

**PERBANDINGAN PERFORMA ALGORITMA *DECISION*
TREE C4.5 DAN *RANDOM FOREST* UNTUK
DETEKSI PENYAKIT DIABETES**

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

MARSA QURRATU'AIN
NIM : 09021282126084

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA**

2025

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN PERFORMA ALGORITMA DECISION TREE C4.5 DAN RANDOM FOREST UNTUK DETEKSI PENYAKIT DIABETES

Sebagai salah satu syarat untuk penyelesaian studi di
Program Studi S1 Teknik Informatika

Oleh:

MARSA QURRATU'AIN

09021282126084

Pembimbing 1 : Alvi Syahrini Utami, S.Si., M.Kom
NIP. 197812222006042003

Mengetahui

Ketua Jurusan Teknik Informatika



Hadipurnawan Satria, Ph.D
198004182020121001

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jumat tanggal 13 Juni 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Marsa Qurratu'ain

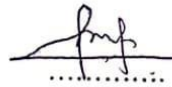
NIM : 09021282126084

Judul : Perbandingan Performa Algoritma *Decision Tree* C4.5 dan *Random Forest* Untuk Deteksi Penyakit Diabetes

dan dinyatakan LULUS.

1. Ketua Penguji

Prof. Dr. Ermatita, M.Kom.
NIP. 196709132006042001



2. Penguji I

Muhammad Ourhanul Rizqie, S.Kom., M.T., Ph.D
NIP. 198712032022031006



3. Pembimbing

Alvi Syahrini Utami, S.Si., M.Kom
NIP. 197812222006042003



Mengetahui,
Ketua Jurusan Teknik Informatika,

Hadipurnawan Satria, M.Sc., Ph.D.
NIP. 198004182020121001

HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan dibawah ini:

Nama : Marsa Qurratu'ain
NIM : 09021282126084
Program Studi : Teknik Informatika
Judul : Perbandingan Performa Algoritma *Decision Tree*
C4.5 dan *Random Forest* Untuk Deteksi
Penyakit Diabetes

Hasil Pengecekan *iThenticate/Turnitin* : 5%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.



Palembang, 24 Juni 2025



ZDC28AMX314587082

Marsa Qurratu'ain
NIM. 09021282126084

“Sesuatu yang kita rasa sulit akan jadi mudah bila kita yakin bahwa Allah SWT itu Maha Penolong.”

“Di balik setiap kesuksesan anak, ada doa dan perjuangan orang tua.”

“Keberhasilan adalah kemampuan untuk melewati dan mengatasi dari satu kegagalan ke kegagalan berikutnya tanpa kehilangan semangat.” – Winston Churchill

Kupersembahkan karya tulis ini kepada:

- ALLAH SWT & Nabi Muhammad SAW
- Orang Tua dan Saudaraku
- Teman Seperjuanganku

ABSTRACT

Diabetes is a chronic disease with a continuously increasing global prevalence. Early detection poses a major challenge because symptoms often appear only when the condition is already severe. The classification of diabetes plays a crucial role in recognizing the disease early to allow for faster and more accurate interventions. According to data from the World Health Organization in 2020, over the past three decades, the prevalence of diabetes has significantly increased in various countries. This research is conducted to compare the performance of two *classification* methods, that are *Decision Tree C4.5* and *Random Forest*, in *classifying* diabetes *disease* based on patients' health data records. This research includes the design stage of *classification* system, software implementation, and the performance evaluation of each method. the dataset used is from processed health data records, even though the quality is considered less than ideal. The test results show that the *Random Forest* algorithm achieved the highest accuracy of 89.61%, while the *C4.5 Decision Tree* algorithm reached an accuracy of 88.63%. In addition, the *Random Forest* algorithm also demonstrated a higher precision compared to the *C4.5 Decision Tree*. This is due to *Random Forest* using an ensemble approach that employs many *Decision Trees*, resulting in more accurate predictions.

