

Support Vector Machine (SVM) For Toddler's Nutritional Classification in Palu City

Andi Hendra¹, Gazali²

Abstract—Toddlers are groups who are vulnerable about the health nutrition problems. Nutritional status of children is one of the indicators that can describes the level of social welfare in the city. Nutritionists are the people that can determined the nutritional status. The problem that arises is the limited number of the nutrition experts in each area, this problem causes the children's malnutrition in the Palu city is detected in very slow condition. The aims of this study is to help the health professionals in the health centers or the hospitals to determine the children's nutritional status computerized, so the malnutrition problem in the Palu city can be detected earlier. Besides that, to help the government in policy making about nutrition of the toddlers in Palu city. This study uses a Support Vector Machine (SVM) which implemented in computer-based software application to analyze nutrition of the toddlers.

Keywords—Nutrition, Software, Support Vector Machine (SVM), Toddlers, Palu city.

I. INTRODUCTION

Toddlers are children between one to five years old, or more common said children below five years old (Muaris, 2006). Toddlers are groups who are vulnerable to health problems, and nutrition such as protein energy malnutrition (PEM). because of that reasons toddlers should get a special attention especially about the toddlers' nutrition. Nutritional's status of children is an indicator that reflect the level of social welfare in the area. Measurement of the children nutritional status can be performed by anthropometry which reference to weight index for age, height for age, or weight for height. The standard reference used in performing anthropometric classification by decree of the Minister of Health, Decree Number 1995/MENKES/SK/XII/2012 About the standard of anthropometric assessment of nutritional status of the children, referring to the standard of the World Health Organization (WHO, 2005). The measurement of nutritional status of children are conducted by nutritionists. The problem that arises is the number of nutritionists who placed in the clinic or hospital is limited. The distribution of nutritionists who are placed in health centers and hospitals in the palu city is still not evenly distributed. This is the evident from the data that there are still 25% of health centers and 66% hospitals that do not have a nutritionist. Besides that, the data also shows that the health centers have the average experts equivalent secondary school education and not educated undergraduate (bachelor). As for hospitals, nutritionists distributed uneven therefore there are many

hospitals that do not have a nutritionist while there is a hospital that has more than one nutritionist.

This limitations of the nutritionist have been a problem to provide data on the nutritional condition of toddlers in central Sulawesi, especially Palu. This circumstances made the government and other relevant parties hard to make their policy and to detect the early symptoms of malnutrition in Palu. Therefore, through this research an implementation of the nutritionist will be done by computer-based system which can be used to determinate the nutritional status of the children. This study uses a support vector machine (SVM) in analyzing the nutritional status of the children by anthropometric index.

Nutritional status of Toddler

Nutritional Status of Toddler is a state of child health defined by the needs of physical energy and other nutrients from food that their physical impact measured anthropometric (Suharjo, 1996), and categorize according to WHO-NCHS standard using W/A, H/A, and W/H indices. There are many ways to measure the nutritional status of a group, one of it is the measurement of human body called anthropometry. Anthropometric number showed and connected with another variable on the use. Such as:

1) Age: age has a very important role to measure the nutritional status, wrong age will cause wrong interpretation of nutritional status. All precise result of weight and height will be useless if the age measurement is wrong. The error that exist mostly because the inclination of choosing the easy number such as 1 year, 1.5 year, 2 years, therefore the age measurement have to be done precisely, example 1 year is 12 months, 1 months is 30 days, so the age measurement using the full month, and the age by days is not count (Minister of Health, 1994).

2) Weight: weight is one of the measurement that shows the mass of the body tissue, like body fluid. Weight is very sensitive to a sudden change of the body cause by infection or the decrease consumption of food. Weight showed in the W/A (Weight for Age) or to predict the weight according to the previous weight. Weight commonly used because it's only use one measurement, but rely on the age but do not quite explain the nutritional status through time (Abunain, 1990).

Height: height shows the growth function. Height is very good to see the past nutritional condition, especially body weight and malnutrition in the past. Height showed in H/U (Height for Age) indices, or W/H (Weight for Height) indices. They are seldom being done because the slow change of height and commonly done once a year. This indices condition commonly shows the bad condition of environment, poverty, and yearly unhealthy incident (Ministry of Health, 2010).

