

Identification of Stunting Disease using Anthropometry Data and Long Short-Term Memory (LSTM) Model

Faris Mushlihul Amin¹, Dian Candra Rini Novitasari^{2*}

¹Department of Information System, Faculty of Science and Technology, UIN Sunan Ampel

²Department of Mathematics, Faculty of Science and Technology, UIN Sunan Ampel

*diancrini@uinsby.ac.id

ABSTRACT

Children with unbalanced nutrition are currently crucial health issues and under the spotlight around the world. One of the terms for malnourished children is stunting. Stunting is a disease of malnutrition found in children aged under 5 years; as many as 70% of stunting sufferers are children aged 0-23 months. There are several ways to diagnose stunting, one of which is using stunting anthropometry. Stunting anthropometry can measure the physique of children so that some of the features that characterize the presence of stunting can be identified. Features resulted from the stunting anthropometry cover age, height, weight, gender, upper arm circumference, head size, chest circumference, and hip fat measurement. The process of identifying stunting can be simplified using an intelligent system called the Computer-Aided Diagnosis (CAD) system. CAD system contains 2 main processes, namely preprocessing and classification. Preprocessing includes normalization and augmentation of data using the SMOTE method. The classification process in this study uses the LSTM method. LSTM is a modification of the Recurrent Neural Network (RNN) method by adding a memory cell so that it can store memory data for a long time and in large quantities. The results of this study compare between the results of models that apply preprocessing and the one without preprocessing. The model that only uses LSTM has the best accuracy of 78.35%; the model with normalization produces an accuracy of 81.53%; the model that uses SMOTE produces an accuracy of 81.66%; and the model that uses normalization and SMOTE produces the best accuracy of 85.79%.

Keywords: Classification, LSTM, Long Short-Term Memory, Stunting, SMOTE.

1. INTRODUCTION

The growing cases of malnourished children is an important indicator that there is a lack of community welfare in one country. This problem occurs across the world, and there are millions of children who fail to reach their optimal growth potential due to inadequate nutrition [1]. Children with unbalanced nutrition are currently crucial health problem and become global concern. One of the terms for malnourished children is stunting [2]. Stunting is a disease of malnutrition found in children aged under 5 years; as many as 70% of stunting sufferers are children aged 0-23 months [3]. In 2014, World Health Organization (WHO) stated that there were 162 million children under the age of 5 who suffered from stunting. WHO also predicts that in 2025 there will be an additional 125 million stunting sufferers if the children are to be left untreated. The stunting has long-term effects and potentially becomes a degenerative disease or inherited disease such as diabetes [4].

Faris Mushlihul Amin, Dian Candra Rini Novitasari
Identification of Stunting Disease using Anthropometry Data and Long Short-Term Memory (LSTM) Model

Indonesia is one of the countries that has worrying rates of stunting, wasting, and over-weight. Based on data from R&D in 2018, there were around 30.8% of children in Indonesia with stunting. According to WHO, if the prevalence of a disease in a country is around 30%-39%, then the disease is in the severe category so that further treatment is needed by the government [5]. There are several ways to diagnose stunting, one of which is using stunting anthropometry. Stunting anthropometry can measure the physique of the children so that some of the significant features that characterize the presence of stunting can be identified. The features that become the focus of the stunting anthropometry are namely age, height, weight, gender, upper arm circumference, head size, chest circumference, and hip fat measurement of the children [6]. The process of identifying stunting can be simplified using an intelligent system called the Computer-Aided Diagnosis (CAD) system [7].

The CAD system is an intelligent system that is used to diagnose a disease. Usually CAD system contains 2 main processes, namely preprocessing and classification. Preprocessing is used to refine data to make it easier to classify. Anthropometric data has an uneven distribution of data so that it may cause the classification process to be less optimal. One of the preprocessing methods to overcome these problems is to apply data normalization. Data normalization makes the data to have an even distribution and can minimize the computation time because these normalized data have a smaller scale [8]. Although normalization can make the data experience a decrease in the value scale, the data still contains information and the normalization does not eliminate the information contained in a data [9]. Data related with stunting are not balanced between data for normal children and stunting sufferers; this problem can be overcome by using the Synthesis Minority Over Sampling Technique (SMOTE) method. This method poises the data by creating synthesis data in the data class that has the least amount [10], [11]. After the data is preprocessed, they are entered into a classification system. In this study, the method chosen for classification is deep learning.

