International Journal of Computer Trends and Technology
ISSN: 2231 – 2803 / https://doi.org/10.14445/22312803/IJCTT-V70I1P104
Volume 70 Issue 1, 20-24, Jan, 2022

Original Article

Malnutrition among Filipino Children: A Predictive Analysis using Data Mining Approach

Racquel L. Pula[1], Rosanna A. Esquivel[2]

[1]*Assistant Professor, Nueva Ecija University of Science and Technology, Cabanatuan City, Philippines*
[2]*Professor, Angeles University Foundation, Angeles City, Philippines*

Received Date: 13 December 2021
Revised Date: 14 January 2022
Accepted Date: 27 January 2022

Abstract — *Despite being one of the fastest-growing economies, the Philippines still experiences pressing health issues such as malnutrition, which is characterized as either overweight or underweight and generally defined as a nutritional deficiency. As a result, various government agencies, in collaboration with Local Government Units (LGUs) and other partner Non-Governmental Organizations (NGOs), are working to address these nutrition-related issues. In Cabanatuan City, the City Nutrition Committee (CNC) prepared the City Nutrition Action Plan (CNAP) to improve the nutritional status of those tested and address the core causes of malnutrition. With this setting in mind, the goal of this study was to use a data mining approach to do a predictive analysis of malnutrition among children aged six (6) years and under in Cabanatuan City. The study was carried out at 19 Day Care Centers (DCC) in District IV of Cabanatuan City. The population of children designated as malnourished – either underweight or overweight – from 2016 to 2018 was used in the prediction analysis. In this work, data mining and clustering techniques were applied. The analysis showed that in the three (3) years following the period under survey, underweight cases would be decreasing while overweight cases would be increasing. This study concludes that data mining and clustering methods, with the help of predictive analysis assessment, are appropriate in predicting certain phenomena, such as malnutrition cases. However, it is recommended that further studies be conducted to determine and explore more variables that bear on the malnutrition among children in the City of Cabanatuan. Moreover, an in-depth predictive analysis of malnutrition in the area may also be conducted. The analysis results help develop policies and regulations that will help overcome the malnutrition problem in the country.*

CCS Concept — *Data Mining → K-means clustering → Predictive Analysis.*

Keywords — *Cabanatuan city, Data mining, K-means, Clustering, Predictive analysis.*

I. INTRODUCTION

Malnutrition is one of the most pressing health issues in the Philippines; it is most commonly characterized as either undernutrition or overnutrition, yet, it can also be defined as a nutritional deficiency [1]. Malnutrition is a severe global health problem, particularly in developing nations, because it is linked to inadequate diet, an unpleasant environment, and, most typically, poverty. Malnutrition is the leading cause of death for about 54 percent of children worldwide, primarily in underdeveloped countries, according to a World Health Organization (WHO) research from 2004.

Since they are the most vulnerable due to their high dietary requirements for growth and development, women and children's nutritional condition is critical [1]. A malnourished mother is thought to have a high chance of having malnourished offspring, who are then more susceptible to illnesses. Undernutrition puts children at risk for severe infections and illnesses, makes them weak to recover from diseases, and even puts them in danger of death. Malnutrition was cited as the cause of 95 children's deaths in the Philippines [6]. In fact, malnutrition considerably increased the death rate among children in the 1990s because malnourished people were more likely to contract infectious infections [7]. As a result, it is believed that malnutrition was responsible for around half of all child deaths during the said period.

Little has changed worldwide since the establishment of the UN International Children's Emergency Fund (UNIC). Malnutrition is measured using various metrics, including the following three indicators: underweight people who are below their weight requirement for their age; wasting people who are underweight for their height, and stunted people who are underweight for their height underweight for their age. However, the most common metric for measuring a child's nutritional condition is their weight.

In the Philippines, a developing country considered one of the fastest-growing economies in the Southeast Asian region and the whole world, the problem of malnutrition is still present [20]. The improvement in the country's economic standing has not translated into improvements in Filipinos' nutritional status. With the Philippines suffering from a triple burden of malnutrition together with other forms of undernutrition (including stunting and wasting), micronutrient deficiencies, along with overweight and obesity, the Department of Health (DOH), National Nutrition Council (NNC), FAO, WHO, and UNICEF jointly call upon the public, civil society organizations, academe, and the private sector to take action to prevent and manage childhood overweight and obesity. [11].