Keywords: Diabetes Classification, Decision Tree C4.5, Random Forest, Data mining, Random under sampling, Min-Max Scaling

ABSTRAK

Diabetes adalah penyakit kronis dengan prevalensi global yang terus meningkat. Deteksi dini menjadi tantangan utama karena seringkali gejala baru muncul saat kondisi sudah parah. Klasifikasi diabetes berperan penting dalam mengenali penyakit sejak awal guna memungkinkan intervensi yang lebih cepat dan tepat. Berdasarkan data dari *World Health Organization* pada tahun 2020, dalam tiga dekade terakhir, prevalensi diabetes mengalami peningkatan yang signifikan di berbagai negara. Penelitian ini bertujuan untuk membandingkan kinerja dua metode klasifikasi, yaitu *Decision Tree C4.5* dan *Random Forest*, dalam mendeteksi penyakit diabetes berdasarkan data kesehatan pasien. Penelitian ini mencakup tahap perancangan sistem klasifikasi, implementasi perangkat lunak, serta evaluasi kinerja masing-masing metode. Dataset yang digunakan berasal dari data kesehatan yang telah diproses, meskipun kualitasnya dinilai kurang ideal. Hasil pengujian menunjukkan bahwa algoritma *Random Forest* memperoleh akurasi tertinggi sebesar 89.61%, sementara algoritma *Decision Tree C4.5* mencapai akurasi sebesar 88.63%. Selain itu, algoritma *Random Forest* juga menunjukkan tingkat ketepatan (*precision*) yang lebih tinggi dibandingkan *Decision Tree C4.5*. Hal ini disebabkan karena *Random Forest* menggunakan pendekatan *ensemble* yang menggunakan banyak *tree* keputusan, sehingga menghasilkan prediksi yang lebih akurat.

Kata Kunci: Deteksi Diabetes, *Decision Tree C4.5*, *Random Forest*, *Data mining*, *Random under sampling*, *Min-Max Scaling*

KATA PENGANTAR

Segala puji dan rasa syukur penulis haturkan kepada Allah SWT atas limpahan nikmat, rahmat, dan petunjuk-Nya, sehingga penulis dapat menyusun dan menyelesaikan Tugas Akhir yang berjudul “Perbandingan Performa Kinerja Algoritma Decision Tree C4.5 dan Random Forest Untuk Deteksi Penyakit Diabetes”. Tugas Akhir ini disusun sebagai bagian dari pemenuhan syarat kelulusan pada Program Studi Teknik Informatika Reguler, Universitas Sriwijaya.

Selama proses penyusunan Tugas Akhir ini, penulis menerima banyak bantuan dari berbagai pihak, baik berupa bimbingan, arahan, nasihat, maupun pemikiran yang sangat membantu. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Allah SWT atas rahmat dan nikmat-Nya penulis dapat menyelesaikan skripsi ini dengan baik.
2. Kedua orangtua dan adikku yang selalu mendoakan, memberikan semangat serta bantuan baik materi maupun non materi sehingga laporan ini dapat terselesaikan.
3. Bapak Prof. Erwin, S.Si., M.Si sebagai Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Hadipurnawan Satria, M.Sc., Ph.D. sebagai Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Alvi Syahrini Utami, S.Si., M.Kom selaku Dosen Pembimbing Tugas Akhir yang telah bersedia meluangkan waktunya untuk memberikan bimbingan, saran dan kritik yang membangun dalam

penyusunan laporan tugas akhir ini.

