

## Segmenting the eating behaviour of university students using the K-means algorithm

Lenin Quiñones Huatangari<sup>1</sup>, María Alina Cueva Ríos<sup>1,2</sup>, Rafaela Teodosia Huerta Camones<sup>3</sup>

<sup>1</sup>Instituto de Investigación de Ciencia de Datos, Universidad Nacional de Jaén, Cajamarca, Perú

<sup>2</sup>Departamento de Ingeniería de la Industria Alimentaria, Universidad Nacional de Jaén, Cajamarca, Perú

<sup>3</sup>Departamento de Psicología y Filosofía, Facultad de Ciencias Sociales y Humanidades, Universidad Nacional Enrique Guzmán y Valle, Lima, Perú

### Article Info

#### Article history:

Received Aug 9, 2022

Revised Nov 2, 2022

Accepted Jan 4, 2023

#### Keywords:

Computer applications

Data mining

Development models

Dietetics

Nutrition

### ABSTRACT

Universities that teach arts education do not only teach how to play an instrument or conduct musical ensembles; they are agents of change in eating behaviour for the praxis of teaching and dissemination of healthy education. The objective of the research was to segment the eating behaviour of students of the artistic education-music speciality of the National University of Education "Enrique Guzmán y Valle" by applying the K-means algorithm. To do this, the methodology consisted of understanding the problem, understanding the data collected, preparing the data, modelling and evaluating the model. For modelling, the free software Weka was used through the K-means clustering technique on a data matrix of 148 instances with forty-three nominal variables collected online based on an instrument designed and validated to assess eating behaviour in university students. Two was determined to be the optimal clustering for eating behaviour in university students, using the elbow method, with a distribution of 49% for the first cluster and 51% for the second cluster. The results of the study population showed that the eating behaviour of university students is adequate.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



### Corresponding Author:

Lenin Quiñones Huatangari

Instituto de Investigación de Ciencia de Datos, Universidad Nacional de Jaén

Carretera Jaén, San Ignacio KM 24, Sector Yanuyacu, Jaén, Cajamarca, Perú

Email: lenin.quinones@unj.edu.pe

## 1. INTRODUCTION

Feeding is a universal act, it is one of the pillars that support and at the same time determine the quality of life, it is the basic element for survival, it is elementary and it sustains the rest of the needs [1]. People opt for unhealthy eating behaviours, promoted primarily to satisfy hunger, regardless of the nutritional value of the food [2]. World Health Organisation warns of diseases linked to poor diet [3]–[6]: type II diabetes, cancer, osteoporosis, anaemia, and cardiovascular problems. Therefore, it is important to pay attention to food and its nutritional content, for example, the consumption of fruit and vegetables reduces the risk of developing diseases [7], [8]. Therefore, healthy eating is about eating what is necessary for good development, without deprivation or limitations, it is about having a stable state of mind, having energy, and feeling good [9].

Connecting the analysis of socialisation processes with food culture allows us to understand the inseparable nature of the nutritional and socio-cultural dimensions of food [10]. Therefore, the study of music develops emotional intelligence, discipline, sensitivity, creativity and cultural self-expression. Students of art education at the National University of Education "Enrique Guzmán y Valle", specialising in music, include

the subject of food education in their curriculum. In this sense, the university not only trains professionals with artistic skills, such as playing an instrument or directing musical ensembles; they are agents of change in nutritional behaviour for the practice of teaching and dissemination of healthy education.

The unhealthy characteristics of university students' eating behaviours are conditioned by underlying beliefs [11]. In addition, the factors that are perceived as influencing their diet are the limited time available during university life; the family group; friends; the high cost and scarce supply of healthy food within the educational institution; situations of academic stress and the place of residence [12]. As a result, they have observed inadequate eating habits [13], [14], such as a low consumption of fruit and vegetables and a high consumption of sweets [15]. It reflects the need to implement effective strategies that contribute to rooting healthy eating and nutrition habits in young people, in accordance with the postulates of the nutritional sciences [16], [17].

