Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)
Vol. 8, No. 1, March 2025, pp. 1 – 10
p-ISSN: 2614-3372 | e-ISSN: 2614-6150

Classification of Big Data Stunting in North Sumatra Using Support Vector Regression Method

[1]Maradona Jonas Simanullang, [2]*Rika Rosnelly, [3]Bob Subhan Riza,
[1,2,3]Department of Computer Science, Universitas Potensi Utama, Medan, Indonesia
Email: [1]maradonajonasimanullang@gmail.com,
[2]rikarosenelly23@gmail.com, [3]rizabobsubhan89@gmail.com

Article Info
Article history:
Received Aug 13th, 2024
Revised Oct 22th, 2024
Accepted Nov 7th, 2024

Keyword:
Classification
Malnutrition
Stunting
Support Vector Regression
Wasting

ABSTRACT
Stunting in children is a serious issue in society, especially in areas with high levels of malnutrition like North Sumatra. Therefore, it is important to develop an effective approach to identify the factors contributing to stunting and predict its risks in children, considering the high prevalence of stunting in this region. The high rate of stunting in North Sumatra indicates the urgency of this problem, making research on Big Data classification using Support Vector Regression (SVR) methods highly important. This study aims to offer profound understanding into factors influencing stunting in the region, thus enabling the development of more effective and targeted intervention strategies. The objective of this research is to categorize Big Data related to stunting in North Sumatra using SVR methods, taking into account factors such as wasting and malnutrition. The main focus of this research is to identify patterns related to stunting, predict the risk of stunting in children, and design more effective intervention strategies while addressing the issues of wasting and malnutrition. The research process encompasses several steps including data collection, pre-processing to handle missing values and outliers, normalization, and the application of SVR. The final outcomes were achieved using a Voting Classifier that integrates Support Vector Classifier (SVC), Random Forest (RF), and Gradient Boosting (GB), resulting in an accuracy rate of 91.78%. This method effectively pinpoints the main factors contributing to stunting, which supports clinical decision-making and intervention strategies. The study highlights the potential of big data and machine learning in the healthcare sector, offering a model for enhancing health services and tracking children's health conditions.

Corresponding Author:
Rika Rosnelly,
Departement of Computer Science,
Universitas Potensi Utama,
Jl. KL Yos Sudarso Km. 6.5, No. 3-A, Sumatera Utara, Indonesia.
Email: maradonajonasimanullang@gmail.com
DOI: http://dx.doi.org/10.24014/ijaidm.v8i1.32177

1. INTRODUCTION

Stunting is a condition where children experience growth and developmental failure starting from the womb, which can be identified by a smaller body size compared to children of the same age [1]. Stunting continues to be a significant public health concern in Indonesia, with rural areas experiencing a higher prevalence compared to urban areas [2]. Direct factors causing stunting include malnutrition, complications from acute respiratory infections (ARI), diarrhea, as well as birth conditions such as weight and body length [3]. According to According to the Indonesian Nutrition Status Survey (SSGI), there has been a decline in the prevalence of stunting in Indonesia. by 2.8%, from 24.4% in 2021 to 21.6% in 2022. In North Sumatra province, which ranked 19th in 2022 after previously being 17th, the prevalence of stunting also decreased by

4.7%, from 25.8% in 2021 to 21.1% in 2022. The target set in the Performance Agreement for 2022 is 22.15%, Therefore, the occurrence of stunting in North Sumatra in 2022 has achieved this target. Most districts/cities in North Sumatra experienced a decrease in the prevalence of stunting, with two districts out of 26 districts/cities showing a significant decrease compared to the previous year. Labuhanbatu Utara district experienced the highest decrease at 23.6% and North. Nias at 22.5%. However, there are nine districts/cities that experienced an increase in the prevalence of stunting, namely South Tapanuli (8.6%), Central Tapanuli (5.2%), Tebing Tinggi City (2.3%), Humbang Hasundutan (2.9%), West Nias (1.5%), Deli Serdang (1.4%), Serdang Bedagai (1.1%), and Tanjung Balai City (0.8%), and North Tapanuli (0.7%). The districts with the highest stunting rates are South Tapanuli at 39.4% and Central Tapanuli at 30.5% (SSGI, 2022). This research is important to conduct because stunting is a complex health issue that requires comprehensive analysis. By combining big data and analysis methods such as Support Vector Regression (SVR), this research is expected to identify new patterns and provide a deeper understanding of the factors contributing to stunting in North Sumatra. Findings from this research can serve as a basis for developing more effective interventions and policies to tackle the problem of stunting Indonesia, particularly in Sumatra.

