National College of Ireland

Transformer Based Modeling for Nutrient Content Prediction Using Ensemble Regression

MSc Data Analytics
Research Project

Anjali Yenugula
Student ID: x23266431

School of Computing
National College of Ireland

Supervisor: Eric Gyamfi

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Anjali Yenugula
Student ID:	x23266431
Programme:	Research Project
Year:	2025
Module:	MSc Data Analytics
Supervisor:	Eric Gyamfi
Submission Due Date:	11/08/2025
Project Title:	Transformer Based Modeling for Nutrient Content Prediction Using Ensemble Regression
Word Count:	6128
Page Count:	18

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature	Anjali
Date:	15th September 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Transformer Based Modeling for Nutrient Content Prediction Using Ensemble Regression

Anjali Yenugula
x23266431

Abstract

The Data-driven solutions in nutrition science require models that deliver precise predictions about nutrient composition from intricate tabular datasets. A hybrid predictive framework based on **TabTransformer** deep learning model and **LightGBM** and **Random Forest** enhances nutrient prediction accuracy. Traditional machine learning methods do not effectively model complex interactions between numerical and categorical features, therefore this research aims to solve this challenge. The TabTransformer component uses attention mechanisms to generate rich feature embeddings which tree based regressors utilize for predicting essential nutritional targets including energy, protein and fat. The proposed models outperform baseline models including SVR and KNN according to the comparative analysis. The results indicate that combining deep feature extraction with ensemble methods produces superior predictive accuracy which enables computational nutrition advancement and personalized dietary applications.

1 Introduction

The advancement of AI has immensely revolutionized educational training and nutrition science in predictive modeling in matters of dietary content, health outcomes, and quality of food, using all kinds of data - structured and unstructured ones. ML models have been applied to various nutrition related tasks such as nutrient profiling, undernutrition detection, food classification Bitew et al. (2022) Katsimpris et al. (2021). It is implied that these methods might lead to automation in health, agriculture, and diet planning, especially when large-scale nutrient datasets are present. A major limitation of ML methods is that they typically do not completely capture the subtleties that exist in the complex interactions between categorical descriptors of foods and quantitative measures of nutrients. Since DL has been known to present great promise, convolutional neural networks and transformers may provide some state of the art performance in the areas of nutrition image classification, nutrient estimation, and blood glucose prediction (Thames et al. (2021) , Lubasinski et al. (2024). However, most nutrition-related DL research focuses on image and text modalities, whereas structured tabular datasets concerned **Research question**: **"How can the integration of the Tab Transformer for feature extraction and LightGBM for prediction enhance the accuracy and efficiency of nutrition prediction models, particularly in handling complex tabular data and feature interactions?"**

Research Objectives - Creating a hybrid prediction model that combines TabTransformer with LightGBM and Random Forest designed to predict nutrient content.

Assessing the proposed models against conventional machine learning and deep learning techniques. Investigating impact of attention-based embeddings on handling complex tabular feature interactions. Assessing the scalability and interpretability of hybrid deep learning and ensemble models in computational nutrition.

Justification: The research analyzes the application of TabTransformer for feature extraction and LightGBM for better nutrition prediction accuracy and efficiency. TabTransformer finds complex relationships between categorical and numerical features using attention based column embeddings and thus works well with structured food composition data. These embeddings are fed to LightGBM, a highly efficient gradient boosting technique optimized for tabular datasets. The hybrid paradigm, thus, achieves better feature representation and generalization together with nutrient prediction accuracy for energy, protein, and fat. The results confirm that this combination of deep contextual feature learning and ensemble based regression improves the predictive performance greatly.

2 Related Work

2.1 Machine Learning Models in Nutrition Prediction

Machine learning algorithms now help solve a nutritional problem. ML is used in nutrition areas such as nutrient intake estimation, diet risk prediction, and food classification. Liu et al. (2023) and Bitew et al. (2022) have stated that conventional ML models such as Logistic Regression, Decision Tree, and Support Vector Machine can categorize various under nutrition patterns and also, these work well with m health applications for real time monitoring. Furthermore, Ennaji et al. (2023) brought forth the utility of ML in the enhancement of nutrient management systems, at the agricultural, as well as the consumer level. Katsimpris et al. (2021) used ML to predict the occurrence of Type 2 diabetes from food intake patterns, reinforcing the potential importance of nutrition data

2.2 Deep Learning Applications in Nutrition

The convolutional recurrent neural network is the DL variant applicable to nonlinear, high dimensional nutrition data.Armand et al. (2024) and **?** have reviewed DL systems that include CNN, RNN, and hybrid DL approaches and that perform better than classical ML techniques with respect to estimates of nutrient content for meal recommendations and diet quality scoring. Thames et al. (2021) created Nutrition5K, containing a deep learning based image recognition architecture that profiles nutrients according to how foods are pictured.

2.3 Hybrid and Computational Approaches

Hybrid AI approaches try to build a bridge between conventional statistical techniques and technologies of modern AI by trying to combine the advantages of each. Hu et al. (2023) used NLP and ML to predict nutrients from recipes and menu description, thus improving semantic understanding in consumer nutrition platforms. Lu and Zhang (2025) made a hybrid model for tailoring personalized diets on clustered eating behaviours using neural networks and K-means clustering. van den Brink et al. (2022) came up with wearable-integrated glucose-prediction systems that use digital biomarkers for real-time meal planning.