6. Bapak Kanda Januar Miraswan, S.Kom., MT. Selaku dosen pembimbing akademik saya.
7. Dosen – dosen Universitas Sriwijaya terutama dosen Teknik Informatika.
8. Teman-teman seperjuangan terkhusus kepada Haji Risky, Keysha Zascha Medina, Nadhifa Faiza, Nadya Aura Binoto Gultom, dan Shirin Nur Fadhillah.
9. dan pihak-pihak lain yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa skripsi ini belum sepenuhnya sempurna. Oleh sebab itu, penulis sangat mengharapkan masukan dan saran yang bersifat membangun demi perbaikan di masa yang akan datang. Harapannya, skripsi ini dapat memberikan manfaat bagi pembaca serta menjadi acuan dalam pengembangan penelitian pada bidang terkait.

Palembang, 19 Juni 2025

Hormat penulis,

Marsa Qurratu'ain

DAFTAR ISI

Halaman

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI	iii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iv
HALAMAN MOTTO DAN PERSEMBAHAN	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xv
BAB I PENDAHULUAN.....	I-1
1.1. Pendahuluan	I-1
1.2. Latar Belakang Masalah	I-1
1.3. Rumusan Masalah.....	I-3
1.4. Tujuan Penelitian.....	I-4
1.5. Manfaat Penelitian.....	I-4
1.6. Batasan Masalah.....	I-4
1.7. Sistematika Penulisan	I-5
BAB I. PENDAHULUAN	I-5
BAB II. KAJIAN LITERATUR	I-5
BAB III. METODOLOGI PENELITIAN.....	I-6
BAB IV. PENGEMBANGAN PERANGKAT LUNAK.....	I-6
BAB V. HASIL DAN ANALISIS PENELITIAN.....	I-6
BAB VI. KESIMPULAN DAN SARAN	I-6
1.8. Kesimpulan.....	I-6
BAB II KAJIAN LITERATUR	II-1
2.1. Pendahuluan.....	II-1
2.2. Landasan Teori.....	II-1
2.2.1. Diabetes.....	II-1
2.2.2 Klasifikasi	II-3
2.2.3 <i>Decision Tree</i>	II-4

2.2.4 <i>Random Forest</i>	I-6
2.2.5 <i>Confusion Matrix</i>	II-9
2.2.6 <i>Random Under Sampling (RUS)</i>	II-11
2.3. Penelitian Lain yang Relevan	II-13
2.4. Kesimpulan	II-14
BAB III METODOLOGI PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Pengumpulan Data.....	III-1
3.3 Tahapan Penelitian	III-5
3.4 Metode Pengembangan Perangkat Lunak	III-10
3.5 Manajemen Proyek Penelitian	III-14
3.6 Kesimpulan.....	III-17
BAB IV PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1 Pendahuluan.....	IV-1
4.2 Rational Unified Process.....	IV-1
4.3 Kesimpulan	IV-94
BAB V HASIL DAN ANALISIS PENELITIAN.....	V-1
5.1 Pendahuluan	V-1
5.2 Data Hasil Percobaan / Penelitian	V-1
5.3 Analisis Hasil Pengujian.....	V-7
5.4 Kesimpulan.....	V-7
BAB VI KESIMPULAN DAN SARAN	VI-1
6.1 Pendahuluan	VI-1
6.2 Kesimpulan.....	VI-1
6.3 Saran	VI-2
DAFTAR PUSTAKA.....	xvii
LAMPIRAN.....	

BAB I

PENDAHULUAN

1.1. Pendahuluan

Bab ini memuat pembahasan mengenai latar belakang masalah, perumusan masalah, tujuan penelitian, manfaat yang diharapkan dari penelitian, serta batasan-batasan yang ditetapkan. Secara umum, bab ini memberikan gambaran menyeluruh tentang keseluruhan isi dan arah penelitian.

1.2. Latar Belakang Masalah

Diabetes adalah gangguan kesehatan yang terjadi ketika tubuh tidak dapat merespons atau menggunakan insulin secara efektif. Jika tidak ditangani dalam waktu lama, kadar gula darah yang tinggi dapat menyebabkan kerusakan pada organ-organ tubuh, memicu disfungsi jaringan, dan meningkatkan risiko komplikasi serius hingga mengancam jiwa ('Syamsudin, 'Handhayani dan 'Syarifudin, Isnaini, 2024). Oleh karena itu, perlu dilakukan deteksi dini penyakit diabetes mengingat diabetes merupakan salah satu penyakit kronis yang mematikan, serta penyakit diabetes di Indonesia semakin meningkat.

