frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 07 January 2025
DOI 10.3389/fpubh.2024.1369041

Multimodal machine learning for analysing multifactorial causes of disease—The case of childhood overweight and obesity in Mexico

Rosario Silva Sepulveda[1] and Magnus Boman[1,2]*

[1]Karolinska Institutet, Department of Medicine Solna, Division of Clinical Epidemiology, Stockholm, Sweden, [2]MedTechLabs, BioClinicum, Karolinska University Hospital, Stockholm, Sweden

OPEN ACCESS

EDITED BY
Maria Oana Sasaran,
"George Emil Palade" University of Medicine, Pharmacy, Sciences and Technology of Târgu Mureș, Romania

REVIEWED BY
Marco Bilucaglia,
IULM University, Italy
Martin Romero-Martínez,
National Institute of Public Health, Mexico
Álvaro Torres,
University of Granada, Spain

*CORRESPONDENCE
Magnus Boman
✉ magnus.boman@ki.se

RECEIVED 11 January 2024
ACCEPTED 16 December 2024
PUBLISHED 07 January 2025

CITATION
Silva Sepulveda R and Boman M (2025) Multimodal machine learning for analysing multifactorial causes of disease—The case of childhood overweight and obesity in Mexico. *Front. Public Health* 12:1369041. doi: 10.3389/fpubh.2024.1369041

Background: Mexico has one of the highest global incidences of paediatric overweight and obesity. Public health interventions have shown only moderate success, possibly from relying on knowledge extracted using limited types of statistical data analysis methods.

Purpose: To explore if multimodal machine learning can enhance identifying predictive features from obesogenic environments and investigating complex disease or social patterns, using the Mexican National Health and Nutrition Survey.

Methods: We grouped features into five data modalities corresponding to paediatric population exogenous factors, in two multimodal machine learning pipelines, against a unimodal early fusion baseline. The supervised pipeline employed four methods: Linear classifier with Elastic Net regularisation, k-Nearest Neighbour, Decision Tree, and Random Forest. The unsupervised pipeline used traditional methods with k-Means and hierarchical clustering, with the optimal number of clusters calculated to be $k = 2$.

Results: The decision tree classifier in the supervised early fusion approach produced the best quantitative results. The top five most important features for classifying child or adolescent health were measures of an adult in the household, selected at random: BMI, obesity diagnosis, being single, seeking care at private healthcare, and having paid TV in the home. Unsupervised learning approaches varied in the optimal number of clusters but agreed on the importance of home environment features when analysing inter-cluster patterns. Main findings from this study differed from previous studies using only traditional statistical methods on the same database. Notably, the BMI of a randomised adult within the household emerged as the most important feature, rather than maternal BMI, as reported in previous literature where unwanted cultural bias went undetected.

Conclusion: Our general conclusion is that multimodal machine learning is a promising approach for comprehensively analysing obesogenic environments. The modalities allowed for a multimodal approach designed to critically analyse data signal strength and reveal sources of unwanted bias. In particular, it may aid in developing more effective public health policies to address the ongoing paediatric obesity epidemic in Mexico.

KEYWORDS

supervised machine learning, unsupervised machine learning, multimodal machine learning, bias, paediatric obesity, obesogenic environment, Mexico

1 Introduction

Paediatric obesity is a systemic, chronic, inflammatory, and recurrent disease characterised by an abnormal or excessive accumulation of body fat in children up until 19 years (1). At least 50% of children and adolescents affected by obesity continue being affected into adulthood (2), compromising their future health, quality of life, and educational or professional performance (3–6). Paediatric obesity prevalence has more than quadrupled globally from 1975 (4%) to 2016 (18%), regardless of strategies and efforts to control it (7). Paediatric obesity's increasing trend has been found in countries irrespective to their gross income classification (8, 9), with low to middle income countries particularly vulnerable due to the economic burden of obesity which represents on average 8.4% of the annual health spending even for wealthy Organisation for Economic Co-Operation and Development (OECD) countries (5). Mexico is an emergent economy with one of the highest incidences of paediatric overweight and obesity in the world (10). Thus, Mexico's case will be the geographic and demographic focus of this study, using the Mexican National Survey of Health and Nutrition (*ENSANUT: Encuesta Nacional de Salud y Nutrición*) as data source.

ENSANUT is a probabilistic national survey that works as a regular strategy to assess the status on health and nutrition of the Mexican population and as basis for Public Health policy-making (11). Public health policies and strategies designed from ENSANUT can be observed in Figure 1, and include establishing guidelines for sale and distribution of foods and beverages in elementary schools (2010); restrictions on the television advertisement of high caloric density foods and sweetened beverages for children audiences (2014); adding a 10% tax to junk food to reduce consumption (2014); regulations on front package labelling (2015); updating clinical guidelines for improving timely diagnosis and treatment (2018); and other diverse campaigns or action plans to promote behaviour change (11, 12). Since 1999, the number of children and adolescents in Mexico suffering from overweight or obesity increased by at least 10%, representing 38.2% of children and 43.8% of adolescents by 2020, which could indicate that the implementation of the policies and strategies designed from ENSANUT have failed to control the epidemic (8, 9). Since, to our knowledge, the findings from which these public health strategies and policies were based (11) involved only traditional mathematical and statistical analyses, it is plausible that other methods such as those of machine learning (ML) could provide additional knowledge on the predictors of paediatric overweight and obesity in Mexico, thanks to its capacity to analyse complex and non-linear patterns (13).

Efforts to understand more comprehensively the aetiology of the disease using ML go back to 1995 (14), and a systematic review on ML models to predict childhood and adolescent obesity conducted in 2020 reported that when comparing to the results reached through statistical methods, ML models gave better prediction performance in internal and external validations (15). Studies like the following three suggest that building prediction model using ML, with environmental and phenotypical factors taken into account, is fruitful. Rehkopt et al. (16) used diet and physical activity alongside parental risk factors on the psychological, social, and health aspects. Wiechman et al. (17) made use of demographics, caregiver feeding style and practices, home environment characteristics, and familial characteristics such as social and spousal support. Kim et al. (18) fitted features coming from behavioural patterns of teenagers like family wealth, smartphone use, amount of pocket money, academic performance, and quality of sleep.

An obesogenic environment is characterised by surroundings, opportunities, or conditions of life, external to the individual, which promote overweight or obesity (19). The paediatric population's environment is almost completely regulated by the adults who are part of their life and their dwelling (10). Hence, exploring exogenous factors and studying them as a whole is key to understanding the aetiology of paediatric obesity more comprehensively (10, 19, 20). These type of studies could also promote creating public health strategies to address obesogenic environmental factors, which are currently rare (11, 19, 20). Out of the different ML methods, multimodal ML models can be particularly useful for studying exogenous factors of paediatric obesity as it lets us process and fuse information from multiple data modalities. Modalities can convey different sources of data: applications in health care frequently use this approach to predict a specific health outcome (21). A supervised multimodal ML pipeline includes early and late fusion approaches (22). Early, or feature, fusion involves concatenating the data from the different modalities after the pre-processing stage and before training the model. Late, or decision, fusion involves selecting the features, training and validating the models individually, and then using a series of techniques to combine the predictions of each modality's best performing methods (23).

Our aim is to explore if multimodal ML can enhance, against a unimodal baseline (see Section 2.6.2), identifying predictive features from obesogenic environments and investigating complex disease or social patterns using the latest ENSANUT available to us (ENSANUT-18) as data source. The study includes two multimodal ML pipelines with supervised and unsupervised approaches, respectively (24). Our approach could help identify clusters that may not be targeted yet by public health initiatives or could aid identifying the most predictive features from the household that promote the development of the disease. Furthermore, the knowledge extracted could be used to design early interventions in the form of clinical guidelines (15). This makes it possible to identify the most predictive features from the child or adolescent's environment and to investigate comprehensively explanatory patterns for the development of paediatric overweight and obesity in Mexico.

2 Materials and methods

2.1 Study outline

As shown in Figure 2, the study consisted of three phases: a preparation phase and two separate multimodal ML training phases (also called *pipelines*). The preparation phase had the goal of producing the datasets to train the models in the consequent phases. The multimodal ML pipelines adhered to Carbonell et al. (24) for its comprehensiveness and scalability

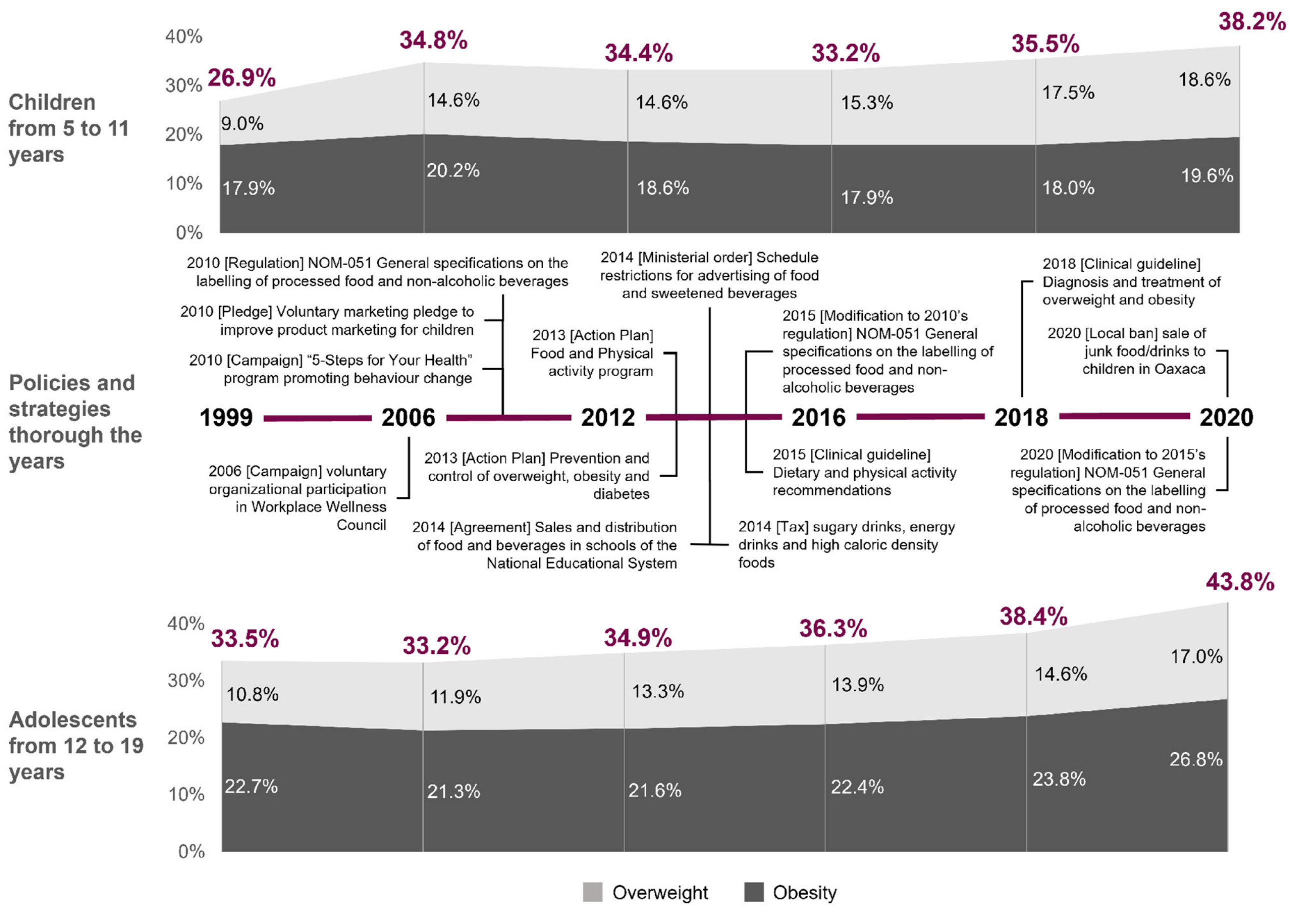


FIGURE 1
Trends of overweight and obesity in children and adolescents from 1999 to 2020 in Mexico and the policies and strategies implemented over the years. Personal elaboration of earlier reported material (11, 20, 44).

