

PENERAPAN METODE CART DAN C5.0 DALAM KLASIFIKASI TUBERKULOSIS INTESTINAL: STUDI KASUS DI RSUPN DR. CIPTO MANGUNKUSUMO

NAZUWA AULIA



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan Metode CART dan C5.0 dalam Klasifikasi Tuberkulosis Intestinal: Studi Kasus di RSUPN Dr. Cipto Mangunkusumo” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Nazuwa Aulia
G1401211033

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

NAZUWA AULIA. Penerapan Metode CART dan C5.0 dalam Klasifikasi Tuberkulosis Intestinal: Studi Kasus di RSUPN Dr. Cipto Mangunkusumo. Dibimbing oleh AAM ALAMUDI dan GERRY ALFA DITO.

Tuberkulosis masih menjadi salah satu penyebab utama kematian secara global, dengan Indonesia menempati peringkat kedua dan menyumbang sekitar 10% dari total kasus dunia. Salah satu bentuk tuberkulosis yang sering sulit didiagnosis adalah tuberkulosis intestinal (TBI) karena gejalanya menyerupai penyakit Crohn. Kemiripan pada gejala klinis, temuan endoskopi, dan hasil radiologi sering kali menyebabkan kesalahan diagnosis, sehingga diperlukan metode deteksi dini yang akurat dan efektif. Penelitian ini bertujuan untuk membandingkan kinerja dua model klasifikasi, yaitu CART dan C5.0, dalam mengidentifikasi kasus tuberkulosis intestinal berdasarkan data klinis dan biomarker dari 143 pasien terduga TBI di RSUPN Dr. Cipto Mangunkusumo. Hasil penelitian menunjukkan bahwa model CART memiliki rata-rata sensitivitas yang lebih tinggi yaitu sebesar 64% dibandingkan model C5.0 yang sebesar 52%, yang berarti model CART lebih baik dalam mendeteksi kasus TBI. Sebaliknya, model C5.0 memiliki rata-rata spesifisitas yang lebih tinggi yaitu 81%, dibandingkan CART yang sebesar 78%, sehingga model C5.0 lebih baik dalam mengenali kasus non-TBI. Dari segi *balanced accuracy*, model CART lebih unggul dengan nilai 71%, sementara C5.0 dengan nilai 67%. Struktur pohon keputusan pada model CART juga lebih sederhana, dengan jumlah cabang dan variabel pemisah yang lebih sedikit. Peubah yang digunakan dalam model CART meliputi laju endap darah dan kadar Interferon Gamma dalam feses. Peubah yang digunakan C5.0 adalah laju endap darah, kadar human beta-defensin 2 dalam feses, gejala berkeringat di malam hari, basofil, kadar Interferon-Gamma dalam feses, batuk, riwayat kontak dengan pasien TB, positif IGRA, gejala penurunan berat badan, kadar ADA dalam darah, eosinofil, demam, dan darah dalam feses.

Kata kunci: C5.0, CART, diagnosis, pohon keputusan, tuberkulosis intestinal



ABSTRACT

NAZUWA AULIA. Application of CART and C5.0 Methods for Classifying Intestinal Tuberculosis: A Case Study at Dr. Cipto Mangunkusumo National General Hospital. Supervised by AAM ALAMUDI and GERRY ALFA DITO.