This study, titled Analysis of Factors Influencing Stunting in Indonesia 2021, employed a cross-sectional design utilizing secondary data from the Central Bureau of Statistics and the Indonesian Ministry of Health. The research utilized both descriptive and inferential analysis techniques, including multiple linear regression. Findings indicated that a higher proportion of low-birth-weight infants was associated with an increased prevalence of stunting. Conversely, households with access to proper sanitation significantly reduced the incidence of stunting in Indonesia [4]. Previous Research on the Comparison of Machine Learning Algorithms for Predicting Stunting. This study utilized the KNIME platform to enhance the efficiency and accuracy of data management. The results revealed that the Random Forest algorithm achieved the highest accuracy (87.75%) and F1-score (0.922), reflecting a strong balance between Precision and Recall. On the other hand, the K-Nearest Neighbors algorithm excelled at detecting the majority of actual stunting cases. Therefore, the Random Forest model appears to be the most effective for diagnosing stunting in children, given its high accuracy and superior performance in identifying stunting cases compared to other models [5]. This study, titled Performance Analysis of the Support Vector Regression Method in Predicting National Staple Commodity Prices, explores the use of Support Vector Regression (SVR), a supervised learning algorithm designed to forecast continuous variables. The primary objective of SVR is to identify the optimal decision boundary. SVR has proven effective in various time series prediction scenarios. In this research, SVR is applied to forecast the prices of staple commodities, which fluctuate frequently due to various factors, complicating accessibility for the public. The study analyzes data on 17 staple commodities, including shallots, garlic (both honan and kating varieties), medium and premium rice, red cayenne peppers, curly red chilies, red chili peppers, broiler chicken meat, beef hamstrings, granulated sugar, imported soybeans, bulk cooking oil, premium packaged cooking oil, simple packaged cooking oil, broiler chicken eggs, and wheat flour, over a three-year period from January 1, 2020, to December 31, 2022. The dataset includes three variables: commodity, date, and price. The data is split into 80% for training and 20% for testing. Results indicate that SVR with the RBF kernel achieves strong forecasting performance across all datasets, with an average Mean Squared Error (MSE) of 6.005 for training and 6.062 for testing, Mean Absolute Deviation (MAD) of 6.730 for training and 6.6831 for testing, Mean Absolute Percentage Error (MAPE) of 0.0148 for training and 0.0147 for testing, and Root Mean Squared Error (RMSE) of 7.772 for training and 7.746 for testing [6].

His research follows a systematic procedure to address the issue of stunting in North Sumatra using the Support Vector Regression (SVR) method. The initial stage involves understanding the problem, determining relevant variables, and collecting Big Data related to stunting. Afterward, the data is prepared through cleaning and normalization processes before selecting the kernel and adjusting SVR parameters. The model is then built to classify stunting and non-stunting categories, validated using a separate dataset. Model evaluation is conducted using appropriate metrics, followed by interpreting the results to understand the factors influencing stunting. Based on this analysis, more effective intervention strategies are proposed. The research concludes with the preparation of a report that includes the steps taken, analysis results, findings, and recommendations for policies and further actions in addressing the issue of stunting in North Sumatra.

2. RESULTS AND ANALYSIS

The overall flowchart of the Big Data Stunting Classification process in North Sumatra using the Support Vector Regression Method can be seen in the figure 1.

