Journal of Positive School Psychology
2022, Vol. 6, No. 4, 11691 – 11697
http://journalppw.com

A Data-Driven Approach to Understanding the Impact of Covid-19 On Dietary Habits Amongst Bangladeshi Students

Jaynab Sultana, Sheikh Elhum Uddin Quadery, Fahad Rahman Amik, Tushar Basak, Sifat Momen

Department of Electrical and Computer Engineering, North South University
Dhaka 1229, Bangladesh;
jaynab.sultana@northsouth.edu, sheikh.quadery@northsouth.edu,
fahad.rahman1@northsouth.edu, tushar.basak@northsouth.edu, sifat.momen@northsouth.edu
Correspondence: fahad.rahman1@northsouth.edu

ABSTRACT

The advent of COVID-19 has brought upon behavioral changes in food habits and overall nutritional lifestyles. It is necessary for nutritionists to analyze how the crisis affected human behavior, particularly general well-being. The purpose of this research is to see how quarantine affected food habits amongst Bangladeshi youngsters. A cross-sectional survey consisting of a three-part questionnaire (personal details, pre-COVID habits, and during COVID habits) was done on 230 students. Exploratory data analysis was done to summarize the resulting dataset. Afterward, K-means clustering was used to find the natural groupings that exist in this dataset. Later on, elbow and silhouette methods were applied to determine the number of optimal clusters so that machine learning classifiers could be run on them. Finally, SHAP was implemented to identify the features' importance. After clustering the data, the silhouette method identified 5 as the optimal cluster number to try the models out on. Out of all the models utilized, Extra Tree Classifier worked the best since it scored the highest in all the evaluation criteria. SHAP revealed that the habit of not eating carrots had the most significant impact in predicting the membership of a cluster class. These findings demonstrated that although there was indeed a negative edge to the crisis, there were slits of positive outcomes among nutritional behaviors which could be leveraged for a healthier societal change. Some of the values are within the consensus (reducing intake of fruit juice) while others are novel (smoking patterns being unchanged). These ideas will help in creating a broader idea as to how nutrition can take a nosedive and can be managed using data.

Keywords: COVID-19, eating habits; physical activity, quarantine, machine learning.

1. Introduction

The massive outburst of the COVID-19 pandemic has brought about a massive number of infections and deaths. As a consequence, governments all around the world have called for drastic measures including travel bans and mass lockdowns. Complete lockdowns have halted population mobility and resulted in socio-economic downturns, causing disruptions in people's daily lives [1] [2]. Lockdown has had a variety of effects on the lives of young people, some of which are favorable and others that are detrimental [3]. Balanced meals are beneficial to the well-being of the mind and body, prevent illnesses and boost immunity [4]. COVID-19's spread has triggered a slew of negative feelings and harmed sleep quality, resulting in an increase in obesity [6]. There is also evidence of the risks of inappropriate eating such as overeating which are brought forth as a form of mental coping mechanism to deal with the restrictions of COVID-19 [7].

Research indicates that COVID-19 impacted deleteriously both on physical and mental health. According to a poll of 8444 German citizens, melancholy and anxiety have increased as a result

of the quarantine's isolation [8] while, according to a poll of 672 Bangladeshis, emotions of loneliness, melancholy, anxiety, and sleep disturbance are on the rise in the general population [9]. It is clear that quarantine has resulted in an array of negative emotions which consequently affect the eating and sleeping habits of youngsters but what is less explored is understanding the mechanism by which this happens. While this is a long-term goal, this article provides a first step toward understanding the influence of quarantine on Bangladeshi youngsters' dietary habits. We investigate this using a machine learning approach combined with explainable AI.

The rest of the article is organized as follows: Related works are presented in section 2 followed by the methodology we adopted in section 3. Section 4 presents the results and the article concludes in section 5.

