



Analysis of Factors Influencing the Event Stunting with Chi-Square Method Approach Automatic Interaction Detection in North Sumatra

Mei Yunina Arianti¹, Ismail Husein², Rina Widyasari³

^{1,2,3}Department of Mathematics, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Article Info

Article history:

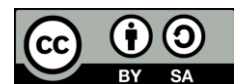
Keywords:

Classification
CHAID
Stunting

ABSTRACT

In statistics, one way to group (classify) a data mathematically is called classification. There are several types of classification methods, one of which is Chi-Square Automatic Interaction Detection, which is abbreviated as CHAID. CHAID method will be applied to North Sumatra stunting data. Stunting is a condition in which a person's height is shorter than that of other people of the same age. The dependent variable in the study was stunting, which was measured in infants aged 24-59 months, while the independent variable was a factor that affected stunting. The factors that influence the incidence of stunting in infants aged 24-59 months in North Sumatra based on the results of the analysis (CHAID) are family income, sanitation and water sources. From the results of the CHAID analysis, 5 different groups were obtained, namely: Infants aged 24-59 months who were stunted were infants with an economic income of IDR 1,000,000 - IDR 2,000,000 (98.4%) and inadequate sanitation (100%), Babies aged 24-59 months who are stunted are babies with an economic income of IDR 1,000,000 - IDR 2,000,000 (98.4%) and proper sanitation (93.2%), babies aged 24-59 months who are not stunted are infants with an economic income of Rp. 3,000,000 - Rp. 5,000,000 (95.5%). Infants aged 24-59 months who are not stunted are infants with an economic income of > Rp. 5,000,000 (99.5%) and adequate water sources (100%) and infants aged 24-59 months who are not stunted with an economic income of > IDR 5,000,000 (99.5%) and inadequate water sources (97.1%).

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Mei Yunina Arianti,

Department of Mathematics, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: meiyuninaart@gmail.com

1. INTRODUCTION

In statistics one way to group (classify) something data is mathematically called classification. Frequent classification problems encountered in everyday life, such as classification of educational data, health, business and other fields. There are several kinds of classification methods viz Classification And Regression Trees abbreviated as CART, Chi-Square Automatic Interaction Detection, abbreviated as CHAID, Support Vector Machine abbreviated as SVM and other classification methods. The CHAID method is a non-parametric statistical method introduced by G. V. Kass in 1980. The procedure for this method is part of the Automatic Interaction Detection (AID) technique with uses the chi-square statistic as its main tool. In general, this method works by studying the relationship between the dependent variable with some the independent variable then classifies the sample based on the relationship the. This method has three stages, namely merging, separating, and termination (Syahdan et al., 2021).

Stunting is a condition where a person is too tall short compared to the height of other people in general who are the same age (Rahayu et al., 2018). Basically stunting is a condition of growth failure in children due to prolonged malnutrition. This nutritional deficiency as the impact of giving food that is not in accordance with nutritional needs. Stunting occurs in all regions in Indonesia, the results of the Status Survey Indonesian Toddler Nutrition (SSGBI) Year 2021, North Sumatra province has 5th highest prevalence in Indonesia. In this study, the province of North Sumatra used as the main focus in viewing the incidence of stunting with numbers stunting prevalence is most similar to Indonesia's national stunting prevalence namely 24.4%. The factors that influence the incidence of stunting in North Sumatra namely, internal and external factors. Internal factors, namely events stunting as measured in infants aged 24-59 months. External factors namely gender, history of breastfeeding, mother's education, economic status, sanitation, source of water and number of family members. The number of cases of stunting is quite large with a variety of causative factors need serious treatment.

Therefore, it will be analyzed which factors are very influential with incidence of stunting using the CHAID method. This method will form a segmentation of stunting data based on the relationship between variables dependent, namely stunting in infants aged 24-59 months with the independent variable namely seven factors that influence stunting. Segmentation results will be shows which factors are very influential on the incidence of stunting so that it has a significant impact on reducing the incidence of stunting in infants aged 24-59 months in North Sumatra.

2. RESEARCH METHODE

Descriptive Analysis

a. Merging stage

- Create a contingency table for each dependent variable on the variable independent.
- Calculating the Chi Squared test to identify the independent variables most significant to serve as the initial separator in the decision tree.

$$\chi^2 = \sum_{i=1}^b \sum_{j=1}^k \frac{(n_{ij} - E_{ij})^2}{E_{ij}}$$

Where:

$$E_{ij} = \frac{n_i \cdot n_j}{n}$$

Information :

n_{ij} : Number of observations that fall into the i-th category of the first variable and the j category of the second variable

E_{ij} : The expected frequency of observations that fall into the i-th category of the first variable and the j category of the second variable

b : The number of categories in the first variable

k : The number of categories in the second variable

b. Stage of separation (Splitting)

c. Choose the independent variable that has the largest χ^2 count and p-value smallest (merging stage) which will be used as a node separator.

d. Stage of termination (Stopping) Repeat the merging and splitting steps to get a knot next, the termination stage if there are no more independent variables significant.

- Draw a stunting CHAID tree diagram.

- Interpret the CHAID stunting tree diagram, namely: determine the order of association of each independent variable to stunting.

3. RESULT AND ANALYSIS

Stunting data analysis uses the Chi-Square Automatic method Interaction Detection is as follows:

a. Merging Stage

Make a contingency table for each independent variable with its dependent variable, then test the chi-square test for each independent variable. Following are the results of the first chi-square test on each independent variable that affect stunting.

Table 4.1 The Results Of The First Chi- Square Test

Variabel	χ^2_{hitung}	χ^2_{tabel}	Keterangan
Jenis Kelamin* Stunting	18,830	3,841	Tidak Signifikan
Riwayat Pemberian ASI* Stunting	726,849	7,8417	Tidak Signifikan
Pendidikan Ibu* Stunting	842,906	7,7794	Tidak Signifikan
Pendapatan Keluarga* Stunting	938,885	7,7794	Signifikan
Sanitasi* Stunting	260,317	3,841	Tidak Signifikan
Sumber Air* Stunting	159,7	3,841	Tidak Signifikan
Jumlah Anggota Keluarga* Stunting	809,257	5,9915	Tidak Signifikan

The next stage is splitting. Of the 7 independent variables on Table 4.16 above, it can be seen that the most significant variable to Stunting based on the Chi-Square value is the largest of the other independent variables, namely family income with χ_{hitung} 938,885. Therefore, this variable will be used as the best sorting or separating variable at the inner root node decision tree so that the initial decision tree is obtained as shown in Figure 4.2 below:

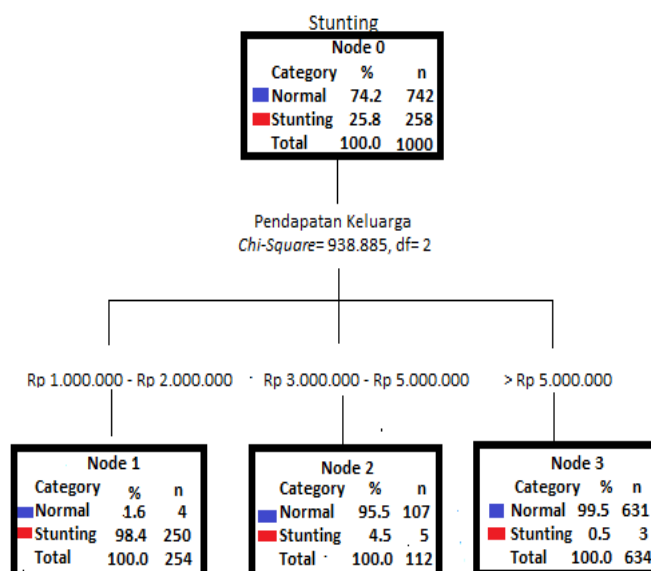


Figure 1. CHAID'S Initial Tree

To see if there are still branches at node 1 variable family income, the Chi-Squared test can be repeated for each category. Following are the results of the second chi-square test:

Table 4.2 The Results Of The Second Chi- Square Test

Variabel	χ^2_{hitung}	χ^2_{tabel}	Keterangan
Jenis Kelamin* Stunting	3,639	3,841	Tidak Signifikan
Riwayat Pemberian ASI* Stunting	8,793	7,8417	Signifikan
Pendidikan Ibu* Stunting	12,512	7,7794	Signifikan
Sanitasi* Stunting	13,432	3,841	Signifikan
Sumber Air* Stunting	3,309	3,841	Tidak Signifikan
Jumlah Anggota Keluarga* Stunting	10,417	5,9915	Signifikan

From Table 4.28 it can be seen that the most significant variable is sanitation variable with χ^2_{count} 13,342. Because Sanitation variable is a significant variable, then sanitation variable is made a branch at node 1. So that the tree is formed by addition the sanitation variable becomes like Figure 4.3 below:

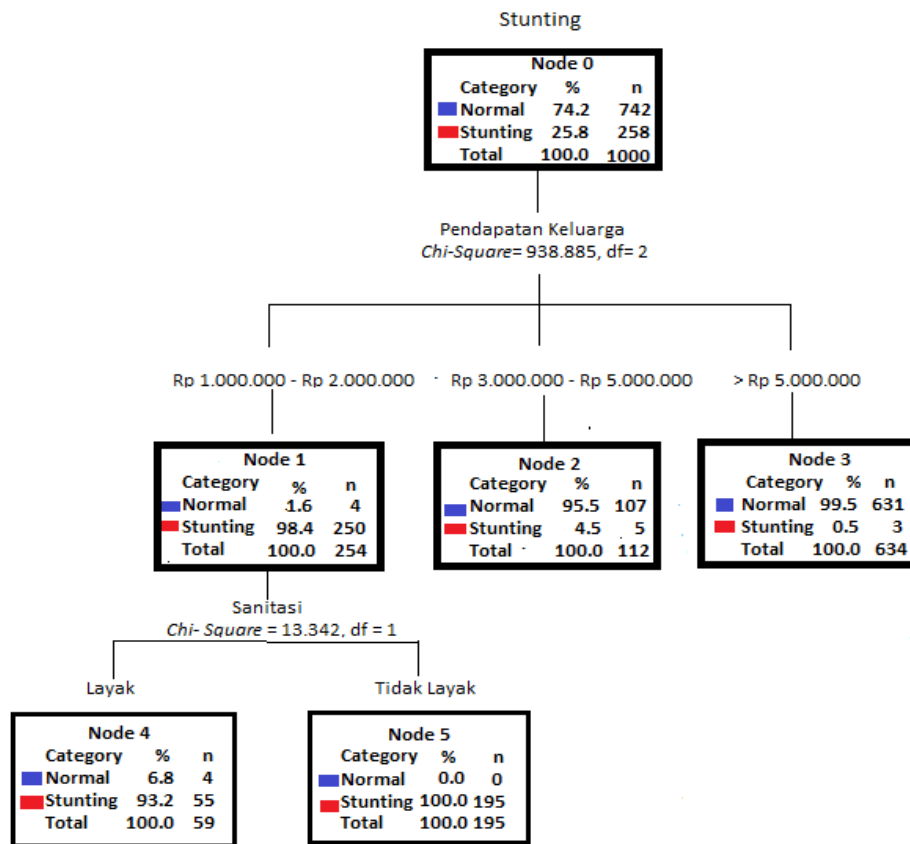


Figure 2. Additional Sanitation Variables

Next, we will look for branches for node 2 by detecting which variable is the most significant by looking at the largest χ^2_{count} value. From the results of the third chi-square test there is no significant variable. Next, we look for the branch node 3. Following are the results of the fourth chi-square test:

Table 4.51 The Results Of The Fourth Chi- Square Test

Variabel	χ^2_{hitung}	χ^2_{tabel}	Keterangan
Jenis Kelamin* Stunting	6,493	3,841	Signifikan
Riwayat Pemberian ASI* Stunting	13,898	7,8417	Signifikan
Pendidikan Ibu* Stunting	13,898	7,7794	Signifikan
Sumber Air* Stunting	15,361	3,841	Signifikan
Jumlah Anggota Keluarga* Stunting	11,247	5,9915	Signifikan

From Table 4.51 it can be seen that the most significant variable is the water source variable with χ^2_{count} 15.361. Therefore, this variable was chosen to be a branch for node 3. So the tree that is formed with the addition of the sanitation variable becomes as shown in Figure 4.4 below:

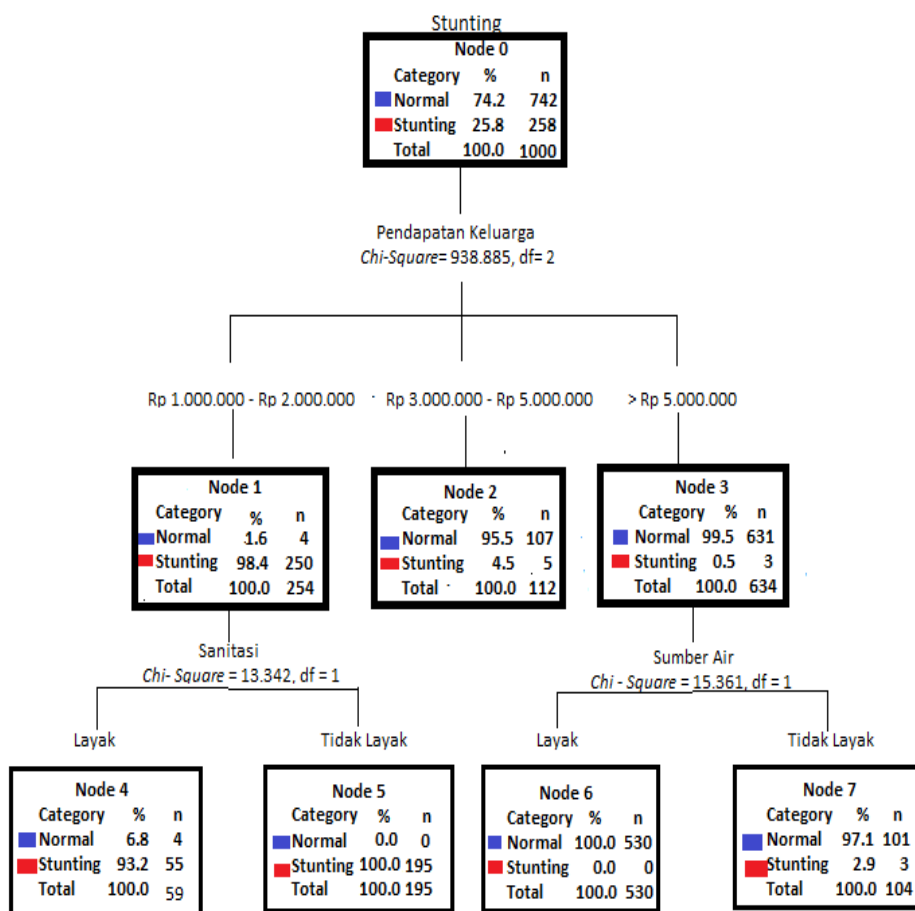


Figure 4.3 Additional Water Source Variables

Next, we will look for branches for nodes 4, 5, 6 and 7 by conducting the fifth to eighth chi-square test. By looking at which variable is the most significant that has the largest χ^2_{count} value. Because the results of the fifth to eighth chi-square test on each independent variable were not significant, a stop was made. Thus, the chi-square test stops right at the fourth chi-square test. The decision tree results remain the same according to Figure 4.4.

4. CONCLUSION

Based on the results of the research and discussion above, several can be drawn conclusion as follows:

1. Factors that influence the incidence of stunting in infants aged 24- 59 months in North Sumatra, namely: Gender, History of Breastfeeding, Mother's Education, Family Income, Sanitation, Water Source and Total Family members.
2. Factors that influence the incidence of stunting in infants aged 24- 59 months in North Sumatra based on Chi-Square Automatic analysis Interaction Detection (CHAID) which was discontinued until the chi- the fourth square is Family Income, Sanitation, and Water Resources.
3. Chi-Square Automatic Interaction Detection (CHAID) analysis results Five different groups were obtained, namely:
 - a. Infants aged 24-59 months who are stunted are infants with state of economic income of IDR 1,000,000 - IDR 2,000,000 (98.4%) and inadequate sanitation (100%).
 - b. Infants aged 24-59 months who are stunted are infants with state of economic income of IDR 1,000,000 - IDR 2,000,000 (98.4%) and proper sanitation (93.2%).
 - c. Infants aged 24-59 months who are not stunted are infants with state of economic income Rp. 3,000,000 - Rp. 5,000,000 (95.5%).
 - d. Infants aged 24-59 months who are not stunted are infants with condition of economic income > IDR 5,000,000 (99.5%) and water source decent (100%).

- e. Infants aged 24-59 months who are not stunted by their circumstances economic income > IDR 5,000,000 (99.5%) and the water source is not feasible (97.1%).

REFERENCES

- [1] Fitri, H., Aniswita, Charles. 2018. Factor Linkage Analysis Thesis Completion Period Using the CHAID Method. *Journal Proceedings of STKIP PGRI West Sumatra*. 4(1). 230-243
- [2] Frogenese, F., Siekmans, K., Kouanda, S., Druetz, T., Ly, A., Diabete, S., & Haddad, S. 2017. Impact Of Contaminated Household Environment On Stunting In Children Aged 12-59 Months In Burkina Faso. 356-363
- [3] Helena, K. Susanti, Y., Respatiulwan. 2019. Application Of Methods Chi-Square Automatic Interaction Detection (CHAID) And Classification And Regression Trees On The Classification Of Working Status In Brebes District. *Journal Of Scientific National Proceedings*. 6(5). 1-9
- [4] Islamy, U., Dini, S. K. 2021. Application of the CHAID Method (Chi-Squared Automatic Interaction Detection) on Student Continuation Join the Tutoring Institute. *Journal of Proceedings Mathematics and Mathematics Education*. 4. 1-7
- [5] Mufdillah. 2017. Guidebook For Empowering Breastfeeding Mothers In The Exclusive Breastfeeding Program. Aisyiyah University Yogyakarta: Yogyakarta
- [6] Nugraha, Jaka. 2016. Introduction to Categorical Data Analysis. Yogyakarta: Deepublish
- [7] Rahayu, A., Yulidasari, F., Angraini, L. 2018. Study Guide – Stunting And Prevention Efforts For Health Students Public. New Banjar: Publisher CV Mine
- [8] Shyntia, F. D., Kurnia, A., Dito, G. A. 2022. The CHAID Method Approach and Logistic Regression in Analyzing Influential Factors on Stunting Incidents in West Java Province. *Journal of Statistics*. 11(1).
- [9] Sugiyono. 2019. Quantitative, Qualitative and R&D Research Methods. Bandung : Alfabeta
- [10] Sulviana, V., Wigena, A. H., Indahwati. 2018. Method Implementation CHAID (Chi-Squared Automatic Interaction Detection) on Segmentation Trend of Sales of Soft Drinks in Indonesia. 2(2). 24-31
- [11] Syahdan, S., Arif, A., Megawati. 2021. Analysis of the Factors that Cause Dengue Hemorrhagic Fever (DHF) Using Chi-Square Automatic Interaction Detection (CHAID). *International Journal of Nature Science and Engineering*. 5(3). 104-113
- [12] Yamin, S. 2021. Statistical Tutorial: SPSS, LISREL, WARPPS and JASP. Depok: PT Dewangga Energi Internasional
- [13] Zubedi, F., Oroh, F. A., Aliu, M.A. 2021. Stunting and Nutrition Modeling Less In Bone Bolango District Using Regression Poisson Generalized . *Journal of Mathematics and Education Mathematics*. 6(2). 1