Diabetes termasuk penyakit kronis yang mengalami peningkatan prevalensi secara global. Tantangan utama dalam penanganan penyakit ini adalah deteksi dini, karena banyak penderita tidak menyadari keberadaan diabetes hingga muncul gejala yang parah. Proses klasifikasi penyakit diabetes berperan penting dalam mengenali kondisi ini sejak awal, sehingga memungkinkan dilakukannya intervensi yang lebih tepat dan efisien

(Nurussakinah dan Faisal 2023).

Perkembangan teknologi modern mempermudah manusia dalam mengakses informasi dan melakukan prediksi terhadap penyakit. Teknologi ini juga berperan dalam mendukung proses pengobatan dan upaya pencegahan diabetes secara lebih efektif melalui penerapan machine learning dengan metode klasifikasi ('Syamsudin, 'Handhayani dan 'Syarifudin, Isnaini, 2024). Algoritma klasifikasi yang akan digunakan penulis untuk memprediksi penyakit diabetes adalah algoritma *Decision Tree* dan algoritma *Random Forest*.

Dalam sebuah penelitian Jogo, Biddinika, dan Fadlil (2023) untuk Klasifikasi Penyakit Diabetes dengan algoritma *Decision Tree* dan *Naive Bayes*. Berdasarkan perhitungan kedua algoritma tersebut, algoritma *Decision Tree* mempunyai hasil yang lebih baik dibandingkan dengan algoritma *Naive Bayes*.

Pada penelitian yang berjudul, 'Perbandingan Performa Algoritma *Random Forest Classifier* dan *Naive Bayes* pada Penyakit Diabetes Melitus' oleh Ibrahim (2022), hasil menunjukkan bahwa algoritma *Random Forest* memiliki akurasi sebesar 91,0%, lebih tinggi dibandingkan *Naive Bayes* yang mencapai 87,4%, sehingga *Random Forest Classifier* lebih unggul dalam memprediksi penyakit diabetes. Begitu pula dengan penelitian yang dilakukan oleh (Aprilia et al., 2021), Hasil yang diperoleh menunjukkan bahwa algoritma klasifikasi *Random Forest* mampu memberikan prediksi penyakit diabetes dengan tingkat akurasi yang lebih tinggi dibandingkan algoritma

Support Vector Machine (SVM) dan *Naive Bayes*.

Meskipun *Random Forest* merupakan pengembangan dari *Decision Tree*, perbandingan antara keduanya tetap relevan dalam penelitian untuk mengevaluasi performa, interpretabilitas, dan efisiensi model. Pemilihan antara C4.5 dan *Random Forest* tergantung pada kebutuhan spesifik dari masalah yang dihadapi, seperti kebutuhan akan interpretabilitas, akurasi, dan sumber daya yang tersedia. *Random Forest* sering dijadikan pembanding dengan algoritma *Decision Tree C4.5* dalam penelitian karena keduanya memiliki perbedaan mendasar dalam pendekatan dan performa. Perbedaan ini memungkinkan peneliti untuk mengevaluasi bagaimana pendekatan *ensemble* (*Random Forest*) dibandingkan dengan pohon tunggal (C4.5) dalam berbagai konteks data.

Berdasarkan uraian di atas, algoritma *Decision Tree* dan *Random Forest* memiliki performa yang lebih baik dibandingkan dengan algoritma yang lain. Sehingga pada penelitian ini penulis akan membandingkan performa *Decision Tree* dan *Random Forest* untuk deteksi penyakit diabetes.