The computational power of computer hardware, coupled with the existence of information in social domains, including education, requires the development and application of algorithms, and software that facilitate obtaining meaningful information from this data and transforming it into knowledge [18]. The art of learning science from data is cultivated through machine learning (ML), which uses statistical and computer science knowledge for pattern recognition and data mining applications [19]. The potential of ML to understand multivariate datasets involving different observation platforms, resolutions, scales, and interconnectedness of physical laws can play a key role in unravelling hidden patterns in data and developing predictive models [20]. Supervised ML uses labelled data, which means that some input data is already labelled with the correct output, and unsupervised ML works with unlabelled data sets to analyse and group them together [21]. Artificial neural networks (ANNs) and decision trees were successfully used to predict one aspect of diet quality. However, further exploration of the use of ANNs and decision trees in dietary pattern analysis is warranted [22]. Also, using logistic regression analysis, they showed that yoghurt consumption was the only independent risk factor for road traffic accidents and, using the decision tree, that yoghurt consumption was the initial divergence variable, i.e. patients who consumed yoghurt regularly, the incidence of road traffic accidents was 6.60%, while that of patients who did not consume yoghurt was 16% [23]. Clusters of university students who have similar dietary habits when using the school cafeteria were detected using clustering techniques. Data were collected from the prepaid meal card that records the history of food purchased [24]. Finally, they have structured recommendations for the development of a food policy based on labour productivity and the economic cost on workers' income [25]. Clustering analysis groups objects (observations and events) based on information found in the data describing the objects or their relationships. The goal is that objects in a group are similar (or related) to each other and different (or unrelated) from objects in other groups. The greater the similarity within a group and the greater the disparity between groups, the better the clustering [26]. However, K-means is the essential and fundamental technique through which data centres are analysed. It is the algorithm used for clustering with known sets of median points; it is also called nearest neighbour clustering [27]. In this context, the research evaluates the eating behaviours of university students through an unsupervised data mining technique, which will allow decisions to be made by the relevant areas such as university welfare and tutoring. Therefore, the objective was to segment the eating behaviour of students at the National University of Education "Enrique Guzmán y Valle" by applying the K-means algorithm.

## **2. METHOD**

### **2.1. Data matrix**

Data were collected in May 2021. From one hundred and forty-eight (148) students of the National University of Education "Enrique Guzmán y Valle" of Peru, using an online questionnaire based on the instrument designed and validated to assess eating behaviour in university students [28]. Thus, forty-three nominal variables have been obtained as shown in Table 1 (see in Appendix).

### **2.2. Method for using K-means**

It was the one proposed by the cross industry standard process for data mining (CRISP-DM) and consisted of five phases [29]. Understanding the problem: a literature review was conducted on the eating behaviours of university students; no studies were found that employed clustering techniques to determine eating behaviours in Peru. The information collected was from students of the professional career of art education through a Google Form. Understanding the data: data from forty-three variables collected were standardised and understood. They were then entered into excel for preparation. Data preparation: data were obtained without null or outliers, which allowed reliable results to be obtained, using mode. The data source was transformed in excel to the attribute relation file format (ARFF) format required by the modelling software. Modelling: Weka is a data mining tool consisting of a set of ML algorithms [30]. The K-means

clustering algorithm has been employed, using the Weka tool. To determine the optimal value of clusters, the elbow method has been considered [31]. Model evaluation: the quality of the clusters obtained was evaluated on the basis of the number of clusters, cluster instances, sum of squared error (SSE) and construction time

$$SSE = \sum_{j=1}^K \sum_{i=1}^n \|x_i^{(j)} - \mu_j\|^2 \quad (1)$$

Where  $K$  is the number of clusters,  $n$  is the number of instances of the data matrix,  $x_i^{(j)}$  is the object in the cluster  $j$ ,  $\mu_j$  the cluster centroid and  $\|\cdot\|$  is the distance function.

### 3. RESULTS AND DISCUSSION

The preferences of university students on eating behaviour has generated a database with forty-three variables as shown in Table 1 (see in Appendix). It shows that more than half of university students (53.4%), considered that the most important factor in choosing a food is to know the nutritional content. In addition, they consume healthy food and in the company of their families. However, less than half consider that they read and understand food nutrition labels as shown in Table 1 (see in Appendix).

Eating habits, health-related behaviours and beliefs of students at a university in Greece have been assessed [32]. The transition to university is a critical developmental window during which eating behaviours are susceptible to dysregulation [33]. In addition, most university students are in the transition to adulthood and tend to adopt unhealthy eating habits characterised by low consumption of fruit and vegetables [34]. Also, only a minority of students showed positive health practices above recommended levels and that the level of excessive and problematic alcohol consumption was high [35]. The study differs from work carried out in three poor communes in Chile, specifically in the Metropolitan region of Santiago: Cerro Navia, Quinta Normal and La Granja, where they found that 45% of the people interviewed read the information on food labelling [36]. This corroborates that the study population consisting of university students is a determining factor in eating behaviour. However, the results ratify the study, where it is mentioned that the lack of knowledge about nutrition labelling does not contribute to the process of selection and purchase of food, other aspects of a personal and cultural nature determine the decision to buy food [37]. Therefore, young people's use of nutrition labelling on food does not contribute to their choice when purchasing food [38].