(25), and for its critical bias detection features. Additionally, for replication purposes, we have made available a GitHub repository containing all scripts used to prepare the datasets and run the ML pipelines: https://github.com/rosarioss/multimodal-paediatric-obesity.

2.2 Scoping review

The first step was to perform a scoping review to facilitate an informed selection of the features to include in the study. The review focused on identifying studies where the ENSANUT database was used to explore feature correlations on topic. Journal articles that used ENSANUT data as their primary source, published in English or Spanish, at any time, were included. Six databases were used: Medline, Web of Science (WoS), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Scientific Electronic Library Online (SciELO), Global Index Medicus from the World Health Organisation (GIM/WHO), and Virtual Health Library from Pan American Health Organisation (VHL/PAHO). The databases were selected for their extensive variety of health informatics-related topics, or because of their inclusion of Mexican or Latin American journals (SciELO and VHL/PAHO). Detailed criteria can be found in the Supplement, Section 1, including a PRISMA chart in Supplementary Figure S1.

2.3 Feature selection and modality definition

Following the scoping review, the features in each ENSANUT questionnaire were examined to identify those that should be "Definitely included" (271 of them), "Definitely excluded," and "Neutral." Since the focus of the study was analysing obesogenic environments, questionnaires that focused directly on the child or adolescent were excluded from the feature selection process. The questionnaires included referred either to a person that could have a direct influence on daily habits or on the home environment *per se*. This could either be represented by the adult who usually oversee the preparation of food or an adult randomly selected from the household by the ENSANUT-18 interviewers. The "Definitely included" features were organised in five modalities: home environment, household expenses and income, health information, biometrics, and knowledge on nutritional information (summarised in Figure 3 and detailed in

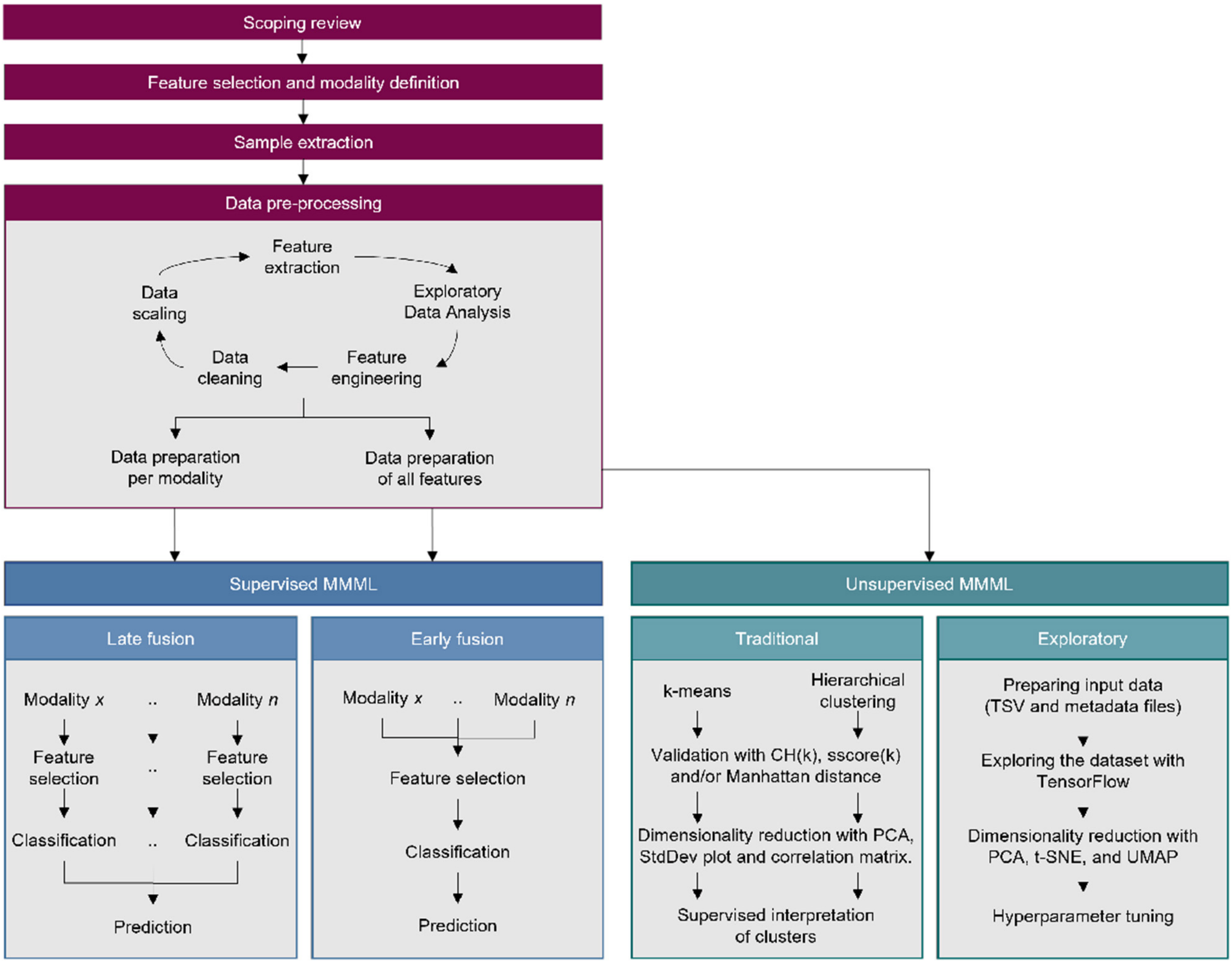

Scoping review
Feature selection and modality definition
Sample extraction
Data pre-processing
Feature extraction
Exploratory Data Analysis
Feature engineering
Data cleaning
Data scaling
Data preparation per modality
Data preparation of all features
Supervised MMML
Unsupervised MMML
Late fusion
Modality x
..
Modality n
Feature selection
Feature selection
Classification
Classification
Prediction
Early fusion
Modality x
Modality n
Feature selection
Classification
Prediction
Traditional
k-means
Hierarchical clustering
Validation with CH(k), sscore(k) and/or Manhattan distance
Dimensionality reduction with PCA, StdDev plot and correlation matrix.
Supervised interpretation of clusters
Exploratory
Preparing input data (TSV and metadata files)
Exploring the dataset with TensorFlow
Dimensionality reduction with PCA, t-SNE, and UMAP
Hyperparameter tuning

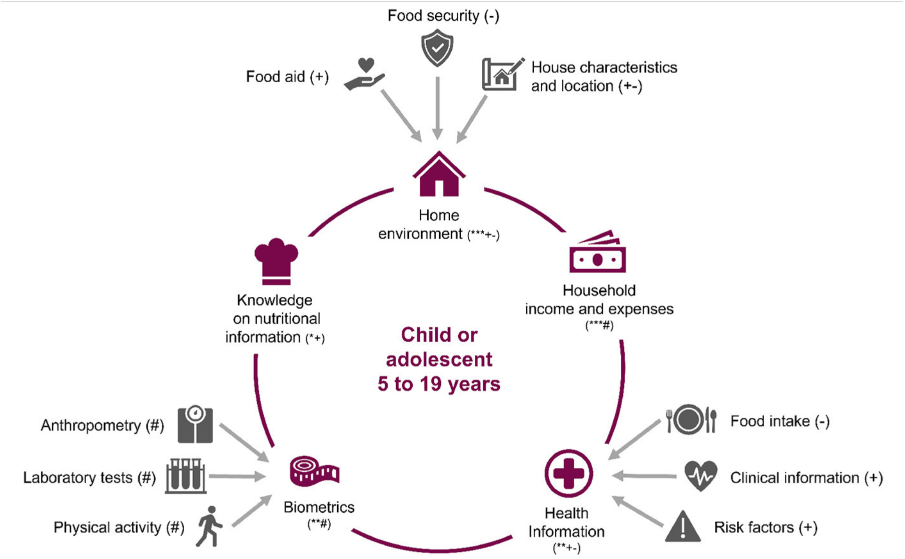
FIGURE 2
Complete study outline which includes a data preparation phase preceding the supervised and unsupervised machine learning pipelines.

Supplementary Figure S3). Each modality is further explained in Section 2.5 to avoid repetition. Criteria for categorising the features summarises as follows:

- Definitely excluded:
  - features unrelated to the environment of the child or adolescent,
  - features that focused directly on the child or adolescent,
  - multicollinear features,
  - features containing more than 35% NaN values;
- Definitely included:
  - features identified in the scoping review,
  - Other environmental features;
- Neutral:
  - Redundant features in view of another “Definitely included” feature,
  - Remainder features with no natural inclusion into the other two categories.

2.4 Sample extraction

The sample population dataset for this study was built from an anthropometric data questionnaire applied to children and adolescents. The study population includes female or male individuals from 5 to 19 years of age (60–228 months) at the time of ENSANUT-18 data collection. The first step in the data selection process was to exclude pregnant individuals, due to potentially altered BMI, and those who were missing anthropometric data. Afterwards, metadata required to identify the study population (such as multi-level unique keys) and features to calculate BMI-for-age (Weight, height, age in months, gender) were extracted and prepared into the sample population dataset. The International BMI-for-age curve formula (26), defined by the World Health Organisation (WHO) to diagnose paediatric overweight and obesity (27), was applied to calculate the BMI-for-age and interpret it. WHO’s BMI-for-age measures for children and adolescents



Source of data:
* Questionnaire answered by the adult in charge of all or most of the cooking in the household.
** Questionnaire answered by or tests taken to a random adult in the household.
*** Questionnaire answered by the chief of the household.

Type of data:
+ Categorical.
# Numerical.
- Binary.

FIGURE 3
The five modalities defined for the study, with their corresponding sources and types.

from 5 to 19 years consist of sex-specific age curves computed by percentiles or z-scores. The interpretation of the z-score cut-offs, based on the number of standard deviations from zero, were used as labels for the model training (28–30):

- Underweight: less or equal to −1SD,
- Normal weight: more than −1SD and less than +1SD,
- Overweight: more or equal to +1SD and less than +2SD,
- Obesity: more or equal to +2SD.

After the initial data cleaning of the sample population dataset, a total $n = 10{,}301$ observations remained. However, the actual dataset size to train each of the models depended on the modal approach. For the late fusion pipeline, each modality had different sample and dimension sizes. For the early fusion pipeline and the unsupervised pipelines, three dataset with equal sample size for different dimension size were required: a numerical dataset, a categorical dataset, and one containing both. The first two were used to process separately when using Mean Squared Error as a dimensionality reduction technique during the unsupervised traditional approach. The latter was used for the supervised early fusion and for both unsupervised approaches. See next section for details on the final dataset size for each modality.