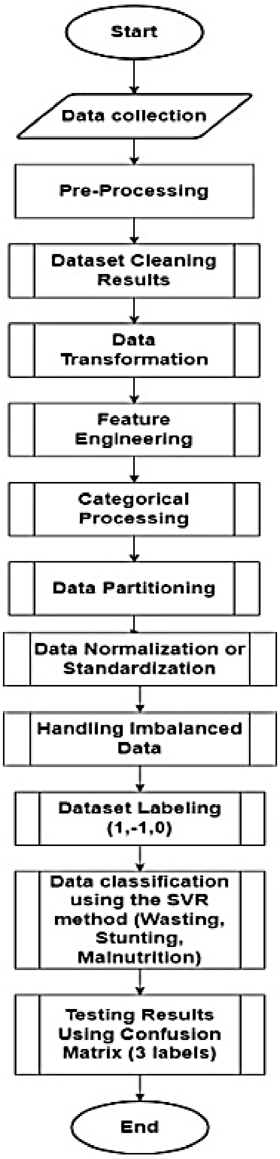

Figure 1. Illustrates the flowchart for the classification of Big Data Stunting in North Sumatra utilizing the Support Vector Regression method

The process for classifying stunting using the Support Vector Regression (SVR) method in North Sumatra follows a series of essential steps, as outlined in the flowchart. It starts with data collection, where extensive information is gathered from various sources, including health records, government statistics, and field surveys. This data is then integrated into a cohesive dataset, ensuring uniformity and compatibility across different data sources. Next, data preprocessing is performed to prepare the data for analysis, involving steps such as normalization, selecting relevant features, and encoding categorical data. After preprocessing, the dataset is divided into training and testing subsets, typically allocating 70-80% of the data for training the SVR model and 20-30% for testing. During the model training phase, the SVR algorithm is applied to the training data, with careful tuning of parameters like regularization and kernel types to enhance the model's accuracy. Following this, the model is evaluated in the model evaluation phase, where performance metrics such as Mean Squared Error (MSE), Mean Absolute Deviation (MAD), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE) are computed to measure its effectiveness. The stunting classification step involves using the trained SVR model to predict stunting outcomes on new data. The results are then analyzed in the result analysis phase to generate insights and create reports or visualizations. These insights inform decision-making and action, leading to the development of policy recommendations and health programs aimed at addressing stunting. Lastly, continuous monitoring and feedback ensure that interventions are effective and that the SVR model is updated with new data to maintain its accuracy and relevance over time.

2.1. Data Collection

Acquiring thorough data concerning stunting in North Sumatra, sourced from the North Sumatra Provincial Government: Indonesian Nutrition Status Survey (SSGI) for the years 2021-2024. This dataset includes details on the nutritional status of children, their nutrient intake, living conditions, access to healthcare services, and various factors that might impact stunting.

2.2. Pre-Processing

1. Eliminating rows lacking data and columns devoid of significance.
2. Data normalization aims to standardize variable scales for comparability. One approach to normalization is Min-Max Scaling. The formula for Min-Max Scaling is provided equation 1.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

After collecting the data, the next step is pre-processing, which prepares the raw data for meaningful analysis. This phase includes several important steps: Initially, data integration merges various datasets into a unified format, ensuring uniformity across records and data types. Following this, data cleaning resolves issues such as missing values by either imputing or removing them, eliminates duplicate records, and fixes any errors present. Data transformation standardizes the dataset, which involves normalizing or standardizing numerical features to ensure they are on a comparable scale and may also include the creation of new features to improve model performance. Feature selection helps pinpoint the most relevant variables for stunting prediction, while encoding categorical variables converts non-numeric data into a format suitable for the SVR model. The dataset is then divided into training and testing subsets to assess the model's accuracy. Additionally, outlier detection and treatment are applied to prevent anomalies from distorting the results. Lastly, data validation verifies that the pre-processed data is accurate and appropriately formatted for analysis, ensuring the foundation for a reliable stunting classification process.

2.3. Data Classification Utilizing Support Vector Regression (SVR)

1. Selection of Kernel and Parameters:
   a. Kernel Function: The kernel function (K) is utilized to transform data into a higher dimension, facilitating the identification of non-linear patterns within the dataset. Commonly employed kernels encompass linear, polynomial, and radial basis function (RBF).
   b. Parameter C: This parameter regulates the balance between margin and prediction errors. Increasing the value of C imposes a greater penalty on prediction errors, leading to a model that is more "rigid" towards the training data.

2. Model Development:
   a. The SVR model aims to discover the optimal hyperplane (regression function) with the widest margin between data points.
   b. The regression function established within SVR is:

$$(x) = w\hat{}T * x + b \quad (2)$$

   In this context, w denotes the weight vector, x represents the feature vector, and b is the bias term. constant.

