# Automatic Classification of High vs. Low Individual Nutrition Literacy Levels from Loyalty Card Data in Switzerland

Marc Blöchlinger
University of St. Gallen
St. Gallen, Switzerland
mbloechlinger@acm.org

Jing Wu
University of St. Gallen
St. Gallen, Switzerland
jing.wu@unisg.ch

Simon Mayer
University of St. Gallen
St. Gallen, Switzerland
simon.mayer@unisg.ch

Klaus L. Fuchs
ETH Zurich
Zurich, Switzerland
fuchsk@ethz.ch

Melanie Stoll
University Hospital of Bern
Bern, Switzerland
melanie.stoll@extern.insel.ch

Lia Bally
University Hospital of Bern
Bern, Switzerland
lia.bally@insel.ch

## ABSTRACT

The increasingly prevalent diet-related non-communicable diseases (NCDs) constitute a modern health pandemic. Higher *nutrition literacy* (NL) correlates with healthier diets, which in turn has favorable effects on NCDs. Assessing and classifying people's NL is helpful in tailoring the level of education required for disease self-management/empowerment and adequate treatment strategy selection. With recently introduced regulation in the European Union and beyond, it has become easier to leverage loyalty card data and enrich it with nutrition information about bought products. We present a novel system that utilizes such data to classify individuals into high- and low- NL classes, using well-known machine learning (ML) models, thereby permitting for instance better targeting of educational measures to support the population-level management of NCDs. An online survey (n = 779) was conducted to assess individual NL levels and divide participants into high- and low- NL groups. Our results show that there are significant differences in NL between male and female, as well as between overweight and non-overweight individuals. No significant differences were found for other demographic parameters that were investigated. Next, the loyalty card data of participants (n = 11) was collected from two leading Swiss retailers with the consent of participants and a ML system was trained to predict high or low NL for these individuals. Our best ML model, which utilizes the XGBoost algorithm and *monthly* aggregated baskets, achieved a Macro-$F_1$-score of .89 at classifying NL. We hence show the feasibility of identifying individual NL levels based on household loyalty card data leveraging ML models, however due to the small sample size, the results need to be further verified with a larger sample size.


*MADiMa '22, October 10, 2022, Lisboa, Portugal*

ACM ISBN 978-1-4503-9502-1/22/10…$15.00
https://doi.org/10.1145/3552484.3555744

## CCS CONCEPTS

• **Applied computing** → **Consumer health**; **Health informatics**; • **Computing methodologies** → *Supervised learning by classification.*

## KEYWORDS

Nutrition literacy; loyalty card data; machine learning algorithms

**ACM Reference Format:**
Marc Blöchlinger, Jing Wu, Simon Mayer, Klaus L. Fuchs, Melanie Stoll, and Lia Bally. 2022. Automatic Classification of High vs. Low Individual Nutrition Literacy Levels from Loyalty Card Data in Switzerland. In *Proceedings of the 7th International Workshop on Multimedia Assisted Dietary Management (MADiMa '22), October 10, 2022, Lisboa, Portugal.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3552484.3555744

## 1 INTRODUCTION

Diet-related non-communicable diseases (NCDs), such as obesity and associated comorbidities including type 2 diabetes, constitute a modern global health pandemic [31, 40]. This so-called *obesity pandemic* is associated with increased consumption of highly-processed energy-dense foods of low nutrition quality [27, 34]. In the World Health Organisation (WHO) European Region, overweight and obesity affect almost 60% of adults and nearly one third of children (29% of boys and 27% of girls) [57]. In Switzerland, 42% of the population aged 15 or above were overweight or obese as of 2017 [8]. The total economic cost related to overweight and obesity was estimated to be at 7.99 billion Swiss Francs (CHF) in 2012, noting that the costs have tripled over the past decade and are expected to continue rising [42]. The prevention of diet-related NCDs is paramount to reduce their burden on individuals and society.

Despite recommendations published by global organizations and national institutes (cf. [56] and [48]) that provide guidelines on healthier diets, current dietary patterns still follow a worrying trend. *Nutrition literacy* (NL)– "understanding nutrition information and acting on that knowledge in ways consistent with promoting nutrition goals" [5] – has been proposed to promote healthy dietary behavior [9, 19, 20, 55], which is believed to decrease the risks of diet-associated disorders.

Today, NL is commonly assessed through standardized validated questionnaires. However, they are limited by the precision of the assessment (due to biases such as recall bias) and can hardly be used for longitudinal or even permanent studies. With the introduction of the General Data Protection Regulation (GDPR) [51] and EU Regulation 1169/2011 [50], loyalty card data is becoming easier to collect and enrich with nutritional information. The objectivity and continuous data flow from loyalty cards would in principle make it a promising long-term NL assessment tool, provided that individual NL is reflected in – and can be inferred from – food purchasing behavior.

To this end, machine learning (ML) offers ways to process such datasets to "detect complex patterns without needing to specify them" [43], which is one of the reasons why it has been increasingly and oftentimes successfully been employed in health research [2, 18, 30, 37, 44].

In this article, we present a proof-of-concept that verifies the feasibility of using loyalty card data to predict individual NL levels leveraging ML models (Random Forest [7], Support Vector Machines [6], Naive Bayes, and Extreme Gradient Boosting [11]). We thereby provide a new approach to automatically assess individual NL levels and continuously monitor them at low cost. Our proof-of-concept provides the basis for larger studies to permit generalization of our results.

## 2 BACKGROUND AND RELATED WORK

### 2.1 Nutrition Literacy

The term *nutrition literacy* (NL) has been used inconsistently in the literature. The scientific community has long relied on *nutrition knowledge* to describe "individual's cognitive processes related to information about food and nutrition" [4]. In our context, we use the definition of NL based on Block et al. [5]: "[*Nutrition literacy*] entails both understanding nutrition information and acting on that knowledge in ways consistent with promoting nutrition goals".

The level of NL in individuals is commonly assessed through validated questionnaires, such as the *NLit Survey* [19, 53] and the *Short Food Literacy Questionnaire* (SFLQ) [14, 25]. Nevertheless, assessment effort is strongly increased in longitudinal or even permanent studies. Besides, individuals might get trained to answer the specific survey questions leading to an overestimation of their NL. Long-term NL assessment tools might be needed.

### 2.2 Food Purchasing Data

While the most prevalent source of dietary data today is self-reported, studies have for a long time used food purchasing data as an alternative [16]. One way to record and retrieve such food purchasing data is through loyalty card programs, thus in principle enabling *automated* collection of people's purchase history.

Switzerland provides an ideal context to conduct automatic loyalty card data collection: First, the two largest Swiss retailers occupy 35.1% and 34.8% of the Swiss retail market, respectively [46], and have 3.2 million and 2.8 million regular customers using their loyalty card systems, respectively [17, 32]. Second, because of the GDPR [51] that was introduced in 2016, companies are required to provide the data they collected to customers upon request. Although Switzerland is not an EU member state, companies in Switzerland still have to comply under certain conditions with the new law. This has given rise to data platforms such as BAM[1] that permit users to request, store, and share their personal data with third parties.

Compared to self-reported data, which takes manual effort to report and has limited accuracy due to under-reporting [12, 28, 39], conformity to social desirability [10], and recall bias [47], loyalty card data is objective and de-burdens subjects as participants only need to create accounts on relevant platforms and consent to terms of service. This permits longitudinal studies or even continuous interventions, while the required efforts using self-reporting are prohibitive [35].

### 2.3 Food Composition Database EatFit

The digital receipts on loyalty cards that are accessible in this study only contain information including article name, quantity, timestamp, supermarket location, price, and discount (if applicable). Therefore, further processing is required to enrich the digital receipts with nutritional information. This is supported by EU Regulation 1169/2011 [50], which mandates the publication of specific information (e.g., macro-nutrient content) about grocery items sold. However, the matching of items based on their article names such as "Zitrone" (lemon) still constitutes a challenge.

