

**PENERAPAN SELEKSI FITUR *INFORMATION GAIN* DAN OPTIMASI
GRID SEARCH PADA ALGORITMA XGBOOST UNTUK KLASIFIKASI
REKURENSI PENYAKIT TIROID**

SKRIPSI



UIN SUNAN AMPEL
S U R A B A Y A

Disusun Oleh
HANA MEIDINA QURROTA A'YUN
09040222056

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2026

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : HANA MEIDINA QURROTA A'YUN
NIM : 09040222056
Program Studi : Matematika
Angkatan : 2022

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "PENERAPAN SELEKSI FITUR *INFORMATION GAIN* DAN OPTIMASI *GRID SEARCH* PADA ALGORITMA XGBOOST UNTUK KLASIFIKASI REKURENSI PENYAKIT TIROID". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 05 November 2025

Yang menyatakan,




HANA MEIDINA QURROTA A'YUN
NIM. 09040222056

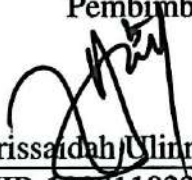
LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : HANA MEIDINA QURROTA A'YUN
NIM : 09040222056
Judul skripsi : PENERAPAN SELEKSI FITUR *INFORMATION GAIN*
DAN OPTIMASI *GRID SEARCH* PADA ALGORITMA
XGBOOST UNTUK KLASIFIKASI REKURENSI
PENYAKIT TIROID

telah diperiksa dan disetujui untuk diujikan.

Pembimbing I

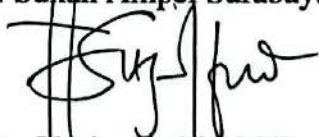

Nurissaidah/Ulinnuha, M.Kom
NIP. 199011022014032004

Pembimbing II


Dr. Lutfi Hakim, M.Ag.
NIP. 197312252006041001

Mengetahui,

Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya


Dr. Yuniar Farida, M.T
NIP. 197905272014032002

PENGESAHAN TIM PENGUJI SKRIPSI

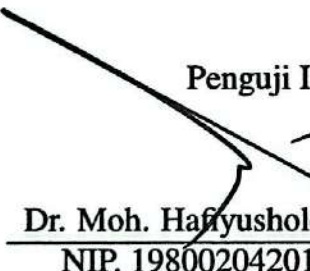
Skripsi oleh

Nama : HANA MEIDINA QURROTA A'YUN
NIM : 09040222056
Judul Skripsi : PENERAPAN SELEKSI FITUR *INFORMATION GAIN*
DAN OPTIMASI *GRID SEARCH* PADA ALGORITMA
XGBOOST UNTUK KLASIFIKASI REKURENSI
PENYAKIT TIROID

Telah dipertahankan di depan Tim Penguji
pada tanggal 5 Januari 2026

Mengesahkan,
Tim Penguji

Penguji I


Dr. Moh. Hafiyusholeh, M.Si., M.PMat.
NIP. 198002042014031001

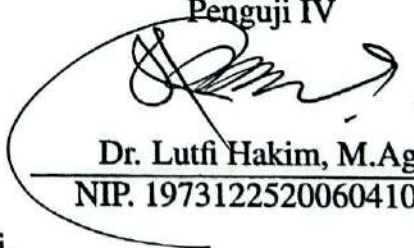
Penguji II


Dian Yulianti, M.Si
NIP. 198707142020122015

Penguji III


Nurissardah Linnuha, M.Kom
NIP. 199011022014032004

Penguji IV


Dr. Lutfi Hakim, M.Ag.
NIP. 197312252006041001

Mengetahui,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya



Dr. A. Saepul Hamdani, M.Pd
NIP. 196507312000031002



UIN SUNAN AMPEL
SURABAYA

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300

E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : HANA MEIDINA QURROTA A'YUN
NIM : 09040222056
Fakultas/Jurusan : SAINS DAN TEKNOLOGI / MATEMATIKA
E-mail address : hanameidina8@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Skripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)

yang berjudul :

PENERAPAN SELEKSI FITUR INFORMATION GAIN DAN
OPTIMASI GRID SEARCH PADA ALGORITMA XGBOOST
UNTUK KLASIFIKASI REKURENSI PENYAKIT TIROID.

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 5 Januari 2026

Penulis

(Hana Meidina Q.A.)
nama terang dan tanda tangan

ABSTRAK

PENERAPAN SELEKSI FITUR *INFORMATION GAIN* DAN OPTIMASI *GRID SEARCH* PADA ALGORITMA XGBOOST UNTUK KLASIFIKASI REKURENSI PENYAKIT TIROID

Penyakit tiroid, khususnya *Differentiated Thyroid Cancer* (DTC), merupakan salah satu gangguan endokrin dengan tingkat kejadian yang tinggi secara global dan berpotensi terus meningkat. Meskipun DTC memiliki tingkat kesembuhan yang relatif baik, kejadian rekurensi masih menjadi permasalahan klinis yang penting karena dapat muncul kembali setelah terapi awal. Variasi karakteristik klinis dan patologis pasien serta kompleksitas data medis menyebabkan proses prediksi rekurensi menjadi sulit dilakukan secara akurat. Tujuan penelitian ini adalah menerapkan seleksi fitur *Information Gain* dan optimasi menggunakan *Grid Search* pada algoritma XGBoost dalam klasifikasi rekurensi penyakit tiroid. Fitur yang digunakan mencakup karakteristik demografis, klinis, patologis, serta respons terhadap terapi, dengan status rekurensi sebagai fitur target. Data penelitian berasal dari rekam medis pasien DTC di Iran yang dikumpulkan selama periode 15 tahun, dengan masa tindak lanjut minimal 10 tahun setelah diagnosis awal. Hasil penelitian menunjukkan bahwa kombinasi seleksi fitur *Information Gain* dan optimasi menggunakan *Grid Search* pada algoritma XGBoost menghasilkan performa terbaik dibandingkan dengan metode XGBoost yang dikombinasikan dengan *Grid Search*, dimana model menghasilkan akurasi sebesar 96,98%, sensitivitas 93,52%, dan spesifisitas 98,44%. Pada model XGBoost yang dikombinasikan dengan seleksi fitur *Information Gain* dan optimasi *Grid Search* menghasilkan akurasi sebesar 97,53%, sensitivitas 93,52%, dan spesifisitas 99,22%, yang menunjukkan peningkatan kinerja klasifikasi. Hasil penelitian ini menunjukkan bahwa kombinasi seleksi fitur *Information Gain* dan optimasi *Grid Search* pada algoritma XGBoost dapat meningkatkan performa klasifikasi rekurensi.

Kata kunci: *Grid Search*, *Information Gain*, Klasifikasi, Rekurensi Penyakit Tiroid, XGBoost

ABSTRACT

APPLICATION OF INFORMATION GAIN FEATURE SELECTION AND GRID SEARCH OPTIMIZATION IN THE XGBOOST ALGORITHM FOR CLASSIFICATION OF THYROID DISEASE RECURRENCE

Thyroid disease, particularly Differentiated Thyroid Cancer (DTC), is one of the endocrine disorders with a high incidence rate globally and has the potential to continue to increase. Although DTC has a relatively good cure rate, recurrence remains an important clinical problem because it can reappear after initial therapy. Variations in the clinical and pathological characteristics of patients and the complexity of medical data make the process of predicting recurrence difficult to perform accurately. The purpose of this study was to apply Information Gain feature selection and optimization using Grid Search in the XGBoost algorithm in the classification of thyroid disease recurrence. The features used included demographic, clinical, pathological, and response to therapy characteristics, with recurrence status as the target feature. The research data came from the medical records of DTC patients in Iran collected over a 15-year period, with a minimum follow-up period of 10 years after the initial diagnosis. The results showed that the combination of Information Gain feature selection and optimization using Grid Search in the XGBoost algorithm produced the best performance compared to the XGBoost method combined with Grid Search, where the model produced an accuracy of 96,98%, sensitivity of 93,52%, and specificity of 98,44%. The XGBoost model combined with Information Gain feature selection and Grid Search optimization resulted in an accuracy of 97,53%, sensitivity of 93,52%, and specificity of 99,22%, indicating an improvement in classification performance. The results of this study indicate that the combination of Information Gain feature selection and Grid Search optimization in the XGBoost algorithm can improve the performance of recurrence classification.

Keywords: Classification, Grid Search, Information Gain, Thyroid Disease Recurrence, XGBoost

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI	iii
HALAMAN PERNYATAAN KEASLIAN	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
ABSTRAK	xiv
ABSTRACT	xv
 I PENDAHULUAN	 2
1.1. Latar Belakang Masalah	2
1.2. Rumusan Masalah	8
1.3. Tujuan Penelitian	8
1.4. Manfaat Penelitian	9
1.5. Batasan Masalah	9
1.6. Sistematika Penulisan	10

II TINJAUAN PUSTAKA	12
2.1. Rekurensi Penyakit Tiroid	12
2.1.1. Jenis Penyakit Tiroid	14
2.1.2. Faktor Risiko	16
2.1.3. Diagnosis	19
2.1.4. Penanganan	21
2.2. Klasifikasi	22
2.3. <i>Encoding</i>	23
2.4. Seleksi Fitur	24
2.5. <i>Boosting</i>	26
2.6. <i>Extreme Gradient Boosting</i> (XGBoost)	26
2.7. <i>Grid Search</i>	40
2.8. <i>K-Fold Cross Validation</i>	41
2.9. Evaluasi	43
2.10. Integrasi Keilmuan	44
III METODE PENELITIAN	48
3.1. Jenis Penelitian	48
3.2. Sumber Data	48
3.3. Variabel Penelitian	49
3.4. Tahapan Penelitian	50
IV HASIL DAN PEMBAHASAN	56
4.1. <i>Preprocessing</i> Data	56
4.1.1. <i>Encoding</i>	57
4.1.2. Deskripsi Data	58
4.2. <i>Information Gain</i>	63
4.3. Perhitungan Manual Metode XGBoost	66
4.3.1. Menghitung Nilai Prediksi Awal dan Nilai Residual	66
4.3.2. Menghitung <i>Similarity Score</i>	68
4.3.3. Menghitung <i>Gain</i> dan <i>Split</i> Pohon	68

4.3.4.	Menghitung <i>Cover</i>	77
4.3.5.	Melakukan <i>Pruning</i> Pohon	78
4.3.6.	Menghitung <i>Output Value</i>	79
4.3.7.	Menghitung Prediksi Baru	79
4.4.	Klasifikasi Rekurensi Penyakit Tiroid Menggunakan Algoritma XGBoost	83
4.4.1.	Klasifikasi XGBoost dengan <i>Default Hyperparameter</i>	83
4.4.2.	Klasifikasi XGBoost dengan <i>Grid Search</i>	86
4.4.3.	Klasifikasi XGBoost dengan <i>Information Gain</i> dan <i>Grid Search</i>	89
4.4.4.	Perbandingan Hasil Klasifikasi	94
4.5.	Klasifikasi dalam Perspektif Islam	97
V	PENUTUP	100
5.1.	Kesimpulan	100
5.2.	Saran	101
	DAFTAR PUSTAKA	101

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

2.10	<i>Confusion Matrix</i>	43
3.3	Deskripsi Fitur Dataset Penyakit Tiroid	49
3.4	<i>Hyperparameter</i> XGBoost	52
4.1	Data Awal	56
4.2	Data Setelah Pembersihan Duplikasi	57
4.3	Data Setelah <i>Encoding</i>	57
4.4	Statistika Deskriptif Fitur Kategorik	58
4.5	Statistika Deskriptif Fitur Numerik	62
4.6	Hasil Perhitungan Nilai <i>Information Gain</i>	64
4.7	Nilai Prediksi Awal	66
4.8	Residual	68
4.9	Nilai <i>Output Value</i>	79
4.10	Nilai Prediksi Baru pada Setiap Leaf Node	80
4.11	Nilai Sigmoid pada Setiap <i>Leaf Node</i>	81
4.12	Hasil Penentuan Kelas pada Setiap <i>Leaf Node</i>	82
4.13	Hasil Prediksi	82
4.14	Hasil <i>Hyperparameter Tuning Grid Search</i>	87
4.15	Hasil Uji Coba Jumlah Fitur	90
4.16	Hasil <i>Hyperparameter Tuning Grid Search</i> dan <i>Information Gain</i>	92
4.17	Perbandingan Kinerja Model	94
4.18	Perbandingan Metode dan Akurasi Penelitian	97

DAFTAR GAMBAR

2.1	Kelenjar Tiroid	13
2.2	<i>Label Encoding dan One-Hot Encoding</i>	24
2.6	XGBoost	27
2.7	<i>Grid Search</i>	41
2.8	<i>K-fold Cross Validation</i>	42
3.4	<i>Digram Alir Penelitian</i>	50
3.4	<i>Diagram Alir Information Gain</i>	50
3.4	<i>Diagram Alir XGBoost</i>	52
4.1	Pohon XGBoost dengan max_depth = 1	70
4.2	Pohon XGBoost dengan max_depth = 2	74
4.3	Pohon XGBoost dengan max_depth = 3	76
4.4	Pohon XGBoost dengan max depth = 6	77
4.5	<i>Confusion Matrix XGBoost dengan Hyperparameter Default</i>	83
4.6	Pohon XGBoost dengan <i>Hyperparameter Default</i>	85
4.7	<i>Confusion Matrix XGBoost dengan Grid Search</i>	86
4.8	<i>Pohon XGBoost dengan Grid Search</i>	88
4.9	<i>Confusion Matrix XGBoost dengan IG dan Grid Search</i>	91
4.10	Pohon XGBoost dengan Seleksi Fitur IG dan Optimasi <i>Grid Search</i>	93

DAFTAR PUSTAKA

- Abraham, M. T., Satyam, N., Lokesh, R., Pradhan, B., & Alamri, A. (2021). Factors Affecting Landslide Susceptibility Mapping Assessing the Influence of Different Machine Learning Approaches, Sampling Strategies and Data Splitting. *Land*, 10(9), 1–24. <https://doi.org/10.3390/land10090989>.
- Al-Mahdawi, F. K. I., Mohammed, M. R., Taher, M. G., Muthanna, A.-M. A. S., Alkalifa, N. F., Kadi, A., Potoroko, I., & Chemek, M. (2024). Thyroid Gland: Physiology and Pathology. *Academic Science Journal*, 2(3), 161–170. <https://doi.org/10.24237/ASJ.02.03.857I>.
- Alawiyah, T., Wibisono, T., & Mulyani, Y. S. (2024). The Prediction of Thyroid Cancer Recurrence with the XGBoost Method: The Clinicopathological Feature-Based Approach. *Journal of Computer Networks, Architecture and High Performance Computing*, 6(3), 1035–1045. <https://doi.org/10.47709/cnahpc.v6i3.4101>.
- Ali, Z. A., Abduljabbar, Z. H., Tahir, H. A., Sallow, A. B., & Almufti, S. M. (2023). Exploring the Power of Extreme Gradient Boosting Algorithm in Machine Learning: A Review. *Academic Journal of Nawroz University (AJNU)*, 12(2), 320–334. <https://doi.org/10.25007/ajnu.v12n2a1612>.
- Amir, N. A. B., Kasprata, H. N., Priyahita, P. D., Syahla, T., & Triani, E. (2023). Clinical Diagnosis and Management of Thyroid Cancer. *Jurnal Kedokteran Unram*, 12(3), 240–246. <https://doi.org/10.29303/jk.v12i3.4493>.
- Andreadi, A., Andreadi, S., Cerilli, M., Todaro, F., Lazzaroni, M., Lodeserto, P., Meloni, M., Ferrari, C., Bellia, A., Coppeta, L., Magrini, A., & Lauro, D. (2025). Occupational Radiation Exposure and Thyroid Nodules in Healthcare Workers: A Review. *International Journal of Molecular Sciences*, 26(13), 1–16. <https://doi.org/10.3390/ijms26136522>.
- Araque, K. A., Gubbi, S., & Klubo-Gwiedzinska, J. (2020). Updates on the Management of Thyroid Cancer. *Hormone and Metabolic Research*, 52(8), 562–577. <https://doi.org/10.1055/a-1089-7870>.

- Ariobimo, B. N., Nujum, N., Puspamaniar, V. A., & Harto Saputro, D. P. (2023). Tata Laksana Perioperatif Pembedahan Tiroid: Sebuah Tinjauan Pustaka. *Journal of The Indonesian Medical Association*, 73(1), 46–54. <https://doi.org/10.47830/jinma-vol.73.1-2023-859>.
- Arista, R. D., Karima, K., Anugrah, M. F., Widyastuti, P., & Triani, E. (2023). Thyroid Cancer: An Overview of Epidemiology, Risk Factors, and Treatment. *Lombok Medical Journal*, 2(3), 90–96. <https://doi.org/10.29303/lmj.v2i2.2791>.
- Astuti, R. A. & Irfani, F. N. (2024). Pemeriksaan Imunologi terhadap Kadar Hormon Thyroid Stimulating Hormone (TSH) pada Pasien Gangguan Tiroid di RSUD Panembahan Senopati Periode 2020–2022. *Jurnal Ilmiah Kedokteran dan Kesehatan (KLINIK)*, 3(1), 280–292. <https://doi.org/10.55606/klinik.v3i1.2631>.
- Ayomi, F. A. J. & Dewi, K. E. (2023). Analisis Emosi pada Media Sosial Twitter Menggunakan Metode Multinomial Naïve Bayes dan Synthetic Minority Oversampling Technique. *KOMPUTA : Jurnal Ilmiah Komputer dan Informatika*, 12(2), 9–19. <https://doi.org/10.34010/komputa.v12i2.9454>.
- Bashir, A. A., El-Zaheri, M. M., Bashir, A. A., Fayyad, L., Obed, A. H., Alkam, D., & Bashir, A. Y. (2025). Papillary Thyroid Microcarcinoma in Thyroid Surgical Practice: Incidental vs Non-Incidental: A Ten-Year Comparative Study. *Cancers*, 17(12), 1–17. <https://doi.org/10.3390/cancers17122029>.
- Bhartariya, A. G., Singh, S. K., & Bharti, A. K. (2025). Next-Gen Business Intelligence: Leveraging Predictive Analytics, AI & Real-Time Decision-Making. *Journal of Theoretical and Applied Information Technology*, 103(8), 3272–3284. <https://doi.org/10.5281/zenodo.17208037>.
- Borzooei, S., Briganti, G., Golparian, M., Lechien, J. R., & Tarokhian, A. (2023). Machine Learning for Risk Stratification of Thyroid Cancer Patients A 15-Year Cohort Study. *European Archives of Oto-Rhino-Laryngology*, 281(2024), 2095–2104. <https://doi.org/10.1007/s00405-023-08299-w>.
- Chen, S., Xu, L., Pan, S., & Chen, G. (2024). Impact of Distant Metastasis on Overall Survival and Cancer-Specific Survival of Elderly Patients

- with Differentiated Thyroid Cancer. *Scientific Reports*, 14(24855), 1–10. <https://doi.org/10.1038/s41598-024-76613-5>.
- Coca-Pelaz, A., Shah, J. P., Hernandez-Prera, J. C., Ghossein, R. A., Rodrigo, J. P., Hartl, D. M., Olsen, K. D., Shaha, A. R., Zafero, M., Suarez, C., Nixon, I. J., Randolph, G. W., Mäkitie, A. A., Kowalski, L. P., Vander Poorten, V., Sanabria, A., Guntinas-Lichius, O., Simo, R., Zbären, P., Angelos, P., Khafif, A., Rinaldo, A., & Ferlito, A. (2020). Papillary Thyroid Cancer—Aggressive Variants and Impact on Management: A Narrative Review. *Advances in Therapy*, 37(20), 3112–3126. <https://doi.org/10.1007/s12325-020-01391-1>.
- De Carvalho, A. Y., Kohler, H. F., Gomes, C. C., Vartanian, J. G., & Kowalski, L. P. (2021). Predictive Factors for Recurrence of Papillary Thyroid Carcinoma: Analysis of 4,085 Patients. *ACTA Otorhinolaryngologica Italica*, 41(3), 236–242. <https://doi.org/10.14639/0392-100X-N1412>.
- Fadilah, I., Rusjdi, D. A., & Aprilia, D. (2021). Gambaran Pemeriksaan Ultrasonografi pada Pasien Struma di Bagian/SMF Radiologi RSUP Dr. M. Djamil Periode Januari–Desember 2019. *Jurnal Ilmu Kesehatan Indonesia*, 2(1), 41–47. <https://doi.org/10.25077/jikesi.v2i1.285>.
- Ferraninda, Kusumajaya, H., & Ardiansyah (2023). Faktor-faktor yang Berhubungan dengan Penyakit Hipertiroid pada Pasien Poliklinik Penyakit Dalam di UPTD RSUD Dr. (HC) Ir. Soekarno Provinsi Bangka Belitung Tahun 2022. *Jurnal Ilmu Kesehatan Bhakti Husada: Health Science Journal*, 14(1), 41–48. <https://doi.org/10.34305/jikbh.v14i01.673>.
- Filetti, S., Durante, C., Hartl, D., Leboulleux, S., Locati, L. D., Newbold, K., Papotti, M. G., Berruti, A., & Committee, G. (2019). Thyroid Cancer: ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up. *Annals of Oncology*, 30(12), 1856–1883. <https://doi.org/10.1093/annonc/mdz400>.
- Firdaus, F. F., Nugroho, H. A., & Soesanti, I. (2020). A Review of Feature Selection and Classification Approaches for Heart Disease Prediction. *International Journal of Information Technology and Electrical Engineering*, 4(3), 75–82. <https://doi.org/10.22146/ijitee.59193>.
- Flor, L. S., Anderson, J. A., Ahmad, N., Aravkin, A., Carr, S., Dai, X., Gil, G. F., Hay, S. I., Malloy, M. J., McLaughlin, S. A., Mullany, E. C.,

- Murray, C. J. L., O'Connell, E. M., Okereke, C., Sorensen, R. J. D., Whisnant, J., Zheng, P., & Gakidou, E. (2024). Health Effects Associated with Exposure to Secondhand Smoke: A Burden of Proof Study. *Nature Medicine*, 30(1213), 149–167. <https://doi.org/10.1038/s41591-023-02743-4>.
- Forma, A., Klodnicka, K., Pajak, W., Flieger, J., Teresinska, B., Januszewski, J., & Baj, J. (2025). Thyroid Cancer: Epidemiology, Classification, Risk Factors, Diagnostic and Prognostic Markers, and Current Treatment Strategies. *International Journal of Molecular Sciences*, 26(11), 1–44. <https://doi.org/10.3390/ijms26115173>.
- Friedenreich, C. M., Ryder-Burbidge, C., & Mcneil, J. (2021). Physical Activity, Obesity and Sedentary Behavior in Cancer Etiology: Epidemiologic Evidence and Biologic Mechanisms. *Molecular Oncology*, 15(3), 790–800. <https://doi.org/10.1002/1878-0261.12772>.
- Fuadah, Y. N., Ubaidullah, I. D., Ibrahim, N., Taliningsing, F. F., Khofiya, N. S., & Pramuditho, M. A. (2022). Optimasi Convolutional Neural Network dan K-Fold Cross Validation pada Sistem Klasifikasi Glaukoma. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 10(3), 728–741. <https://doi.org/10.26760/elkomika.v10i3.728>.
- Giovanella, L., D'Aurizio, F., Ovčariček, P. P., & Görges, R. (2024). Diagnostic, Theranostic and Prognostic Value of Thyroglobulin in Thyroid Cancer. *Journal of Clinical Medicine*, 13(9), 1–16. <https://doi.org/10.3390/jcm13092463>.
- Gomez-Izquierdo, J., Filion, K. B., Boivin, J.-F., Azoulay, L., Pollak, M., & Yu, O. H. Y. (2020). Subclinical Hypothyroidism and the Risk of Cancer Incidence and Cancer Mortality: A Systematic Review. *BMC Endocrine Disorders*, 20(83), 1–10. <https://doi.org/10.1186/s12902-020-00566-9>.
- Grani, G., Lamartina, L., Durante, C., Filetti, S., & Cooper, D. S. (2018). Follicular Thyroid Cancer and Hurthle Cell Carcinoma: Challenges in Diagnosis, Treatment, and Clinical Management. *THE LANCET Diabetes & Endocrinology*, 6(6), 1–15. [https://doi.org/10.1016/S2213-8587\(17\)30325-X](https://doi.org/10.1016/S2213-8587(17)30325-X).

- Halomoan, M. S. (2023). Epidemiologi Kanker Tiroid. <https://www.alomedika.com/penyakit/onkologi/kanker-tiroid/epidemiologi>.
- Haugen, B. R. (2017). 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer What Is New and What Has Changed. *Cancer*, 123(3), 372–381. <https://doi.org/10.1002/cncr.30360>.
- Henisaniyya, N., Pertiwi, C., Desiani, A., Amran, A., & Arhami, M. (2025). Klasifikasi Class Thyroid menggunakan Algoritma ID3 dan Artificial Neural Network (ANN) Classification of Thyroid Class using ID3 Algorithm and Artificial. 14, 1–14. <https://doi.org/10.32520/stmsi.v14i1.3440>.
- Hoque, R., Das, S., Hoque, M., & Haque, E. (2024). Breast Cancer Classification Using XGBoost. *World Journal of Advanced Research and Reviews*, 21(2), 1985–1994. <https://doi.org/10.30574/wjarr.2024.21.2.0625>.
- Houten, P. V., Netea-Maier, R. T., & Smit, J. W. (2023). Differentiated Thyroid Carcinoma: An Update. *Best Practice & Research Clinical Endocrinology & Metabolism*, 37(1), 1–15. <https://doi.org/10.1016/j.beem.2022.101687>.
- Irawan, R. N., Hindrayani, K. M., & Idhom, M. (2024). Penerapan Cross Validation sebagai Analisis Sentimen Pelayanan Publik Kereta Api Lokal Daop 8 Menggunakan Metode Multinomial Naïve Bayes. *G-Tech: Jurnal Teknologi Terapan*, 8(2), 954–963. <https://doi.org/10.33379/gtech.v8i2.4117>.
- Ismunandar, D., Firdaus, M. R., & Alkhalifi, Y. (2024). Penerapan Hyperparameter Machine Learning dalam Prediksi Gagal Pinjam. *Inti Nusa Mandiri*, 19(1), 62–70. <https://doi.org/10.33480/inti.v19i1.5612>.
- Juliana, I. B., Sibarani, Loy, K. M., & Suharjito (2024). Enhancing Predictive Accuracy for Differentiated Thyroid Cancer (DTC) Recurrence Through Advanced Data Mining Techniques. 5(1), 11–22. <https://doi.org/10.47065/tin.v5i1.5237>.

- Karim, A. A., Prasetyo, M. A., & Saputro, M. R. (2023). Perbandingan Metode Random Forest, K-Nearest Neighbor, dan SVM dalam Prediksi Akurasi Pertandingan Liga Italia. *Seminar Nasional Teknologi & Sains*, 2(1), 377–382. <https://doi.org/10.29407/stains.v2i1.2877>.
- Khaffah, N., Wibawa, G. N. A., Arman, Somayasa, W., Ruslan, & Gubu, L. (2024). Penerapan Metode Naive Bayes untuk Penentuan Jurusan Siswa–Siswi pada SMAN 9 Kendari. *Jurnal Matematika, Komputasi dan Statistika*, 4(1), 615–623. <https://doi.org/10.33772/jmks.v4i1.73>.
- Khansa, S. F. A., Ulinnuha, N., & Utami, W. D. (2025). Grid Search and Random Search Hyperparameter Tuning Optimization in XGBoost Algorithm for Parkinson’s Disease Classification. *BAREKENG: Journal of Mathematics and Its Applications*, 19(3), 1609–1624. <https://doi.org/10.30598/barekengvol19iss3pp1609-1624>.
- Krajewska, J., Chmielik, E., & Jarzab, B. (2017). Dynamic Risk Stratification in the Follow-Up of Thyroid Cancer: What Is Still to Be Discovered in 2017? *Endocrine-Related Cancer*, 24(11), 1–16. <https://doi.org/10.1530/ERC-17-0270>.
- Kusumadewi, R. C., Pratiwi, R. S., Exhasna, A. T. I., Azizaturrahmah, G. A., Sastrawan, I. G. P. D. A., Wenilia, R., Muttaqin, S., Supriantari, D., Ilham, M. A. R., Fatahillah, W. R., & Triani, E. (2024). Understanding Thyroid Carcinoma Clinical Manifestations, Diagnosis, and Management. *Jurnal Biologi Tropis*, 24(4), 645–656. <https://doi.org/10.29303/jbt.v24i4.7664>.
- Kuzudisli, C., Bakir-Gungor, B., Bulut, N., Qaqish, B., & Yousef, M. (2023). Review of Feature Selection Approaches Based on Grouping of Features. *PeerJ*, 11(15666), 1–37. <https://doi.org/10.7717/peerj.15666>.
- Leni, D., Dwiharzandis, A., Sumiati, R., & Afriyani, H. S. (2023). Seleksi Fitur Berdasarkan Korelasi Pearson dalam Pemodelan Efisiensi Energi Bangunan. *Jurnal Teknik Sains*, 8(2), 103–115. <https://doi.org/10.24967/teksis.v8i2.2525>.
- Li, Y., Tian, J., Jiang, K., Wang, Z., Gao, S., Wei, K., Yang, A., & Li, Q. (2023). Risk Factors and Predictive Model for Recurrence in Papillary Thyroid Carcinoma: A Single-Center Retrospective Cohort Study Based on 955 Cases. *Frontiers in Endocrinology*, 14(1268282), 1–10. <https://doi.org/10.3389/fendo.2023.1268282>.

- Luvhengo, T. E., Bombil, I., Mokhtari, A., Moeng, M. S., Demetriou, D., Sanders, C., & Dlamini, Z. (2023). Multi-Omics and Management of Follicular Carcinoma of the Thyroid. *Biomedicines*, 11(4), 1–28. <https://doi.org/10.3390/biomedicines11041217>.
- Miranda-Filho, A., Lortet-Tieulent, J., Bray, F., Cao, B., Franceschi, S., Vaccarella, S., & Dal Maso, L. (2021). Thyroid Cancer Incidence Trends by Histology in 25 Countries: A Population-Based Study. *The Lancet Diabetes and Endocrinology*, 9(4), 225–234. [https://doi.org/10.1016/s2213-8587\(21\)00027-9](https://doi.org/10.1016/s2213-8587(21)00027-9).
- Murdiansyah, D. T. (2024). Prediksi Stroke Menggunakan Extreme Gradient Boosting. *JIKO (Jurnal Informatika dan Komputer)*, 8(2), 419–426. <http://dx.doi.org/10.26798/jiko.v8i2.1295>.
- Nadya, F. E. P., Ferdiansyah, M. F. I., Nastiti, V. R. S., & Aditya, C. S. K. (2024). Implementation of Feature Selection Strategies to Enhance Classification Using XGBoost and Decision Tree. *Scientific Journal of Informatics*, 11(1), 187–194. <https://doi.org/10.15294/sji.v11i1.48145>.
- Negara, A. B. P., Muhardi, H., & Putri, I. M. (2020). Analisis Sentimen Maskapai Penerbangan Menggunakan Metode Naive Bayes dan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(3), 599–606. <https://doi.org/10.25126/jtiik.2020711947>.
- Nurhidayat, R. & Dewi, K. E. (2023). Penerapan Algoritma K-Nearest Neighbor dan Fitur Ekstraksi N-Gram dalam Analisis Sentimen Berbasis Aspek. *KOMPUTA : Jurnal Ilmiah Komputer dan Informatika*, 12(1), 91–100. <https://doi.org/10.34010/komputa.v12i1.9458>.
- Pardede, J. & Dwianto, R. (2025). The Effect of Feature Selection on Machine Learning Classification. *International Journal on Informatics Visualization*, 9(4), 1419–1429. <http://dx.doi.org/10.62527/joiv.9.4.2926>.
- Park, J., An, S., Bae, J. S., Kim, K., & Kim, J. S. (2024). Impact of Tumor Size on Prognosis in Differentiated Thyroid Cancer with Gross Extrathyroidal Extension to Strap Muscles: Redefining T3b. *Cancers*, 16(14), 1–13. <https://doi.org/10.3390/cancers16142577>.
- Paschke, R., Ghaznavi, S., Imran, S. A., Jacquier, J., Lochnan, H., Massicotte, M.-H., Mitmaker, E., Morrison, D., Munro, V., Sveistrup, M., Wiseman,

- S. M., & Zahedi, A. (2025). Follow-Up and Transition of Care for Low Recurrence Risk Thyroid Cancer Patients in Canada. *European Thyroid Journal*, 14(3), 1–18. <https://doi.org/10.1530/ETJ-25-0072>.
- Pattiasina, S., Susilo, D. H., & Wibowo, M. D. (2025). Predictive Factors for the Recurrence of Papillary Thyroid Cancer Following Resection at Dr. Soetomo General Hospital, Indonesia. *South Eastern European Journal of Public Health*, 27(6706), 1604–1612. <https://doi.org/10.70135/seejph.vi.6706>.
- Pietsch, K. & Dhillon, V. K. (2023). Early Assessment of Laryngeal Dysfunction After Thyroid Surgery: A Narrative Review on Detection and Treatment Considerations and the Role of Speech Language Pathology. *Annals of Thyroid*, 8(2), 1–11. <https://doi.org/10.21037/aot-22-11>.
- Putri, R. C. T. W., Irfani, F. N., & Bimantara, A. (2024). Gambaran Kadar Pemeriksaan Thyroid Stimulating Hormone (TSH) dan Free Thyroxine (FT4) pada Pasien dengan Gangguan Tiroid di RSUD Bagas Waras Klaten. *Jurnal Kesehatan Tambusai*, 5(4), 10698–10704. <https://doi.org/10.31004/jkt.v5i4.34511>.
- Rizky, P. S., Hirzi, R. H., & Hidayaturohman, U. (2022). Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang. *Jurnal Ilmiah Teori dan Aplikasi Statistika*, 15(2), 228–236. <https://doi.org/10.36456/jstat.vol15.no2.a5548>.
- Rusman, J., Haryati, B. Z., & Aryo, M. (2023). Optimisasi Hiperparameter Tuning pada Metode Support Vector Machine untuk Klasifikasi Tingkat Kematangan Buah Kopi. *Jurnal Informatika dan Komputer*, 11(2), 195–202. <https://doi.org/10.35508/jicon.v11i2.12571>.
- Saenko, V. & Mitsutake, N. (2024). Radiation-Related Thyroid Cancer. *Endocrine Reviews*, 45(1), 1–29. <https://doi.org/10.1210/endrev/bnad022>.
- Salgado, S. A., Kaye, E. R., Sargi, Z., Chung, C. H., & Papaleontiou, M. (2023). Management of Advanced Thyroid Cancer Overview, Advances, and Opportunities. *Head and Neck Cancer*, 43(389708), 1–10. https://doi.org/10.1200/EDBK_389708.
- Schlumberger, M. & Leboulleux, S. (2020). Current Practice in Patients

With Differentiated Thyroid Cancer. *Nature Reviews Endocrinology*.
<http://dx.doi.org/10.1038/s41574-020-00448-z>.

Setiawati, R., Wulanhadarini, T., Hayati, F., Erawati, D., Jaya, M. G., Thoriq, A. A., Wijaya, T. M., Ismiyati, G. N., Kusumaningrum, D. W., Koesmarsono, B., Basja, A. T., Nugroho, M. I., Yuliana, S., Shedyta, S. Z., & Situmorang, H. B. (2022). Clinical and Ultrasonography Evaluation of Thyroid Tumor Screening in Symptomatic Patients of Bajulmati Primary Care Center, Banyuwangi, East Java, Indonesia. *Medicine*, 101(52), 1–5. <https://doi.org/10.1097/MD.00000000000032546>.

Shalih, M. G., Utami, M. R., Adam, M. I., & Shadrina, J. A. (2023). Edukasi Hormon Tiroid dan Antitiroid terhadap Penyakit Gondok di SMK Wirasaba Karawang. *Jurnal Dorkes (Dedikasi Olahraga dan Kesehatan)*, 1(2), 50–57. <https://doi.org/10.35706/dorkes.v1i2.10475>.

Shank, J. B., Are, C., & Wenos, C. D. (2022). Thyroid Cancer: Global Burden and Trends. *Indian Journal of Surgical Oncology*, 13(1), 40–45. <https://doi.org/10.1007/s13193-021-01429-y>.

Shen, J., Yan, M., Chen, L., Ou, D., Yao, J., Feng, N., Zhou, X., Lei, Z., & Xu, D. (2023). Prognosis and Influencing Factors of Follicular Thyroid Cancer. *Cancer Medicine*, 13(6727), 1–12. <https://doi.org/10.1002/cam4.6727>.

Shurbaji, S., Al Tamimi, F., Al Ghwairi, M. M., El Chaar, D., Younes, S., Majdalawieh, A. F., Pintus, G., Al-Dewik, N., & Nasrallah, G. K. (2023). High-Sensitive Detection and Quantitation of Thyroid-Stimulating Hormone (TSH) from Capillary/Fingerstick and Venepuncture Whole-Blood Using Fluorescence-Based Rapid Lateral Flow Immunoassay (LFIA). *Heliyon*, 9(10), 1–11. <https://doi.org/10.1016/j.heliyon.2023.e20589>.

Singh, G. A., Mane, V. V., Anand, R., & Sakoglu, U. (2025). ScienceDirect ScienceDirect Thyroid Disease Multi-Class Classification Thyroid Disease Multi-Class Classification. *Procedia Computer Science*, 258(2024), 3649–3658. <https://doi.org/10.1016/j.procs.2025.04.620>.

Singhavi, H. R., Mair, M., Qayyumi, B., Singh, A., & Chaturvedi, P. (2023). Thyroid Papillary Microcarcinoma – Carcinoma with Clinically Benign Behaviour. *Journal of Head & Neck Physicians and Surgeons*, 11(2), 95–101. https://doi.org/10.4103/jhnps.jhnps_44_23.

