Data mining for dengue hemorrhagic fever (DHF) prediction with naive Bayes method

To cite this article: Ria Aralfyah and Fariani Hermin 2018 J. Phys.: Conf. Ser. 948 012077

View the article online for updates and enhancements.
Knowledge about DDB is lacking [7]. Data mining techniques with various methods are widely used to solve problems in social and economic impacts. There are several topics on this public health problem that are not considered in other studies; this issue is caused by the omission of critical factors that can influence DDB. Therefore, it is necessary to reduce the number of mosquitoes using the spread of dengue fever knowledge among the public about the importance of wearing protective clothing during the day and avoiding exposure to mosquitoes. This study aims to analyze the spread of dengue fever knowledge among the public about the importance of wearing protective clothing during the day and avoiding exposure to mosquitoes. The study focuses on the impact of education on the prevention and control of Dengue Fever (DDB). The study also measures the effectiveness of education programs in reducing the spread of dengue fever in Indonesia. The study population consists of residents of Surakarta City, Central Java, Indonesia. The study used a questionnaire survey method to collect data. The data analysis was conducted using multiple regression analysis.

Introduction

One of the problems of prevention and eradication of infectious diseases, which until now is still a problem in Indonesia, is Dengue Fever. This disease is transmitted by Aedes mosquitoes, which bite humans and birds. Dengue Fever is a viral disease that affects all age groups, but it is more common in children and young adults. The disease is caused by the dengue virus, which is transmitted to humans through the bite of infected mosquitoes. The virus is present in various countries around the world, including Indonesia. The spread of dengue fever is not only a problem in urban areas but also in rural areas.

Data mining for dengue hemorrhagic fever (DHF) prediction

With naive Bayes method
predictdisease [6]. Diagnosing heart disease by using Naive Bayes algorithm [7][8] offers more than 86% of accuracy. [9][10] It offers more than 79% of accuracy in a work to predict diabetes disease. This study aims to predict dengue fever with Naive Bayes method. It is used because it has advantages: It enhances the classification performance by eliminating the un-related features. It has a good performance and takes less computational time [11]. Parameter of the input variable is the patient’s medical data (temperature, spot, bleeding, and torruine test) and the output variable suffers from DHF or not.

2. Research Method
This research was conducted from April 2017 until November 2017. Field research was conducted in DKI Jakarta. Laboratory research was conducted in Computer Laboratory, Faculty of Mathematics and Natural Sciences of State University of Jakarta and Computer Science Department of Bogor Agricultural University. The research data is secondary data from research of M. Syafii in 2006. The sample of research was taken from medical record (medical record) hospital Dr. Sardjito Yogyakarta on 13-16 December 2005. The sample consists of data of patients suffering from dengue fever and dengue fever (DD) as many as 212 patients. According to international classification of diseases tenth revision (ICD 10) DHF disease is coded A.91 and DD disease with code A.90. Samples were taken from all hospitalized patients from January to November 2005. The requirements of the medical records taken as samples were in the medical record that there were 4 (four) notes on the clinical criteria: fever (heat), patches, spontaneous bleeding signs (nosebleeds, bleeding gums, bloody vomitus and black colored stools) and results torukuet test also noted criteria laboratory results of blood platelet examination. DHF was used as a comparison in this study due to consideration of 4 clinical criteria and 1 laboratory criteria. Processing and data analysis were using software Orange 3.4.5.

3. Research Result
By using patient medical data input, that is; fever (temperature), presence of spotting, the presence of bleeding, and torukuet test and output data, diagnose of the patient whether or not having dengue fever is made. Data processing using Naive bayes method with tools Orange, obtained in figure 1.

![Figure 1. Widget data processing with Orange.](image-url)
The performance of the algorithm is bad, as the resulting curve approaches the
frequency. The performance classification of Naive Bayes algorithm can be seen from the shape
From Figure 4, the performance classification of Naive Bayes algorithm can be seen as follows.
From Figure 3, the model is able to predict false of 55 and true equal to 92, in other words accuracy
between the predicted and actual class shows in confusion matrix.

Figure 2. The results of the Naive Bayes prediction test.
Acknowledgments

The authors would like to thank the directors of research and community service of the ministry of research technology and higher education Directorate General of Higher Education for funding and supporting the institutes of research and community service, Jakarta State University, Jakarta.

References