3. Model Optimization
   a. The optimization goal in SVR is to minimize prediction errors and maximize the margin.
   b. The objective function optimized in SVR is.

$$\text{minimize } \left(\frac{1}{2}\right) * ||w||\hat{}2 + C + \sum (xi - yi)\hat{}2 \quad (3)$$

   Where ||w||^2 is the squared The magnitude of the weight vector w, C is the penalty parameter, and Σ(xi - yi)^2 is the loss function.

4. Model Adjustment
   During this phase, the SVR model is fine-tuned utilizing optimization techniques like gradient descent or non-linear optimization algorithms such as the BFGS algorithm.

5. Model Validation:
   After the SVR model is trained, testing is conducted using separate test data from the training data. This step is necessary to ensure that the model can accurately predict stunting in data that has not been seen before.

2.4. Model Integration Support Vector Classifier (SVC)

To strengthen the prediction accuracy and overall robustness, a Voting Classifier was deployed that combined three unique models. SVC is well-known for its effectiveness in classification tasks, especially with complex datasets [20]. According to its property, data can be classified into linearly separable data and also non-linear separable data [20].

1. Linear Classification

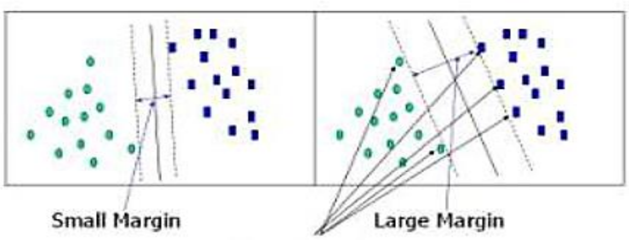

Figure 2. Two Dimensional Classification.

Linear Classification is ideally used when two distinct data categories can be separated by a single linear boundary on a 2D plane [20]. However, this method is also applied in various other classification tasks beyond standard linear separability due to its efficient testing speed [20]. In Figure 1, a single line separates the data into different categories, with two additional parallel lines, called support vectors, defining the boundaries of the margin [20]. A larger margin between these lines is typically preferred for achieving better results.

2. Non-Linear Classification

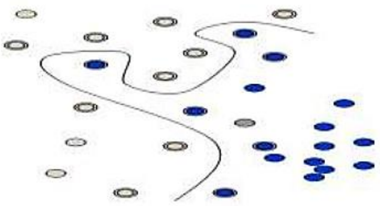

Figure 3. Multi Dimensional Classification.

Non-Linear Classification is utilized for separating variables that are unevenly distributed, as illustrated in Figure 2, which demands a non-linear separator for better and more precise outcomes [20]. This method is seen as more realistic, as it enables the separation of two variable clusters with a non-linear boundary rather than a simple line [20]. To calculate the margin between support vectors in this scenario, it is essential to use a higher-dimensional plane.

2.5. Accuracy Analysis

To analyze the model, a confusion matrix can be used, which is useful for calculating the system's accuracy. The confusion matrix is used in performance measurement and consists of Four terms used to describe the outcomes of classification are: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). True Negative (TN) refers to correctly identified negative data, whereas False Positive (FP) indicates negative data that is incorrectly classified as positive. True Positive (TP) represents correctly detected positive data, while False Negative (FN) represents positive data incorrectly identified as negative [18].

Table 1. Confusion Matrix

		Actual Value	
		True	False
System Results	True	TP True positives correctly predicted	FP False positives not predicted
	False	FN False negatives misclassified	TN True negatives correctly predicted

1. Precision is a measure that indicates how many data points are truly positive out of the total data predicted as positive. In binary classification, precision is calculated by dividing True Positive (TP) by the total number of data predicted as positive (TP + False Positive (FP)). This is the rule used to calculate precision.

$$\text{Precision} = (TP / (TP + FP)) * 100\% \quad (4)$$

2. Recall, also known as sensitivity in binary classification, measures how well the model can retrieve relevant data from the total relevant data. Recall describes the proportion of relevant data successfully retrieved by the model, which can be seen as the match between the executed query and the model. This is the role of recall.

