

MACHINE LEARNING MODEL FOR STUNTING PREDICTION

Sutarmi, Warijan Warijan, Tavip Indrayana, Dwi P. Putro b, Indra Gunawan
Poltekkes Kemenkes Semarang
Sekolah Tinggi Teknologi Ronggolawe
Email: sutarmisaja15@gmail.com

Keywords:
Machine Learning,
Stunting, Detection,
Prediction

ABSTRACT

This study aims to find the best Supervised Machine Learning (SML) model for stunting prediction. This research was conducted using an experimental approach using 192 infant data with a composition of 183 normal infant data and 9 stunted infant data using a custom dataset. The conclusion obtained from this study can be concluded that the combination of the Random Forest classification algorithm with Support Vector Machine Weighting and the Genetic Algorithm Feature Selection has the best performance. The parameters with the best performance are: The training and testing data distribution is 90% of the training data and 10% of the testing data. The number of trees in the random forest algorithm is 100, and the Gain Ratio criterion and max_depth is 10. In the Genetic Algorithm, the best parameters are: The Roulette Wheel selection method, the population is 20, the mutation value is 0.03, and the crossover value is 0.9. The validation method uses k-fold cross validation with a value of k = 10. Another conclusion is that there are 44 supporting factors for stunting, which, if we take a ranking of 10 in order of magnitude from largest to smallest, the supporting factors for stunting are 1. Baby's weight at birth. 2. Baby's Height at Birth. 3. Number of meal per day. 4. Breast Milk. 5. Diarrhe times per 3 month. 6. Child development examination during covid by Health Worker at home. 7. Mother's age at birth. 8. Mother height at birth. 9. Number of sibling. 10. Age when the first food was given. This research has the disadvantage of no test on other datasets. So researchers do not know the reliability of findings is on different datasets.

INTRODUCTION

ML (Machine Learning) is a branch of intelligence science in software technology that utilizes a specific artificial intelligence algorithm based on data knowledge to solve problems independently when given new data. Currently, the topic of ML is hitting a sharp rise in the research world (Jansen et al. 2023). Research topics related to ML are in information technology and other fields such as health, military, marketing, social, politics, and others (Fallah, Karami, and Kojouri 2023). In the health sector, reports state that 86% of health organizations use ML solutions to help their business (Aminizadeh et al. 2023). Furthermore, we searched using the Publish or Perish application using the keywords machine learning, healthcare with search restrictions in 2022 and only covering Google Scholar and English searches, obtained no less than 3779 article citations, and predicted the trend would continue to increase.

In line with the advancement of computer hardware technology, It is also a lot of datasets of various fields that exist on the internet (Mishra and Tyagi 2022). One

of the uses of ML in the health sector is as a tool for detecting factors that cause stunting (Berhe et al. 2019). ML works by recognizing patterns of stunting data from existing stunting cases in the dataset, and then ML can predict stunting from the new data entered (Vaivada et al. 2020). As a promising technology, ML can solve problems related to: regression, prediction, association, classification, and clustering (Lee et al. 2017). However, with all the advantages and benefits several, ML has several drawbacks (Abid et al. 2022). For example, many algorithms can be chosen according to the parts of the problem, and many algorithms can be selected according to the features of the problem and data (Kotthoff 2016). Each algorithm has different accuracy according to the given dataset (Dogan and Tanrikulu 2013). So the algorithm with the best accuracy value, does not necessarily produce the same accuracy for other problems (Bierman, Li, and Lu 2023). The problems don't just stop here. If you have chosen an algorithm, the accuracy value will also vary according to the parameters (Zhang et al. 2023). The culmination of the problem is that there is no parameter value to get the best results, and the problem goes through a process of trial and error (Sohrabi et al. 2023). So there is research opportunity to find algorithms and parameters for each case (Salachoris et al. 2023).

The contribution of this study is to propose an ML model that performs better for stunting detection (Ndagijimana et al. 2023). This ML model includes algorithm selection, oversampling, feature selection, feature weighting, data distribution, and determining algorithm parameters (Sultana and Islam 2023). Another contribution is formulating the causes of stunting (Simatupang, Gultom, and Rahman 2023).

1.1 Related Paper

Based on our search, several studies are similar to this research. The first study analyzed and conducted an analysis of ML performance on stunting data in Zambia. The study concluded that the Random Forest algorithm had the best performance and succeeded in selecting 13 factors that caused stunting from 58 input variables. The second the causes of stunting using ML. The results of this study state that the Extreme Gradient Boost algorithm has the best performance. This algorithm is able to select 5 factors supporting stunting through the Ethiopian stunting dataset. The third study from. This research concludes that the Random Forest algorithm has the best performance. This study uses 5 input variables and 1 output label.

