

**PERBANDINGAN *DECISION TREE* ALGORITMA ID3 DAN
C4.5 UNTUK KLASIFIKASI KANKER SERVIKS
BERDASARKAN *SAMPLE BOOTSTRAPPING***

SKRIPSI

Oleh:

TRI AJENG FADILAH

08011381924084



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA
2023**

LEMBAR PENGESAHAN

**PERBANDINGAN *DECISION TREE* ALGORITMA ID3 DAN
C4.5 UNTUK KLASIFIKASI KANKER SERVIKS
BERDASARKAN *SAMPLE BOOTSTRAPPING***

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Matematika**

Oleh

TRI AJENG FADILAH

NIM. 08011381924084

Pembimbing Kedua



Dra. Ning Eliyati, M.Pd
NIP.195911201991022001

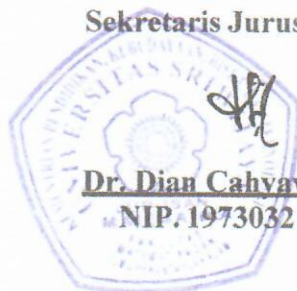
Indralaya, Agustus 2023
Pembimbing Utama



Novi Rustiana Dewi, S.Si., M.Si
NIP.197011131996032002

Mengetahui,
a.n. Ketua

Sekretaris Jurusan Matematika



Dr. Dian Cahyawati S. S.S., M.Si
NIP. 197303212000122001

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan dibawah ini:

Nama Mahasiswa : Tri Ajeng Fadilah
NIM : 08011381924084
Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan strata satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat dalam skripsi ini yang berasal dari penulis lain baik yang dipublikasikan atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 08 Agustus 2023

Penulis



Tri Ajeng Fadilah

NIM. 08011381924084

HALAMAN PERSEMBAHAN

MOTTO

“Sabarlah pada setiap prosesmu, karena Tuhan tau kapasitas yang kamu miliki dan Tuhanmu percaya bahwa kau mampu”

(Tri Ajeng Fadilah)

“Percayalah pada Tuhanmu, Tuhan selalu tau apa yang kamu inginkan”

(Mamas)

Kupersembahkan skripsi ini untuk :

- ❖ Allah Subhanahu Wa Ta’ala,**
- ❖ Kedua orang tuaku tersayang,**
- ❖ Kakak – kakak ku,**
- ❖ Keluarga besarku,**
- ❖ Guru dan dosenku,**
- ❖ Sahabat-sahabatku,**
- ❖ Almamaterku**

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Dengan mengucapkan Alhamdulillah, segala puji bagi Allah Subhanahu wa Ta'ala yang telah memberikan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Perbandingan *Decision Tree* Algoritma ID3 dan C4.5 untuk Klasifikasi Kanker Serviks Berdasarkan *Sample Bootstrapping***” sebagai salah satu syarat memperoleh gelar sarjana sains studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Penulis menyadari bahwa proses pembuatan skripsi ini sebagai proses pembelajaran yang sangat berharga yang tak lepas dari kekurangan dan keterbatasan. Dengan segala hormat dan kerendahan hati, penulis mengucapkan terima kasih dan penghargaan yang tulus penulis sampaikan kepada :

1. Kedua orang tuaku **M. Afrizal** dan **Sayuti** yang sudah membantu baik secara moril maupun materil, terkhususnya Ibu saya yang tak pernah lelah mendidik, membimbing, menasehati, dan mendukung serta tak henti untuk mendoakan anaknya.
2. Bapak **Prof. Hermansyah, S.Si., M.Si., Ph.D.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
3. Bapak **Drs. Sugandi Yahdin, M.M** selaku ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

4. Ibu **Dr. Dian Cahyawati Sukanda, M.Si** selaku Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
5. Ibu **Novi Rustiana Dewi, S.Si., M.Si** selaku Dosen Pembimbing Utama yang telah bersedia meluangkan waktu, tenaga, pikiran untuk memberikan bimbingan dan arahan dengan penuh perhatian, pengertian, dan kesabaran sehingga skripsi ini dapat diselesaikan.
6. Ibu **Dra. Ning Eliyati, M.Pd** selaku Dosen Pembimbing Kedua yang telah bersedia meluangkan waktu, tenaga, pikiran untuk memberikan bimbingan dan arahan dengan penuh perhatian, pengertian, dan kesabaran sehingga skripsi ini dapat diselesaikan.
7. Bapak **Drs. Endro Setyo Cahyono, M.Si** dan Ibu **Endang Sri Kresnawati, S.Si., M.Si** selaku Dosen Pembahas yang telah memberikan tanggapan, kritik, dan saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.
8. Bapak **Drs. Ali Amran, M.T** selaku Dosen Pembimbing Akademik yang sangat baik telah memberikan saran, membimbing, membantu, dan mengarahkan urusan akademik penulis setiap semester.
9. **Seluruh Dosen di Jurusan Matematika FMIPA** yang telah memberikan ilmu, nasihat, motivasi, serta bimbingan selama proses perkuliahan.
10. Pak **Irwansyah** selaku admin dan Ibu **Hamidah** selaku pegawai tata usaha Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah banyak membantu penulis selama perkuliahan.