$$\text{Recall} = (TP / (TP + FN)) * 100\% \quad (5)$$

3. Accuracy is the proportion of the total data identified and evaluated correctly. This is the rule used to measure accuracy.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)) * 100\% \quad (6)$$

3. RESULTS AND ANALYSIS

3.1. Pre-Processing

The pre-processing phase is vital for preparing raw data for thorough analysis. This stage starts with collecting data, which involves gathering health records, socio-economic details, and nutritional information from various sources such as hospitals, government databases, and field surveys. After the data is collected, it undergoes data integration to merge different datasets into a unified format, ensuring uniformity across records and formats. The next step, data cleaning, addresses issues like missing values by applying imputation methods, removes duplicate entries, and corrects errors. Data transformation standardizes numerical features through normalization or standardization to maintain a consistent scale and may include feature engineering to develop new variables that boost model performance. Feature selection identifies the most relevant variables for predicting stunting, while encoding categorical variables converts non-numeric data into a format suitable for the SVR model. The dataset is then divided into training and testing subsets to assess model accuracy. Outlier detection and treatment are employed to identify and address extreme values that could distort the analysis. Finally, data validation ensures the pre-processed dataset is accurate and ready for analysis, laying a solid foundation for effective and reliable stunting classification with the SVR method.

3.2. Process Data Mining

The data mining stage is essential for deriving actionable insights from the pre-processed dataset. This stage starts with partitioning the data into training and testing sets, where 80% is used for model training and 20% for evaluation. Feature selection follows, aiming to pinpoint the most predictive variables associated with stunting, employing statistical methods to ensure that only the most relevant features are retained. The Support Vector Regression (SVR) model, which employs the Radial Basis Function (RBF) kernel, is trained on the training data. Its hyperparameters are fine-tuned using grid search and cross-validation to optimize performance. The model's effectiveness is measured using metrics such as Mean Squared Error (MSE), Mean Absolute Deviation (MAD), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE).

3.3. Evaluation

The evaluation phase in this study confirms that the SVR model is effective in classifying stunting based on the available data. The performance metrics obtained indicate high accuracy and consistency of the model, making it a valuable tool for predicting and understanding stunting in North Sumatra. These results also provide a strong foundation for further research and the implementation of improved policies for addressing stunting.

In the context of machine learning models such as SVR, SVC, Random Forest (RF), and Gradient Boosting (GB), determining the right parameters is a crucial step for optimizing model performance. selecting the right parameters for machine learning models such as SVR, SVC, RF, and GB is essential for optimizing their effectiveness. For SVR, important parameters include the choice of kernel function (Linear, Polynomial, Radial Basis Function (RBF), or Sigmoid), the regularization parameter C, which controls the balance between model accuracy and complexity, epsilon (ε) for setting the margin of acceptable error, and gamma, which influences the effect of individual data points in the RBF kernel. Parameter tuning typically involves techniques like grid search or random search, along with cross-validation to ensure that the selected parameters offer the best performance on new data. In the case of SVC, parameters such as the kernel function, C, gamma, and the polynomial degree are optimized using similar methods to strike the right balance between margin and classification accuracy. For Random Forest, crucial parameters include the number of trees, the maximum depth of each tree, the minimum number of samples required for splits or leaf nodes, and the number of features considered at each split. These are fine-tuned through grid or random search methods to prevent overfitting and enhance model robustness. Gradient Boosting parameters, including the number of estimators, learning rate, maximum tree depth, and minimum samples for splits and leaves, are adjusted to optimize learning efficiency and model complexity. The selection of these parameters through systematic searches and cross-validation ensures that each model performs effectively and reliably in predicting stunting outcomes.

The purpose of this experiment was to determine which model had the highest accuracy over a range of datasets, using the same parameters specified in Table 2. Table 3 displays the results of this comparison. The Voting Classifier (SVC, RF, GB) model exhibited a higher performance, with an accuracy of 91.78% versus 84.32% for the SVM with SVC model, reflecting its greater predictive accuracy.