The synthesis result collection concludes that there has been no research using datasets with a total of 73 input variable columns. The use of survey methods to obtain research data. The research location is the Blora Regency, Central Java Province. Retrieval of this dataset through the SITEKSTAGI (Nutrition Status Detection System) application in our previous research. This dataset has 192 rows with details of 183 with normal status and 9 with stunting status. This dataset has 78 input variables and 1 output label.

1.2 Machine Learning

There are four types of machines: Unsupervised Learning, Supervised Learning, Semisupervised Learning, and Reinforcement learning. Unsupervised learning is a type of ML used when the available data does not have an output label. Supervised learning is a type of ML whose data has output label assistance. Semisupervised Learning is ML used to solve combination problems of Unsupervised Learning and

Supervised Learning. Reinforcement Learning is the use of ML types to solve problems so that they can find the best solution.

1.3 Supervised Learning

Supervised learning is a form of ML category that requires the help of output labels on the dataset it processes. In this study, we used nine Supervised Learning classification algorithms. There are many Supervised Learning classification algorithms available. Logistic Regression (LR), is a classification algorithm that limits each label/class, Then, each variable will be searched based on its proximity to these boundaries to find a predicted relationship between the variable and the label/class. LR is generally used in applied statistics to solve discrete analysis problems. Support Vector Machine (SVM) is a classification algorithm that finds the maximum hyperplane value. To separate the table can use SVM. K-Nearest Neighbor (KNN), is a classification algorithm that finds the similarity or distance of each input variable to the output variable(label/class). A Decision Tree (DT), is a classification algorithm that makes a decision tree diagram and considers each component to find the relationship between each input variable and output label. Random Forest(RF), is a development classification algorithm from DT. This Algorithm form a collection of decision trees using a random sample as the basis for the decision tree, Comparison of the results of each decision tree to get the best value. Gradient Boost(GB), is also an algorithm derived from DT. This algorithm works the way. Naïve Bayes(NB), is an algorithm that. Neural Network (NN), is an algorithm that imitates the workings of neuron cells in the human brain. Deep Learning (DL), is an algorithm developed from the principle of the NN algorithm. The difference between DL and NN principally lies in the existence and number of hidden layers used [8].

1.4 Feature Weighting

Variable input weighting can improve the performance of the feature selection process. In the Rapidminer application, there are several weighting methods provided, namely: Information Gain(IG), Gain Ratio(GR), Correlation(Corr), Chi Square (Chi), Gini Index(Gini), and Support Vector Machine(SVM) [8].

1.5 Feature Selection

Feature selection plays a critical, crucial role in classifying ML to improve its performance. In classification problems, not all input variables/features in the dataset affect the output label. The existence of variables that do not have this influence decreases the overall performance of the ML algorithm. Feature selection is in charge of selecting which variables trial impact powering do not affect influencing the output label and eliminating other variables that do not affect the output label so that it can improve ML performance [9]. This study uses the Rapidminer 9.10 application as an ML experimental device. The Rapidminer application provides several feature selection methods, Forward Selection, Backward Selection, and Optimized Selection. In this study, Optimized Selection (Evolutionary) feature selection was selected using a Genetic Algorithm. This Algorithm was chosen because it is proven, based on existing literature, to improve ML accuracy. Another advantage of GA is that its performance can be optimized based on its parameters.

The Genetic Algorithm (GA) is a heuristic algorithm that mimics the process of the genetic evolution of living things. The GA will generate as many random values as the population values. Then crossbreed with other populations with a percentage based on the crossover value, calculate the fitness value, then cross again with another random population to produce new generations as many as the generation value to find the best fitness value. The GA can be optimal by changing the parameters of feature selection, population size, number of generations, mutation percentage, crossover percentage, and the crossover method [10].

RESEARCH METHODS

This research applied an experimental approach. The dataset used comes from the SITEKSTAGI application from our previous study [7]. The beginning of this research goes through the pre-processing stage. Next, in the pre-processing step, researchers select datasets based on irrelevant data such as addresses, names, and others. Then, the researcher performs a correction stage on the wrong data when filling in the data. Next is the encoding stage to change the nominal data type to numeric. The encoding stage applies by breaking each variable column into several variables based on the content of the data. So that each column of the input variable has only the same data, normalize the data to ensure all data has the correct distribution of values. After cleaning the data, we leave 73 independent variables in the form of numbers and 1 dependent variable/output label in the form of stunting and normal labels.