Tuberculosis remains one of the leading causes of death globally, with Indonesia ranking second and contributing approximately 10% of global cases. One form of tuberculosis that is often challenging to diagnose is intestinal tuberculosis (ITB), due to its clinical resemblance to Crohn's disease. The similarity in clinical symptoms, endoscopic findings, and radiological results frequently leads to misdiagnosis, emphasizing the need for accurate and effective early detection methods. This study aims to compare the performance of two classification models, namely CART and C5.0, in identifying ITB cases using clinical data and biomarkers from 143 suspected ITB patients at Dr. Cipto Mangunkusumo National General Hospital. The results show that the CART model achieved a higher average sensitivity of 64% compared to 52% in the C5.0 model, indicating better performance in detecting ITB. In contrast, the C5.0 model demonstrated a higher average specificity of 81% compared to 78% in CART, suggesting better recognition of non-ITB cases. In terms of balanced accuracy, CART slightly outperformed C5.0 with a score of 71% versus 67%. Furthermore, the decision tree structure of the CART model was simpler, with fewer branches and splitting variables. The predictor variables used in the CART model included erythrocyte sedimentation rate and fecal Interferon-Gamma levels, while the C5.0 model utilized a broader range of predictors, including erythrocyte sedimentation rate, fecal human beta-defensin 2 levels, night sweats, basophils, fecal Interferon-Gamma levels, cough, history of contact with TB patients, positive IGRA, weight loss, blood ADA levels, eosinophils, fever, and blood in the stool.

Keywords: C5.0, CART, decision tree, diagnosis, tuberculosis intestinal



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

**PENERAPAN METODE CART DAN C5.0 DALAM KLASIFIKASI
TUBERKULOSIS INTESTINAL: STUDI KASUS DI RSUPN
DR. CIPTO MANGUNKUSUMO**

NAZUWA AULIA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Judul Skripsi : Penerapan Metode CART dan C5.0 dalam Klasifikasi
Tuberkulosis Intestinal: Studi Kasus di RSUPN Dr. Cipto
Mangunkusumo

Nama : Nazuwa Aulia

NIM : G1401211033

Disetujui oleh

Pembimbing 1:

Ir. Aam Alamudi M.Si.



Pembimbing 2:

Gerry Alfa Dito S.Si., M.Si.



Diketahui oleh

Ketua Program Studi:

Dr. Bagus Sartono, S.Si., M.Si.

NIP 197804112005011002



Tanggal Ujian:
15 Agustus 2025

Tanggal Lulus:



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Januari 2025 sampai bulan Agustus 2025 ini ialah klasifikasi, dengan judul “Penerapan Metode CART dan C5.0 dalam Klasifikasi Tuberkulosis Intestinal: Studi Kasus di RSUPN Dr. Cipto Mangunkusumo”.

Terima kasih penulis ucapkan kepada semua pihak yang telah memberikan doa, dukungan dan bantuan dalam penyusunan karya ilmiah ini, diantaranya :

1. Bapak Ir. Aam Alamudi M.Si. dan Bapak Gerry Alfa Dito S.Si, M.Si. selaku dosen pembimbing yang telah membimbing dan banyak memberi saran serta masukan dalam penulisan karya ilmiah ini.
2. Ibu Laily Nissa Atul Mualifah, S.Si., M.Si. selaku dosen penguji yang telah memberikan saran serta masukan dalam penulisan karya ilmiah ini.
3. Ibu Dr. Yenni Angraini S.Si., M.Si. selaku dosen moderator kolokium dan Ibu Akbar Rizki S.Stat., M.Si. selaku dosen moderator seminar hasil atas saran serta masukan dalam penulisan karya ilmiah ini
4. Seluruh dosen dan staff akademik Program Studi Statistika dan Sains Data yang telah banyak memberi ilmu dan membantu selama proses perkuliahan selesai serta staff Tata Usaha Program Studi Statistika dan Sains Data yang telah membantu penulis dalam proses administrasi perkuliahan.
5. Orang tua penulis yaitu Bapak B.A Sanuri dan Ibu Jamilah serta kedua adik penulis yaitu Irzi Yazid dan Qamira Nur Syifa yang selalu memberikan doa dan dukungan kepada penulis.
6. Teman – teman terdekat penulis yaitu Adis, Salwa, Nurul , Hanifa, Amal, Zafira, Pingkan, Rani, Dindana, Azra, Kaylila, Mega, dan Runi serta teman – teman satu bimbingan yang telah menjadi tempat berbagi ilmu, pengalaman serta dukungan selama penulis menempuh pendidikan di IPB.
7. Ibu Nuri Dyah Indrasari dan rekan - rekan yang data hasil penelitiannya dapat penulis gunakan dalam karya ilmiah ini.
8. Semua pihak yang tidak dapat disebutkan satu persatu yang telah memberikan doa, bantuan, dan dukungan dalam penyelesaian karya ilmiah ini