2.5 Data pre-processing

Data pre-processing included a cyclical process of feature extraction, exploratory data analysis, feature engineering, data cleaning, and data scaling until the dataset was considered ready for data preparation. Feature extraction consisted of importing the “Definitely included” features from the corresponding questionnaire and concatenating them to form one dataset for each modality. The dataset then underwent a general exploratory data analysis that looked for general data patterns such as distribution of each feature and identifying NaN values, with the purpose of identifying Feature engineering requirements. Feature engineering involved different strategies, such as computation of numerical values or creation of dummy variables for categorical features, depending of the modality (see Sections 2.5.1 through 2.5.5 for details). Data cleaning included an additional step to identify features that should be “Definitely excluded” based on insufficient data quality. This was defined as having a Pearson correlation of ≥ 0.70 (multicollinearity) or containing more than 35% NaN values. If a feature was discovered to have a multicollinear relationship with another feature, one of the features were dropped. The criteria for keeping features considered the amount of information one questionnaire item provided as compared to the other one, see Sections 2.5.1 through 2.5.5 for details. Ambiguous answers such as “I do not know” or “I do not remember” were processed as NaN values as well. The dropping criteria for NaN values followed these rules: if the feature had more than 35% NaN values, the whole feature (column) was dropped. If it had <35% NaN values, then only the rows that had NaN values were dropped. The exception to dropping rows was if the dataset forced n below the acceptable level given by the Cochran equation for infinite values (31):

$$n = \frac{z^2 \times \rho(1-\rho)}{\epsilon^n} \tag{1}$$

where: z = z-score, ϵ = margin of error, n = population size, and ρ = population proportion.

For the Mexican population of ∼32.5 million 5–19 year olds, with a margin of error of 2%, a confidence interval of 95% and an overweight- or obesity-affected population proportion of 42.3%, the minimum acceptable size of each dataset was n = 2,344.

Data preparation consisted once more on dropping every observation that had at least one NaN value (imputation was avoided to prevent generating bias), assigning each observation a random group number from 1 to 5 in preparation for using Leave-one-group-out (LOGO) Cross-validation (CV), and performing feature scaling (for numerical features) or creating dummy features (for categorical features). The randomised groups are needed for LOGO-CV so that one of the groups can be excluded from the training pipeline and the model can be tested against and independent and unseen fold. This helps prevent overfitting and inflated performance estimates, even when there's correlation or dependency within the groups. In contrast, k-fold cross-validation randomly splits the dataset into k parts without regard for any inherent group structure, which may lead to data leakage when groups are present, since multiple records from the same subject may end up in both the training and test sets.

2.5.1 Modality 1: home environment

This modality included features about the environment where the child or adolescent was growing up. It comprised sociodemographic data, data about subscription to food aid programs, food safety situation (results of ELCSA), physical characteristics of the house (e.g., conditions in which food is prepared), as well as location (locality size and type). Feature engineering involved three main transformations, the most significant of which was the calculation and interpretation of the ELCSA into food safety, as well as mild, moderate, and severe insecurity in the home. Two features ("fuel type used to cook" and "stove type") were dropped due to multicollinearity with another feature ("has gas stove"); only the latter feature was kept. Since no dimensions had more than 35% missing values, only rows with NaN values were removed. Label distribution remained similar to the sample dataset, with suffering from overweight or obesity at 42.5%, as well as the representation of the sample demographic metadata after groups were created for LOGO-CV. After pre-processing, the data set included 39 dimensions and 9585 observations.

2.5.2 Modality 2: household expenses and income

This modality included features about the monthly income of the household and food or health expenses per category of product or service consumed. Feature engineering involved three main transformations, the most significant of which was the calculation of the proportion that a specific product or service represented of the total household expenses, measured in either food or health expense. Other aspects were analogous to those of Modality 1. Because Modality 2 contained only numerical features and the distribution of the label in its SD format was non-Gaussian, the feature scaling strategy involved standardising the features. After pre-processing, 21 dimensions and 7, 819 observations remained.

2.5.3 Modality 3: health information

‘ This modality included features about current health status, personal and familial health history, health risk factors (Table 1), and food intake patterns of a random adult living in the same household as the child or adolescent. The demographic features of the adults selected at random included showed that 62.3% were between 30 and 49 years old, predominantly female (57.6%), and predominantly the parent of the child or adolescent (68.7%). Demographic details are provided in Supplementary Figure S7, whereas feature engineering is detailed in Supplementary Figures S12, S13. The latter involved calculation of the kinship. Worth noting is that since the relationship registered in the database was that of a resident of the household towards the "household chief," a series of functions needed to be created to translate that into kinship towards the study population. Due to this situation, certain limitations when computing the kinship were identified, for instance if the adult selected at random was the son or daughter of the chief and the child or adolescent was the grandchild, it was not possible to discern if the adult was the parent or uncle/aunt. Nevertheless, kinships obtained provided interesting information presented in the Section 4.

The feature containing the number of working hours per week, which was originally a numeric feature, was converted to binary for two reasons. First, to evaluate if the person was working more or less than the law-regulated maximum of 40 h per week. Second, because through the experimental data analysis it was discovered that most answers were either 40 or 50, giving the data a semi-categorical aspect from the start. Twelve features regarding the question "Which healthcare institution are you affiliated to?" were dropped due to multicollinearity with their correspondent items on the question "Where do you go to receive healthcare?" (see list of features dropped in Supplementary Figures S12, S13). The latter features were kept since, by default, included the information on the healthcare institution affiliation. After pre-processing, the data set included 113 dimensions and 3, 383 observations.

TABLE 1 Documented risk factors for paediatric overweight and obesity, per category.

Category	Risk factors
Dietary patterns	Fast food consumption (4, 32, 52), sugary beverages consumption (2, 53), snack foods consumption, portion size, and availability and accessibility of unhealthy food (4, 32, 33, 52, 54)
Physical activity	Level of physical activity, sedentary behaviour, basal metabolic rate, time in front of screens (1, 4)
Genetics	Genetic predisposition (4), genetic conditions such as hyperphagia (55)
Family dynamics	Parenting style, parental physical exercise habits, parental feeding style, family mealtime structure (4, 34, 56), grandparent co-residence in the household (57–59)
Environment	School policies (4, 60), parents' work-related demands (42, 56), fast-food industry (2, 61), regulation of marketing strategies for unhealthy foods (4), safety of walking/cycling routes (4), store density (54)
Socio-cultural influence	Food-as-reward related beliefs (4), perception of being a normal stage of growth (1)

2.5.4 Modality 4: biometrics

This modality included the numerical features about the anthropometric measurements and laboratory results of the same randomly selected adult as in Modality 3. Feature engineering involved six main transformations (see Supplementary Figures S12, S13). The feature regarding "Total cholesterol level" was dropped due to multicollinearity with "HDL cholesterol" and "LDL cholesterol" levels. The feature on "Total cholesterol" was dropped since the type of cholesterol provided more information than the total cholesterol. Pre-processing produced 19 dimensions and 4, 419 observations.

2.5.5 Modality 5: knowledge on nutritional information

This modality included the results of a test about the nutritional information on the packaging. The responding adult was the one that oversaw most of the cooking in the household. The test included questions about usage of different types of nutritional information, labels or legends printed in commercial products, as well as questions on the supposed importance of different nutrients such as carbohydrates, proteins, salt, fats, trans-fats, total energy, etc. About the food intake patterns, previous studies had found that a pattern of healthy eating could not be defined based on the numerical self-reported portions in the survey (32–34). Reasons for this could be under- or over-reporting of food intake based on social pressure (8, 34), or simply the challenge of answering the survey, as it required to remember the portion sizes based in grammes or millilitres per food group (e.g., fruits) and food item (e.g., bananas) eaten during the past 7 days (8). Hence, to avoid possible sources of noise or bias, only the binary version of those questions, based on WHO's daily food intake recommendations, were included.

The demographic features of the adults in this modality shows that 65.1% were between 30 and 49 years old, predominantly female (58.7%), and predominantly the parent (70.5%). The largest challenge in the feature engineering was that questions about the correct identification of three products had to be dropped, due to two thirds of the values missing. The feature regarding the question "Do you use the nutritional information of a product to compare their nutritional value?" had to be dropped due to multicollinearity with the question "Does the nutritional information of a product affect your buying decision?". The latter was kept since our main interest was to know if the product reached the home environment of the children or adolescent. After the pre-processing, the dataset included 70 dimensions and 4, 399 observations.

2.6 Supervised methods

The supervised multimodal ML pipelines used in this study included early and late fusion approaches (Figure 2); Sections 2.6.1, 2.6.2 detail these, respectively. The pipelines display some differences to the classical approach to multimodal ML. Firstly, in the latter, each data modality corresponds to different sources that capture the same reality (e.g., video and audio from the same recording). The modalities defined in this study instead correspond to different environmental sources that could affect the study population's behaviour. Second, the classical approach typically uses data from a micro-level of social analysis (individual identity, motives, and cognition) to predict an individual's label in the model. Since the present study aims to study obesogenic environments, the data used corresponds to a meso-level of social analysis (organisations and groups, such as family members in the same dwelling) to predict an individual's label (35).

Both early and late fusion pipelines used the same four ML methods; which are supported by a review on ML for obesogenic environments among others (15). Decision Trees and Random Forest methods can both handle continuous features that have been scaled and categorical data encoded as binary features. The k-Nearest Neighbour (kNN) was appropriate because as a non-parametric algorithm it makes no assumptions on the distribution of the data. Finally, a Linear Classifier with Elastic Net Regularisation (Elastic Net) offered the calculation of a Logistic Regression Coefficient that could also be used to evaluate feature importance and its positive or negative predictive association (36). Logistic Regression coefficients were used to analyse predictive patterns of the features and compare their predictive importance when using early fusion (unimodal) and late fusion (multimodal) approaches.

For analysing the performance of the supervised models, several metrics were employed, including Area Under the Receiver Operating Characteristic (AUROC), Recall, and greater-than-chance analysis. AUROC was used to compare performance between classifiers and between the training and validation sets of the same model and represents the probability that if given a randomly chosen positive and negative sample, the model will rank the positive higher (36). Recall is one of the four common performance measurements calculated from a confusion matrix (37), measuring the ability to identify all the true positive predictions. Greater-than-chance analysis was used to compare the Recall value, representing the random probability of being a True Positive as 0.25.

The classifier hyperparameters were tuned using Exhaustive Grid Search (GridSearchCV) method from Scikit learn. This method tries every possible combination for a set of given hyperparameters, hence it is both comprehensive and effective in finding near-optimal hyperparameters. Details and the search grids can be found in the Section 5 in Supplementary methods and Supplementary Figures S8, S11. All methods were tested using a Jupyter Notebook running Python 3.9 and the Pandas, NumPy, scikit-learn, Pickle, Seaborn and Matplotlib libraries, and the code is publicly available. In particular, NumPy was used to run mathematical, statistical, logical, and random simulation operations that required using a multidimensional array object, for randomly selecting households and members of those households. Since the ENSANUT-18 database is openly available, this allows for reproduction and also for later databases to be tested in the future.

2.6.1 Multimodal data experiments

Multimodal data experiments exclusively involved the Supervised Late Fusion pipeline. Late fusion, or decision-level fusion, involves selecting the features, training, and validating the models individually and then using a series of techniques to combine the predictions of each modality's best performing classifiers. One of the benefits of late fusion is that it allows for having different size of observations and data types in each modality, while early fusion require the same number of observations before concatenating the data and building the models. The techniques, or rules, used to combine the predictions were the following:

- Maximum rule: get max value between classifiers,
- Sum rule: sum the scores from each modality and calculate the average,
- Product rule: get product value between classifiers,
- Weight criterion: get ratio of each modality and compute the weighted scores,
- Rule-based: choose classifier according to validation confusion matrix.

