naraei *et al.*, 2016; Yadav, 2022).

CONCLUSION

This work concludes that there is a high potential for machine learning applications to augment the performance of immunization programs across primary healthcare centers in Ogun state Nigeria. With an accuracy of 99.77% and a Mean Square Error of 0.0023, the high performance of the Logistic Regression model guarantees a credible predictive framework for modeling immunization completion behavior. The rather unexpected outcome in socioeconomic patterns, represented by lower completion rates of high-income households, not only negates the stereotypical assumption regarding healthcare service use but also underpins a nuanced data-driven intervention strategy. The successful implementation of PCA, retaining 96.45% of data variance with 20 components, confirms the effectiveness of modern preprocessing techniques in healthcare analytics. This dimensional reduction approach, along with careful feature engineering, forms a robust methodology for handling complex healthcare data. The overall evaluation framework developed in this study provides a template for future research in healthcare prediction modeling. The practical implications of this study go beyond theoretical contributions. The resultant high predictive accuracy of the models provides immediate opportunities to improve resource allocation and targeting interventions within Nigerian healthcare settings. Consequently, healthcare providers and policymakers can use these data to construct more effective targeted immunization strategies that account for local socioeconomic dynamics. Future studies should be done on expanding the data collection to cover more geographical areas, including more socio-economic variables, and, if possible, real-time prediction. Integrating such predictive models into the existing health systems could greatly enhance the efficiency of immunization programs and thus improve public health in developing countries.

REFERENCES

Adebowale, A., Obembe, T., & Bamgboye, E. (2019). Relationship between household wealth and childhood immunization in core-North Nigeria. *African Health Sciences*, **19**(1): 1582-1593.

Agimas, M. C., Belew, A. K., Sisay, M., Daniel Baffa, L., Gashaw, M., Yiheyis Abriham, Z., & Mengistu, B. (2024). Spatial variations and determinants of timely completion of vaccination in Ethiopia using further analysis of EDHS 2019 data: Spatial and multilevel analysis. *PLoS One*, **19**(4): e0301409.

Ali, H. A., Hartner, A. M., Echeverria-Londono, S., Roth, J., Li, X., Abbas, K., & Gaythorpe, K. A. (2022). Vaccine equity in low and middle-income countries: a systematic review and meta-analysis. *International Journal for Equity in Health*, **21**(1): 82.

De Cantuária Tauil, M., Sato, A. P. S., & Waldman, E. A. (2016). Factors associated with incomplete or delayed vaccination across countries: a systematic review. *Vaccine*, **34**(24): 2635-2643.

Eze, C. E., Igwama, G. T., Nwankwo, E. I., & Victor, E. (2024). Predictive modeling for healthcare needs in the aging US population: A conceptual exploration. *Global Journal of Research in Science and Technology*, **2**(02): 094-102.

Forshaw, J., Gerver, S. M., Gill, M., Cooper, E., Manikam, L., & Ward, H. (2017). The global effect of maternal education on complete childhood vaccination: a systematic review and meta-analysis. *BioMed Central Infectious Diseases*, **17**(1): 1-16.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **374**(2065): 20150202.

Mbengue, M. A. S., Sarr, M., Faye, A., Badiane, O., Camara, F. B. N., Mboup, S., & Dieye, T. N. (2017). Determinants of complete immunization among senegalese children aged 12–23 months: evidence from the demographic and health survey. *BioMed Central Public Health*, **17**(1): 1-9.

Naraei, P., Abhari, A., & Sadeghian, A. (2016). Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data. In *2016 Future Technologies Conference (FTC)* (pp. 848-852).

Nusinovici, S., Tham, Y. C., Yan, M. Y. C., Ting, D. S. W., Li, J., Sabanayagam, C., & Cheng, C. Y. (2020). Logistic regression was as good as machine learning for predicting major chronic diseases. *Journal of Clinical Epidemiology*, **122**(1): 56-69.

Ogero, M., Orwa, J., Odhiambo, R., Agoi, F., Lusambili, A., Obure, J., & Ngugi, A. (2022). Pentavalent vaccination in Kenya: coverage and geographical accessibility to health facilities using *data from a community demographic and health surveillance system in Kilifi County. BioMed Central Public Health*, **22**(1): 826.

Orueta, J. F., Nuño-Solinis, R., Mateos, M., Vergara, I., Grandes, G., & Esnaola, S. (2013). Predictive risk modelling in the Spanish population: a cross-sectional study. *BioMed Central Health Services Research*, **13**(1): 1-9.

Prabhod, K. J., & Gadhiraju, A. (2019). Reinforcement Learning in Healthcare: Optimizing Treatment Strategies and Patient Management. *Distributed Learning and Broad Applications in Scientific Research*, **5**(1): 67-104.

Satya Sahisnu, J., Natalia, F., Vincenttius Ferdinand, F., Sudirman, S., & Seong Ko, C. (2020). Vaccine prediction system using ARIMA method. *International Conference on Innovative Computing Express Letters Part B: Applications*, **11**(6): 567-575. doi:10.24507/icicelb.11.06.567

Segar, M. W., Hall, J. L., Jhund, P. S., Powell-Wiley, T. M., Morris, A. A., Kao, D., & Pandey, A. (2022). Machine learning–based models incorporating social determinants of health vs traditional models for predicting in-hospital mortality in patients with heart failure. *Journal of the American Medical Association Cardiology*, **7**(8): 844-854.

Shiferie, F., Gebremedhin, S., Andargie, G., Tsegaye, D. A., Alemayehu, W. A., Mekuria, L. A., & Fenta, T. G. (2023). Vaccination dropout and wealth related inequality among children aged 12–35 months in remote and underserved settings of Ethiopia: a cross-sectional evaluation survey. *Frontiers in Pediatrics*, **11**(1): 1280746.

Smith, P. J., Kennedy, A. M., Wooten, K., Gust, D. A., & Pickering, L. K. (2006). Association between health care providers' influence on parents who have concerns about vaccine safety and vaccination coverage. *Pediatrics*, **118**(5): e1287-e1292.

Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, **10**(5): e1379.

Taye, M. M. (2023). Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers*, **12**(5): 91.

Toma, M., & Wei, O. C. (2023). Predictive modeling in medicine. *Encyclopedia*, **3**(2): 590-601.

Wang, Y., & Li, H. (2022). Analysis on the Balance of Health Care Resource Allocation Based on Improved Machine Learning. In Sinha, A. & Kumar, S. (Eds.), *IoT and Big Data Technologies for Health Care* (1st ed., pp. 102-116). New York, NY: Springer Nature.

World Health Organization. (2019). Surveillance for Vaccine-Preventable Diseases. WHO Technical Report Series. Geneva: WHO Press.

Yadav, V. (2022). Machine Learning for Predicting Healthcare Policy Outcomes: Utilizing Machine Learning to Forecast the Outcomes of Proposed Healthcare Policies on Population Health and Economic Indicators. *Journal of Artificial Intelligence & Cloud Computing*, **1**(2): 2-10.

Yaya, A. (2017). *Reducing Under-five Childhood Mortality using IMCI/e-IMCI: Implementation Approaches in Nigeria.* https://doi.org/10.17615/hveb-7b20