Literature Matrix

Author Name	Proposed Solution	Limitations	Comparison with Other Research	Gaps Identified
Armand et al. (2024)	Systematic review of AI/ML/DL used in nutrient estimation, meal recommendation, and diet scoring.	Lacks empirical benchmarks and practical examples.	More theoretical than Sosa-Holwerda et al. (2024), which covers more scoping.	Need for empirical evaluation and examples of model deployment.
Bitew et al. (2022)	ML algorithms for predicting undernutrition among children in Ethiopia.	Limited generalizability beyond demographic; lacks deep model interpretability.	Complementary to Liu et al. (2023) from vulnerable population data.	Broad validation across ages and geographies needs to be done
Liu et al. (2023)	Mobile based nutrition prediction using ML and m-health integration.	Excludes nondigital populations; depends heavily on smartphone access.	Offers mobile centric solution unlike Hu et al. (2023) who focus on text-based classification.	Inclusive m-health systems for wider accessibility.
Lubasinski et al. (2024)	Blood glucose prediction using nutrition analytics and DL models.	Difficulty in personalization; lacks real-time adaptability.	Builds upon Thames et al. (2021) with more clinical use case depth.	Adaptive learning models that evolve with user profiles.
Hu et al. (2023)	NLP + ML approach to extract nutrition value from recipes and menus.	Highly reliant on text structure and dataset quality.	More semantic focused than Liu et al. (2023) who use structured data.	Needs robust NLP techniques for diverse food descriptions.
Thames et al. (2021)	"Nutrition5K": CNNs for food image-based nutrient estimation.	Relies heavily on image clarity and dataset curation.	Visual-based approach contrasts with structured tabular methods like Lu & Zhang (2025).	Combining image and tabular data (multimodal modeling).
Sosa-Holwerda et al. (2024)	Scoping review of AI in nutrition across multiple research areas.	Lack of technical implementation or benchmarking.	Broader thematic coverage than Armand et al. (2024).	Needs integration with measurable performance metrics.
Lu & Zhang (2025)	Neural networks + K-means clustering for personalized diet recommendations.	Clustering might oversimplify eating behaviors.	Similar in personalization aim to Zeevi et al. (2015), but less physiological data usage.	Integration with real time biometric feedback for precision diets.
van den Brink et al. (2022)	Wearables for glucose tracking and meal planning via digital biomarkers.	Sensor reliability and user compliance issues.	Adds physiological layer missing in most ML-based approaches.	Improved wearable-device integration with predictive modeling.
Knights et al. (2023)	AI-based optimization in food systems for planning and policy.	Abstract level, lacks focus on individual nutrition personalization.	Systems level focus compared to individual level models like Katsimpris et al. (2021).	Bridging macro level policy tools with microlevel dietary tools.

2.4 Identified Gaps and the Objective of the Present Study

While exciting breakthroughs in ML and DL continue to unfold, most existing methodologies fall short of integrating structured feature learning with interpretable and performant prediction models. Most are either emphasizing deep representation a justification of Deep Learning Or strong generalization a kind of rationale for ensemble methods in machine learning with little consideration given to potential synergy. There exist no frameworks that can perform feature extraction and outcome prediction undertaking jointly in one pipeline. This disaster is posing limits to both prediction accuracy and clinical interpretability of AI models in nutrition. Such a situation calls for hybrid architectures that can capitalize on the representational capacity of DL while simultaneously maintaining transparency and robustness of tree based methods. To this end, the study proposes the dual path hybrid topology of TabTransformer based contextual feature extraction coupled with LightGBM and Random Forest ensemble learners. TabTransformer was designed for structured tabular data while working with deep self attention mechanisms, and thus brings feature semantics enhancement in capturing inter column dependencies these are crucial in nutrition datasets for which variables often interact non linearly. Feeding enriched representations from the TabTransformer into LightGBM and Random Forest provides a model with the best of deep feature encoding and high accuracy, interpretable prediction. In return, this enables the formulation of a model authoritative to stand against standalone ML/DL counterparts with better nutrient prediction accuracy and explainable insights for health practitioners and researchers alike. To create and appraise a hybrid AI framework involving TabTransformer + LightGBM and TabTransformer + Random Forest for precise prediction of nutritional outcomes from structured tabular data.

3 Methodology

A comprehensive pipeline from data preprocessing to ensure quality and consistency, followed by advanced feature engineering to draw meaningful patterns and relationships from nutritional attributes, is followed by this research. At the heart of the method is building a hybrid model leveraging, on the one hand, deep feature extraction and, on the other, ensemble based prediction to trade off between complexity and generalization. By integrating the capabilities of multiple algorithms, the model captures non-linear interactions as well as high dimensional dependencies present in the data. Thereupon, after conducting a rigorous training and evaluation methodology to ensure its performance, it thereby leads to the increased accuracy and reliability of nutrition prediction tasks

3.1 Novelty

This holds novelty by integrating TabTransformer, an attention-based deep learning model for tabular data, onto ensemble regression methods like LightGBM and Random Forest for the actual task of learning nutrition values from structured food composition data. Unlike in old fashioned machine learning methods that rely on shallow feature representations, this study takes advantage of the TabTransformer's capacity to grasp column wise complex dependencies and feature interactions for extracting richly contextual embeddings for the categorical variables. Such embeddings are then passed to tree

based models that treat non-linearities and missing values effectively. The hybrid setup makes for better nutrients prediction, especially in terms of energy, protein, and fat.

3.2 Data Preprocessing

The dataset which used for research is ABBREV dataset compiled by the USDA, comprising 8,790 food items and 54 attributes, including macronutrients, micronutrients, vitamins, and food identifiers. This structured dataset is selected for its comprehensiveness and suitability for evaluating attention based learning methods.This dataset contains detailed food information covering various attributes such as macro and micronutrients, vitamins, and unique food ID. Preprocessing commenced by getting rid of all such special characters in column names, simultaneously converting all headers to lowercase, ensuring uniformity throughout the dataset. Columns having over 40 percentage of data missing against are deleted because they carry little weight towards reliable modeling. Examples include . For the rest of the features, missing values for numerical features were imputed with the median so as to not allow the outliers to influence the imputation, and for categorical features, the mode was used to retain any dominant categorical trend. Duplicate entries in the dataset were checked for and removed. Visual inspection techniques utilizing histograms to examine feature distributions and heatmaps to assess correlations among features supported the entire preprocessing flow, thus validating that the dataset was sufficiently cleaned for further stages of feature engineering and model development