According to the Philippine Food and Nutrition Research Institute (FNRI) [5], about 31.2 percent of Filipino children aged 5 to 10 are underweight and stunted. In comparison, 26 percent of children aged 0 to 2 years old are underweight and stunted. The Philippines was ranked 88th out of 132 nations with the highest frequency of stunting, according to the 2016 Global Nutrition Report. As a result, various government agencies, in collaboration with Local Government Units (LGUs) and other partner Non-Governmental Organizations (NGOs), are working to address these nutrition-related issues by monitoring children's first 1,000 days, which correspond to the period when they require the most attention due to the increasing nutritional needs for growth and development.

Cabanatuan City, with a registered population of around 314,318 people, a total workforce of 182,235 people, and an employment rate of 89.20 percent as of 2017, is also working to improve its nutritional status and is willing to address the city's core causes of malnutrition. The city's leading source of income is farming, with a gross harvest of 98,577.87 metric tons of rice every year, and is subdivided into districts, namely District I, District II, District III, District IV, and District V [3].

The City Nutrition Committee (CNC) prepared the City Nutrition Action Plan (CNAP) to improve the nutritional status of those tested and address the core causes of malnutrition. The CNAP is a collection of nutrition programs and projects based on reports from Cabanatuan City member agencies and information obtained by volunteers from the City's various barangays. As a result, the LGU of Cabanatuan City is implementing many programs to improve children's nutritional status. The City Health Office (CHO) revealed that the City's child malnutrition rate is lower than other cities in the Philippines [2]. With this setting in mind, the goal of this study was to use a data mining approach to do a predictive analysis of malnutrition among children aged six (6) years and under in Cabanatuan City. This research aimed to provide answers and outcomes to the following questions:

A. How may data mining be used to create data for the study?

B. How can the k-means clustering algorithm be used to group the identified data retrieved by data mining?

C. How can the clustered data be utilized to forecast the rate of malnutrition among Cabanatuan City children?

II. METHODS

The research was carried out at 19 Day Care Centers (DCC) in District IV of Cabanatuan City. The population of children designated as malnourished – either underweight or overweight – from 2016 to 2018 was used in the prediction analysis.

In this work, data mining and clustering techniques were applied. Data mining, also called knowledge discovery or data discovery [22], is the process of extracting a vast amount of data to create a data pattern [13, 14]. Association rule mining, data clustering, and data categorization are all data mining examples [16]. In addition, since the healthcare system uses and produces a huge amount of data, a well-speedy health decision-making tool should be used, where data mining can be applied and used for this purpose [17]. On the other hand, predictive analysis is one of one the emerging field of analytics that uses statistical algorithms to make predictions and forecasts based on historical data [15].

Clustering was the most appropriate data mining strategy for this study, which attempted to treat malnutrition among children aged 1-6 years old. As defined by [8, 18], clustering is the process of arranging data into clusters based on similar concepts, patterns, and other factors. Clustering is done to combine two separate data sets into a single one. K-means clustering is one of the most successful approaches for clustering data, with good clustering results in various applications [19]. Because of its simplicity in operations, the K-means data mining method is a popular and efficient approach that may be used in any field where information needs to be extracted. Furthermore, K-means clustering is used to analyze unlabeled data [21, 23]. This strategy was utilized in this study because it allows for the creation of predefined numbers of clusters, which reduces the size of large data sets [24, 25].

K-means can be utilized using the following steps:

Step 1: Define the k, or the number of clusters, for k-means clustering.

Step 2: A random sample of the data is used to create the initial partition that classifies the data into cluster k. This can also be done systematically by clustering the initial data sample k, collecting the remaining training samples (N-k) into the nearest centroid, then recomputing the new centroid using the newly acquired cluster.

Step 3: Perform each sample in order and calculate the distance from each group's centroid centre. If a sample is incompatible with the cluster closest to the centroid, it must be replaced with a new cluster, with the new sample and the sample loss cluster being used to update the centroid point.

Step 4: Repeat step 3 until the target value is obtained, i.e. the training sample matches and no new job is assigned.

The k-means method is based on the following [13]:

$$arg_s min \sum_{i=1}^{k} \sum_{X_j \ni S_i} \left\| X_j - \mu_i \right\|^2$$

The equation for a k-means algorithm.

Where: **(X_1, X_2…X_n)** are the observation results that represent a cluster element with a real d dimension vector;

n is the number of observations where the observed value to k set (k <= n) S = {S_1, S_2, ... S_k}; and

μ_i is the mean value of the point at S_i.

