ISSN(O): 2320-5407 | ISSN(P): 3107-4928

International Journal of Advanced Research

Publisher's Name: Jana Publication and Research LLP

www.journalijar.com

REVIEWER'S REPORT

Manuscript No.: IJAR-58780

Title: Growth-supporting food consumption in 6-23 months children in northern Benin and minimum diet diversity prediction with supervised machine learning approaches

Recommendation:

Accept as it is .....

Accept after minor revision.....

Accept after major revision .....x.....

Do not accept (*Reasons below*) .....

Rating	Excel.	Good	Fair	Poor
Originality		x		
Techn. Quality		x		
Clarity		x		
Significance			x	

Reviewer's ID: JPR-142

Detailed Reviewer's Report

This manuscript presents a longitudinal study of complementary feeding among children initially aged 6–13 months in Banikoara, northern Benin, followed through five assessment rounds from November 2023 to July 2024. The authors assess Egg and Flesh Food (EFF), Fruit and Vegetable (FV), and Minimum Dietary Diversity (MDD), using multivariable Generalized Estimating Equations and seven supervised machine-learning algorithms for MDD prediction. The final analytical sample comprises 149 children.

The topic is relevant and potentially publishable. The use of subject-level splitting to reduce information leakage from repeated observations is a positive methodological feature. The manuscript also appropriately acknowledges its single-commune setting, caregiver-reported dietary data, omitted determinants, and limited nine-month follow-up.

However, major revision is necessary, particularly concerning the statistical tables, machine-learning methodology, interpretation of prediction performance, small subgroup sizes, and consistency between the stated conclusions and the actual predictive task.

Comments

1. **Major inconsistency in Table 3 requires correction**
2. Table 3 contains an apparent structural/data-labeling error. Under **"Social cultural groups"**, the categories include Bariba, Gando, Gourmantche, Zerma and **Founougo**, whereas Founougo is a living district rather than a sociocultural group. Subsequently, the living-district section includes Banikoara, Kokiborou, Somperekou and Soroko but does not list Founougo.

This is a substantive problem rather than merely a typographical issue because it creates uncertainty about the actual model specification and estimated ORs.

ISSN(O): 2320-5407 | ISSN(P): 3107-4928
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

REVIEWER'S REPORT

Required revision:-The authors should carefully verify the original statistical output and correct Table 3. They should confirm:

- all sociocultural categories;
- all five living districts;
- the reference categories;
- corresponding β coefficients;
- ORs;
- 95% CIs; and
- p-values.

The corrected table should be cross-checked against the underlying statistical analysis.

2. The extremely large FV odds ratio for Soroko requires statistical investigation

Table 3 reports an OR of approximately **4.7×10^{10}** for FV consumption in Soroko, with no meaningful confidence interval reported.

Given that only **10 children (6.7%)** came from Soroko, this estimate strongly suggests possible **complete or quasi-complete separation, sparse-data bias, or instability of the GEE estimate**.

Required revision:-The authors should:

1. examine the underlying cell frequencies for FV consumption in Soroko;
2. report whether separation occurred;
3. explain why the CI could not be estimated;
4. consider an appropriate alternative estimation strategy if the conventional GEE estimate is unstable; and
5. avoid interpreting this OR as a reliable magnitude of association unless adequately supported.

This issue is particularly important because the manuscript currently describes the Soroko result as evidence of markedly elevated FV consumption.

3. The machine-learning analysis needs substantially greater methodological detail

The manuscript states that seven algorithms were evaluated using 104 children for training and 45 for prediction, with 10-fold cross-validation in the training set.

However, important details are missing:

- How were categorical variables encoded?
- How was the outcome class distribution handled?
- Were missing data present?
- How were hyperparameters selected?
- Was feature scaling performed for kNN and SVM?
- Was preprocessing performed separately within each cross-validation fold?
- Was feature selection performed?
- Was class imbalance assessed?
- Were the same hyperparameter settings used across algorithms?

ISSN(O): 2320-5407 | ISSN(P): 3107-4928
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

- Was cross-validation stratified?
- Were repeated cross-validation procedures used?
- Were confidence intervals for AUC or other performance metrics calculated?

Required revision:-The ML methods should be rewritten in sufficient detail to permit reproducibility.

4. The training/test sample size is small for the number of algorithms evaluated

Only **149 children** were included, with 104 assigned to training and 45 to testing. The manuscript then evaluates seven different algorithms.

This creates a substantial risk of unstable model performance and optimistic conclusions. The authors should explicitly acknowledge that the reported AUC values are based on a relatively small independent test sample.

Required revision:-The authors should provide:

- number of observations per outcome class in training and test datasets;
- class distribution;
- justification for the 104/45 split;
- uncertainty estimates around performance measures where possible; and
- a stronger limitation statement concerning model instability and lack of external validation.

5. The distinction between longitudinal statistical inference and prediction should be clearer

The manuscript combines GEE-based association analysis with machine-learning prediction. These are different analytical objectives.

