JURNAL RISET INFORMATIKA **P-ISSN: 2656-1743 |E-ISSN: 2656-1735**
Vol. 5, No. 3. June 2023 DOI: https://doi.org/10.34288/jri.v5i3.228
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Stunting Early Warning Application Using KNN Machine Learning Method

Nani Purwati[1*], Gunawan Budi Sulistyo[2]

Information System[1,2]
Universitas Bina Sarana Informatika [1,2]
nani.npi@bsi.ac.id, [1*)], gunawan.gnw@bsi.ac.id[2]
Jakarta, Indonesia

(*) Corresponding Author

Abstract

Stunting in toddlers is defined as a condition of failure to thrive due to chronic malnutrition in the long term. The problem of stunting in Indonesia is an issue that is still a concern for the Indonesian government. The prevalence of stunting in Indonesia is still relatively high, coupled with the COVID-19 pandemic, which has impacted the economic sector. For this reason, research on stunting is still a critical topic. This study aims to classify toddler stunting using the k-Nearest Neighbor classification algorithm and build a website-based early detection application for toddler stunting cases. The research results using the k-Nearest Neighbor Algorithm trial obtained a relatively high accuracy of 92.45%. Implementing an early detection system for stunting cases has proven to help health workers classify toddlers as stunted or not. This application is also helpful as an archive and facilitates data reporting. The application has eight main menus: the Puskesmas data menu, Posyandu data, toddler data, weighing, weighing results, development menu, and stunting early warning menu, which contains malnourished and stunted toddlers.

Keywords: Application; Classification; k-NN; Stunting.

Abstrak

Stunting pada balita diartikan sebagai kondisi gagal tumbuh akibat kekurangan gizi kronik dalam jangka waktu yang panjang. Masalah stunting di Indonesia merupakan isu yang masih menjadi perhatian pemerintah Indonesia. Prevelensi Stunting di Indonesia masih cukup tinggi, ditambah lagi akibat pandemi covid 19 yang cukup berdampak pada sektor ekonomi. Untuk itu, penelitian mengenai stunting masih menjadi topik yang sangat penting. Penelitian ini bertujuan melakukan klasifikasi stunting balita menggunakan algoritma klasifikasi k-Nearest Neighbor, serta membangun aplikasi deteksi dini kasus stunting balita berbasis website. Hasil dari penelitian menggunakan uji coba Algoritma k-Nearest Neighbor diperoleh akurasi yang cukup tinggi sebesar 92,45%. Implementasi sistem deteksi dini kasus stunting terbukti membantu dan memudahkan petugas kesehatan dalam mengklasifikasikan balita yang termasuk kedalam kategori stunting atau tidak. aplikasi ini juga bermanfaat sebagai arsip dan memudahkan dalam pelaporan data. Pada aplikasi terdapat 8 menu utama yaitu, menu data Puskesmas, data Posyandu, data balita, penimbangan, hasil penimbangan, menu perkembangan, menu stunting early warning yang berisi balita gizi buruk, balita stunting.

Kata kunci: Aplikasi; Klasifikasi; K-nn; Stunting

INTRODUCTION

Reducing stunting is the leading indicator of Indonesia's sustainable development goal as a 2025 global nutrition target. In 2018, before the Covid outbreak hit, the prevalence of stunting in Indonesia was around 37% (Beal et al., 2018). Stunting is a failure to thrive in children due to chronic malnutrition, and children are too short for their age (Wahyuningsih et al., 2022). Stunting can impact children's performance at school (Ponum et al., 2020), which is less than optimal, so this is of great concern to the government. In addition to impacting children's intelligence levels, stunting also has the potential for economic loss for the country (Renyoet et al., 2016). Several indicators of stunting include lower birth weight (LBW), mother's education level, household income, lack of household sanitation, and the greater risk of toddlers becoming stunted (Apriluana & Fikawati,

2018). Efforts that the government has made to overcome the problem of stunting are providing education and knowledge to the public regarding stunting. This education and knowledge are provided to the community through outreach programs and classes for pregnant women and toddlers (Alamsyah et al., 2023).

