



**UNDERGRADUATE THESIS**

**PERFORMANCE ANALYSIS OF THE KNN AND  
SVM ALGORITHMS IN STUNTING  
CLASSIFICATION BASED ON  
ANTHROPOMETRIC STANDARDS**

**SALMA FATHIYATUR RIZKY MUNIR**  
NPM 20081010025

**THESIS ADVISORS**

Yisti Vita Via, S.ST., M.Kom.  
Eka Prakarsa Mandyartha, S.T., M.Kom

**MINISTRY OF HIGHER EDUCATION, SCIENCE, AND TECHNOLOGY  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FACULTY OF COMPUTER SCIENCE  
INFORMATIC'S STUDY PROGRAM  
SURABAYA  
2026**

## APPROVAL SHEET

### PERFORMANCE ANALYSIS OF THE KNN AND SVM ALGORITHMS IN STUNTING CLASSIFICATION BASED ON ANTHROPOMETRIC STANDARDS

By:  
SALMA FATHIYATUR RIZKY MUNIR  
NPM. 20081010025

Has been defended before, and accepted by, the Board of Assessors of the Thesis Examination of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, on May 21, 2026:

Approved

Yisti Vita Via, S.ST., M.Kom.  
NIP. 19860425 202121 2 001



.....

(Advisor I)

Eka Prakarsa Mandyartha, S.T., M.Kom  
NIP. 19880525 201803 1 001



.....

(Advisor II)

Dr. Intan Yuniar Purbasari, S.Kom. MSc.  
NIP. 19800602 202521 2 029



.....

(Head Assessor)

Eva Yulia Puspaningrum, S.Kom., M.Kom  
NIP. 19890705 202121 2 002

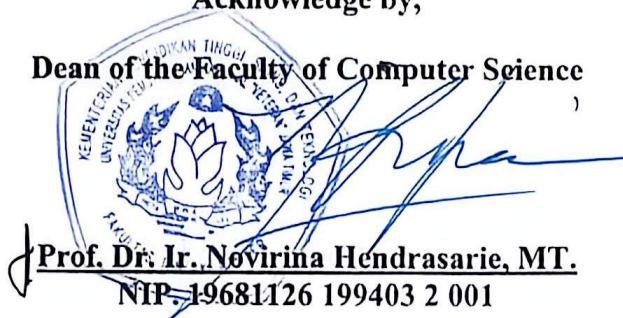


.....

(Assessor I)

Acknowledge by,

Dean of the Faculty of Computer Science



Prof. Dr. Ir. Novirina Hendrasarie, MT.  
NIP. 19681126 199403 2 001

**APPROVAL SHEET**

**PERFORMANCE ANALYSIS OF THE KNN AND SVM ALGORITHMS IN  
STUNTING CLASSIFICATION BASED ON ANTHROPOMETRIC  
STANDARDS**

By:  
SALMA FATHIYATUR RIZKY MUNIR  
NPM. 20081010025

Approved to proceed to the Thesis Examination



Approved by,  
**Coordinator of Informatics Study Program  
Faculty of Computer Science**

**Dr. Intan Yuniar Purbasari, S.Kom. MSc.**  
**NIP. 19800602 202521 2 029**

## STATEMENT OF ORIGINALITY

I am the undersigned:

Student Name : Salma Fathiyatur Rizky Munir

NPM : 20081010025

Degree Program: Bachelor (S1)

Study Program : Informatics

Faculty : Faculty of Computer Science

Hereby declares that this undergraduate thesis contains no part of any other scientific work that has been submitted to obtain an academic degree at any higher education institution. Furthermore, it does not contain any work or opinions previously written or published by others, except for those which are explicitly cited in this thesis and listed completely in references.

And I declare that this scientific document is free from elements of plagiarism. If in the future indications of plagiarism are found in this Thesis, I am willing to accept sanctions in accordance with the applicable laws and regulations.

Thus, I made this statement without any coercion from anyone and to be used as it should.