1.3. Rumusan Masalah

Rumusan-rumusan masalah yang akan dibahas dalam penelitian ini berdasarkan latar belakang masalah yang telah dibahas sebelumnya adalah:

1. Bagaimana mengimplementasikan algoritma *Decision Tree* dan *Random Forest* untuk melakukan klasifikasi penyakit diabetes?
2. Bagaimana hasil perbandingan performa klasifikasi diabetes

dengan algoritma *Decision Tree* dan *Random Forest*?

1.4. Tujuan Penelitian

Tujuan dilakukannya penelitian ini adalah sebagai berikut:

1. Membuat model machine learning klasifikasi penyakit diabetes menggunakan algoritma *Decision Tree* dan *Random Forest*.
2. Mengetahui hasil perbandingan performa algoritma *Decision Tree* dan *Random Forest* untuk klasifikasi penyakit diabetes untuk mengetahui model mana yang lebih baik untuk klasifikasi penyakit diabetes.

1.5. Manfaat Penelitian

Adapun manfaat yang diharapkan bisa diambil dari penelitian ini adalah:

1. Diharapkan dapat membantu dalam melakukan deteksi dini penyakit diabetes.
2. Dapat dijadikan sebagai acuan atau referensi untuk penelitian-penelitian selanjutnya.

1.6. Batasan Masalah

1. Dataset yang digunakan bersumber dari situs website *Kaggle* tepatnya <https://www.kaggle.com/datasets/iammustafatz/diabetes-Prediction-dataset/data> yaitu *Diabetes Prediction dataset* yang terdiri dari 8 variabel prediktor ('gender', 'age', 'hypertension', 'heart_disease', 'smoking_history', 'bmi', 'HbA1c_level', dan

- 'blood_glucose_level')* serta 1 *class / target* (diabetes atau tidak diabetes).
2. Metrik evaluasi yang digunakan untuk mengevaluasi performa model di antaranya *accuracy, confusion matrix, precision, recall, f1- score, dan ROC-AUC score*.
 3. Pembuatan model *machine learning* menggunakan *python programming language*.
 4. Beberapa teknik *balancing data* yang digunakan adalah *Random undersampling*.

1.7. Sistematika Penulisan

Sistematika penulisan skripsi ini mengikuti standar penulisan skripsi yaitu sebagai berikut:

BAB I. PENDAHULUAN

Bab ini membahas latar belakang masalah, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, sistematika penulisan, serta kesimpulan terkait judul penelitian yang akan dilaksanakan.

BAB II. KAJIAN LITERATUR

Bab ini mengulas berbagai artikel dan jurnal penelitian sebelumnya yang digunakan sebagai acuan dan dasar teori dalam penelitian ini, mencakup proses Data *Pre-processing*, prinsip kerja algoritma Decision Tree dan Random Forest, teknik-teknik yang digunakan dalam pembuatan model machine learning, serta beberapa literatur terkait lainnya yang relevan dengan penelitian.

BAB III. METODOLOGI PENELITIAN

Bab ini menguraikan langkah-langkah dan prosedur yang akan dijalankan selama penelitian, termasuk pengumpulan data, analisis data, pengolahan data, dan perancangan perangkat lunak. Penjelasan tiap tahap didasarkan pada kerangka kerja yang telah ditetapkan.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini menguraikan analisis, tahapan, dan perancangan perangkat lunak yang akan dibuat dengan memanfaatkan algoritma Decision Tree dan Random Forest. Desain pengembangan perangkat lunak tersebut direncanakan menggunakan framework Rational Unified Process (RUP).

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini menyajikan hasil pengujian sesuai dengan tahapan yang telah direncanakan. Analisis hasil pengujian tersebut digunakan sebagai dasar dalam menarik kesimpulan dari penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Bab ini memaparkan kesimpulan yang diperoleh dari pembahasan pada bab-bab sebelumnya, serta menyajikan saran-saran yang diberikan berdasarkan hasil penelitian yang telah dilaksanakan.