2. Literature review

Following the lockdown, a variety of studies were conducted to see how COVID-19 affected people's physical activity, diets, mental health, and overall well-being. Corby et al. [12] conducted a study based on a cross-sectional self-critical online survey conducted before and during COVID-19 induced closures. To investigate temporal differences, various statistical methods were performed. The study found that physical activity involvement had plummeted, sedentary behavior had increased, and food habits had changed in a negative way. According to studies conducted in a few European countries [13], COVID-19 quarantine played a role in lifestyle changes such as increased poor eating habits, a fall in athletic activity, and a spike in weight gain. Shaun and colleagues [14], in a separate study, discovered that students who gained weight, during the post-shutdown lockdown period, lacked physical activity and the intake of homemade food. Bertrand et al. [15] focused their research on 125 university students' nutritional consumption, physical activity, and sedentary behavior. The findings were discouraging: only one-tenth fulfilled the physical activity systems, while only one-third met the sedentary behavior standards. Mattioli et al. [16] explored how isolation affects one's lifestyle, specifically nutrition and physical activity. Quarantine, according to them, causes self-isolation, which might lead to a lack of fruit and vegetable consumption. Isolative habits also cause people to abstain from physical activity since this may require one to leave the house. Bağcı et al. [17] conducted an investigation, using an online survey, on 536 medical interns to see if there were any changes in their nutritional and physical activities during the lockdown period. They discovered that after the start of the lockdown, the average BMI has increased irrespective of their gender. Surprisingly, the number of students with good eating habits climbed up by 18.8%, while the number of students with bad eating habits declined by 3.2%. Unfortunately, two-thirds of the respondents reported that they were doing less physical exercise while another third said their sleep was of poor quality. Taeymans J et al. [18] evaluated lifestyle patterns such as physical activity, resting period, food habits, alcohol intake, and sleep behavior during a confinement period of two-months owing to the COVID-19 outbreak. They collected data through an anonymous online survey. Low sleeping quality was found in 44.7% of the sample, although the median healthy sleep duration for adults is considered to be 8 hours. The data were analyzed using frequency analyses and nonparametric statistical approaches.

3. Methodology

This section provides details on the data acquisition process, data pre-processing, and the application of machine learning algorithms to understand the physical and dietary habits of Bangladeshi university students.

3.1. Data Collection

This study's dataset is focused on Bangladeshi students and was gathered via an online poll. The study included 230 students (57.83% male and 42.17% female) from various universities, colleges, and schools in Dhaka, Bangladesh. The data was collected between July 25th and August 30th, 2021.

The majority of responders (75.22%) are between the ages of 21 and 25, while 17.39% of students are between the ages of 15 and 20. There are also very few responses aged 26 and up, as well as those aged 15 and under (a total of 6.52%). The number of unmarried respondents was high, at roughly 94.78%, because the majority of the respondents were university students. Prior to COVID-19, the bulk of the student's household income was less than 40,000 BDT, accounting for 40.43% of the total. Furthermore, 61.74% of respondents stated that their family's income had been reduced as a result of COVID-19. Despite the fact that 80% of respondents were aware of the benefits of a well-balanced diet, only 57% of them adhered to it. Women were found to be more concerned about their health than men, resulting in a twofold rise in the number of women consuming a balanced diet. The majority of the respondents got a restful night's sleep, with 59.6% sleeping for 6-8 hours per day.

