Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

RESEARCH PAPER

OPEN ACCES Check for updates

Early Detection of Stunting Based on Feature Engineering and Machine Learning Algorithm Approach

Ayu Ahadi Ningrum[1], Yunia Ikawati[2]

[1] Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia
[2] Department of Informatics, Politeknik Elektronika Negeri Surabaya, Surabaya, Indonesia

ABSTRACT

The stunting problem in Indonesia is still an extensive issue for the government. Around 22% of cases of stunting affect brain development, resulting in reduced intellectual capacity and permanent disruption of the structure and function of nerves and brain cells. This research early detection of stunting using an feature selection approach. So, datasets related to stunting are valuable in providing complete insight or information in detecting early symptoms of stunting in toddlers. The data collection starts from 2021 to 2024. The data consists of fourteen features, one target label and ten thousand dataset rows. This study use three kind of Machine learning models and feature engineering method for selecting the best features. From fourteen features, only breastfeeding, parental education, and birth weight and length are critical factors in ensuring child nutrition. We continue *to* test the dataset using Machine Learning algorithms with various variants, such as Random Forest, *Support Vector Machine* and Multilayer Perceptron. *T*he Multilayer Perceptron algorithm can produce an accuracy value of 98%. Based on the approach using feature engineering, several important features were obtained which were proven to influence stunting. Having these very influential features can help in initial prevention steps to prevent stunting in children.

PAPER HISTORY

Received April 02, 2024
Revised May 31, 2024
Accepted May 31, 2024

KEYWORDS

Feature Engineering;
Feature Selection;
Machine learning;
Stunting;

1. INTRODUCTION

The problem of stunting in Indonesia is still an extensive issue for the Indonesian government [1]. Centre for Health Data and Information of the Republic of Indonesia Based on data from SSGI in 2022, there is a 22% prevalence of stunting in Indonesia. Worldwide in 2022, approximately 149 million children younger than five years old were classified as stunted, meaning they were shorter than expected for their age. Additionally, an estimated 45 million children in this age group were categorized as wasted, indicating they were excessively thin for their height. Especially in South Kalimantan, there were 2490 cases of malnutrition. In the short term, stunting causes growth failure, hampered motor and cognitive growth and suboptimal growth [2]. In the long term, stunting affects brain development, permanently reducing performance, nerve cells, and function. This causes a decrease in the ability to absorb lessons and can even further affect the level of productivity in adulthood [3].

One of the factors underlying stunting is inadequate calorie intake [4]. When dissected more deeply, stunting can be caused by poverty, limited knowledge regarding the adequacy of breast milk for babies and toddlers, the role of providing animal protein in MPASI, congenital heart disease, metabolic disorders, and chronic infections. If it is not addressed seriously, it will gradually affect the quality of human resources in Indonesia's young generation [1]. Especially for toddler, proper nutrition is essential for maintaining good health and fostering the best possible physical development [5]. Adequate calorie intake is essential to prevent malnutrition, which can have serious consequences for health, including an increased risk of heart disease, stroke, diabetes, and other health issues[6]. This goal relies heavily on proper nutrition and malnutrition prevention. Low to moderate malnutrition is observed to have an early impact on the growth and development of most children. Diagnosing malnutrition at an early age reduces healthcare costs and helps to maintain good health [7].

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

The Indonesian government is committed to addressing the issue of stunting. The National Medium-Term Development Plan (RPJMN) 2020-2024, officially enacted through Government Regulation No. 18 of 2020, explicitly sets a target of reducing the prevalence of stunting in Indonesia to 24% by 2024 [8]. In general, data is a limited form of description of reality. Each piece of data provides pieces of information. However, measurement noise is always present, and data needs to be added [5]. Feature Engineering is formulating features that best suit the data, model and task. Selecting the right features is crucial to ensure the effectiveness of model building in Machine Learning. The choice of features can have a significant impact on the accuracy and efficiency of the model [6]. Feature Engineering involves transforming raw data into valuable features to form a case detection model. It is tasked with determining the best features to represent data samples to train a solution to a problem.