To this end, the study team has been building a food composition database, *EatFit*[2], which matches ambiguous article names on digital receipts to unique Global Trade Item Numbers (GTINs). On this basis, EatFit provides nutrition information of food products, including content of energy, total fat, saturated fat, carbohydrates, sugar, sodium, protein and dietary fiber per 100 gram product, and major (e.g., "Vegetables") and minor categories (e.g., "Fresh Vegetables") that the product belongs to. This makes it particularly suitable for studies using digital receipts from loyalty cards in Switzerland [29, 33, 58]. As of June 2022, EatFit comprises 53'780 products categorised in 21 major categories and 126 minor categories. The article names of the 6'472 most frequently purchased items have been matched to corresponding GTINs.

### 2.4 Dietary Pattern Assessment Methods

Some existing diet quality indices are commonly used to gain an overview of people's diet, such as the British Food Standards Agency Nutrient Profiling System Dietary Index (FSA-NPS DI) [23]. The FSA-NPS DI allocates points for "harmful" components and deducts points for "good" components, thereby creating a scale that ranges from -15 (healthiest) to +40 (unhealthiest) [24]. Energy-weighted means can be used to aggregate all rated food items. The FSA-NPS DI has been shown in multiple studies to validly express dietary quality, even based on digital receipts [3, 23, 24, 58] and is therefore also employed in this study. Expertise- and heuristics-based index design of these indices brings the benefit that dietary patterns are easily computable and comparable [38], but also introduces subjectivity in the index design.

In contrast to these hypothesis-oriented "a priori" diet quality indices, "a posteriori" methods, such as many ML algorithms, rely on statistical procedures to detect patterns from (dietary) data [38]. One advantage of ML is its ability to "detect patterns of variables that are 'diagnostic' of a particular outcome, over the conventional approach

[1] https://bitsabout.me
[2] https://eatfit-service.foodcoa.ch/static/swagger/, it requires log-in credentials.

of examining isolated, statistically independent relationships that are specified a priori" [43]. However, the most far-reaching implication of such algorithms is the trade-off between their ability to work with complex patterns and their "black box" nature of arriving at these conclusions [43]. How and why trained models weigh their input variables is sometimes - depending on the algorithms - hard to extract or interpret, which limits their usability for generalization or rule generation. Nevertheless, as long as the result but not the path to it is the primary goal, even such "black box" algorithms are viable tools to use.

## 3 METHODOLOGY

Ethics approval[3] was obtained for this study on July 06, 2020.

### 3.1 Participants

The project team recruited a convenience sample through various channels, predominantly by distributing leaflets with a quick response (QR) code linked to the survey at the metabolic outpatient clinic of the University Hospital Bern (Inselspital Bern)[4], online announcements, and word-of-mouth in Switzerland. To be eligible for the survey, participants needed to be German-speaking and at least 18 years old. The donation of food purchasing data additionally required participation in at least one of the two programs, i.e., Migros Cumulus and Coop Supercard.

The participant flow can be seen in Figure 1. In total, 779 participants completed the NL survey. The socio-demographic characteristics of these participants are summarized in Table 1. Due to the complexity of donating digital receipts and technological challenges related to the communication between the survey and the data collection platform, only 11 participants (8 female, 3 male) managed to finish the survey and donate their food purchasing data. There were 5 participants aged between 18-30, 1 aged between 31-39, 2 aged between 40-49 and 3 aged between 50-59. The median weight, height and BMI were 74.0 kilograms (kg) (interquartile range (*IQR*) = 14.5 kg), 170.0 centimeters (cm) (*IQR* = 4.5 cm) and 23.74 kilograms per square meter (kg/m²) (*IQR* = 6.88 kg/m²) respectively. Most (7) participants finished university or other tertiary, 1 finished vocational apprenticeship, 1 finished high school and 2 finished higher vocational education. None of these 11 participants self-reported to be diabetic, but 2 indicated themselves suffering from obesity and 2 from other chronic disease(s).

The purchase data of these 11 participants totalled in 1176 distinct baskets containing 13'299 items (mean basket size: 11 items, standard deviation (*SD*) = 12). Roughly 58.65% of all purchases were covered by EatFit. For the data to be used in the ML part, only baskets contained at least 5 identifiable items by EatFit were included to increase the expressiveness of basket compositions, thereby shrinking the number of baskets to 789 baskets (mean basket size: 13 items, *SD* = 10).

[3]Reference Number: Req-2020-00836
[4]https://www.insel.ch/de/

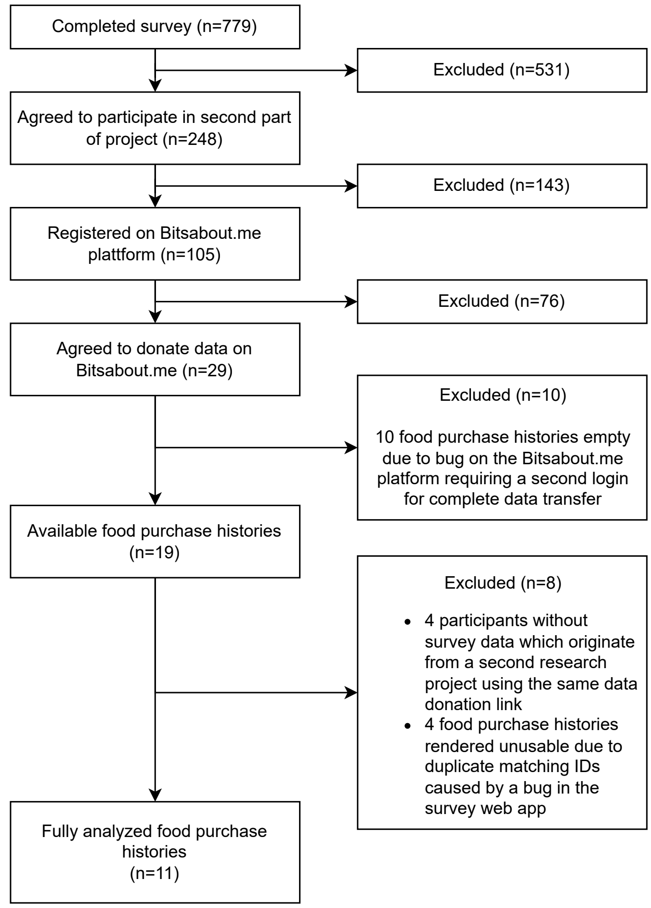


**Figure 1: Participant flow through the study**

### 3.2 Collection of Nutrition Literacy Survey Data and Loyalty Card Data

The data collection of the study consisted of two parts: completing a NL survey and donating loyalty card data from Migros Cumulus and/or Coop Supercard.

