

FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG

Program Studi: Sistem Informasi

Tugas Akhir Sarjana Komputer

Semester Ganjil Tahun 2019/2020

**IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI
PENYAKIT LIVER BERBASIS WEB**

Mochammad Ibnu Fachruddin

NIM : 16.01.55.0020

Abstrak

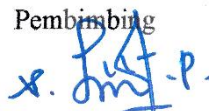
Penyakit liver merupakan peradangan didalam sel-sel hati, yang umumnya diakibatkan oleh infeksi (parasit, bakteri dan virus.), lemak yang berlebih, konsumsi obat-obatan dan alkohol berlebih, serta penyakit autoimmune. Pada stadium awal, penyakit liver cenderung sulit untuk di temukan. Hati dapat menjaga fungsinya bahkan ketika ada bagian yang mengalami kerusakan. Sehingga penanganan penyakit liver pada stadium awal akan memperpanjang usia pasien, untuk itu diperlukan suatu sistem komputerasi yang mampu membantu dalam memprediksi penyakit liver.

Dalam penelitian ini prediksi dilakukan menggunakan teknik data mining dan metode klasifikasi dengan algoritma *decision tree* C4.5. Serta menggunakan data Indian Liver Patient Dataset (ILPD Dataset) yang diperoleh dari database *UCI Machine Learning Repository*, dengan jumlah data sebanyak 583 record dengan 11 atribut. Metode pengembangan sistem yang digunakan terdiri dari Studi Literatur, Pengumpulan Data, Analisis Model Sistem, Perancangan Sistem, Pembuatan Program serta, Pengujian. Sedangkan perancangan sistem menggunakan UML (*Unified Modeling Language*) dan pembuatan website menggunakan PHP sebagai bahasa pemrograman dan MySQL sebagai media database.

Penelitian ini telah menghasilkan aplikasi untuk memprediksi penyakit liver menggunakan metode Decision Tree dengan Algoritma C4.5 berbasis web. Aplikasi dapat mengklasifikasi data pasien penyakit liver untuk menghasilkan keputusan apakah pasien termasuk penderita penyakit liver atau bukan. Pengujian algoritma C4.5 dilakukan menggunakan metode *10 fold cross validation*, dan dihasilkan rata-rata pola prediksi penyakit liver yang terbentuk sebanyak 63 pola (rules) dengan atribut *direct_bilirubin* sebagai atribut berpengaruh dan atribut *sgpt*, *tot_bilirubin*, *albumin* serta *alkphos* sebagai atribut node ke-2. Sedangkan rata-rata nilai akurasi yang dihasilkan sebesar 58,64% dengan akurasi tertinggi 74,14% dan akurasi terendah 41,38%.

Kata Kunci : Penyakit liver, Data Mining, Klasifikasi, C4.5, cross validation

Pembimbing



(Dra. Sulastri, M.Kom)

FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG

Program Studi: Sistem Informasi

Tugas Akhir Sarjana Komputer

Semester Ganjil Tahun 2019/2020

**IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI
PENYAKIT LIVER BERBASIS WEB**

Mochammad Ibnu Fachruddin

NIM : 16.01.55.0020

Abstrak

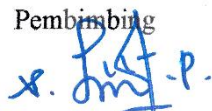
Liver disease is inflammation in the cells of the liver, which is generally caused by infection (parasites, bacteria and viruses,), excess fat, excessive consumption of drugs and alcohol, and autoimmune disease. In the early stages, liver disease tends to be difficult to find. The liver can maintain its function even when parts are damaged. So that the treatment of liver disease at an early stage will extend the life of the patient, for that we need a computerized system that can help in predicting liver disease.

In this study predictions were made using data mining techniques and classification methods with the decision tree C4.5 algorithm. And using Indian Liver Patient Dataset (ILPD Dataset) data obtained from the UCI Machine Learning Repository database, with a total of 583 records with 11 attributes. System development methods used consist of Literature Studies, Data Collection, Analysis of System Models, System Design, Programming and Testing. While the system design uses UML (Unified Modeling Language) and website creation uses PHP as a programming language and MySql as a media database.

This research has produced applications to predict liver disease using the Decision Tree method with the web-based C4.5 Algorithm. The application can classify the data of liver disease patients to produce a decision whether the patient is a sufferer of liver disease or not. The C4.5 algorithm was tested using the 10 fold cross validation method, and an average of 63 patterns of liver disease prediction were formed with the direct_bilirubin attribute as the influential attribute and the sgpt, tot_bilirubin, albumin and alkphos attributes as the node attribute 2 While the average accuracy value is 58.64% with the highest accuracy of 74.14% and the lowest accuracy of 41.38%.

Kata Kunci : Penyakit liver, Data Mining, Klasifikasi, C4.5, cross validation

Pembimbing



(Dra. Sulastri, M.Kom)