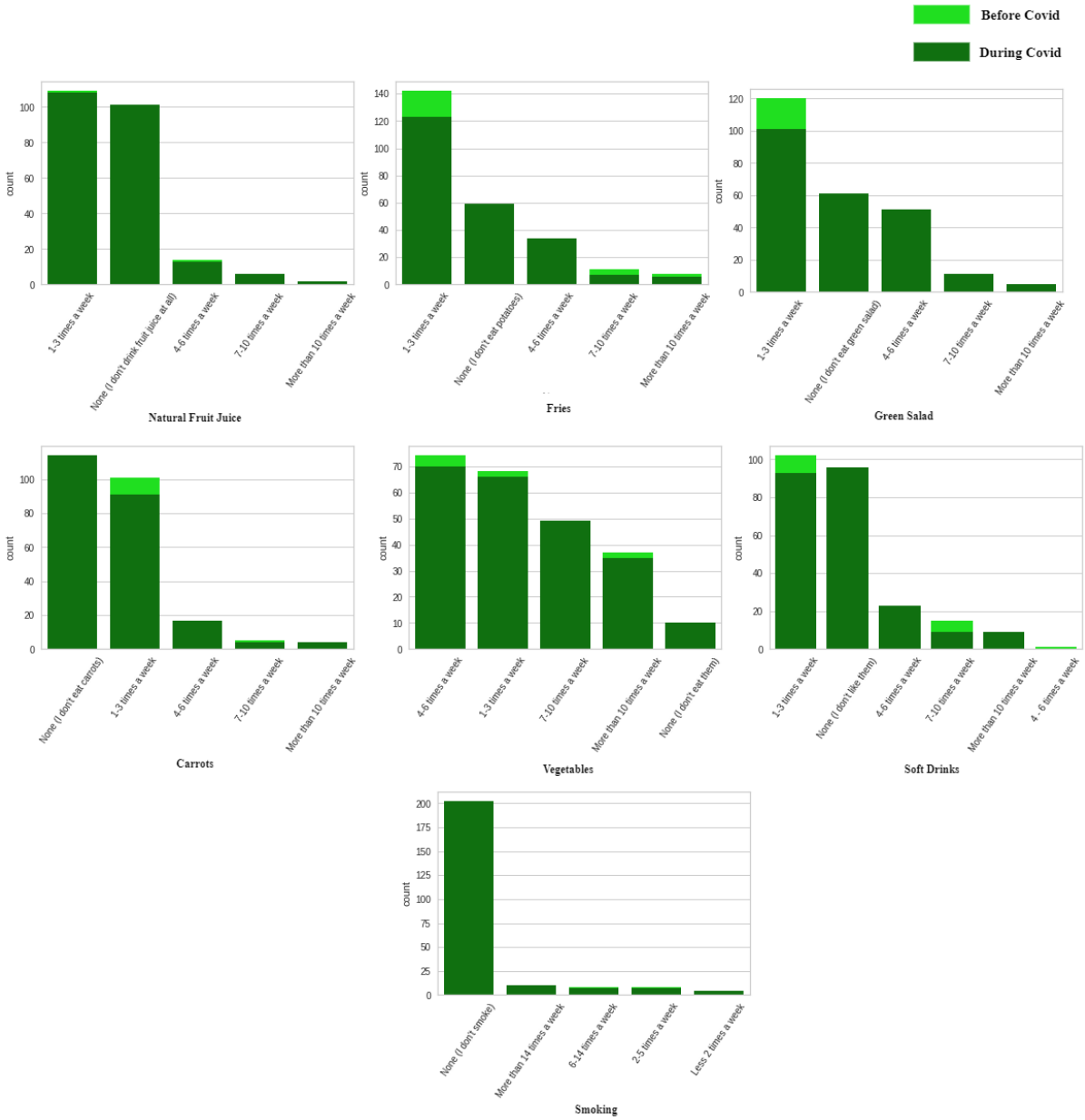

Figure 1. Comparison of eating patterns before and during COVID.

3 .2. Data preprocessing and exploratory data analysis

We developed a total of 30 questions for the questionnaire, which were separated into three categories: personal details, eating habits (before Covid-19 time), and eating habits (during Covid-19 period). The data type was checked at the start of data preparation. After that, null or missing data were reviewed for deletion. Duplicate entries were sought and, if found, were deleted. The number of unique values in each column was tallied. After counting the values in the gender field, just one record was found with binary gender, which was later eliminated since it is an underrepresented data and no statistics can be derived from it.

A number of observations can be made from Figure 1. Firstly, we see that people are less likely to drink natural fruit juice during COVID than they were before COVID. Prior to the pandemic, 142 responders ingested fries, but this number had reduced to 123 by the time the pandemic began. Surprisingly, just 38 people indicated they did not eat fries before the pandemic, but that number jumped to 59 just after the declaration of lockdown. Green salad consumption, on the other hand, has not reduced as much as fries consumption. Vegetable consumption has increased for responders who generally consume high levels of vegetables (7-10 times a week). Those who used to take soft drinks 1-3 times each week (before the lockdown) reduced their consumption to zero while amongst those who, in general, do not consume soft drinks have increased their soft-drink consumption. The smoking pattern has remained unchanged.