The NL survey can be found on our supplemental GitHub repository[5] and consisted of three parts. The first part collected sociodemographic information about the participants. Second, there were in total 6 questions where participants needed to judge the content of sugar, dietary fiber, protein, energy density, unsaturated fatty acids and salt in displayed packaged food items. Third, 8 random pictures out of a catalogue of 70 pictures of dishes on a plate were displayed. Here, the specific tasks were dependent on participants' diabetic statuses. Type 1 diabetic participants had to estimate the carbohydrate content in grams, because type 1 diabetic patients are supposed to do it before each meal in order to determine the appropriate dose of insulin. Participants with non-type-1 diabetes needed to select the carbohydrate content from a list of options (e.g., quinoa, egg and vegetables: How many grams of carbohydrates

[5]https://github.com/Interactions-HSG/2022madima

**Table 1: Demographics of survey participants**

| Characteristic | n | *Median* (*IQR*[1]) or % |
|---|---|---|
| **Age (years)** | 779 | |
| 18-30 | 259 | 33.25 % |
| 31-39 | 183 | 23.49 % |
| 40-49 | 148 | 19.00 % |
| 50-59 | 128 | 16.43 % |
| 60-69 | 45 | 5.78 % |
| 70+ | 16 | 2.05 % |
| **Gender** | 779 | |
| female | 582 | 74.71 % |
| male | 194 | 24.90 % |
| other | 3 | 0.39 % |
| **Weight (kg [2])** | 779 | 65.0 (17.0) |
| **Height (cm [3])** | 779 | 169.0 (11.0) |
| **BMI (kg/m² [4])** | 779 | 22.59 (4.23) |
| **Education level** | 779 | |
| Compulsory education | 12 | 1.54 % |
| Vocational apprenticeship | 122 | 15.66 % |
| Gymnasium - high school level | 73 | 9.37 % |
| Higher vocational education | 119 | 15.28 % |
| University/other tertiary | 453 | 58.15 % |
| **Diabetes status** | 779 | |
| Type 1 diabetes | 66 | 8.47 % |
| Type 2 diabetes | 12 | 1.54 % |
| Other type of diabetes | 6 | 0.77 % |
| None | 695 | 89.22 % |
| **Obesity & chronic diseases** | 779 | |
| Obesity | 57 | 7.32 % |
| Other chronic disease(s) | 79 | 10.14 % |
| None | 643 | 82.54 % |

[1] Interquartile range. [2] Kilogram.
[3] Centimeter. [4] Kilograms per square meter.

does this meal contain? Options: 26, 31, 36, 41). Non-diabetic participants needed to select the *energy content* from a list, as calorie control is the cornerstone of any weight-control strategy..

After finishing the survey, participants received a personalized sign-up link for the BAM platform via e-mail, if they agreed to participate in this part. Using the user-specific BAM links, participants could sign up, connect their loyalty cards, and share their loyalty card data to this project anonymously. Note-worthily, users needed to login again to BAM after creating the account to enable the data flow. This was not communicated to the study group in the beginning and caused some user attrition as shown in Figure 1. Once participants had completed these steps, their loyalty card data were shared with the study group via the BAM service.

## 3.3 Measures and Data

Participants (n = 779) were grouped into high- and low- NL groups, based on their performance in the NL survey. To counter the influence of different question designs as described in Sec. 3.2, the division into the high- and low- NL groups was done proportionally within each group with different diabetes statuses. For non-type-1 diabetic and non-diabetic participants, the NL levels were determined based on the combined share of correct answers from both the packaged foods section and the food on a plate section of the survey. For instance, if a participant correctly answered 3/6 of the questions about the packaged foods and 2/8 of the questions about the plates, he/she correctly answered (3+2)/(6+8) = 35.71% of the questions. For type 1 diabetic participants, it was solely dependent on their carbohydrate estimation errors in the food on a plate section. The smaller the error, the more literate in carbohydrate counting.

To understand the differences between the high- and low- NL groups, the following metrics were computed. First, the adherence to the WHO guidelines regarding the intake of sugar, saturated fats, and sodium was indicated by the nutrient content in grams per 1'000 kilocalories (g/1'000kcal) of purchased food. Next, based on the entire shopping history, the FSA-NPS DI of participants were calculated, indicating the food shopping healthiness.

Three different basket-level nutritional profiles were then calculated to be used as input for the ML model training. These profiles were compositional profiles of the energy share in kcal per 1'000 kcal (kcal/1000kcal) of major food categories, minor food categories, and nutrients, based on *single* baskets, *monthly* baskets and *quarterly* baskets. Baskets were identified by the user id and timestamp and only included those with at least 5 items identified by the EatFit database, because baskets with more items are expected to be more expressive of people's purchasing habits. To accommodate for possible basket profile distortions due to infrequent items (e.g., cooking oils), baskets were additionally grouped by month and quarter.

## 3.4 Statistical Analysis

Prior to conducting any other statistical tests, we always verified whether the data was normally distributed by D'Agostino's normality tests when the number of samples in a given group was 50 or above, or by Shapiro-Wilk's tests when sample sizes were smaller.

The first part of the analysis only concerned the data from the NL surveys (n=779). We explored the differences in NL scores between groups with different socio-demographic factors, such as age and gender. Comparing two resulting groups, we used Welch's t-tests for parametric data, and Mann–Whitney U tests for non-parametric data. With more than two groups, Welch's one-way analysis of variance (Welch's ANOVA) was used. If groups with small sample sizes (n<25) exhibited non-normally distributed data, a Kruskal-Wallis H test was applied instead. In case of ANOVA or Kruskal–Wallis H test revealing significant differences between groups, post-hoc analysis was subsequently conducted using Bonferroni-adjusted Dunn's test to identify the significantly different groups.

The correlation between BMI and NL scores was assessed by calculating Pearson's correlation coefficient. Spearman's rank correlation coefficient was used alternatively if data was not normally distributed. We used Pearson's $\chi^2$ tests to analyze the differences in proportion of correct answers for specific questions in the second part of the survey (see Sec. 3.2).

In the second part of data analysis, the food purchasing data of the 11 participants and its associations with NL were analyzed. Welch's t-tests were used to compare the FSA-NPS DI, as well as energy-adjusted intake of sugar, sodium, and saturated fats respectively, of participants with high- and low- NL levels. Additionally, multiple linear regression was used to investigate the correlation between individual socio-economic variables and the FSA-NPS DI scores while adjusting for NL levels. Only BMI and gender were used as predictive variables, because using other variables, such as educational level and age, will lead to too imbalanced groups with sample size of 1.

Lastly, the relative food category popularity in high- and low- NL groups was assessed by the energy contribution to total purchased energy (in kcal per 1'000 kcal). To efficiently identify significant differences between high- and low- NL groups without verifying normality, a Mann-Whitney U test was applied.

## 3.5 Prediction Models

The calculated basket-level dietary profiles as described in Sec. 3.3 of n = 11 participants with food purchase histories were used to train multiple ML models with the objective of predicting individual NL levels, as indicated by the survey. Due to the small sample size, no prediction based on complete participants' dietary profiles, i.e., comparing entire purchase histories, was feasible. Based on their popularity in the literature, Random Forests (RF), Support Vector Machines (SVM), Naive Bayes (NB), and Extreme Gradient Boosting (XGBoost) were chosen as the algorithms to compare [36, 52].

The input dataset was divided into train (80%) and test (20%) data randomly with stratification. After the train-test-split, all features were standardized to zero mean and unit variance. Because of the higher ratio (8 out of 11) of high NL participants, the training data was artificially balanced through random oversampling. During the training phase of each algorithm, 5-fold cross validation was used to minimize over-fitting [22]. Hyper-parameter tuning by using grid search was applied during the training phase to try and find the optimally performing parameters for each model[6].

The primary metric to assess the performance of the models was the macro-averaged $F_1$ score, which considers both precision and recall and is particularly suited for imbalanced datasets [15]. This is necessary as the testing set is not randomly over-sampled like the training set, and a simple majority-class-predicting model would otherwise still achieve a high score.