11. **Seluruh guru** yang telah memberikan ilmu yang bermanfaat hingga mengantarkan penulis pada pendidikan ini.
12. Kakakku **Aldino Octariansyah, S.IP** dan **Dwi Alfiqih, S.IP** yang selalu mendoakan, memberikan nasihat, dan keluarga besar yang selalu mendukung penulis.
13. **SM** dan **sahabat – sahabat terbaik** yang pernah penulis miliki dari masa sekolah hingga saat ini, terima kasih sudah menjadi orang-orang hebat yang selalu membantu, mengingatkan, dan mendukung dengan tulus selama ini.
14. **Keluarga Matematika 2019, BPH Himastik Gelora Karya dan Akselerasi**, dan **rekan-rekan** selama perkuliahan.
15. **UKM Harmoni** yang menjadi wadah untuk mengembangkan minat dan bakat saya serta telah mewarnai kehidupan di dalam dunia kampus.
16. Kakak-kakak tingkat Angkatan 2017 dan 2018 serta adik-adik tingkat Angkatan 2020 dan 2021, terima kasih atas segala kebaikan dan bantuan.
17. Semua pihak yang tidak dapat penulis sebutkan satu per satu. Semoga segala kebaikan yang diberikan mendapatkan balasan terbaik dari Allah.

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya dan seluruh pihak yang membutuhkan.

Indralaya, Agustus 2023

Penulis

**COMPARISON OF *DECISION TREE* ID3 ALGORITHM AND
C4.5 FOR CLASSIFICATION OF CERVICAL CANCER
BASED ON *SAMPLE* BOOTSTRAPPING**

By:

TRI AJENG FADILAH

08011381924084

ABSTRACT

Cervical cancer is one of the malignant tumors in women that attacks the female reproductive organs. There are various methods of preventing cervical cancer, one of which is conducting a pap smear screening test which is done by taking or collecting cervical cell *samples* on the cervical uterine wall. This study discusses the comparison of classification methods on cervical cancer data using the *decision tree* method of ID3 and C4.5 algorithms based on *sample* bootstrapping. *Decision tree* is a method in machine learning used for data classification and prediction that uses *decision tree* creation to find solutions to a problem. The precision produced by the ID3 algorithm is 98,44%, 95,05%, 94,90%, and 94,96% while for C4.5 it is 95,72%, 87,05%, 86,63%, and 86,72%. Based on the results obtained it was concluded that the ID3 algorithm is better than the C4.5 algorithm.

Keywords : Cervical cancer, *Decision tree*, ID3, C.45, *Sample* bootstrapping

**PERBANDINGAN *DECISION TREE* ALGORITMA ID3 DAN
C4.5 UNTUK KLASIFIKASI KANKER SERVIKS
BERDASARKAN *SAMPLE BOOTSTRAPPING***

Oleh:

TRI AJENG FADILAH

08011381924084

ABSTRAK

Kanker serviks merupakan salah satu tumor ganas pada wanita yang menyerang organ reproduksi wanita. Terdapat berbagai macam metode pencegahan kanker serviks salah satunya melakukan uji skrining *pap smear* yang dilakukan dengan pengambilan atau pengumpulan sampel sel serviks pada dinding rahim bagian serviks. Penelitian ini membahas perbandingan metode klasifikasi pada data kanker serviks dengan menggunakan metode *decision tree* algoritma ID3 dan C4.5 berdasarkan *sample bootstrapping*. *Decision tree* merupakan metode dalam pembelajaran mesin yang digunakan untuk klasifikasi data dan prediksi yang menggunakan pembuatan pohon keputusan untuk menemukan pemecahan suatu masalah. Ketepatan yang dihasilkan oleh algoritma ID3 adalah 98,44%, 95,05%, 94,90%, dan 94,96% sedangkan untuk C4.5 adalah 95,72%, 87,05%, 86,63%, dan 86,72%. Berdasarkan hasil yang didapat disimpulkan bahwa algoritma ID3 lebih baik dibandingkan algoritma C4.5.

