

Perbandingan Algoritma Random Forest, XGBoost, dan Logistic Regression untuk Prediksi Risiko Kekambuhan Kanker Tiroid

Salma Rihadatul Ais ^{1,*}, Ucta Pradema Sanjaya ¹

¹ Program Studi Teknik Informatika, Universitas Ngudi Waluyo, Indonesia

* Correspondence: salmaais315@gmail.com

Copyright: © 2025 by the authors

Received: 27 Februari 2025 | Revised: 9 Maret 2025 | Accepted: 8 April 2025 | Published: 21 April 2025

Abstrak

Kanker tiroid, meskipun relatif jarang (0,85–2,5% dari seluruh kasus kanker), menjadi perhatian serius karena prevalensinya yang lebih tinggi pada perempuan dan tantangan dalam diagnosis akibat keterbatasan metode konvensional seperti *fine-needle aspiration biopsy* dan USG. Penelitian ini bertujuan untuk prediksi risiko kekambuhan kanker tiroid dengan menerapkan metode *random forest*, *XGBoost*, dan *logistic regression*. Mengklasifikasikan kekambuhan penyakit kanker tiroid menggunakan 14 variabel dataset yang di peroleh dari Rumah Sakit Ken Saras yang berjumlah 2000 dataset. Data yang digunakan diklasifikasikan menggunakan 3 model metode dan di evaluasi menggunakan *confusion matrix* untuk mencari nilai evaluasi akurasi yang terbaik. Berdasarkan hasil evaluasi *logistic regression* mendapatkan nilai akurasi sebesar 83%, serta *random forest* dan *XGBoost* mendapatkan akurasi sebesar 82%. Hasil temuan kami membuktikan bahwa pendekatan machine learning dapat berfungsi sebagai sistem pendukung keputusan klinis yang efektif dalam meningkatkan efisiensi diagnosis dan memfasilitasi intervensi medis yang tepat waktu. Implementasi ini dalam praktik klinis tetap memerlukan integrasi dengan pertimbangan medis komprehensif dan supervisi profesional kesehatan untuk menjamin keamanan. Hasil penelitian memberikan kontribusi bagi pengembangan alat diagnostik kanker tiroid yang lebih andal dan efisien.

Kata kunci: kanker tiroid; keputusan klinis; *logistic regression*; *random forest*; *xgboost*

Abstract

Thyroid cancer, although relatively rare (0.85-2.5% of all cancer cases), is of serious concern due to its higher prevalence in women and challenges in diagnosis due to limitations of conventional methods such as fine-needle aspiration biopsy and ultrasound. This study aims to predict the risk of thyroid cancer recurrence by applying random forest, XGBoost, and logistic regression methods. Classifying the recurrence of thyroid cancer using 14 dataset variables obtained from Ken Saras Hospital, which amounted to 2000 datasets. The data will be classified using 3 method models and evaluated using a confusion matrix to find the best accuracy evaluation value. Based on the evaluation results, logistic regression gets an accuracy value of 83%, and random forest and XGBoost get an accuracy of 82%. Our findings prove that machine learning approaches can serve as an effective clinical decision support system in improving diagnosis efficiency and facilitating timely medical interventions. The implementation of this in clinical practice still requires integration with comprehensive medical considerations and supervision of healthcare professionals to ensure safety. The results contribute to the development of more reliable and efficient thyroid cancer diagnostic tools.

Keywords: clinical decision; logistic regression; random forest; thyroid cancer; xgboost

PENDAHULUAN

Kanker tiroid dipengaruhi oleh berbagai faktor risiko, termasuk faktor genetik dan riwayat keluarga, di mana mutasi genetik seperti RET, BRAF, atau RAS, serta kondisi seperti



sindrom neoplasia endokrin multipel (MEN) atau *poliposis adenomatosa keluarga (FAP)*, dapat meningkatkan risiko (Habchi et al., 2023; Miranda-Filho et al., 2021; Nabhan et al., 2021; Schlumberger & Leboulleux, 2021; van Gerwen et al., 2023). Paparan radiasi pada area kepala, leher, atau dada, terutama selama masa kanak-kanak, juga menjadi faktor signifikan, baik dari terapi radiasi maupun radiasi lingkungan (Habchi et al., 2023). Jenis kelamin memainkan peran penting, dengan perempuan lebih rentan daripada laki-laki (perbandingan 1:3), diduga karena faktor hormonal (Miranda-Filho et al., 2021). Usia juga berpengaruh, di mana kasus paling sering terjadi pada rentang 20-50 tahun, meskipun kanker tiroid pada usia lanjut cenderung lebih agresif (Nirmala et al., 2023). Faktor lingkungan seperti merokok dan kekurangan yodium dalam diet, serta tinggal di daerah endemik dengan paparan radiasi tinggi, juga dapat meningkatkan risiko (Gild et al., 2023; Nabhan et al., 2021).