Students point to sensory and hygiene factors, cost of food, availability of food and time available for eating as the main factors influencing choice. The influence of friends and family, among others, was identified as influencing choices in this aspect [39]. University students who live alone prioritise their social status, education and career growth over their health. Compared to eating with family, eating with friends is seen as fun and less controlled as it represents independence and freedom to eat without restrictions [40]. However, because the questionnaire was carried out under quarantine due to a global health emergency within the framework of COVID-19, the pandemic has forced families into confinement [41]. It is therefore considered to be a cause for students to eat their meals in the company of their families, see Table 1 (see in Appendix). As for the acquisition of healthy dietary habits, it does not seem sufficient that the knowledge is correct, it is also necessary that the modification or abandonment of these unhealthy and erroneous eating habits takes place, in order to achieve a healthy and balanced diet [42]. Their good eating habits would allow an adequate level of adherence to the Mediterranean diet [43]; the increased scientific evidence of its benefits and effectiveness in terms of longevity, quality of life and disease prevention, and the recognition of UNESCO, have brought this dietary pattern to an unprecedented level [44].

With the increasing presence of large and complex data, the need arises to aggregate them. The need arises to group them. Since these are difficult or impossible to process with traditional methods. These instances present a number of objects of a multidimensional nature [45]. K-means is a method that attempts to find a specific number of groups, which are represented by their centroids, applicable to a group of objects; it is one of the oldest and most widely used clustering algorithms [46]. It has the disadvantage that it requires that the number of clusters to be obtained is specified a priori, therefore some domain knowledge is necessary and in many cases, the quality of the final partition depends on a good selection of the initial partition [47]. It was determined that two clusters were the optimal value of the preferences of university students on eating behaviour and their graphical representation obtained through the elbow method, Figure 1. The evaluation of the quality of the clusters shows that the iterations used were seven, the SSE was 2.521, the distribution of instances for each group was 72 and 76, and finally the time used by the software to obtain the results was 0.02 s.

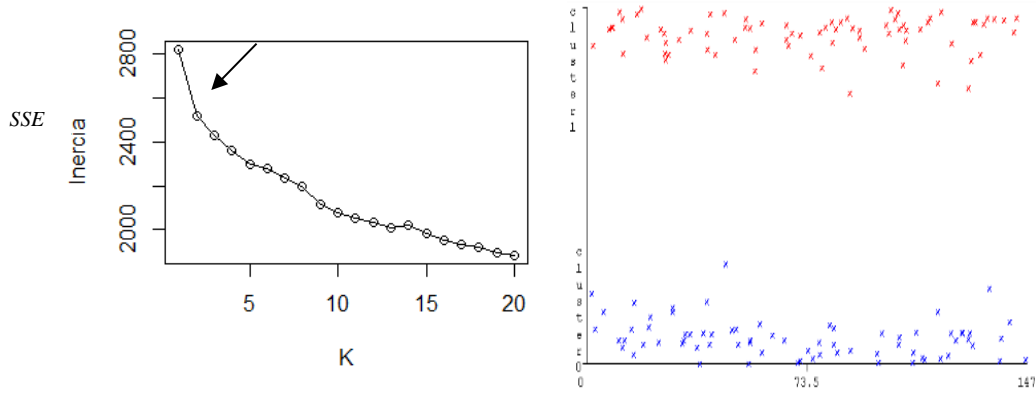


Figure 1. Optimal number of clusters and their graphical representation

Two groups of characteristics have been determined, which show the behaviour of eating habits in university students. However, these cannot be generalised simply because they vary, for example, by geographical area, season and cultural factors from one locality to another. University students have high rates of health risk behaviours, and these may be predictive of academic success [48].

**4. CONCLUSION**

Taking into account that the research was carried out in the context of the COVID-19 pandemic and that the population evaluated were university students. The results show that eating behaviour is adequate. The K-means clustering algorithm was used to segment the eating behaviour of 148 university students, whose data were collected on the basis of a validated instrument. It has been determined that two is the optimal number of clusters, with the SSE of 2.521. Furthermore, the segmentation of eating behaviour into two clusters is given for the type of population assessed, which would not be appropriate to generalise to the entire university student population. However, this type of study allows for the determination of patterns that facilitate decision-making by state or private institutions. It is recommended that multidisciplinary teams be formed to adequately evaluate the multiple factors of university students' eating behaviour.

**ACKNOWLEDGEMENTS**

For To the students of the National University of Education "Enrique Guzmán y Valle" in Peru, for their support in answering the online questionnaire on eating behaviour.