2.6.2 Unimodal data experiments

Unimodal data experiments involved both the Supervised Early Fusion pipeline and the Unsupervised pipelines. Early fusion, or data-level fusion, involves concatenating the data from the different modalities after the pre-processing stage and before training the model. In terms of training and validation, this can be considered a unimodal strategy since all the features are used together to train the model, as a single modality. In this case, all the features from Modality 1 through Modality 5 were included and feature engineering involved all of the transformations described above, in Sections 2.5.1, 2.5.2, 2.5.3, 2.5.4, 2.5.5.

The features were first concatenated and then processed together in terms of dropping columns or rows of missing values. The biggest difference was that the threshold for dropping columns with missing values had to be lowered to 30% because the representativeness of the size of *n* was compromised (see Section 2.5). Label distribution remained similar to the sample dataset, as well as the representation of the sample demographic metadata after random groups were created for LOGO-CV. Three different manipulations were required after this point, creating three separate datasets: one with only numerical features, another one with only categorical features, and the last one containing both. The first two were used to process separately when using Mean Squared Error as a dimensionality reduction technique during the unsupervised classical approach. The last was used for the supervised early fusion and for both unsupervised approaches. For the categorical features, dummy features were created, while the numerical features were normalised. After the pre-processing, the three datasets contained 2, 467 observations and varied in dimension size: the categorical consisted of 183, the numerical of 28, and the multimodal thus of 211 dimensions.

2.7 Unsupervised methods

This study's unsupervised ML pipeline includes classical clustering methods and visual examination using TensorFlow Embedding Projector, where the adjustable parameters were manually tuned until interesting patterns were found. We have made everything publicly available so that readers can visualise and interact with the data themselves. In this way, the dimensionality reduction algorithms produce separations in the cluster analyses (38). Traditional clustering methods included k-Means and hierarchical clustering. The former clusters data trying to separate samples in *n* centroids of equal variance, while the latter represents the hierarchy of clusters as a tree with the root containing all the samples and the leaves containing only one sample (36).

Dimensionality reduction was used to adjust the models in terms of pattern recognition. The optimal number of clusters was calculated using two techniques measuring cluster tightness and separation: the Calinski-Harabaz score (CH Index) and the silhouette score. The CH Index has reduced sensibility to monotonicity, varied cluster densities, subclusters, and skewed distributions (39). The silhouette score offers a more robust technique for handling noisy data than CH Index, since it considers intra-cluster distance and different-nearest-cluster distance (24, 40). Dimensionality reduction methods used were Principal Component Analysis (PCA), t-distributed Stochastic Neighbour Embedding (t-SNE), Mean Squared Error (MSE), and Uniform Manifold Approximation and Projection (UMAP). PCA aids in identifying the most meaningful patters among redundant features or noise, while t-SNE improves the visualisation of high-dimensional data. UMAP is a manifold technique that competes with t-SNE in visualisation quality but with a better run-time performance and arguably in mathematical grounding. Finally, MSE measures how close a regression line is to a set of data points: we dropped features that had variability close to zero ($\leq$0.14), to reduce noise in the data. As stated above, the numerical and categorical datasets were used to process separately when using Mean Squared Error as a dimensionality reduction technique during the unsupervised classical approach.

2.8 Ethics approval statement

All information in ENSANUT-18 has been pre-processed to remove sensitive data before made publicly available and does not include any open text answers that may give the identity of the participants away. Special consideration has been taken to limit the impact of unwanted bias, including handling multi-collinearity, avoiding imputation of missing values, and analysing the context in the original Spanish in which the question was phrased and what answers were offered, when selecting the features to include in the models.

3 Results

3.1 Scoping review

The 21 articles that met the inclusion criteria of the scoping review were analysed in four categories: dietary patterns (38.1%); physical, social, cultural, economic, or political environment (28.6%); epidemiology (23.8%), and family dynamics (9.5%). The article that analysed the highest number of features (45) out of all reviewed was in the first category. All the studies into dietary patterns found were based on the Food Frequency Intake questionnaire used in ENSANUT 2006, 2012, 2016, and 2018. Most

TABLE 2 Performance of the four machine learning methods used for the early fusion approach, during training and testing.

Classifier	Training AUROC	Testing AUROC
Elastic Net	0.6429 ± 0.0070	0.6021 ± 0.0293
k-NN	0.5662 ± 0.0048	0.5575 ± 0.0065
Decision Tree	**0.6335 ± 0.0093**	**0.6163 ± 0.0328**
Random Forest	0.9750 ± 0.0012	0.6088 ± 0.0234

The best quantitative score was obtained by the Decision Tree method, in bold. Random Forest was overfitting heavily.

studies concluded that the questionnaire overreported healthy foods and underreported unhealthy foods, especially so in the cases where the person answering the questionnaire was suffering from overweight or obesity. In the second category, relevant sociodemographic features included the region where the subject lived, type of locality (rural or urban), size of locality, Socio-Economic Strata (SES), enrolment in governmental food aid programs, and the result from level of Food Security obtained from the Latin American and the Caribbean Food Security Survey (*ELCSA: Encuesta Latinoamericana y del Caribe de Seguridad Alimentaria*). In the third category, descriptive statistics and logistic regression models were the most common analysis strategies. Only one study proposed a stratified approach of interpreting reduced or excess weight. Regrettably, it was more beneficial for identifying the causes of children with an underweight status than for those with excess weight. In the fourth category, we only found studies that analysed the relationship of a social factor with maternal characteristics (professional situation or breastfeeding practices). The full results of our scoping review can be found in structured form in Supplementary Figure S2.

3.2 Supervised multimodal pipelines

3.2.1 Early fusion

The best performing classifier in the early fusion approach was the Decision Tree (Table 2 and Figure 4). Out of the top five highest logistic regression coefficient feature importance scores, four were features of the randomly selected adult in the household (BMI, obesity diagnosis, being single, and seeking private care), while the fifth was related to the house environment (having paid TV). Note that ~85% of the single-person households in Mexico are single-mother ones.

3.2.2 Late fusion

The best performing classifier (Figure 5) varied per modality (Figure 6): Elastic Net for Modality 1, Decision Tree for Modality 3 and Random Forest for Modality 2, 4, and 5. In terms of the recall confusion matrix, Modality 1 had the highest performing predictor for the Overweight or Obese label (0.61), while Modality 3 had the highest value for the Normal label (0.78). Compared to the early fusion situation, more Logistic Regression coefficient feature importance scores with positive or negative values were found per modality, possibly indicating that some potentially relevant features are less hidden by noise when modelled in smaller batches. The highest metric value was the Maximum Criterion, which involved taking the best performing classifier of all, which was Modality 4 with a validation AUROC of 0.6340 ± 0.0152 and recall values of 0.68 for the normal label and 0.52 values for the Overweight or Obese label. In this way, a better performance than the early fusion approach was achieved.

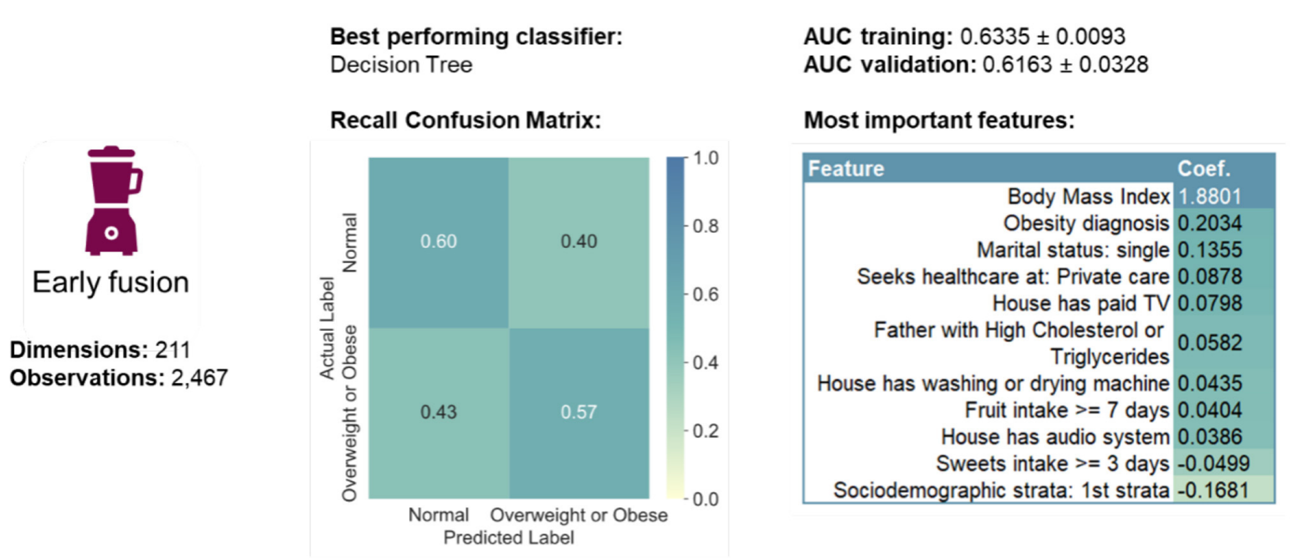
3.3 Unsupervised multimodal pipelines

The multimodal dataset with encoded and normalised features was evaluated before and after dimensionality reduction (Figure 7). Although variability between clusters before and after dimensionality reduction decreased, clear patterns could still not be identified. Following Carbonell et al. (24), a decision tree was used to interpret the clusters in terms of underlying features (Figure 8). Note that the first three nodes are composed predominantly of home environment features. Looking at the first decision node, the instances that had urban locality type smaller than 0.5 ended up in the second cluster. The second node's features were having Internet and being in the first socio-economic stratum, while the third mostly had features about food expenses (corn and dairy) and a feature indicating not using the nutritional information on the front-packaging labels of processed food as buying advice. Three different dimensionality reduction techniques were employed: PCA, UMAP and t-SNE. Even if no clear pattern emerged when colouring by label, colouring by the ternary normal/overweight/obese instead of the binary used to train the supervised models produced results (Figure 9). A pattern can be seen for the smaller localities (<2,500 inhabitants) where a considerably smaller number of children or adolescents suffering from overweight or obesity can be observed, a pattern that cannot be seen in any of the other locality sizes. Analogous results were obtained from UMAP and t-SNE (see Supplementary Figures S5, S6).

