



**IMPLEMENTASI ALGORITMA C5.0 MENGGUNAKAN *FEATURE SELECTION CHI-SQUARE* TERHADAP DATA PASIEN UNTUK DETEKSI
DINI PENYAKIT HEPATITIS C**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

MAHMUD

NIM 2011016110001

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
MARET 2024**



IMPLEMENTASI ALGORITMA C5.0 MENGGUNAKAN *FEATURE SELECTION CHI-SQUARE* TERHADAP DATA PASIEN UNTUK DETEKSI DINI PENYAKIT HEPATITIS C

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

MAHMUD

NIM 2011016110001

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

MARET 2024

SKRIPSI

IMPLEMENTASI ALGORITMA C5.0 MENGGUNAKAN *FEATURE SELECTION CHI-SQUARE* TERHADAP DATA PASIEN UNTUK DETEKSI DINI PENYAKIT HEPATITIS C

Oleh:

Mahmud

NIM. 2011016110001

Telah dipertahankan di depan Dosen Penguji pada tanggal 26 Maret 2024.

Susunan Dosen Penguji:

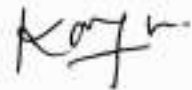
Pembimbing I



Irwan Budiman, S. T., M. Kom.

NIP. 197703252008121001

Dosen Penguji I



Dwi Kartini, S.Kom., M.Kom.

NIP. 198704212012122003

Pembimbing II



Fatma Indriani, S. T., M. I. T., Ph. D.

NIP. 198404202008122004

Dosen Penguji II



Mohammad Reza Faizal, S.Si S.T M.T Ph. D.

NIP. 197612202008121001

Banjarmasin, 26 Maret 2024

Koordinator Program Studi Ilmu Komputer



Irwan Budiman, S. T., M. Kom.

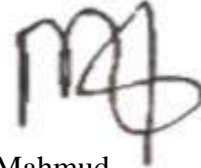
NIP. 197703252008121001

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam jurnal ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 26 Maret 2024

Yang Menyatakan,

A handwritten signature in black ink, appearing to be the name 'Mahmud' written in a stylized, cursive script.

Mahmud

NIM. 2011016110001

3. ABSTRAK

4. IMPLEMENTASI ALGORITMA C5.0 MENGGUNAKAN *FEATURE SELECTION CHI – SQUARE* TERHADAP DATA PASIEN UNTUK DETEKSI DINI PENYAKIT HEPATITIS C

(Oleh: Mahmud; Pembimbing: Irwan Budiman, S.T., M.Kom. dan Fatma Indriani, S. T., M. I. T., Ph. D.; 2024; halaman)

Hepatitis C, sebuah tantangan kesehatan global yang signifikan, mempengaruhi 71 juta orang di seluruh dunia, dengan komplikasi parah seperti sirosis dan karsinoma hepatoseluler. Meskipun prevalensi dan ketersediaannya dalam tes diagnostik cepat (RDT), kebutuhan akan metode deteksi dini yang akurat masih tetap penting. Penelitian ini bertujuan untuk meningkatkan akurasi klasifikasi virus hepatitis C dengan mengintegrasikan algoritma C5.0 dengan pemilihan fitur *Chi-Square*, mengatasi keterbatasan pendekatan diagnostik saat ini dan berpotensi mengurangi kesalahan diagnostik. Penelitian ini mengeksplorasi pengembangan model pembelajaran mesin untuk prediksi hepatitis C, memanfaatkan kumpulan data yang tersedia untuk umum dari Kaggle. Ini mencakup teknik pra-pemrosesan seperti pengkodean label, penanganan nilai yang hilang, normalisasi, pemilihan fitur, pengembangan model, dan evaluasi untuk memastikan kemanjuran dan keakuratan model dalam mendiagnosis hepatitis C. Temuan penelitian ini mengungkapkan bahwa penerapan pemilihan fitur *Chi-Square* secara signifikan meningkatkan efektivitas algoritma pembelajaran mesin. Secara khusus, kombinasi algoritma C5.0 dan pemilihan fitur *Chi-Square* menghasilkan akurasi luar biasa sebesar 96,75%, melampaui tolok ukur penelitian sebelumnya. Hal ini menyoroti sinergi yang kuat antara teknik pemilihan fitur tingkat lanjut dan algoritme pembelajaran mesin dalam meningkatkan presisi diagnostik. Studi ini secara meyakinkan menunjukkan bahwa pembelajaran mesin adalah alat yang efektif untuk mendeteksi hepatitis C, dan menunjukkan potensi untuk meningkatkan akurasi diagnostik secara signifikan. Sebagai rekomendasi di masa mendatang, penggunaan AutoML disarankan untuk mengotomatiskan pemilihan algoritme optimal secara berkala, sehingga menjanjikan peningkatan lebih lanjut dalam kemampuan deteksi.