Deep learning is one of the artificial neural network methods that is processed using a hidden layer that has a very large number; it can even reach hundreds of layers [12]. Deep learning methods are increasingly complex and accurate when the number of layers increases so that they can recognize data features better [13]. There are several studies related to stunting that have been identified using the deep learning method [14], [15] which were investigated by Shahriar, et al. In the study, one of the deep learning methods was applied, namely Artificial Neural Network (ANN) using survey data on children aged 0-59 months in Bangladesh. The data were classified using 3 classes of cases of malnutrition, namely stunting, wasting, and underweighting. This research compared the deep learning ANN method with the Support Vector Machine (SVM), Naïve Bayes and random forest. The best accuracy result in this study was using the ANN method, which was 84.93% [16]. Momand, et al. also applied deep learning methods to identify malnutrition in children in Afghanistan. The data used in this study were anthropometric data. This research resulted in higher accuracy than other machine learning methods such as Naïve Bayes and Logistic Regression [17].

One of the most popular deep learning methods is Long Short Term Memory (LSTM). LSTM is a modification of the Recurrent Neural Network (RNN) method by adding a memory cell so that it can store memory data for a long time and in large quantities [18]. Several studies have applied the LSTM method to identify diseases, one of which was carried out by Zheng, et al. In the study, the LSTM method was applied to identify arrhythmias disease using Electrocardiogram (ECG) data. The

study classified the data into 2 classes, namely normal and arrhythmias. The accuracy obtained in this study was 99.01% [19]. Bellone, et al. once did a comparison between CNN and LSTM to find out the best deep learning method. The case used in the study was to model a virtual sensor on a machine. LSTM was chosen as the best method with a small error value of 2.40% for fuel, 2.80% for NO_x, and 18.19% for soot sensors [20].

Based on the studies on LSTM which can perform almost accurate classification, prediction, and modeling, this study applies the LSTM deep learning method to classify stunting. It is expected that this study provides better results than some previous studies related to stunting.

2. MATERIAL AND METHODS

2.1 STUNTING

Stunting is a disease caused by lack of nutrition received by children, causing the children experience slow growth. In general, stunting is characterized by several aspects such as a lack of weight and height in children at a certain age and gender [21]. Most stunting occurs in children under 5 years old [22]. Stunting disease can also be caused by genetic, hormonal, and lack of nutrition [23]. Stunting can be identified using stunting anthropometry which can measure the physique of the children in terms of their age, height, weight, gender, upper arm circumference, head size, chest circumference, and hip fat measurement [24]. Stunting must be addressed immediately because stunting has long-term effects and has the potentials to become a degenerative disease or hereditary disease [4]

2.2 DATA COLLECTION AND ANALYSIS

The data used were numerical data obtained from one of the hospitals in Madura, East Java. The data consisted of 3 variables, namely gender, age, and height. There were 100 data obtained, with the detail of 62 normal data and 38 stunting data. Because the data were not balanced between normal and stunting data, it was necessary to apply the SMOTE method to balance the data. The data were divided into training data and testing data using K-Fold Cross Validation and classified using the LSTM method. This research includes quantitative descriptive research because the results of the research are in the form of an analysis of the computational results.

2.3 DATA NORMALIZATION

Normalization is a data mining technique that is used to overcome certain data that is randomly distributed. Normalization can facilitate the computational process because it has a smaller intermediate value [25]. There are several techniques for normalization, one of which is the min-max method. Min-max is a normalization technique that maps a data distribution into a 0 to 1 [26]. The min-max normalization equation can be seen in Equation (1):

$$x_i^{j'} = \frac{x_i^j - \min^j}{\max^j - \min^j} \quad (1)$$

where x_i^j is the origin data, $x_i^{j'}$ is the result of the normalization process, and max^j , min^j are the maximum and minimum values of the entire data used.

2.4 SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)

SMOTE is a method to overcome the problem of unbalanced data [10]. The SMOTE method proposed by Chawla utilizing an oversampling system using artificial data. This method duplicates data that has a small amount by creating random data but still in the same data distribution [27]. Illustration of SMOTE can be seen in Figures 1 and 2.