In this study, Machine Learning is utilized to predict whether a child is stunted or not. The machine learning variants applied include Random Forest, Support Vector Machine, and Multilayer Perceptron. Previous research related to stunting using computer algorithms was related to the classification of stunted children using the C4.5 algorithm [5] ; stunting that occurs when toddlers is a risk factor for non-communicable diseases in the future [6]; detection of diagnoses of overweight and obesity using Machine Learning [7]; the performance of Machine Learning in classifying stunting cases in children under five years in Zambia [14]; children's measurement using machine learning [15]; Machine Learning approach for predicting the Impact of food Insecurity [16].

Several previous studies utilized various variants of machine learning algorithms to perform out machine learning processes, but no one has examined in depth at the features in them that have an important effect on prediction results. The study seeks to the most influential factors contributing to stunting, enabling timely interventions to prevent the condition and enhance the quality of the future workforce. Following are the contributions of this research that outline the most important features that may affect stunting. It hopes that by recognizing the primary causes of stunting, early detection and intervention can be implemented. In the long term, this may reduce or possibly reduce the prevalence of stunting. This research continues to focus on studying features from previous records rather than in real time. Future research should use time series data to improve toddler health monitoring. The goal is to begin the intervention process as soon as possible after identifying the child's nutritional status.

2. MATERIALS AND METHOD

A. Stunting

Stunting is a condition where a person's growth failure occurs when compared with others based on WHO reference [17]. During the first 1000 days of life, a persistent malnutrition causes stunting [4]. Babies in the LBW (low birth weight) category have a body weight of less than 2500 gr (2.5 kg) or under 1.5 kg. If the height of a kid is less than two standard deviations of the WHO, they are considered short. Table 1 shows stunting indicators based on WHO standards.

Tabel 1. Stunting Indicators

No	Category	Z-score
1	Very Short	< -3 SD
2	Short	-3 SD s.d < -2 SD
3	Normal	-2 SD s.d +3 SD
4	Tall	> +3 SD

In the long term, stunting can reduce a child's cognitive function, hamper the balance of body functions, and prevent body posture from growing optimally as an adult [15]. The factors mentioned reduce the quality of human resources in terms of productivity, quality, and competitiveness with other countries [3].

B. Machine Learning

Machine learning is a combination of mathematical algorithms and computer applications. Sources of knowledge resulting from learning based on data and the results can predict the future [10]. Machine Learning algorithms are designed to process large amounts of data and identify valuable patterns for making predictions or decisions [11]. The learning process comes from training and testing. Features in Machine Learning are part of the model for making predictions or decisions. Machine Learning requires data as a learning source for optimal output results. Various kinds of machine learning algorithms can be used to create models. The model in this study was evaluated using Random Forest, SVM and Multilayer Perceptron.

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

We use these three types of Machine Learning algorithms because they can handle both numerical and categorical data. These three algorithms are also well-suited for dealing with large datasets and have a variety of features. Random Forest offers flexibility in handling classification and regression cases [12]. This algorithm can handle large and complex datasets and prevent overfitting. In its application, this method can improve prediction results because it can generate child nodes at each node randomly [13]. This concept is also usually used in building decision tree methods. In the initial stage of this algorithm, we have to determine the entropy value using Equation 1 and the gain value using Equation 2.

$$Entropy\ (Y) = \sum_{i} p(c|Y) log_2 .p(c|Y) \quad (1)$$

where Y is a set of cases, and p(c|Y) is the insertion of the Y value into class c.

$$Entropy(Y) = -\sum_{vt\ Value\ (a)} \frac{Y_v}{Y_\propto} Entopy\ (Y_v) \quad (2)$$

where value (a) indicates all possible values in the case set a, and Y_v shows subclass of Y with class v connected to class a.

Support Vector Machine is an algorithm compatible with solving classification and prediction cases by finding the most optimal separator in different classes. Equation 3 shows the formulation of SVM.

$$f(x_d) = \sum_{n=1}^{ns} \alpha_i y_i \underset{x\ i}{\rightarrow} \underset{x\ d}{\rightarrow} + b \quad (3)$$

where ns shows number of support vector, α_i shows the weight value for each data point, y_i shows data class, $\underset{x\ i}{\rightarrow}$ shows variable support vector, $\underset{x\ d}{\rightarrow}$ shows data to be classified, and b shows bias.