Furthermore, researchers separate the data into parts, namely training and testing data. The use of training data to perform algorithm learning. The use of data testing to perform algorithm testing. To ensure that the calculation results are valid and consistent, researchers conduct testing using the k-cross validation method with a value of $k = 10$. Figure 1 below describes the flow of research performed.

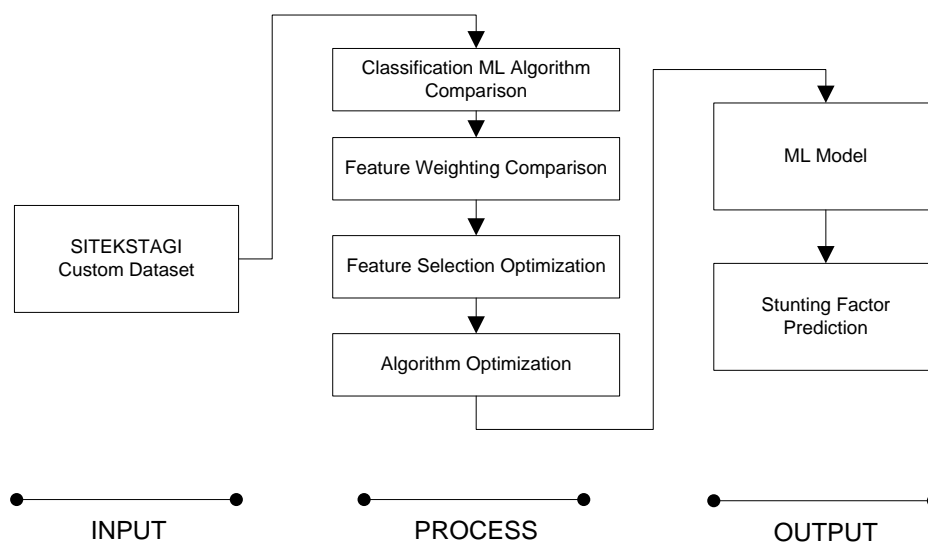


Figure 1. Research Workflow.

RESULTS AND DISCUSSION

Distribution of Data on the Dataset

The SITEKSTAGI dataset has 192 rows containing 183 normal baby data and 9 stunted baby data. Due to the imbalance in the amount of data, the machine learning algorithm will be inaccurate. On the other hand, we wanted to keep the synthetic data from getting too much. So the application of the SMOTE technique (Sampling Minority Over) by adding 300% synthetic data for stunted babies. In addition, the dataset remained in an unbalanced position with stunting percentages and normal data from 1:24 to 1:7, The results as follows:

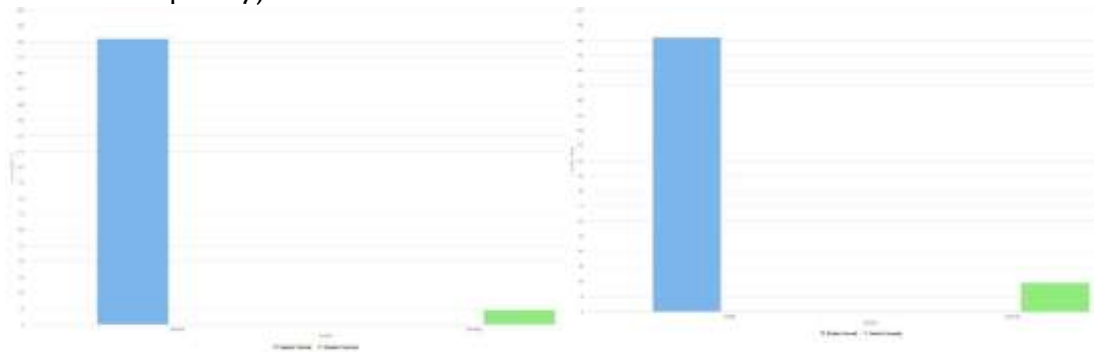


Figure 2. Comparison of Original Data vs SMOTE Data

Algorithm Selection with Feature Selection

The first step is to compare the performance of nine classification algorithms, namely: k-Nearest Neighborhood(KNN), Naives Bayes(NB), Random Forest(RF), Decision Tree(DT), Support Vector Machine(SVM), Logistic Regression(LR) , Neural Network(NN), Deep Learning(DL) and Gradient Boost(GB). The results of this comparison are as follows:

Table 1 Performance of Nine ML Algorithm

	DL	LR	SVM	NN	KNN	NB	DT	RF	GB
Accuracy	0.95	0.93	0.96	0.96	0.95	0.63	0.95	0.02	0.93
AUC	0.77	0.77	0.47	0.76	0.66	0.5	0.63	0.40	0.73
Precision	0.5	0.36	1.0	0.8	0	0.05	0.5	0	0.33
Recall	0.11	0.44	0.22	0.44	0	0.33	0.33	0	0.33
F_measure	0.18	0.4	0.36	0.57	0	0.07	0.4	0	0.33

The performance assessment in table 1 above was carried out on the original data with a comparison of stunting data compared to normal, which was 1:20 with details of 9 stunting and 182 normal. The results showed the score had poor scores for all algorithms and five scoring elements. It proves that if ML gives an unbalanced set on each label of the output data and there is too little data, then the ML performance will be poor.