Kata Kunci : Kanker serviks, *Decision tree*, ID3, C.45, *Sample bootstrapping*

DAFTAR ISI

HALAMAN JUDUL	1
LEMBAR PENGESAHAN	ii
HALAMAN PERSEMBAHAN.....	iii
KATA PENGANTAR.....	iv
ABSTRACT	vii
ABSTRAK	viii
Daftar Isi	ix
Daftar Gambar	xi
Daftar Tabel.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Kanker Serviks.....	6
2.2 Pap smear	7
2.3 Klasifikasi.....	8
2.4 Diskritisasi Data	9
2.5 <i>Decision tree</i>	10
2.6 Algoritma <i>Iterative Dichotomiser 3 (ID3)</i>	11
2.7 Algoritma C4.5	13
2.8 <i>Sample Bootstrapping</i>	14
2.9 Himpunan <i>Fuzzy</i>	15
2.10 Fungsi Keanggotaan <i>Fuzzy</i>	15
2.11 <i>Multi-Class Confusion matrix</i>	18
BAB III METODOLOGI PENELITIAN	21
3.1 Tempat	21
3.2 Waktu	21

3.3	Alat dan Bahan	21
3.4	Metode Penelitian.....	21
BAB IV HASIL DAN PEMBAHASAN.....		25
4.1	Deskripsi Data	25
4.2	Diskritisasi Data	26
4.3	<i>Sample Bootstrapping</i>	29
4.4	Klasifikasi Menggunakan Metode <i>Decision tree</i> ID3	33
4.4.1	Perhitungan Manual Metode <i>Decision Tree</i> ID3.....	33
4.4.2	Evaluasi <i>Machine Learning</i> Metode <i>Decision Tree</i> ID3.....	38
4.5	Klasifikasi Menggunakan Metode <i>Decision Tree</i> C4.5	46
4.5.1	Contoh Perhitungan Manual Metode <i>Decision Tree</i> C4.5.....	46
4.5.2	Evaluasi <i>Machine Learning</i> Metode <i>Decision Tree</i> C4.5.....	49
4.6	Analisis Perbandingan Hasil.....	52
BAB V KESIMPULAN DAN SARAN		53
5.1	Kesimpulan	53
5.2	Saran	53
DAFTAR PUSTAKA		55

Daftar Gambar

Gambar 1. Kurva S-Pertumbuhan	16
Gambar 2. Kurva S-Penyusutan	17
Gambar 3. Kurva Segitiga.....	18
Gambar 4. Ilustrasi Pohon Keputusan Algoritma ID3	37
Gambar 5. Ilustrasi Pohon Keputusan Algoritma C4.5.....	49

Daftar Tabel

Tabel 2.1 Tabel <i>Multi-Class Confusion matrix</i>	19
Tabel 3.1 Karakteristik sel <i>pap smear</i>	22
Tabel 3.2 Variabel Prediktor Penelitian	23
Tabel 4.1 <i>Dataset</i> Penelitian	25
Tabel 4.2 Diskritisasi Data.....	26
Tabel 4.3 Hasil Diskritisasi Data	29
Tabel 4.4 Data Untuk Menghitung Manual <i>Sample Bootstrapping</i>	30
Tabel 4.5 Pengambilan <i>Sample</i> Secara Acak.....	30
Tabel 4.6 Data Baru pada Data Contoh <i>Sample Bootstrapping</i>	32
Tabel 4.7 <i>Dataset</i> Setelah <i>Sample Bootstrapping</i>	32
Tabel 4.8 Data Contoh untuk <i>Decision tree</i> ID3.....	33
Tabel 4.9 Hasil Perhitungan <i>Entropy</i> masing – masing Variabel	35
Tabel 4.10 Nilai Gain dari Contoh Data	36
Tabel 4.11 <i>Confusion matrix Multiclass</i> Metode <i>Decision tree</i> ID3	38
Tabel 4.12 Nilai Split Information	47
Tabel 4.13 Hasil Gain Ratio.....	48
Tabel 4.14 <i>Confusion matrix Multiclass</i> Metode <i>Decision tree</i> C4.5.....	50
Tabel 4.15 Hasil Visualisasi <i>Confusion matrix</i>	51
Tabel 4.16 Hasil Evaluasi <i>Confusion matrix</i>	51
Tabel 4.17 Perbandingan Hasil Ketepatan ID3 dan C4.5	52