After the Covid 19 pandemic, the issue of stunting became an interesting research topic to carry out. Technology continues to develop and penetrate several sectors, one of which is the health aspect. Various studies that focus on the problem of stunting are not only explored in the field of health sciences but are starting to synergize with computer science. Research applying Bayes' Theorem to Detect Stunting in Toddlers (Sapriatin & Sianturi, 2021). In his research, he applied the Naive Bayes algorithm, resulting in an accuracy of 99%. This study detected five diseases and 36 symptoms in detecting stunting. It's just that it hasn't reached the application development stage yet. Several previous studies focused on the issue of stunting, including research (Titimeidara & Hadikurniawati, 2021). Using the Naive Bayes algorithm, his research (Titimeidara & Hadikurniawati, 2021)classifies the stunting status. The results of the Naive Bayes algorithm trials produce an average performance accuracy of 88%. At the same time, other studies with the same research focus were conducted by (Arisandi et al., 2022). In this study, trials were carried out using the Naive Bayes Classifier algorithm with the K-Fold Cross Validation test obtained an accuracy of 94.39%. Subsequent research (Lonang & Normawati, 2022) applies the K-Nearest Neighbor-based selection feature Backward Elimination method. The results of his research show that the average accuracy produced by the K-Nearest Neighbor algorithm at k = 5 is 91.90% with nine attributes, and the average accuracy produced by the K-Nearest Neighbor algorithm with the addition of Backward Elimination is 92.20% with eight attributes (Lonang & Normawati, 2022). Other studies use a support vector machine algorithm to integrate toddler anthropometric data. The average accuracy of the SVM method is 85.1% (Widodo & -, 2022). Research on stunting was also carried out by (Wiraguna et al., 2022) in his research using the Support Vector Machine method. The dataset used consists of 22 variables.

This research is different from previous research. The author wants to know how accuracy is generated from a dataset of 331 toddler data consisting of 11 variables using the k-Nearest Neighbor algorithm. Suppose in the previous study, the results of the k-NN method with adding the Backward Elimination method with eight variables obtained a reasonably high accuracy. What if this study used a different dataset and number of variables without the additional Backward Elimination method?

RESEARCH METHODS

The background of this research is that the problem of stunting in toddlers is still a significant problem. Meanwhile, after the Covid-19 pandemic, it had quite an impact on various sectors of life. This research aims to survey the problem of stunting in toddlers. The survey results are in the form of toddler data, which is processed into a dataset that can be useful for research in data mining. The resulting dataset will be subject to data analysis. Before data processing, preprocessing is carried out first, one of which is setting target attributes or labels.

The method or model proposed in this study is the k-Nearest Neighbor (k-NN). At the same time, the proposed research method uses the CRISP-DM method, which aims to obtain new patterns or knowledge from the data used in solving a problem which is also used for making the right decisions. The test results using the K-NN algorithm are used to develop stunting early detection applications. This application was built using the PHP programming language with the CodeIgniter framework. The Research Framework is contained in Figure 1.

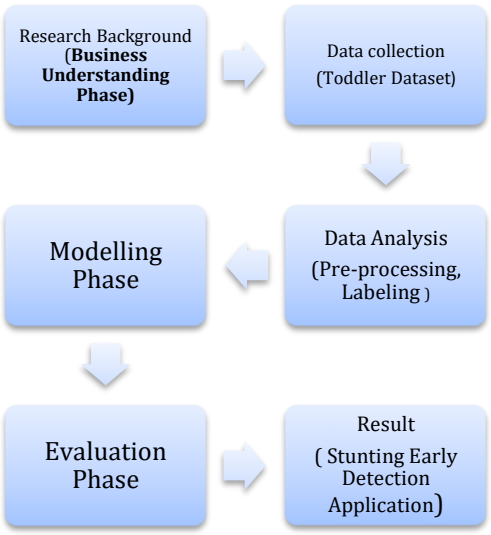

Figure 1. The Research Framework

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v5i3.228
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Data Collection Method

The dataset used in this study was obtained from several Posyandu in rural areas. The results of collecting data on toddlers during the periods of January, February, March, April, and May 2023 obtained 331 toddler data. Variables used in this study included gender, birth order, age, birth weight, birth height, body weight, height, arm circumference, head circumference, exclusive breastfeeding, and nutritional status.