The next stage is to perform ML optimization to increase the scores of the five assessment elements in all algorithms. Then the addition of stunting data by 300% artificial data with the SMOTE technique. Then the composition of the stunting and normal data ratio becomes 1:7 with details of 19 edits and 182. The next step is to evaluate the performance, are as follows:

Table 2 Performance of nine algorithm to SMOTE dataset

	DL	LR	SVM	NN	KNN	NB	DT	RF	GB
Accuracy	0.93	0.86	0.90	0.96	0.91	0.66	0.91	0.96	0.92
AUC	0.96	0.5	0.68	0.98	0.98	0.5	0.85	0.98	0.94
Precision	0.88	0	0.76	0.87	0.86	0.30	0.73	0.90	0.74
Recall	0.66	0.0	0.48	0.86	0.43	0.86	0.68	0.90	0.79
F_measure	0.73	0	0.56	0.86	0.59	0.43	0.66	0.87	0.73

The conclusion from the performance test results of nine machine learning algorithms is that the Random Forest algorithm has the best performance. Then the selection of the RF algorithm for the optimization process is in the next step.

Selection of Feature Weights

This stage tests the five input variable weighting methods to know which input variable influences stunting or normal. The results are as follows:

Table 3 Feature Weighting Performance on RF Algorithm

	Chi	IG	Corr	Look	SVM
Accuracy	0.97	0.96	0.96	0.97	0.97
AUC	0.99	0.99	0.99	0.99	0.99
Precision	0.94	0.96	0.94	0.96	0.91
Recall	0.87	0.80	0.84	0.86	0.90
F_measure	0.89	0.86	0.86	0.89	0.90

Table 3 shows that the SVM method scored the highest for the five assessment elements. Then, the application of the SVM method goes through the following testing stage.

Genetic Algorithm Optimization at the Feature Selection stage

This study uses a genetic algorithm as the method selected at the feature selection stage. The performance of the genetic Algorithm depends on the given parameter values. The parameter to be tested at this stage is the feature selection method. There are 5 selection methods, Tournament, Roulette Wheel, Boltzman, Stochastic, and Non-Dominated Sorting. From the test results, in table 3 the

Roulette Wheel method has the best performance with an accuracy value of 0.96, AUC of 0.99, precision of 0.93, recall of 0.83, and f_measure of 0.86. Furthermore, we will select the Roulette Wheel method in the next test.

Table 4 Feature Selection Performance on GA Algorithm

	Non	TM	BM	RW	Uniform	ST
Accuracy	0.96	0.95	0.95	0.96	0.95	0.96
AUC	0.99	0.99	0.99	0.99	0.99	0.99
Precision	0.90	0.90	0.87	0.93	0.88	0.96
Recall	0.84	0.80	0.74	0.83	0.83	0.77
F_measure	0.86	0.83	0.80	0.86	0.83	0.83

After finding the best feature selection parameters on GA, the next step is to test the crossover method on GA to see the best performance. Approval of three crossover methods: One Point, Uniform and Shuffle. The test results are as follows

Table 5 Crossover method Performance on GA Algorithm

	OP	Uniform	Shuffles
Accuracy	0.95	0.95	0.96
AUC	0.99	0.99	0.99
Precision	0.89	0.88	0.95
Recall	0.83	0.80	0.80
F_measure	0.83	0.82	0.84

Test results in table 5 show that the Shuffle method has the best performance with an accuracy value of 0.96, AUC of 0.99, precision of 0.95, recall of 0.80, and f_measure of 0.84. Next, the use of the Shuffle method goes through the following stages of testing. The next test for Genetic Algorithm (GA) parameters is the percentage value of crossover, mutation, and population. At this stage, tested four value variations were:

- a) Crossover = 0.9, mutation = 0.03, and population = 5
- b) The default value of the rapidminer application is crossover = 0.5, mutation = 1.0, and population = 5
- c) Crossover = 0.9, mutation = 0.03, and population = 20
- d) Crossover 0.5 mutation = 1.0, and population = 20