Surabaya, May 21, 2026

Declarant,



SALMA FATHIYATUR RIZKY MUNIR  
NPM. 20081010025

## ABSTRACT

Student Name / NPM : Salma Fathiyatur Rizky Munir / 20081010025  
Thesis Title : Performance Analysis of the KNN and SVM Algorithms in Stunting Classification Based on Anthropometric Standards  
Supervisor : 1. Yisti Vita Via, S.ST., M.Kom.  
2. Eka Prakarsa Mandyartha, S.T., M.Kom

The purpose of this study is to evaluate how well the KNN and SVM algorithms in toddlers at the Patianrowo Community Health Center. Since stunting is a long-term nutritional issue that can impair children's growth and development, categorization techniques are essential for accurately identifying stunting cases. This study compares the two algorithms performance using evaluation metrics and examines how parameter changes affect classification performance. The Patianrowo Community Health Center provided primary data on newborns between the ages of 0 and 60 months. After preprocessing, 1,067 of the original 1,102 data points were recovered these were then split into 20% test data and 80% training data. Gender, age, height, weight, and Z-score values are among the features that are employed. While the SVM algorithm employed linear, polynomial, and radial basis function (RBF) kernels, the KNN algorithm testing was tested with several K values, specifically 3, 5, 7, 9, and 11. A confusion matrix comprising assessment metrics including accuracy, precision, recall, and F1-score was used to assess the model's performance.

According to the test findings, KNN with  $K = 5$  generated an F1-score of 77.52%, recall of 67.73%, accuracy of 96.72%, and precision of 91.25. With an accuracy of 97.47%, precision of 90.82%, recall of 78.96%, and an F1-score of 82.55%, the polynomial kernel produced the best result for the SVM algorithm. These findings indicate that the SVM technique with a polynomial kernel outperforms the others in stunting classification.

**Kata kunci :** *stunting, klasifikasi, KNN, SVM*

## ACKNOWLEDGEMENTS

All praise be to Allah SWT for His mercy, blessings, and the ease He has granted, enabling the author to successfully complete this thesis titled “Analysis of the Performance of the KNN and SVM Algorithms in Stunting Classification Based on Anthropometric Standards”.

During the process of writing this thesis, the author received a great deal of assistance, support, and prayers from various parties. Therefore, the author would like to take this opportunity to express his deepest gratitude to:

1. Mr. Prof. Dr. Ir. Akhmad Fauzi, MMT. as the rector of the National Development University “Veteran” East Java.
2. Mrs. Prof. Dr. Ir. Novirina Hendrasarie, MT. as the Dean of the Faculty of Computer Science, National Development University “Veteran” East Java.
3. Mrs. Dr. Intan Yuniar Purbasari, S.Kom., M.Sc. as the coordinator of the Informatics Study Program at the Faculty of Computer Science, National Development University “Veteran” East Java, for providing guidance and support throughout the academic program.
4. Mrs. Yisti Vita Via, S.ST., M.Kom., as the primary thesis advisor, who consistently provided guidance, motivation, and feedback to the author throughout the process of writing this thesis.
5. Mr. Eka Prakarsa Mandyartha, S.T., M.Kom., as the advisor who provided guidance, criticism, and suggestions for the refinement of this thesis.
6. All faculty members of the Computer Science Program, Faculty of Computer Science, National Development University “Veteran” East Java, who shared their knowledge and experience throughout the author’s academic journey.
7. My beloved parents, Father and Mother, who have always offered prayers, love, support, and unwavering encouragement to the author throughout every stage of life and education.
8. My younger siblings, who have consistently supported encouragement and support throughout the process of writing this thesis.

9. The author's closest friends Sintya, Ais, Natalia, and Indy who have been by her side, offering encouragement, support, and companionship throughout her college years and while writing her thesis.
10. All those who cannot be named individually but who have helped and supported the author in completing this thesis.

The author acknowledges that this thesis still has shortcomings and is far from perfect. Therefore, constructive criticism and suggestions are greatly appreciated for future improvements. May this thesis be of benefit to readers and contribute to the advancement of knowledge.

Surabaya, May 21 2026

Author

## TABLE OF CONTENTS

<b>APPROVAL SHEET .....</b>	<b>iii</b>
<b>APPROVAL SHEET .....</b>	<b>v</b>
<b>STATEMENT OF ORIGINALITY.....</b>	<b>vii</b>
<b>ABSTRACT .....</b>	<b>ix</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>xi</b>
<b>TABLE OF CONTENTS .....</b>	<b>xiv</b>
<b>LIST OF FIGURES.....</b>	<b>xvii</b>
<b>LIST OF TABLES .....</b>	<b>xx</b>
<b>LIST OF NOTATIONS .....</b>	<b>xxi</b>
<b>CHAPTER I INTRODUCTION.....</b>	<b>1</b>
1.1. Background.....	1
1.2. Research Question .....	3
1.3. Research Objectives.....	3
1.4. Research Benefits .....	3
1.5. Scope of the Study .....	4
<b>CHAPTER II LITERATURE REVIEW.....</b>	<b>5</b>
2.1. Previous Research.....	5
2.2. Stunting.....	7
2.3. Classification .....	8
2.4. Data Normalization.....	8
2.5. K-Nearest Neighbor (KNN) Algorithm .....	9
2.6. Support Vector Machine (SVM) Algorithm.....	12
2.7. Python .....	16
2.8. Pandas .....	16
2.9. NumPy .....	17
2.10. Scikit-learn.....	17
2.11. Confusion Matrix.....	18
<b>CHAPTER III RESEARCH METHODOLOGY.....</b>	<b>20</b>
3.1. Research Stages .....	21
3.1.1. Problem Identification .....	21

