IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10793

An Innovative Application to Predict Malnutrition and Anemia using ML

Kavya Priya M L[*1], Chaithra[2], Ganavi M[3], Shreyaswini P[4], Prathibha B[5]

[*1] Assistant Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India

[2, 3, 4, 5] UG Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India

Abstract Malnutrition is regarded as a crucial health issue of a nation considering today's children as tomorrow's youngsters and leaders who directly influence the fruitful development of the nation. The exact amount of nutrition is the most required element for the survival, development, growth of the children in this society. The foremost goal of this proposed phenomenon in children less than five years is suffering from a lack of nutrition. The problem of Malnutrition and anemia is majorly found in under-developed and developing countries. To overcome this problem we make use of various Machine learning and data mining approaches to predict the malnutrition condition of a child less than five years based on the training data-sets. Training data-sets downloaded from www.kaagle.com.Various factors such as Gender, Age, HAZ, WAZ, etc…are extracted. Classification techniques used for malnutrition status prediction. We use algorithms such as "Bayesian classifier" and "K-nearest neighbor" for prediction. The results will be compared and an efficient algorithm will be identified. With the help of testing and Checking knowledge, precautionary measures will be led with the aid of medical practitioners to minimize the anemia and malnutrition condition in a child. We build this as a real-time software application useful for society. To build the real-time application we use technologies such as "Visual Studio" for the front end and "SQL server" for the back end. Both of these tools are powerful tools to work with the real-time application. This system specifies how the classification methods can classify anemia and malnutrition condition of children below five years of age. In General, our proposed method is designed in such a way that, the most accurate results are obtained to find the malnutrition and anemia status based on data sets collected.

Keywords Malnutrition, Anemia, KNN, Naive Bayes, Data mining, Classification rules

I.INTRODUCTION

In today's world nutrition is most important and children should take proper nutrition for proper growth, development, and survival. The current system is a manual process of child analysis and leads to less accurate results and not an appropriate method for malnutrition prediction. The current system is involved tedious tasks and involves more time and more expensive. Many factors affect children's malnutrition such as age, gender, height, weight, WAZ, HAZ, WHZ, etc… Malnutrition detection is important in today's world. The current system is a manual process of child analysis and leads to less accurate results and not an appropriate method for malnutrition prediction. The current system involves tedious tasks and involves more time and more expensive. There is no automation for malnutrition status prediction of children.

The system classifies children into stunted, underweight, wasted, and nutritional oedema statuses. We use parameters such as age, gender, height, weight, WAZ, HAZ, WHZ, etc.. classifiers are used for malnutrition and anemia prediction of children. We use efficient classifiers such as Bayesian classifier, KNN classifier, or Random Forest classifier. We collect data sets from online sources such as "kaagle.com", "Dataworld.com" etc... We can develop this as an application useful for doctors. We use efficient technologies such as "visual studio" and "SQL server" for application development.

In the traditional software approach, no written code indicates the system how to decide to pick the right object among different objects since it's difficult to manage each situation on an object. It uses the provided data to make intelligent decisions, imply future predictions, and recognize anomalies. These days it is being used in several applications such as search engine results, virtual assistance, identifying objects, advertising, predictive analysis, autonomous cars, etc... There exists 3 ML strategies viz Supervised Learning, Unsupervised Learning, and Reinforcement Learning. The proposed system uses data science algorithms for prediction.

II.RELATED WORK

[1] METHODOLOGY: Many algorithms such as RBF-SVM, Decision tree, K nearest neighbor, Randomforest, K-star.
LIMITATIONS:

- RF (Random Forest) algorithm is efficient but the execution is slow.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10793

- It is not suitable for real-time prediction.

[2] METHODOLOGY: Applying different classification techniques like PART rule induction, RF, logistic regression, and Naïve Bayes to find the malnutrition condition.
LIMITATIONS:

- As seen in the previous paper RF algorithm has its limitation in real-time execution.
- The method only predicts the malnutrition status of the child but how to overcome the problem remains a question.

[3] METHODOLOGY: Four different techniques namely Naïve Bayes, Bayesian Network, Multi-Layer Perceptron, and Logistic Regression are being used.
The prediction is done based on previously available data that is 539 samples with 10 different parameters.
LIMITATIONS:

- Only two types of anemia that is iron deficiency and vitamin deficiency anemia are being considered here.
- The data set considered here is quite small. Hence the results are not reliable.
- Sometime MLP technique may result in redundancy and inefficiency.