A Typical Machine Learning Problem: Class Imbalance

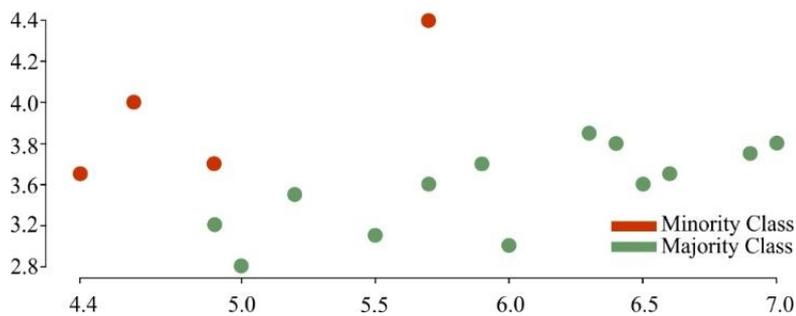


FIGURE 1. Original Distribution Data

Adressing Class Imbalance Problems of ML via SMOTE

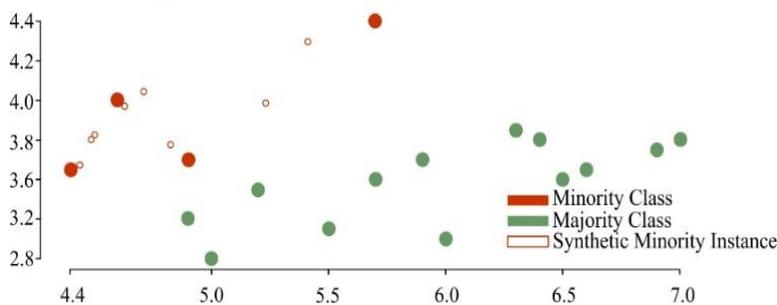


FIGURE 2. SMOTE Result

Several stages in the SMOTE process are prosed [11]:

1. Determine the nearest neighbor of the minor data. In Figure 1 minor data is shown with a red border color.
2. Calculating the neighboring distance between data using Equation (2)

$$Distance = \sqrt{\sum (x_i - x_j)^2} \quad (2)$$

3. Calculating the transition matrix that occurs using Equation (3).

$$Transition = rand(0,1)Distance(x,y) \quad (3)$$

4. Make data synthesis on minor data using Equation (4). The illustration of the results of the synthesis data can be seen in Figure 2 with small red dots.

$$\text{Synthetic Data} = \text{Data}(x, y) + \text{Transition}(x, y) \quad (4)$$

2.5 LONG SHORT-TERM MEMORY (LSTM)

LSTM is a deep learning method developed from the RNN method. The main idea behind this method is the memory cell and gate unit [28]. The memory cell applies a system similar to the brain system in recognizing an object, while the gate unit aims to decide which data processing to use. LSTM has 3 types of layers, namely input layer, hidden layer, and output layer. The hidden layer in the LSTM has the same system as the brain network that can sort out the data that should be forgotten and the data that should be remembered [29]. The hidden layer system used in LSTM can be seen in Figure 3.

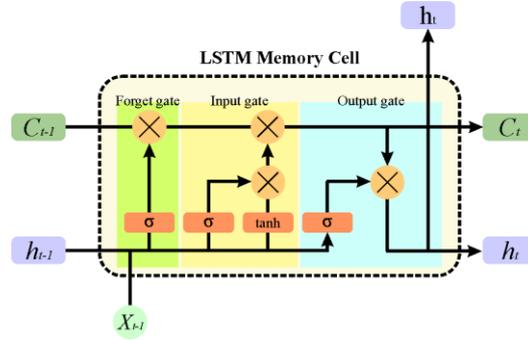


FIGURE 3. LSTM Architecture

In Figure 4 it can be shown that there is a forget gate to forget data that is not considered important, the input gate is used to process the data used and the output gate is useful for carrying out the final process before determining the classification results. In the LSTM x is the input data with the terms sequencing with $x = [x_1, \dots, x_t]$. Based on Figure 4, the equation used in the LSTM can be formed, namely [18]:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

$$\bar{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (7)$$

$$C_t = C_{t-1} * f_t + i_t * \bar{C}_t \quad (8)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t * \tanh(C_t) \quad (10)$$

where f_t is the result of the forget gate, σ is the activation of the sigmoid function, W_f is the weight of the forget output, h_{t-1} is the output of the $t - 1$ data, x_t is the input to the t data, b_f is the bias on the forget gate, i_t is the input gate, W_i is the weight value on the input gate, b_i is the bias that exists at the input gate, \bar{C}_t is the new value in the memory cell state, W_c is the weight that is in the memory cell state, o_t is the output gate, and h_t is the output on the t data.