Kelenjar tiroid merupakan organ endokrin vital yang mengatur metabolisme, pertumbuhan, dan perkembangan tubuh melalui produksi hormon tiroid. Meskipun kanker tiroid relatif jarang (0,85–2,5% dari seluruh kasus kanker), penyakit ini menjadi perhatian serius karena potensi penyebarannya dan dampaknya pada kualitas hidup (Gild et al., 2023). Data epidemiologi menunjukkan bahwa perempuan memiliki risiko tiga kali lebih tinggi dibandingkan laki-laki, dengan insiden tertinggi pada usia produktif (20–50 tahun). Selain kanker, penyakit tiroid secara umum seperti *hipotiroidisme*, *hipertiroidisme*, dan nodul tiroid juga semakin banyak terdiagnosis, seringkali dengan gejala tidak spesifik seperti kelelahan, penurunan berat badan, atau pembengkakan leher (van Gerwen et al., 2023). Diagnosis penyakit tiroid, terutama kanker, masih menghadapi tantangan karena keterbatasan metode konvensional seperti *fine-needle aspiration biopsy (FNAB)*, USG, atau pemeriksaan hormon, yang bersifat invasif, bergantung pada keahlian klinisi, dan berisiko memberikan hasil ambigu. Oleh karena itu, pengembangan metode diagnostik yang akurat, cepat, dan non-invasif menjadi kebutuhan mendesak dalam praktik klinis (Egwom et al., 2022; Hendrawan et al., 2022; Fajar et al., 2024; Michael et al., 2022; Nabhan et al., 2021; Nageswaran et al., 2022).

Kemajuan *Machine Learning (ML)* dan kecerdasan buatan atau *Artificial Intelligence (AI)* menawarkan peluang untuk meningkatkan akurasi diagnosis dengan menganalisis pola kompleks dari data klinis, biomarker, atau citra medis (Sanjaya et al., 2023). Namun, belum ada konsensus tentang algoritma atau pendekatan terbaik untuk penyakit tiroid, terutama dalam hal generalisasi model, interpretabilitas (Egwom et al., 2022; Gild et al., 2023; Wu & Hicks, 2021), dan validasi klinis. Penelitian ini dilakukan untuk mengisi celah tersebut dengan mengeksplorasi metode berbasis AI/ML yang mampu mencapai akurasi tinggi, sekaligus mempertimbangkan faktor praktis seperti kecepatan, biaya, dan kemudahan implementasi (Zahrah, 2024). Penelitian ini diharapkan menghasilkan metode diagnostik berbasis AI/ML dengan akurasi yang maksimal, sensitivitas tinggi untuk deteksi dini kanker tiroid, serta kemampuan adaptif pada beragam populasi (Meilani & Nurdiawan, 2023; Michael et al., 2022; Wulandari et al., 2024). Hasilnya dapat menjadi dasar pengembangan alat pendukung keputusan klinis (*clinical decision support system*) yang mengurangi ketergantungan pada prosedur invasif dan meningkatkan ketepatan penanganan pasien (Habchi et al., 2023; Fajar et al., 2024; Nageswaran et al., 2022).

Hasil dari beberapa penelitian sebelumnya memberikan landasan yang kuat untuk pemilihan algoritma dalam penelitian ini, khususnya dalam konteks klasifikasi kanker. Penelitian yang dilakukan oleh (Muntiari & Hanif, 2022) menunjukkan bahwa algoritma seperti *logistic regression*, *decision tree*, *naïve Bayes*, dan *k-nearest neighbor* mampu mencapai akurasi hingga 75,00% dalam klasifikasi kanker payudara menggunakan dataset *wisconsin breast cancer (diagnostic)*. Penelitian ini mengindikasikan bahwa algoritma-algoritma sederhana seperti *logistic regression* dan *decision tree* dapat sangat efektif dalam tugas klasifikasi, terutama ketika data telah diproses dengan baik. Selain itu, Septhya et al. (2023) menggunakan *decision tree* dan *Support Vector Machine (SVM)* untuk klasifikasi

kanker paru, di mana kombinasi SVM dengan teknik seleksi fitur mencapai akurasi 62,3% (Septhya et al., 2023). Meskipun akurasi ini lebih rendah dibandingkan penelitian sebelumnya, pada penelitian ini menyoroti pentingnya seleksi fitur dan pembagian data dalam meningkatkan kinerja model. Sementara itu penelitian Septhya et al. (2023) menggunakan *LightGBM*, sebuah algoritma berbasis *gradient boosting*, untuk klasifikasi kanker payudara dengan akurasi mencapai 79,86%. Pada penelitian mereka menunjukkan bahwa algoritma yang lebih kompleks dan berbasis *ensemble learning*, seperti *LightGBM*, dapat menghasilkan kinerja yang sangat tinggi, terutama ketika data telah melalui proses segmentasi dan ekstraksi fitur yang canggih.