# 4 RESULTS

## 4.1 Nutrition Literacy Survey Performance

Regarding the first 6 questions about packaged foods (see Sec. 3.2), the mean of correct questions was 42.54% (*SD* = 18.22%). Overall, female participants achieved higher scores (*mean* = 44.3%, *SD* = 17.75%) than male participants (*mean* = 37.5%, *SD* = 18.56%), $t(315.14) = 4.43$, $p < .001$. Also, women showed significantly higher NL scores than men in estimating sugar content, $\chi^2(2, N = 725) = 7.51$, $p = .023$, estimating energy density, $\chi^2(2, N = 735 = 6.22$, $p = .045$, and estimating sodium content, $\chi^2(2, N = 761) = 10.63$, $p = .005$. Welch's ANOVA showed a statistically significant difference in scores between education levels, $F(4,71.17) = 4.16$, $p = .004$. Subsequent post-hoc analysis with Bonferroni-adjusted Dunn's test indicated significantly higher scores for the "higher vocational education" group (*mean* = 45.94%, *SD* = 17.08%) compared to the "vocational apprenticeship" group (*mean* = 38.53%, *SD* = 17.94%), $p = .017$. Overweight participants (BMI ≥ 25) achieved lower scores (*median* = 33.33%, *IQR* = 16.67%) than non-overweight participants (*mean* = 43.44%, *SD* = 17.98%). A Mann-Whitney U test revealed this difference to be statistically significant, $U(N_{\text{overweight}} = 196, N_{\text{non-overweight weight}} = 575) = 49'055.0$, $p = .005$. Similarly, obese participants (BMI ≥ 30) reached lower scores (*median* = 33.33%, *IQR* = 29.17%) than non-obese participants (*median* = 50.00%, *IQR* = 16.67%), $U(N_{\text{obese}} = 58, N_{\text{non-obese}} = 713) = 17'358.5$, $p = .034$. Differences in NL scores between other characteristics of participants, namely age groups, diabetic *vs.* non-diabetic, and diabetes types failed to show significance at the $p < 0.05$ threshold. Overweight (obese) participants showed a significantly lower proportion of correct answers to sodium content estimating questions in comparison to non-overweight (non-obese) participants, $\chi^2(2, N = 764) = 6.37$, $p = .041$ ($\chi^2(2, N = 764) = 16.46$, $p < .001$). Diabetic participants scored significantly higher in the question about protein content, $\chi^2(2, N = 754) = 12.054$, $p = .002$. There was no significant difference in correct answers per question between different education levels, age groups, or different types of diabetes. Spearman's rank correlation showed that BMI and the ratios of correct answers had small negative correlations, $r_s = -0.12$, $p = .034$, $N = 779$.

In the third part of the NL survey (see Sec. 3.2), 33.33% (*SD* = 17.96%) and 25.0% (*SD* = 20.56%) of the questions were correctly answered by non-diabetic and non-type-1 diabetic participants accordingly. The average estimation error was 61.77% (*SD* = 47.66%) for type 1 diabetic participants.

## 4.2 Difference in Food Purchasing Healthiness between High- and Low- NL Levels

Table 2 shows detailed characteristics of the food purchasing data of all 11 participants of the second project phase and Table 3 summarizes the nutrient breakdown of the same data.

There was no significant difference between mean FSA-NPS DI of participants with high (*mean* = 5.01, *SD* = 2.92) and low NL (*mean* = 6.74, *SD* = 0.31) levels. The energy-adjusted sodium content among purchased items (in g/1'000 kcal) was lower for participants with low NL (*mean* = 1.12, *SD* = 0.36) compared to the participants with high NL level (*mean* = 3.12, *SD* = 1.72), $t(8.63) = -2.95$, $p = .020$. There was no statistically significant difference in energy-adjusted sugar or saturated fat content between the two groups. The relative amount of energy (in kcal/1'000 kcal) from items in the major food category "meat and sausages" was significantly lower in the high NL group (*median* = 46.82, *IQR* = 33.66) compared to the low NL group (*median* = 97.79, *IQR* = 14.84), $U(N_{\text{high NL}} = 9, N_{\text{low NL}} = 3) = 1.0$, $p = .024$.

Comparing the relative amount of energy (in kcal/1'000 kcal) from minor food categories and applying Mann-Whitney U tests revealed three significant differences. First, the relative amount of energy from "butter and margarine" was lower in the high NL group (*median* = 43.96, *IQR* = 46.89) compared to the low NL group (*median* = 145.67, *IQR* = 49.09), $U(N_{\text{high NL}} = 7, N_{\text{low NL}} = 3) = 1$, $p = .033$. Second, the relative amount of energy from "legumes" was significantly lower as well in the high NL group (*median* = 5.30, *IQR* = 4.00) compared to the other group (*median* = 10.44, *IQR* =

[6]The final parameter settings are available from the code at https://github.com/Interactions-HSG/2022madima/blob/main/code/3_ml_model.ipynb

**Table 2: Characteristics of food shopping data for participants with high (n = 8) and low (n = 3) *nutrition literacy* (NL) levels**

| Characteristics | *Mean* (*SD*),[1] *Median* (*IQR*)[2]*, or % |
|---|---|
| **Purchased items** [3] | |
| Total | 1'209 (1'139) |
| Low NL | 1'112 (780) |
| High NL | 1'245 (1'246) |
| **Items per basket** | |
| Total* | 8 (11) |
| Low NL* | 8 (13) |
| High NL* | 8 (11) |
| **Share of vegetables in items (%)** | |
| Total* | 16.85 (7.80) |
| Low NL | 22.08 (7.18) |
| High NL* | 17.44 (3.94) |
| **Share of fruits in items (%)** | |
| Total | 12.42 (4.65) |
| Low NL | 10.82 (6.50) |
| High NL | 13.02 (4.17) |
| **Purchased energy (kcal)**[4] | |
| Total | 654'793.53 (549'434.11) |
| Low NL | 711'982.66 (429'581.25) |
| High NL* | 422'476.95 (572'589.68) |
| **Energy per basket (kcal)*** | |
| Total* | 3'459.63 (6'982.35) |
| Low NL* | 5'269.94 (7'473.71) |
| High NL* | 2'959.70 (6'389.13) |
| **Energy share of vegetables (%)** | |
| Total* | 4.39 (3.28) |
| Low NL | 4.33 (0.52) |
| High NL* | 3.31 (3.31) |
| **Energy share of fruits (%)** | |
| Total* | 6.94 (6.49) |
| Low NL | 5.50 (3.61) |
| High NL | 7.98 (4.17) |
| **FSA-NPS DI** | |
| Total | 5.49 (2.58) |
| Low NL | 6.74 (0.31) |
| High NL | 5.01 (2.92) |

[1] Standard deviation. [2] Interquartile range.
[3] All items were purchased from Feburary 02 2018 to August 24 2021. [4] Kilocalories.
* The *median* (*IQR*) is reported for non-normally distributed data.

3.11), $U(N_{\text{high NL}} = 6, N_{\text{low NL}} = 3) = 1$, $p = .048$. Third, the high NL participants purchased significantly more relative energy from "yogurt and sour milk" (*median* = 26.41, *IQR* = 20.53), compared to low NL participants (*median* = 5.81, *IQR* = 1.65), $U(N_{\text{high NL}} = 8, N_{\text{low NL}} = 3) = 24$, $p = .012$.

**Table 3: Nutrient breakdown of food shopping data of participants with high and low *nutrition literacy* (NL) levels**

| Nutrient (g/1'000kcal)[1] | Total (n = 11) *Mean* (*SD*)[2] | High NL (n = 8) *Mean* (*SD*)[2] | Low NL (n = 3) *Mean* (*SD*)[2] |
|---|---|---|---|
| Carbohydrates | 108.7 (16.39 ) | 113.32 (16.57) | 96.37 (8.27) |
| Sugar | 48.32 (16.96) | 52.55 (17.93) | 37.03 (7.00) |
| Protein | 37.05 (5.08) | 38.13 (5.58) | 34.17 (1.76) |
| Fat | 46.21 (9.25) | 43.64 (9.43) | 53.07 (4.44) |
| Saturated fat | 20.29 (4.94) | 18.67 (4.50) | 24.64 (3.52) |
| Sodium | 2.57 (1.81) | 3.12 (1.84) | 1.12 (0.45) |
| Dietary fiber | 3.49 (1.16) | 3.61 (1.33) | 3.17 (0.60) |

[1] Grams per 1000 kilocalories. [2] Standard deviation.