[4] METHODOLOGY:
This system uses a deep autoencoder network with an accuracy of 80%.
LIMITATIONS:

- The autoencoder network uses the images as input as it is a deep learning algorithm.

[5] METHODOLOGY:
The 2-Varogram algorithm programmed using the IDL language is used.

LIMITATIONS:

- The system only talks about the special analysis of the nutrition epidemiology and it gives only the distribution of a variety of data.

III.PROPOSED WORK

Essential nutrition is the most required element for the growth, survival, and development of a child. Both malnutrition and anemia are worldwide issues in the present life. The proposed system's major objective is to forecast the anemia and malnutrition condition in beneath 5 years olds. The system classifies children into wasted, underweight, stunted, and nutritional oedema statuses. We use parameters such as age, gender, height, weight, WAZ, WHZ, HAZ, etc.. classifiers used for malnutrition prediction of children. We use efficient classifiers such as Bayesian classifier, KNN classifier, or Random Forest classifier. We collect data sets from online sources such as "kaagle.com", "Dataworld.com" etc... We can develop this as an application useful for doctors. We use efficient technologies such as "visual studio" and "SQL server" for application development.

A. PARAMETER LIST

TABLE II. FINAL DATASET ATTRIBUTES

#	Feature	Description
1	Sex	Sex of child
2	Birthdate	Date of birth of child
3	Visit_date	Date of survey
4	Age	Age of child in month
5	Weight	Weight of child in Kg
6	Height	Height of child in cm
7	MUAC	Mid-Upper Arm Circumference
8	Age_group	Child age group, 5 distinct groups of age
9	WAZ	Weight-for-Age Z-Score
10	HAZ	Height-for-Age Z-Score
11	WHZ	Weight-for-Height Z-Score
12	Province	Region of the child
13	Stunted	Indicates whether a child is stunted or not
14	Underweight	Indicates if child is underweighted
15	Wasted	Indicates whether a child is wasted or not
16	Oedema	Bilateral pitting oedema in both feet

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10793

B. METHODOLOGY

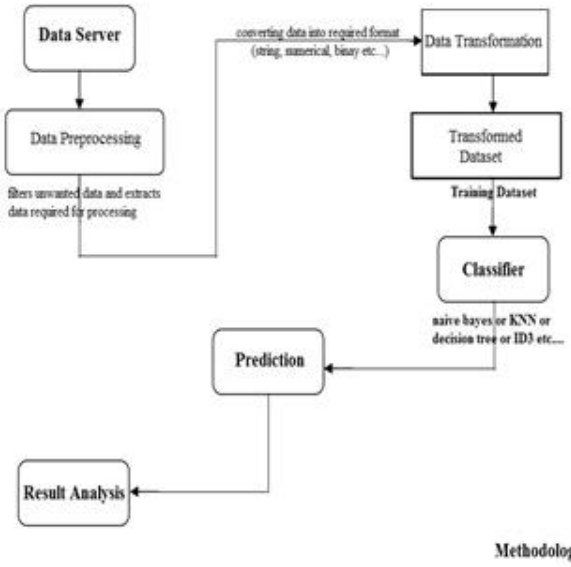
a. DATA MINING

Data mining refers back to the method of studying data from numerous views and extracting convenient information from the processed data. Data mining is applied to n number of fields and used to solve real-world problems. Data mining supports many techniques.

In the project we use "Data Mining Classification Rules" to process data and for prediction.

b. CLASSIFICATION RULES (CLASSIFIERS)

Classification facilitates categorizing every object in a group of data into one of the predetermined sets of classes, instances, otherwise additional groups. These techniques use mathematical strategies for problem-solving.
Ex: Employee statuses in a company (leaves or stay)
To forecast which working employees are most likely to leave the company in the future.
Here we use either "Naive Bayes" or "KNN" or "Decision tree" classifier to process previous data and for prediction. These specified algorithms are most efficient and take slighter time for processing the data. These algorithms work well and good for n number of parameters.