**APPENDIX**

Table 1. Name, description, values and number of variables collected

Variable	Description	Values	Quantity
Decision_Factor	Most important factor when choosing a food to eat	Its taste	30
		Its price	13
		Pleasing to the eye	4
		Expiry date	22
		Nutritional content	79
Tags	Difficulty in reading nutrition labels	Lack of time	41
		Not interested	5
		I don't understand them	27
		Laziness	12
Consumption_Restriction	Reason for food avoidance	I read and understand them	63
		I don't like it	38
		For taking care of me	67
		It's bad for me	26
Food_Preparation	Regular preparation of your food	I don't avoid any food	17
		Fried	18
		Steamed or boiled	33
		Roasted or grilled	20
		Baked	5
		Stewed or sautéed	72

Table 1. Name, description, values and number of variables collected (continue)

Variable	Description	Values	Quantity
Responsible_processing	Prepare your food more frequently during the week	My person	62
		My mother/relative	82
		Domestic worker	3
		I buy them ready-made	1
Satisfaction	What it does when you feel satisfied	I stop eating without a problem	112
		I stop eating but I find it difficult to do so	15
		I keep eating without problem	18
		I keep eating but I feel bad about it	3
Fat_action	Actions with visible meat fat	I remove all of it	59
		I remove most of it	41
		I remove a little	34
		I remove nothing	14
Chewing_food	Opinion on chewing each mouthful more than twenty-five times	I fully agree	47
		Agree	44
		Neither agree nor disagree	49
		Disagree	7
		Strongly Disagree	1
Schedule_Monday_Friday	Respect meal times from Monday to Friday	Yes	55
		Sometimes	82
		No	11
Monday_Friday_Monday_Friday	People with whom you usually eat during the day from Monday to Friday	Family	121
		Colleagues, partners, friends	3
		Only	24
Week_Friday_times	Respect meal times on Saturdays and Sundays	Yes	59
		Sometimes	67
		No	22
Snacks_week_Friday	People with whom you usually eat during the day on Saturday or Sunday	Family	129
		Colleagues, partners, friends	8
		Only	11
Agrada_fruits	Pleases to eat fruit	I like it very much	75
		I like it	59
		I neither like it nor dislike it	14
		I dislike it	0
		I dislike it very much	0
Agrada_vegetables	Delight in eating vegetables	I like it very much	53
		I like it	76
		I neither like it nor dislike it	16
		I dislike it	3
		I dislike it very much	0
Agrada_meat	Delights in eating meats (chicken, beef, and pork)	I like it very much	50
		I like it	80
		I neither like it nor dislike it	14
		I dislike it	3
		I dislike it very much	1
Agrada_fish	Delights in eating fish and seafood	I like it very much	61
		I like it	69
		I neither like it nor dislike it	11
		I dislike it	5
		I dislike it very much	2
Agrada_dairy_products	Delights in consuming milk or dairy products	I like it very much	34
		I like it	81
		I neither like it nor dislike it	25
		I dislike it	5
		I dislike it very much	3
Agrada_flour	Delight in eating bread, tortillas, potatoes, pasta and cereals	I like it very much	35
		I like it	83
		I neither like it nor dislike it	21
		I dislike it	8
		I dislike it very much	1
Agrada_legumes	Delights in eating dried legumes (beans, chickpeas and lentils)	I like it very much	49
		I like it	77
		I neither like it nor dislike it	17
		I dislike it	5
		I dislike it very much	0
Agrada_sweetfoods	Enjoys eating sweet foods (rice pudding, and mazamorra)	I like it very much	36
		I like it	80
		I neither like it nor dislike it	28
		I dislike it	3
		I dislike it very much	1

Table 1. Name, description, values and number of variables collected (continue)