2.6 CONFUSION MATRIX

Confusion matrix is a system evaluation method to determine the success rate of a CAD system [30]. The confusion matrix produces 3 parameters, namely accuracy to measure the success of classification, sensitivity to measure the success of identification of sick patients, and specificity to measure the success of identification of healthy or normal patients. These results can be calculated using several variables, namely True Positive (TP), Positive False (FP), False Negative (FN), and True Negative (TN) which can be seen in Table 1 [31].

TABLE 1.
Confusion Matrix

| Actual | Classification | |
|--------|----------------------------|-----------------------------|
| | + | - |
| + | <i>True Positive (TP)</i> | <i>False Negatives (FN)</i> |
| - | <i>False Positive (FP)</i> | <i>True Negatives (TN)</i> |

3. RESULTS AND DISCUSSION

Several stages used in this research are preprocessing and classification. The preprocessing used was data normalization using min-max and data augmentation using the SMOTE method. The classification used in this study is the LSTM method. The data used had 3 variables, namely age, height, and gender. The sample data used in this study can be seen in Table 2.

TABLE 2.
Sample Data

| Normal | | | Stunting | | |
|--------|--------|--------|----------|--------|--------|
| Age | Height | Gender | Age | Height | Gender |
| 28 | 90.1 | F | 24 | 69.6 | F |
| 30 | 93 | F | 25 | 70.1 | M |
| 31 | 91.8 | F | 29 | 73.6 | M |
| 32 | 88.8 | M | 33 | 84.5 | M |
| 34 | 90 | F | 35 | 86.7 | F |

The data on gender were converted into numerical form to make it easier to process. Data F is female patients and M for male patients. Gender F is defined as number 0 and gender M is defined as number 1.

3.1 PRE-PROCESSING

Preprocessing in the CAD system serves to improve data so that it is easier to process and produces higher accuracy. Preprocessing in this study contained 2 processes, namely normalization and data augmentation. The first process was normalization using the min-max method with equation (1). The results of the normalization process can be seen in Table 3.

TABLE 3.
Sample of Normalization Result

| Original Data | | | Normalization Data | | |
|---------------|--------|--------|--------------------|--------|--------|
| Age | Height | Gender | Age | Height | Gender |
| 24 | 69.6 | 0 | 0.400 | 0.321 | 0 |
| 25 | 70.1 | 1 | 0.418 | 0.329 | 1 |
| 29 | 73.6 | 1 | 0.490 | 0.388 | 1 |
| 33 | 84.5 | 1 | 0.563 | 0.571 | 1 |
| 35 | 86.7 | 0 | 0.600 | 0.608 | 0 |

The data generated from the normalization process were data with a distribution of 0 to 1. The data were easier to process because they did not have a high value so that they can speed up the classification process. The next preprocessing process was data balancing. Unbalanced data are often a problem in the classification system and causes poor accuracy. One way to balance data is to use the SMOTE method. This method forms the synthesis data to be used in the training process and helps the formation of the optimum model. The data carried out by the SMOTE process is only data that has the least amount of data in a certain class. In this study, the data used were stunting data because it has a smaller amount than normal data. The results of the formation of the synthesis data can be seen in Table 4.

TABLE 4.
Sample of SMOTE Result

| Age | Height | Gender |
|----------|----------|--------|
| 0.095923 | 0.101586 | 0 |
| 0.185068 | 0.241741 | 1 |
| 0.210882 | 0.233819 | 0 |
| 0.396355 | 0.546965 | 1 |
| 0.419433 | 0.367583 | 0 |

Table 4 is the result of the SMOTE process where the synthesis data were formed randomly with the same distribution limit as the stunting data. The resulted gender data were actually a decimal data between 0 to 1; as there are only 2 genders, namely male and female, the data were rounded off so that the data generated were only a number 0 or 1. The SMOTE process carried out was as much as 100% so that the comparison of the data which was originally 62 normal data and 38 stunting data became 62 normal data and 72 stunting data. Changes in SMOTE results can be seen in Figure 5.