The results are as follows:

Table 6 Crossover, Mutation, Population Performance on GA Algorithm

	c:0.9/ m:0.03 /p:5	c:0.5/ -1.0/ p:5	c:0.9/ m:0.03 /p:20	c:0.5 /-1.0/ p:20
Accuracy	0.96	0.98	0.97	0.96
AUC	0.99	0.99	0.99	0.99
Precision	0.91	0.96	0.96	0.90
Recall	0.87	0.90	0.83	0.83
F_measure	0.88	0.93	0.87	0.83

The test results in table 6 conclude that the combination of crossing values of 0.9, mutations of 0.03, and a population of 20 produced the best scores with accuracy,

AUC, precision, recall, and f_measure. Furthermore, the researchers will enter this value in the next stage of testing.

Split Data Testing

The division of Training and Testing data (Data splitting) is an essential factor in machine learning theory. The researchers divide the dataset into two parts with a certain percentage. Machine Learning uses the training data to carry out the learning process, while Machine Learning uses the testing data to carry out the testing process. The percentage distribution of training and testing data is very influential on the results of algorithm performance. Then the researchers tested the data separation of the five methods and derived its value from literature studies. The results of these tests are as follows:

Table 7 Split Data Percentage Performance on RF Algorithm

	08:02	06:04	09:01	07:03	0677:0377
Accuracy	0.96	0.95	0.98	0.94	0.94
AUC	0.99	0.98	0.99	0.98	0.98
Precision	0.94	0.90	0.96	0.87	0.90
Recall	0.80	0.80	0.90	0.73	0.67
F_measure	0.84	0.82	0.93	0.79	0.76

The conclusion of the split data percentage test results in table 6 is that the training data rate compared to the best-performing data test is 0.9:0.1. Researchers will use this percentage in the following testing stage.

Optimize the Random Forest Algorithm

The performance of machine learning algorithms depends on the given parameters. Each algorithm has different parameters. So it is necessary to test these parameters to determine which produces the best performance value. At this stage, four criteria assessment methods will be tested on RF, namely: Gain Ratio (GR), Correlation (Corr), Gini Index (Gini), and Support Vector Machine (SVM). The results of these tests are as follows:

Table 8 Criteria Assessment Method Performance on RF Algorithm

	Gr	Corr	Look	SVM
Accuracy	0.96	0.96	0.97	0.97
AUC	0.99	0.99	0.99	0.99
Precision	0.96	0.94	0.96	0.91
Recall	0.80	0.84	0.86	0.90
F_measure	0.86	0.86	0.89	0.90

The test results in table 7, the highest accuracy values are in the Gini and SVM methods, Still, the SVM method has the highest recall score, which means that the detection of stunting labels is better. The highest f_measure value, which means a

balance between the number of stunting detections (recalls) and the accuracy of stunting detection (precision), is better than other methods.

Model Testing on Nine Algorithms

After getting all the optimal parameters, namely the feature weighting method, feature selection, selection scheme on GA, crossover method on GA, crossover value, mutation and population on GA, and the percentage of split data, Researchers use this parameter to test other algorithms. The results are as follows:

Table 9 Performance of Nine Algorithm using Chosen Parameter

	DL	LR	SVM	NN	KNN	NB	DT	RF	GB
Accuracy	0.94	0.94	0.96	0.95	0.94	0.64	0.96	0.98	0.93
AUC	0.99	0.99	0.98	0.99	0.99	0.5	0.99	0.99	0.97
Precision	0.84	0.84	0.93	0.94	1.0	0.28	0.89	0.97	0.80
Recall	0.84	0.86	0.8	0.74	0.60	0.89	0.83	0.90	0.83
F_measure	0.81	0.82	0.83	0.79	0.72	0.42	0.83	0.92	0.78

The conclusion of the test results of the nine ML algorithms in table 8 uses the best parameters from the previous stage; namely, the RF algorithm has the best performance and excels at five accuracies, AUC, precision, recall, and f_measure score. Then the researcher chooses the RF algorithm for parameter optimization to get the best performance.