3.1.2.	Literature Review.....	22
3.1.3.	Data Collection.....	22
3.1.4.	Data Preprocessing.....	27
3.1.5.	Model Design.....	30
3.1.5.1.	KNN Model Training.....	30
3.1.5.2.	SVM Model Training.....	34
3.1.5.3.	Model Training and Testing.....	40
3.2.	Testing Scenarios.....	43
3.2.1.	Testing K Values in the KNN Algorithm.....	44
3.2.2.	Kernel Testing on the SVM Algorithm.....	44
3.3.	Performance Measurement Methods.....	45
3.4.	System Design.....	47
3.4.1.	Use Case Diagram.....	47
3.4.2.	Activity Diagram.....	48
3.4.3.	Sequence Diagram.....	49
3.4.4.	User Interface Design.....	50
3.4.4.1.	Input Form.....	50
3.4.4.2.	Prediction Result.....	51
<b>CHAPTER IV RESULT AND DISCUSSION .....</b>		<b>52</b>
4.1.	Program Implementation.....	53
4.1.1.	Importing Libraries.....	53
4.1.2.	Dataset Preparation.....	54
4.1.3.	Data Preprocessing.....	55
4.1.4.	Separate Features and Labels.....	57
4.1.5.	Data Split.....	57
4.1.6.	Data Normalization.....	58
4.1.7.	KNN Implementation.....	59
4.1.8.	SVM Implementation.....	65
4.1.9.	Prediction Result.....	70
4.1.10.	The Five Most Influential Factors on Classification Results.....	71
4.2.	KNN Experiment Scenario Results.....	73
4.2.1.	Value of K = 3.....	74
4.2.2.	Value of K = 5.....	75

4.2.3.	Value of K = 7.....	76
4.2.4.	Value of K = 9.....	77
4.2.5.	Value of K = 11.....	78
4.2.6.	KNN Results.....	79
4.3.	SVM Experiment Scenario Results .....	79
4.3.1.	Linear Kernel.....	79
4.3.2.	Polynomial Kernel.....	80
4.3.3.	Radial Basis Function (RBF) Kernel.....	81
4.3.4.	SVM Results.....	82
4.4.	The Result Using Synthetic Data.....	83
4.5.	Comparative Analysis of KNN and SVM.....	91
4.5.1.	Comparative Analysis of Classification Models .....	91
4.5.2.	Class Imbalance.....	92
4.5.3.	Analysis of the Influence of Model Parameters .....	93
4.5.4.	Interpretation of Research Results.....	93
4.5.5.	The Effect of Data Distribution on Model Performance .....	94
4.6.	System Implementation .....	95
4.6.1.	Best Model.....	95
4.6.2.	Model Storage .....	96
4.6.3.	Implementing Flask .....	96
4.6.4.	Website Implementation .....	97
4.6.5.	Website Interface .....	97
<b>CHAPTER V CONCLUSION.....</b>		<b>99</b>
5.1.	Conclusions.....	99
5.2.	Recommendations.....	100
<b>REFERENCES .....</b>		<b>101</b>

## LIST OF FIGURES

Figure 2. 1 Illustration of KNN .....	10
Figure 2. 2 Hyperplane dan Margin.....	12
Figure 3. 1 Research Workflow .....	21
Figure 3. 2 Stunting Data.....	28
Figure 3. 3 Flowchart of the KNN Method .....	31
Figure 3. 4 Flowchart of the SVM Method .....	34
Figure 3. 5 Flowchart of Model Training and Testing.....	40
Figure 3. 6 Use Case Diagram.....	47
Figure 3. 7 Activity Diagram.....	48
Figure 3. 8 Sequence Diagram .....	49
Figure 3. 9 Input Form Wireframe.....	50
Figure 3. 10 Prediction Result Wireframe .....	51
Figure 4. 1 Displaying Result .....	54
Figure 4. 2 Data Count .....	55
Figure 4. 3 Data Statistics.....	56
Figure 4. 4 Result After Encoding .....	57
Figure 4. 5 Data Split Result .....	58
Figure 4. 6 After Normalization .....	59
Figure 4. 7 KNN Result.....	64
Figure 4. 8 SVM Result.....	69
Figure 4. 9 Prediction Result .....	71
Figure 4. 10 5 KNN Features .....	72
Figure 4. 11 5 SVM Features.....	73
Figure 4. 12 Result for K=3.....	74
Figure 4. 13 Confusion Matrix fK=3 .....	74
Figure 4. 14 Result for K=5.....	75
Figure 4. 15 Confusion Matrix K=5 .....	75
Figure 4. 16 Result for K=7.....	76
Figure 4. 17 Confusion Matrix K=7 .....	76
Figure 4. 18 Result for K=9.....	77
Figure 4. 19 Confusion Matrix K=9 .....	77