Multilayer Perceptron is a Neural Network model with many hidden layers between the input and output layers. Multilayer Perceptron has the advantage of detecting very complex problems. Equation 4 shows the linearity of the MLP.

$$f(x) = \sqrt{2x^2 + 1} \quad (4)$$

A. Feature Engineering

Selecting the right features is critical to ensuring the effectiveness of a machine learning model [14]. The accuracy and effectiveness of the model may be significantly impacted by this procedure. By removing unnecessary features, feature selection techniques lower the final model's complexity. The ultimate objective is to create a model that can be computed more quickly while maintaining a level of prediction accuracy [15]. In implementing feature engineering, several feature selection techniques require increasing the overall data training time and reducing the model assessment time.

In this process, the data is entered into the imputation process. This method functions to handle missing values. The primary function of imputation is to improve the quality of incomplete or damaged data for further analysis. There are several techniques used to fill in missing values. There are three techniques for handling missing data in a dataset: filling in with the average value, filling in with the most significant data value, and deleting records with empty data. All the data has been filled in, and then the data transformation process is carried out again. This step aims to uniformly distribute the range of values for each column. The StandardScaller technique was used in this study. The scaling process is carried out before the model training process. This aims to improve the model's performance and increase the resulting accuracy values. Equation 5 shows the data transformation formula.

$$x' = \frac{x - \min(x)1}{max(x) - \min(x)} \quad (5)$$

The subset data is then passed into the evaluation process using three feature selection techniques: filtering, wrapping, and embedding. The filtering technique works by removing features that are not useful for the model, wrapping treats the model as a black box that provides a quality score from a subset of proposed features, and the embedded technique performs feature selection as part of the model training process [24]. Embedded methods provide a balance between computational cost and result quality. After finding the desired feature criteria, the feature selection process stops and produces a score value for each feature. Good features facilitate additionally modeling stages and increase the resulting model's ability to accomplish the intended task. For an insufficient feature to perform at the same level, a much more complicated model might

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

be needed. Figure 1 illustrates the Feature Engineering process in the Machine Learning cycle.

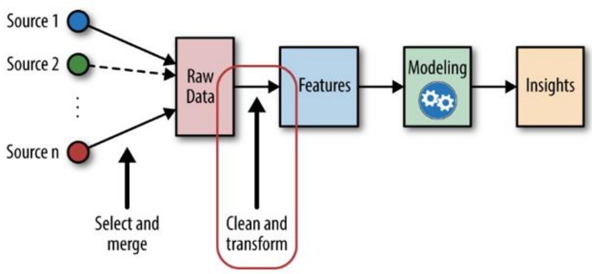

Figure 1. Machine Learning Cycle

B. Design System

The research's design is shown in Figure 2. It is capable of being broadly separated into two processes: the feature engineering process comes after feature formation. where the following machine learning variant models Random Forest, Support Vector Machine, and Multilayer Perceptron are employed during the testing phase.

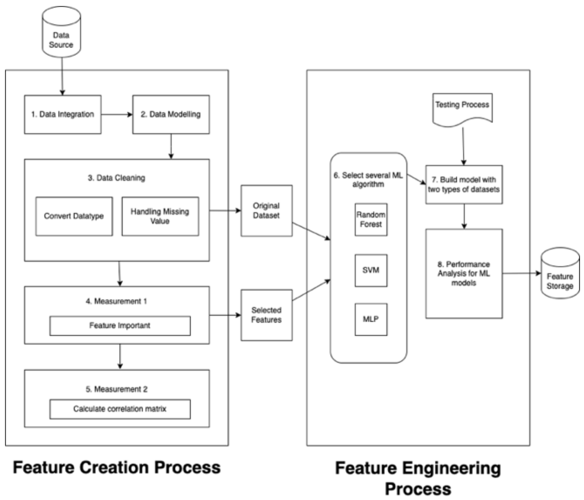

Figure 2. Design System

C. Dataset

The data source in this research comes from measurements in Puskesmas the Kelayan District area, South Banjarmasin. The data collection starts from 2021 to 2024. The data consists of fourteen features, one target label and ten thousand dataset rows. Table 2 shows an explanation of the features used in the research. This is a semi-observational study based on cases that occur in the service area. The subjects of observation were toddler patients aged 0 months to 3 years, with varying educational and socioeconomic backgrounds.