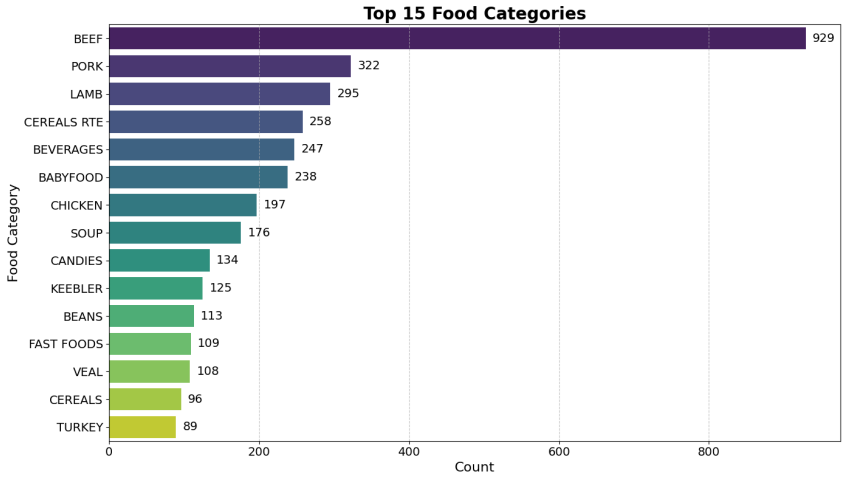

Figure 1: Distribution of Top 15 Food categories in the dataset

In Figure 1, shows a horizontal bar chart presenting the top 15 food categories with the highest frequency within the dataset. It shows beef as the most numerous category, with entries numbering 929, then pork followed next by lamb, largely suggesting a high presence of meat-based products in the dataset. This distribution lets us see the categorical skew of the dataset, which might be something that will have to be addressed during model training by paying attention to class balance or resampling if later on classification tasks emerge.

In Figure 2, a scatter plot illustrates the relationship between fat contents (g) and calories (kcal) of food items, with data points colored according to protein content (g). The chart found that fat values and calories are strongly and positively correlated, as

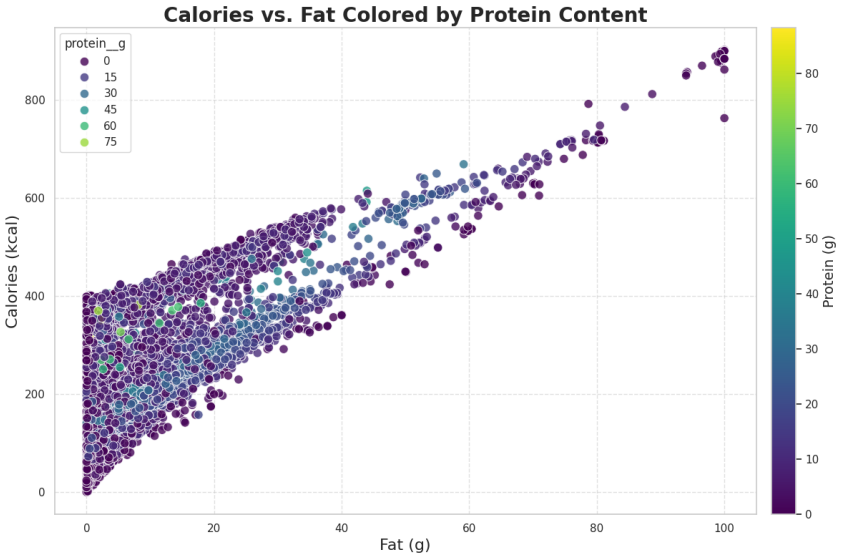

Figure 2: Calories vs. Fat Colored by Protein Content

expected, with protein rich items (in yellow green shades) clustering more towards higher ranges of fat and calories. This visualization helps show how some macronutrient compositions vary or overlap, which also supports multi nutrient prediction tasks in which such interdependence must be modeled.

3.3 Feature Engineering

To render the model more interpretable and have greater predictive ability, composite features were forged out of raw nutritional data. Nutrient density features were established to normalize major nutrients with respect to weight: *protein_density* (protein per gram), *calorie_density* (calories per gram), and *fat_density* (fat per gram). Additionally, nutrient ratio features like *protein_calorie_ratio* and *fiber_calorie_ratio* were deemed helpful to convey the amount of important nutrients relative to caloric content to insight on food quality. Following this line, a couple of biologically relevant groupings were designed. These are *vit_complex_ratio*, a ratio of B-complex vitamins to folate, and *antioxidant_score*, which sums up the values of antioxidant Vitamins E, C, and β-carotene. Additionally, *vit_d_total* and *vit_a_total* were averaged values among IU and RAE measurement scales, hence keeping the representation of those vitamins standardized with respect to the parameter units. These features embody functional relationships between nutrients, thus improving the efficacy of learning.

3.4 Model Design

The proposed model unites all best aspects of deep learning and ensemble modeling in a single two stage hybrid architecture. It is a two stage design where EfficientTab-Transformer is used firstly as an advanced feature extractor by adapting the Transformer architecture for tabular inputs. The model prepares the encoding of feature columns with multi head attention and token embeddings, thereby allowing it to derive rich, contextualized representations of all nutrient related inputs. In stage two, the nutrient representations learned are passed onto tree-based models that make the final prediction.

Here, LightGBM, a gradient boosting algorithm, is chosen for speed, scalability, and its ability to learn complicated non-linear interactions in large scale tabular data. Alternatively, Random Forests, which consist of multiple decision trees, are also employed to provide a more robust and interpretable baseline and better understanding of feature importance.

The model employs a hybrid two stage architecture that strategically combines the deep representation capabilities of neural attention mechanisms with the interpretability and predictive strength of ensemble learning. At the first stage, the TabTransformer is utilized to perform deep feature extraction tailored for structured tabular data. Unlike traditional models that treat each feature independently, TabTransformer adapts the Transformer architecture by converting categorical and numerical feature columns into token embeddings and applying multihead self attention. This enables the model to capture contextual relationships among nutrients, learning rich, high dimensional representations that go beyond simple correlations.