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kanker serviks merupakan salah satu tumor ganas pada wanita yang menyerang organ reproduksi wanita, yaitu selaput vagina dan sekitar selaput saluran serviks, atau biasa disebut bagian bawah antara alat kelamin wanita dan rahim (Erniawati *et al.*, 2020). Terdapat berbagai macam metode pencegahan kanker serviks salah satunya melakukan uji skrining *pap smear* (Endriani *et al.*, 2021). Pemeriksaan *pap smear* dilakukan pengambilan atau pengumpulan sampel sel serviks pada dinding rahim bagian serviks, setelah itu sampel ini akan diteliti menggunakan mikroskop selanjutnya didiagnosis secara manual oleh ahli patologi. Dengan hasil yang diperoleh yaitu sel yang normal atau abnormal (Wiastini *et al.*, 2019). Pada pemeriksaan kanker serviks menggunakan analisis *pap smear* rawan terjadi kesalahan, terutama kesalahan manusia atau *human error* hal ini karena proses analisis dilakukan secara manual dan menggunakan waktu yang cukup lama (William *et al.*, 2019).

Dataset yang digunakan merupakan hasil sitologi *pap smear* serviks dari Rumah Sakit Universitas Herlev. Beberapa peneliti telah melakukan penelitian terhadap *dataset* Herlev ini, seperti membahas klasifikasi sel serviks pada citra *pap smear* berdasarkan fitur bentuk deskriptor regional dan fitur tekstur *uniform rotated local binary pattern* yang menggunakan metode klasifikasi *Fuzzy K-Nearest Neighbour* (K-NN) memperoleh nilai akurasi sebesar 91,59% pada dua kategori dan 67,89% pada tujuh kelas sel (Sholik & Fatichah, 2017).

Klasifikasi adalah teknik yang digunakan dalam data mining, untuk menganalisis data yang kemudian dijadikan kedalam beberapa kategori sesuai dengan variabel – variabel yang terkait (Limantara & Nababan, 2019). Salah satu metode yang digunakan dalam klasifikasi yaitu metode *decision tree*. *Decision tree* merupakan metode dalam pembelajaran mesin yang digunakan untuk klasifikasi data dan prediksi yang menggunakan pembuatan pohon keputusan untuk menemukan pemecahan suatu masalah (Lutfi, 2018).

Algoritma *decision tree* yang sering digunakan yaitu *Iterative Dichotomiser 3* (ID3) dan *C4.5* (Novyana *et al.*, 2012). *Decision tree* yang menggunakan algoritma ID3 dirancang bahwa dalam pembuatan pohon keputusan hanya dapat dilakukan dengan data yang memiliki tipe kategorik sehingga tidak bisa digunakan untuk data yang memiliki tipe numerik (Lutfi, 2018). *Decision tree* algoritma C4.5 merupakan algoritma pembaharuan dari ID3, pembaharuan yang dilakukan yaitu pada algoritma C4.5 dirancang mampu menangani data yang memiliki tipe numerik (Lutfi, 2018). Terdapat beberapa penelitian mengenai perbandingan akurasi dari kedua algoritma *decision tree*. Penelitian berjudul Perbandingan Kinerja Algoritma ID3 Dan C4.5 Dalam Klasifikasi Spam-Mail pada penelitian tersebut membahas perbandingan kinerja dari kedua algoritma yang memberikan hasil bahwa algoritma ID3 memiliki hasil yang lebih baik dibandingkan dengan C4.5 (Defiyanti, 2019). Pada penelitian lain yang membandingkan performa dari *decision tree* algoritma ID3, C4.5 dan CART. Pada penelitian tersebut menggunakan *dataset* readmisi pasien diabetes yang menyimpulkan bahwa algoritma C4.5 menghasilkan nilai akurasi paling tinggi

yaitu 54,13% sedangkan ID3 memiliki nilai akurasi rendah sebesar 47,82% (Yusa *et al.*, 2016).