Figure 4. 20 Result for K=11 .....	78
Figure 4. 21 Confusion Matrix K=11 .....	78
Figure 4. 22 Linear Kernel Result.....	79
Figure 4. 23 Confusion Matrix for the Linear Kernel.....	80
Figure 4. 24 Result of the Polynomial Kernel.....	80
Figure 4. 25 Confusion Matrix for the Polynomial Kernel.....	81
Figure 4. 26 RBF Kernel Results .....	81
Figure 4. 27 Confusion Matrix of the RBF Kernel .....	82
Figure 4. 28 Synthetic Data Results for K=3 .....	83
Figure 4. 29 Confusion Matrix for Synthetic Data K=3 .....	83
Figure 4. 30 Synthetic Data result for K=5 .....	84
Figure 4. 31 Confusion Matrix for Synthetic Data K=5 .....	84
Figure 4. 32 Synthetic Data Result for K=7.....	85
Figure 4. 33 Confusion Matrix for Synthetic Data K=7 .....	85
Figure 4. 34 Synthetic Data Result for K=9.....	85
Figure 4. 35 Confusion Matrix for Synthetic Data K=9 .....	86
Figure 4. 36 Synthetic Data Result for K=11.....	86
Figure 4. 37 Confusion Matrix for Synthetic Data K=11 .....	87
Figure 4. 38 Results of Synthetic Data with Linear Kernel .....	88
Figure 4. 39 Confusion Matrix for Synthetic Data Linear Kernel .....	88
Figure 4. 40 Results of Synthetic Data with Polynomial Kernel .....	89
Figure 4. 41 Confusion Matrix for Synthetic Data Polynomial Kernel .....	89
Figure 4. 42 Results of Synthetic Data with RBF Kernel .....	89
Figure 4. 43 Confusion Matrix for Synthetic Data RBF Kernel .....	90
Figure 4. 44 Stunting Prediction Input Form .....	97
Figure 4. 45 Non-Stunting Result .....	98
Figure 4. 46 Stunting Result.....	98

## LIST OF TABLES

Table 2. 1 Confusion Matrix.....	18
Table 3. 1 Primary Data Split Table.....	22
Table 3. 2 Data Table .....	23
Table 3. 3 Variable Descriptions.....	23
Table 3. 4 Synthetic Data Split Table .....	27
Table 3. 5 Example Data for the KNN Method.....	31
Table 3. 6 Calculating Euclidean Distance .....	32
Table 3. 7 Sorting by Closest Distance.....	33
Table 3. 8 Example of SVM Calculation.....	34
Table 3. 9 Normalization Result.....	35
Table 3. 10 Transformation Result .....	36
Table 3. 11 Output Values (y).....	36
Table 3. 12 KNN Testing Scenarios .....	44
Table 3. 13 SVM Testing Scenarios .....	44
Table 4. 1 KNN Method Testing Results.....	79
Table 4. 2 SVM Method Testing Result .....	82
Table 4. 3 Result of Synthetic Data Testing Using the KNN Method.....	87
Table 4. 4 Result of Synthetic Data Testing Using the SVM Method.....	90

## LIST OF NOTATIONS

$D$	:	Proximity distance between the test data and the training data (Euclidean distance)
$x$	:	Training data vector
$y$	:	Test data vector
$n$	:	Number of attributes in the data
$i$	:	Index of individual attributes
$K$	:	Number of nearest neighbors selected in the KNN method
$x_i$	:	i-th data point
$y_i$	:	Class label of the i-th data point
$w$	:	Weight vector
$b$	:	Bias (relative distance of the hyperplane to the coordinate point)
$E_i$	:	Error function for the i-th data point in SVM
$K(x_i, x_j)$	:	The value of the kernel function between the i-th and j-th data points
$f(x)$	:	The decision function in the SVM prediction process
$TP$	:	True Positive – Positive data correctly classified
$TN$	:	True Negative – Negative data correctly classified
$FP$	:	False Positive – Negative data classified as positive
$FN$	:	False Negative – Positive data classified as negative