Variable	Description	Values	Quantity
Agrad_eggs	Delights in eating eggs	I like it very much	51
		I like it	76
		I neither like it nor dislike it	15
		I dislike it	5
		I dislike it very much	1
Agrada_dryfruit	Delights in eating nuts (almonds, walnuts, sultanas, and seeds)	I like it very much	53
		I like it	69
		I neither like it nor dislike it	21
		I dislike it	3
		I dislike it very much	2
Agrada_alcoholic_beverages	Delights in drinking alcoholic beverages	I like it very much	7
		I like it	27
		I neither like it nor dislike it	62
		I dislike it	25
		I dislike it very much	27
Agree_food_packaging	Delights in eating packaged foods	I like it very much	9
		I like it	49
		I neither like it nor dislike it	69
		I dislike it	11
		I dislike it very much	10
Drink_preference	Drinks in larger quantities during the day	Water	98
		Soft drinks (chicha and passion fruit)	40
		Industrialised soft drinks (nectars and juices)	3
		Milk	4
		Carbonated and energy drinks (soft drinks and rehydrating drinks)	
Intake_between_meals	Habitually ingests between meals	Sweets	13
		Fruit or vegetable	101
		Biscuit	17
		Yoghurt	9
		Chips, chips, and fried foods	6
		Candied peanuts	2
Main_soup_main_meal	Usually include soup or broth in your main meal time during the day	Yes	91
		No	57
Second_main_meal	Usually include main course (main dish of a meal that is larger or more substantial) in your main meal time during the day	Yes	100
		No	48
Meat_main_meal	Regularly include meat, fish, poultry or seafood in your main meal time during the day	Yes	127
		No	21
Arrpasfrij_main_meal	Usually include rice, pasta or beans in your main meal time during the day	Yes	126
		No	22
Garnish_verdental	Regularly include vegetables or salads in your main meal time during the day	Yes	26
		No	122
Side_dish_flour	Regularly include tortillas, savoury bread, toast in your main meal time during the day	Yes	79
		No	69
Dessert_main_meal	Regularly include dessert in your main meal time during the day	Yes	110
		No	38
Fruit_main_food	Regularly include fruit in your main meal time during the day	Yes	119
		No	29
Appetizer_main_food	Regularly include cocktail or snack in your main meal time during the day	Yes	122
		No	26
Refreshment_main_food_main_drink	Habitually includes sweetened drink in your main meal time during the day	Yes	62
		No	86
Drinksinendul_main_food_main	Habitually includes unsweetened drink in your main meal time during the day	Yes	58
		No	90

Table 1. Name, description, values and number of variables collected (continue)

Variable	Description	Values	Quantity
Home_frequency	Frequency of eating food away from home	Every day	11
		5 to 6 times a week	3
		3 to 4 times a week	9
		1 to 2 times a week	36
		Once a fortnight	25
		Once a month	34
		Less than once a month	30
Excess_frequency	Frequency of overeating	Every day	10
		5 to 6 times a week	6
		3 to 4 times a week	16
		1 to 2 times a week	35
		Once a fortnight	30
		Once a month	34
		Less than once a month	17
Care_body_care	Willingness to take care of your body	Taking care of my diet	49
		Follow a temporary dietary regime	7
		Exercising	43
		Watching my diet and exercising	49
		Taking dietary or herbal supplements	0
Improve_food	Willingness to improve your diet	More information	33
		Social support	8
		Money	26
		Personal commitment or motivation	63
		Time	18
Diet	Consideration of diet	Different every day	60
		Different only a few times during the week	55
		Different only on weekends	9
		Monotonous	24
Nutrition_advice	Uses nutrition advice to improve health status	Strongly disagree	29
		Disagree	14
		Neither agree nor disagree	10
		Agree	52
		Strongly agree	43