Optimization of the Random Forest Classification Algorithm

Random Forest is an ML algorithm whose performance can be adjusted using existing parameters. Parameters that can affect the results of the implementation of the RF algorithm are the value of trees(T), the value of max_depth(D), criterion(C), and pruning(P). The results of testing the four parameters are as follows:

Table 10 RF Optimization Result

	ACC	GI	IG	default(T=100, D=10, C=GR, P=false)	T=110, D=10, C=GR, P=false	T=115, D=15, C=GR, P=false	T=120, D=20, C=GR, P=false	T=120, D=20, C=GR, P=true
accuracy	96.21%	95.74%	96.69%	0.98	0.96	0.96	0.95	0.95
AUC	96.21%	0.998	0.998	0.99	0.99	0.99	1	0.99
precision	96.21%	94.17%	94.17%	0.97	0.96	0.96	0.95	0.96
recall	80.83%	77.50%	84.17%	0.90	0.80	0.84	0.77	0.70
f_measure	84.81%	82.81%	86.81%	0.92	0.86	0.87	0.81	0.78

The table above concludes that the best random forest parameters are in tress = 100, depth = 10, criteria = gain ratio, and pruning disabled.

Feature Rank

Feature rank/variable input rank is the results of learning and pattern recognition of datasets using ML results of learning and pattern recognition. The result of the

Feature rank is the weight of the relationship between each input variable and the output label (Stunting / Normal). The results of ranking each variable using the SVM method can be as follows:

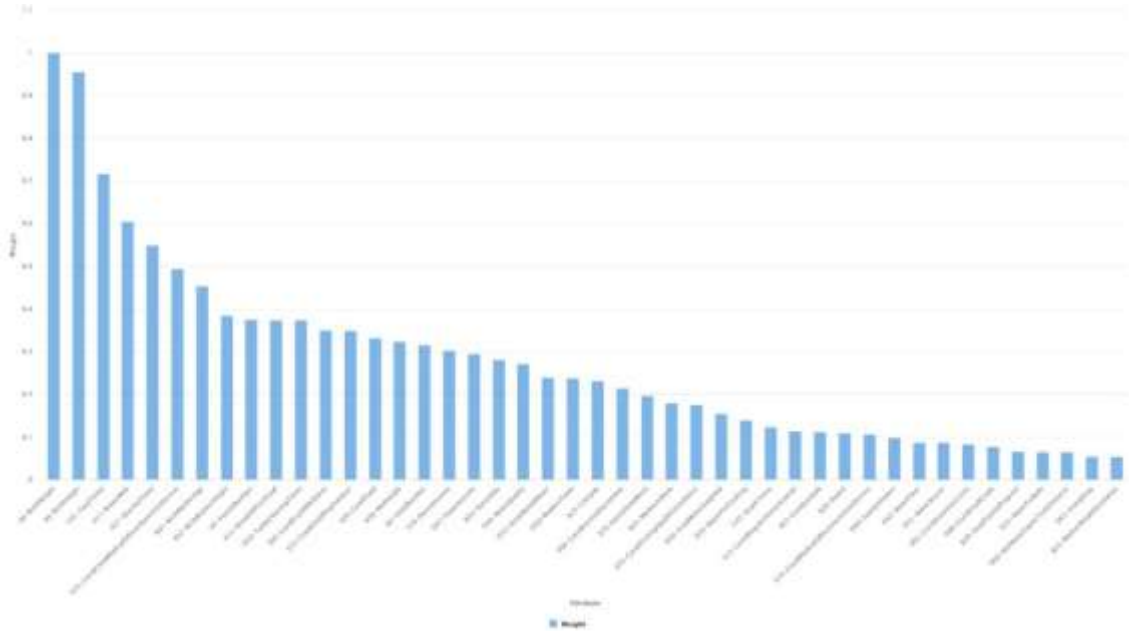


Figure 1 Feature Rank

The conclusion from figure 3 is that 44 of the 73 factors affect stunting status. The ranking of each of these variables is as follows:

Table 11. Feature Rank

Rank	Weight	Attribute
1	1,0	A8- Baby's weight at birth
2	1,0	A9- Baby's Height at Birth
3	0,7	C41-Number of meal per day
4	0,6	A11-Breast Milk
5	0,5	A27-Diarrhe times per 3 month
6	0,5	E75- Child development examination during covid by Health Worker at home
7	0,5	B31- Mother's age at birth
8	0,4	B32- Mother height at birth
9	0,4	A5- Number of sibling
10	0,4	A12- Age when the first food was given
11	0,4	D59- Toilet cleaning status at least 2-3 times a week
12	0,3	E69- Provision of nutritious food during covid
13	0,3	E73-Kids playing outdoor during covid
14	0,3	A30-Provision of nutritious food
15	0,3	B36-Mother job
16	0,3	A4-Order in family
17	0,3	F78-Parent Income
18	0,3	D47-Hypertension
19	0,3	B34-Parent Education
20	0,3	D50-Quality of Drinking water
21	0,2	D53-Drinking Boiled Water
22	0,2	D58-Using Modern Toilet
23	0,2	A25-Measles vaccinated
24	0,2	E66-Decrement of income during covid
25	0,2	E76- Handwashing equipment at home
26	0,2	B35- Does Mother Work
27	0,2	E70-Parent visiting medical facility during covid
28	0,2	D56-Clean water supply
29	0,1	D54- Using clean water for cooking
30	0,1	C42-Snack Times per day
31	0,1	E72-Immunization during covid
32	0,1	E67-Parents lost their job during covid
33	0,1	A20-Polio4 Vaccinated
34	0,1	E74-There are health workers who provide home services during covid
35	0,1	D49- There are smokers in the family
36	0,1	D52-Using Water Filter
37	0,1	D51- Does the family use dug wells
38	0,1	E65- Has the family experienced an economic impact due of covid
39	0,1	E68-Grade food supply during covid
40	0,1	B39-Family planning and birth control program
41	0,1	D55-Use cleaned water for bathing
42	0,1	D60-Distance of septic tank and water source less than 10 meteres
43	0,1	D63-Using Soap
44	0,1	B33- Escalation Mother's weight during pregnant