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2025

Nazuwa Aulia



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR ISI

DAFTAR TABEL	xii
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	3
II TINJAUAN PUSTAKA	4
2.1 Classification and Regression Tree (CART)	4
2.2 C5.0	6
2.3 Peubah Biomarker dan Klinis	8
2.4 Nested Cross-Validation	12
2.5 Class Weight	12
2.6 Evaluasi Model	13
III METODE	15
3.1 Data	15
3.2 Prosedur Analisis	16
IV HASIL DAN PEMBAHASAN	19
4.1 Eksplorasi Data	19
4.2 Analisis Klasifikasi CART	28
4.3 Analisis Klasifikasi C5.0	31
4.4 Perbandingan Kinerja Model CART dan C5.0	34
V SIMPULAN DAN SARAN	36
5.1 Simpulan	36
5.2 Saran	36
DAFTAR PUSTAKA	37
LAMPIRAN	40
RIWAYAT HIDUP	44



DAFTAR TABEL

1	Ketepatan hasil klasifikasi	13
2	Daftar peubah	15
3	Evaluasi model CART berdasarkan nilai CP	28
4	Performa model klasifikasi CART pada setiap <i>fold</i>	29
5	Evaluasi model C5.0 berdasarkan nilai CF	31
6	Performa model klasifikasi C5.0 pada setiap <i>fold</i>	32
7	Perbandingan kinerja model CART dan C5.0	34

DAFTAR GAMBAR

1	Distribusi hasil diagnosis pasien terduga TBI	19
2	Distribusi jenis kelamin pasien berdasarkan diagnosis	19
3	Distribusi kelompok usia pasien berdasarkan diagnosis	20
4	Distribusi riwayat tuberkulosis berdasarkan diagnosis : (a) Riwayat kontak dengan penderita tuberkulosis dan (b) Riwayat memiliki tuberkulosis.	20
5	Distribusi gejala gastrointestinal berdasarkan diagnosis : (a) Nyeri perut non-spesifik, (b) Diare kronis, (c) Sembelit, dan (d) Diare dan sembelit.	21
6	Distribusi gejala yang berkaitan dengan kondisi feses berdasarkan diagnosis : (a) Darah dalam feses, (b) Lendir dalam feses, dan (c) Darah dan lendir dalam feses.	22
7	Distribusi gejala tubuh berdasarkan diagnosis : (a) Penurunan berat badan, (b) Demam, (c) Keringat berlebih di malam hari, dan (d) Kehilangan nafsu makan.	23
8	Distribusi gejala batuk berdasarkan diagnosis	24
9	Distribusi biomarker tinja dan darah berdasarkan diagnosis : (a) Kadar Interferon-gamma, (b) Kadar human beta-defensin-2, (c) Jumlah leukosit, (d) Basofil, (e) Eosinofil, dan (f) Neutrofil.	25
10	Distribusi biomarker darah berdasarkan diagnosis : (a) Limfosit, (b) Monosit, (c) Rasio neutrofil terhadap limfosit. (d) Rasio monosit terhadap limfosit, (e) Laju endap darah, dan (f) Aktivitas adenosine deaminase.	26
11	Distribusi hasil IGRA terhadap diagnosis	27
12	Pohon klasifikasi CART	30
13	Pohon klasifikasi C5.0	33

DAFTAR LAMPIRAN

1	Lampiran 1 Perhitungan bobot masing-masing kelas	40
2	Lampiran 2 Rincian kesalahan klasifikasi per fold dan rata-rata pada model CART	40
3	Lampiran 3 Rincian kesalahan klasifikasi per fold dan rata-rata pada model C5.0	40
4	Lampiran 4 Ringkasan aturan klasifikasi model C5.0	40