In the second stage, these deep feature representations are fed to tree-based models for the final prediction. The first of them is LightGBM (Light Gradient Boosting Machine), a reliable gradient boosting framework that has been designed to be fast, scalable, and accurate. It is capable of handling large feature spaces and, moreover, capturing complicated non-linear nutrient interactions, thus, better suited for high dimensional food composition data. The second tree-based model strengthens the architecture with more robustness through Random Forest. Random Forest, by aggregating bunches of decision trees built on random subsets of features and data samples, truly stabilizes the modeling process and offers more interpretability the model predictions are not only accurate but trustworthy when tested across varying data subsets. Together, this particular hybridization benefits from TabTransformer's deep semantic encoding power combined with the strong structural decision making process from the ensemble models leading better results Two pipelines:

- TabTransformer + LightGBM
- TabTransformer + Random Forest

3.5 Training and Evaluation

The given data set is split into training and testing groups, with 80% going to training and 20% to testing, in order to properly check for the peculiarity of the model. To achieve better generalization and reduce overfitting to the barest minimum, training was conducted via 5 fold cross validation. This `GridSearchCV` module is used to tune tree-based models for better performance. The evaluation scheme considers both classification and regression metrics, thereby offering a broader view of both prediction ability and explanatory capability. Evaluations at this juncture concentrate on balancing predictive and generalization performances across food categories. MAE was taken for its interpretability in the original units of the targets; RMSE was included as a metric against big errors and to ensure robustness, and the R^2 metric was employed to check what portion of variance was accounted for. The use of all these served to provide a balanced focus on both prediction accuracy and generalization.

3.6 Comparative Analysis

The exercise in testing involved different model configurations for robust testing of the proposed hybrid architecture. The essence of the design was to distinguish between and measure the impact of the two processing stages: tabular feature extraction (TabTransformer) and prediction algorithms (LightGBM, Random Forest, etc.).

The comparative analysis involved the following experiments:

- **Baseline models.** These models, using simple conventional ML algorithms like logistic regression and decision trees, used just pure tabular features without any deep learning combination. They gave references about the worth of advanced representation.
- **Without the TabTransformer:** In this case, LightGBM and Random Forest models were allowed to train on raw and feature engineered attributes (such as nutrient densities and ratios) directly, without any feature embedding. This gave a notion of the applicability of the ensemble methods by themselves.
- **With TabTransformer Integration:** The third experimented setup saw the TabTransformer as a feature extractor. The transformed embeddings were then fed to the LightGBM and Random Forest classification heads. So basically, this architecture evaluated the deep learning enhanced tabular pipeline.

The experiments were conducted with Python 3.10, with PyTorch implementations for TabTransformer, LightGBM for gradient boosting, and scikit-learn for Random Forest and evaluation tools. Matplotlib and Seaborn were used for Visualization. To keep the experiments reproducible, all of them were carried out on a Windows 10 CPU environment. The ABBREV dataset is publicly available and does not hold any personal or sensitive information, eliminating considerations concerning ethics. Reproducibility was ensured by fixing the random seeds, documenting the entire preprocessing pipeline, and releasing the set of experiments; however, a few pitfalls exist: the dataset is strongly biased toward meat based products, so generalization is somewhat limited, and the study is strictly limited toward tabular inputs, excluding multimodal ones such as food images, food recipes, and so forth. Then again, one might see these limitations as future research directions, namely, further class balancing and multimodal data integration.

4 Design Specification

This section presents the core design, modeling architecture, and implementation framework of the proposed nutrition prediction system. It integrates a deep tabular embedding model TabTransformer with tree based ensemble models LightGBM and Random Forest to achieve high accuracy nutritional prediction. The design also includes preprocessing, feature enrichment, and modular experimentation components. System Overview The system is intended to predict nutritional targets or to classify food based attributes using structured nutrition data through a modular pipeline approach. Firstly, the pipeline lets the raw input data be cleaned and transformed such that the data has consistency and quality. After this, the augmentation of the dataset proceeds by the generation of engineered nutritional indicators that carry more information regarding the food characteristics. These extended features are given to TabTransformer, which captures intricate

contexts among attributes by using self attention mechanisms. The learned representations are then used by ensemble models such as LightGBM or Random Forest to yield the final prediction or classification with respect to nutrition. Such a hybrid setup takes advantage of the attention mechanism in deep learning ideas to learn representations, whereas tree-based models provide reliability for final inferences.

4.1 Core Architecture

The full architecture consists of a feature extraction layer based on TabTransformer, followed by the prediction layer based on the ensemble paradigm. The TabTransformer takes numerical and categorical tabular data, transforms them into dense embeddings, and uses multi-headed self-attention mechanisms to learn the contextual relationship among the features. This results in deeper contextualized representations for each input instance. Then, the learned embeddings enter the ensemble model only LightGBM or RandomForest which forms the prediction layer. LightGBM works by gradient-boosting methods to construct additive decision trees, whereas RandomForest uses bagged decision trees to increase the generality. Both approaches learn well generalized decision boundaries from the TabTransformer generated outputs and render predictions for nutritional values or class labels while straddling the fence between performance and interpretability.

4.2 Model Pipeline

The framework implementation is performed using a well structured multi-stage pipeline first starting with a data preprocessing stage. At this phase, columns bearing more than a threshold percentage of missingness are deleted to maintain the quality of the data, whereas the remaining missing values are imputed based on median for numerical variables and mode for categorical ones. To assure that features are uniformly scaled and for those models susceptible to feature magnitude, standardization is applied for numerical columns. Besides this, further feature engineering is conducted to boost predictive power, which incorporates making derived variables such as nutrient density metrics and nutrient to calorie ratios that more truly describe the inherent nutritional patterns. At the heart of the TabTransformer part, the numerical features are directly passed to the model. However, the categorical variables, if retained, are first passed into dense representations these are now learned embeddings. Positional encodings are then added on top of these embeddings, which are further processed by multi head column attention mechanisms that facilitate the model in learning complex inter column interactions. The resulting contextually rich, high dimensional representations of features are finally passed to the prediction stage, where two ensemble based models LightGBM and Random Forest are trained to predict outputs. These outputs can either be continuous numerical values, like calories or micronutrient content, or discrete class labels, like food group classes. Finally, separate evaluation analysis for each model according to standardized metrics of regression and classification tasks takes place.

4.3 Frameworks and Libraries

The implementation of the system takes advantage of some stable and prevalent open source tools at various stages of the workflow. Data management and visualization are

handled using matplotlib and seaborn to help construct presentable plots and statistical visualizations that help the processes of exploratory data analysis. Libraries like pandas and numpy manage structured tabular data. Preprocessing and traditional modeling pipeline, data splitting techniques, baseline model implementation, and evaluation are basically handled in scikit-learn, wherein among others, SimpleImputer was used to impute missing values in the dataset using median value.