Tabel 1. Toddler Datasets

Data Name	Data Source	Amount of Data	Number of Attributes	Attributes
Toddler datasets	Posyandu	331	11	gender, birth order, age, birth weight, birth height, weight, height, arm circumference, head circumference, exclusive breastfeeding, and nutritional status.

Data Analysis

At this stage, the data preprocessing stage is carried out. Data preprocessing is a stage that starts from the data selection stage, then performs data cleaning, eliminates data duplication, checks inconsistent data, and corrects errors in data (Saputra et al., 2023). After the data preprocessing stage is complete, the labeling stage continues. The label used in this dataset is an attribute of nutritional status. The attribute of nutritional status is an attribute that indicates whether a toddler is at risk of stunting or not.

The method or model proposed in this study is the k-Nearest Neighbor (k-NN) classification method. K-Nearest Neighbor (K-NN) is an instance-based learning group (Azis et al., 2020). This method aims to classify new objects using attributes and training samples (Waliyansyah & Fitriyah, 2019). The KNN method uses neighborhood classification in determining predictions (M. Lestari, 2014). Several previous studies use the k-NN method, among others, by (Z. D. Lestari et al., 2019). His research proved that k-NN has an accuracy of 82% in determining the type of banana (Z. D. Lestari et al., 2019). The final result of this research is an application for early detection of toddler stunting. The development of this application uses the CodeIgniter framework. CodeIgniter is an open-source application for building dynamic PHP applications (Muqorobin & Rozaq Rais, 2022). CodeIgniter can speed up developers for developing PHP-based web applications (Anggraini et al., 2020).

RESULTS AND DISCUSSION

K-NN Method Test Results

Testing the performance of the machine learning method was carried out with a sample of toddler data with variables in the form of gender, birth order, age, birth weight, birth height, body weight, height, arm circumference, head circumference, exclusive breastfeeding, and nutritional status. The nutritional status attribute is an attribute that is set to be a label—algorithm testing using Rapidminer 10.1 software. Figure 2 is the design model in this study. Training and testing data are tested using the k-NN algorithm and then validated using the performance operator.

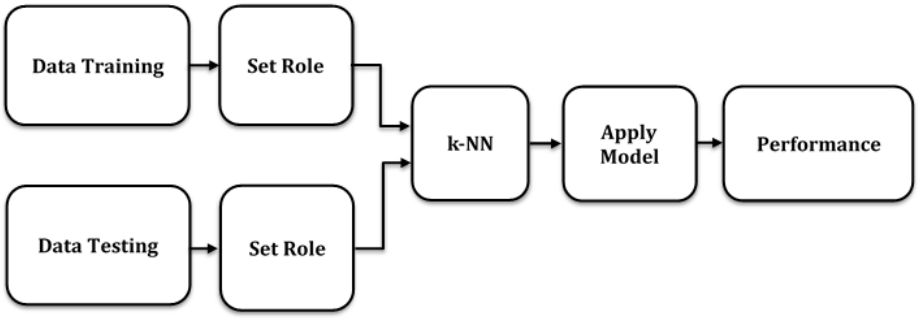

Figure 2. Model Design.

Whereas Figure 3 is the performance result of the design of the k-NN algorithm model after being tested for validation. Table 2 explains that the class precision for the prediction of ordinary is 93.97%, the class precision for the prediction of malnutrition is 84.62%, and the class precision for the prediction of stunting is 82.61%.