## REFERENCES

- [1] S. S. Sergio *et al.*, "Aspectos psicológicos del hombre y su alimento: transitando de la naturaleza a la biotecnología en pro de la calidad de vida," *Revista Científica Electrónica de Psicología*, vol. 8, pp. 1–15, 2009.
- [2] S. S. Musa *et al.*, "The challenges of addressing the cholera outbreak in Cameroon," *Public Health in Practice*, vol. 4, pp. 1–3, 2022, doi: 10.1016/j.puhip.2022.100295.
- [3] S. Acosta, A. Johansson, and I. Drake, "Diet and lifestyle factors and risk of atherosclerotic cardiovascular disease—a prospective cohort study," *Nutrients*, vol. 13, no. 11, pp. 1–10, 2021, doi: 10.3390/nu13113822.
- [4] N. McKay, J. Przybysz, A. Cavanaugh, E. Horvatits, N. Giorgianni, and K. Czajka, "The effect of unhealthy food and liking on stress reactivity," *Physiology & Behavior*, vol. 229, pp. 1–41, 2021, doi: 10.1016/j.physbeh.2020.113216.
- [5] N. A. F. A. Bakar *et al.*, "Association between a dietary pattern high in saturated fatty acids, dietary energy density, and sodium with coronary heart disease," *Scientific Reports*, vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-17388-5.
- [6] Y. Wang *et al.*, "The role of diet and nutrition related indicators in biliary diseases: an umbrella review of systematic review and meta-analysis," *Nutrition & Metabolism*, vol. 19, no. 1, pp. 1–33, 2022, doi: 10.1186/s12986-022-00677-1.
- [7] D. F. Pais, A. C. Marques, and J. A. Fuinhas, "The cost of healthier and more sustainable food choices: do plant-based consumers spend more on food?," *Agricultural and Food Economics*, vol. 10, no. 1, pp. 1–21, 2022, doi: 10.1186/s40100-022-00224-9.
- [8] S. Gurusamy, C. S. Vidhya, B. Y. Khasherao, and A. Shanmugam, "Pulses for health and their varied ways of processing and consumption in India—a review," *Applied Food Research*, vol. 2, no. 2, pp. 1–9, 2022, doi: 10.1016/j.afres.2022.100171.
- [9] O. Arisukwu, D. Olaosebikan, A. J. Asaley, and F. Asamu, "Feeding habit and the health of undergraduate students: evidence from Nigeria," *The Journal of Social Sciences Research*, no. 52, pp. 498–506, 2019, doi: 10.32861/jssr.52.498.506.
- [10] E. Z. Aparici, "Aprender a comer: procesos de socialización y 'trastornos del comportamiento alimentario,'" Universitat de Barcelona, 2007.
- [11] J. Sánchez, A. Martínez, G. Nazar, C. Mosso, and L. D. -Muro, "Creencias alimentarias en estudiantes universitarios mexicanos: una aproximación cualitativa," *Revista chilena de nutrición*, vol. 46, no. 6, pp. 727–734, 2019, doi: 10.4067/S0717-75182019000600727.
- [12] C. T. Pantoja, "Percepción de condicionantes del comportamiento alimentario en estudiantes universitarios con malnutrición por exceso," *Perspectivas en Nutrición Humana*, vol. 16, no. 2, pp. 135–144, 2014, doi: 10.17533/udea.penh.v16n2a02.
- [13] A. V. -Martí, I. Elío, and S. S. -Cano, "Eating behavior during first-year college students, including eating disorders—RUVIC-RUNEAT-TCA project. Protocol of an observational multicentric study," *International Journal of Environmental Research and Public Health*, vol. 18, no. 18, pp. 1–10, 2021, doi: 10.3390/ijerph18189457.
- [14] E. F. Sprake *et al.*, "Dietary patterns of university students in the UK: a cross-sectional study," *Nutrition Journal*, vol. 17, no. 1, pp. 1–17, 2018, doi: 10.1186/s12937-018-0398-y.
- [15] F. B. Bulla, G. P. Villate, and M. V. -Zárate, "Prácticas alimentarias de un grupo de estudiantes universitarios y las dificultades percibidas para realizar una alimentación saludable," *Revista de la Facultad de Medicina*, vol. 63, no. 3, pp. 457–453, 2015, doi: 10.15446/revfacmed.v63n3.48516.
- [16] J. M. P. -Chaves and M. D. Á. S. -Meléndez, "Análisis de los hábitos alimenticios con estudiantes de décimo año de un Colegio Técnico en Pérez Zeledón basados en los temas transversales del programa de tercer ciclo de educación general básica de Costa




