https://jurnal.seaninstitute.or.id/index.php/jukomi
Jurnal Komputer Indonesia (JU-KOMI)
Volume 3 No 01, 2024

JU-KOMI
E-ISSN: 2963-0460

Analysis of Malaria Disease Classification Based on Age in the Work Area of Idanogawo Health Center, Idanogawo District, Nias Regency Using the Decision Tree Algorithm

Sipra Barutu
Magister Teknologi Informasi, Unversitas Pembangunan Pancabudi, Medan, Indonesia

ARTICLE INFO

Keywords:
Malaria, age group, Decision Tree algorithm, malaria incidence

ABSTRACT

Malaria is an infectious disease caused by the Plasmodium parasite, which is transmitted through the bite of the Anopheles mosquito. This disease can affect individuals of all age groups, but its prevalence varies between age groups. This study aims to analyze the relationship between age and malaria incidence in the working area of the Idanogawo Health Center, Nias Regency, using the Decision Tree algorithm. Secondary data collected from medical records at the Idanogawo Health Center were processed through preprocessing to obtain age and malaria incidence attributes. The results of the analysis showed that the age group <30 years had the highest percentage of malaria incidence (80%), followed by the age groups 30-39 years and 40-49 years, each of which recorded an incidence rate of around 73%. The age group >=50 years also showed a high incidence of malaria (100%), although the sample size was small. These findings indicate that malaria attacks younger people more, but older age groups are also significantly affected.

Email :
barutusipra@gmail.com

INTRODUCTION

Malaria is an infectious disease caused by the Plasmodium parasite that lives and reproduces in human red blood cells, transmitted by female malaria mosquitoes (Anopheles), can attack everyone, both men and women in all age groups from infants, children and adults. Malaria is also caused by the plasmodium parasite that lives and reproduces in human red blood cells. The types of plasmodium that are commonly found in Indonesia are plasmodium falciparum and plasmodium vivax, while plasmodium mamariae can be found in several provinces including: Lampung, NTT and Papua (North Sumatra Health Office, 2019).

The percentage of districts/cities with low endemic status by province shows that most districts/cities in Indonesia have malaria-free status or have API <1 per 1,000 population. Only 11 provinces have not met these two criteria, namely North Sumatra, West Sumatra, Southeast Sulawesi, Lampung, Riau Islands, North Maluku, East Kalimantan, Maluku, East Nusa Tenggara, West Papua, and Papua (Ministry of Health of the Republic of Indonesia, 2020).

On the malaria distribution map according to the 2019 health profile, the malaria morbidity rate is depicted by the Annual Parasite Incidence (API) indicator per 1,000 population, which is the proportion of positive malaria patients to the population at risk in the area with a constant of 1,000. The malaria API in Indonesia in 2019 increased compared to 2018, from 0.84 to 0.93 per 1,000 population. However, the malaria API in Indonesia has shown a downward trend since 2009.

In 2019, there were 14 districts/cities with a percentage of examination of suspected Malaria in the Laboratory reaching 100%, namely Nias, Gunung Sitoli, Pemantang Siantar, Sibolga, West Nias, North Nias, North Padang Lawas, Serdang Berdagai, South Nias, Asahan, Toba Samosir, North Tapanuli, Central Tapanuli, South Tapanuli, Langkat, 2 districts reached a figure above 90%, namely Langkat Regency, Mandailing Natal, while there were 14 districts/cities that

did not have a percentage of 0%, namely Padang Sidempuan Regency/City, Medan, Binjai (North Sumatra Health Office, 2019).

Nias is one of the areas endemic to malaria from the data on the percentage of malaria suspects examined in the laboratory according to Regency/City in 2019 from the 2019 North Sumatra health profile, it states that the incidence of malaria reached 100% in Nias Regency/City, Gunung Sitoli, South Nias, West Nias, North Nias (North Sumatra Health Office, 2019).

Based on the results of an initial survey conducted by researchers at the Idanogawo Health Center in 2021, there were 10 diseases and malaria was the first disease based on data recording obtained from the Idanogawo Health Center on July 23, 2021, malaria was in the 10 largest diseases over the past year.