3.3. Clustering the data using K-means analysis

As it is an unlabeled data, we have planned to find the natural grouping in the dataset. The K-means clustering technique was used to accomplish this. K-means clustering algorithm starts with K clusters, with the centroid of each cluster set at a random location in the space [20]. The Euclidean distance metric was used to determine the membership of each data point. This re-evaluates the centroid for each cluster and the process is repeated until the membership of the data points remains unchanged. Determining the optimal number of clusters (i.e. the value of k) is an NP-hard problem. One way of estimating an appropriate value of K is using the Elbow method.

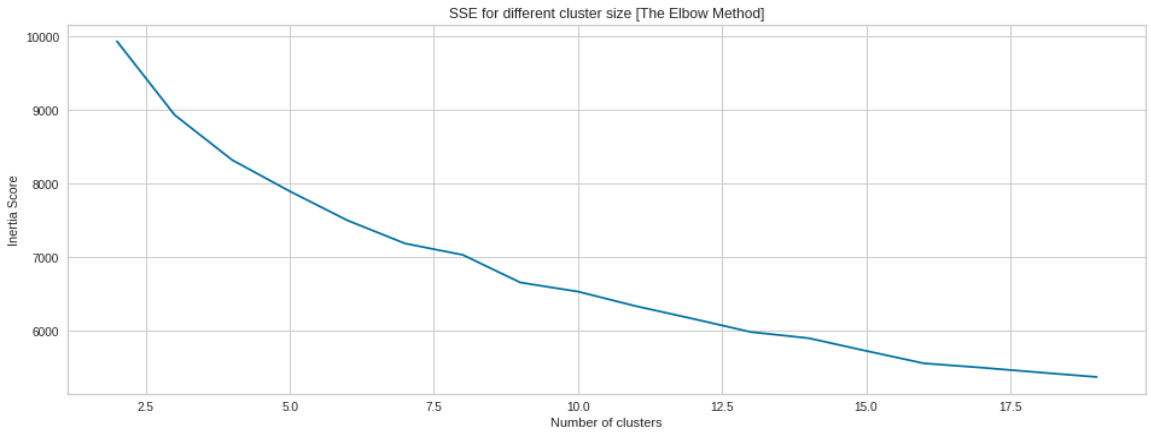

Figure 2. Determination of the ideal cluster number using the Elbow method.

A graph (as shown in Figure 2) of inertia score (on the y-axis) is plotted against candidate clusters (between 2 and 20). Since the plot did not reveal a sharp elbow, it was difficult to decide on the suitable number of clusters using this approach.

Silhouette methods were then applied to identify the required number of clusters for this dataset.

Table 1. Average silhouette scores.

Number of Clusters	2	3	4	5	6	7
Average Silhouette Score(%)	13.54	13.40	12.09	13.64	12.87	12.16

$$a(i) = \frac{1}{|C_I| - 1} \Sigma_{j \in C_I, i \neq j} d(i,j) \quad \text{(1)}$$

$$b(i) = min_{J \neq I} \frac{1}{|C_J|} \Sigma_{j \in C_J} d(i,j) \quad \text{(2)}$$

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}, if\ |C_I| > 1 \quad \text{(3)}$$

Silhouette score for each data point is calculated using equations (1) - (3). Here, *a(i)* gives the average distance between a given data point *i* and all other points in the same cluster (except itself) and *b(i)* is the smallest average distance between the datapoint *i* and the datapoint in the nearest cluster. Equation 3 provides the silhouette score for the datapoint *i*. If clustering for the data point *i* is done correctly, we should expect intra-class similarity to be higher than inter-class similarity. This will be true if *b(i) > a(i)*. As a result, a higher positive value for the silhouette score indicates better clustering goodness. This is then used to calculate the average silhouette score. Table 1 shows that the highest average silhouette score was obtained for k = 5; thus, this suggests that a total of 5 clusters is the best cluster size for this dataset.

4. Evaluation

Once the natural grouping is derived (5 clusters), the dataset is labeled for each data point. The dataset is split into train and test in the ratio of 70:30 respectively. A total of 5 machine learning classifiers were applied and their corresponding hyperparameters were tuned. Of the five classifiers, Extra Trees Classifier had the best performance (shown in Table 2).

Table 2. Performance of machine learning classifiers.