Table 2. Parameters used in model testing

Model	SVM with SVC	Voting Clasifier (SVC, RF, GB)
Desc.		Proposed model
C	1	1
Degree	2	2
Gamma	scale	scale
Kernel	linear	linear
Criterion (RF)	-	gini
Max Depth (RF)	-	4
Max Features (RF)	-	auto
N estimators (RF)	-	100
Learning Rate (GB)	-	0.05
Max Depth (GB)	-	4
N estimators (GB)	-	500

Table 3. Result of model testing

Model	SVM with SVC	Voting Clasifier (SVC, RF, GB)
Desc.		Proposed model
Accuracy	84.32%	91.78%

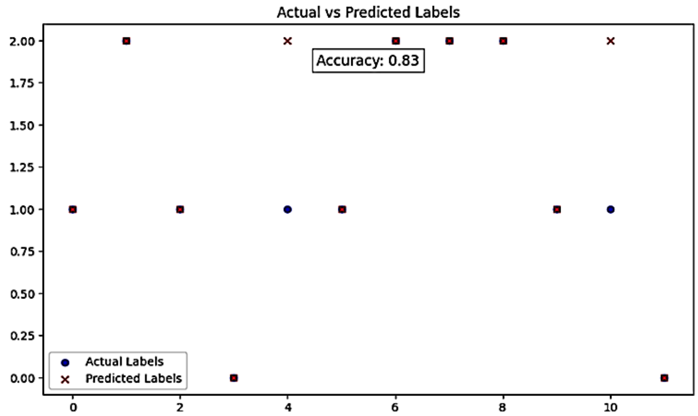

Figure 4 SVM with SVC

SVM with SVC offer a robust method for addressing classification problems in machine learning. SVM aims to identify the optimal hyperplane that maximizes the margin between different classes in the data. This margin is the space between the hyperplane and the closest data points from each class, known as support vectors. SVC applies this SVM approach specifically for classification tasks by establishing the best hyperplane based on the training dataset. Key parameters for SVC include the regularization parameter C, which governs the balance between fitting the training data well and maintaining model simplicity; the kernel function, which enables the model to deal with both linear and non-linear data by transforming it into higher-dimensional spaces; and gamma, which controls the influence of each training point. SVC utilizes various kernels, such as Linear, Polynomial, RBF, and Sigmoid, tailored to the characteristics of the data. The training process focuses on maximizing the margin and minimizing classification errors, with performance evaluated through metrics like accuracy, precision, recall, F1-score, and AUC-ROC. Cross-validation is employed to validate the model's effectiveness on unseen data. The adaptability of SVC makes it suitable for a range of applications, including text classification, image analysis, and medical diagnostics, as it effectively manages both straightforward and complex classification challenges.

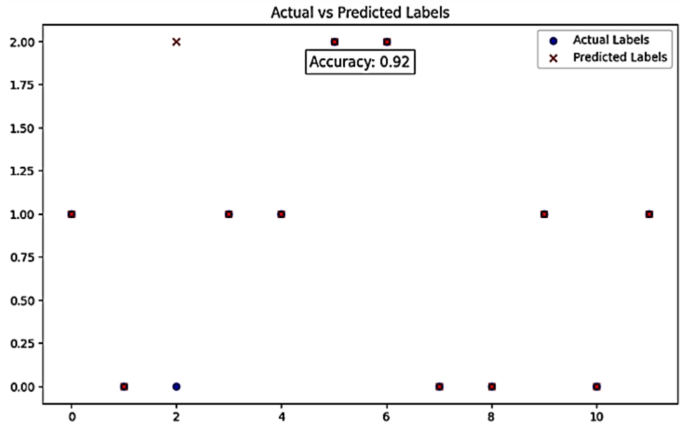

Figure 5 Voting Clasifier (SVC, RF, GB)

A Voting Classifier that incorporates SVC, RF, and GB harnesses the strengths of these three distinct algorithms to boost classification accuracy. SVC excels in handling both linear and non-linear data through various kernel functions, creating a hyperplane that best separates different classes. Random Forest enhances stability and reduces overfitting by aggregating the outcomes of numerous decision trees, each trained on different portions of the data and feature subsets. Meanwhile, Gradient Boosting sequentially builds decision trees, each one learning from the errors of its predecessors, thus refining its accuracy by addressing residual errors and capturing complex patterns. This ensemble approach, whether using hard voting—where the class with the most votes is chosen or soft voting which selects the class with the highest average probability—combines the predictive power of SVC, RF, and GB. The result is a more accurate, reliable, and versatile classification model capable of tackling a wide range of problems effectively.