Knowing the age prediction in malaria incidence has significant benefits in regional health policies. By understanding the health profile of the community, service managers can adjust the types of actions and approaches that are more effective for prevention in each age group. Therefore, this study will focus on age prediction for people experiencing malaria who visit the Idanogawo Health Center, Idanogawo District, Nias Regency using the Decision Tree algorithm, with the aim of providing relevant recommendations as a source of information, evaluation, and attention for consideration in making a decision or policy and action in controlling malaria as well as input for improving performance in the field. As a contribution of information, motivation, and evaluation to improve efforts to control malaria.

METHOD

Data collection

Data collection used in this study is secondary data. Secondary data is data that already exists, both in literature books and sources from publications that are carried out by reading, studying, citing and summarizing data related to the problem to be discussed. Secondary data in this study were obtained through reports and documents from medical records in the working area of the Idanogawo Health Center, Idanogawo District, Nias Regency.

Preprocessing

The data obtained is preprocessed, Preprocessing is an important step in the data mining process. The data preprocessing stage removes data that has null or empty values and changes the data to be more structured by using the implementation of data cleaning and data transformation. This step functions so that the data to be processed becomes more structured and facilitates the modeling process. So that the dataset is of high quality and in accordance with the rules for implementing predictions, namely by eliminating Education, work, knowledge, community actions, and home environmental conditions, so that what is used is the age attribute and malaria incidence.

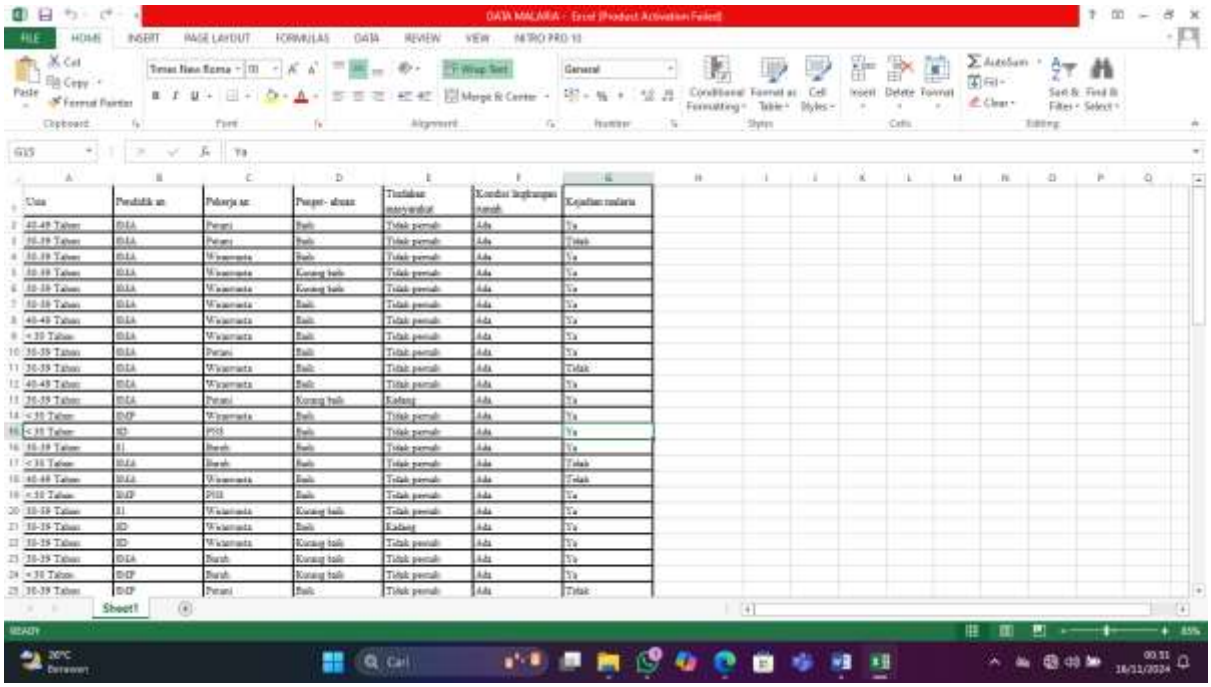
Modeling and analysis

The processed data is modeled using rapid miner and the modeling results. The analysis is carried out. Rapid Miner is an application or software that functions as a learning tool in data mining science. The platform was developed by a company dedicated to all steps involving large amounts of data in commercial business, research, education, training and learning.