1.8. Kesimpulan

Bab ini telah menguraikan latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, serta sistematika penelitian yang akan menjadi landasan pemahaman dalam penelitian ini.

DAFTAR PUSTAKA

- Amaliah, S, Nusrang, dan Aswi. 2022. Penerapan Metode *Random Forest* Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng. *Journal of Statistics and Its Application on Teaching and Research (online)*, vol. 4, No. 2, 121 – 127.
- Apriliah, W., Kurniawan, I., Baydhowi, M., & Haryati, T. 2021. Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi *Random Forest*. *Sistemasi*, 10 (1), 163.
<https://doi.org/10.32520/stmsi.v10i1.1129>
- Arifin, M. F. dan Devi, F. 2017. Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus: PT Atria Artha Persada. Teknik Informatika, Universitas Mercu Buana, Jakarta.
IncomTech, Jurnal Telekomunikasi dan Komputer, Vol. 8, no. 2.
- Aziz, W. A. 2021. Implementasi Metode *Random Forest* pada Klasifikasi Data Ulasan Konsumen Perusahaan (Studi Kasus: Aplikasi KAI Access). Skripsi. Teknik Informatika Fakultas Sains dan Teknologi UIN Syarif Hidayatullah. Jakarta.
- BK Taruna. 2024. Algoritma Klasifikasi pada *Data mining*.
<https://bktaruna.uma.ac.id/algoritma-klasifikasi-pada-data-mining/>
- Buani. 2016. Optimasi Algoritma *Naive Bayes* dengan Menggunakan Algoritma Genetika untuk Prediksi Kesuburan (*Fertility*). *Evolusi*, 4 (1), 55 – 64.
<https://doi.org/10.2311/evo.v4i1.277>
- Fadli, F dan Belsana, B. 2019. Penerapan *Decision Tree* Menggunakan Algoritma C4.5 Untuk Deteksi Demam Berdarah pada RS. IMC Bintaro. *IJSE – Indonesian Journal on Software Engineering*, Vol. 5, No. 1, Juni, 75 – 86.
- Hidayatulloh, Kurnia, P. K., dan Dimas, F. H. P. 2022. Implementasi Algoritma C4.5 Untuk Memprediksi Capaian Pembelajaran Daring (Studi Kasus Siswa MAN 3 Blitar). *Jurnal Algoritme*, Vol. 3, No. 1, Oktober. Hal 33 – 47.
- Ibrahim, F. 2024. Perbandingan Performa Algoritma *Random Forest Classifier* dan *Naive Bayes* pada Penyakit Diabetes Melitus. Skripsi Program Sarjana (S1) Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah. Jakarta.
- Irawan, E., dan Wahono, R. S. 2015. Penggunaan Random Under Sampling untuk Penanganan Ketidakseimbangan kelas pada Prediksi Cacat *Software* Berbasis *Neural Network*. *Journal of Software Engineering*, 1 (2), 92 – 100.
- Kementerian Kesehatan RI. 2020. Infodatin tetap produktif, cegah, dan atasi Diabetes Melitus 2020. Pusat Data dan Informasi Kementerian Kesehatan RI. pp. 1-10.