¹Department of Information Technology, Faculty of Engineering, Tadulako University, Jalan Soekarno Hatta Km.9 Palu, Central Sulawesi, Indonesia.

²Faculty of Education and Teacher Training, Tadulako University, Jalan Soekarno Hatta Km.9 Palu, Central Sulawesi, Indonesia.

3) Weight and Height of the body is one of important parameters to human health, especially related to nutritional status of human. The using of W/A, H/A, and W/H indices is an indicator of nutritional status to knows the error of growth function, and body composition (M. Khumidi, 1994).

Support Vector Machine

Support Vector Machine (SVM) developed by Boser, Guyon, Vapnik, and was first presented in 1992 at the Annual Workshop on Computational Learning Theory. The basic concept of SVM is actually a harmonious combination of computational theories that have been there for decades before, such as margin hyper plane discussed by Duda & Hart in 1973, Cover Vapnik in 1965 and 1964. Kernel concepts introduced by Aronszajn 1950, and also the other supporting concepts. But until 1992, there has never been an effort assembling these components (Vapnik, 1999).

In the learning process, SVM introduced a new strategy by finding the best hyperplane in the input space, through a strategy called Structural Risk Minimization (Tsuda, 2000). The basic principle of SVM is linear classifier, and then developed in order to work on a non-linear problem by inserting the concept of kernel trick on high-dimensional workspace.

4) *Linear SVM*: SVM concepts can be explained simply as an attempt to find the best hyperplane which serves as a separator to two classes in the input space. Figure 1 shows some pattern that is a member of two classes: Positive (denoted by +1) and negative (denoted by -1). Joined pattern to the negative class symbolized by a box, while the pattern on the positive class symbolized by circles. The learning process in the classification problem interpreted as an attempt to find a line (hyperplane) that separates between the two groups. Various alternatives dividing line is shown in Figure 1.

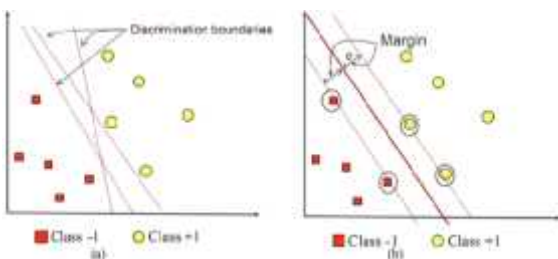


Fig. 1. The concept hyperplane on Support Vector Machine (SVM).

The best hyperplane separates between the two classes can be found by measuring the margin hyperplane, and find the maximum point. Margin is the distance between the hyperplane to the closest data from each class. The closest subset of data training set is called a support vector. Solid line in [1b] shows the best hyperplane, which is located in the middle of both class, whereas the squares point and circles that are in the black circle is a support vector. The efforts to locate the optimal hyperplane is the core of the learning process on SVM.

1) *Non-Linear SVM*: In general, the problems in the real world domains rarely are linearly separable. Most of these problems are nonlinear. To resolve the non-linear case, the calculation SVM divided into two phases, in which it used a concept called Kernel trick. Kernel trick providing

various facilities, because in the SVM learning process, to determine the support vector, we just simply knowing the used kernel function, and does not need to know the form of the nonlinear function. Various types of kernel functions commonly known, as summarized in (1). In non-linear SVM, first of all the data is mapped by the function to a higher dimensional vector space. In this new vector space, hyperlinks that linearly separates these two classes can be constructed. This is in line with the theory of Cover which states that if a transformation is nonlinear and dimensions of the feature space is high enough, then the data in the input space can be mapped to the new feature space, where the patterns have the high probability can be separated linearly (Haykin, 1998). The illustration of this concept can be seen in [2]. In [2a] shows that data in the negative class and positive class exist in two dimensional input space cannot be separated linearly. Furthermore, [2b] shows that the function maps each data on input space into new higher dimensional vector space (3-dimensional), where both classes can be separated linearly by a hyperplane.

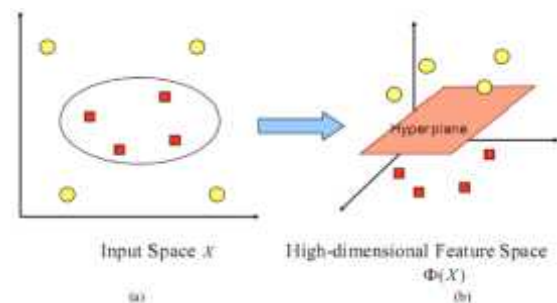


Fig. 2. Mapping Data Into The Higher Dimensional Space Vector.