Methodology

NAIVE BAYES ALGORITHM

Step 1: Examine datasets (storage servers)
Retrieving essential data for mining from various servers such as excel sheets, clouds, databases, etc.
Step 2: Compute the probabilities of every attribute value. [n, n_c, m, p]
In this, for each attribute calculate the probability of incedence using the subsequent formula. (mentioned in step 3). For each class(i.e disease) apply the formulae.
Step 3: Applying the formula
P(attributevalue(ai)/subjectvaluevj)=(n_c + mp)/(n+m)
Where:
n = number of training examples for which v = vj
nc = number of training examples to whichv = vj and a = ai
p = a priori estimationto P(aijvj)
m = equivalent sample size
Step 4: Multiplying probabilities by p
For every single class, more than one the consequences of every attribute with p and ending consequences are used for classification purpose.
Step 5: Differentiate the values and characterize the characteristic values into one of the predetermined group of class values.

KNN ALGORITHM

The working of KNN is explained in the following steps:
Step1: Choose the K value.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10793

Step2: Compute Euclidean distance of K number of neighbors
Step3: Extract K's nearest neighbors according to the computed Euclidean distance.
Step 4: Out of those k neighbors, sum up the wide variety of data points in each class.
Step5: Place new points to the class to which the range of the neighbor is most.
Step6: Finally, the model is ready.

C. EXPERIMENTS RESULTS

a. MALNUTRITION PREDICTION - NAIVE BAYES ALGORITHM RESULTS

Constraint	Naive Bayes Algorithm
Accuracy	94.4736842105263%
Efficiency (milli secs)	37253
Precision	94.4736842105263%
Recall	5.526315789473 69%

b. ANEMIA PREDICTION - KNN ALGORITHM RESULTS

Constraint	Algorithm
Accuracy	96.92857142857 14%
Time (milli secs)	14616
Correctly Classified	96.92857142857 14%
InCorrectly Classified	3.071428571428 57%

CONCLUSION

Essential nutrition is a required element for the survival, growth, and development of children. Malnutrition is a international issue in this era of lifestyles. The proposed system's major objective is forecasting malnutrition and anemia conditions of a child below five years of age. The system classifies children into nutritional oedema, stunted underweight, wasted statuses. We use parameters such as age, gender, height, weight, WAZ, HAZ, WHZ, etc... Classifiers used for malnutrition prediction of children. The system also predicts anemia and suggests suitable dietary recommendations for the users. The system is a real-time application useful for the medical sector.

FUTURE ENHANCEMENTS

Additionally, more algorithms can be used and can compare the algorithms to identify the efficient algorithm. More training datasets can be used for prediction.

REFERENCES

[1] Classification of Pathological Disorders in Children using RF Algorithm. Sujit Bebortta, Manoranjan Panda, Shradhanjali Panda.
[2] Data Mining Based Prediction of Malnutrition in Afghan Children. Ziaullah Momand, Pornchai Mongkolnam, Pichai Kositpantavong, Jonathan H. Chan.
[3] Analysis of Anemia Using Data Mining Techniques with Risk Factors Specification. Mohammed Sami Mohammed, Arshed A. Ahmad, Murat Sari.
[4] Towards Computer Vision Powered Color-Nutrient Assessment of Pureed Food. Kaylen J Pfisterer, Robert Amelard, Braeden Syrnyk, Alexander Wong.
[5] Spatial Analysis Applied to Nutritional Epidemiology. Eliana Marina Alvarez Di Fino, Maria Daniela Defago, Carlos Marcelo Scavuzzo.
[6] "Malnutrition in Children". (n.d.). Retrieved June 13, 2019, from https://data.unicef.org/topic/nutrition/malnutrition/.
[7] "Joint child malnutrition estimates - Levels and trends (2019 edition)". (2019, June 06). Retrieved June 13, 2019, from https://www.who.int/nutgrowthdb/estimates2018/en/.
[8] Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., & Sachdev, H. S. (2008). "Maternal and child undernutrition: consequences for adult health and human capital". The Lancet, 371(9609), 340–357.
[9] Ministry of Public Health, Afghanistan. (2014). "Integrated Guidelines for the Management of Acute Malnutrition".
[10] Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., De Onis, M., Ezzati, M., & Maternal and Child Undernutrition Study Group. (2008). "Maternal and child undernutrition: global and regional exposures and health consequences". The Lancet, 371(9608), 243-260.
[11] Nandy, S., Irving, M., Gordon, D., Subramanian, S. V., & Smith, G. D. (2005). "Poverty, child undernutrition and morbidity: new evidence from India". Bulletin of the World Health Organization, 83, 210-216.
[12] Maleta K. (2006). "Undernutrition". Malawi medical journal: the journal of Medical Association of Malawi, 18(4), 189–205.
[13] "Malnutrition: Afghanistan's silent emergency". Retrieved June 13, 2019, from https://www.unicef.org/afghanistan/nutrition.
[14] Ministry of Public Health, Afghanistan. (2019). "Afghanistan Nutrition Cluster, 2018 annual report".
[15] Ministry of Public Health (MOPH) - Afghanistan. (2004). 2004 "Afghanistan National Nutrition Survey".
[16] Ministry of Public Health (MOPH) - Afghanistan. (2013). 2013 "Afghanistan National Nutrition Survey".
[17] Ministry of Public Health (MoPH), ICF, Central Statistics Organization. (2017). "Afghanistan Demographic and Health Survey 2015".
[18] Higgins-Steele, A., Mustaphi, P., Varkey, S., Ludin, H., Safi, N., & Bhutta, Z. A. (2016). "Stop stunting: situation and way forward to improve maternal, child and adolescent nutrition in Afghanistan". Maternal & Child Nutrition, 12, 237–241.
[19] Akseer, N., Bhatti, Z., Mashal, T., Soofi, S., Moineddin, R., Black, R. E., & Bhutta, Z. A. (2018). "Geospatial inequalities and determinants of nutritional status among women and children in Afghanistan": observational study. The Lancet GlobalHealth, 6(4), e447–e459.