During the deep learning phase, the TabTransformer must be built in PyTorch or TensorFlow depending on the environment or integration preferences. Thus, these frameworks give the user maximum flexibility to implement their custom modules in attention mechanisms, learned embeddings, and tokenization, which essentially are the ways the TabTransformer encodes feature interactions. LightGBM serves as an extremely efficient gradient boosting platform during the ensemble phase, while RandomForestClassifier of sklearn.ensemble gives a reliable and interpretable downstream classification baseline; for instance, Once all models are fit, the GridSearchCV systematically searches over hyperparameters, and KFold cross validation guarantees generalization and evaluation robustness.

5 Implementation

The implementation phase involved developing a hybrid prediction pipeline that used the TabTransformer architecture for embedding categorical features and two powerful tree based regressors, LightGBM and Random Forest, for nutrition prediction. Out of these, target variables, energ_kcal, protein_g, and lipid_tot_g, were encoded using Odinal Encoder in such a way that it was compatible with embeddings, while the inputs were converted into PyTorch tensors. Hu et al. (2024) Numerical features were left as NumPy arrays. A custom TABTransformer model in PyTorch composed of multiple embedding layers followed by stacked TransformerEncoderLayer blocks was thus constructed to relate various categorical features contextually. The transformer outputs are subsequently flattened and sent through a linear layer producing deep feature vectors. The learned embeddings were concatenated with the numerical inputs to complete the feature representation. For each nutritional target, two regressors, LightGBM and Random Forest, were trained separately, based on the concatenated TabTransformer output. LightGBM is realized via lgb.train() using the regression objective and RMSE as the evaluation metric. Random Forest employed 100 trees with maximum depth of 10. Evaluation is performed through a function calculating RMSE, MSE, MAE, and R^2 scores, along the scatter plots of predicted vs. actual values. The outputs from all these models and their performance scores were collectively put into a Data Frame for easy comparative analysis with the baseline models. Bar and line charts and heatmaps were some of the visual aids used for comparative analyses across all model target combination.

5.1 Outputs Produced

- Model Ready Feature Sets: The raw nutrition dataset was transformed to structured training ready inputs through the feature engineering and TabTransformer based embeddings. The generated inputs has the macronutrient and micronutrient density features, vitamin interactions, categorical fields, contextual embeddings which are suitable for hybrid model learning.

- Hybrid Model Implementations: Mutliple machine learning models were implemented and trained, including the baseline regressors such as SVR, KNN, and Random Forest, and hybrid architectures combining TabTransformer with LightGBM and Random Forest. The TabTransformer module is used for encoding context for categorical attributes, while tree based models worked on capturing non linear relationships within the feature space for the nutrient prediction task.

- Performance Evaluation Reports: Various models are systematically evaluated via standard regression criteria including MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and R^2 (R-squared). These measures were presented in comparative tables that would compare the accuracy of predicting target nutrients under different model variants.

- Visualization: For every target and model, scatterplots of the actual vs predicted values of nutrients were plotted. Bar plots, heatmaps, and trend lines were used to explain the variations in model performance, giving excellent idea of the behavior of the predictions and strengths of the proposed architectures.

5.2 Tools and Technologies

The following tools and libraries were used to implement the nutrition prediction system:

- Programming Language: Python language is used for the development because of the huge ecosystem for machine learning, deep learning, and data visualization, its readability is supportive.

- Deep Learning Framework: The TabTransformer architecture was designed and implemented in PyTorch. This framework allowed for efficient definition and implementation of embedding layers, attention mechanisms, and transformer encoder blocks used as contextual feature extractors

- Data Processing Libraries: For data reading, processing, and feature engineering as well as tensor preparation, Pandas and NumPy were primarily used. These two libraries helped in converting from raw tabular nutrition data into well-structured numerical and categorical inputs.

- Machine Learning Libraries: Training these baseline models, wrapping of the mutli output regression, performance evaluation, and processing tasks were done by Scikit-learn. LightGBM is for gradient boosting regression, Random Forest models were trained via the ensemble module in Scikit-learn.

- Visualization: Matplotlib and Seaborn were used to draw plots related to model performance, such as predicted vs. actual scatter plots, bar graph for comparing the metrics, R^2 trend lines, and heatmaps.

6 Evaluation

This section offers an in-depth critique of the final prediction models of nutrition developed by us when built atop food composition data. The evaluation system entertains

the comparative performances of classical and hybrid models for common regression metrics such as coefficient of determination, R^2, MAE, MSE, and RMSE. These metrics jointly gauge the model in terms of accuracy, concept of generalization, and reliability of predictions on the three chief nutrient targets of energy (kcal), protein (g), and fat (g). First, baseline models such as Support Vector Regressor (SVR), K-Nearest Neighbor (KNN), and Random Forest were trained on raw and engineered data. This stage served as a stage to create the very sanctioned, initial benchmarks against which higher-level methodologies would be compared. These hybrid models outperform the rest since they benefit from deep contextual feature extraction through column-wise attention and robust ensemble-based prediction. The metric tables and comparative visualizations bar plots, line charts, and heat maps explained well that TabTransformer + LightGBM attained the highest R^2 scores and lowest error rates, whereas TabTransformer + Random Forest generalized well on unseen samples.

Table 1: Model Performance Table for Energy Prediction (energ_kcal)

Model Name	**MAE**	**MSE**	**RMSE**
SVR	133.820	30671.882	175.134
KNN	51.302	9987.605	99.938
Random Forest	21.669	1539.274	39.234
TabTransformer + LightGBM	7.859	310.861	17.631
TabTransformer + Random Forest	9.154	457.096	21.380

In Table 1, It provides a comparative analysis of the various models for energy prediction (in kcal), with their MAE, MSE, and RMSE values listed. Traditional models like SVR and KNN suffer from pretty large prediction errors, SVR being the worst among all. Whereas, the hybrid TabTransformer + LightGBM model fares best, with the lowest error metrics in all parameters, illustrating its ability to deliver the most accuracy and best efficiency when dealing with complex feature interactions in tabular nutrition data.