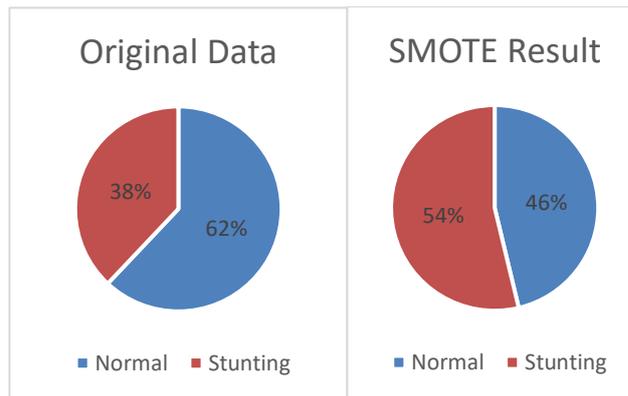


FIGURE 5. SMOTE Result

3.2 CLASSIFICATION

The results of SMOTE are classified using the LSTM method. Previously, the data had to be divided into two, namely training data and testing data. The data separation process was carried out using the K-Fold Cross Validation method using $k = 5$ so that it can divide the data into 80% training data and 20% testing data. The test results using K-Fold Cross Validation can be seen in Table 5.

Table 5 explains that each hidden node has a different accuracy value for each fold. In the hidden 50 and 100 the best results were achieved in the second Fold with an accuracy value of 81.9% and 83.22%, respectively. At hidden nodes of 150, 300, and 500, the best accuracy results were in the fifth fold with an accuracy value of 83.83%, 84.39%, and 85.79%, respectively. At the hidden node of 1000, the best results were obtained in fold 4 with an accuracy value of 84.81%. This concludes that many hidden nodes affect the accuracy results. More hidden nodes do not guarantee better accuracy, and vice versa. In knowing the best number of hidden nodes, trials must be carried out. The best results in this study were obtained at the number of hidden nodes as much as 500. These results produce an accuracy graph that can be seen in Figure 6 and an error graph that can be seen in Figure 7.