4 Discussion

4.1 Study findings

Our study found that integrally analysing the environment where the child or adolescent is growing up is essential to better understand the aetiology of paediatric obesity. Imagining a home environment built on some of the most important features identified, it is plausible to identify it as an obesogenic environment; Hobbs et al. (19) describe them as being characterised by surroundings, opportunities, or conditions of life, external to the individual, which promote overweight or obesity. Further studies should focus on adding features about behavioural patterns of the children and adolescents *per se*, analysing them in the context of obesogenic environments. Additionally, studies should investigate the daily habit patterns of adults that serve as behavioural examples for the paediatric population. Features to explore could include factors that overall contribute to having less time for building or maintaining healthy habits, such as exercise and leisure, the amount of time at work, the time spent commuting, the availability of spaces to exercise in, the safety situation in those spaces, and the economic and logistics accessibility of healthy foods. Finally, studies could



Dimensions: 211
Observations: 2,467

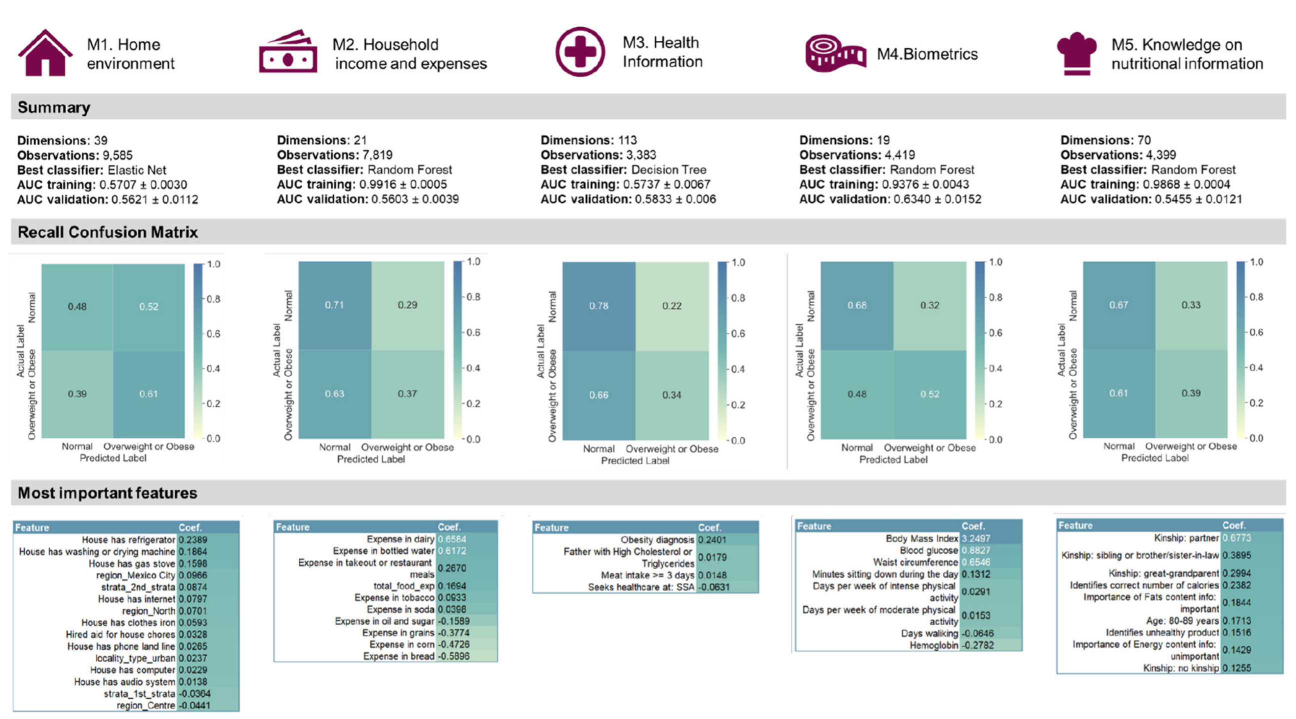
Best performing classifier: Decision Tree

AUC training: 0.6335 ± 0.0093
AUC validation: 0.6163 ± 0.0328

Most important features:

Feature	Coef.
Body Mass Index	1.8801
Obesity diagnosis	0.2034
Marital status: single	0.1355
Seeks healthcare at: Private care	0.0878
House has paid TV	0.0798
Father with High Cholesterol or Triglycerides	0.0582
House has washing or drying machine	0.0435
Fruit intake >= 7 days	0.0404
House has audio system	0.0386
Sweets intake >= 3 days	-0.0499
Sociodemographic strata: 1st strata	-0.1681

FIGURE 4
Summary of early fusion performance measures.



M1. Home environment
Dimensions: 39
Observations: 9,585
Best classifier: Elastic Net
AUC training: 0.5707 ± 0.0030
AUC validation: 0.5621 ± 0.0112

M2. Household income and expenses
Dimensions: 21
Observations: 7,819
Best classifier: Random Forest
AUC training: 0.9916 ± 0.0005
AUC validation: 0.5603 ± 0.0039

M3. Health Information
Dimensions: 113
Observations: 3,383
Best classifier: Decision Tree
AUC training: 0.5737 ± 0.0067
AUC validation: 0.5833 ± 0.006

M4.Biometrics
Dimensions: 19
Observations: 4,419
Best classifier: Random Forest
AUC training: 0.9376 ± 0.0043
AUC validation: 0.6340 ± 0.0152

M5. Knowledge on nutritional information
Dimensions: 70
Observations: 4,399
Best classifier: Random Forest
AUC training: 0.9868 ± 0.0004
AUC validation: 0.5455 ± 0.0121

Feature	Coef.
House has refrigerator	0.2389
House has washing or drying machine	0.1864
House has gas stove	0.1598
region_Mexico City	0.0966
strata_2nd_strata	0.0874
House has internet	0.0797
region_North	0.0701
House has clothes iron	0.0593
Hired aid for house chores	0.0328
House has phone land line	0.0265
locality_type_urban	0.0237
House has computer	0.0229
House has audio system	0.0138
strata_1st_strata	-0.0364
region_Centre	-0.0441

Feature	Coef.
Expense in dairy	0.6584
Expense in bottled water	0.6172
Expense in takeout or restaurant meals	0.2670
total_food_exp	0.1694
Expense in tobacco	0.0933
Expense in soda	0.0398
Expense in oil and sugar	-0.1589
Expense in grains	-0.3774
Expense in corn	-0.4726
Expense in bread	-0.5896

Feature	Coef.
Obesity diagnosis	0.2401
Father with High Cholesterol or Triglycerides	0.0179
Meat intake >= 3 days	0.0148
Seeks healthcare at: SSA	-0.0631

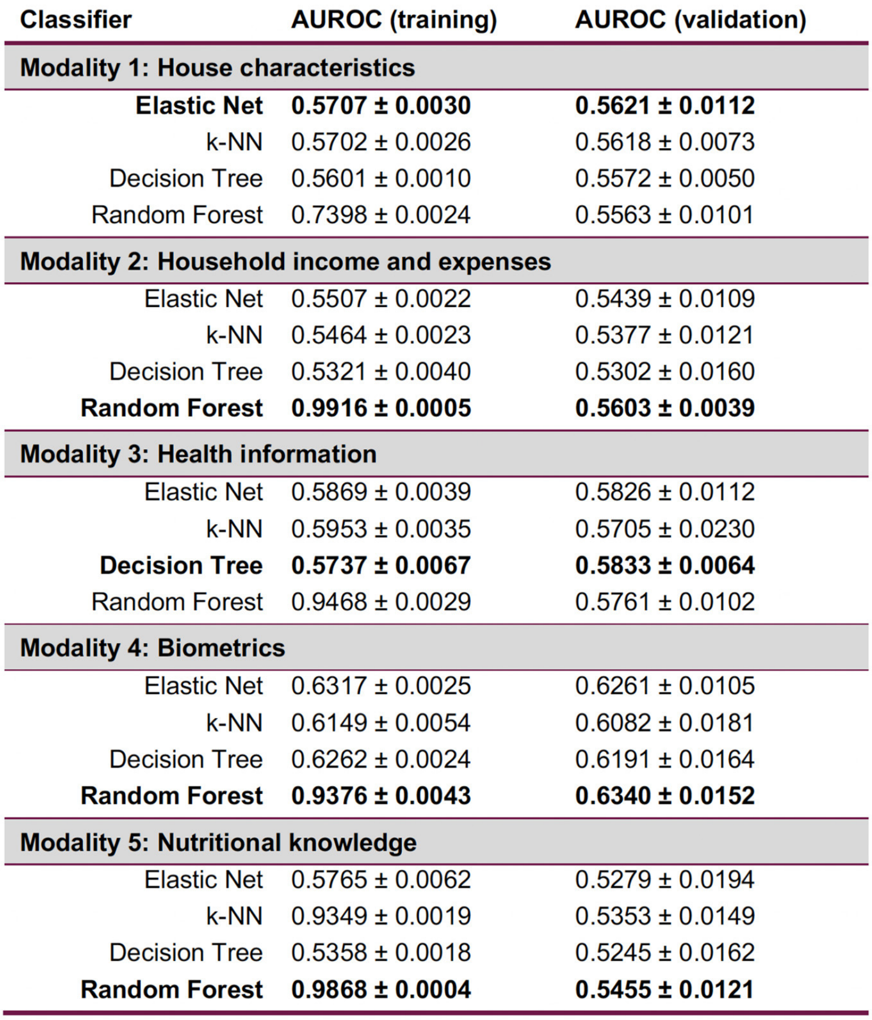
Feature	Coef.
Body Mass Index	3.2497
Blood glucose	0.8827
Waist circumference	0.6546
Minutes sitting down during the day	0.1312
Days per week of intense physical activity	0.0291
Days per week of moderate physical activity	0.0153
Days waliking	-0.0646
Hemoglobin	-0.2782

Feature	Coef.
Kinship: partner	0.6773
Kinship: sibling or brother/sister-in-law	0.3895
Kinship: great-grandparent	0.2994
Identifies correct number of calories	0.2382
Importance of Fats content info: important	0.1844
Age: 80-89 years	0.1713
Identifies unhealthy product	0.1516
Importance of Energy content info: unimportant	0.1429
Kinship: no kinship	0.1255

FIGURE 5
Summary of the quantitative late fusion approach and its feature importance results.

be designed with other algorithms, other hyperparameters than the one we used (see Supplementary Figure S8), or ML methods in mind. In particular, larger samples could make possible the use of novel deep learning algorithms (41) to explore patterns even further.

After data pre-processing, 183 categorical and 28 numerical features had been extracted. Datasets were then prepared for each supervised and unsupervised multimodal ML pipeline, but the best performing late fusion approach did not outdo the best early fusion by much, as measured by AUROC. Rather than the quantitative predictive performance in itself, however, the truly meaningful aspect of the comparison was the analysis the differences of feature importance using the logistic regression coefficient (Figure 5). In the home environment modality, it was concluded that having a refrigerator, washing or drying machine, gas stove, a Mexico City dwelling, and belonging to the second socio-economic stratum



Classifier	AUROC (training)	AUROC (validation)
Modality 1: House characteristics		
Elastic Net	**0.5707 ± 0.0030**	**0.5621 ± 0.0112**
k-NN	0.5702 ± 0.0026	0.5618 ± 0.0073
Decision Tree	0.5601 ± 0.0010	0.5572 ± 0.0050
Random Forest	0.7398 ± 0.0024	0.5563 ± 0.0101
Modality 2: Household income and expenses		
Elastic Net	0.5507 ± 0.0022	0.5439 ± 0.0109
k-NN	0.5464 ± 0.0023	0.5377 ± 0.0121
Decision Tree	0.5321 ± 0.0040	0.5302 ± 0.0160
Random Forest	**0.9916 ± 0.0005**	**0.5603 ± 0.0039**
Modality 3: Health information		
Elastic Net	0.5869 ± 0.0039	0.5826 ± 0.0112
k-NN	0.5953 ± 0.0035	0.5705 ± 0.0230
Decision Tree	**0.5737 ± 0.0067**	**0.5833 ± 0.0064**
Random Forest	0.9468 ± 0.0029	0.5761 ± 0.0102
Modality 4: Biometrics		
Elastic Net	0.6317 ± 0.0025	0.6261 ± 0.0105
k-NN	0.6149 ± 0.0054	0.6082 ± 0.0181
Decision Tree	0.6262 ± 0.0024	0.6191 ± 0.0164
Random Forest	**0.9376 ± 0.0043**	**0.6340 ± 0.0152**
Modality 5: Nutritional knowledge		
Elastic Net	0.5765 ± 0.0062	0.5279 ± 0.0194
k-NN	0.9349 ± 0.0019	0.5353 ± 0.0149
Decision Tree	0.5358 ± 0.0018	0.5245 ± 0.0162
Random Forest	**0.9868 ± 0.0004**	**0.5455 ± 0.0121**

FIGURE 6
Late fusion results per modality, with each best-performing validation ML method in bold. Random Forest is consistently overfitting. Expressions of worst-case time complexity are detailed in the Supplementary Section 4; see Supplementary Figure S9.

were the top five most important features. In the *household income and expenses* modality, we were able to identify the expenses in dairy, bottled water, takeout or restaurant meals, tobacco, as well as the total food expense mattered as the most important. The third and fourth modalities concerned the adult randomly selected in each household and the former pointed to having an obesity diagnosis, having a father with high cholesterol or triglycerides, and having a meat intake of more than three days per week. The latter pointed to BMI, blood glucose, waist circumference, the number of minutes sitting down during the day, and the days per week of intense physical activity. By comparison, early fusion identified a total of eleven features as important, six of which were also present in the late fusion approach. The top five most important were BMI, having an obesity diagnosis, being single, seeking health care at private care, and having paid TV in the house. Significantly, early fusion results are incomplete, most likely because the selected features' interplay has a low signal-to-noise ratio. Therefore, considering results from two supervised approaches was beneficial.