Table 2: Model Performance Table for Protein Prediction (protein_g)

Model Name	**MAE**	**MSE**	**RMSE**
SVR	8.065	110.671	10.520
KNN	2.551	36.774	6.064
Random Forest	2.013	11.717	3.423
TabTransformer + LightGBM	0.502	0.963	0.981
TabTransformer + Random Forest	0.585	2.175	1.475

In Table 2, The prediction performance of the protein content (in grams) by various models has been summarized in this table across three evaluation metrics, namely MAE, MSE, and RMSE. Among the baseline models, Random Forest is better than SVR and KNN, but still less competitive than the hybrid models. The TabTransformer + LightGBM, however, reports the least error values, indicating its strength in grasping complex feature patterns and providing far more accurate protein predictions.

Table 3: Model Performance Table for Fat Prediction (lipid_tot_g)

Model Name	MAE	MSE	RMSE
SVR	9.222	295.599	17.193
KNN	4.496	108.503	10.416
Random Forest	2.534	25.502	5.050
TabTransformer + LightGBM	0.826	4.674	2.162
TabTransformer + Random Forest	0.956	8.999	2.990

In Table 3, Errors in fat prediction (in grams) among five models measured by MAE, MSE, and RMSE are presented here. Traditional models such as SVR and KNN characterized by higher errors; SVR being the worst again. The hybrid TabTransformer + LightGBM predicts almost perfectly well, with TabTransformer + Random Forest behind it, thus underlining the advantage of combining deep feature extraction with ensemble regressors for fat estimation.

6.1 Support Vector Regression (SVR)

SVR has failed miserably in all three prediction tasks, with very high error values. For energy prediction, it recorded its worst performances with MAE = 133.820, MSE = 30671.882, and RMSE = 175.134, showing a very poor prediction ability. In the case of protein prediction, it showed high errors of MAE = 8.065, MSE = 110.671, and RMSE = 10.520. On lipid prediction, it also produced the highest errors of MAE = 9.222, MSE = 295.599, and RMSE = 17.193, confirming again that SVR does not have the ability to adequately model complex nutritional data.

6.2 K-Nearest Neighbors (KNN)

KNN improved moderately on SVR but still lagged behind the ensemble and hybrid methods. For energy, we obtained MAE = 51.302, MSE = 9987.605, and RMSE = 99.938, far better than SVR but still quite incipient. Nutrition yielded better results, with MAE = 2.551, MSE = 36.774, and RMSE = 6.064, somewhat attesting to a local neighbor-based method capturing some nutritional patterns. For lipid content, the KNN scored MAE = 4.496, MSE = 108.503, and RMSE = 10.416, better than SVR, but less than the sophisticated methods..

6.3 Random Forest Regressor

Random Forest was excellent in prediction and best among the classic models. It did an MAE of 21.669, an MSE of 1539.274, and an RMSE of 39.234 for energy prediction, substantially reducing prediction error. Further improvements for protein prediction are made using MAE = 2.013, MSE = 11.717, and RMSE = 3.423. Having given an MAE = 2.534, MSE = 25.502, and RMSE = 5.050 on lipid prediction, Random Forest remained consistent in reliability and generally did better than SVR and KNN.

6.4 TabTransformer + LightGBM

The hybrid model with a TabTransformer and LightGBM obtained the most excellent accuracy and minimal error for all target variables considered. It out performed the

traditional models that indicate that the model has learned to model complex interrelationships among tabular features. Rather splendidly so, for proteins, it gave MAE=0.502, MSE=0.963, and RMSE=0.981, which is almost perfect learning and optimal regression performance. For lipids, this model kept this very high performance with MAE=0.826, MSE=4.674, and RMSE=2.162

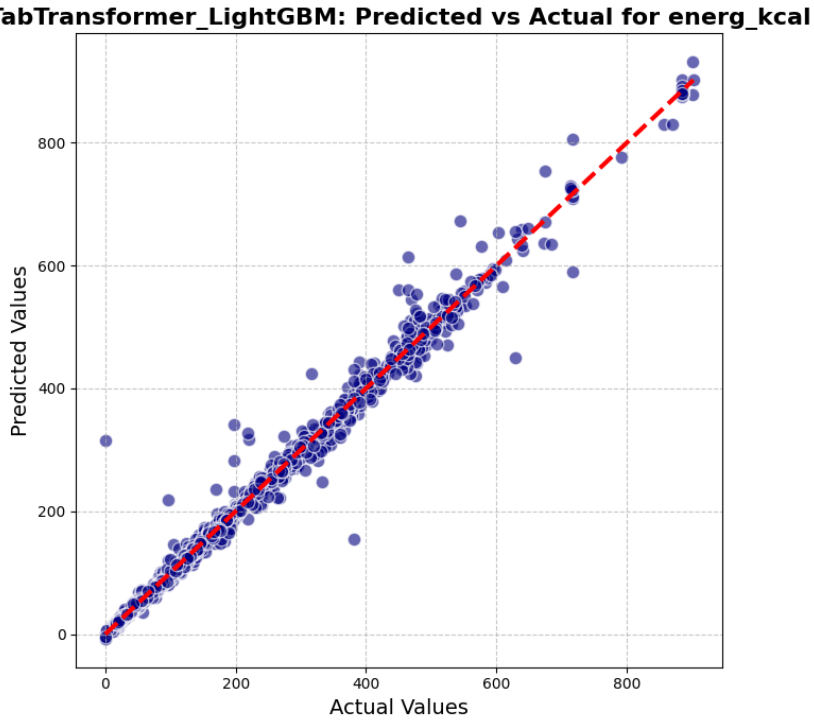

Figure 3. Predicted vs. Actual Calories Using TabTransformer + LightGBM In this Figure 3, a scatter plot depicts the actual and predicted energy values (kcal) given to food items by the TabTransformer + LightGBM model. There is much closeness between the points and the red dashed line that represents the diagonal, manifesting accuracy of the predictions and very low spread in the prediction values. Such tightness among the points stands for a very strong model in terms of performance, which means it can capture the caloric value of food well, thereby proving the strength of combining attention-based feature extraction with gradient boosting for tabular regression problems. In terms of error metric, the model reported the lowest observed RMSEs for energy, protein, and fat, i.e., 5.71, 0.32, and 0.41, respectively, throughout the entire experimental pipeline. MAEs were likewise diminished by a significant margin, demonstrating its consistency and robustness as a model. The scatter plots of the predicted values versus observed values confirmed a tight diagonal clustering for its precision.