The clustered data were treated to regression analytics to develop a model that accounts for the malnutrition rate among children aged 6 and under in the study area. Regression analysis is a way of mathematically sorting out which of those variables does indeed have an impact. A linear regression model was utilized to do this, which establishes a linear relationship between the variables studied.

III. RESULTS AND DISCUSSION

Data were collected from the City Social Welfare and Development Office of Cabanatuan City through their Day Care Centers. These data report the number of underweight and overweight per year. Data mining was also utilized in this study to predict the various procedures and data validity, and it improves decision-making using patterns and trends [10].

The table below shows the total cases of malnutrition (in terms of overweight and underweight) in the District IV of Cabanatuan City for the period covered by the study.

These samples were grouped according to their weight status and clustered using the K-means clustering method.

K-means clustering is the simplest and most common algorithm used to group objects by attributing or featuring into k number of clusters, where k is a positive integer and will be defined by the user. The grouping minimized the sum of squares distance between the data and the appropriate centroid cluster [13].

Table 1. Distribution of children with malnutrition

DCC No.	Barangay Day Care Centers	Malnutrition Case		
		2016	2017	2018
1	Macatbong (Cherubin DCC)	4	3	4
2	Bagong Sikat (Pineda DCC)	5	0	0
3	Communal (Communal DCC)	0	0	1
4	Valle Cruz (Blessed DCC)	0	0	0
5	Pangatian (St. Vincent DCC)	11	5	3
6	Obrero (Divine Grade DCC)	4	2	1
7	Bantug Norte (Children of the Lord DCC)	12	0	10
8	Patalac (Patalac DCC)	1	1	1
9	Bakod Bayan (Little Lamb DCC)	5	0	5
10	Bagong Buhay (St. Mary DCC)	0	0	4
11	San Isidro (Love of Angel DCC)	0	0	1
12	Cabu (St. Claire DCC)	0	0	3
14	Cabu Bana (Sto Nino DCC)	0	0	0
14	Kalikid Sur (Crowning Glory DCC)	6	0	2
15	Bangad I (Angelus DCC)	4	5	2
16	Bangad (St. Joseph DCC)	3	0	4
17	San Isidro (Shining Star DCC)	0	0	3
18	Bakod Bayan 2 (Bulilit DCC)	0	0	0
19	Calauagan (Little Heart DCC)	0	0	3

The figure that follows shows the cases of malnutrition in a graphical form.

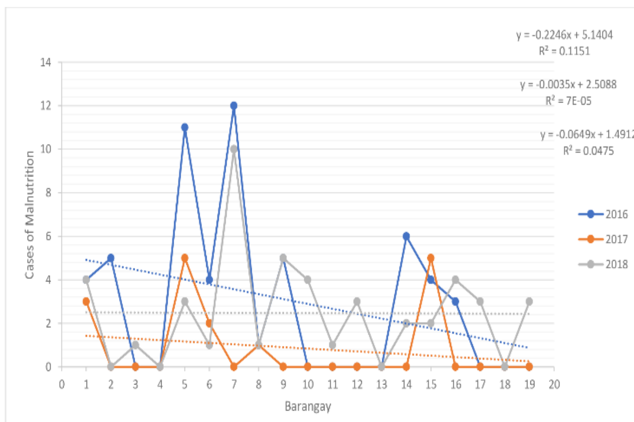

Fig. 1 Malnutrition cases in cabanatuan city - district IV

It can be seen from the figure the highest number of recorded malnutrition cases in District IV of Cabanatuan City was in 2016. The case recorded reached a peak of 12. On the other hand, in 2017, there was a decrease in the number of cases of malnutrition recorded in the research area. Unfortunately, it can also be seen in the graph that in 2018, there was an increasing number of malnourished children.

The clustered data were subjected to regression analysis, where a linear model was utilized to arrive at the malnutrition rate in the research area.

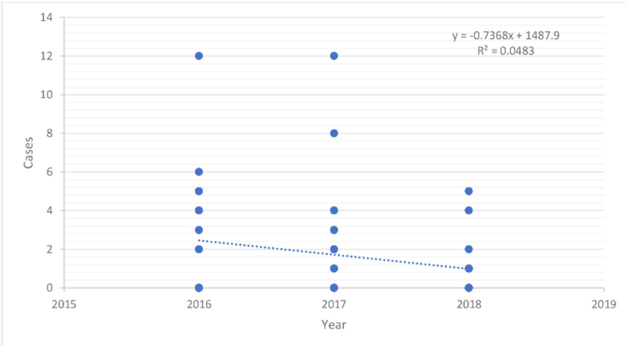

Fig. 2 Regression analysis of underweight cases

Figure 2 indicates the regression analysis of the underweight cases in Cabanatuan City-District IV. It is reflected in the figure that the variations were more pronounced during 2016. This can be attributed to Typhoon Sarika and Super Typhoon Haima (local names: Karen and Lawin, respectively), which hit the City and affected almost all of its barangays by bringing floodwaters that led to the destruction of agricultural crops, which includes rice crops.