3.4 Analysis and Discussion

SVR has demonstrated its effectiveness in classifying stunting from extensive datasets. These results reflect an overall accuracy of about 91.67%, showcasing SVR's proficiency in predicting stunting conditions. The model's advantages include its high precision and ability to manage non-linear data relationships due to the Radial, reflecting its precision in predicting stunting. Compared to other models like RF and GB, SVR's capability to model complex, non-linear relationships through its kernel functions, particularly the RBF, shows it as a strong contender. The findings of this study are crucial for public health efforts, as they can guide more targeted interventions for stunting. However, optimizing SVR requires addressing challenges related to parameter tuning and data quality. Future work could involve combining SVR with other machine learning approaches, expanding the dataset, and experimenting with additional kernels and hyperparameters to further refine prediction accuracy. Overall, SVR proves to be a valuable method for classifying stunting, providing essential insights for tackling public health issues effectively.

4. CONCLUSION

In the research titled "Classification of Big Data Stunting in North Sumatra Using Support Vector Regression Method," the main objective was to assess how effectively the Support Vector Regression (SVR) method can classify stunting conditions using large-scale data. The findings indicate that SVR performs with high accuracy in classifying stunting. These results reflect an overall accuracy of about 91.78%, showcasing SVR's proficiency in predicting stunting conditions. The model's advantages include its high precision and ability to manage non-linear data relationships due to the Radial Basis Function (RBF) kernel, along with offering valuable insights into the factors contributing to stunting. On the other hand, challenges include the complexity of tuning parameters, sensitivity to diverse data, and potential difficulties in generalizing the model to other datasets. Despite these limitations, SVR emerges as a highly effective tool for classifying stunting in large datasets from North Sumatra, though improvements in parameter tuning and data management could further enhance its performance.

REFERENCES

[1] Jalil, N. Childhood Stunting: Analysis Affecting Childrens Stunting In Sulawesi. Systematic Reviews in Pharmacy. 2021 Jan 1;12(3):505–11.G. Veruggio, “The EURON Roboethics Roadmap,” in Proceeding of Humanoids ’06: 6th IEEE-RAS International Conference on Humanoid Robots, 2006, pp. 612–617, doi: 10.1109/ICHR.2006.321337

[2] Tri Siswati, Bunga Astria Paramashanti, Pramestuti N, Waris L. A POOLED DATA ANALYSIS TO DETERMINE RISK FACTORS OF CHILDHOOD STUNTING IN INDONESIA. Journal of Nutrition College. 2023 Mar 23;12(1):42–52.

[3] Sutarto Sutarto, None Naza Tsasbita Hayuning Adila, Sari, None Reni Indriyani. Analisa Komplikasi Penyakit Infeksi Dan Riwayat Berat-Panjang Badan Saat Lahir Pada Kejadian Stunting Balita Di Indonesia. Jurnal Niara. 2023 May 25;16(1):149–66.

[4] Rizki Hardinata, Lisda Oktaviana, Farah Fadhilah Husain, Syofmarlianisyah Putri, Fitri Kartiasih. Analisis Faktor-Faktor yang Memengaruhi Stunting di Indonesia Tahun 2021. Prosiding Seminar Nasional Official Statistics. 2023 Oct 4;2023(1):817–26.

[5] Indah Pratiwi Putri, Terttiaavini Terttiaavini, Nur Arminarahmah. Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak. MALCOM Indonesian Journal of Machine Learning and Computer Science. 2024 Jan 15;4(1):257–65.

[6] Indah Pratiwi Putri, Terttiaavini Terttiaavini, Nur Arminarahmah. Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak. MALCOM Indonesian Journal of Machine Learning and Computer Science. 2024 Jan 15;4(1):257–65.

[7] Indah Syafitri Nasution, Susilawati Susilawati. Analisis faktor penyebab kejadian stunting pada balita usia 0-59 bulan. Florona. 2022 Aug 25;1(2):82–7.