Metode *sample bootstrapping* merupakan metode yang akan dipakai untuk mengurangi jumlah data training yang diproses selanjutnya menghitung validitas dari data training. Keakuratan hasil klasifikasi diverifikasi dengan metode *confusion matrix* (Agus *et al.*, 2017). Pada penelitian terdahulu peneliti menggunakan metode *sample bootstrapping* untuk menentukan nilai akurasi *naïve bayes* pada klasifikasi citra *pap smear* dan menghasilkan 85,24% untuk klasifikasi tujuh kelas dan untuk dua kelas memiliki nilai akurasi 93,24% (Dewi & Sariasih, 2019). Ada juga yang membahas optimasi penggunaan *sample bootstrapping* dan *particle swarm optimization* untuk membantu keakuratan algoritma C4.5 dan *naïve bayes*, sehingga didapat nilai akurasi 94,44% untuk C4.5 dan 83,09% untuk *naïve bayes* (Agustian & Bisri, 2019). Berdasarkan penelitian ini dapat dikatakan bahwa *sample bootstrapping* mampu meningkatkan nilai akurasi klasifikasi.

Berdasarkan penelitian sebelumnya terlihat bahwa dua algoritma, ID3 dan C4.5 memiliki performa yang baik berdasarkan nilai akurasi yang didapat untuk melakukan klasifikasi pada suatu *dataset*. Maka penulis bertujuan untuk melakukan penelitian perbandingan performa ketepatan melalui nilai akurasi pada *dataset* klasifikasi hasil *pap smear* untuk mengidentifikasi kanker serviks menggunakan *decision tree* C4.5 dan ID3 berdasarkan *sample bootstrapping*. Untuk mengukur kinerja tingkat kesesuaian klasifikasi penelitian ini dibatasi dengan adanya hasil akurasi, presisi, *recall* dan *f1-score*.

1.2 Perumusan Masalah

Rumusan masalah pada penelitian ini yaitu sebagai berikut.

1. Bagaimana model klasifikasi kanker serviks yang didapat berdasarkan metode *decision tree* algoritma ID3 dan C4.5?
2. Bagaimana nilai akurasi, presisi, *recall*, dan *F1-score* pada klasifikasi kanker serviks yang didapat berdasarkan metode *decision tree* algoritma ID3 dan C4.5?
3. Bagaimana perbandingan performa *decision tree* antara algoritma ID3 dan C4.5 ?

1.3 Batasan Masalah

Batasan masalah penelitian ini yaitu sebagai berikut :

1. Data yang digunakan berasal dari *dataset* hasil ekstraksi citra *pap smear* kanker serviks Rumah Sakit Universitas Herlev.
2. Data yang digunakan sebanyak 917 dengan 20 atribut yang terbagi menjadi 7 kelas.
3. Menggunakan 3 fungsi keanggotaan *fuzzy* yaitu kurva-S penyusutan, kurva-S pertumbuhan dan kurva segitiga.
4. Data numerik diolah menggunakan diskritisasi pendekatan fungsi keanggotaan *fuzzy*.
5. Pada penelitian ini dibantu dengan metode *sample bootstrapping*.
6. Presentasi ketepatan klasifikasi dibatasi dengan adanya nilai akurasi, presisi, *recall* dan *F1-score*.

1.4 Tujuan Penelitian

Tujuan penelitian ini yaitu sebagai berikut :

1. Mendapatkan model *confusion matrix* klasifikasi kanker serviks yang didapat berdasarkan metode *decision tree* algoritma ID3 dan C4.5
2. Mendapatkan nilai akurasi, presisi, *recall*, dan *F1-score* klasifikasi kanker serviks yang didapat berdasarkan metode *decision tree* algoritma ID3 dan C4.5
3. Mendapatkan perbandingan performa *decision tree* antara algoritma ID3 dan C4.3.

1.5 Manfaat Penelitian

Manfaat penelitian ini yaitu sebagai berikut :