- Rica," *Revista Electronica Educare*, vol. 21, no. 3, pp. 1–23, 2017, doi: 10.15359/ree.21-3.12.
- [17] A. Sadiq, E. Strodl, N. G. Khawaja, R. Kausar, and M. J. Cooper, "Understanding eating and drinking behaviours in Pakistani university students: a conceptual model through qualitative enquiry," *Appetite*, vol. 161, pp. 1–10, 2021, doi: 10.1016/j.appet.2021.105133.
- [18] F. M. -Abad, H. -Ramos, and J. Pablo, "Técnicas de minería de datos con software libre para la detección de factores asociados al rendimiento," *Revista de Estudios y Experiencias en Educación*, vol. 2, no. 2, pp. 163–173, 2018, doi: 10.21703/rexe.Especial3201812514512.
- [19] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer, 2009, doi: 10.1007/978-0-387-84858-7.
- [20] K. J. Bergen, P. A. Johnson, M. V. d. Hoop, and G. C. Beroza, "Machine learning for data-driven discovery in solid Earth geoscience," *Science*, vol. 363, pp. 1–12, 2019, doi: 10.1126/science.aau0323.
- [21] V. M. Tiwari, "Need for machine learning applications in solid Earth geosciences in India," *Current Science*, vol. 122, no. 6, pp. 647–648, 2022, doi: 10.1111/ner.12135.
- [22] W. F. W. Yaacob, N. M. Sobri, S. A. M. Nasir, W. F. W. Yaacob, N. D. Norshahidi, and W. Z. W. Husin, "Predicting student drop-out in higher institution using data mining techniques," *Journal of Physics: Conference Series*, vol. 1496, pp. 1–13, 2020, doi: 10.1088/1742-6596/1496/1/012005.
- [23] T. Kawaguchi *et al.*, "An association between dietary habits and traffic accidents in patients with chronic liver disease: a data-mining analysis," *Biomedical Reports*, vol. 4, no. 5, pp. 615–622, 2016, doi: 10.3892/br.2016.640.
- [24] K. Nomura, H. Nambo, and H. Kimura, "Classification of school cafeteria users using the meal prepaid card data," *IEEE Transactions on Electronics, Information and Systems*, vol. 135, no. 10, pp. 1262–1269, 2015, doi: 10.1541/ieej.iss.135.1262.
- [25] D. A. P. Gavilanes, P. A. P. Silva, and B. C. Segovia, "Productividad laboral y su costo económico, relacionado con los hábitos alimenticios de los trabajadores de la Universidad Técnica Estatal de Quevedo," *Revista Científica de la Universidad de Cienfuegos*, vol. 11, no. 2, pp. 232–236, 2019.
- [26] N. Sharma, A. Bajpai, and R. Litoriya, "Comparison the various clustering algorithms of weka tools," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 5, pp. 73–80, 2012.
- [27] S. Dwivedi and L. K. P. Bhaiya, "A systematic review on K-means clustering techniques," *International Journal of Scientific Research Engineering Trends*, vol. 5, no. 3, pp. 750–752, 2019.
- [28] Y. F. M. -Sandoval *et al.*, "Diseño y validación de un cuestionario para evaluar el comportamiento alimentario en estudiantes mexicanos del área de la salud," *Nutricion Hospitalaria*, vol. 30, no. 1, pp. 153–164, 2014, doi: 10.3305/nh.2014.30.1.7451.
- [29] L. Q. Huatangari, D. M. Jara, N. Alvarado, M. E. Milla, and O. A. Gamarra, "Modelo para la estimación de la deserción estudiantil Awajún y Wampis empleando minería de datos," *Revista de Ciencia y Tecnología*, vol. 34, no. 1, pp. 45–50, 2020.
- [30] P. Singh and A. Surya, "Performance analysis of clustering algorithms in data mining in Weka," *International Journal of Advances in Engineering & Technology*, vol. 7, no. 6, pp. 1866–1873, 2015.
- [31] P. M. Hasugian, H. D. Hutahaean, B. Sinaga, Sriadhi, and S. Silaban, "Design of big data technology prototype for classification of village status based on village development index involves k-means algorithm to support village ministry Pdt work programs," *Journal of Physics: Conference Series*, vol. 1811, no. 1, pp. 1–12, 2021, doi: 10.1088/1742-6596/1811/1/012012.
- [32] I. Tiroidimos, I. Georgouvia, T. N. Savvala, E. Karanika, and D. Noukari, "Healthy lifestyle habits among Greek university students: differences by sex and faculty of study," *Eastern Mediterranean Health Journal*, vol. 15, no. 3, pp. 722–728, 2009, doi: 10.26719/2009.15.3.722.
- [33] S. Y. Lee, V. B. Agocha, P. R. Hernandez, C. L. Park, M. Williams, and L. M. Carney, "Coping styles moderate the relationship between perceived discrimination and eating behaviors during the transition to college," *Appetite*, vol. 168, pp. 1–39, 2022, doi: 10.1016/j.appet.2021.105699.
- [34] G. L. Bernardo *et al.*, "Association of personal characteristics and cooking skills with vegetable consumption frequency among university students," *Appetite*, vol. 166, pp. 1–33, 2021, doi: 10.1016/j.appet.2021.105432.
- [35] W. E. Ansari *et al.*, "Health promoting behaviours and lifestyle characteristics of students at seven universities in the UK," *Central European Journal of Public Health*, vol. 19, no. 4, pp. 197–204, 2011, doi: 10.21101/cejph.a3684.
- [36] M. A. B. and E. A. S., "Factores que determinan la selección de alimentos en familias de sectores populares," *Revista chilena de nutrición*, vol. 29, no. 3, pp. 308–315, 2002, doi: 10.4067/S0717-75182002000300006.
- [37] L. A. L. -Cano and S. L. R. -Mesa, "Etiquetado nutricional, una mirada desde los consumidores de alimentos," *Perspectivas en Nutrición Humana*, vol. 16, no. 2, pp. 145–158, 2014, doi: 10.17533/udea.penh.v16n2a03.
- [38] G. E. G. Morillo, J. B. R. Rodríguez, V. A. C. Vaca, K. A. P. Estévez, and J. C. Burbano, "Conocimientos, opiniones y uso del etiquetado nutricional de alimentos procesados en adolescentes ecuatorianos según tipo de colegio," *Perspectivas en Nutrición Humana*, vol. 21, no. 2, pp. 145–157, 2019, doi: 10.17533/udea.penh.v21n2a02.
- [39] C. Y. D. Cuervo, D. M. R. Caballero, A. C. L. Guapo, and P. N. G. Robayo, "Factores relacionados con las prácticas alimentarias de estudiantes de tres universidades de Bogotá," *Revista de Salud Pública*, vol. 17, no. 6, pp. 925–937, 2016, doi: 10.15446/rsap.v17n6.38368.
- [40] C. L. Marroquín, G. C. Díaz, C. A. Hurtarte, and R. Pardo, "La influencia de los grupos sociales en la alimentación de estudiantes universitarios mexicanos," *Psicumex*, vol. 11, pp. 1–21, 2021, doi: 10.36793/psicumex.v11i1.346.
- [41] L. D. Conejo, P. C. -Chaves, and S. L. -González, "Las familias y la pandemia del COVID-19," *Revista Electrónica Educare*, vol. 24, no. 1, pp. 1–4, 2020, doi: 10.15359/ree.24-S.10.
- [42] A. M. Bravo, N. Ú. Martín, and A. G. González, "Evaluación de los hábitos alimentarios de una población de estudiantes universitarios en relación con sus conocimientos nutricionales," *Nutricion Hospitalaria*, vol. 21, no. 4, pp. 466–473, 2006.
- [43] R. M. Lacoba, I. P. García, M. E. A. Saus, and F. E. Sotos, "Factores socioeconómicos asociados a los hábitos alimenticios de los jóvenes en la provincia de Albacete," *Al-Basit*, vol. 63, pp. 133–174, 2018.
- [44] L. S. -Majem and A. O. -Andrelluch, "La dieta mediterránea como ejemplo de una alimentación y nutrición sostenibles: enfoque multidisciplinar," *Nutricion Hospitalaria*, vol. 35, no. 4, pp. 96–101, 2018, doi: 10.20960/nh.2133.
- [45] J. P. -Ortega *et al.*, "Una heurística eficiente aplicada al algoritmo K-means para el agrupamiento de grandes instancias altamente agrupadas," *Computacion y Sistemas*, vol. 22, no. 2, pp. 607–619, 2018, doi: 10.13053/CyS-22-2-2548.
- [46] G. L. T., J. A. R., and O. P. B., "Segmentación de imágenes médicas digitales mediante técnicas de clustering," *APORTE SANTIAGUINO*, vol. 3, no. 1, pp. 108–116, 2010, doi: 10.32911/as.2010.v3.n1.428.
- [47] L. Arco, R. Bello, J. M. Mederos, and Y. Pérez, "Agrupamiento de documentos textuales mediante métodos concatenados," *Revista Iberoamericana de Inteligencia Artificial*, vol. 10, no. 30, pp. 43–53, 2006.
- [48] C. K. Y. Ong, M. J. Hutchesson, A. J. Patterson, and M. C. Whatnall, "Is there an association between health risk behaviours and academic achievement among university students?," *International Journal of Environmental Research and Public Health*, vol. 18, no. 16, pp. 1–14, 2021, doi: 10.3390/ijerph18168314.