Model	**Accuracy**	**AUC**	**Recall**	**Prec.**	**F1**
Extra Trees Classifier	**0.8812**	**0.9869**	**0.8753**	**0.9027**	**0.8788**
Logistic Regression	0.8688	0.9830	0.8603	0.8943	0.8652
Random Forest Classifier	0.8625	0.9854	0.8443	0.8767	0.8569
Light Gradient Boosting Machine	0.8438	0.9742	0.8180	0.8692	0.8352
Gradient Boosting Classifier	0.8062	0.9612	0.7990	0.8413	0.8030
Decision Tree Classifier	0.7500	0.8415	0.7133	0.7777	0.7448

By applying the best model (Extra Trees classifier) to SHapley Additive exPlanations(SHAP) [19], we captured features' importance. From Figure 3, we see that the feature 'ccarrots_none' has the most significant impact in predicting the class a record belongs to.

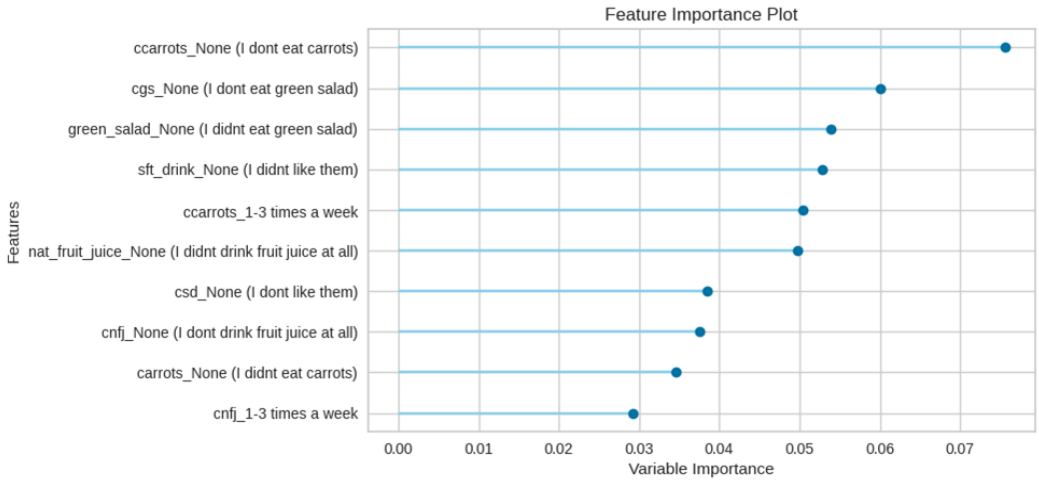

Figure 3. Feature importance graph.

5. Conclusion

The essential aim of this study is to quantitatively analyze how quarantine impacts dietary habits amongst Bangladeshi university students. A dataset of a total of 230 respondents was collected. After relevant data pre-processing and exploratory data analysis, the K-means clustering approach was used to determine the natural groupings of the dataset. Silhouette technique reveals a cluster of 5 to be the best cluster size. SHAP was used to determine the importance and dominance of features in each cluster. Unsupervised learning is followed by classification to predict the dietary habit to one of the classes (out of five). Extra Trees Classifier provides the best performance in terms of accuracy, precision, recall, f1-score, and the area under the ROC curve.

References

6.1. Journal Article

[1] De Palma, A., Vosough, S., & Liao, F. (2022). An overview of effects of COVID-19 on mobility and lifestyle: 18 months since the outbreak. *Transportation Research Part A: Policy and Practice*,V159, 2022

[2] M. Pellegrini, V. Ponzo, R. Rosato,E. Scumaci,I. Goitre,A. Benso,S. Belcastro, C. Crespi,F. D. Michieli, E. Ghigo,F. Broglio,and S. Bo1,7,12(2020).COVID-19 phobia, mindful eating, eating habits and body weight change among university students during pandemic: A pilot study.*Food and Health*, 8(2).

[3] M. Bodrud-Doza, M. Shammi, A. Islam, and M. Rahman(2020).Exploring COVID-19 stress and its factors in Bangladesh: A perception-based study.*Heliyon*,6(7).

[4] D. Reyes-Olavarría, P. Á. Latorre-Román, I. P. Guzmán-Guzmán, D. Jerez-Mayorga,F. Caamaño-Navarrete, and P. Delgado-Floody (2020).*International Journal of Environmental Research and Public Health*,17(15).