- Lestari, Zulkarnain, ST. Aisyah Sijid. 2021. Diabetes Melitus: *Review* Etiologi, Patofisiologi, Gejala, Penyebab, Cara Pemeriksaan, Cara Pengobatan dan Cara Pencegahan. Program Studi Biologi, Fakultas Sains dan Teknologi, UIN Alauddin Makassar.
<http://journal.uin-alauddin Makassar>
- Maulana, D. 2019. Implementasi Algoritma *Naïve Bayes* Untuk Klasifikasi Penderita Penyakit Jantung Di Indonesia Menggunakan Rapid Miner. Jurnal SIGMA, Vol. 10. Hlm. 191 – 197.
- Mukminin, A., & Riana, D. 2017. Komparasi Algoritma C4.5, *Naive Bayes* dan *Neural Network* untuk Klasifikasi Tanah. Jurnal Informatika, 4 (1), 21 - 31.
- Nurussakinah. 2023. Klasifikasi Penyakit Diabetes Menggunakan Algoritma *Decision Tree*. Jurnal Informatika 10 (2): 143-149.
- Pambudi, R. H., Budi, D. S., dan Indriati. 2018. Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah berdasarkan Faktor Eksternal. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 2, No. 7, Juli, hlm. 2637 – 2643.
- Purnajaya, A. R. Dan F. D. Hanggara. 2021. Perbandingan Performa Teknik *Sampling Data* untuk Klasifikasi Pasien Terinfeksi Covid-19 Menggunakan Rontgen Dada. *Journal of Applied Informatics and Computing* (JAIC), vol. 5, No. 1, Juli 2021, pp. 37 – 42.
e – ISSN: 2548 – 6861.
- Rosandy, T. 2016. Perbandingan Metode *Naive Bayes Classifier* dengan Metode *Decision Tree* C4.5 Untuk Menganalisa Kelancaran Pembiayaan (Studi Kasus: KSPPS/BMT Al-Fadhila. Jurnal Teknologi Informasi Magister Darmajaya, 2(1)), 52 – 62.
- Saifullah, M. Zarlis, Zakaria, Rahmat. 2017. Analisa Terhadap Perbandingan Algoritma *Decision Tree* dengan Algoritma *Random Tree* untuk *Pre-Processing* Data. Jurnal Sains Komputer & Informatika (J-SAKTI), Volume (1) No.2, September.
- Samodro, M. M. J., M. K. Biddinika, dan A. Fadlil. 2023. Klasifikasi Penyakit Diabetes dengan Algoritma *Decision Tree* dan *Naive Bayes*. Magister Informatika Universitas Ahmad Dahlan, Yogyakarta.
- Sari, D. P. 2019. Diabetes Melitus. Fakultas Ilmu Kesehatan UPN Veteran. Jakarta.
<https://www.honestdocs.id/penyakit-diabetes>
- Suyanto. 2017. *Data mining* untuk Klasifikasi dan Klasterisasi Data. INFORMATIKA. Bandung.
- Syamsudin, T., T. Handayani, M. I. Syaifudin. 2024. Perbandingan Klasifikasi Penyakit Diabetes Menggunakan Metode *Machine Learning*. Jurnal Ilmu Komputer dan Sistem Informasi Fakultas Teknologi Informasi Universitas Tarumanegara. Jil. 12 No. 1.

Trisnawati, Sri, Widarsa, dan Ketut Suastika. 2013. Faktor Risiko Diabetes melitus Tipe 2 Pasien Rawat Jalan di Puskesmas Wilayah Kecamatan Denpasar Selatan. *Public Health and Preventive Medicine Archive (PHPMA)* 2013, Volume 1, Number 1: 69-73

Trivusi. 2022. <https://www.trivusi.web.id/2022/08/algoritma-random-forest.html?m=1>

World Health Organization (WHO). 2020. *Fact Sheet of Diabetes*.

Yolanda, A. R. 2022. Perbandingan Metode *Random Forest* dengan Metode *K-Nearest Neighbor* untuk Klasifikasi Penyakit Diabetes Melitus [skripsi]. Fakultas Ilmu Komputer. Universitas Sriwijaya. Palembang.

Zuhairah, A. 2022. Penerapan Algoritma *Random Forest*, *Support Vector Machines* (SVM) dan *Gradient Boosted Tree* (GBT) untuk Deteksi Penipuan (*Fraud Detection*) pada Transaksi Kartu Kredit. Skripsi. Teknik Informatika, Fakultas Sains dan Teknologi. UIN Syarif Hidayatullah. Jakarta.