Tabel 2. Dataset Feature

No	Feature	Description
1	Jk	Genre
2	P_OT	Education level of the parent
3	PK_OT	Parent occupation
4	Age	Age
5	TTD	Consum blood tablet status
6	PDN	Birth in hospital status
7	Birth Weight	Birth Weight
8	Birth Lenght	Birth Height
9	Body Weight	Current Weight
10	Body Height	Current Height
11	LKN	Head circumference
12	IML	Vaccine status
13	BF	Breast Feeding status
14	Status	Label

D. Metric Evaluation

Several parameters are used to evaluate the matrix results in this research, namely, Pearson's correlation, which shows the degree of relationship between two variables. It carries out natural measurements and measures the strength of the relationship between two quantitative variables, symbolized by r.

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{(\sum x^2 - \frac{(\sum x)^2}{n}).(\sum y^2 - \frac{(\sum y)^2}{n}).}} \quad (6)$$

The relationship value ranges between (-1) and (+1). The r value expresses the strength of the association, as illustrated in Figure 3.

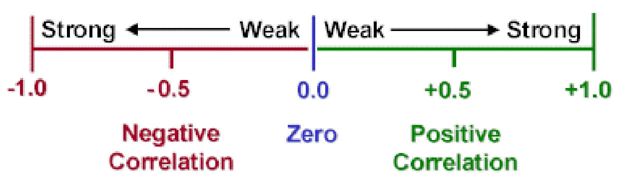

Figure 3. Relation Values

Based on figure 5 it show that If r = 0 it means there is no correlation between two variable. If 0 < r

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

< 0.5 = weak correlation. If 0.5 ≤ r < 1 = strong correlation. If r = 1 = perfect true correlation.

This research also uses several parameters to analyze algorithm performance. The indicators used are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). From the indicators mentioned, an equation can be formed that states their accuracy, as in equation 7. The accuracy value shows the comparison between the number of children indicated to be stunted or not from the total data.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

Recall shows the actual value positively. Precision indicates the estimated value of the positive probability of being correct. F1 Score shows the weighted average comparison value of precision and recall.

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (8)$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (9)$$

$$\mathrm{F1\ Score} = 2\mathrm{x}\,\frac{\mathrm{Precission\ x\ Recall}}{\mathrm{Precision} + \mathrm{Recall}} \quad (10)$$

Where True Positive (TP) means the model accurately identified an instance as belonging to the positive class, True Negative (TN) means the model accurately identified an instance as belonging to the negative class, False Positive (FP) occurs when the model incorrectly identifies an instance as belonging to the positive class, and False Negative (FN) occurs when the model incorrectly identifies an instance as belonging to the negative class.

C. RESULTS

Table 3 shows the feature selection results using the feature importance method. Based on the initial 13 features, seven features show a value of more than 0.080 as essential features. We chose only seven main features because they are the most important in producing a good score. High-scoring features increase the prediction model's accuracy. This is because these features contain the most useful information for predicting the target value. We can reduce the risk of overfitting by selecting only the most relevant features. Overfitting occurs when a model is overly complex and learns noise from the data, resulting in poor performance on new data.

Tabel 3. Feature Selection Result

No	Feature	Score Value
1	Gender	0.007
2	P_OT	0.214
3	PK_OT	0.028
4	Age	0.046
5	TTD	0.008
6	PDN	0.002
7	Birth Weight	0.082
8	Birth Height	0.078
9	Body Weight	0.080
10	Body Length	0.073
11	LKN	0.115
12	IML	0.007
13	BF	0.399

Seven features have dominant value in detecting stunting cases: BF, P_OT, Birth BB, Birth TB, LKN, BW, and BL. Figure 4 shows the visualization results of the feature selection process. Figure 7 shows that the BF or breastfeeding feature, parental occupation, birth length, birth weight are very dominant features in detecting stunting.