[5]Q. Feng, Q.-le Zhang, Y. Du, Y.-ling Ye, and Q.-qiang He, (2014).Associations of physical activity, screen time with depression, anxiety and sleep quality among Chinese college freshmen.*PLOS ONE*, 9(6).

[6]J.-P. Chaput and C. Dutil, (2016), Lack of sleep as a contributor to obesity in adolescents: Impacts on eating and activity behaviors,*International Journal of Behavioral Nutrition and Physical Activity*, 1 (13)

[7]C. Evers, A. Dingemans, A. F. Junghans, and A. Boevé, (2018), Feeling Bad Or Feeling Good, Does Emotion Affect Your Consumption Of Food? A Meta-Analysis Of The Experimental Evidence. Neuroscience & Biobehavioral Reviews,*Neuroscience and Biobehavioral*

Reviews, 92
[8]A. Fiske, A. Schneider, S. McLennan, S. Karapetyan, and A. Buyx, (2021), Impact of COVID-19 on patient health and self-care practices: a mixed-methods survey with German patients, *BMJ Open*, 9(11)
[9]R. Das, M. R. Hasan, S. Daria, and M. R. Islam, (2021), A nationwide dataset on the mental health of the Bangladeshi population due to the COVID-19 pandemic,*ELSEVIER*, 4(11)
[10]Pop, L-M, Iorga M, Muraru I-D, Petrariu F-D, (2021), Assessment of Dietary Habits, Physical Activity and Lifestyle in Medical University Students. *Sustainability*, 6(13).
[11]T. Deliens, B. Deforche, I. De Bourdeaudhuij & P. Clarys, (2014), Determinants of eating behavior in university students: a qualitative study using focus group discussions, *BMC Public Health*, 1(14).
[12]C. Sidebottom, S. Ullevig, K. Cheever, T. Zhang, 2021, Effects of COVID-19 pandemic and quarantine period on physical activity and dietary habits of college-aged students. *Sports Medicine and Health Science*, 2(3).
[13]C. Arianna, S. Umberto, R. Laura, (2021), Lifestyle Changes Related to Eating Habits, Physical Activity, and Weight Status During COVID-19 Quarantine in Italy and Some European Countries, *Frontiers in Nutrition* 8.
[14]Md M. A. Shauna,Md W. R. Nizum, S. Munny,F. Fayeza,S. K. Mali,M. T. Abid, A.- R. Hasand, (2021) Eating habits and lifestyle changes among higher studies students post-lockdown in Bangladesh: A web-based cross-sectional study. *Heliyon*, 8(7).
[15]L. Bertrand, K. A. Shaw, J. Ko, D. Deprez, P. D. Chilibeck, and G.A. Zello gordon, (2021),The impact of the coronavirus disease 2019 (COVID-19) pandemic on university students' dietary intake, physical activity, and sedentary behavior. Applied Physiology, *Nutrition, and Metabolism*, 3(46).
[16]A. V. Mattioli, S. Sciomer, C. Cocchi, S. Maffei, and S. Gallinae, (2020), Quarantine during COVID-19 outbreak: Changes in diet and physical activity increase the risk of cardiovascular disease. *Nutrition ,metabolism, and cardiovascular disease:NMCD*, 9(30).
[17]T. A. B. Bağcı , A. Kanadıkırık, E. Somyürek, G. Gerçek, H. B. Tanrıkulu, E. Öntaş, S. Uzun, (2021), Impact of COVID-19 on eating habits, sleeping behavior and physical activity status of final-year medical students in Ankara, Turkey. *Public Health Nutrition*, 18(24).
[18]J. Taeymans, E. Luijckx,S. Rogan, (2021), Physical activity, nutritional habits and sleeping behavior in health professions students and employees of a Swiss University of Applied Sciences during and after the COVID-19 confinement period: A protocol for a questionnaire survey study. *JMIR Research Protocols*, 4(7).
[19]Lundberg, Scott M., and Su-In Lee. (2017), A unified approach to interpreting model predictions. *Advances in neural information processing systems* (30).
[20]Likas A, Vlassis N, Verbeek JJ. (2003), The global k-means clustering algorithm. *Pattern recognition*, 36(2).