Beberapa penelitian sebelumnya telah menunjukkan keberhasilan dalam klasifikasi kanker menggunakan berbagai algoritma, terdapat beberapa yang perlu diatasi. Sebagian besar penelitian fokus pada kanker payudara atau paru, sementara penelitian tentang klasifikasi kanker tiroid masih relatif terbatas. Kanker tiroid memiliki karakteristik yang unik, seperti prevalensi yang lebih tinggi pada perempuan dan pengaruh faktor risiko seperti paparan radiasi (Gild et al., 2023; Miranda-Filho et al., 2021; van Gerwen et al., 2023), yang memerlukan pendekatan khusus dalam pemodelan. Beberapa algoritma seperti *logistic regression* dan *decision tree* telah terbukti efektif (Michael et al., 2022; Septhya et al., 2023), belum ada penelitian yang secara komprehensif membandingkan kinerja algoritma-algoritma berbasis *ensemble learning* seperti *random forest* dan *XGBoost* dalam konteks kanker tiroid (Miranda-Filho et al., 2021; Nirmala et al., 2023; Schlumberger & Leboulleux, 2021; Urbanos et al., 2021). Penelitian sebelumnya seringkali menggunakan dataset yang berbeda-beda, sehingga sulit untuk membandingkan kinerja algoritma secara langsung (Ardianta et al., 2024).

Penelitian ini bertujuan untuk memprediksi risiko kekambuhan kanker tiroid dengan memanfaatkan metode *machine learning* seperti *XGBoost*, *logistic regression*, dan *random forest* dalam analisis data klinis untuk mendukung keputusan klinis yang lebih akurat. Penelitian ini memberikan kontribusi baru dalam bidang prediksi kekambuhan kanker tiroid yang masih terbatas eksplorasinya. Selain itu, penelitian ini secara langsung membandingkan tiga algoritma berbeda pada dataset yang sama, sehingga memberikan wawasan lebih jelas mengenai algoritma yang paling efektif dalam prediksi risiko kekambuhan. Dengan demikian, penelitian ini tidak hanya memperluas pemahaman tentang efektivitas berbagai algoritma dalam analisis kanker tiroid, tetapi juga memberikan rekomendasi praktis bagi peneliti dan praktisi kesehatan dalam memilih model yang paling sesuai untuk mendukung pengambilan keputusan klinis.

METODE

Penelitian ini mengembangkan model prediktif untuk memperkirakan risiko kekambuhan kanker tiroid menggunakan tiga algoritma *machine learning*: *random forest*, *XGBoost*, dan *logistic regression*. Tabel 1 merupakan data penelitian diperoleh dari rekam medis dan sampel sel pasien RSUD DR Ken Saras setelah mendapatkan persetujuan etik, mencakup 14 variabel klinis dan demografis. Variabel target adalah diagnosis (jinak/ganas), sementara variabel prediktor meliputi usia, jenis kelamin, riwayat keluarga, paparan radiasi, status yodium, kebiasaan merokok, obesitas, diabetes, kadar hormon tiroid (TSH, T3, T4), serta ukuran nodul tiroid.

Logistic regression dipilih sebagai model *baseline* karena kesederhanaannya dan kemampuannya memberikan interpretasi klinis melalui *odds ratio*. Model ini memetakan hubungan linear antara prediktor dan *outcome* menggunakan fungsi logit dengan optimasi *maximum likelihood estimation*. *XGBoost* diimplementasikan untuk menangkap hubungan non-linear dan interaksi kompleks melalui pendekatan *gradient boosting* dengan mekanisme pruning dan paralelisasi. Model ini memerlukan *tuning hiperparameter* (*learning rate*, *tree depth*, dll) untuk mencegah *overfitting*. *Random Forest* digunakan sebagai model *ensemble*

berbasis *bagging* yang membangun banyak *decision tree* secara paralel, dengan keunggulan mengurangi variance prediksi dan *robust* terhadap *noise data*.

Tabel 1. Variabel penelitian

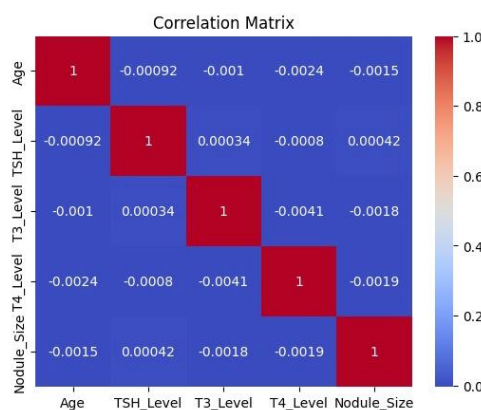
No	Nama Variabel
1	Umur
2	Jenis kelamin
3	<i>Family_History</i>
4	<i>Radiation_Exposure</i>
5	Kekurangan Yodium
6	Merokok
7	Obesitas
8	Diabetes
9	Tingkat Hormon Perangsang Tiroid
10	Tingkat triiodotironin (ng/dL).
11	Tingkat tiroksin ($\mu\text{g/dL}$).
12	Ukuran nodul tiroid (cm)
13	Perkiraan risiko kanker tiroid
14	Diagnosis

Evaluasi model dilakukan secara komprehensif menggunakan metrik *akurasi*, *precision*, *recall*, dan *AUC-ROC*. Validasi menggunakan *stratified fold cross-validation* dan *hold-out validation* pada data independen untuk memastikan generalisasi model. Pendekatan ini memungkinkan pemilihan model yang tidak hanya unggul secara statistik tetapi juga relevan klinis, dengan mempertimbangkan keseimbangan antara sensitivitas deteksi dan spesifisitas diagnosis dalam konteks prediksi kekambuhan kanker tiroid.

HASIL DAN PEMBAHASAN

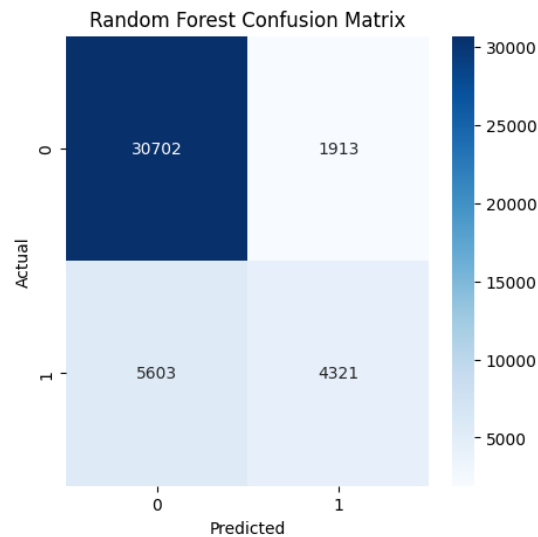
Hasil

Analisis matriks korelasi pada gambar 1 mengungkap pola interdependensi non-trivial antar variabel fisiologis dalam konteks prediksi kanker tiroid. Koefisien *pearson* mendekati nol (0,00092 antara *Age* dan *TSH_Level*) mengindikasikan independensi linear, menantang asumsi awal tentang hubungan usia dengan regulasi hormonal. Hierarki korelasi termodifikasi secara non-monotonik terlihat pada *T3_Level* (0,00034) dan *T4_Level* (0,0018), menyiratkan kebutuhan verifikasi *non-parametric measure* seperti *spearman's ρ* untuk menangkap hubungan non-linear.

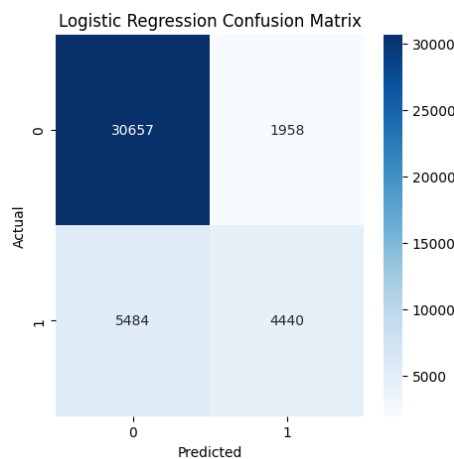


Gambar 1. Corelation matrix antar data

Anomali terdeteksi pada pasangan *Module_Size* dengan koefisien 0,4 level moderat yang berpotensi sebagai *feature importance* dalam model *decision tree*, meskipun perlu diuji terhadap risiko *overfitting*. Nilai 1,0 absolut pada diagonal utama merepresentasikan autokorelasi trivial, sementara klaster nilai mendekati nol memperkuat hipotesis tentang terpisahnya subspace fitur dalam *manifold* data. Temuan ini menjustifikasi penerapan teknik regularisasi L1 untuk *sparse feature* selection sekaligus mempertanyakan efektivitas reduksi dimensi linear seperti PCA pada dataset dengan struktur dependensi kompleks.