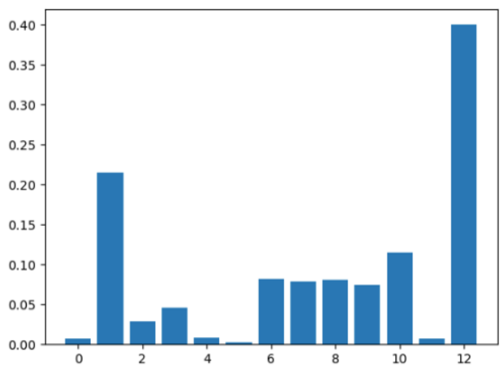

Figure 4. Feature Important

Still based on Figure 5, it shows that the features of immunization conditions, help from

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

health workers, signs, gender, parents' occupation do not appear to have much influence on a child's stunting status.

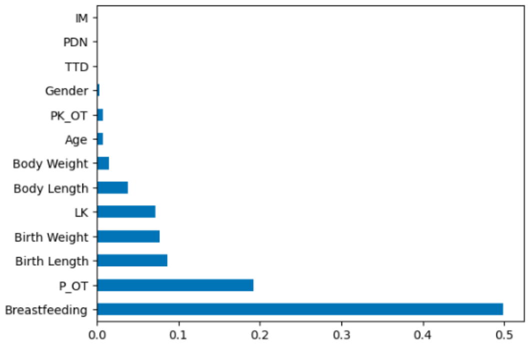

Figure 5. Feature Important Rank

Furthermore, Figure 8 depicts the correlation matrix between the seven selected main features and nutritional status in children. The correlation value ranges from – 1.0 to 1.0.

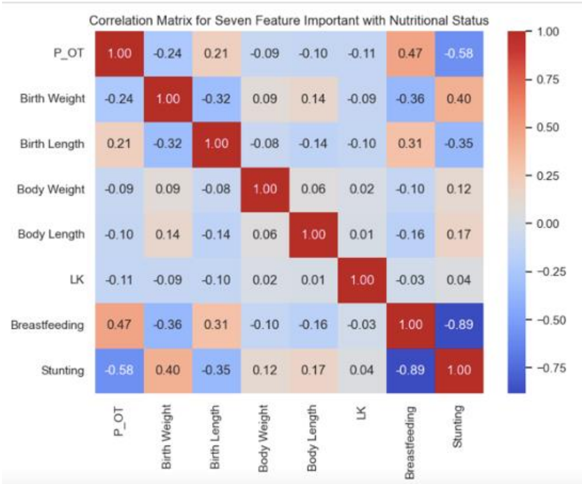

Figure 6. Correlation Matrix

The correlation between the breastfeeding status feature and stunting status has a very strong negative correlation value, as can be seen by looking at the correlation matrix in Figure 6. This may be proof that a child's risk of stunting decreases with the amount of breast milk the mother gives them [21]. According to data collected by the World Health Organization (WHO), exclusive breastfeeding is an effective way to reduce the prevalence of stunting in Indonesia [16]. However, unfortunately, many mothers do not provide exclusive breastfeeding [17]. This is because many mothers do not know about the importance of breastfeeding during the first two years of a child's life [18].

In the next position, the feature of parental education level can also be a strong feature in influencing stunting status in children. The parental education level feature has a neutral negative correlation value in influencing children's nutritional status. Furthermore, the birth weight feature has a normal positive correlation value with the nutritional status of children.

E. Performance Analysis

The feature engineering process stages have been completed, then we continue with the model evaluation process using various algorithms. This research uses 3 types of Machine Learning algorithms to test the performance of the resulting model. Table 4 shows the results of model evaluation using complete features from the dataset. Where the Random Forest algorithm can produce an accuracy of 78% in predicting early stunting status in children. The Support Vector Machine algorithm can only produce an accuracy value of only 58%, finally the Multilayer Perceptron algorithm can produce an accuracy value of 88% in predicting early stunting status in children.