Proposed Machine Learning Models

The contribution of this study is to propose a Supervised Learning classification model for detecting stunting cases using seven stages, namely: 1. SMOTE Upsampling. 2.Variable encoding. 3. Data Normalization. 4. SVM Feature Weighting 5.

Genetic Algorithm Feature Selection. 6. Data Splitting, and 7. Random Forest Classification.

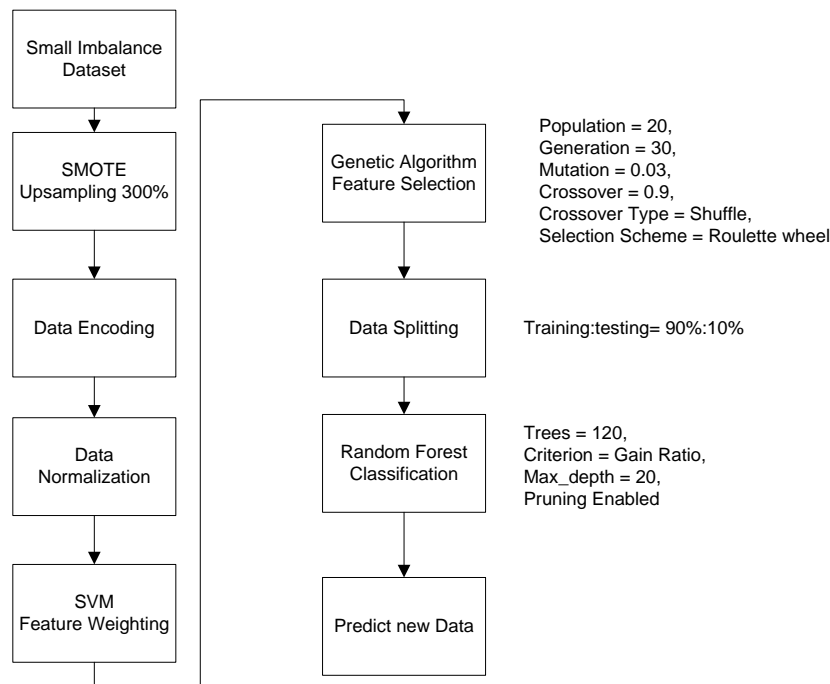


Figure 4. Proposed Model of ML

CONCLUSION

This study concludes that machine learning can predict stunting status in infants according to the test results on small imbalance datasets in several machine learning classification algorithms and their parameters. Thus, combining the Random Forest classification algorithm, weighting with a Support Vector Machine, and feature selection with the Genetic Algorithm has the best performance. The parameters with the best performance are: The training and testing data distribution is 90% of the training data and 10% of the testing data. The number of trees in the random forest algorithm is 100, and the Gain Ratio criterion and max_depth are 10. In the Genetic Algorithm, the best parameters are: the Roulette Wheel selection method, the population is 20, the mutation value is 0.03, and the crossover value is 0.9. The validation method uses k-fold cross validation with a value of $k = 10$. Another conclusion is that there are 44 supporting factors for stunting. If we take a ranking of 10 in order of magnitude from most significant to most negligible, the supporting factors for stunting are: 1. Baby's weight at birth. 2. Baby's Height at Birth. 3. Number of meal per day. 4. Breast Milk. 5. Diarrhe times per 3 month. 6. Child development examination during covid by Health Worker at home. 7. Mother's age at birth. 8. Mother height at birth. 9. Number of sibling. 10. Age when the first food was given. This research has the disadvantage of no test on other datasets. So that researchers do not consider the reliability of findings in different datasets