1. Untuk meningkatkan wawasan serta kemampuan peneliti dalam menyelesaikan karya ilmiah
2. Sebagai tambahan ilmu pengetahuan bagi para pembaca dan masyarakat mengenai klasifikasi kanker serviks.
3. Dapat membantu tenaga medis dalam mendiagnosis pasien kanker serviks lebih awal.
4. Menjadi bahan referensi untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- Agus, T., Adib, S. M., & Karomi, A. (2017). Penerapan Metode *Sample Bootstrapping* Untuk Meningkatkan Performa Knearest Neighbor Pada *Dataset* Berdimensi Tinggi. *Ic-Tech, Xii*(1, April), 9–14.
- Agustian, A. A., & Bisri, A. (2019). Data Mining Optimization Using *Sample Bootstrapping* And Particle Swarm Optimization In The Credit Approval Classification. *Indonesian Journal Of Artificial Intelligence And Data Mining*, 2(1), 18–27. <https://doi.org/10.24014/ijaidm.V2i1.6299>
- Aldino, A. A., & Sulistiani, H. (2020). *Decision tree C4 . 5 Algorithm For Tuition Aid Grant Program Classification (Case Study : Department Of Information System , Universitas Teknokrat Indonesia)*. 7(1), 40–50.
- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *Ilkom Jurnal Ilmiah*, 10(2), 160–165. <https://doi.org/10.33096/ilkom.V10i2.303.160-165>
- Anwariyah, K., & Sofyan, D. (2019). Pemodelan Segmentasi Sel Epitel Serviks Pada Citra Digital Pap Smear. *Jtim : Jurnal Teknologi Informasi Dan Multimedia*, 1(2), 96–105. <https://doi.org/10.35746/jtim.V1i2.25>
- Argina, A. M. (2020). Penerapan Metode Klasifikasi K-Nearest Neighbor Pada *Dataset* Penderita Penyakit Diabetes. *Indonesian Journal Of Data And Science*, 1(2), 29–33. <https://doi.org/10.33096/ijodas.V1i2.11>
- Budiyantara, A., Wijaya, A. K., Gunawan, A., & Rolland, M. (2021). Analisis Data Mining Hotel Booking Menggunakan Model Id3. *Jbase - Journal Of Business And Audit Information Systems*, 4(1), 1–12. <https://doi.org/10.30813/jbase.V4i1.2728>
- Davvaz, B., & Mukhlash, I. (2021). *Himpunan Fuzzy Dan Rough Sets*. 18(1), 79–94.
- Defiyanti, S. And D. L. C. P. (2019). Perbandingan Kinerja Algoritma Id3 Dan C4.5 Dalam Klasifikasi Spam-Mail Sofi. *Gunadarma University Repository*, 9(2), 62–68.
- Dewi, Y. N., & Sariasih, F. A. (2019). *Metode Sample Bootstrapping Untuk Meningkatkan Performa*. 12(1).
- Endriani, T. I., Budiana, I. N. G., Mahendra, I. N. B., & Surya, I. G. N. H. W. (2021). Karakteristik Hasil Pap Smear Menurut Bethesda Di Rsup Sanglah Denpasar Tahun 2016-2017. *Jurnal Medika Udayana*, 10(1), 77–81. <https://ojs.unud.ac.id/index.php/eum>

- Erniawati, Purnamasari, D., Risnayanti, Jusni, & Kamaruddin, M. (2020). Tingkat Pengetahuan Remaja Putri Tentang Kanker Serviks Di Desa Bialo Kecamatan Gantarang Kabupaten Bulukumba Tahun 2020. *Medika Alkhairaat: Jurnal Penelitian Kedokteran Dan Kesehatan*, 2(3), 131–135. <https://doi.org/10.31970/ma.v2i3.61>
- Faizah, T. (2021). Perbandingan Algoritma C4.5 Dan Id3 Untuk Prediksi Ketepatan Waktu Lulus Mahasiswa. *Jatisi (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(2), 980–990. <https://doi.org/10.35957/jatisi.v8i2.593>
- Farida, F., & Nurhidayah, F. O. (2017). Pengetahuan Kanker Serviks Dalam Tindakan Melakukan Pap Smear Pada Wanita Usia Subur. *Journal Of Nursing Practice*, 1(1), 40–47. <https://doi.org/10.30994/jnp.v1i1.21>
- Fatmasari, A. R. (2017). Pap Smear Berbasis Kombinasi Fitur The Cervical Cancer Classification Through Feature Combination Based Pap Smear Images Andi Rezky Fatmasari Program Pascasarjana Universitas Hasanuddin. *The Cervical Cancer Classification Through Feature Combination Based Pap Smear Images*.
- Globocan. (2020). Cancer Incident In Indonesia. *International Agency For Research On Cancer*, 858, 1–2. <https://gco.iarc.fr/today/data/factsheets/populations/360-indonesia-factsheets.pdf>
- Hadianto, N., Novitasari, H. B., Rahmawati, A., & Komputer, I. (2019). *Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode Neural Network*. 15(2), 163–170. <https://doi.org/10.33480/pilar.v15i2.658>
- Hayuningtyas. (2021). *Hang Tuah Medical Journal, Profil Kejadian Kanker Serviks Di Rspal Ramelan Periode Januari-Desember 2018 Vol. 19 No. 1*. 19(1), 25–40. www.journal-medical.hangtuah.ac.id
- Imelda, F., & Santosa, H. (2020). Deteksi Dini Kanker Serviks Pada Wanita. In *Journal Endurance*. <https://repository.usu.ac.id/bitstream/handle/123456789/2300/fulltext.pdf?sequence=1&isallowed=Y>
- Khairina, N. (2019). *Analisis Fungsi Keanggotaan Fuzzy Tsukamoto Dalam Menentukan Status Kesehatan Tubuh Seseorang*. 1, 19–24.
- Limantara, C., & Nababan, D. (2019). Klasifikasi Kepribadian Menggunakan Algoritma *Decision tree* Berdasarkan Ten Item Personality Inventory. *Jurnal Coreit: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 5(1), 8–12.
- Lutfi, M. (2018). Klasifikasi Kualitas Mutu Jagung Dengan Menggunakan