Testing was then continued again by implementing the results of the feature engineering process which can be seen in table 5.

Tabel 4. Before Feature Engineering Process

No	Algoritma	Precision	Recall	F1-score	Accuracy
1	RF	0.82	0.83	0.84	0.78
2	SVM	0.61	0.64	0.65	0.58
3	MLP	0.85	0.87	0.90	0.88

After implementing the feature engineering process, the Random Forest algorithm can produce an accuracy value of 85%. The Support Vector Machine algorithm produces an accuracy value of 84% and the Multi Layer Perceptron can

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

produce 98% in predicting the value of early stunting in children.

Based on the Tables 3 to 5, feature engineering can improve the performance of machine learning models by creating relevant and informative features from raw data. With engineered tools, ML models can make more accurate predictions, handle complex and distributed data, reduce overfitting, and extract valuable insights from categorical and numerical data. For further results in Table 4, the training results using a Machine Learning algorithm are displayed before the Feature Engineering process. Table 5 shows the test results after going through the feature engineering process. It can be proven that applying feature selection can optimize model formation and increase accuracy in predicting the results of Machine Learning algorithm work.

Tabel 5. After Feature Engineering Process

No	Algoritma	Precision	Recall	F1-score	Accuracy
1	RF	0.87	0.88	0.89	0.85
2	SVM	0.81	0.86	0.89	0.84
3	MLP	0.94	0.96	0.95	0.98

Based on Table 4, the Multilayer Perceptron algorithm produced the highest accuracy value before the Feature Engineering was carried out. However, after the Feature Engineering process, the model formation process was carried out again using various variants of the Machine Learning Algorithm, which can be seen in Table 5, which can further improve the accuracy results. According to the results of this study, the multilayer perceptron algorithm variant can perform exceedingly well in modeling complex non-linear relationships between features and targets. MLP, on the other hand, has several drawbacks, including the risk of overfitting if not properly tuned and the need for a longer training time than other algorithms. This study found that the Multilayer Perceptron is better suited to the dataset used.

This research is still primarily concerned with studying features from previous records rather than in real time. Future research should use time series data to make it easier to monitor the health of toddlers. The goal is that by identifying the child's nutritional status as soon as possible, the intervention process can begin.

D. DISCUSSION

The findings of this study lend support to the theory about the most powerful factors that cause stunting. According to the top three rankings of the feature selection in this study, breastfeeding, parental education, and birth weight and length are critical factors in ensuring child nutrition. Looking deeper, the higher the parents' education, the more aware they are of the importance of providing good nutrition for their children by providing vitamins and nutrition to pregnant women throughout the breastfeeding process to prevent stunting. The algorithm best suited to this case and dataset is machine learning with a multilayer perceptron variant. This can result in a prediction accuracy of 98%. According to similar research on stunting detection using machine learning algorithms, Ula can predict nutritional status in children using multiple linear regression with an 80% accuracy [23], in their study on stunting categorization, Banurea et al. also discovered a similar effect, demonstrating 89% accuracy using the SVM and Random Forest methods [24]. This research continues to focus on studying features from previous records rather than in real time. Future research should use time series data to improve toddler health monitoring. The goal is to begin the intervention process as soon as possible after identifying the child's nutritional status.

E. CONCLUSION

Based on the research above, we conclude that feature engineering can enhance the performance of machine learning models by generating insightful and pertinent features from unprocessed data. With feature engineering, ML models can make more accurate predictions, handle complex and distributed data, reduce overfitting, and extract valuable insights from categorical and numerical data.

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

This study uses machine learning modelling to detect stunting early. Of the 14 features that predict the value of detecting stunting, only 3 are the most influential based on their correlation value. The top of important feature is breastfeeding, it has a very strong negative correlation value with stunting status. This could be evidence that the more children are given breast milk by their mothers, the lower their chances of getting stunting. Unfortunately, many mothers do not provide exclusive breastfeeding. This is because many mothers do not know about the importance of breastfeeding during the first two years of a child's life. When testing continues using Machine Learning algorithms with various Multilayer Perceptron variants, it produces an accuracy value of 98%. Future research plans can be continued with the application of Deep Learning algorithms in the process of observing children's development status to prevent stunting using a time series approach.