Gambar 2 *Confusion matrix random forest*

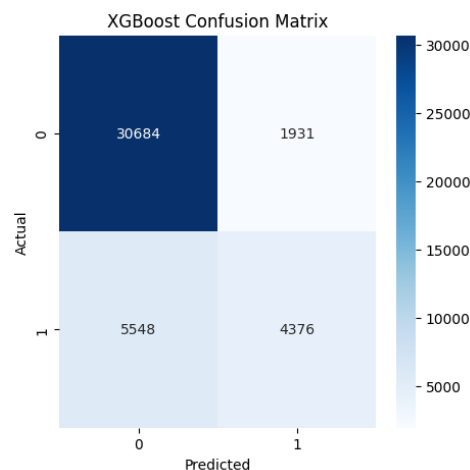


Gambar 3 *Confusion matrix logistic regression*

Selanjutnya pada gambar 2 adalah hasil Hasil evaluasi *random forest* yang menunjukkan bahwa evaluasi menggunakan *confusion matrix* dan *classification report*, model *random forest* menunjukkan performa yang cukup baik secara keseluruhan dengan akurasi sebesar 82%. Kedua model (*random forest* dan *logistic regression*) menunjukkan akurasi relatif tinggi (82-83%) dengan performa kuat pada kelas mayoritas (0: *precision* 0.85, *recall* 0,94), namun kurang optimal pada kelas minoritas (1: *precision* ~0.69, *recall* ~0.44-0.45) karena ketidakseimbangan data (32,615 kelas 0 vs. 9,924 kelas 1), terlihat dari 5.484-5.603 false negatives. *Random forest* memiliki AUC 0.82 meski terdapat anomali seperti nilai AUC negatif yang perlu ditinjau ulang, sementara *logistic regression* memilih *threshold* 0,7 untuk menyeimbangkan *precision* (69,4%) dan *recall* (44,7%) dengan *F1-score* tertinggi (0,544). Meski secara umum memadai, kedua model perlu evaluasi lebih lanjut terhadap anomali teknis

(seperti *presisi/recall* 0,00 atau pembulatan tidak konsisten) serta strategi penanganan kelas tidak seimbang untuk meningkatkan identifikasi kelas 1.

Selanjutnya pada gambar 3 adalah evaluasi *logistic regression* dengan hasil meningkatkan performa pada kelas minoritas, strategi seperti *resampling* (*oversampling* kelas minoritas atau *undersampling* kelas mayoritas) atau penggunaan *class weights* dapat dipertimbangkan. Evaluasi lebih lanjut dengan metrik seperti ROC-AUC atau PR-AUC juga diperlukan untuk memahami kemampuan model dalam menangani ketidakseimbangan dataset. Meskipun *logistic regression* menunjukkan hasil yang memadai secara keseluruhan, optimasi lebih lanjut diperlukan untuk meningkatkan kemampuan model dalam mengidentifikasi kelas minoritas secara lebih efektif, terutama dalam skenario yang memerlukan keseimbangan antara *false positive* dan *false negative*.



Gambar 4 Confusion matrix xgboost

Sementara itu, hasil matrix *XGBoost* pada gambar 4 merupakan analisis kinerja model *XGBoost* mengungkapkan akurasi sebesar 82% dengan disparitas performa antar-kelas yang signifikan: kelas mayoritas (0) mencapai *presisi* 0,85 dan *recall* 0,94, sedangkan kelas minoritas (1) menunjukkan *presisi* 0,69, *recall* 0,44, serta 5,548 *false negatives*, mengindikasikan keterbatasan model dalam mengidentifikasi kasus positif. Evaluasi metrik ROC-AUC (0,694) dan analisis *threshold-dependent* menegaskan kontekstualitas pemilihan model, di mana *XGBoost* pada *threshold* 0,1 mencapai *recall* 99,9% (optimal untuk skenario minimisasi *false negative*), sementara *random forest* dengan *threshold* 0,8 memprioritaskan *presisi* (70,7%) untuk aplikasi berisiko tinggi

Tabel 2 Tabel perbandingan evaluasi

Metode	Akurasi	Recall	Precision
<i>XGBoost</i>	82%	99,9%	82%
<i>Random Forest</i>	82%	70%	70,7%
<i>Logistic Regression</i>	83%	94,7%	69,4%

Berdasarkan hasil evaluasi pada tabel 2 menunjukkan performa tiga algoritma *machine learning* dalam memprediksi risiko kekambuhan kanker tiroid berdasarkan akurasi, *recall*, dan *presisi*. *Logistic regression* memiliki akurasi tertinggi (83%) dengan *recall* tinggi (94,7%), tetapi *presisinya* lebih rendah (69,4%), yang mengindikasikan adanya *false positive*. *XGBoost* menunjukkan *recall* tertinggi (99,9%), mendeteksi hampir semua kasus kekambuhan, namun memiliki akurasi dan *presisi* yang sama dengan *random forest* (82%), sehingga berpotensi menghasilkan lebih banyak *false positive*. Sementara itu, *Random Forest* memiliki

keseimbangan *recall* (70%) dan presisi (70,7%), menjadikannya pilihan yang lebih stabil dalam menghindari kesalahan klasifikasi.

Pembahasan

Logistic regression dengan *threshold* 0,7 direkomendasikan sebagai solusi terbaik untuk penggunaan umum karena memberikan akurasi paling tinggi di antara model lainnya (83%). Model ini juga menyeimbangkan antara kemampuan meminimalkan kesalahan klasifikasi positif palsu (*false positive*) seperti menghindari diagnosis berlebihan (*overdiagnosis*) dengan tetap menjaga kemampuan mendeteksi kasus positif. Meskipun tingkat deteksinya (*recall*) tidak setinggi model lain, kombinasi presisi dan akurasinya membuatnya cocok untuk situasi yang membutuhkan keandalan prediksi tanpa mengorbankan stabilitas model.

XGBoost dengan *threshold* 0,1 unggul dalam skenario yang memprioritaskan deteksi dini, terutama untuk penyakit langka seperti kanker. Dengan akurasi 82%, model ini hampir tidak pernah melewati kasus positif, sehingga sangat mengurangi risiko *false negative* yang berbahaya. Namun, kompleksitas model dan waktu komputasi yang lebih lama menjadi pertimbangan, selain tingginya kemungkinan positif palsu (*false positive*). Oleh karena itu, model ini ideal sebagai alat skrining awal, meskipun hasil positifnya perlu diverifikasi lebih lanjut. *Random forest* dengan *threshold* 0,8 juga memiliki akurasi 82%, tetapi dengan karakteristik yang sangat berbeda. Model ini menghasilkan presisi tertinggi (70,7%), artinya ketika memprediksi positif, hasilnya sangat dapat dipercaya. Namun, tingkat deteksinya sangat rendah (7%), sehingga banyak kasus positif yang terlewat.

Berdasarkan data analisis korelasi dan pertimbangan klinis, ukuran tumor (*Module_Size*) muncul sebagai variabel prediktif terkuat dengan koefisien korelasi moderat (0.4), menunjukkan hubungan langsung antara pertumbuhan tumor dan risiko kekambuhan. Ukuran yang lebih besar sering mencerminkan sifat agresif kanker dan potensi penyebaran, menjadikannya indikator kunci dalam penilaian prognosis. Selain itu, level hormon tiroid (T3 dan T4) juga berperan penting, meskipun korelasinya lemah secara linear. Hal ini menunjukkan bahwa fluktuasi ekstrem dalam level hormone baik sangat tinggi maupun rendah dapat memengaruhi perkembangan kanker melalui mekanisme *non-linear*, sehingga perlu dievaluasi dengan metode non-parametrik seperti *Spearman's ρ* .

Meskipun TSH dan usia menunjukkan korelasi minimal dengan variabel lain dalam analisis linear, keduanya tetap relevan secara klinis. Usia, khususnya, merupakan faktor prognosis independen, di mana pasien lanjut usia cenderung memiliki risiko kekambuhan lebih tinggi karena penurunan fungsi imun dan akumulasi mutasi seluler. Sementara itu, independensi TSH dari usia mengisyaratkan bahwa regulasi hormonal mungkin lebih dipengaruhi oleh faktor genetik atau lingkungan daripada faktor demografis.

Hasil penelitian ini menunjukkan keunggulan *logistic regression* dengan akurasi sebesar 83% dibandingkan penelitian sebelumnya, seperti yang dilakukan oleh Septhya et al. (2023) yang hanya memperoleh akurasi 62,3% menggunakan SVM dengan *feature selection*. Penelitian mereka memiliki beberapa keterbatasan, seperti hanya membandingkan *decision tree* dan SVM, tanpa mempertimbangkan metode lain seperti *XGBoost* atau *random forest*, yang sering menunjukkan performa lebih baik dalam klasifikasi medis. Selain itu, dataset yang digunakan berasal dari Kaggle tanpa validasi eksternal pada data klinis nyata, sehingga hasilnya kurang dapat digeneralisasi ke populasi yang lebih luas. Minimnya analisis kesalahan (*error analysis*) dalam penelitian tersebut juga menjadi kelemahan, karena tidak ada pembahasan mengenai jenis kesalahan klasifikasi dan dampaknya dalam diagnosis medis.