https://jurnal.seaninstitute.or.id/index.php/jukomi
Jurnal Komputer Indonesia (JU-KOMI)
Volume 3 No 01, 2024
JU-KOMI
E-ISSN: 2963-0460

RESULTS AND DISCUSSION

Data collection

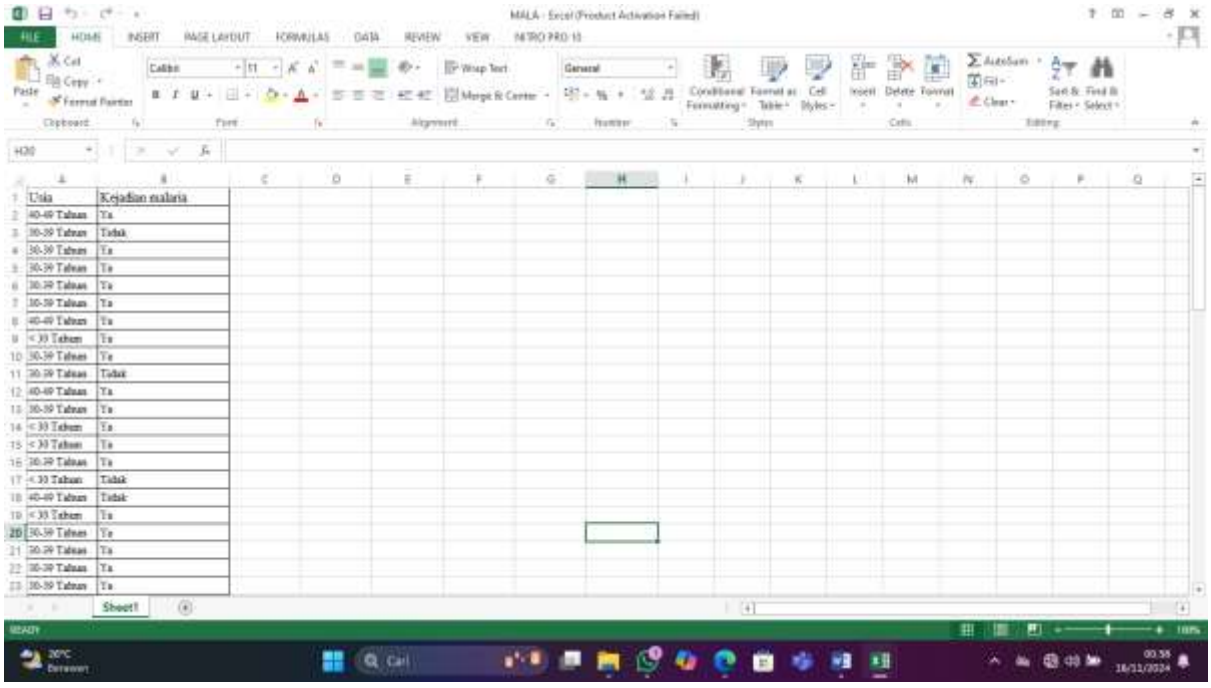


Usia	Pendidikan	Pekerjaan	Pengetahuan	Tindakan masyarakat	Kondisi lingkungan rumah	Kejadian malaria
40-49 Tahun	SMA	Petani	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Petani	Baik	Tidak pernah	Ada	Tidak
30-39 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Wiraswasta	Kurang baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Wiraswasta	Kurang baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Ya
40-49 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Ya
< 30 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Petani	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Tidak
40-49 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Petani	Kurang baik	Kadang	Ada	Ya
< 30 Tahun	SMP	Wiraswasta	Baik	Tidak pernah	Ada	Ya
< 30 Tahun	SD	PNS	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SI	Buruh	Baik	Tidak pernah	Ada	Ya
< 30 Tahun	SMA	Buruh	Baik	Tidak pernah	Ada	Tidak
40-49 Tahun	SMA	Wiraswasta	Baik	Tidak pernah	Ada	Tidak
< 30 Tahun	SMP	PNS	Baik	Tidak pernah	Ada	Ya
30-39 Tahun	SI	Wiraswasta	Kurang baik	Tidak pernah	Ada	Ya
30-39 Tahun	SD	Wiraswasta	Baik	Kadang	Ada	Ya
30-39 Tahun	SD	Wiraswasta	Kurang baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMA	Buruh	Kurang baik	Tidak pernah	Ada	Ya
< 30 Tahun	SMP	Buruh	Kurang baik	Tidak pernah	Ada	Ya
30-39 Tahun	SMP	Petani	Baik	Tidak pernah	Ada	Tidak