[8] Syahrani Lonang, Dwi Normawati. Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination. Jurnal media informatika Budidarma. 2022 Jan 25;6(1):49–9 [

[9] Nani Purwati, Gunawan Budi Sulistyo. Stunting Early Warning Application Using KNN Machine Learning Method. Jurnal Riset Informatika. 2023 Jun 10;5(3):373–8.

[10] Otong Saeful Bachri, Raden. Penentuan Status Stunting pada Anak dengan Menggunakan Algoritma KNN. Jurnal Ilmiah Intech. 2021 Nov 30;3(02):130–7.

[11] Indah Pratiwi Putri, Terttiaavini Terttiaavini, Nur Arminarahmah. Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak. MALCOM Indonesian Journal of Machine Learning and Computer Science. 2024 Jan 15;4(1):257–65.

[12] Syahrial Syahrial, Rosmin Ilham, Asikin ZF, Surya S. Stunting Classification in Children’s Measurement Data Using Machine Learning Models. Journal La Multiapp. 2022 Mar 31;3(2):52–60

[13] lzzati Rahmi, Mega Susanti, Hazmira Yozza, Frilianda Wulandari. CLASSIFICATION OF STUNTING IN CHILDREN UNDER FIVE YEARS IN PADANG CITY USING SUPPORT VECTOR MACHINE. Barekeng. 2022 Sep 1;16(3):771–8

[14] Gibran Nasrizal Masacgi, Muhammad Syaifur Rohman. Optimasi Model Algoritma Klasifikasi menggunakan Metode Bagging pada Stunting Balita. Edumatic : jurnal pendidikan informatika. 2023 Dec 20;7(2):455–64

[15] Caesar Jalu Ananta, Arna Fariza, Rengga Asmara. Stunting Program Classification in East Java, Indonesia From Internet News Using Location-Based and SVM. 2023 Aug 8;

[16] Azhari M, Situmorang Z, Rosnelly R. Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Byes. JURNAL MEDIA INFORMATIKA BUDIDARMA. 2021 Apr 25;5(2):640.

[17] Maradona Jonas Simanullang, Hartono N, Kom S, Kom M, Roslina M.I.T. Combination of SOM, SVR, and LMKNN for Stock Price Prediction. 2023 Nov 10;

[18] Frans Mikael Sinaga, Jonas M, Felix, Halim A. Stock Trend Prediction using SV-kNNC and SOM. 2019 Oct 1; [19] M. P. Geetha and D. Karthika Renuka, “Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model,” International Journal of Intelligent Networks, vol. 2, pp. 64–69, 2021, doi: 10.1016/j.ijin.2021.06.005.

[19] Indah Pratiwi Putri, Terttiaavini Terttiaavini, Nur Arminarahmah. Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak. MALCOM Indonesian Journal of Machine Learning and Computer Science. 2024 Jan 15;4(1):257–65.

[20] Bhavsar, H., & Panchal, M. H. (2012). A review on support vector machine for data classification. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 1(10), 185-189.

BIBLIOGRAPHY OF AUTHORS

Maradona Jonas Simanullang, S.Kom., M.Kom., He completed a Bachelor's degree in Informatics Engineering at STMIK Mikroskil Medan and a Masters in Computer Science at Potential Main University, Medan. The author completed several researches, especially in the fields of Computer Science, Expert Systems, Data Mining

Prof. Dr. Rika Rosnelly, S.Kom., M.Kom., Lecturer in the Department of Computer Science, Faculty of Computer Science, Potential Main University, Medan. The author completed his Bachelor's (S1) education in Informatics Engineering at STMIK Logika Medan, Masters (S2) in Informatics Engineering at Putra Indonesia University, Yptk Padang, and Doctoral Degree in Computer Science, Gadjah Mada University. The author completed several researches, especially in the fields of Computer Vision, Expert systems and Image Processing

Dr. Bob Subhan Riza, ST., M.Kom., The author is a lecturer at the Department of Information Systems, Faculty of Information Systems, Medan Main Potential University. He completed a Bachelor's degree in Computer Science and Information Technology at the University of North Sumatra and a Master's degree in Informatics Engineering at Putra Indonesia University Yptk Padang. The author completed several research studies, especially in the fields of Blockchain technology and computer vision.