4.2 Comparative study

Our study corroborates previous studies in finding that the lower prevalence of children suffering from obesity and adolescents in the lowest socio-economic stratum, majorly located in rural areas with <2,500 inhabitants (42, 43). This prevalence has also been reported about children from households with high food insecurity, who on average weighed less than their peers from food-secure households. However, food insecurity and obesity/overweight did not appear to have a close correlation in our study. Another similar finding concerns the correlation of the child or adolescent's BMI with that of an adult in their immediate environment (15, 44). However, previous studies that also used ENSANUT data (43, 45) exclusively studied the correlation between maternal BMI classified as obesity, and children or adolescents with obesity. By contrast, our study found that any adult selected in the household having either a high BMI or the diagnosis of obesity were the strongest predictor in the population. Furthermore, the only kinship-related features appeared in Modality 5, and the most

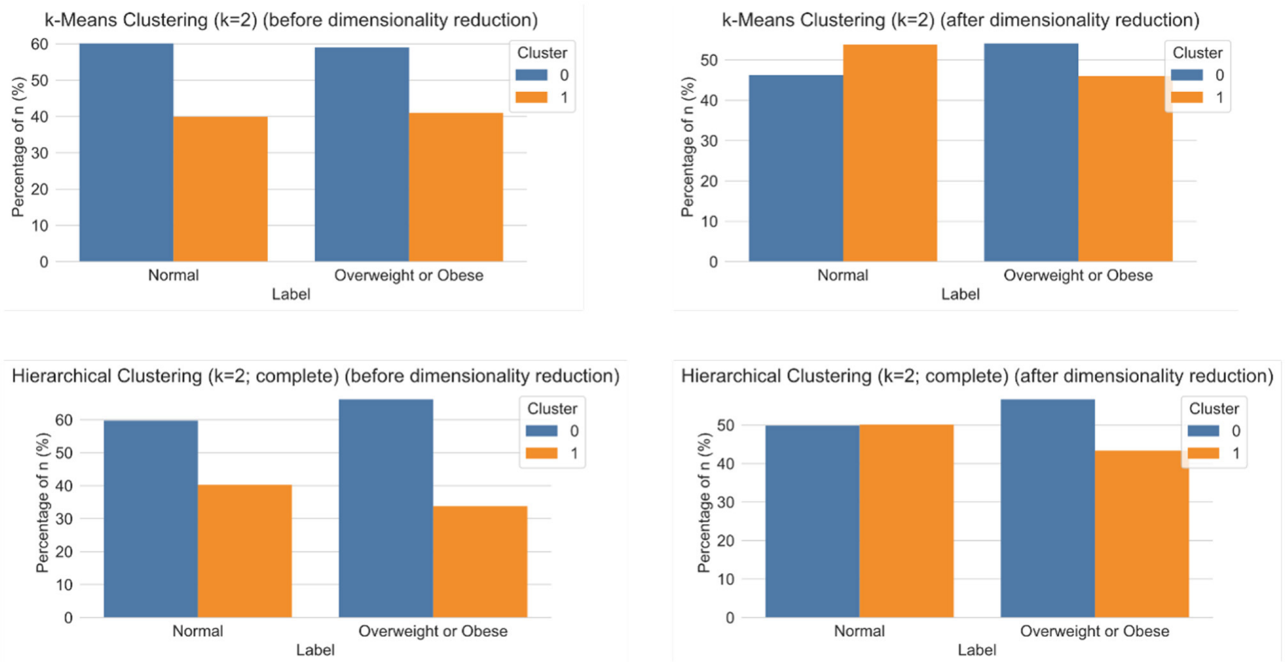


FIGURE 7
Results of clustering analyses for $k = 2$ using k-Means **(top)** and hierarchical clustering with complete distance method **(bottom)**, before **(left)** and after **(right)** dimensionality reduction.

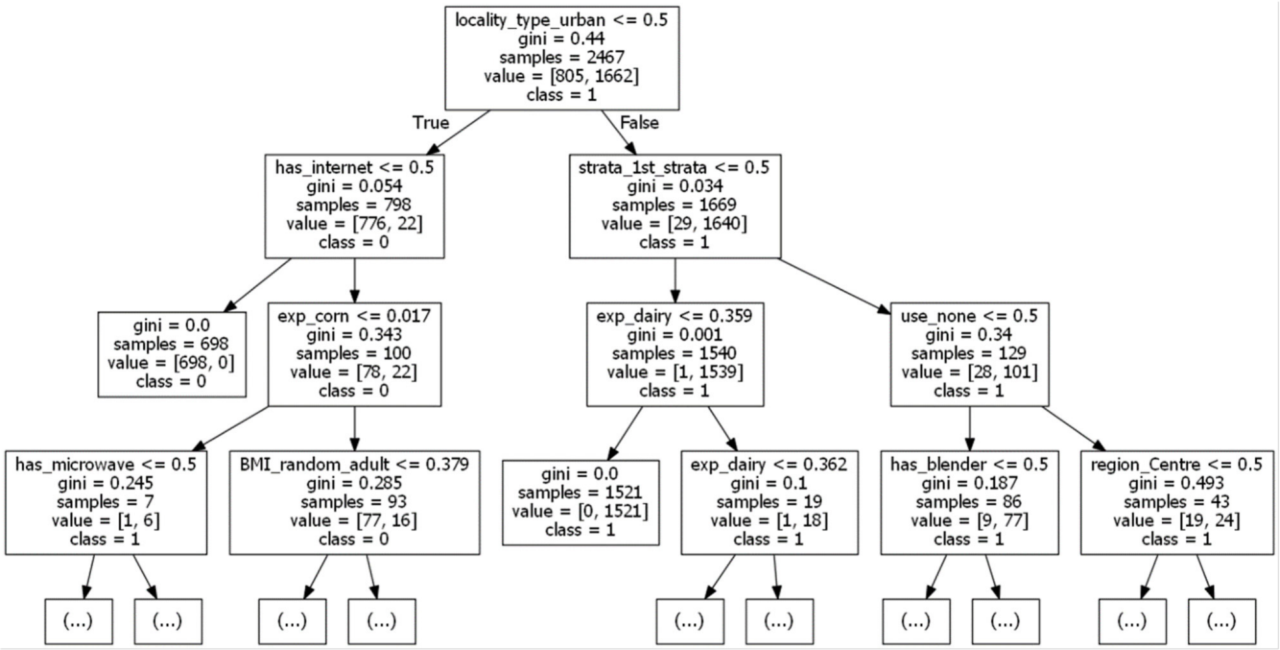


FIGURE 8
A relevant fragment of the decision tree used to interpret the clustering ($k = 2$ using k-Means, before dimensionality reduction).

important feature concerned a "partner" relationship with the adult whose nutritional knowledge was evaluated. This finding echoes studies that found a positive correlation between overweight or obesity and female adolescents living with a partner (46). It is worth repeating that our study did not find the sex of the adult a relevant feature.

The multimodal ML analysis may help reduce cultural or social bias on the selection of specific features. We noticed a

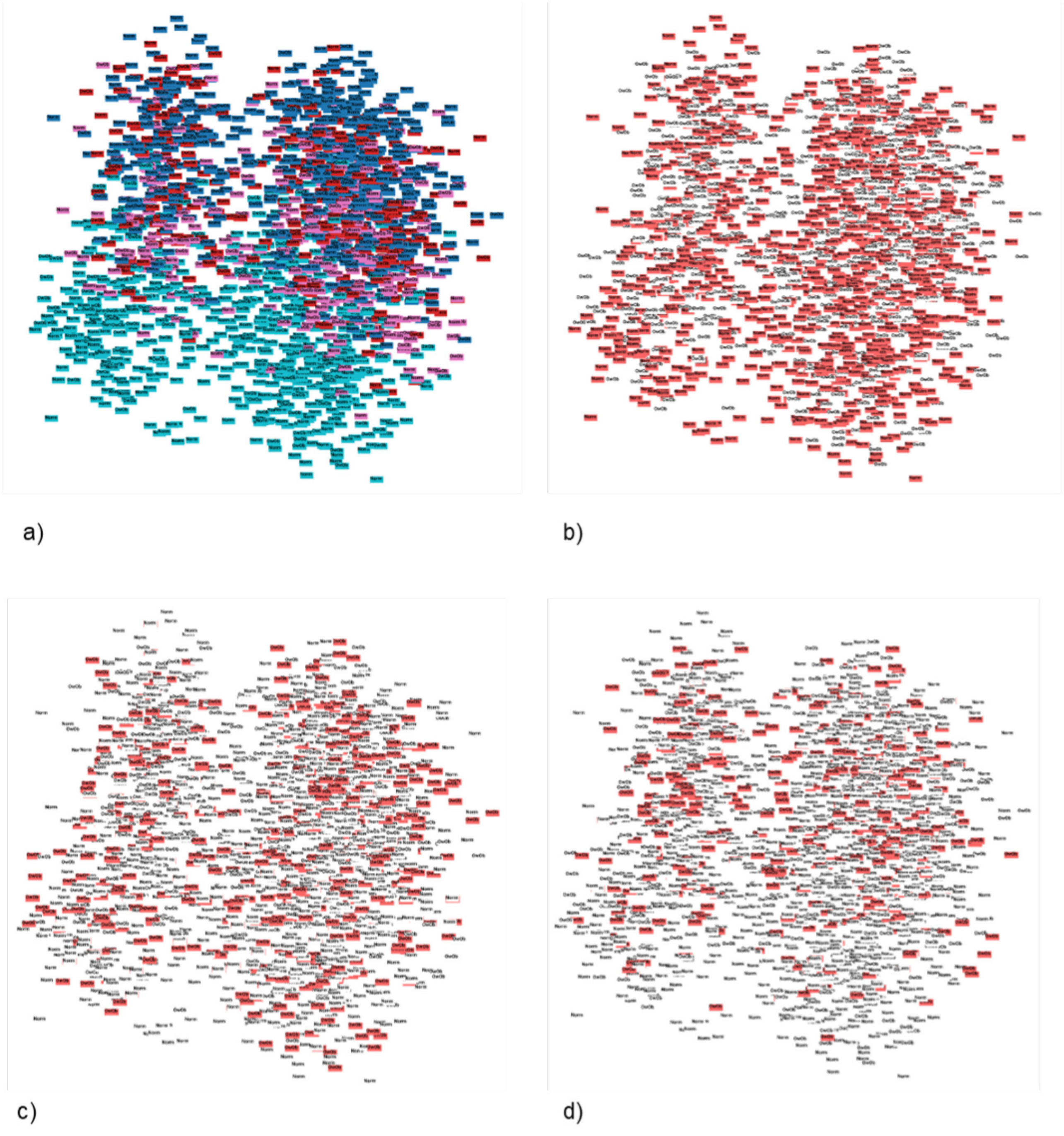


FIGURE 9
Patterns found using Principal Component Analysis (PCA) in 2D, constraining the search by the three-layered label and colouring by locality size. **(A)** PCA 2D label coloured by locality_size. **(B)** PCA 2D label coloured by locality_size with normal three-layered label marked. **(C)** PCA 2D label coloured by locality_size with overweight three-layered label marked. **(D)** PCA 2D label coloured by locality_size with obese three-layered label marked.

tendency towards studying almost exclusively maternal influence on paediatric obesity. In fact, a search in PubMed using the query "maternal" AND "paediatric obesity" yielded 1, 410 results, while "paternal" AND "paediatric obesity" yielded only 124 with all of them, without exception, used the term "paternal" to refer to both sexes and studied the relationship to risk factors of both parents. By contrast, studies focusing on obesogenic environments that built ML models for paediatric obesity prediction had a higher variety of features and included both sexes as equally relevant features, leaving the model to indicate importance.