Figure 1. Malaria data collection

In Figure 1 it can be seen that the sample size in this study was 71 people from a population of 250 people. The sample in this study was obtained using the Slovin formula. The incidence of diarrhea based on the table is influenced by age, education, occupation, knowledge, community actions, and home environmental conditions.

Process



Usia	Kejadian malaria
40-49 Tahun	Ya
30-39 Tahun	Tidak
30-39 Tahun	Ya
30-39 Tahun	Ya
30-39 Tahun	Ya
30-39 Tahun	Ya
40-49 Tahun	Ya
< 30 Tahun	Ya
30-39 Tahun	Ya
30-39 Tahun	Tidak
40-49 Tahun	Ya
30-39 Tahun	Ya
< 30 Tahun	Ya
< 30 Tahun	Ya
30-39 Tahun	Ya
< 30 Tahun	Tidak
40-49 Tahun	Tidak
< 30 Tahun	Ya
30-39 Tahun	Ya
30-39 Tahun	Ya
30-39 Tahun	Ya

In Figure 2, the data has been preprocessed, resulting in Age data, and Malaria incidence.

Decision Tree Analysis

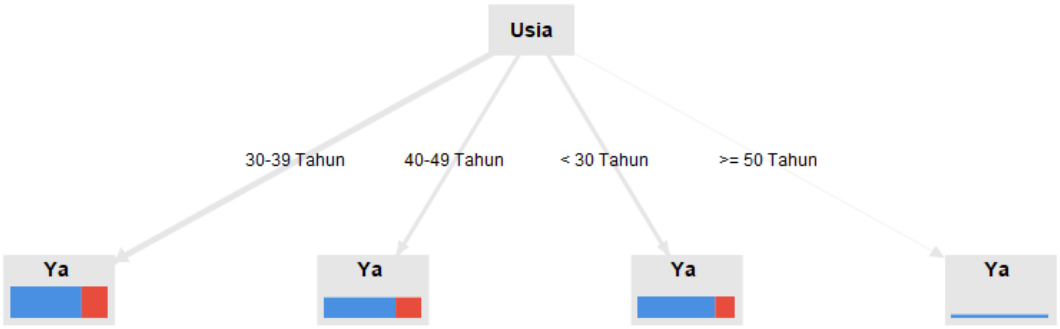


Figure 3 After the data has been preprocessed, the data is modeled using a decision tree.