## 4.3 Model Performance

Overall, the best ML model at classifying participants' NL levels based on their food shopping data was XGBoost using minor food category composite characteristics trained on *monthly* aggregated baskets, judged by the macro-averaged $F_1$-score (*Macro-$F_1$-score* = 0.89, *Recall* = 0.85, *Precision* = 0.95, *Accuracy* = 0.92). A comparison of each model's testing performance, based on different types of input data (major / minor food category popularity or nutrient composition in the baskets), is shown in Figure 2.

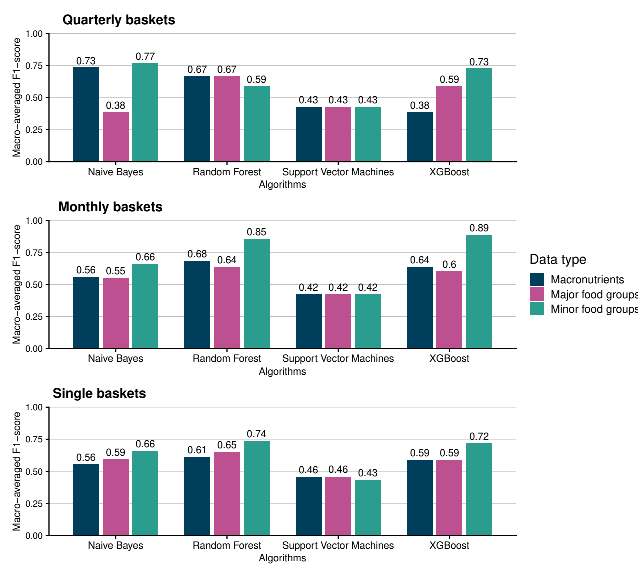


**Figure 2: Performance comparison of all tested machine learning algorithms with different dietary profiles and basket averaging time-frames as inputs. XGBoost on monthly basket data derived from minor food categories performed best and achieved a Macro-$F_1$-score of 0.89**

For basket data aggregated on *quarterly* basis, NB trained on minor food category characteristics performed the best (*Macro-$F_1$-score* = 0.77, *Recall* = 0.79, *Precision* = 0.75, *Accuracy* = 0.74). RF achieved an identical $F_1$-score of 0.67 here using both major food category characteristics and nutrient compositions as types of input data (*Recall* = 0.67, *Precision* = 0.67, *Accuracy* = 0.75). The best XGBoost implementation on *quarterly* baskets reached a $F_1$-score of 0.73 (*Recall* = 0.71, *Precision* = 0.76, *Accuracy* = 0.81). Lastly, scores of SVM were identical across all three input data types (*Macro-$F_1$-score* = 0.43, *Recall* = 0.5, *Precision* = 0.38, *Accuracy* = 0.75).

With *monthly* aggregated baskets, the best version of RF's macro-averaged $F_1$-score was 0.85 and used features of minor food category characteristics as input as well (*Recall* = 0.83, *Precision* = 0.89, *Accuracy* = 0.89). NB's best-performing version for *monthly* basket data reached a macro-averaged $F_1$-score of 0.66 and used minor food category characteristics as predictive features (*Recall* = 0.66, *Precision* = 0.66, *Accuracy* = 0.74). SVM achieved identical scores across all three different characteristic input data sets (*Macro-$F_1$-score* = 0.42, *Recall* = 0.5, *Precision* = 0.37, *Accuracy* = 0.74).

RF scored the highest at classifying based on *single (non-averaged)* baskets, reaching a Macro-$F_1$-score of 0.75 (*Recall* = 0.71, *Precision* = 0.79, *Accuracy* = 0.84). The best model of XGBoost using single baskets attained a macro-averaged $F_1$-score of 0.83 (*Recall* = 0.69, *Precision* = 0.79, *Accuracy* = 0.84). NB's best-performing version for this type of basket data reached a macro-averaged $F_1$-score of 0.7 (*Recall* = 0.71, *Precision* = 0.66, *Accuracy* = 0.70). The best performing version of SVM used major food category characteristics as input data and achieved a macro-averaged $F_1$-score of 0.46.

# 5 DISCUSSION

## 5.1 Interpretation of Results

The primary goal of this article was to verify the feasibility of estimating people's NL levels based on food shopping data using ML models. In total, 779 participants finished the survey but only 11 of them donated their food purchasing data. Therefore, utmost caution is suggested when it comes to the interpretation of results based on the food shopping data.

The results indicate that NL levels can potentially be binarily classified using ML models based on fragmentary data such as aggregated basket compositions (macro-averaged $F_1$-score = 0.89) or even single basket compositions (macro-averaged $F_1$-score = 0.75). It needs to be further verified with a larger sample size. The only significant difference between the food purchasing behavior of high- and low- NL participants, besides a few differences in food category popularity, was in energy-adjusted sodium content (in g/1'000 kcal), where the high NL participants purchased more sodium, counter-intuitively. The analysis of the survey data (n = 779) revealed that gender, weight status (overweight or not, obese or not), and BMI are factors that seem to be significantly related to NL. These results provide evidence that food shopping data could be a viable tool to assess NL using ML models among the customers of loyalty programs and identify consumers with high- and low-NL levels.

*Survey Performance* Female participants scored higher in the second part of the survey than their male counterparts, indicating a higher NL level in regard to the areas covered by those questions. This difference is in turn also present in one specific question in the third part of the survey, where the share of correct answers at identifying the most energy-dense item was significantly higher in female participants. These results are in line with previous findings that found women to score significantly higher than men in NL surveys [9, 25, 54].

Only between two of the five present education levels ("vocational apprenticeship" scoring lower than "higher vocational education") could a difference be detected, which were not the two levels that were the furthest apart. Although some past studies have found a relationship between educational level and NL [9, 19, 54], there is also a study [25] that fails to find a disparity, but offers a potential explanation based on the skewed distribution of education levels in their sample population. A similar bias is present in this study, where more than half of the participants had a tertiary education. This might explain the lack of more pronounced differences.

Overweight participants had a lower NL score than non-overweight participants. Lower NL might, as the literature suggests, result in a worse diet with excessive energy intake [9, 19, 45]. The negative correlation between BMI and NL has been found in multiple previous studies as well [13, 19, 49]. Diabetic participants also had a lower NL score compared to non-diabetic participants. Similar differences with diabetic participants scoring lower in NL were obtained in other studies [19, 49], although in the latter of the two studies the trend was only nearly significant (p = 0.06). Contrarily, age did not appear to significantly correlate with NL, though some previous studies indicating age to be a significant factor in NL [19, 54].

Overall, these results were mostly in line with previous studies that examined factors that appear to influence NL. Our understanding of potential risk factors of low NL can be deepened and validated by such consistent results, with the ultimate goal of providing a directive of where and how to improve NL.

*Difference in Food Purchasing between Groups with High and Low Nutrition Literacy Levels* This part of discussion only regards the 11 participants who donated their loyalty card data. The lack of a correlation between NL levels and dietary quality was inconsistent with a previous study, where a significant association between NL and higher dietary quality was found [9]. While the mean FSA-NPS DI in this study was (insignificantly) lower in the high NL group (indicating a better diet), the SD was very large in this group. The small sample size and the limitations of loyalty card data (see Sec. 5.2) might explain the non-significance in the statistical tests.

Only purchased energy-adjusted sodium purchase showed a correlation with NL. Participants with low NL purchased less sodium (in g/1'000kcal) compared to those with high NL. As the WHO recommends to limit sodium consumption to 2g per day [56], a lower sodium amount signifies stronger adherence to these guidelines. Hence, it is counter-intuitive that the low NL group adhered more strongly to this recommendation. One possible explanation is that pure salt items, such as containers of table salt, were included in the analysis. Participants with high NL purchased three times as many items containing "Salz" ("salt") in their names as in another group, which might increase the purchased sodium content substantially. There is no association that links NL with sodium intake in the Swiss population [26]. Similarly, a literature review found conflicting results regarding the link between knowledge about salt and intake of it [41]. It might be because that awareness about sodium recommendations itself is the first step, but acting on it appropriately is trickier.

Examining the differences in food category popularity relative to purchased energy revealed four differences: Fewer energy was purchased from minor categories "meat and sausages", "butter and margarine", "legumes", and more energy from "yogurt and sour milk" by participants with high NL. Especially the lower amount of purchased meat is interesting, with the high NL group buying nearly half as much energy from meats and sausages. This could potentially be explained by environmental or health-related concerns. A higher NL could mean that these individuals are more capable of understanding the indirect consequences of their dietary choices and acting on such concerns.
No significant difference in the energy from fruits and vegetables could not be found in the two groups with different NL levels. The high NL group purchased less "legumes", which are generally considered as healthy, compared to the low NL group. This is contrary to previous findings that the intake of fruit and vegetables significantly correlate with participants' nutrition knowledge [45]. If correct, this may indicate that the difference in the consumption of these healthy foods between the NL levels may not be as clear as presumed , or the sample size was too small.

*Model Performance in Predicting Nutrition Literacy* The prediction of NL using ML models applied to food shopping data is a novelty. The results are therefore not directly comparable to the existing literature. Overall, there was no clear pattern of any given algorithm or input data type strictly dominating all the others. Still, it seems food profiles based on minor food categories were the best basis for the most successful prediction models. This could possibly be explained by the extra granularity that minor categories bring compared to major categories.
The best model used *monthly* baskets, instead of *quarterly* or *single* baskets as input. This might be explained by the trade-off between basket timeframe and number of available samples.

## 5.2 Limitations and Outlook

We highlight the following limitations that affect the general interpretability and validity of the results provided here and offer an outlook.
First, the major limitation of the study comes from the small number of participants who successfully donated their food purchasing data (N = 11). The tedious and complex process of donating the loyalty card data (see Sec. 3.2) can be the main cause of the high attrition rate from survey participation to complete data donation (see Fig. 1). With no strong enough intrinsic motivation or value to gain from participation, participants tend not to power through such a process. Improvements here would arguably enhance sample sizes in future studies, reduce dropout rates, and potentially move towards creating a representative sample of the broader Swiss population.
Second, loyalty card data has intrinsic limitations, as discussed in previous studies [33, 58]. The two most predominant ones are a) loyalty card data does not equal dietary intake or even food shopping. The coverage is up to the frequency of shopping outside Migros/Coop and the loyalty card usage frequency b) loyalty cards are oftentimes shared among household members. However, the NL is an individual-level topic. Although we used energy-adjusted measures (e.g. sodium in g/1'000kcal) to counter the influences, there are still gaps in between.
Third, we can only identify around 58.65% of food products with EatFit in the study. Not different from most food databases [21], Eatfit can potentially (most likely) contain inaccuracies. These might twist the food purchasing behavior of participants drastically. The study team has been dedicated to improving the data quality of the EatFit database to narrow the gaps. More robust results can be expected with improved data quality.
Fourth, the presented study only classifies participants into high vs. low NL levels. Our results suggest that a larger sample size should permit more nuanced segmentation of NL levels. Nonetheless, even a binary classification is helpful in tailoring the level of education required for disease self-management/empowerment and selecting adequate treatment strategies and goals.
Last, the NL survey questions were primarily designed to assess the quantitative nutritional component estimation skills of participants, to evaluate the need for automated nutrient assessment applications. However, assessing NL using not validated tools is a recurring theme in the literature [45]. In addition, responses in such surveys are heavily influenced by factors like participant literacy, which correlates with education and socio-economic status [1]. This influences the accuracy of estimated NL. Nonetheless, this is a general limitation of survey design and out of the sphere of this study.
As this proof-of-concept study is to verify the feasibility of using ML models to predict NL levels, these limitations should not restrain the future prospects and implications of this exploratory study that much. We expect more solid results with a larger sample size and better data quality in the future.
If, as the preliminary results suggest, it is possible to interpolate NL based on loyalty card data, it could open new possibilities in health surveillance which can directly be applied to customers of food retailers. The ease and availability of interventions based on such large datasets would be remarkable. However, such hypothetical programs are still far away and more research is needed to explore the relationship between NL and food purchasing data and verify the feasibility of predicting NL based on such data.

## 5.3 Conclusion

This article discusses a proof-of-concept to classify individuals' NL levels based on their loyalty card data using ML algorithms, with a small sample size of 11. It proposes a new way of identifying individual NL levels and could help target those with low NL to improve their NL and then their health status. Although the generality of the current results must be validated by future research, the present study has shown that this is indeed a feasible approach. Furthermore, this survey data (n = 779) validates existing literature about the differences in NL between different genders and overweight statuses (overweight or non-overweight).

# 6 ACKNOWLEDGMENTS

We would like to acknowledge the Swiss National Science Foundation for supporting this research under grant #188402, *FoodCoach*[7].

[7] https://p3.snf.ch/project-188402

# REFERENCES


[1] Robert J Adams, Sarah L Appleton, Catherine L Hill, Mark Dodd, Christopher Findlay, and David H Wilson. 2009. Risks associated with low functional health literacy in an Australian population. *Medical Journal of Australia* 191, 10 (2009), 530–534. https://doi.org/10.5694/j.1326-5377.2009.tb03304.x

[2] Luca Maria Aiello, Rossano Schifanella, Daniele Quercia, and Lucia Del Prete. 2019. Large-scale and high-resolution analysis of food purchases and health outcomes. *EPJ Data Science* 8, 14 (2019), 1–22. https://doi.org/10.1140/epjds/s13688-019-0191-y

[3] Carukshi Arambepola, Peter Scarborough, and Mike Rayner. 2008. Validating a nutrient profile model. *Public health nutrition* 11, 4 (2008), 371–378. https://doi.org/10.1017/S1368980007000377

[4] Marta L Axelson and David Brinberg. 1992. The measurement and conceptualization of nutrition knowledge. *Journal of Nutrition Education* 24, 5 (1992), 239–246. https://doi.org/10.1016/S0022-3182(12)81238-6

[5] Lauren G Block, Sonya A Grier, Terry L Childers, Brennan Davis, Jane EJ Ebert, Shiriki Kumanyika, Russell N Laczniak, Jane E Machin, Carol M Motley, Laura Peracchio, et al. 2011. From nutrients to nurturance: A conceptual introduction to food well-being. *Journal of Public Policy & Marketing* 30, 1 (2011), 5–13. https://doi.org/10.1509/jppm.30.1.5

[6] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. 1992. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*. Association for Computing Machinery, New York, NY, USA, 144–152. https://doi.org/10.1145/130385.130401

[7] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32. https://doi.org/10.1023/A:1010933404324

[8] Bundesamt für Statistik (BFS). 2020. Schweizerische Gesundheitsbefragung 2017: Übergewicht und Adipositas. Retrieved July 1, 2022 from https://www.bfs.admin.ch/asset/de/14147720

[9] Elise Carbonneau, Benoît Lamarche, Véronique Provencher, Sophie Desroches, Julie Robitaille, Marie-Claude Vohl, Catherine Bégin, Mathieu Bélanger, Charles Couillard, Luc Pelletier, et al. 2021. Associations Between Nutrition Knowledge and Overall Diet Quality: The Moderating Role of Sociodemographic Characteristics—Results From the PREDISE Study. *American Journal of Health Promotion* 35, 1 (2021), 38–47. https://doi.org/10.1177/0890117120928877

[10] Jacopo Cerri, John Thøgersen, and Francesco Testa. 2019. Social desirability and sustainable food research: A systematic literature review. *Food Quality and Preference* 71 (2019), 136–140. https://doi.org/10.1016/j.foodqual.2018.06.013

[11] Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. Association for Computing Machinery, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785

[12] Nikhil V Dhurandhar, Dale Schoeller, Andrew W Brown, Steven B Heymsfield, Diana Thomas, Thorkild IA Sørensen, John R Speakman, Madeline Jeansonne, and David B Allison. 2015. Energy balance measurement: when something is not better than nothing. *International journal of obesity* 39, 7 (2015), 1109–1113. https://doi.org/10.1038/ijo.2014.199

[13] James J Diamond. 2007. Development of a reliable and construct valid measure of nutritional literacy in adults. *Nutrition journal* 6, 1 (2007), 1–4. https://doi.org/10.1186/1475-2891-6-5

[14] Hasan Durmus, Mehmet Enes Gökler, and Suzan Havlioğlu. 2019. Reliability and validity of the Turkish version of the short food literacy questionnaire among university students. *Prog Nutr* 21, 2 (2019), 333–338. https://doi.org/10.23751/pn.v21i2.7094

[15] Alberto Fernández, Salvador García, Mikel Galar, Ronaldo C Prati, Bartosz Krawczyk, and Francisco Herrera. 2018. *Learning from imbalanced data sets*. Vol. 10. Springer Cham. https://doi.org/10.1007/978-3-319-98074-4

[16] Ronald E Frank. 1967. Correlates of Buying Behavior for Grocery Products. *Journal of marketing* 31, 4 (1967), 48–53. https://doi.org/10.1177/002224296703100410

[17] Coop Genossenschaft. 2021. Unternehmensprofil der Coop-Gruppe 2021. Retrieved July 1, 2022 from https://report.coop.ch/app/uploads/Coop_Unternehmensprofil_2021_DE.pdf

[18] Philippe J Giabbanelli and Jean Adams. 2016. Identifying small groups of foods that can predict achievement of key dietary recommendations: data mining of the UK National Diet and Nutrition Survey, 2008–12. *Public health nutrition* 19, 9 (2016), 1543–1551. https://doi.org/10.1017/S1368980016000185

[19] Heather D Gibbs, Edward F Ellerbeck, Byron Gajewski, Chuanwu Zhang, and Debra K Sullivan. 2018. The Nutrition Literacy Assessment Instrument (NLit) is a valid and reliable measure of nutrition literacy in adults with chronic disease. *Journal of nutrition education and behavior* 50, 3 (2018), 247–257. https://doi.org/10.1016/j.jneb.2017.10.008

[20] Eun-Jeong Ha and Natalie Caine-Bish. 2009. Effect of nutrition intervention using a general nutrition course for promoting fruit and vegetable consumption among college students. *Journal of nutrition education and behavior* 41, 2 (2009), 103–109. https://doi.org/10.1016/j.jneb.2008.07.001

[21] G.G. Harrison. 2004. Fostering Data Quality in Food Composition Databases: Applications and Implications for Public Health. *Journal of Food Composition and Analysis* 17, 3-4 (2004), 259–265. https://doi.org/10.1016/j.jfca.2004.03.014

[22] Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. 2009. *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. Springer, New York, NY. https://doi.org/10.1007/978-0-387-84858-7

[23] Chantal Julia, Caroline Méjean, Mathilde Touvier, Sandrine Péneau, Camille Lassale, Pauline Ducrot, Serge Hercberg, and Emmanuelle Kesse-Guyot. 2016. Validation of the FSA nutrient profiling system dietary index in French adults—Findings from SUVIMAX study. *European journal of nutrition* 55, 5 (2016), 1901–1910. https://doi.org/10.1007/s00394-015-1006-y

[24] Chantal Julia, Mathilde Touvier, Caroline Méjean, Pauline Ducrot, Sandrine Péneau, Serge Hercberg, and Emmanuelle Kesse-Guyot. 2014. Development and validation of an individual dietary index based on the British Food Standard Agency nutrient profiling system in a French context. *The Journal of nutrition* 144, 12 (2014), 2009–2017. https://doi.org/10.3945/jn.114.199679

[25] Corinna Gréa Krause, Sigrid Beer-Borst, Kathrin Sommerhalder, Stefanie Hayoz, and Thomas Abel. 2018. A short food literacy questionnaire (SFLQ) for adults: Findings from a Swiss validation study. *Appetite* 120 (2018), 275–280. https://doi.org/10.1016/j.appet.2017.08.039

[26] Xhyljeta Luta, Stefanie Hayoz, C Gréa Krause, K Sommerhalder, E Roos, P Strazzullo, and S Beer-Borst. 2018. The relationship of health/food literacy and salt awareness to daily sodium and potassium intake among a workplace population in Switzerland. *Nutrition, Metabolism and Cardiovascular Diseases* 28, 3 (2018), 270–277. https://doi.org/10.1016/j.numecd.2017.10.028

[27] Mirko Marino, Federica Puppo, Cristian Del Bo', Valentina Vinelli, Patrizia Riso, Marisa Porrini, and Daniela Martini. 2021. A systematic review of worldwide consumption of ultra-processed foods: findings and criticisms. *Nutrients* 13, 8 (2021), 2778. https://doi.org/10.3390/nu13082778

[28] Marianne S Markussen, Marit B Veierød, Giske Ursin, and Lene F Andersen. 2016. The effect of under-reporting of energy intake on dietary patterns and on the associations between dietary patterns and self-reported chronic disease in women aged 50–69 years. *British Journal of Nutrition* 116, 3 (2016), 547–558. https://doi.org/10.1017/S000711451600218X

[29] Simon Mayer, Klaus Fuchs, Dominic Brügger, Jie Lian, and Andrei Ciortea. 2021. Improving customer decisions in web-based e-commerce through guerrilla modding. *Nature Machine Intelligence* 3, 12 (2021), 1008–1010. https://doi.org/10.1038/s42256-021-00424-w

[30] Susan E McCann, James R Marshall, John R Brasure, Saxon Graham, and Jo L Freudenheim. 2001. Analysis of patterns of food intake in nutritional epidemiology: food classification in principal components analysis and the subsequent impact on estimates for endometrial cancer. *Public health nutrition* 4, 5 (2001), 989–997. https://doi.org/10.1079/PHN2001168

[31] David R Meldrum, Marge A Morris, and Joseph C Gambone. 2017. Obesity Pandemic: Causes, Consequences, and Solutions—but Do We Have the Will? *Fertility and Sterility* 107, 4 (2017), 833–839. https://doi.org/10.1016/j.fertnstert.2017.02.104

[32] Migros-Genossenschafts-Bund. 2017. Cumulus: Die beliebteste Kundenkarte der Schweiz feiert ihren 20. Geburtstag. https://corporate.migros.ch/de/medien/mitteilungen/show/news/medienmitteilungen/2017/cumulus-20-geburtstag~id=0015b299-a5e4-4957-9c2a-f140a91fb402~.html

[33] Annette Mönninghoff, Klaus Fuchs, Jing Wu, Jan Albert, and Simon Mayer. 2021. The Effect of a Future-Self Avatar mHealth Intervention on Physical Activity and Food Purchases: The FutureMe Randomized Controlled Trial. (2021).