Table 2. Results

Prediksi	True Normal	True Malnutrition	True Stunting	Class Precision
Pred. Normal	265	9	8	93.97%
Pred. Malnutrition	4	22	0	84.62%
Pred. Stunting	3	1	19	82.61%
Class Recall	97.43%	68.75%	70.37%	**Accuracy =** 92.45%

In Table 2, it can be seen that the accuracy of using the k-NN model is 92.45%. The accuracy shown in the results of this model test is higher than that of previous studies using the k-NN method with the addition of Backward Elimination is 92.20% (Lonang & Normawati, 2022). This proves that the k-NN method can provide accurate results without needing additional models to optimize. However, k-NN does not produce higher accuracy than the Naive Bayes method, with an accuracy of 94.39% in previous research (Arisandi et al., 2022). Even so, the difference in accuracy is minimal and not too much different, so the k-NN method is still in the excellent category for predicting stunting cases.

System Application Development and Testing

Designing the application in this study is to make an initial design for modeling. After passing through the user interface design stage, programming is developed by implementing designs designed in computer programming languages. The programming language used to build applications is PHP with the CodeIgniter framework.

CONCLUSIONS AND SUGGESTIONS

Conclusion

Based on the results of applying the classification method using the k-Nearest Neighbor algorithm, it can be concluded that this algorithm has a high level of accuracy for the classification of toddler stunting. The resulting accuracy is 92.45%. This result is higher even though it does not use additional methods as was done in previous studies using the additional method of Backward Elimination. It can be concluded that the k-NN algorithm can provide high accuracy in predicting toddler stunting even without additional methods or optimization. Using 11 variables in the k-NN algorithm is also proven to increase accuracy even without adding the Backward Elimination method or other optimization methods. Furthermore, the results of implementing programs built using the CodeIgniter framework produce applications that are easy to use with a user-friendly interface. Implementing the stunting early warning system helps make it easier for health workers to archive data and detect toddlers with the potential to be stunted.

Suggestion

For further research, it can involve the same or different datasets. Future research can also use other classification methods and compare classification algorithms, k-NN, SVM, decision trees, Naive Bayes, etc. Of these several methods, additional models such as PSO optimization can also be used. Then the model test results can be selected as the best model using a different t-test.

ACKNOWLEDGMENTS

The author would like to thank the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia for funding the 2023 PDP scheme grant. In addition, the author also thanks Bina Sarana Informatika University for its support.

REFERENCES

Alamsyah, P. J., Khair, U., Bengkulu, U. M., Tadulako, U., Suroso, A., Bella, K. S., Takallar, G. K., Desa, B., Pemuda, T., Adat, T., & Agama, T. (2023). Upaya pemerintah dan masyarakat dalam penanganan dan pencegahan stunting. *Jurnal*

Pengabdian Kolaborasi Dan Inovasi IPTEKS, *1*(3), 191–198. https://journal.ppmi.web.id/index.php/JPKI2/article/view/33

Andryan, M., Faisal, M., & Kusumawat, R. (2023). K-Means Binary Search Centroid With Dynamic Cluster for Java Island Health Clustering. *Jurnal Riset Informatika*, *5*(3), 269–276. https://doi.org/10.34288/jri.v5i3.511

Anggraini, Y., Pasha, D., Damayanti, D., & Setiawan, A. (2020). Sistem Informasi Penjualan Sepeda Berbasis Web Menggunakan Framework Codeigniter. *Jurnal Teknologi Dan Sistem Informasi*, *1*(2), 64–70. https://doi.org/10.33365/jtsi.v1i2.236

Apriluana, G., & Fikawati, S. (2018). Analisis Faktor-Faktor Risiko terhadap Kejadian Stunting pada Balita (0-59 Bulan) di Negara Berkembang dan Asia Tenggara. *Media Penelitian Dan Pengembangan Kesehatan*, *28*(4), 247–256. https://doi.org/10.22435/mpk.v28i4.472

Arisandi, R. R. R., Warsito, B., & Hakim, A. R. (2022). Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation. *Jurnal Gaussian*, *11*(1), 130–139. https://doi.org/10.14710/j.gauss.v11i1.33991

Azis, H., Purnawansyah, P., Fattah, F., & Putri, I. P. (2020). Performa Klasifikasi K-NN dan Cross Validation Pada Data Pasien Pengidap Penyakit Jantung. *ILKOM Jurnal Ilmiah*, *12*(2), 81–86. https://doi.org/10.33096/ilkom.v12i2.507.81-86