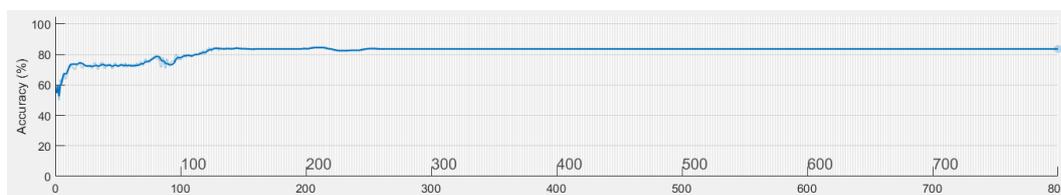


FIGURE 6. Graph of Accuracy

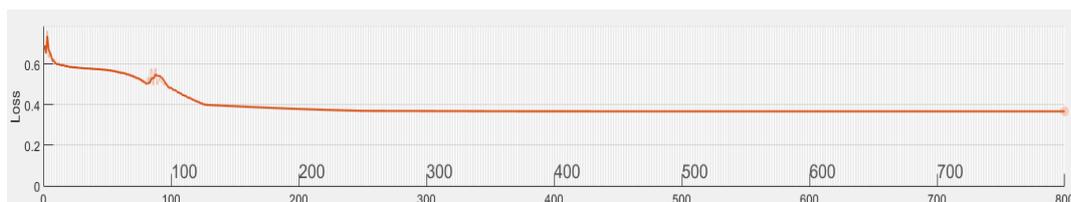


FIGURE 7. Graph of Error

TABLE 4.
Sample of SMOTE Result

| Number of Hidden Node | Kernel | Accuracy | Sensitivity | Specificity |
|-----------------------|--------|----------|-------------|-------------|
| 50 | Fold 1 | 79.28% | 81.2% | 77.37% |
| | Fold 2 | 81.9% | 82.71% | 81.1% |
| | Fold 3 | 74.15% | 73.64% | 74.67% |
| | Fold 4 | 80.1% | 82.1% | 78.1% |
| | Fold 5 | 72.3% | 73% | 71.62% |
| Max | | 81.9% | 82.71% | 81.1% |
| 100 | Fold 1 | 82.84% | 80.93% | 84.74% |
| | Fold 2 | 83.22% | 84.44% | 82% |
| | Fold 3 | 80.93% | 80.67% | 81.19% |
| | Fold 4 | 81% | 81.55% | 80.46% |
| | Fold 5 | 82.9% | 81.28% | 84.52% |
| Max | | 83.22% | 84.44% | 82% |
| 150 | Fold 1 | 83.1% | 82.38% | 83.69% |
| | Fold 2 | 81.46% | 82.9% | 80.1% |
| | Fold 3 | 81.37% | 80.91% | 81.82% |
| | Fold 4 | 81.82% | 81.94% | 81.7% |
| | Fold 5 | 83.83% | 82.89% | 84.78% |
| Max | | 83.83% | 82.89% | 84.78% |
| 300 | Fold 1 | 81.68% | 82.54% | 80.82% |
| | Fold 2 | 82.32% | 80.97% | 83.67% |
| | Fold 3 | 82.97% | 82.1% | 83.89% |
| | Fold 4 | 83.18% | 82.92% | 83.44% |
| | Fold 5 | 84.39% | 85.74% | 83.1% |
| Max | | 84.39% | 85.74% | 83.1% |
| 500 | Fold 1 | 84.13% | 91.88% | 80.37% |
| | Fold 2 | 84.67% | 87.33% | 82.1% |
| | Fold 3 | 85% | 80.1% | 91.87% |
| | Fold 4 | 84.89% | 89.53% | 80.26% |
| | Fold 5 | 85.79% | 83.5% | 89.31% |
| Max | | 85.79% | 83.5% | 89.31% |
| 1000 | Fold 1 | 84.74% | 81.25% | 90.22% |
| | Fold 2 | 84% | 82.69% | 85.3% |
| | Fold 3 | 81.52% | 80.85% | 82.19% |
| | Fold 4 | 84.81% | 83.4% | 85.22% |
| | Fold 5 | 83.5% | 82.17% | 85.82% |
| Max | | 84.81% | 83.4% | 85.22% |

Although the results obtained were not perfect, the results in Figures 5 and 6 are the best results that can be produced by LSTM on the stunting data used. It can be seen that the accuracy results start to be constant at the 300th iteration so that even if the number of hidden nodes are added, it would not affect the results because the results cannot change anymore unless retraining is done with a different data composition. This also applies to the error graph because the graph has an effect on each other. The best results from each number of hidden nodes can be seen in Table 6.

Faris Mushlihul Amin, Dian Candra Rini Novitasari
Identification of Stunting Disease using Anthropometry Data and Long Short-Term
Memory (LSTM) Model

TABLE 6.
 Best Result in Each Number of Hidden Node

| Number of Hidden Node | Accuracy | Sensitivity | Specificity |
|-----------------------|----------|-------------|-------------|
| 50 | 81.9% | 82.71% | 81.1% |
| 100 | 83.22% | 84.44% | 82% |
| 150 | 83.83% | 82.89% | 84.78% |
| 300 | 84.39% | 85.74% | 83.1% |
| 500 | 85.79% | 83.5% | 89.31% |
| 1000 | 84.81% | 83.4% | 85.22% |

In Table 6, it can be seen that all accuracy has reached above 80% but still none has reached 90%. This is because the LSTM system was originally devoted to sequencing data such as video data or data for the prediction process, so it is not suitable when using feature data such as stunting data. Despite the fact it is not suitable, LSTM is still able to recognize data well and produce high accuracy. Furthermore, the results of the accuracy of the training time will be compared if preprocessing is applied. The results of the comparison can be seen in Table 7.

TABLE 7.
 Comparison Between Several Model

| Method | Accuracy | Training Time |
|------------------------------|----------|---------------|
| LSTM | 78.26% | 6 m 39 s |
| SMOTE + LSTM | 84.32% | 9 m 21 s |
| Normalization + LSTM | 82.56% | 2 m 56 s |
| Normalization + SMOTE + LSTM | 85.79% | 4 m 12 s |

Based on the results from Table 7, it can be concluded that normalization can increase accuracy and reduce training time so that it is faster. The SMOTE method can also increase accuracy, but increase the computational time of the system used so that the application of normalization. The SMOTE method can increase the accuracy of the LSTM and increase the speed in LSTM training time. The best results obtained are 85.79% with a training time of 4 minutes 12 seconds.