Table 2. The volume of palay production in Nueva Ecija: 2016-2018 (in metric tons) [8]

Year	Palay Production
2016	1,684,352
2017	1,884,091
2018	1,886,709

Table 2 shows that the palay production for 2016 for the province of Nueva Ecija, where Cabanatuan City is situated, was significantly lower compared to 2017 and 2018. This factor would have caused significant impacts on the availability of the production of rice products. In addition, it negatively impacts the nutrition of children in the research locale and affects the weight status of the children, thus increasing the incidence of malnutrition in the area. Malnutrition incidence is higher among children who were weaned at a very young age and were introduced to new diets.

When the data were subjected to regression analysis, malnutrition cases, specifically underweight occurrences, are predicted to decrease in the years following the period in a survey. This is shown by the table below that statistically regressed the data collected from the CSWDO.

Table 3. Prediction analysis of underweight cases

Year	Predicted Cases
2019	0.3008
2020	-0.436
2021	-1.1728

On the other hand, overweight cases were also subjected to regression analysis.

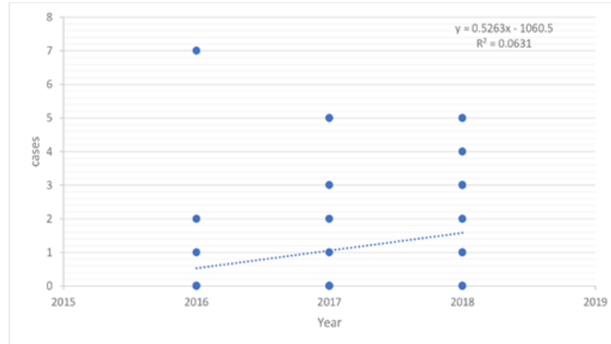

Fig. 3 Regression analysis of overweight cases

Figure 3 shows the trends of overweight cases in the District IV of Cabanatuan City that indicate an increasing trend of overweight occurrences. In 2016, there was a report where overweight cases in District IV reached a peak of 11, while in 2017 and 2018, it only reached 8 cases. However, what matters here is the number of cases occurring, and as the figure shows, 2018 has the most significant number of malnutrition cases.

The regression analysis also showed the predicted cases of overweight occurrences in the area in the coming years. Based on the statistics, overweight cases in Cabanatuan City - District IV will increase in the future.

Table 4. Prediction analysis of overweight cases

Year	Predicted Cases
2019	2.0997
2020	2.626
2021	3.1523

IV. CONCLUSION

In summary, data mining and clustering were utilized in this study as the first two steps of analysis of the malnutrition status of children in the District IV of Cabanatuan City. This most suited clustering method for the study is the k-means clustering, where samples were grouped according to their weight status. Predicted analyses were also utilized to determine the trends of malnutrition cases in the City using the 2016 to 2018 data collected from the City Social Welfare and Development Office. The analysis showed that in the three (3) years following the period under survey, underweight cases would be decreasing while overweight cases would be increasing. This study concludes that data mining and clustering methods, with the help of predictive analysis assessment, are appropriate in predicting certain phenomena, such as malnutrition cases. However, it is recommended that further studies be conducted to determine and explore more variables that bear on the malnutrition among children in the City of Cabanatuan. Moreover, an in-depth predictive analysis of malnutrition in the area may also be conducted. The analysis results help develop policies and regulations that will help overcome the malnutrition problem in the country.

REFERENCES

[1] Blosser, M., Onis, M., Malnutrition: quantifying the health impact at national and local levels. World Health Organization (2005).

[2] Cabanatuan City Government, City Nutrition Plan. Technical Report. Cabanatuan City Health Office (2019).

[3] Cabanatuan City Government, Socio-Economic Profile(2019).

[4] Chakraborti S, Gupta SB, Chaturvedi B, Chakraborty SK. A study of protein-energy malnutrition (PEM) in children (0 to 6 years) in a rural population of Jhansi district (UP). Indian J Community Med. 31(4) (2006) 291-2

[5] Food and Nutrition Research Institute, 2015. National nutrition survey.