Metode *Decision tree* Pada Dinas Pertanian Bojonegoro Classification Of Corn Quality Using The *Decision tree* Method At The Bojonegoro Agriculture Office. *Jurnal Agromix*, 8(1), 2599–3003.

- Manliguez, C. (2016). *Generalized Confusion matrix For Multiple Classes Generalized Confusion matrix For Multiple Classes The Total Numbers Of False Negative (Tfn), False Positive (Tfp), And True Negative (Ttn) For Each Class I Will Be Calculated Based On The Generalized*. November, 2–4. <https://doi.org/10.13140/Rg.2.2.31150.51523>
- Marina Yanti, Eka Rahmawati, Putu Lusita, & Tuti Farida. (2021). Hubungan Pendidikan, Paritas Dan Dukungan Kader Dengan Deteksi Dini Kanker Serviks Pada Pemeriksaan Iva Di Puskesmas Nagaswidak Palembang Tahun 2021. *Jurnal Kebidanan : Jurnal Medical Science Ilmu Kesehatan Akademi Kebidanan Budi Mulia Palembang*, 11(2), 193–204. <https://doi.org/10.35325/Kebidanan.V11i2.269>
- Nadiah, N., Soim, S., & Sholihin, S. (2022). Implementation Of *Decision tree* Algorithm Machine Learning In Detecting Covid-19 Virus Patients Using Public Datasets. *Indonesian Journal Of Artificial Intelligence And Data Mining*, 5(1), 37. <https://doi.org/10.24014/Ijaidm.V5i1.17054>
- Nasrullah, A. H. (2021). Implementasi Algoritma *Decision tree* Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, 7(2), 45–51. <https://doi.org/10.35329/Jiik.V7i2.203>
- Nlerum, P. A., Komputer, I., Harcourt, U. P., Komputer, I., & Federal, U. (2020). *Iterative Dichotomizer 3 (Id3) Decision tree : Algoritma Pembelajaran Mesin Untuk Klasifikasi Data Dan Analisis Prediktif Edward E . 6495*, 514–521.
- Novyana, A., Martaleli, B., & Hendra, K. (2012). *Penerapan Algoritma Id3 Dan C.45 Dalam Hubungan Data Awal Masuk Mahasiswa Dengan Prestasi Akademik*. 3, 10.
- Ogheneovo, E. E., & Nlerum, P. A. (2020). Iterative Dichotomizer 3 (Id3) *Decision tree: A Machine Learning Algorithm For Data Classification And Predictive Analysis*. *International Journal Of Advanced Engineering Research And Science*, 7(4), 514–521. <https://doi.org/10.22161/Ijaers.74.60>
- Pratiwi, B. P., & Silvia, A. (2020). *Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion matrix*. 6(2), 66–75.
- Qadrini L, Sepperwali A, & Aina A. (2021). *Decision treedan Adaboostpada Klasifikasi Penerima Program Bantuan Sosial*. *Decision tree Dan Adaboost Pada Klasifikasi Penerima Program Bantuan Sosial*, 2(7), 1959–1966.