Beal, T., Tumilowicz, A., Sutrisna, A., Izwardy, D., & Neufeld, L. M. (2018). A review of child stunting determinants in Indonesia. *Maternal and Child Nutrition*, *14*(4), 1–10. https://doi.org/10.1111/mcn.12617

Lestari, M. (2014). Penerapan Algoritma Klasifikasi Nearest Neighbor (K-NN) untuk Mendeteksi Penyakit Jantung. *Faktor Exacta*, *7*(September 2010), 366–371. http://journal.lppmunindra.ac.id/index.php/Faktor_Exacta/article/view/290

Lestari, Z. D., Nafi'iyah, N., & Susilo, P. H. (2019). Sistem Klasifikasi Jenis Pisang Berdasarkan Ciri Warna HSV Menggunakan Metode K-NN. *Seminar Nasional Teknologi Informasi Dan Komunikasi*, 11–15. http://prosiding.unipma.ac.id/index.php/SENATIK/article/view/880

Lonang, S., & Normawati, D. (2022). Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination. *Jurnal Media Informatika Budidarma*, *6*(1), 49. https://doi.org/10.30865/mib.v6i1.3312

Muqorobin, M., & Rozaq Rais, N. A. (2022). Comparison of PHP Programming Language with Codeigniter Framework in Project CRUD. *International Journal of Computer and Information System (IJCIS)*, *3*(3), 94–98. https://doi.org/10.29040/ijcis.v3i3.77

Ponum, M., Khan, S., Hasan, O., Mahmood, M. T., Abbas, A., Iftikhar, M., & Arshad, R. (2020). Stunting diagnostic and awareness: Impact assessment study of sociodemographic factors of stunting among school-going children of Pakistan. *BMC Pediatrics*, *20*(1), 1–9. https://doi.org/10.1186/s12887-020-02139-0

Renyoet, B. S., Martianto, D., & Sukandar, D. (2016). Potensi Kerugian Ekonomi Karena Stunting Pada Balita Di Indonesia Tahun 2013. *Jurnal Gizi Dan Pangan*, *11*(3), 247–254. https://journal.ipb.ac.id/index.php/jgizipangan/article/view/16461

Sapriatin, B., & Sianturi, A. F. (2021). Penerapan Teorema Bayes Mendeteksi Stunting pada Balita. *Jurnal Media Informatika [Jumin]*, *3*(1), 24–37. http://ejournal.sisfokomtek.org/index.php/jumin/article/view/203

Titimeidara, M. Y., & Hadikurniawati, W. (2021). Implementasi Metode Naïve Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita. *Jurnal Ilmiah Informatika*, *9*(01), 54–59. https://doi.org/10.33884/jif.v9i01.3741

Wahyuningsih, W., Bukhari, A., Juliaty, A., Erika, K. A., Pamungkas, R. A., Siokal, B., Saharuddin, S., & Amir, S. (2022). Stunting Prevention and Control Program to Reduce the Prevalence of Stunting: Systematic Review Study. *Open Access Macedonian Journal of Medical Sciences*, *10*(F), 190–200. https://doi.org/10.3889/oamjms.2022.8562

Waliyansyah, R. R., & Fitriyah, C. (2019). Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *5*(2), 157. https://doi.org/10.26418/jp.v5i2.32473

Widodo, S. W., & -, A. S. (2022). Development of Intelligent Software for Early Detection of Stunting in Toddlers Based on Anthropometry. *Proceeding of International Conference on Science, Health, And Technology*, 63–69. https://doi.org/10.47701/icohetech.v3i1.22

79
Wiraguna, I. K. A., Setyati, E., & Pramana, E. (2022). Prediksi Anak Stunting Berdasarkan Kondisi Orang Tua Dengan Metode Support Vector Machine Dengan Study Kasus Di Kabupaten Tabanan-Bali. *Smatika Jurnal*, *12*(01), 47–54. https://doi.org/10.32664/smatika.v12i01.662