Information Systems Student Research Journal

Thesis Digest DOI: ISSN:

Predicting Postpartum Contraceptive Utilization and Identifying Its Determinants Among Reproductive Age Women in Selected Sub-Saharan African Countries, Using a Machine Learning Algorithm

Eliyas Addisu Taye

MPH in Health Informatics from University of Gondar, Ethiopia

eliyasaddisu12@gmail.com

Aim of the Study

This study aimed to predict postpartum contraceptive (PPC) use and identify its determinants among women of reproductive age in 26 Sub-Saharan African countries. Driven by the urgent need to reduce unplanned pregnancies and maternal mortality, the research sought to answer the research question: What factors most significantly influence postpartum contraceptive use, and how can machine learning (ML) models enhance prediction accuracy to inform policy?

Area of Concern

PPC use remains critically low in Sub-Saharan Africa, contributing to high rates of maternal and infant mortality (Akinyemi et al., 2015; WHO, 2024; Yemane et al., 2021). Despite global efforts, 26% of women in the region use PPC, with barriers to use including limited healthcare access, low literacy, and socioeconomic disparities. Prior studies have relied on traditional statistical methods, often overlooking complex variable interactions. This study bridges existing knowledge gaps by applying ML algorithms to uncover complex patterns and systematically analyze their key determinants, including sociodemographic, socioeconomic, and maternal health-related factors (Agrawal et al., 1993; Dhar, 2013; Sarker, 2021).

Investigative Approach

Using demographic health survey data (2016-2023) from 62,030 women, the study applied seven supervised ML algorithms (e.g., XGBoost and LightGBM) to predict PPC use. Data preprocessing included SMOTE oversampling to address so-called class imbalance, a common issue where one class is significantly underrepresented compared to another, leading to biased model predictions. By generating synthetic samples of the minority class, SMOTE helped improve the model's ability to learn meaningful patterns from both classes. SHAP analysis was employed to enhance interpretability by quantifying the contribution of each feature to the model's predictions, ensuring transparency in decision-making. Additionally, association rule mining was conducted to identify underlying relationships between key determinants of PPC use, revealing co-occurring factors influencing contraceptive uptake. A design science approach guided the integration of theoretical knowledge and practical solutions, with model deployment via a Flask API for real-time predictions (Grus, 2019; Holmström, 2009).

Results

The XGBoost model emerged as the most effective predictor of PPC use in Sub-Saharan Africa, achieving 86% accuracy and an Area Under the Curve of 93% on balanced data. SHAP analysis identified the key determinants. Giving birth at a health facility, maternal literacy, antenatal care (ANC) visits, higher education, and media exposure were strong positive predictors. In contrast, home delivery, illiteracy, lack of media exposure, and lower socioeconomic status negatively influenced PPC use. Association rule mining further revealed significant patterns, showing that women who delivered at home without ANC visits had an 83% likelihood of PPC non-use. In comparison, illiterate women with limited healthcare access had an 81% probability of not using PPC. Additionally, socioeconomic disparities, such as poverty and rural residence, compounded the risk, with a 76% confidence for PPC non-use among low-income households.

Contribution

This study advances maternal health research by demonstrating ML's superiority over traditional methods in predicting PPC use. It confirms the critical role of healthcare access and education while introducing novel insights into media exposure and household dynamics. Policymakers can use these findings to design targeted interventions, such as enhancing rural healthcare infrastructures and literacy programs, to meet Sustainable Development Goal 3.1.

More Information

More information about the author:

- LinkedIn profile: https://www.linkedin.com/in/eliyas-a-taye-42a17b276/
- CDHi website: https://cdhi.uog.edu.et/eliyas-addisu/

Access the entire thesis: https://submission.springernature.com/submission-details/80b704f9-db1c-4587-8540-e04d83a0f152

Declaration of AI

No generative AI tools were employed for data analysis, interpreting results, and writing text during the original thesis research and manuscript preparation. However, for the creation of this thesis digest, DeepSeek was used to assist in summarizing key findings and structuring the content.

References

Agrawal, R., Imieliński, T., & Swami, A. (1993). *Mining association rules between sets of items in large databases.* Paper presented at the Proceedings of the 1993 ACM SIGMOD international conference on Management of data.

Akinyemi, A., Adedini, S., Hounton, S., Akinlo, A., Adedeji, O., Adonri, O., . . . Barros, A. J. (2015). Contraceptive use and distribution of high-risk births in Nigeria: a sub-national analysis. *Glob Health Action, 8*, 29745.

Dhar, V. (2013). Data science and prediction. *Communications of the ACM, 56*(12), 64-73.

Grus, J. (2019). *Data science from scratch: first principles with python*: O'Reilly Media.

Holmström, J., Ketokivi, M., & Hameri, A. P. (2009). Bridging practice and theory: A design science approach. *Decision sciences, 40*(1), 65-87.

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science, 2*(3), 160.

WHO. (26 April 2024). Maternal mortality. Retrieved from https://www.who.int/news-room/fact-sheets/detail/maternal-mortality

Yemane, T. T., Bogale, G. G., Egata, G., & Tefera, T. K. (2021). Postpartum Family Planning Use and Its Determinants among Women of the Reproductive Age Group in Low-Income Countries of Sub-Saharan Africa: A Systematic Review and Meta-Analysis. *Int J Reprod Med, 2021*, 5580490.