The GEE analysis estimates associations between predictors and EFF/FV/MDD, whereas the ML component attempts to predict MDD. The manuscript sometimes moves from association to prediction without clearly distinguishing the two.

Required revision:-The authors should explicitly separate:

1. **inferential analysis** — determinants/associations using GEE; and
2. **predictive analysis** — performance of ML models.

The Discussion should avoid presenting predictive variables as causal determinants.

6. The ML prediction target and clinical/public-health interpretation need correction

The manuscript states that the models could support "two-stage screening of at-risk children," while the outcome being predicted is **achievement of MDD**. The conclusion similarly proposes identifying children "likely to meet MDD."

ISSN(O): 2320-5407 | ISSN(P): 3107-4928
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

REVIEWER'S REPORT

This is conceptually confusing. If the objective is to identify children **at risk of inadequate dietary diversity**, the positive class should presumably represent **failure to achieve MDD**, rather than achievement of MDD. Conversely, if the positive class is MDD achievement, the application should be framed as identifying children likely to achieve adequate dietary diversity or potential positive-deviance households.

Required revision:-Clearly define:

- positive class;
- negative class;
- screening objective;
- whether the model is intended to identify children with inadequate MDD or children achieving MDD.

.7. The reported specificity is relatively low and should be discussed

The best-performing models have sensitivity around **0.89**, but specificity is only approximately **0.42**. Table 5 shows similar performance for Tree, EGB and Naive Bayes.

Therefore, although sensitivity is relatively high, the models generate a substantial number of false-positive classifications.

The manuscript acknowledges the false-positive burden but concludes that the operational trade-off is acceptable.

Required revision:-This conclusion should be more cautious. The authors should explain the consequences of the low specificity and ideally provide the confusion matrix or counts of true positives, false positives, true negatives and false negatives for the final test set.

8. The claim that Decision Tree, EGB and Naive Bayes are clearly "best" is overstated

Table 5 shows:

- Tree: AUC 0.74, accuracy 0.74, F1 0.83, sensitivity 0.89;
- EGB: AUC 0.75, accuracy 0.74, F1 0.83, sensitivity 0.89;
- Naive Bayes: AUC 0.75, accuracy 0.74, F1 0.83, sensitivity 0.89.

These differences are very small. Without statistical comparison or uncertainty intervals, it is difficult to establish that one algorithm truly outperforms another.

Required revision:-The authors should state that these algorithms demonstrated **similar performance**, rather than implying a definitive superiority of one model. Confidence intervals or resampling-based uncertainty estimates would strengthen the comparison.

9. The use of "moderate" AUC should be retained, but claims of practical readiness should be tempered

The reported AUC values of approximately 0.74–0.75 represent moderate discrimination. The manuscript itself describes the models as having moderate performance.

ISSN(O): 2320-5407 | ISSN(P): 3107-4928
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

REVIEWER'S REPORT

However, the conclusion suggests that community monitoring programmes could adopt a two-stage model-based approach.

This is premature without:

- external validation;
- prospective validation;
- calibration assessment;
- evaluation across different seasons;
- validation in other communities; and
- assessment of real-world consequences.

Required revision:-The proposed community application should be presented as a **potential future application**, not as a ready-to-deploy intervention.

10. External validation is absent

The model was developed and tested within the same study population. There is no external dataset from another commune, region, or country.

Because the authors explicitly propose implementation in community nutrition programmes, external validation is particularly important.

Required revision:-The authors should explicitly state that the models are preliminary and require external and prospective validation before implementation.

11. Temporal structure needs further consideration

The study contains repeated measurements at five time points: November 2023, January 2024, March 2024, May 2024 and July 2024.

The GEE approach appropriately accounts for repeated observations using an AR(1) working correlation structure. However, the ML analysis should clarify how repeated measurements were handled after the subject-level split.

The manuscript states that splitting was performed at subject level to prevent information leakage, which is commendable.

Nevertheless, the authors should clarify whether:

- all five observations per child were retained in the training/test partitions;
- children with incomplete observations existed;
- temporal ordering was preserved;
- prediction represents prediction for a new child or prediction of a future time point in an existing child.

12. The age definition requires clarification

ISSN(O): 2320-5407 | ISSN(P): 3107-4928
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

REVIEWER'S REPORT

The title describes children aged **6–23 months**, while the inclusion criterion at baseline was children aged **6–13 months**.

This may be scientifically appropriate because children age during follow-up, but the manuscript should make this explicit.

Suggested clarification:-State that children were **6–13 months at enrollment and were followed longitudinally as they aged through the 6–23-month IYCF period**.

13. The exclusion of participants who did not complete all interviews may introduce selection bias

The exclusion criteria included children whose parents did not participate in all planned interviews.

This complete-case approach may preferentially retain families who are more engaged, accessible, or stable.

Required revision:-Report whether any participants were lost after enrollment and provide reasons for non-completion. Discuss the potential selection bias introduced by excluding incomplete follow-up.