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10793

[20] Mashal, T., Takano, T., Nakamura, K., Kizuki, M., Hemat, S., Watanabe, M., &Seino, K. (2008). "Factors associated with the health and nutritional status of children under 5 years of age in Afghanistan: family behavior related to women and past experience of war-related hardships". Public Health, 8(1).
[21] Ministry of Public Health, Afghanistan. (2018). "Integrated Management of Acute Malnutrition National Guidelines". Kabul, Afghanistan: Public Nutrition Directorate.
[22] Khare, S., Kavyashree, S., Gupta, D., & Jyotishi, A. (2017). "Investigation of Nutritional Status of Children based on Machine Learning Techniques using Indian Demographic and Health Survey Data". Procedia Computer Science, 115(2017), 338–349.
[23] Das, S., Hossain, M. Z., & Islam, M. A. (2008). "Predictors of child chronic malnutrition in Bangladesh". Proceedings of the Pakistan Academy of Sciences 45(3): 137-155.
[24] Shahriar, M., Iqubal, M. S., Mitra, S., & Das, A. K. (2019, July). "A Deep Learning Approach to Predict Malnutrition Status of 0-59 Month's Older Children in Bangladesh". In 2019 IEEE International Conference on Communications Technology (IAICT) (pp. 145-149). IEEE.
[25] Ariyadasa, S. N., Munasinghe, L. K., Senanayake, S. H., & Fernando, M. G. N. A. S. (2013, January). "Knowledge Extraction to Mitigate Child Malnutrition in Developing Countries (Sri Lankan Context)". In 2013 4th International Conference on Intelligent Systems, Modelling, and Simulation (pp. 321-326). IEEE.
[26] Reis, R., Peixoto, H., Machado, J., & Abelha, A. (2017). "Machine Learning in Nutritional Follow-up Research". Open Computer Science, 7(1), 41–45.
[27] Markos, Z. (2014). "Predicting Under Nutrition Status of Under-Five Children Using Data Mining Techniques: The Case of 2011 Ethiopian Demographic and Health Survey". Journal of Health & Medical Informatics, 5(2).
[28] Thangamani, D., & Sudha, P. (2014). "Identification of malnutrition with use of supervised data mining techniques–decision trees and artificial neural networks". Int J Eng Comput Sci, 3(September 9), 8236-824.
[29] Cios, K. J., Pedrycz, W., Swiniarski, R. W., & Kurgan, L. A. (2007). "Data Mining: A Knowledge Discovery Approach".
[30] "WHO Anthro Survey Analyser and other tools". (2019, June 7). Retrieved September 3, 2019, from https://www.who.int/childgrowth/software/en/.
[31] Nitesh V. Chawla et. al. (2002). "Synthetic Minority Over-sampling Technique". Journal of Artificial Intelligence Research. 16:321-3.
[32] World Health Organization. (2010). "Nutrition Landscape Information System. Geneva, Switzerland": WHO.