REFERENCES

[1] K. Lingga , V. Aulia and M. Hadhikul, "Risks of Stunting and Interventions to prevent Stunting," *Journal of Community Engagement in Health,* pp. 241-245, 2023.

[2] D. Permana, S. Anantanyu and A. Nanda Priyatama, "Stunting Incidence in Toddlers Aged 24-59 Months in Kuburaya District Viewed From Feeding Patterns," *Global Health Science Group,* pp. 259-266, 2023.

[3] M. S. Haris, M. Anshori and A. N. Khudori, "Prediction Of Stunting Prevalence In EastAST Java Province With Random Forest Algorithm," *Jurnal Teknik Informatika (JUTIF),* pp. 11-13, 2023.

[4] M. E. Sumiwi, Intervensi Spesifik untuk Percepatan Penurunan Stunting, Jakarta: Direktur Jenderal Kesehatan Masyarakat, 2023.

[5] S. Ankalaki, V. G. Biradar, K. Kumar Naik P and G. S. Hukkeri, "A Deep Learning Approach for Malnutrition Detection," *International Journal of Online and Biomedical Engineering,* vol. 20, no. 6, 2024.

[6] M. M. Shahriar, M. S. Iqubal and S. Mitra, "A deep learning approach to pre- dict malnutrition status of 0–59 month's older children in Bangladesh," in *rtificial Intelligence, and Communications Technology (IAICT)*, Bali, 2019.

[7] Pradhananga and Priza, "A Trend Analysis of Protein-Energy Malnutrition and High Body Mass Index Using The Data From Global Burden of Disease 2010–2019," *Plos One,* vol. 17, no. 9, 2022.

[8] B. Soetono and A. Siti Barokah, "Trends in Stunting Prevelance Reduction: An Examination Of Data Toward Achieving The 2024 Target in Indonesia," *The Social Perspective Journal,* vol. 3, no. 1, pp. 51-68, 2024.

[9] P. Duboue, The Art of Feature Engineering Essentials for Machine Learning, United Kingdom: Cambridge University Press, 2020.

[10] T. Rawat and V. Khemchandani, "Feature Engineering (FE) Tools and Techniques for Better Classification Performance," in *International Journal of Innovations in Engineering and Technology (IJIET)*, India, 2019.

[11] M. Yunus, M. K. Biddinika and A. Fadlil, "Classification of Stunting in Children Using the C4.5 Algorithm," *Join (Jurnal Online Informatika),* pp. 99-106, 2023.

[12] H. Kurniati, R. Djuwita and M. Istiqfani, "Stunting Saat Balita sebagai Salah Satu Faktor Risiko Penyakit Tidak Menular di Masa Depan," *Jurnal Epidemiologi Kesehatan Indonesia,* pp. 59 - 68, 2022.

[13] M. Calderón-Díaz and L. J. Serey-Castillo, "Detection of variables for the diagnosis of overweight and obesity in young Chileans using machine learning techniques," *International Workshop on Human-Centric Innovation and Computational Intelligence,* 2023.

[14] O. Nchimunya Chilyabanyama, R. Chilengi and M. Simuyandi, "Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia," *Children,* pp. 2-11, 2022.

[15] S. R. Ilham and Z. F Asikin, "Stunting Classification in Children's Measurement Data Using Machine Learning Models," *Journal La Multiapp,* 2022.

[16] R. Qasrawi and S. Sgahir, "Machine Learning Approach for Predicting the Impact of Food Insecurity on Nutrient Consumption and Malnutrition in Children Aged 6 Months to 5 Years," *Children,* vol. 11, p. 810, 2024.

[17] M. Yanto and F. Hadi, "Determination of children's nutritional status with machine learning classification analysis approach," *Indonesian Journal of Electrical Engineering and Computer Science,* 2024.

[18] V. Yanry Lameky, "Stunting in Indonesia: Current progress and future directions," *Journal of Healthcare Administration,* 2024.