**BIOGRAPHIES OF AUTHORS**

**Lenin Quiñones Huatangari**    is a Bachelor's degree in Mathematics, Master's degree in Computer Science and Ph.D in Industrial Engineering. Work and professional experience in undergraduate university teaching in Chile and Peru and postgraduate in the Master of Science with mention in Environmental Management of the Faculty of Ecology of the Universidad Nacional de San Martín. Currently teaching undergraduate and graduate courses at the National University of Jaén. The lines of research that his develop: construction and use of mathematical models for Engineering, application of artificial intelligence, computation as a tool for solving research problems and teaching of mathematics. Organizer and speaker at national and international conferences, thesis advisor, reviewer of scientific journals, author of scientific articles and books. He can be contacted at email: lenin.quinones@unj.edu.pe.



**María Alina Cueva Ríos**    is a Principal Professor at the Department of Food Industry Engineering, Science and Food Technology institute member, Data Science Institute member of National University of Jaen, Perú. Master of Science in Agricultural Production Systems, Doctor of Educational Sciences, doctoral student in food engineering; food safety specialist, trinorma assessor. With more than 30 years of experience in food industry, and university teaching. She has been developing work on public health policies with emphasis on healthy eating, tele-education, organization of canteens and food sales. Development of new food products and application of essential oils in food preservation. She can be contacted at email: maria\_cueva@unj.edu.pe.



**Rafaela Teodosia Huerta Camones**    current principal investigator in Education, appointed principal professor with experience in undergraduate and postgraduate university teaching, attached to the Department of Psychology and Philosophy of the Faculty of Sciences Social and Humanities of the National University of Education Enrique Guzmán y Valle de Lima, Peru, with more than 36 years of official services. Degree in Initial Education, Degree in Primary Education, Degree in Secondary Education: Psychology-Mathematics, Master in Educational Technology, Doctor in Education, second approved cycle of Master in Educational Administration, Lawyer with CAL N 6961. She can be contacted at email: rhuerta@une.edu.pe.