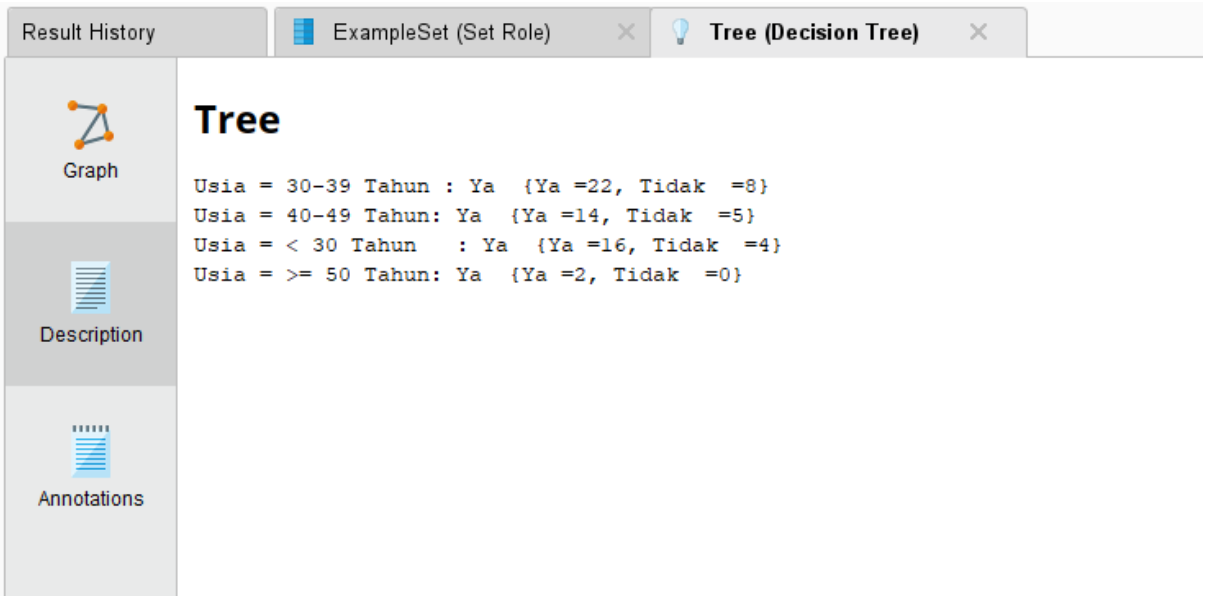


Figure 4 Description Tree

Based on Figures 3 and 4 tree diagrams depicting the relationship between age and malaria incidence, it can be seen that the age group <30 years has the highest percentage of malaria incidence, with 16 out of 20 people (80%) experiencing malaria. The age groups 30-39 years and 40-49 years have similar malaria incidence rates, each around 73%, with 22 out of 30 people aged 30-39 years and 14 out of 19 people aged 40-49 years infected. Meanwhile, in the age group >=50 years, all individuals (100%) experienced malaria, although the sample was very small (only 2 people), which shows a very high incidence of malaria despite the limited number of samples. Overall, these data show a tendency for higher malaria incidence at a young age, but the incidence rate in older age groups is also significant although it cannot be generalized due to the small sample size in the age group >=50 years.

CONCLUSION

This study shows that the Decision Tree algorithm can be effectively used to predict malaria incidence based on age in the working area of Idanogawo Health Center, Nias Regency. The results of the analysis show that the age group <30 years has the highest percentage of malaria incidence (80%), followed by the age group 30-39 years and 40-49 years, with a percentage of incidence of

around 73%. The age group >=50 years also recorded a high incidence rate (100%), although the number of samples was very small. These findings indicate that malaria attacks individuals more at a young age, but older age groups are also significantly affected. This study provides important information that can be used to design malaria control policies that are more focused and in accordance with the age profile of the community, as well as contributing to increasing efforts to prevent malaria in endemic areas such as Nias Regency.

REFERENCE

Sumut, D. (2019). Profil sumatera utara. 1-379.

RI, k. (2020). PROFIL KESEHATAN Indonsesia TAHUN 2019. 1-497.

Banyal, N. A., & Surianti, S. (2023). Implementasi Data Mining Untuk Prediksi Penyakit Malaria Tropika Dengan Algoritma C4. 5. Jurnal Ilmiah MATRIK, 25(3), 236-242.

Ramadhan, N. G., & Khoirunnisa, A. (2021). Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine. Jurnal Media Informatika Budidarma, 5(4), 1580-1584.

Febriana, U., Furqon, M. T., & Rahayudi, B. (2018). Klasifikasi penyakit typhoid fever (tf) dan dengue haemorhagic fever (dhf) dengan menerapkan algoritma decision tree C4. 5 (Studi kasus: Rumah Sakit Wilujeng Kediri). Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2(3), 1275-1282.

Charisma, R. A., Pamungkas, S., Saputra, R. A., Ramadhan, N. G., & Adhinata, F. D. (2022). Analisis Penerapan Metode Ensembled Learning Decision Tree Pada Klasifikasi Virus Hepatitis C. Journal of Computer System and Informatics (JoSYC), 3(4), 405-409.