BIBLIOGRAPHY

- Abid, Namra et al. 2022. "Synthesis of Nanomaterials Using Various Top-down and Bottom-up Approaches, Influencing Factors, Advantages, and Disadvantages: A Review." *Advances in Colloid and Interface Science* 300: 102597.
- Aminizadeh, Sarina et al. 2023. "The Applications of Machine Learning Techniques in Medical Data Processing Based on Distributed Computing and the Internet of Things." *Computer Methods and Programs in Biomedicine*: 107745.
- Berhe, Kidanemaryam et al. 2019. "Risk Factors of Stunting (Chronic Undernutrition) of Children Aged 6 to 24 Months in Mekelle City, Tigray Region, North Ethiopia: An Unmatched Case-Control Study." *Plos one* 14(6): e0217736.
- Bierman, Joel, Yingzhou Li, and Jianfeng Lu. 2023. "Improving the Accuracy of Variational Quantum Eigensolvers with Fewer Qubits Using Orbital Optimization." *Journal of Chemical Theory and Computation* 19(3): 790–98.
- Dogan, Neslihan, and Zuhul Tanrikulu. 2013. "A Comparative Analysis of Classification Algorithms in Data Mining for Accuracy, Speed and Robustness." *Information Technology and Management* 14: 105–24.
- Fallah, M, T Karami, and K M Kojouri. 2023. "THE ART OF HUMAN PARTICIPATION AGAINST AERTIFICAL INTELLIGENCE WITH GEOSPATIAL INFORMATION SYSTEM IN THE TOURISM INDUSTRY." *GeoJournal of Tourism and Geosites* 49(3): 919–33.
- Jansen, Justin J P et al. 2023. "Scaling-up: Building, Leading and Sustaining Rapid Growth Over Time." *Journal of Management Studies* 60(3): 581–604.
- Kotthoff, Lars. 2016. "Algorithm Selection for Combinatorial Search Problems: A Survey." *Data mining and constraint programming: Foundations of a cross-disciplinary approach*: 149–90.
- Lee, June-Goo et al. 2017. "Deep Learning in Medical Imaging: General Overview." *Korean journal of radiology* 18(4): 570–84.
- Mishra, Shashvi, and Amit Kumar Tyagi. 2022. "The Role of Machine Learning Techniques in Internet of Things-Based Cloud Applications." *Artificial intelligence-based internet of things systems*: 105–35.
- Ndagijimana, Similien, Ignace Habimana Kabano, Emmanuel Masabo, and Jean Marie Ntaganda. 2023. "Prediction of Stunting Among Under-5 Children in Rwanda Using Machine Learning Techniques." *Journal of Preventive Medicine and Public Health* 56(1): 41.
- Salachoris, Georgios Panagiotis et al. 2023. "Evolutionary Numerical Model for Cultural Heritage Structures via Genetic Algorithms: A Case Study in Central Italy." *Bulletin of Earthquake Engineering*: 1–35.
- Simatupang, Mei Yati, Syawal Gultom, and Arif Rahman. 2023. "Midwife Competency Training Management Models in Health Promotion in Stunting-Based Prevention Collaborative Innovative Participation (Pilar) in West Nias District."

International Journal of Advanced Multidisciplinary 2(2): 359–68.

Sohrabi, Arvin, Nima Asgari, Muhammad Imran, and Muhammad Wakil Shahzad. 2023. “Comparative Energy, Exergy, Economic, and Environmental (4E) Analysis and Optimization of Two High-Temperature Kalina Cycles Integrated with Thermoelectric Generators for Waste Heat Recovery from a Diesel Engine.” *Energy Conversion and Management* 291: 117320.

Sultana, Azrin, and Rakibul Islam. 2023. “Machine Learning Framework with Feature Selection Approaches for Thyroid Disease Classification and Associated Risk Factors Identification.” *Journal of Electrical Systems and Information Technology* 10(1): 1–23.

Vaivada, Tyler et al. 2020. “Stunting in Childhood: An Overview of Global Burden, Trends, Determinants, and Drivers of Decline.” *The American journal of clinical nutrition* 112(Supplement_2): 777S-791S.

Zhang, Kaihua et al. 2023. “Data Processing for Simultaneous Inversion of Emissivity and Temperature Using Improved CABCSMA and Target-to-Best DE Algorithms in Multispectral Radiation Thermometry (MRT).” *Optics Express* 31(20): 32684–703.

Copyright holder:

Sutarmi, Warijan Warijan, Tavip Indrayana, Dwi P. Putro b, Indra Gunawan (2023)

First publication right:

Jurnal Health Sains

This article is licensed under:

