

ABSTRAK

Klasifikasi merupakan proses untuk dapat menemukan model atau fungsi yang menjelaskan dan membedakan suatu konsep atau kelas data. Klasifikasi bertujuan untuk dapat memperkirakan kelas dari objek yang labelnya tidak dapat diketahui. Salah satu penyakit yang sangat berbahaya yang dapat diklasifikasikan adalah Demam Berdarah Dengue (DBD). Menurut Kementerian Kesehatan jumlah penderita DBD di Bangkalan jumlah kasus meninggal pada 2017 sebanyak 493 jiwa, tahun berikutnya 467 jiwa, pada 2019 telah mencapai mencapai 917 jiwa. Bahkan di Juli 2020 sudah tercatat sebanyak 459 jiwa yang meninggal akibat DBD. DBD adalah suatu penyakit yang dapat ditularkan oleh nyamuk lewat gigitannya yang mengandung virus *flaviviridae* serta dapat menyebabkan kematian. Gejala yang ditimbulkan dari DBD sama atau mirip dengan penyakit lain yaitu adanya ruam, nyeri kepala, nyeri otot dan demam, yang mengakibatkan instansi puskesmas Socah lebih lama dalam mendiagnosis pasien. Dibutuhkan waktu selama tiga sampai empat hari sesuai hasil laboratorium dari Puskesmas Socah. Untuk mempermudah proses deteksi dini, maka diciptakan klasifikasi DBD menggunakan teknik *Data mining* dengan metode *Decision Tree* dan *Random Forest*. Data set yang digunakan untuk mengolah data berasal dari Puskesmas Socah dengan total 200 data pasien yang terindikasi DBD dan tidak. Dari set data akan dibagi menjadi 10 atribut dengan dua kelas data yaitu DBD dan Sehat. Fokus utama dalam penelitian ini adalah menghasilkan klasifikasi dengan akurasi paling tinggi dari metode *Decision Tree* dan *Random Forest*. Akurasi paling tinggi yaitu *Random Forest* mencapai *presisi* 96% , *recall* 97%, dan *F1 score* 96% dengan rata-rata waktu komputasi sebesar 0.233980 sebagai *output* dalam pembuatan aplikasi deteksi dini penyakit DBD. Parameter yang digunakan adalah gini indeks.

Kata kunci : Klasifikasi, *Decision Tree*, *Random Forest*

ABSTRACT

Classification is the process of finding a model or function that explains and distinguishes a concept or class of data. Classification aims to be able to estimate the class of an object whose label cannot be known. One very dangerous disease that can be classified is Dengue Fever (DHF). According to the Ministry of Health, the number of DHF sufferers in Bangkalan, the number of cases died in 2017 was 493 people, the following year 467 people, in 2019 it reached 917 people. Even in July 2020, 459 people were recorded to have died from DHF. DHF is a disease that can be transmitted by mosquitoes through their bites that contain flaviviridae viruses and can cause death. The symptoms of DHF are the same or similar to other diseases, namely a rash, headache, muscle pain and fever, which results in the Socah puskesmas agency taking longer to diagnose patients. It takes three to four days according to laboratory results from the Socah Health Center. To facilitate the early detection process, the classification of DHF was created using data mining techniques with the Decision Tree and Random Forest methods. The data set used to process data comes from the Socah Health Center with a total of 200 patient data indicating DHF and not. The data set will be divided into 10 attributes with two data classes, namely DHF and Healthy. The main focus in this research is to produce a classification with the highest accuracy of the Decision Tree and Random Forest methods. The highest accuracy is Random Forest reaching 96% precision, 97% recall, and F1 score 96% with an average computation time of 0.233980 as an output in making dengue disease early detection applications. The parameter used is gini index.

Keywords: Classification, Decision Tree, Random Forest