- Rachman, F. P. A. Putra, Goejantoro, R., & Hayati, M. N. (2018). Penentuan Jumlah Replikasi Bootstrap Menggunakan Metode Pretest Pada Independent Sampel T Test (Pendapatan Asli Daerah Kabupaten / Kota Di Provinsi Kalimantan Timur Dan Kalimantan Utara Tahun 2015). *Jurnal Eksponensial*, 9(1), 35–40.
- Rindengan, A. J., & Langi, Y. A. R. (2019). *Altien J. Rindengan Yohanes A.R. Langi*.
- Rosyidah, D. U. (2021). Knowledge Improvement Of Cervical Cancer And Early Detection With Papsmear Test. *Jurnal Pengabdian Masyarakat Medika*, 1(1), 12–16. <https://doi.org/10.23917/jpmm>
- Sarno, R., Buliali, J. L., & Maimunah, S. (2020). Pengembangan Metode. *Makara*, 6(1989), 46–55.
- Setianingsih, E., Astuti, Y., & Aisyaroh, N. (2022). Literature Review : Faktor-Faktor Yang Mempengaruhi Terjadinya Kanker Serviks. *Jurnal Ilmiah Pannmed (Pharmacist, Analyst, Nurse, Nutrition, Midwivery, Environment, Dentist)*, 17(1), 47–54. <https://doi.org/10.36911/Pannmed.V17i1.1231>
- Sholik, M., & Fatichah, C. (2017). Klasifikasi Sel Serviks Pada Citra Pap Smear Berdasarkan Fitur Bentuk Deskriptor Regional Dan Fitur Tekstur Uniform Rotated Local Binary Pattern. *Juti: Jurnal Ilmiah Teknologi Informasi*, 15(2), 214. <https://doi.org/10.12962/J24068535.V15i2.A669>
- Siswanto, E., Suprapedi, & Purwanto. (2018). Metode *Sample* Bootstrapping Pada K-Nearest Neighbor Untuk Klasifikasi Status Desa. *Jurnal Teknologi Informasi*, 14(1), 13–23.
- Sompawong, N., Mopan, J., Pooprasert, P., Himakhun, W., Suwannarurk, K., Ngamvirojcharoen, J., Vachiramon, T., & Tantibundhit, C. (2019). Automated Pap Smear Cervical Cancer Screening Using Deep Learning. *Proceedings Of The Annual International Conference Of The Ieee Engineering In Medicine And Biology Society, Embs*, 7044–7048. <https://doi.org/10.1109/Embc.2019.8856369>
- Tinggi, S., Informatika, M., Komputer, D., & Mandiri, N. (2020). *Integrasi Metode Diskritisasi Dan Gain Ratio Pada Prediksi Cacat Perangkat Lunak Berbasis Naive Bayes Tesis*.
- Toha, A., Purwono, P., & Gata, W. (2022). Model Prediksi Kualitas Udara Dengan Support Vector Machines Dengan Optimasi Hyperparameter Gridsearch Cv. *Buletin Ilmiah Sarjana Teknik Elektro*, 4(1), 12–21. <https://doi.org/10.12928/Biste.V4i1.6079>
- Ulfah Nuraini Karim, A. D. (2021). Analisa Faktor Resiko Kanker Serviks

Dikaitkan Dengan Kualitas Hidup Pasien Di Rsia Bunda Jakarta. *Ilmu Keperawatan Dan Kebidanan Universitas Binawan*.

- Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining Dan Reduksi Atribut Pada Data Set Penyakit Jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437. <https://doi.org/10.30865/Mib.V4i2.2080>
- Visa Sofia, D. (2011). *Confusion matrix-Based Feature Selection* Sofia Visa. *Confusionmatrix-Based Feature Selection Sofia*, 710(January), 8.
- Wiasitini, N. P. A. O., Putra, I. K. G. D., & Wibawa, K. S. (2019). Klasifikasi Sel Nukleus Pap Smear Menggunakan Metode Backpropagation Neural Network. *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, 7(3), 224. <https://doi.org/10.24843/Jim.2019.V07.I03.P06>
- William, W., Ware, A., Basaza-Ejiri, A. H., & Obungoloch, J. (2019). A Pap-Smear Analysis Tool (Pat) For Detection Of Cervical Cancer From Pap-Smear Images. *Biomedical Engineering Online*, 18(1), 1–22. <https://doi.org/10.1186/S12938-019-0634-5>
- Yandi Saputra, A., & Primadasa, Y. (2018). Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbour Implementation Of Classification Method To Predict Student Graduation Using K-Nearest Neighbor Algorithm. *Techno.Com*, 17(4), 9.
- Yusa, M., Utami, E., & Luthfi. Taufiq, E. (2016). Evaluasi Performa Algoritma Klasifikasi *Decision tree* Id3, C4.5, Dan Cart Pada *Dataset* Readmisi Pasien Diabetes. *Infosys (Information System) Journal*, 4(1), 23–34.
- Zhang, S., Xu, H., Zhang, L., & Qiao, Y. (2020). Cervical Cancer: Epidemiology, Risk Factors And Screening. *Chinese Journal Of Cancer Research*, 32(6), 720–728. <https://doi.org/10.21147/J.Issn.1000-9604.2020.06.05>