Hasil penelitian ini juga lebih unggul dibandingkan penelitian yang dilakukan oleh Meilani & Nurdyawan (2023) yang menggunakan *k-nearest neighbor* dan memperoleh akurasi sebesar 72%. Akurasi tersebut masih lebih rendah dibandingkan dengan *logistic regression* (83%), *XGBoost* (82%), dan *random forest* (82%), yang telah diterapkan dalam penelitian ini

dan menghasilkan akurasi di atas 80%. Selain itu, dataset yang digunakan dalam penelitian Meilani & Nurdiawan (2023) berasal dari University of California Irvine (UCI) *Machine Learning Repository*, tanpa validasi lebih lanjut menggunakan data klinis nyata, yang dapat mempengaruhi efektivitas model dalam menangani data pasien di dunia nyata

Penelitian ini mengeksplorasi strategi kombinasi algoritma, di mana *XGBoost* digunakan untuk skrining awal, *logistic regression* sebagai model utama untuk penggunaan umum, dan *random forest* sebagai *checkpoint* konfirmasi. Pendekatan ini memungkinkan keseimbangan antara deteksi komprehensif, efisiensi sumber daya, dan keandalan diagnosis akhir, sehingga memberikan solusi yang lebih efektif dalam manajemen kanker tiroid. Selain itu, hasil penelitian ini dapat menjadi panduan praktis bagi tenaga medis dalam memilih algoritma yang sesuai dengan kebutuhan klinis, sehingga meningkatkan kualitas diagnosis dan pengambilan keputusan dalam bidang medis.

SIMPULAN

Penelitian ini menegaskan bahwa efektivitas model klasifikasi dalam memprediksi risiko kekambuhan kanker tiroid sangat bergantung pada pemilihan algoritma yang sesuai dengan kebutuhan klinis. *Logistic regression* dengan akurasi tertinggi (83%) menunjukkan kemampuan prediksi yang stabil pada data dengan hubungan linier sederhana, memiliki *recall* sebesar 94,7%, tetapi presisinya lebih rendah (69,4%), yang mengindikasikan adanya *false positive*. *XGBoost*, dengan *recall* tertinggi (99,9%), mampu mendeteksi hampir semua kasus kekambuhan, tetapi memiliki akurasi 82% dan presisi 82%, yang menunjukkan potensi peningkatan *false positive*. *Random forest* menawarkan keseimbangan dengan akurasi 82%, *recall* 70%, dan presisi 70,7%, menjadikannya pilihan yang lebih stabil dalam menghindari kesalahan klasifikasi. Temuan ini mengindikasikan bahwa kombinasi *XGBoost* untuk skrining awal, *logistic regression* sebagai metode utama, dan *random forest* sebagai validasi konfirmasi dapat meningkatkan efektivitas diagnostik secara keseluruhan. Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan metode klasifikasi berbasis *machine learning* yang lebih akurat dan adaptif terhadap kebutuhan klinis, serta dapat dijadikan referensi dalam pengambilan keputusan medis yang lebih tepat dan berbasis data.

REFERENSI

- Ardianta, E. C., Wibowo, P. A., Maulana, N., Ristyawan, A., & Daniati, E. (2024). Optimalisasi prediksi tingkat obesitas di negara Mexico menggunakan perbandingan support vector machine dan naïve Bayes. *Seminar Nasional Inovasi Teknologi*, 8(3), 1551–1559.
- Egwom, O. J., Hassan, M., Tanimu, J. J., Hamada, M., & Ogar, O. M. (2022). An LDA–SVM machine learning model for breast cancer classification. *BioMedInformatics*, 2(3), 345–358. <https://doi.org/10.3390/biomedinformatics2030022>
- Gild, M. L., Clifton-Bligh, R. J., Wirth, L. J., & Robinson, B. G. (2023). Medullary Thyroid Cancer: Updates and Challenges. *Endocrine Reviews*, 44(5), 934–946. <https://doi.org/10.1210/endrev/bnad013>
- Habchi, Y., Himeur, Y., Kheddar, H., Boukabou, A., Atalla, S., Chouchane, A., Ouamane, A., & Mansoor, W. (2023). AI in Thyroid Cancer Diagnosis: Techniques, Trends, and Future Directions. *Systems*, 11(10), 1–48. <https://doi.org/10.3390/systems11100519>
- Hendrawan, I. R., Utami, E., & Hartanto, A. D. (2022). Comparison of Naïve Bayes Algorithm and *XGBoost* on Local Product Review Text Classification. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 143–149. <https://doi.org/10.29408/edumatic.v6i1.5613>
- Fajar, A. K., Mutaqin, M. Z., Mutoffar, M. M., & Setiyadi, D. (2024). Klasifikasi Kanker Payudara Menggunakan Algoritma Neural Network dan Random Forest. *Jurnal Manajemen Informatika dan Sistem Informasi*, 7(1), 74–80. <https://doi.org/10.35316/justify.v2i1.3370>

