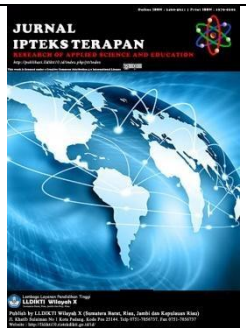

Vol. 17 No.4 2023

ANALYSIS OF STUNTING SYMPTOMS IN EARLY CHILDHOOD WITH CLASSIFICATION TECHNIQUES

Aulia Fitrul Hadi*[1], Silfia Andini[2], Silvia Gusriyeni[3]
[12] Universitas Putra Indonesia - Indonesia
[3] Kazan Federal University - Russia
*Email: fitrulhadi@upiyptk.ac.id

Submit:
17/10/2023

Accept:
27/12/2023

Publish:
31/12/2023

Abstract

WHO estimates that the total number of disease cases that occur at a certain time in an area (prevalence) of Stunting (dwarf toddlers) throughout the world is 22 percent or 149.2 million people in 2020. In Indonesia, based on data from the Asian Development Bank, in In 2022, the percentage of prevalence of stunting among children under 5 years of age in Indonesia will be 31.8 percent. This number causes Indonesia to be in 10th place in the Southeast Asia region. Furthermore, in 2022, based on data from the Ministry of Health, Indonesia's stunting rate has decreased to 21.6 percent. According to the World Health Organization (WHO), stunting is a growth disorder in children due to poor nutrition, recurrent infections and inadequate psychosocial stimulation. The classification method is used to group factors that greatly influence the incidence of stunting in children. Using the C.45 Algorithm, this research found the main factors that can influence children's growth capacity at an early age. By using 7 attributes, namely nutritional status, age, gender, poor status, weight, height, category and obtained an accuracy rate of 88%. This study was able to produce an average accuracy value of 97.87% using 10-fold cross validation

Keywords: Sleep Patterns, Screen Time, Toddlers, Technology, Education

http://publikasi.lldikti10.id/index.php/jit DOI : https://doi.org/10.22216/jit.v17i4 PAGES : 862-868

ISSN : 1979-9292 / 2460-5611
JIT is Sinta 3 Journal (https://sinta3.kemdikbud.go.id/journals/profile/2143) accredited by the Ministry of Research & Technology, Republic of Indonesia

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V17.i4 (862-868)

ISSN : 1979-9292

E-ISSN : 2460-5611

INTRODUCTION

The rapid development of information technology in all areas of life is directly proportional to the data it generates. Starting from the industrial sector, health and various other fields. With the application of information technology in the world of health and medical can generate abundant data. These data can be in the form of data about a disease or a certain medical condition. As is the case with the PuskemasHealth Center which carries out posyandu activities every month to provide convenience to the community, especially mothers and toddlers, to obtain basic health services, besides that, posyandu is also an activity for monitoring and collecting data to determine the threshold of nutritional status in children. Quoted from the official website of the Ministry of Health of the Republic of Indonesia, in 2018 the percentage of stunting in Indonesia was at 30.8%, where Indonesia's percentage was far above other ASEAN countries, which ranged from 4% -17%. Established as a stunting locus, the Head of the Padang City Health Service was stunned. Kadiskes Feri Mulyani is determined to overcome stunting in the near future. Whereas in Padang City itself in 2019 quoted from the portal https://sumbarprov.go.id the percentage of stunting was at 25.51% where this figure was far higher than West Sumatra Province which was at 25%. The problem of malnutrition is caused partly because inadequate consumption is seen as an ecological problem which is not only caused by insufficient availability of food and certain nutrients but is also influenced by poverty, poor environmental sanitation and ignorance about nutrition[1] – [3]. The PuskemasCommunity Health Center performs data collection and input manually using Microsoft Excel. Where the resulting size becomes large and the computational process becomes heavier and the community can only check it within a period of once a month when Posyandu activities are carried out. Based on the problems that occur, it is necessary to design a system to monitor the condition of toddlers, especially stunting, with several indicators to support innovation and public understanding of the importance of adequate nutrition for toddlers. With this innovation, it is hoped that parents can measure nutritional indicators for toddlers, especially those related to stunting conditions on a regular basis and do not need to wait once a month for measurements to be taken by local health center staff. Stunting is a condition where a person's height (TB) is not suitable for age, which can be determined by calculating the Z-index score for height according to age (TB/U). Someone is said to be stunted if the Z-index score for TB/U is below -2 SD (standard deviation) [4] - [6].