[6] Lina, Joey D., Malnutrition: a gnawing crisis needing urgent action. Retrieved from https://www.pressreader.com/ (2018).

[7] Pelletier, D. L., Frongillo, E. A., Habicht, J. P., Epidemiologic evidence for a potentiating effect of malnutrition on child mortality (1993).

[8] Philippine Statistics Authority, Volume of Palay Production in Central Luzon, July to December (2018). Retrieved from http://rsso03.psa.gov.ph/article/volume-palay-production-central-luzon-july-december-2018

[9] Shukla N, Toppo N.A., Thakur A, Kasar P.K. A study of malnutrition and associated risk factors among children of age 06-59 months in an urban area of Jabalpur district (M.P.). Int J Community Med Public Health (5) (2018) 296-301.

[10] Singh, S., S. S. Gill, Analysis and study of a k-means clustering algorithm. International Journal of Engineering Research and Technology. 2(7) (2013).

[11] UNICEF, Philippines. Poor diets, failing food systems, and lack of physical activity are causing overweight and obesity in children. Retrieved from https://www.unicef.org/philippines/press-releases/poor-diets-failing-food-systems-and-lack-physical-activity-are-causing-overweight

[12] United Nations International Children's Emergency Fund, Malnutrition rates remain alarming: stunting is declining too slowly while wasting still impacts the lives of far too many young children (2018). Retrieved from https://data.unicef.org/topic/nutrition/malnutrition/.

[13] Winiarti, Yulianshi, Purnama, Identification of toddlers' nutritional status using data mining approach. International Journal of Advanced Computer Science and Application, 9(1) (2018)

[14] S. M. Metevand V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag. 5(3) (1998) 300-320.

[15] Swani, L., Tyagi, P., Predicted Analytics through Data Mining. International Journal of Engineering and Technology. 04(09) (2017) ISSN 2395-0072

[16] Vijayukumar, V., Predictive Learning Analytics for Student Success towards Outcome Based Education Perspective. International Journal of Computer Trends and Technology, 69(9) (2021) 7-11, ISSN 2231 – 2803

[17] Purnapatre, H., Shukla, S., Pareek, S., Predictive Analysis, Predictive Analysis as a Solution to Branding. International Journal of Computer Trends and Technology, 68(12) (2020) 47-63. ISSN 2231 – 2803

[18] Daniel, A. O., Uwem, E. E., Tululope, O. D., Jesupelumi, O. S., 2020. Prediction and Diagnosis of Liver Disease using Machine Learning. International Journal of Computer Trends and Technology, 68(8) (2020). ISSN 2231 – 2803

[19] Sharma, A., Sharma, N., Text Summarization using K-Means, Tanimoto Distance and Jacard Similarity. International Journal of Computer Trends and Technology, 68(6) (2020). ISSN 2231 – 2803

[20] Singh, K., Malik, D., Sharma, N., Evolving Limitations in K-Means Algorithm in Data Mining and their Removal. International Journal of Computation Engineering and Management, 12 (2011). ISSN 2230-7893

[21] Inter-Agency Regional Analysts Network, Asia Report: Socio-Economy of Chronic Malnutrition in the Philippines: A Preliminary Key Trends Analysis by 2030 (2016).

[22] Shalini, L., Goplai Naga Sravya, Analysis of Health-Tweets using K-means Clustering. International Journal of Engineering and Technology. 05(08) (2018). ISSN 2395-0072

[23] Narang, B., Verma, P., Kochar, P., Application-Based, Advantageous K-Means Clustering Algorithm in Data Mining – A Review. International Journal of Latest Trends in Engineering and Technology. 7(2) (2016). ISSN 2278-621X

[24] Ogbuabor, G., Ugwoke, F. N., Clustering Algorithm for a Healthcare Data Set using Silhouette Score Value. International Journal of Computer Science and Information Technology, 10(2) (2018).

[25] Haraty, R. A., Dimishkieh, M., Masud, M., An Enhanced k-Means Clustering Algorithm for Pattern Discovery in Healthcare Data. International Journal of Distributed Sensor Networks, 15 (2015). http://dx.doi.org/10.1155/2015/615740

[26] Kanungo, T., Netanyahu, N. S., Wu, A. Y., An Efficient k-MeNS Clustering Algorithm: Analysis and Implementation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7) (2002).