TABLE I. GENERALLY USED KERNEL IN SUPPORT VECTOR MACHINE

Kernel Type	Definitions
Polynomial	$K(\vec{X}_i, \vec{X}_j) = (\vec{X}_i \cdot \vec{X}_j + 1)^p$
Gaussian	$K(\vec{X}_i, \vec{X}_j) = \exp\left(-\frac{\ \vec{x}_i - \vec{x}_j\ ^2}{2\sigma^2}\right)$
Sigmoid	$K(\vec{X}_i, \vec{X}_j) = \tanh(a\vec{X}_i \cdot \vec{X}_j + \beta)^p$

II. RESEARCH METHODS

Design systems generally consist of two parts, they are training process and the testing process. General overview of the classification system of nutritional status can be seen in [3]. The Stage that will be conducted in this study were divided into two, they are the stage for the training process and the testing process on the system.

In the training process, the factors which determining the nutritional status used as the feature. Then the SVM classifier is trained to partitioning the feature space. While in the testing phase, SVM classifier that has been trained to put the input pattern into one of the patterns class based on the features being measured.

In general, the classification process using SVM can be seen in [4].

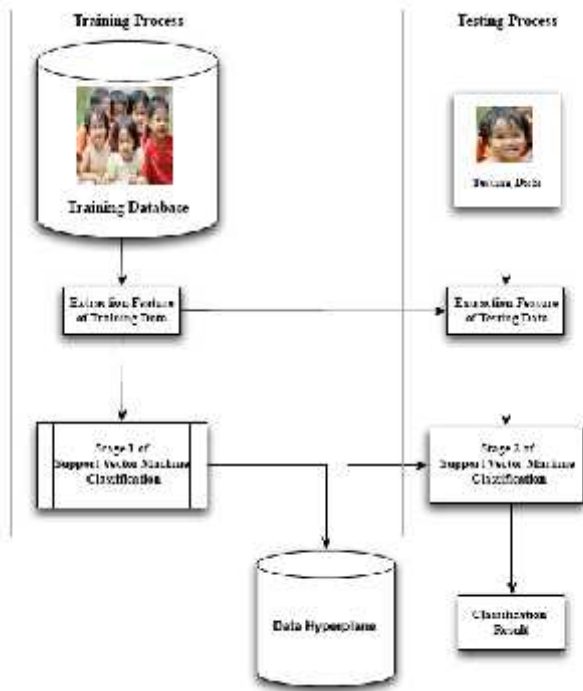


Fig. 3. General overview System.

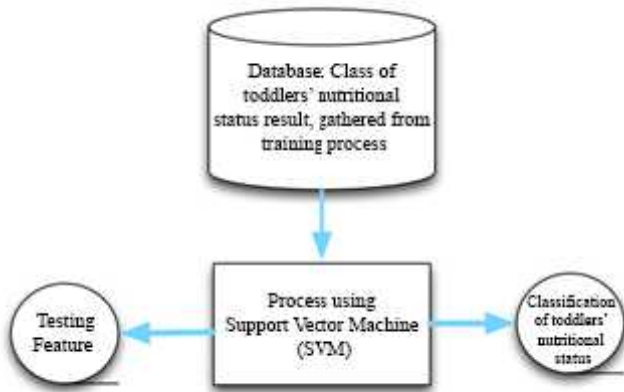


Fig. 4. Classification process of toddler nutritional status using Support Vector Machine.

III. RESULT AND DISCUSSION

Result

1) *Data Research Sample:* The following table are the 20 samples of data research that obtained from one of the health centers in the Palu city.

TABLE II. TODDLER DATA SAMPLE IN SINGGANI HEALTH CENTER

N.	Gender (M/F)	Date of Birth	W (kg)	H (cm)	HC	Age (Month)
1	F	09 th December 2009	11.5	95.5	18	33
2	F	28 th August 2012	2.5	53	32.8	1
3	F	28 th August 2012	2.5	48	33	0
4	F	25 th August 2012	3.5	49	33	0
5	M	12 th September 2012	3.4	53	34	0
6	M	25 th July 2012	4.8	57	34	1
7	F	16 th July 2012	4.5	46	35	2
8	F	07 th September 2012	3.3	48	35	0
9	M	15 th August 2012	3.3	49	35	0
10	F	08 th August 2012	3.7	48	36	1
11	M	10 th July 2012	3.5	51	36	2
12	F	09 th August 2012	3.5	53	36	1
13	F	19 th May 2012	4.5	58	36	3
14	M	04 th August 2012	3	53	36.5	1
15	F	27 th July 2012	5	55	39	2
16	M	07 th August 2012	4.3	58	39	1
17	M	10 th June 2012	5	59	39	3
18	F	21 st April 2012	6.2	60	39	4
19	F	10 th February 2012	6.8	60	39	7
20	F	01 st June 2012	6.2	64	39	3

2) *Clustering Research Data:* Once the data is obtained, the next is the fifth stage which is clustering (division) of the data. Overall data which obtained clustered into two groups of data that test data (training), and the training data (testing). Data sharing into training data and testing data, using fivefold cross validation method.

3) *Data normalization:* Normalization or scaling (change in the range of attribute values) of attributes valued real numbers generally also needed for the easier processes and ensures the attributes with a great value ranges do not dominate the attributes with smaller value ranges. The results of normalized training and testing data sample is shown as (3,4) below.

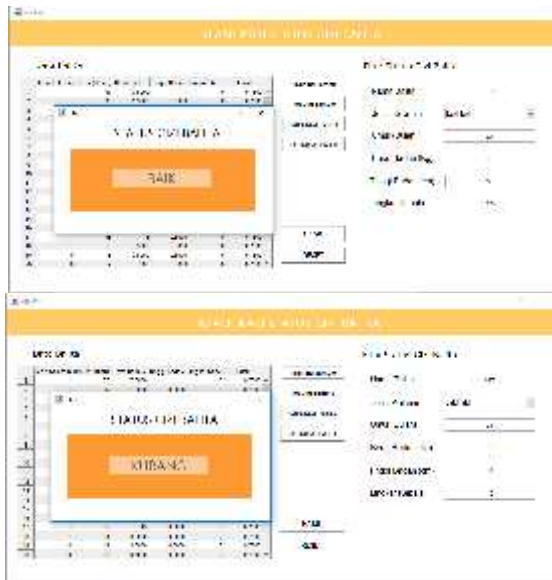


Fig. 8. System testing

Based on test results from total sample of 640 toddler's data from eight health centers in the Palu city (320 data is used as a training and 320 other data as test data) showed that the level of system accuracy of 92.18%. This result showed that out of total 320 samples from the whole health centers found that there are 295 data from the introduction system accordance to the real conditions, and the remaining 25 amount of data are not.

IV. CONCLUSION

Research to determinate the classification of nutritional status using a support vector machine provides the accuracy of 92.18%. These results indicate that the performance of the system is quite good but development of algorithms and completeness of the data has to be done.

ACKNOWLEDGMENT

This research was funded by The Indonesian Ministry of Research, Technology, and Higher Education (Kemenristek DIKTI RI). The writer would thanks also to LPPM UNTAD for support and assistance during the preparation of this paper.

REFERENCES

- [1] Abunain Djumadias, 1990, Aplikasi Antropometri sebagai Alat Ukur Status Gizi, Puslitbang Gizi Bogor.
- [2] Anggraeni,R dan Indrarti,. A, 2010, Klasifikasi Status Gizi Balita Berdasarkan Indeks Antropometri (BBU) Menggunakan Jaringan Saraf Tiruan, SNASTI
- [3] Burges Christopher J.C, 1998, A Tutorial on Support Vector Machine for Pattern Recognition, Kluwer, Academic Publisher, Boston.
- [4] Haykin, S.,1998, Neural Network:A comprehensive Foundation Ed, Prentice Hall
- [5] Kementrian Kesehatan RI, 2011, Keputusan Menteri Kesehatan RI No. 1995/MENKES/SK/XII/2010, Jakarta
- [6] Muaris, H, 2006, Breakfast Healthy Childhood, New York:Scholastic Press Main
- [7] Suharjo, 1996, Gizi dan Pangan, Kanisius, Yogyakarta
- [8] Scholkopf B dan Smola, 2002, Learning with Kernel, MIT Press
- [9] WHO, 2005.Measuring Change in Nutritional Status, Genewa.