Comparing it with a country with the lowest level of malnutrition in the world, namely Russia. In the 2023 Global Hunger Index, Russia ranked 26th out of 125 countries with sufficient data to calculate a 2023 GHI score. With a score of 5.8 on the 2023 Global Hunger Index, the Russian Federation has low levels of hunger. The Global Hunger Index (GHI) is a tool to comprehensively measure and track hunger at the global, regional and national levels. The GHI score is based on the values of four component indicators:

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V17.i4 (862-868)

ISSN : 1979-9292

E-ISSN : 2460-5611

Stunting In Children: the number of children under five years of age who are low in height for their age, reflecting chronic malnutrition.
Child Wasting: the number of children under five years of age who are low in weight for their height, reflecting acute malnutrition.
Child Mortality: the number of children who die before their fifth birthday, partly due to a combination of poor nutrition and an unhealthy environment.

Based on the values of these four indicators, the GHI score is calculated on a 100-point scale reflecting the severity of hunger, where 0 is the best score (no hunger) and 100 is the worst score. Each country's GHI score is classified by severity, from low to very concerning.

Parameter measurements are usually carried out once a month through the Posyandu held by the Puskemas Health Center. From the posyandu, measurements will be taken on toddlers which will later obtain parameters for village name, posyandu name, toddler name, gender, age, weight, and height[3]. In Indonesia, the measurement of the nutritional status of under-fives mostly uses the z-socre or Z-index. Z score is a result that shows the measurement of the median [7] – [8]. A random forest is defined as a classification group from a regression tree, trained from the training data use choice feature random in process generate tree . After a number big tree has in- generate , every tree voting For get class Which most popular. Procedure voting tree in a manner collective This defined as random forest . For technique random classification forest This need two parameters ie amount tree and the number of attributes used [9] – [10]. Many trees are grown to form a forest (forest), then analysis done on gathering tree the. Response something predicted with combine results prediction k tree. On problem classification done based on majority vote (voice the most).

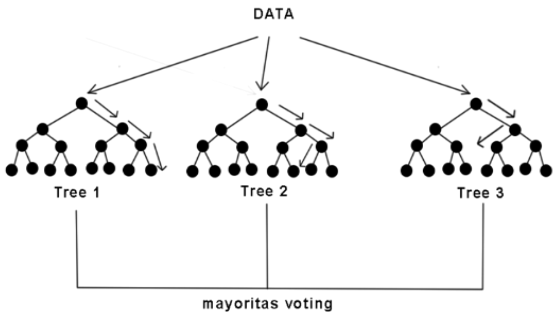

Figure 1. Random Forest Algorithm.

RESEARCH METHODS

Random forest algorithm is inseparable from the decision tree because the random forest algorithm is a collection of the voting majority of several decision trees. Decision tree itself is a collection algorithm conditional nest For find splitting with mark best based on mark entropy (size information in A state) For look for mark maximum from information gain.

$$Entrophy = \sum -pi \log(pi) \quad \ldots\ldots \quad (2)$$

$$IG = (parent) - \sum wi\ E(child\ i) \quad \ldots\ldots \quad (3)$$

Pi in the entrophy formula is the possibility of class I while IG is the information gain is the result of subtracting the entropy of the parent with the entropy of the leaf node (child). Where in random forest several decision trees will be formed , with each result being taken for voting (results with highest number) will made output from algorithm

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V17.i4 (862-868)

ISSN : 1979-9292

E-ISSN : 2460-5611

random forest. K-Fold Cross Validation test is a validation technique by dividing data randomly into k parts and each part will be classified. Where in k is used in this study as many as 3, 5, and 10. Each trial will use one testing data and k-1 part will become data training , Then data testing That will exchanged with One fruit data training so that For each experiment will obtained data testing Which different [11].

RESULTS AND DISCUSSION

Figure 2. fold cross validation

The research method used in this project leads to absolute research that holds on the impact resulting from the experiment [12]. This project uses a random forest algorithm for make an application for predicting stunting conditions in toddlers where previously stunting conditions were measured toddlers in the Puskemasmesih health center use manual calculations. This stage is the stage of understanding the research and the results to be got. Stages data understanding. Stage For gather data.

Number	Index	Category status nutrition	Z *scores*
1.	Weight according age (BBU)	Malnutrition, Nutrition Not Enough, Good, More	<-3SD -3SD up to <-2SD -2SD until with 2SD> 2SD
2.	Long body according to age (TB/U)	Very short Short Normal Tall	<-3SD -3SD up to <-2SD -2SD until with 2SD >2SD
3.	Weight according Long body (BB/TB)	Very short Short Normal Fat	<-3SD -3SD up to <-2SD -2SD until with 2SD >2SD

Table 1. Statuses child nutrition based index[3]

Stage data preprocessing so that data Which got can processed with algorithm Which used in case this time use algorithm random forest.

a. Data cleaning
   Is which stage data will be cleaned from data blank

b. Data labeling
   Data labeling serves to determine the target classification in the form of class labels that have been made.

c. Featured Selection
   Used to select the main features to be used in a machine learning model made.

Based on data stunting Public health center Pit And activity Integrated Healthcare Center Which every month done. data collection Still done with counting manuals with use median reference Which take a long time. So with that developed a classification model using an algorithm random forests for predict do toddlers

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V17.i4 (862-868)

ISSN : 1979-9292

E-ISSN : 2460-5611

experience condition stunting or not. Data stunting taken from results measurement Integrated Healthcare Center in 10 Village in Subdistrict Pit Regency Padang. Data Which used is data Integrated Healthcare Center month January until October 2020. The data that has been collected will be further processed stage of data preprocessing before the data is available used in the learning process using the random forest algorithm . Collected raw data as many as 22855 lines with still empty data or missing value. Needs cleaning data and labeling. It can be seen that the amount of data is reduced. But the data is still not well used in problem classification because data Still Not yet balanced This Can seen from amount each class labels that available. At this stage the data that has gone through the preprocessing process is then "fit" with using the random forest algorithm where the data will be separated into 70:30 where 70% of the data is used as a train set and 30% of the data is used as a test set. And the last step model will save and will applied into the application Which will be made. The evaluation carried out on this project uses two events, namely Train-test split confusion matix by dividing the data three times, namely 50:50, 70:30, and 90:10 and k- fold cross validation with distribution 3-fold, 5-folds, and 10-fold.

```
classification:
              precision    recall  f1-score   support

           0       0.98      0.97      0.98      1945
           1       0.97      0.98      0.98      1956

    accuracy                           0.98      3901
   macro avg       0.98      0.98      0.98      3901
weighted avg       0.98      0.98      0.98      3901

confusion:
[[1890   55]
 [  38 1918]]
Accuracy:
97.61599589848757
Precision:
97.21236695387735
Recall:
98.05725971370143
F1-Score:
97.63298549249174
```

Figure 3. Confusion matrix splitting dataset 50:50

```
classification:
              precision    recall  f1-score   support

           0       0.99      0.97      0.98      1176
           1       0.97      0.99      0.98      1165

    accuracy                           0.98      2341
   macro avg       0.98      0.98      0.98      2341
weighted avg       0.98      0.98      0.98      2341

confusion:
[[1145   31]
 [  17 1148]]
Accuracy:
0.9794959419051688
Precision:
0.9737065309584394
Recall:
0.9854077253218884
F1-Score:
0.9795221843003413
```

Figure 4. Confusion matrix splitting dataset 70:30

```
classification:
              precision    recall  f1-score   support

           0       0.98      0.98      0.98       381
           1       0.98      0.98      0.98       400

    accuracy                           0.98       781
   macro avg       0.98      0.98      0.98       781
weighted avg       0.98      0.98      0.98       781

confusion:
[[374   7]
 [  9 391]]
Accuracy:
97.95134443021767
Precision:
98.24120603015075
Recall:
97.75
F1-Score:
97.9949874686717
```

Figure 5. Confusion matrix splitting dataset 90:10

Evaluate model with 3- fold cross validation. R average mark testing. 3- fold cross validation = **97.70576405524501**. Evaluate model with 5- fold cross validation. R average mark testing 5- fold cross validation = **97.88529706466926**. Evaluate model with 10- fold cross validation. Evaluate model with 10- fold cross validation. 10- fold cross validation = **97.87236941462292**.

At this stage the model has been made and has gone through the testing process and the results obtained considered worthy so model will used in application with the help of flasks framework. Where flasks

ISSN : 1979-9292
E-ISSN : 2460-5611

used as a bridge for applications to access the model that has been made so that the website nor application android can use which models already made.

CONCLUSION

Random Forest algorithm has an average testing accuracy using the confusion matrix of **97.83%**. and on evaluation testing using k-fold cross validation the average accuracy obtained as big **97.821%** where the results of each test evaluate the accuracy produced by the random forest algorithm stable in the **97%** range where it can be concluded that the random forest algorithm can be used in predict condition stunting in toddlers.