Kata kunci: Algoritma C5.0, Seleksi Fitur, Penyakit Hepatitis C, Pembelajaran Mesin

5. ABSTRACT

IMPLEMENTATION OF THE C5.0 ALGORITHM USING THE CHI – SQUARE SELECTION FEATURE ON PATIENT DATA FOR EARLY DETECTION OF HEPATITIS C DISEASE

(By: Mahmud; Supervisors: Irwan Budiman, S.Kom., M.Kom. and Fatma Indriani, S. T., M. I. T., Ph. D.; 2023; page)

Hepatitis C, a significant global health challenge, affects 71 million people worldwide, with severe complications such as cirrhosis and hepatocellular carcinoma. Despite its prevalence and availability in rapid diagnostic tests (RDTs), the need for accurate early detection methods remains critical. This research aims to enhance hepatitis C virus classification accuracy by integrating the C5.0 algorithm with Chi-Square feature selection, addressing the limitations of current diagnostic approaches and potentially reducing diagnostic errors. This research explores the development of a machine learning model for hepatitis C prediction, utilizing a publicly available dataset from Kaggle. It encompasses preprocessing techniques such as label encoding, handling missing values, normalization, feature selection, model development, and evaluation to ensure the model's efficacy and accuracy in diagnosing hepatitis C. The findings of this study reveal that implementing Chi-Square feature selection significantly enhances the effectiveness of machine learning algorithms. Specifically, the combination of the C5.0 algorithm and Chi-Square feature selection yielded a remarkable accuracy of 96.75%, surpassing previous research benchmarks. This highlights the potent synergy between advanced feature selection techniques and machine learning algorithms in improving diagnostic precision. The study conclusively demonstrates that machine learning is an effective tool for detecting hepatitis C, showcasing the potential to enhance diagnostic accuracy significantly. As a future recommendation, adopting AutoML is suggested to periodically automate the selection of the optimal algorithm, promising further improvements in detection capabilities.

Keywords: C5.0 Algorithm, Feature Selection, Hepatitis C Disease, Machine learning

KATA PENGANTAR

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan jurnal yang berjudul *“Implementation of C5.0 Algorithm using Chi-Square Feature Selection for Early Detection of Hepatitis C Disease”* untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

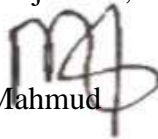
Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada semua pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan jurnal ini, adapun yang dimaksud adalah sebagai berikut:

1. Diri saya sendiri yang tidak pernah patah semangat walaupun banyak menemui kesulitan baik disebabkan oleh diri sendiri maupun hal lain.
2. Keluarga besar yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian jurnal ini.
3. Bapak Irwan Budiman, S.T., M.Kom. selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
4. Ibu Fatma Indriani, S.T., M. I. T., Ph. D. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
5. Hasri Akbar Awal Rozaq selaku mentor MSBI saya yang turut membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
6. Oktay Yildiz selaku hoca mentor saya tekni komputer di turkey yang turut membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
7. Bapak Irwan Budiman, S.T., M. Kom. selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau jurnal ini dapat diselesaikan.
8. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
9. Teman Futsal Bedom 20, dan Habibie yang selalu mendukung dalam berbagai hal.
10. Teman-teman dan sahabat-sahabat keluarga Ilmu Komputer yang memberikan dukungan dan selalu mengingatkan serta mendoakan dalam proses mengerjakan jurnal.

11. Serta semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian jurnal ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 26 Maret 2024


Mahmud