4. CONCLUSION

Research on the identification of stunting is very important because this disease is a measure of the development of one country. This disease must also be addressed immediately because it can be inherited based on genes. This study uses the LSTM method as a means of an intelligent system to identify stunting disease automatically. With two processes of preprocessing and classification, the results of this study compare the results of models that apply preprocessing and the one without preprocessing. The model that only uses LSTM has the best accuracy of 78.35%, the model with normalization produces an accuracy of 81.53%, the model that uses SMOTE produces an accuracy of 81.66%, and the model that uses normalization and SMOTE produces the best accuracy of 85.79%.

REFERENCES

- [1] M. De Onis and F. Branca, "Childhood stunting: a global perspective," *Matern.*

- Child Nutr.*, vol. 12, pp. 12–26, 2016.
- [2] J. L. Leroy and E. A. Frongillo, “Perspective: what does stunting really mean? A critical review of the evidence,” *Adv. Nutr.*, vol. 10, no. 2, pp. 196–204, 2019.
 - [3] B. J. Akombi, K. E. Agho, J. J. Hall, D. Merom, T. Astell-Burt, and A. M. N. Renzaho, “Stunting and severe stunting among children under-5 years in Nigeria: A multilevel analysis,” *BMC Pediatr.*, vol. 17, no. 1, p. 15, 2017.
 - [4] World Health Organization, “Global nutrition targets 2025: stunting policy brief. 2014.” 2018.
 - [5] J. Sugiyanto, S. S. Raharjo, and Y. L. R. Dewi, “The effects of exclusive breastfeeding and contextual factor of village on stunting in bontang, east kalimantan, indonesia,” *J. Epidemiol. Public Heal.*, vol. 4, no. 3, pp. 222–233, 2019.
 - [6] J. R. Batubara, S. Bardosono, A. Irawati, and S. Salimar, “Comparing The Anthropometric Measurements of Intra-Extra Uterine Period between Stunting and Non-stunting Children Aged 6-24 Months Old in Bogor Tengah Subdistrict, Bogor City, West Java,” *World Nutr. J.*, vol. 3, no. 1, pp. 1–7, 2019.
 - [7] A. Z. Foeady, D. C. R. Novitasari, A. H. Asyhar, and M. Firmansjah, “Automated Diagnosis System of Diabetic Retinopathy Using GLCM Method and SVM Classifier,” *Proceeding Electr. Eng. Comput. Sci. Informatics*, vol. 5, no. 5, pp. 154–160, 2018.
 - [8] S. Weiss *et al.*, “Normalization and microbial differential abundance strategies depend upon data characteristics,” *Microbiome*, vol. 5, no. 1, p. 27, 2017.
 - [9] S. J. Mukhyber and A. D. M. DhahirAbdulahdeAbdulah, “Effect Z-score Normalization on Accuracy of classification of liver disease,” *Turkish J. Comput. Math. Educ.*, vol. 12, no. 14, pp. 658–662, 2021.
 - [10] A. Fern and S. Garc, “SMOTE for Learning from Imbalanced Data : Progress and Challenges , Marking the 15-year Anniversary,” vol. 61, pp. 863–905, 2018.
 - [11] D. C. R. Novitasari *et al.*, “Whirlwind Classification with Imbalanced Upper Air Data Handling using {SMOTE} Algorithm and {SVM} Classifier,” *J. Phys. Conf. Ser.*, vol. 1501, p. 12010, Mar. 2020, doi: 10.1088/1742-6596/1501/1/012010.
 - [12] M. B. Bora, D. Daimary, K. Amitab, and D. Kandar, “Handwritten Character Recognition from Images using CNN-ECOC,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2403–2409, 2020, doi: 10.1016/j.procs.2020.03.293.
 - [13] S. H. Wang, K. Muhammad, J. Hong, A. K. Sangaiah, and Y. D. Zhang, “Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization,” *Neural Comput. Appl.*, vol. 32, no. 3, pp. 665–680, 2020, doi: 10.1007/s00521-018-3924-0.
 - [14] M. N. Aulia, M. L. Khodra, and A. P. Koesoema, “Predicting Macronutrient of Baby Food using Near-infrared Spectroscopy and Deep Learning Approach,” 2020.
 - [15] M.-L. How and Y. J. Chan, “Artificial Intelligence-Enabled Predictive Insights for Ameliorating Global Malnutrition: A Human-Centric AI-Thinking Approach,” *AI*, vol. 1, no. 1, pp. 68–91, 2020.
 - [16] M. Shahriar, M. S. Iqubal, S. Mitra, and A. K. Das, “A Deep Learning Approach to Predict Malnutrition Status of 0-59 Month’s Older Children in Bangladesh,” in *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 2019, pp. 145–149.