Previous studies (32–34), while suggesting different numbers and types of dietary pattern, concurred that no pattern could be classified as healthy, since vegetables and fibre intake did not reach the minimum recommended portion, as stated by the

WHO. Similarly, our study could not find a relevant pattern between the label and the consumption of food categorised as unhealthy. Some possible explanations of this phenomenon could be that self-reported surveys are affected by social pressure towards underreporting unhealthy and overreporting healthy food, or that the person being interviewed is unaware of their own eating habits. The most worrisome situation would be if both overreporting of healthy foods and underreporting of unhealthy foods were happening simultaneously.

Within the top ten most important features about the household, expenses in *takeout or restaurant meals* and in *soda* were identified. At the same time, other than a reduction of prevalence of children and adolescents suffering from obesity in the lowest socio-economic stratum, both overweight and obesity were undistinguishably spread among the second through fourth strata. This indicates that the choice to buy healthy vs. unhealthy food may not necessarily be related to accessibility, but rather linked to buying decisions. A study that investigated the price trends of healthy and unhealthy foods from 2011 to 2018 in Mexico corroborates this (47). Our study found that the expense in dairy was the most important feature for Modality 2 and appeared within the first three decision nodes used during the unsupervised traditional approach to explain the clustering decisions. From 1990 to 2004, some unfavourable changes happened with consequences persisting until 2018, including an increase in the price of grains and milk and a decrease of the price for sweet bread, cookies, and potato chips.

Comparing the study results to other studies that used ML to analyse paediatric obesity was a difficult endeavour. Only one study could be found that was built with data from Mexico (48). It used 16 anthropometric features to predict paediatric obesity in children from 6 to 13 years old from an indigenous community in Guerrero. Even though the model was considered to have good performance for identifying overweight and obesity, the sample consisted of only 221 children, and methods not recommended for small samples were used, with neither feature importance nor overfitting evaluated. Therefore, their conclusions are hard to compare with the present study. Hammond et al. performed a similar study in terms of using features from EHRs and publicly available data to build a model for predicting obesity (49). However, they did not study the Mexican population, their study population were under five years old, and only statistical feature associations, rather than feature importance of the ML models were presented. Neither studies that used multimodal ML to study paediatric obesity, nor used a combination of supervised and unsupervised methods were identified.

4.3 Implications for public health policy

One of the ongoing public health initiatives aims to promote long-term behavioural changes, focusing on improving the delivery of nutritional content information with new labels, to influence healthier buying decisions (11). A study in 2018 showed that most consumers did not use the new labels, however, because they had difficulties interpreting them (50). This finding concurs with our study. We showed that not using the nutritional information on the front of packaging to make a buying decision were within the first three decision nodes used during the unsupervised traditional approach to explain clustering decisions.

We have demonstrated that using multimodal ML methods could be worthy of consideration for developing public health policies, especially since the prevalence in Mexico has been increasing in the past decades regardless of the policies and strategies being implemented to control it, such as the 10% tax on sugar-sweetened beverages. Researchers concluded that the tax imposed is very low compared to the 66.9% for tobacco, which was effective in reducing the smoking behaviour of adolescents (51). The efficiency of using ML methods requires data in quality and quantity, which can be challenging to accomplish through sporadic efforts. Consequently, evolving the national health monitoring strategy towards a continuing monitoring health information system could help attain both enough data and data of sufficient quality. Furthermore, it could include multimodal data that could affect individuals seeking to develop and maintain healthy patterns. For instance, it could include health data reported by healthcare professionals, results of laboratory or diagnostic tests, information about working situation, social demographics, and other types of data relevant to individuals (cf. Supplementary Figure S7). Such a system could allow for continuous studies into obesogenic environments in Mexico.

In conclusion, our recommendations to change or improve the paediatric obesity situation include to transform the current periodic survey-based strategy to a continuous monitoring health information system, analyse other risk factors contributing to having less time for building or maintaining healthy habits, and explore the design of policies that may help families have healthy daily habits, such as investing in active public transportation. Our contribution to data-driven health informatics is that we present an alternative to analysing national information from a standardised survey, and that we seem to reach less biassed and more thoroughly analysed conclusions than previously reported from work without learning approaches. Finally, our work shows that practical use of multimodal ML for public health policy making is feasible and potentially leads to improving population health.

Three categories of limitations of our study must be noted, these are dataset-, sample-, and ML-related. First, self-reported information always comes with challenges with respect the credibility and quality of the data. The underlying study itself suffered from missing observations. We also have noted risks of over- as well as underreporting. Because the majority of articles selected in our scoping used traditional statistical methods for their feature selection, there is a risk of unwanted bias in our predictive features. To detect such bias, one may use fairness metrics to cheque whether the model's predictions are equitable across different stratifications of groups or features. An attractive bias removal strategy is testing with adversarial datasets to challenge model robustness. Secondly, the sample-related limitations include that obesity is so widespread in Mexico that it may bias the results to interpret them as causality. We also do not always differentiate between overweight and obesity. Some differences in factors that affect the development of paediatric obesity could be related to the age group of the subject. Thirdly, the distribution of the sample used was complex to determine, before we reached the conclusion that it was a mixed Gaussian. Principle component analysis assumes a linear relationship and re-expresses the data as linear combinations, hence it can skew non-linear relations.

Analogously, the inertia parameter k-Means algorithm assumes that the clusters have a convex form and responds weirdly to elongated clusters.

These limitations notwithstanding, multimodal ML provided a comprehensive approach to analysing obesogenic environments. Our approach thus effectively presents an alternative approach to analysing national information from a standardised survey, which would seem to risk reaching biassed and less thorough conclusions.

Data availability statement

Publicly available datasets were analysed in this study. This data can be found here: ENSANUT: Encuesta Nacional de Salud y Nutrición (https://ensanut.insp.mx/encuestas/ensanut2018/index.php).

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

RS: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. MB: Conceptualization, Methodology, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2024.1369041/full#supplementary-material

References

1. Arriaga-Cázares H, Vergara-Castañeda A, García-Saldaña L, Higareda-Rojas B. *Guía de Evidencias y Recomendaciones: Guía de Práctica Clínica: Prevención y diagnóstico de sobrepeso y obesidad en niños y adolescents.* Ciudad de México: CENETEC (2018).

2. Torres-Álvarez R, Barrán-Zubaran R, Canto-Osorio F, Sánchez-Romero LM, Camacho-García-Formentí D, Popkin BM, et al. Body weight impact of the sugar-sweetened beverages tax in Mexican children: a modeling study. *Pediatr Obes.* (2020) 15:e12636. doi: 10.1111/ijpo.12636

3. World Health Organization. *Report of the Commission on Ending Childhood Obesity.* Geneva: WHO (2016).

4. Sahoo K, Sahoo B, Choudhury AK, Sofi NY, Kumar R, Bhadoria AS. Childhood obesity: causes and consequences. *J Fam Med Prim Care.* (2015) 4:187. doi: 10.4103/2249-4863.154628

5. Cecchini M, Vuik S. *The heavy burden of obesity: The Economics of Prevention.* Paris: OECD Health Policy Studies (2019). doi: 10.1787/3c6ec454-en

6. World Health Organization. *Global nutrition targets 2025: childhood overweight policy brief.* Geneva: WHO (2014).

7. World Health Organization. *Obesity and overweight.* (2021). Available at: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (accessed October 15, 2024).

8. Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Romero-Martínez M, Mojica-Cuevas J, Cuevas-Nasu L, et al. *Encuesta Nacional de Salud y Nutrición 2018-19: Resultados Nacionales.* (2020). Available at: https://ensanut.insp.mx/encuestas/ensanut2018/doctos/informes/ensanut_2018_informe_final.pdf

9. Shamah-Levy T, Romero-Martínez M, Barrientos-Gutiérrez T, Cuevas-Nasu L, Bautista-Arredondo S, Colchero M, et al. *Encuesta Nacional de Salud y Nutrición 2021 sobre COVID-19. Resultados nacionales.* Cuernavaca: Instituto Nacional de Salud Pública (2022). doi: 10.21149/12580

10. Turnbull B, Gordon SF, Martínez-Andrade GO, González-Unzaga M. Childhood obesity in Mexico: a critical analysis of the environmental factors, behaviours and discourses contributing to the epidemic. *Health Psychol Open.* (2019) 6:2055102919849406. doi: 10.1177/2055102919849406

11. Shamah-Levy T, Romero-Martínez M, Cuevas-Nasu L, Méndez Gómez-Humaran I, Antonio Avila-Arcos M, Rivera-Dommarco JA. The Mexican national health and nutrition survey as a basis for public policy planning: overweight and obesity. *Nutrients.* (2019) 11:1727. doi: 10.3390/nu11081727

12. World Obesity Federation Global Obesity Observatory. *Mexico: Policies, Interventions and Actions.* London (2022). Available at: https://data.worldobesity.org/country/mexico-139/#data_policies

13. DeGregory KW, Kuiper P, DeSilvio T, Pleuss JD, Miller R, Roginski JW, et al. A review of machine learning in obesity. *Obes Rev.* (2018) 19:668–85. doi: 10.1111/obr.12667

14. Novak B, Bigec M. Application of artificial neural networks for childhood obesity prediction. In: *Proceedings 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems.* Dunedin: IEEE (1995), p. 377–80. doi: 10.1109/ANNES.1995.499512

15. Colmenarejo G. Machine learning models to predict childhood and adolescent obesity: a review. *Nutrients.* (2020) 12:2466. doi: 10.3390/nu12082466

16. Rehkopf DH, Laraia BA, Segal M, Braithwaite D, Epel E. The relative importance of predictors of body mass index change, overweight and obesity in adolescent girls. *Int J Pediatr Obes.* (2011) 6:e233–42. doi: 10.3109/17477166.2010.545410