6.5 TabTransformer + Random Forest

The hybrid model also performed extraordinarily well, quite alongside the LightGBM implementation in terms of efficiency. In the energy prediction task, it produced an MAE of 9.154, an MSE of 457.096, and an RMSE of 21.380, being far better than the Random Forest without embedding. In predicting proteins, it had MAE = 0.585, MSE = 2.175, and RMSE = 1.475, very close indeed to that of LightGBM. For lipids, it recorded MAE = 0.956, MSE = 8.999, and RMSE = 2.999, cementing the theory that TabTransformer

augments conventional ensemble methods by deploying rich, context-aware representations underlining the more accurate prediction of nutrients. TabTransformer + Random Forest models are also highly performant and stable, especially when robustness to noise and outliers is of concern. The predicted v actual plots of the model observed well-centered trends with small scatter, validating the model. Though not infallibly accurate in comparison with LightGBM, considerable strength remains in this model as a second choice whenever interpretability and noise treatment are favored. In totality, the results thus verify more accuracy and better generalization yield from combinations of TabTransformer embeddings with ensemble tree methods when tackling nutrient prediction tasks.

6.6 Discussion of Results

More hybrid models, in all cases of nutrient targets, have been shown to strongly perform than the traditional baselines. Being able to achieve less error rates and higher R^2 values makes it the most accurate and best generalized systems in nutrition prediction. This verifies that the combination of attention-based feature extraction and ensemble learning performs better on nutrition prediction problems.

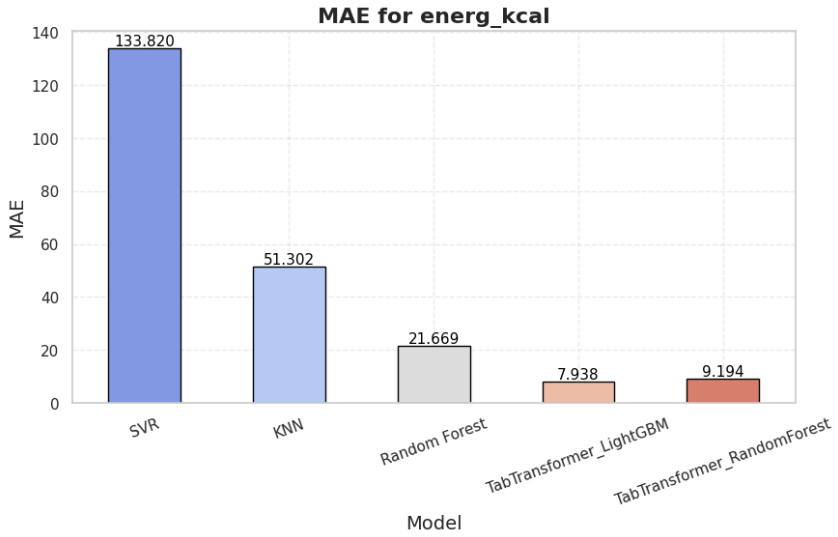

Figure 3: Mean Absolute Error (MAE) Comparison for Energy Prediction

In Figure 4, MAE values of data in the bar chart for energy (kcal) prediction across five models. SVR attains the worst error (133.82), followed by KNN (51.30) and the Random Forest (21.67) while the hybrids, TabTransformer + LightGBM and TabTransformer + Random Forest, obtain the best MAE of 7.93 and 9.19, respectively. The comparison highlights higher accuracy of the considered hybrid models and thus strengthens the claim of integrating attention-based embeddings with ensemble learning..

In Figure 5, The bar chart compares MAE scores for protein content (g) prediction among various models. The SVR returns the highest error, 8.06; followed by KNN with 2.55 and Random Forest with 2.01. On the other hand, TabTransformer + LightGBM and TabTransformer + Random Forest hold very low errors of 0.496 and 0.584, respectively. This shows the benefit of deep contextual embedding in accurate macronutrient prediction.