RECOMMENDATION

Based on research that has been conducted, there is a relationship between screen time (the time children spend using electronic screens such as television, tablet and smartphones) and the sleep patterns of children aged toddler so it can be concluded that excessive screen time can also affect the quality of a child's sleep. Children who spend longer periods of time in front of electronic screens tend to have more disturbed sleep and poor quality sleep. These sleep disorders can have a negative impact on a child's overall development and well-being (Chen, et al, 2019). Although the research results show a relationship between screen time and the sleep patterns of toddler- aged children, it is important to remember that other factors such as healthy sleep habits and a good sleep environment can also influence children's sleep patterns. It is recommended to reduce screen time and create a consistent and supportive sleep routine for toddler- aged children to ensure adequate and quality sleep. After reviewing several literatures about screen time in toddlers, the American Academy of Pediatrics issued a policy in 2016 asking families to avoid providing screen time or any digital media to children aged 18 to 24 months and limit digital media use to only 1 hour a day. The widespread exposure to digital devices identified in some of this literature indicates a lack of awareness of the negative potential of digital media on children's health. Nurses need to provide parent education to limit children's use of screen time so that children can grow and develop optimally.

BIBLIOGRAPHY

[1] Chilyabanyama, O. N., Chilengi, R., Simuyandi, M., Chisenga, C. C., Chirwa, M., Hamusonde, K., ... & Bosomprah, S. (2022). Performance of machine learning classifiers in classifying stunting among under-five children in Zambia. Children, 9(7), 1082.

[2] Rahman, S. J., Ahmed, N. F., Abedin, M. M., Ahammed, B., Ali, M., Rahman, M. J., & Maniruzzaman, M. (2021). Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. Plos one, 16(6), e0253172.

[3] Darnila, E., Maryana, M., Mawardi, K., Sinambela, M., & Pahendra, I. (2022). Supervised models to predict the Stunting in East Aceh. International Journal of Engineering, Science and Information Technology, 2(3), 33-39.

[4] Haris, M. S., Anshori, M., & Khudori, A. N. (2023). PREDICTION OF

ISSN : 1979-9292
E-ISSN : 2460-5611

STUNTING PREVALENCE IN EAST JAVA PROVINCE WITH RANDOM FOREST ALGORITHM. Jurnal Teknik Informatika (Jutif), 4(1), 11-13.

[5] Purbasari, A., Rinawan, F. R., Zulianto, A., Susanti, A. I., & Komara, H. (2021, September). CRISP-DM for data quality improvement to support machine learning of stunting prediction in infants and toddlers. In 2021 8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA) (pp. 1-6). IEEE.

[6] Bitew, F. H., Sparks, C. S., & Nyarko, S. H. (2022). Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. Public health nutrition, 25(2), 269-280.

[7] Khan, J. R., Tomal, J. H., & Raheem, E. (2021). Model and variable selection using machine learning methods with applications to childhood stunting in Bangladesh. Informatics for Health and Social Care, 46(4), 425-442.

[8] Talukder, A., & Ahammed, B. (2020). Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. Nutrition, 78, 110861.

[9] Aryuni, M., Miranda, E., Kumbangsila, M., Zakiyyah, A. Y., Sano, A. V. D., & Bhatti, F. M. (2023, April). Comparison of Nutritional Status Prediction Models of Children Under 5 Years of Age Using Supervised Machine Learning. In Proceeding of the 3rd International Conference on Electronics, Biomedical Engineering, and Health Informatics: ICEBEHI 2022, 5–6 October, Surabaya, Indonesia (pp. 265-277). Singapore: Springer Nature Singapore.

[10] Byna, A. (2020, July). Comparative Analysis of Machine Learning Algorithms for classification about Stunting Genesis. In Proceedings of the First National Seminar Universitas Sari Mulia, NS-UNISM 2019, 23rd November 2019, Banjarmasin, South Kalimantan, Indonesia.

[11] Van, V. T. S., Antonio, V. A., Siguin, C. P., Gordoncillo, N. P., Sescon, J. T., Go, C. C., & Miro, E. P. (2022). Predicting undernutrition among elementary schoolchildren in the Philippines using machine learning algorithms. Nutrition, 96, 111571.

[12] Islam, M. M., Rahman, M. J., Islam, M. M., Roy, D. C., Ahmed, N. F., Hussain, S., ... & Maniruzzaman, M. (2022). Application of machine learning based algorithm for prediction of malnutrition among women in Bangladesh. International Journal of Cognitive Computing in Engineering, 3, 46-57