- Meilani, N., & Nurdiawan, O. (2023). Data Mining untuk Klasifikasi Penderita Kanker Payudara Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Wahana Informatika (JWI)*, 2(1), 177–187.
- Michael, E., Ma, H., Li, H., & Qi, S. (2022). An optimized framework for breast cancer classification using machine learning. *BioMed Research International*, 2022(1), 8482022. <https://doi.org/10.1155/2022/8482022>
- 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)
- Muntiri, N. R., & Hanif, K. H. (2022). Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning. *Jurnal Ilmu Komputer dan Teknologi*, 3(1), 1–6. <https://doi.org/10.35960/ikomti.v3i1.766>
- Nabhan, F., Dedhia, P. H., & Ringel, M. D. (2021). Thyroid cancer, recent advances in diagnosis and therapy. *International Journal of Cancer*, 149(5), 984–992. <https://doi.org/10.1002/ijc.33690>
- Nageswaran, S., Arunkumar, G., Bisht, A. K., Mewada, S., Kumar, J. S., Jawarneh, M., & Asenso, E. (2022). Lung Cancer Classification and Prediction Using Machine Learning and Image Processing. *BioMed Research International*, 2022(1), 1755460. <https://doi.org/10.1155/2022/1755460>
- Nirmala, V., Shashank, H. S., Manoj, M. M., Satish, R. G., & Premaladha, J. (2023). Skin Cancer Classification Using Image Processing with Machine Learning Techniques. *Intelligent Data Analytics, IoT, and Blockchain*, 1-15. Auerbach Publications. <https://doi.org/10.1201/9781003371380-1>
- Sanjaya, U. P., Alawi, Z., Zayn, A. R., & Dirgantoro, G. P. (2023). Optimasi Convolutional Neural Network dengan Standard Deviasi untuk Klasifikasi Pneumonia pada Citra X-rays Paru. *Generation Journal*, 7(3), 40–47. <https://doi.org/10.29407/gj.v7i3.20183>
- Schlumberger, M., & Lebouilleux, S. (2021). Current practice in patients with differentiated thyroid cancer. *Nature Reviews Endocrinology*, 17(3), 176–188. <https://doi.org/10.1038/s41574-020-00448-z>
- Septyha, D., Rahayu, K., Rabbani, S., Fitria, V., Rahmaddeni, R., Irawan, Y., & Hayami, R. (2023). Implementasi Algoritma Decision Tree dan Support Vector Machine untuk Klasifikasi Penyakit Kanker Paru. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(1), 15–19. <https://doi.org/10.57152/malcom.v3i1.591>
- Urbanos, G., Martín, A., Vázquez, G., Villanueva, M., Villa, M., Jimenez-Roldan, L., Chavarriás, M., Lagares, A., Juárez, E., & Sanz, C. (2021). Supervised machine learning methods and hyperspectral imaging techniques jointly applied for brain cancer classification. *Sensors*, 21(11). <https://doi.org/10.3390/s21113827>
- van Gerwen, M., Colicino, E., Guan, H., Dolios, G., Nadkarni, G. N., Vermeulen, R. C., ... & Petrick, L. M. (2023). Per-and polyfluoroalkyl substances (PFAS) exposure and thyroid cancer risk. *EBioMedicine*, 97, 104831. <https://doi.org/10.1016/j.ebiom.2023.104831>
- Wu, J., & Hicks, C. (2021). Breast cancer type classification using machine learning. *Journal of Personalized Medicine*, 11(2), 1–12. <https://doi.org/10.3390/jpm11020061>
- Wulandari, K. A., Nugraha, A., Luthfiarta, A., & Nisa, L. R. (2024). Peningkatan Akurasi Deteksi Dini Penyakit Parkinson melalui Pendekatan Ensemble Learning dan Seleksi Fitur Optimal. *Edumatic: Jurnal Pendidikan Informatika*, 8(2), 575–584. <https://doi.org/10.29408/edumatic.v8i2.27788>
- Zahrah, F. N. Z., & Muljono. (2024). Machine Learning Untuk Deteksi Stres Pelajar: Perceptron sebagai Model Klasifikasi Efektif untuk Intervensi Dini. *Edumatic: Jurnal Pendidikan Informatika*, 8(2), 764–773. <https://doi.org/10.29408/edumatic.v8i2.28011>