[34] Carlos A Monteiro, Geoffrey Cannon, Renata B Levy, Jean-Claude Moubarac, Maria LC Louzada, Fernanda Rauber, Neha Khandpur, Gustavo Cediel, Daniela Neri, Euridice Martinez-Steele, et al. 2019. Ultra-processed foods: what they are and how to identify them. *Public health nutrition* 22, 5 (2019), 936–941. https://doi.org/10.1017/S1368980018003762

[35] Jaakko Nevalainen, Maijaliisa Erkkola, Hannu Saarijärvi, Turkka Näppilä, and Mikael Fogelholm. 2018. Large-scale Loyalty Card Data in Health Research. *Digital health* 4 (2018), 1 – 10. https://doi.org/10.1177/2055207618816898

[36] Didrik Nielsen. 2016. *Tree boosting with xgboost-why does xgboost win" every" machine learning competition?* Master's thesis. Norwegian University of Science and Technology. Retrieved July 1, 2022 from https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2433761/16128_FULLTEXT.pdf

[37] Dimitris Panaretos, Efi Koloverou, Alexandros C Dimopoulos, Georgia-Maria Kouli, Malvina Vamvakari, George Tzavelas, Christos Pitsavos, and Demosthenes B Panagiotakos. 2018. A comparison of statistical and machine-learning techniques in evaluating the association between dietary patterns and 10-year cardiometabolic risk (2002–2012): the ATTICA study. *British Journal of Nutrition* 120, 3 (2018), 326–334. https://doi.org/10.1017/S0007114518001150

[38] George Pounis. 2019. *Analysis in Nutrition Research*. Elsevier.

[39] Dafiny Rodrigues Silva Praxedes, Isabele Rejane Oliveira Maranhão Pureza, Laís Gomes Lessa Vasconcelos, André Eduardo da Silva Júnior, Mateus de Lima Macena, Telma Maria de Menezes Toledo Florêncio, Ingrid Sofia Vieira de Melo, and Nassib Bezerra Bueno. 2021. Association between energy intake under-reporting and previous professional nutritional counselling in low-income women with obesity: A cross-sectional study. *Nutrition Bulletin* 46, 3 (2021), 310–320. https://doi.org/10.1111/nbu.12513

[40] Pouya Saeedi, Inga Petersohn, Paraskevi Salpea, Belma Malanda, Suvi Karuranga, Nigel Unwin, Stephen Colagiuri, Leonor Guariguata, Ayesha A Motala, Katherine Ogurtsova, et al. 2019. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. *Diabetes research and clinical practice* 157 (2019), 107843. https://doi.org/10.1016/j.diabres.2019.107843

[41] Rani Sarmugam and Anthony Worsley. 2014. Current levels of salt knowledge: a review of the literature. *Nutrients* 6, 12 (2014), 5534–5559. https://doi.org/10.3390/nu6125534

[42] Heinz Schneider and Werner Venetz. 2014. *Cost of Obesity in Switzerland in 2012*. Technical Report.

[43] Arielle S Selya and Drake Anshutz. 2018. Machine Learning for the Classification of Obesity from Dietary and Physical Activity Patterns. In *Advanced Data Analytics in Health*. Vol. 93. Springer, Cham, 77–97. https://doi.org/10.1007/978-3-319-77911-9_5

[44] Sara Shariaty, Amirhossein Bahonaran, Parvin Ayremlou, Lida Hoseinalizdeh, and Mohammad Alizadeh. 2019. Comparison of dietary patterns, food groups, nutrients intake, cardio-metabolic biomarkers, and liver enzymes in successful and unsuccessful weight loss maintainers. *Obesity Medicine* 16 (2019), 100147. https://doi.org/10.1016/j.obmed.2019.100147

[45] Inge Spronk, Charina Kullen, Catriona Burdon, and Helen O'Connor. 2014. Relationship between Nutrition Knowledge and Dietary Intake. *British journal of nutrition* 111, 10 (2014), 1713–1726. https://doi.org/10.1017/S0007114514000087

[46] Statista. 2022. Market share of the leading companies in retail in Switzerland in 2020. Retrieved July 1, 2022 from https://www.statista.com/statistics/787298/switzerland-market-share-of-food-retailers/

[47] Amy F Subar, Laurence S Freedman, Janet A Tooze, Sharon I Kirkpatrick, Carol Boushey, Marian L Neuhouser, Frances E Thompson, Nancy Potischman, Patricia M Guenther, Valerie Tarasuk, et al. 2015. Addressing current criticism regarding the value of self-report dietary data. *The Journal of nutrition* 145, 12 (2015), 2639–2645. https://doi.org/10.3945/jn.115.219634

[48] Swiss Society for Nutrition (SSN). 2011. Swiss Food Pyramid. Retrieved July 1, 2022 from https://www.sge-ssn.ch/media/sge_pyramid_E_basic_20161.pdf

[49] Matthew K Taylor, Debra K Sullivan, Edward F Ellerbeck, Byron J Gajewski, and Heather D Gibbs. 2019. Nutrition literacy predicts adherence to healthy/unhealthy diet patterns in adults with a nutrition-related chronic condition. *Public health nutrition* 22, 12 (2019), 2157–2169. https://doi.org/10.1017/S1368980019001289

[50] The European Parliament and the Council of the European Union. 2011. Regulation (EU) No 1169/2011 of the European Parliament and of the Council. Retrieved July 1, 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32011R1169

[51] The European Parliament and the Council of the European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council. Retrieved July 1, 2022 from https://eur-lex.europa.eu/eli/reg/2016/679/oj

[52] Shahadat Uddin, Arif Khan, Md Ekramul Hossain, and Mohammad Ali Moni. 2019. Comparing different supervised machine learning algorithms for disease prediction. *BMC medical informatics and decision making* 19, 1 (2019), 1–16. https://doi.org/10.1186/s12911-019-1004-8

[53] Virginia Vettori, Chiara Lorini, Heather D Gibbs, Francesco Sofi, Vieri Lastrucci, Gino Sartor, Ilaria Fulvi, Duccio Giorgetti, Giuseppe Cavallo, and Guglielmo Bonaccorsi. 2021. The nutrition literacy assessment instrument for italian subjects, NLit-IT: Exploring validity and reliability. *International Journal of Environmental Research and Public Health* 18, 7 (2021), 3562. https://doi.org/10.3390/ijerph18073562

[54] Jane Wardle, Kathryn Parmenter, and Jo Waller. 2000. Nutrition knowledge and food intake. *Appetite* 34, 3 (2000), 269–275. https://doi.org/10.1006/appe.1999.0311

[55] Lauren Crawford Watson, Junehee Kwon, David Nichols, and Martha Rew. 2009. Evaluation of the nutrition knowledge, attitudes, and food consumption behaviors of high school students before and after completion of a nutrition course. *Family and Consumer Sciences Research Journal* 37, 4 (2009), 523–534. https://doi.org/10.1177/1077727X08329002

[56] World Health Organization. 2020. Healthy Diet. Retrieved July 1, 2022 from https://www.who.int/news-room/fact-sheets/detail/healthy-diet

[57] World Health Organization. 2022. *WHO European Regional Obesity Report 2022*. Technical Report. World Health Organization, Copenhagen, Denmark. https://www.who.int/europe/publications/i/item/9789289057738

[58] Jing Wu, Klaus Fuchs, Jie Lian, Mirella Lindsay Haldimann, Tanja Schneider, Simon Mayer, Jaewook Byun, Roland Gassmann, Christine Brombach, and Elgar Fleisch. 2021. Estimating Dietary Intake from Grocery Shopping Data—A Comparative Validation of Relevant Indicators in Switzerland. *nutrients* 14, 1 (2021), 159. https://doi.org/10.3390/nu14010159