- Stains, A. P., Singh, A., Roy, R., & Suryakala, S. V. (2023). Advancement in Thyroid Detection Using XGBoost Algorithm. *International Journal of Applied Engineering & Technology*, 5(4), 106–119. <https://romanpub.com/resources/Vol%205%20%2C%20No%20S3%20-%2011.pdf>.
- Sudibia, N. K. M. P. (2024). Analisis Pemeriksaan Laboratorium pada Penyakit Tiroid. *Action Research Literate*, 8(8), 2193–2198. <https://doi.org/10.46799/ar.v8i8.482>.
- Sumoro, d. W. (2020). Teknik USG Tiroid. <https://www.alomedika.com/tindakan-medis/radiologi/ultrasonografi-tiroid/teknik>.
- Sunarya, U. & Haryanti, T. (2022). Perbandingan Kinerja Algoritma Optimasi pada Metode Random Forest untuk Deteksi Kegagalan Jantung. *Jurnal Rekayasa Elektrika*, 18(4), 241–247. <https://doi.org/10.17529/jre.v18i4.26981>.
- Taylor, P. N., Lansdown, A., Witczak, J., Khan, R., Rees, A., Dayan, C. M., & Okosieme, O. (2023). Age-Related Variation in Thyroid Function – A Narrative Review Highlighting Important Implications for Research and Clinical Practice. *Thyroid Research*, 16(7), 1–12. <https://doi.org/10.1186/s13044-023-00149-5>.
- Turner, N., Hamidi, S., Ouni, R., Rico, R., Henderson, Y. C., Puche, M., Alekseev, S., Colunga-Minutti, J. G., Zafereo, M. E., Lai, S. Y., Kim, S. T., Cabanillas, M. E., & Nurieva, R. (2024). Emerging therapeutic options for follicular-derived thyroid cancer in the era of immunotherapy. *Frontiers in Immunology*, 15(May), 1–18. <https://doi.org/10.3389/fimmu.2024.1369780>.
- Valerio, L., Dalmiglio, C., Maino, F., Mattii, E., Trimarchi, A., Cartocci, A., & Castagna, M. G. (2023). Dynamic Risk Stratification Integrated with ATA Risk System for Predicting Long-Term Outcome in Papillary Thyroid Cancer. *Cancers*, 15(18), 1–12. <https://doi.org/10.3390/cancers15184656>.
- Wang, J., Yu, F., Shang, Y., & Liu, L. (2020). Thyroid cancer : Incidence and Mortality Trends in China , 2005 – 2015. *Endocrine*, 2005, 2005–2015. <https://doi.org/10.1007/s12020-020-02207-6>.

- Watson-Brown, P. & Anderson, D. (2023). Differentiated Thyroid Cancer: A Guide to Survivorship Care. *Australian Journal of General Practice*, 52(1), 47–51. <https://doi.org/10.31128/AJGP-02-22-6327>.
- Wibawa, P. B. A. & Wibawa, A. P. (2022). Algoritma Klasifikasi pada Pengenalan Pola Citra. *Jurnal Inovasi Teknik dan Edukasi Teknologi*, 2(12), 557–566. <https://doi.org/10.17977/um068v2i122022p557-565>.
- Wiens, M., Verone-Boyle, A., Henscheid, N., Podichetty, J. T., & Burton, J. (2025). A Tutorial and Use Case Example of the Extreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications. *Clinical and Translational Science*, 18(3), 1–14. <https://doi.org/10.1111/cts.70172>.
- Wilimitis, D. & Walsh, C. G. (2023). Practical Considerations and Applied Examples of Cross-Validation for Model Development and Evaluation in Health Care A Tutorial. *JMIR AI*, 2(49023), 1–16. <https://doi.org/10.2196/49023>.
- Woo, J., Kim, H., & Kwon, H. (2021). Impact of Multifocality on the Recurrence of Papillary Thyroid Carcinoma. *Journal of Clinical Medicine*, 10(21), 1–8. <https://doi.org/10.3390/jcm10215144>.
- Xie, J., Wang, J., & Cui, X. (2025). Research Progress on Estrogen and Estrogen Receptors in the Occurrence and Progression of Autoimmune Thyroid Diseases. *Autoimmunity Reviews*, 24(6), 1–14. <https://doi.org/10.1016/j.autrev.2025.103803>.
- Yang, H., Wang, J., Wang, W., Shi, S. S., Liu, L., Yao, Y., Tian, G., Wang, P., & Yang, J. (2025). MMsurv: A Multimodal Multi-Instance Multi-Cancer Survival Prediction Model Integrating Pathological Images, Clinical Information, and Sequencing Data. *Briefings in Bioinformatics*, 26(3), 1–14. <https://doi.org/10.1093/bib/bbaf209>.
- Yang, J. C. (2023). The Prediction and Analysis of Heart Disease Using XGBoost Algorithm. *Applied and Computational Engineering*, 41(711), 61–68. <https://doi.org/10.54254/2755-2721/41/20230711>.
- Yu, J., Deng, Y., Liu, T., Zhou, J., Jia, X., Xiao, T., Zhou, S., Li, J., Guo, Y., Wang, Y., Zhou, J., & Chang, C. (2020). Lymph Node Metastasis Prediction of Papillary Thyroid Carcinoma Based on

- Transfer Learning Radiomics. *Nature Communications*, 11(4807), 1–10. <https://doi.org/10.1038/s41467-020-18497-3>.
- Yulian, E. D. & Prawirodihardjo, M. M. P. (2024). Recurrence Analysis of Pediatric Thyroid Carcinoma in Dr. Cipto Mangunkusumo General Hospital and Its Associated Factors. *The New Ropanasuri Journal of Surgery*, 9(1), 8–13. <https://doi.org/10.7454/nrjs.v9i1.1203>.
- Yusri, T. A. S., Rudhistiar, D., & Ratnasari, A. P. (2025). Klasifikasi Varietas Kacang Menggunakan XGBoost dengan Penyesuaian Class Weighting. *Jurnal MNEMONIC*, 8(1), 115–122. <https://doi.org/10.36040/mnemonic.v8i1.12788>.
- Zhang, J. & Xu, S. (2024). High Aggressiveness of Papillary Thyroid Cancer: From Clinical Evidence to Regulatory Cellular Networks. *Cell Death Discovery*, 10(378), 1–15. <https://doi.org/10.1038/s41420-024-02157-2>.
- Zhang, P., Jia, Y., & Shang, Y. (2022). Research and Application of XGBoost in Imbalanced Data. *International Journal of Distributed Sensor Networks*, 18(6), 1–10. <https://doi.org/10.1177/15501329221106935>.
- Zhao, X., Zhao, M., & Zhang, G. (2022). Risk Factors of Lymph Node Metastasis in the Central Region in Papillary Thyroid Microcarcinoma. *Nepalese Journal of Cancer*, 6(2), 32–35. <https://doi.org/10.3389/fendo.2023.1288527>.

UIN SUNAN AMPEL
SURABAYA