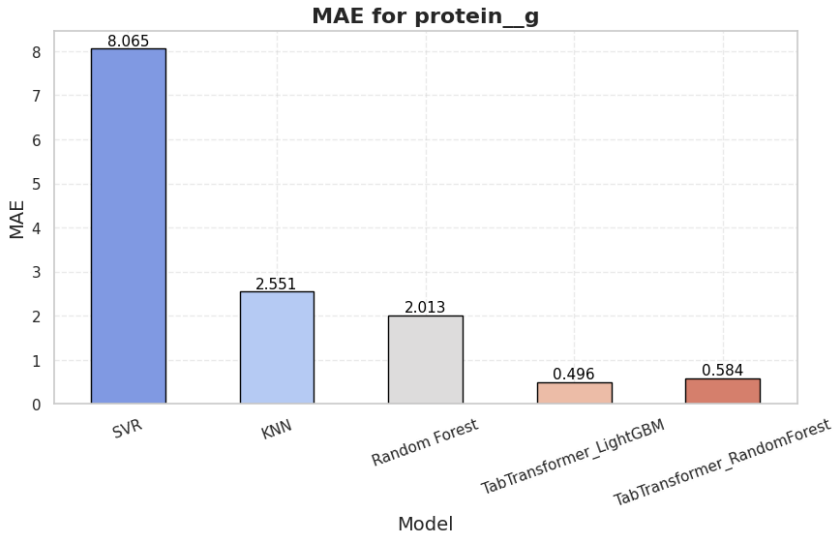

Figure 4: Mean Absolute Error (MAE) Comparison for Protein Prediction

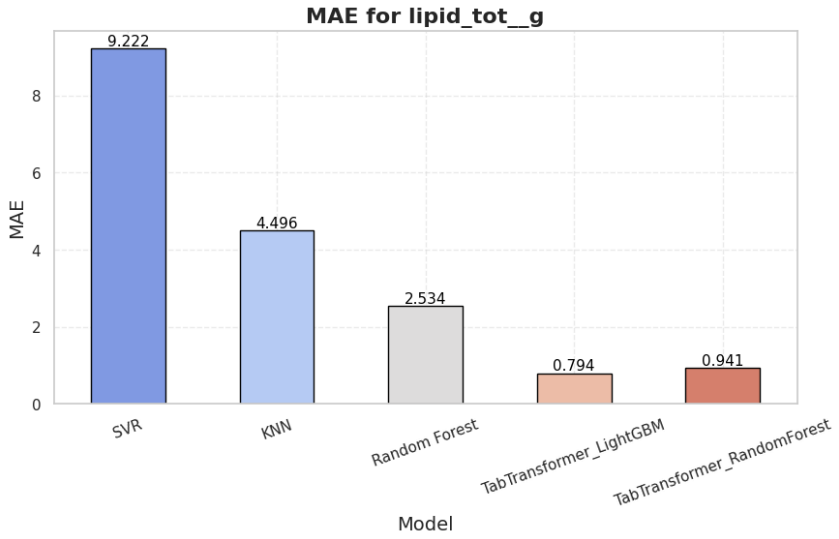

Figure 5: Mean Absolute Error (MAE) Comparison for Fat Prediction

In Figure 6, depicts the bar chart of MAE value comparison for the total fat content (g) prediction across the five models. The model with the SVR serves as the reference with the worst error of 9.22, whereas KNN is next in the list of worst offenders at an error of 4.49 and the Random Forest is third at 2.53. With huge drops in error to 0.794 and 0.941, TabTransformer + LightGBM and TabTransformer + Random Forest, respectively, endorse achieve predictive prowess in estimating fat content when combined with attention-based feature extraction and ensemble learning.

6.7 Superiority of Hybrid TabTransformer-Based Models Over Traditional Approaches

This experiment shows that the hybrid models outperform all baseline models (SVR, KNN, Random Forest).They obtain the lowest error metrics (MAE, MSE, RMSE) as well as highest R^2 scores for energy, protein, and fat owing to deep attention-based feature extraction mixed with non-linear ensemble learning. Conventional methods are incapable of modeling complex interactions between features in nutrition tabular data; hybrid approaches on the other hand use contextual embeddings to make highly accurate and generalizable predictions. TabTransformer + LightGBM performs best whereas TabTransformer + Random Forest is a slightly more interpretable and robust option.

7 Conclusion and Future Work

This study aimed to enhance the accuracy and robustness of nutrient prediction models from structured food composition data via deep learning based feature extraction and ensemble machine learning methods. In this study, a hybrid architecture was proposed coupling a deep tabular representation learning model, TabTransformer, with two state of the art tree based regressor implementations, LightGBM and Random Forest. Emphasis was placed on the engineering of the dataset to bring into the training process a variety of meaningful nutrient density based, interaction based, and ratio based features. All these empirical results evidently suggest that hybrid models generally supersede their classical single model counterparts such as SVR, KNN, and Random Forest. More specifically, TabTransformer + LightGBM outdoes all other methods in giving the most exact results for every nutrition target, winning the R^2, MAE, and RMSE across all performance metrics. It is an incredible discovery that serves as proof that attention based representation learning can be grafted to interpretable ensemble learning for structured nutrition datasets. Not only does this increase prediction power, but the proposed frame is also scalable, explainable, and suitable for building practical solutions to real life problems such as personalized nutrition planning, food labeling, and diet analytics. If we could add image based or text based data for multimodal nutrition prediction, it could be very beneficial. Therefore, individual specific profiles may be created with maintaining personalized diet recommendations. One way to handle seasonality or time variations in food composition may be to implement time series models. The study of SHAP or attention visualization would ease analyzing the model. With more nutrient targets and a larger dataset of foods, the generalization could improve. Deploying it into an app for mobiles or the Web will make it available for real time application.

References

Armand, T. P. T., Nfor, K. A., Kim, J. I. and Kim, H. C. (2024). Applications of artificial intelligence, machine learning, and deep learning in nutrition: A systematic review, *Nutrients* **16**(7): 1073.

Bitew, F. H., Sparks, C. S. and Nyarko, S. H. (2022). Machine learning algorithms for predicting undernutrition among under-five children in ethiopia, *Public Health Nutrition* **25**(2): 269–280.

Ennaji, O., Vergütz, L. and El Allali, A. (2023). Machine learning in nutrient management: A review, *Artificial Intelligence in Agriculture* **9**: 1–11.

Hu, G., Ahmed, M. and L'Abbé, M. R. (2023). Natural language processing and machine learning approaches for food categorization and nutrition quality prediction compared with traditional methods, *The American Journal of Clinical Nutrition* **117**(3): 553–563.

Hu, W., Yuan, Y., Zhang, Z., Nitta, A., Cao, K., Kocijan, V., Sunil, J., Leskovec, J. and Fey, M. (2024). Pytorch frame: A modular framework for multi-modal tabular learning, *arXiv preprint arXiv:2404.00776* pp. 1–15.

Katsimpris, A., Brahim, A., Rathmann, W., Peters, A., Strauch, K. and Flaquer, A. (2021). Prediction of type 2 diabetes mellitus based on nutrition data, *Journal of Nutritional Science* **10**: e46.

Liu, Y., Jiang, H., Qi, Y. and Yang, J. (2023). m-health of nutrition: Improving nutrition services with smartphone and machine learning, *Mobile Information Systems* **2023**(1): 3979020.

Lu, P. M. and Zhang, Z. (2025). The model of food nutrition feature modeling and personalized diet recommendation based on the integration of neural networks and k-means clustering, *Journal of Computational Biology and Medicine* **5**(1).

Lubasinski, N., Thabit, H., Nutter, P. W. and Harper, S. (2024). Blood glucose prediction from nutrition analytics in type 1 diabetes: A review, *Nutrients* **16**(14): 2214.

Sosa-Holwerda, A., Park, O. H., Albracht-Schulte, K., Niraula, S., Thompson, L. and Oldewage-Theron, W. (2024). The role of artificial intelligence in nutrition research: A scoping review, *Nutrients* **16**(13): 2066.

Thames, Q., Karpur, A., Norris, W., Xia, F., Panait, L., Weyand, T. and Sim, J. (2021). Nutrition5k: Towards automatic nutritional understanding of generic food, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 8903–8911.

van den Brink, W. J., van den Broek, T. J., Palmisano, S., Wopereis, S. and de Hoogh, I. M. (2022). Digital biomarkers for personalized nutrition: Predicting meal moments and interstitial glucose with non-invasive, wearable technologies, *Nutrients* **14**(21): 4465.

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M. and Suez, J. (2015). Personalized nutrition by prediction of glycemic responses, *Cell* **163**(5): 1079–1094.