[19] M. Nur Faiz, O. Somantri and A. W. Muhammad, "Machine Learning-Based Feature Engineering to Detect DDoS Attacks," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi,* vol. 11, no. 3, pp. 176-182, 2022.

[20] P. Hasan Putra, A. Azanuddin, B. Purba and Y. Agustina Dalimunthe, "Random forest and decision tree algorithms for car price prediction," *Jurnal Matematika Dan Ilmu Pengetahuan Alam LLDikti Wilayah 1 ,* 2023.

[21] R. Supriyadi, W. Gata, N. Maulidah and A. Fauzi, "Penerapan Algoritma Random Forest Untuk Menentukan

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia

Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
August 2024, Vol. 6, No. 3, pp. 147-155, e-ISSN: 2656-8624, https://ijeeemi.ijahst.org
https://doi.org/10.35882/ijeeemi.v6i3.6

Kualitas Anggur Merah," *JURNAL ILMIAH EKONOMI DAN BISNIS,* pp. 67-75, 2020.

[22] C. Ikeda, K. Ouazzane and Q. Yu, "A New Framework of Feature Engineering for Machine Learning in Financial Fraud Detection," in *Computer Science & Information Technology (CS & IT)*, London, 2020.

[23] T. Verdonck, B. Baesens and M. Óskarsdóttir, "Special issue on feature engineering editorial," in *Springer Nature*, Belgia, 2021.

[24] R. Gustriansyah and N. Suhandi, "Machine Learning Method to Predict the Toddlers' Nutritional Status," *JURNAL INFOTEL,* vol. 16, 2024.

[25] N. Arfiah, "THE RELATIONSHIP BETWEEN FOOD INSECURITY AND THE INCIDENCE OF STUNTING," in *Proceedings of the International Conference on Nursing and Health Sciences*, Surakarta, 2024.

[26] A. Sampe, R. Claurita Toban and M. Anung Madi, "Relationship between Exclusive Breastfeeding and Stunting in Toddlers," *Jurnal Ilmiah Kesehatan Sandi Husada,* vol. 11, no. 1, pp. 448-455, 2020.

[27] Z. Zikrina, E. Halifah and D. Ardhia , "Factors That Influence In Exclusive Breastfeeding," *Idea Nursing Journal,* vol. XIII, no. 3, pp. 7-14, 2022.

[28] D. Ambarwati, "Pemanfaatan Teknologi Artificial Intelligence (AI) dalam Mengurangi Kejadian Stunting pada 1000 Hari Pertama Kelahiran: A literature Review," in *Seminar Nasional Penelitian dan Abdimas*, Jambi, 2023.

[29] M. Ula, A. Faridhatul Ulva and Mauliza, "Application of Machine Learning in Predicting Children's Nutritional Status with Multiple Linear Regression Models," *MULTICA SCIENCE AND TECHNOLOGY,* vol. 2, no. 2, pp. 42-48, 2022.

[30] M. Banurea, D. Betaria Hutagaol and O. Sihombing, "KLASIFIKASI PENYAKIT STUNTING DENGAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST," *Jurnal TEKINKOM,* vol. 6, no. 2, 2023.

AUTHOR BIOGRAPHY

Ayu Ahadi Ningrum received her bachelor degree 2012 from STIE Pancasetia Banjarmasin and 2018 from Politeknik Elektronika Surabaya. Her graduate from applied master degree from Departements Informatics and Computer in Politeknik Elektronika Surabaya. Since 2021, she has been the lecturer at Universitas Muhammadiyah Banjarmasin. Her research concern in Data Science, Deep Learning and how to optimize the algorithm.

Yunia Ikawati received her bachelor degree 2011 from Politeknik Elektronika Surabaya. Her graduate from applied master degree from Departements Informatics and Computer in Politeknik Elektronika Surabaya. Since 2022, she has been the lecturer at Politeknik Elektronika Negeri Surabaya

Corresponding author: Ayu Ahadi Ningrum, ayuahadi@umbjm.ac.id, Department of Informatics, Universitas Muhammadiyah Banjarmasin, Banjarmasin, Indonesia