Faris Mushlihul Amin, Dian Candra Rini Novitasari
Identification of Stunting Disease using Anthropometry Data and Long Short-Term
Memory (LSTM) Model

- [17] Z. Momand, P. Mongkolnam, P. Kositpanthavong, and J. H. Chan, "Data Mining Based Prediction of Malnutrition in Afghan Children," in *2020 12th International Conference on Knowledge and Smart Technology (KST)*, 2020, pp. 12–17.
- [18] A. H. Asyhar, A. Z. Foady, M. Thohir, A. Z. Arifin, D. Z. Haq, and D. C. R. Novitasari, "Implementation LSTM Algorithm for Cervical Cancer using Colposcopy Data," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, 2020, pp. 485–489.
- [19] Z. Zheng, Z. Chen, F. Hu, J. Zhu, Q. Tang, and Y. Liang, "An Automatic Diagnosis of Arrhythmias Using a Combination of CNN and LSTM Technology," *Electronics*, vol. 9, no. 1, p. 121, 2020.
- [20] M. Bellone, E. Faghani, and Y. Karayiannidis, "Comparison of CNN and LSTM for Modeling Virtual Sensors in an Engine," SAE Technical Paper, 2020.
- [21] K. Astarani, D. I. S. H. Poernomo, D. N. T. Idris, and A. R. Oktavia, "Prevention of Stunting Through Health Education in Parents of Pre-School Children," *Str. J. Ilm. Kesehatan.*, vol. 9, no. 1, pp. 70–77, 2020.
- [22] S. Habimana and E. Biracyaza, "Risk factors of stunting among children under 5 years of age in the eastern and western provinces of Rwanda: analysis of Rwanda demographic and health survey 2014/2015," *Pediatr. Heal. Med. Ther.*, vol. 10, p. 115, 2019.
- [23] A. K. Yadav and S. T. Karki, "Short Stature in Children Visiting Endocrine Out Patient Department of Kanti Children's Hospital, Nepal," *J. Coll. Med. Sci.*, vol. 17, no. 1, 2021.
- [24] H. da Silva Ferreira, "Anthropometric assessment of children's nutritional status: a new approach based on an adaptation of Waterlow's classification," *BMC Pediatr.*, vol. 20, no. 1, pp. 1–11, 2020.
- [25] S. Kappal, "Data normalization using median median absolute deviation MMAD based Z-score for robust predictions vs. min-max normalization," *London J. Res. Sci. Nat. Form.*, 2019.
- [26] B. N. Silva *et al.*, "Urban planning and smart city decision management empowered by real-time data processing using big data analytics," *Sensors*, vol. 18, no. 9, p. 2994, 2018.
- [27] A. Fernández, S. Garcia, F. Herrera, and N. V Chawla, "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," *J. Artif. Intell. Res.*, vol. 61, pp. 863–905, 2018.
- [28] D. Z. Haq *et al.*, "Long Short-Term Memory Algorithm for Rainfall Prediction Based on El-Nino and IOD Data," *Procedia Comput. Sci.*, vol. 179, pp. 829–837, 2021.
- [29] J. Zhang, P. Wang, R. Yan, and R. X. Gao, "Long short-term memory for machine remaining life prediction," *J. Manuf. Syst.*, vol. 48, pp. 78–86, 2018.
- [30] S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion Matrix-based Feature Selection," *MAICS*, vol. 710, pp. 120–127, 2011.
- [31] M. Thohir, A. Z. Foady, D. C. R. Novitasari, A. Z. Arifin, B. Y. Phiadelvira, and A. H. Asyhar, "Classification of Colposcopy Data Using GLCM-SVM on Cervical Cancer," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, 2020, pp. 373–378.