17. Wiechmann P, Lora K, Branscum P, Fu J. Identifying discriminative attributes to gain insights regarding child obesity in hispanic preschoolers using machine learning techniques. In: *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*. Boston, MA: IEEE (2017), p. 11–5. doi: 10.1109/ICTAI.2017.00014

18. Kim C, Costello FJ, Lee KC Li Y, Li C. Predicting factors affecting adolescent obesity using general Bayesian network and what-if analysis. *Int J Environ Res Public Health*. (2019) 16:4684. doi: 10.3390/ijerph16234684

19. Hobbs M, Radley D. Obesogenic environments and obesity: a comment on 'Are environmental area characteristics at birth associated with overweight and obesity in school-aged children? Findings from the SLOPE (Studying Lifecourse Obesity PrEdictors) population-based cohort in the south of England'. *BMC Med*. (2020) 18:1–3. doi: 10.1186/s12916-020-01538-5

20. Jia P. Obesogenic environment and childhood obesity. *Obes Rev*. (2021) 22:e13158. doi: 10.1111/obr.13158

21. Rajappa P, Eng KW, Bareja R, Bander ED, Yuan M, Dua A, et al. Utility of multimodality molecular profiling for pediatric patients with central nervous system tumors. *Neurooncol Adv*. (2022) 4:vdac031. doi: 10.1093/noajnl/vdac031

22. Lahat D, Adali T, Jutten C. Multimodal data fusion: an overview of methods, challenges, and prospects. *Proc IEEE*. (2015) 103:1449–77. doi: 10.1109/JPROC.2015.2460697

23. Torres-Martos Á, Bustos-Aibar M, Ramírez-Mena A, Cámara-Sánchez S, Anguita-Ruiz A, Alcalá R, et al. Omics data preprocessing for machine learning: a case study in childhood obesity. *Genes*. (2023) 14:248. doi: 10.3390/genes14020248

24. Carbonell MF, Boman M, Laukka P. Comparing supervised and unsupervised approaches to multimodal emotion recognition. *PeerJ Comput Sci*. (2021) 7:e804. doi: 10.7717/peerj-cs.804

25. Bark D, Boman M, Depreitere B, Wright D, Lewén A, Enblad P, et al. Refining outcome prediction after traumatic brain injury with machine learning algorithms. *Sci Rep*. (2024) 14:8036. doi: 10.1038/s41598-024-58527-4

26. World Health Organization. *Computation of centiles and z-scores for height-for-age, weight-for-age and BMI-for-age*. Geneva: WHO (2007).

27. de Castro JAC, de Lima TR, Silva DAS. Body composition estimation in children and adolescents by bioelectrical impedance analysis: a systematic review. *J Bodyw Mov Ther*. (2018) 22:134–46. doi: 10.1016/j.jbmt.2017.04.010

28. De Onis M. World Health Organization child growth standards. In: Karakochuk CD, Whitfield KC, Green TJ, Kraemer K, editors. *The Biology of the First 1,000 Days*. Boca Raton, FL: CRC Press (2017), p. 17–32. doi: 10.1201/9781315152950-2

29. World Health Organization. *BMI-for-age (5-19 years)*. (2022). Available at: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age (accessed October 15, 2024).

30. Onis Md, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ*. (2007) 85:660–7. doi: 10.2471/BLT.07.043497

31. Cochran WG. *Sampling Techniques*. Hoboken, NJ: John Wiley & Sons (1977).

32. Rodríguez-Ramírez S, Mundo-Rosas V, García-Guerra A, Shamah-Levy T. Dietary patterns are associated with overweight and obesity in Mexican school-age children. *Arch Latinoam Nutr*. (2011) 61:270–8.

33. Zárate-Ortiz AG, Melse-Boonstra A, Rodríguez-Ramírez S, Hernández-Cordero S, Feskens EJ. Dietary patterns and the double burden of malnutrition in Mexican adolescents: results from ENSANUT-2006. *Nutrients*. (2019) 11:2753. doi: 10.3390/nu11112753

34. García-Chávez CG, Rodríguez-Ramírez S, Rivera JA, Monterrubio-Flores E, Tucker KL. Sociodemographic factors are associated with dietary patterns in Mexican schoolchildren. *Public Health Nutr*. (2018) 21:702–10. doi: 10.1017/S1368980017003299

35. Serpa S, Ferreira CM. Micro, meso and macro levels of social analysis. *Int'l J Soc Sci Stud*. (2019) 7:120. doi: 10.11114/ijsss.v7i3.4223

36. Géron A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. Sebastopol, CA: O'Reilly Media Inc. (2017).

37. Fawcett T, Provost F. *Data Science for Business*. Sebastopol, CA: O'Reilly Media Inc. (2013).

38. McInnes L, Healy J, Melville J. Umap: uniform manifold approximation and projection for dimension reduction. *arXiv*. [Preprint] arXiv:1802.03426. (2018). doi: 10.48550/arXiv:1802.03426

39. Caliński T, Harabasz J. A dendrite method for cluster analysis. *Commun Stat-Theory Methods*. (1974) 3:1–27. doi: 10.1080/03610927408827101

40. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. (1987) 20:53–65. doi: 10.1016/0377-0427(87)90125-7

41. Benítez-Andrades JA, Alija-Pérez JM, Vidal ME, Pastor-Vargas R, García-Ordás MT. Traditional machine learning models and bidirectional encoder representations from transformer (BERT)-based automatic classification of tweets about eating disorders: algorithm development and validation study. *JMIR Med Inf*. (2022) 10:e34492. doi: 10.2196/34492

42. Shamah-Levy T, Mundo-Rosas V, Morales-Ruan C, Cuevas-Nasu L, Méndez-Gómez-Humarán I, Pérez-Escamilla R. Food insecurity and maternal-child nutritional status in Mexico: cross-sectional analysis of the National Health and Nutrition Survey 2012. *BMJ Open*. (2017) 7:e014371. doi: 10.1136/bmjopen-2016-014371

43. Shamah-Levy T, Campos-Nonato I, Cuevas-Nasu L. Hernández-Barrera L, del Carmen Morales-Rúan M, Rivera-Dommarco J, et al. Overweight and obesity in Mexican vulnerable population results of Ensanut 100k. *Salud Pública Méx*. (2020) 61:852–65. doi: 10.21149/10585

44. Aceves-Martins M, Llauradó E, Tarro L, Solà R, Giralt M. Obesity-promoting factors in Mexican children and adolescents: challenges and opportunities. *Glob Health Action*. (2016) 9:29625. doi: 10.3402/gha.v9.29625

45. Shamah-Levy T, Cuevas-Nasu L, Humarán IMG, Morales-Ruán C, Valenzuela-Bravo DG, Gaona-Pineda EB, et al. Prevalence and susceptibility to obesity in a national sample of school-age children and adolescents in Mexico. *Salud Pública México*. (2020) 62:725–33. doi: 10.21149/11552

46. Medina-Zacarías MC, Shamah-Levy T, Cuevas-Nasu L, Gómez-Humarán IM, Hernández-Cordero SL. Risk factors associated with overweight and obesity among female Mexican teenagers. *Salud Pública México*. (2020) 62:125–36. doi: 10.21149/10388

47. Batis C, Gatica-Domínguez G, Marrón-Ponce JA, Colchero MA, Rivera JA, Barquera S, et al. Price trends of healthy and less healthy foods and beverages in Mexico from 2011-2018. *J Acad Nutr Diet*. (2022) 122:309–19. doi: 10.1016/j.jand.2021.08.105

48. Ríos-Julián N, Alarcón-Paredes A, Alonso G, Hernandez-Rosales D, Guzman-Guzman I. Feasibility of a screening tool for obesity diagnosis in Mexican children from a vulnerable community of Me'Phaa ethnicity in the State of Guerrero, Mexico. In: *2017 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*. Tuxtla Gutierrez: IEEE (2017), p. 1–6. doi: 10.1109/GMEPE-PAHCE.2017.7972105

49. Hammond R, Athanasiadou R, Curado S, Aphinyanaphongs Y, Abrams C, Messito MJ, et al. Predicting childhood obesity using electronic health records and publicly available data. *PLoS ONE*. (2019) 14:e0215571. doi: 10.1371/journal.pone.0215571

50. Romero-Martínez M, Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Mojica-Cuevas J, Cuevas-Nasu L, et al. Encuesta Nacional de Salud y Nutrición (Ensanut 2018): metodología y perspectivas. *Salud Pública México*. (2019) 61:917–23. doi: 10.21149/11095

51. Illescas-Zárate D, Batis C, Ramírez-Silva I, Torres-Álvarez R, Rivera JA, Barrientos-Gutiérrez T. Potential impact of the nonessential energy-dense foods tax on the prevalence of overweight and obesity in children: a modeling study. *Front Public Health*. (2021) 8:591696. doi: 10.3389/fpubh.2020.591696

52. Arango-Angarita A, Shamah-Levy T, Rodríguez-Ramírez S. Dietary energy density is associated with body mass index-for-age in Mexican adolescents. *Matern Child Nutr*. (2019) 15:e12664. doi: 10.1111/mcn.12664

53. Afeiche M, Koyratty B, Wang D, Jacquier E, Lê KA. Intakes and sources of total and added sugars among 4 to 13-year-old children in China, Mexico and the United States. *Pediatr Obes*. (2018) 13:204–12. doi: 10.1111/ijpo.12234

54. Arango-Angarita A, Méndez-Gómez-Humarán I, Guerrero-López CM, Shamah-Levy T. Is store density associated with sugar-sweetened beverages consumption and overweight or obesity in Mexican adolescents? *Pediatr Obes*. (2022) 17:e12838. doi: 10.1111/ijpo.12838

55. Styne DM, Arslanian SA, Connor EL, Farooqi IS, Murad MH, Silverstein JH, et al. Pediatric obesity–assessment, treatment, and prevention: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. (2017) 102:709–57. doi: 10.1210/jc.2016-2573

56. Espinosa AM. Maternal employment and Mexican school-age children overweight in 2012: the importance of households features. *Bol Med Hosp Infant Mex*. (2018) 75:104–12. doi: 10.24875/BMHIM.M18000015

57. Katzow MW, Messito MJ, Mendelsohn AL, Scott MA, Gross RS. Grandparent coresidence and risk of early child overweight and obesity in low-income, hispanic families in New York City. *Acad Pediatr*. (2022) 22:90–7. doi: 10.1016/j.acap.2021.05.020

58. Davis SP, Northington L, Kolar K. Cultural considerations for treatment of childhood obesity. *J Cult Divers*. (2000) 7:128.

59. Campos AP, Vilar-Compte M, Hawkins SS. Cross-national comparison of breastfeeding, malnutrition and associated risk factors among Mexican-origin children living in Mexico and the US. *Cross-Cult Res*. (2021) 55:287–318. doi: 10.1177/10693971211021558

60. Cerdán C, del Carmen C, Polanco Medina IP, Cortés Salazar CS, Morteo Ortiz E, Acosta Cervantes M. Evaluación de la calidad nutricional del refrigerio y estado nutricio de una población de estudiantes de primaria de Veracruz, México. *Nutr Clín Diet Hosp*. (2018) 38:85–92. doi: 10.12873/383caballero

61. Basto-Abreu A, Barrientos-Gutiérrez T, Vidaña-Pérez D, Colchero MA, Hernández-F M, Hernández-Ávila M, et al. Cost-effectiveness of the sugar-sweetened beverage excise tax in Mexico. *Health Aff*. (2019) 38:1824–31. doi: 10.1377/hlthaff.2018.05469