

Penerapan Algoritma *Random Forest* untuk Klasifikasi Tingkat Keparahan Penyakit pada Data Rekam Medis

¹Fitri Aini Nasution, ²Angga Putra Juledi

¹Manajemen Informatika, Fakultas Sains dan Teknologi, Universitas Labuhanbatu

²Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Labuhanbatu

Email : fitriaininasution689@gmail.com, anggapjl9@gmail.com

Corresponding Author: fitriaininasution689@gmail.com

Abstract

Accurate determination of disease severity is an important step in supporting medical decision-making. This study aims to classify the severity of patients' diseases into three categories—Mild, Moderate, and Severe—using the Random Forest algorithm. The data used were obtained from patients' medical records containing structured clinical parameters and have undergone a preprocessing stage, including data cleaning, variable transformation, and splitting into training data (80%) and testing data (20%). The test results show that the Random Forest model achieved an accuracy of 74.77%. The best performance was obtained in the Mild class with a recall value of 0.95 and an f1-score of 0.84. The Moderate class achieved a recall of 0.71 and an f1-score of 0.73, while the Severe class showed perfect precision (1.00) but a low recall (0.12), indicating the model's limited ability to detect cases in this class. The macro average values for precision, recall, and f1-score were 0.83, 0.60, and 0.59 respectively, while the weighted average values were 0.78, 0.75, and 0.71 respectively. These findings indicate that Random Forest can be used to classify disease severity based on medical records with relatively good performance for the majority class, but further optimization—such as data balancing or parameter adjustment—is needed to improve sensitivity toward classes with fewer samples.

Keywords : Medical Records, Disease Severity, Random Forest, Classification, Machine Learning.

1. Pendahuluan

Kemajuan teknologi informasi dan analisis data telah memberikan kontribusi yang signifikan dalam bidang kesehatan, khususnya pada proses pengambilan keputusan medis. Rekam medis pasien, yang berisi informasi terstruktur seperti hasil pemeriksaan laboratorium, tekanan darah, suhu tubuh, serta riwayat medis, merupakan sumber data yang kaya dan potensial untuk dianalisis. Pemanfaatan data ini dengan pendekatan *machine learning* dapat membantu tenaga medis mengidentifikasi tingkat keparahan penyakit pasien secara lebih cepat, akurat, dan objektif. Penentuan tingkat keparahan penyakit merupakan aspek penting dalam manajemen layanan kesehatan. Pasien dengan tingkat keparahan *Ringan*, *Sedang*, dan *Berat* membutuhkan penanganan yang berbeda baik dari segi sumber daya medis maupun prioritas perawatan. Kesalahan dalam mengklasifikasikan tingkat keparahan dapat berdampak pada keterlambatan penanganan atau penggunaan sumber daya yang tidak efisien. Oleh karena itu, dibutuhkan sistem pendukung keputusan yang mampu melakukan klasifikasi tingkat keparahan dengan performa yang andal. Metode *Random Forest* dipilih dalam penelitian ini karena memiliki kemampuan untuk menangani data berukuran besar, memproses variabel numerik dan kategorikal secara bersamaan, serta mengurangi risiko *overfitting* melalui penggunaan banyak pohon

keputusan (*decision trees*) yang digabungkan. Selain itu, *Random Forest* dikenal memiliki kinerja yang baik pada berbagai permasalahan klasifikasi, termasuk pada data medis yang sering kali memiliki kompleksitas tinggi. Penelitian ini menggunakan data rekam medis pasien yang telah melalui tahap pra-pemrosesan, meliputi pembersihan data, normalisasi, dan pembagian menjadi data latih serta data uji. Model *Random Forest* kemudian dibangun dan dievaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *f1-score*. Hasil awal menunjukkan akurasi sebesar 74,77%, dengan performa yang bervariasi pada setiap kelas. Temuan ini mengindikasikan bahwa meskipun *Random Forest* mampu memberikan hasil prediksi yang cukup baik, diperlukan upaya optimasi lebih lanjut, khususnya untuk meningkatkan sensitivitas pada deteksi kasus dengan tingkat keparahan tinggi. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan sistem klasifikasi tingkat keparahan penyakit berbasis *machine learning*, yang dapat digunakan sebagai alat bantu dalam pengambilan keputusan medis dan manajemen layanan kesehatan.

2. Landasan Teori

Metode *Random Forest*

Metode *Random Forest* adalah salah satu algoritma machine learning yang termasuk dalam kategori ensemble learning. Ensemble learning melibatkan penggabungan hasil dari beberapa model untuk meningkatkan kinerja dan ketepatan prediksi dibandingkan dengan penggunaan satu model tunggal. Dalam konteks *Random Forest*, model yang digunakan adalah pohon keputusan (*decision trees*). Saat ini, teknologi sedang berkembang pesat (Siregar et al., 2023).

Rekam Medis

Rekam medis merupakan berkas yang berisi catatan dan dokumen identitas pasien, hasil pemeriksaan, pengobatan yang telah diberikan serta tindakan dan pelayanan lain yang telah diberikan kepada pasien. Menurut Permenkes No.269/MenKes/Per/III/2008 tentang rekam medis, adalah berkas yang berisikan catatan dan dokumen tentang identitas pasien, pemeriksaan, pengobatan, tindakan dan pelayanan lain yang telah diberikan kepada pasien (Sanggale et al., 2018).

Rekam Medis

Rekam medis adalah dokumen yang memuat catatan tentang identitas, riwayat penyakit, hasil pemeriksaan, diagnosis, pengobatan, dan tindakan medis yang diberikan kepada pasien. Berdasarkan Permenkes RI No. 269/Menkes/Per/III/2008, rekam medis berfungsi sebagai bukti pelayanan medis, sumber informasi untuk pengambilan keputusan, dan dasar analisis data kesehatan. Data rekam medis dapat berupa data terstruktur (angka, kategori) dan tidak terstruktur (teks keluhan, catatan medis).

Machine Learning

Machine learning adalah cabang dari kecerdasan buatan (*artificial intelligence*) yang memungkinkan komputer mempelajari pola dari data untuk membuat prediksi atau klasifikasi. Dalam bidang kesehatan, *machine learning* banyak digunakan untuk prediksi diagnosis, analisis citra medis, serta penentuan tingkat keparahan penyakit. Pemanfaatan teknologi ini dapat membantu meningkatkan kecepatan dan akurasi pengambilan keputusan medis (Araf et al., 2024).

Klasifikasi

Klasifikasi merupakan salah satu metode dalam supervised learning yang bertujuan untuk mengelompokkan data ke dalam kelas-kelas tertentu berdasarkan pola yang dipelajari dari data latih. Dalam konteks penelitian ini, klasifikasi digunakan untuk membedakan tingkat keparahan penyakit menjadi tiga kelas: Ringan, Sedang, dan Berat. Tingkat keparahan penyakit menggambarkan sejauh mana suatu penyakit memengaruhi kondisi pasien dan kebutuhan akan penanganan medis. Umumnya, tingkat keparahan dibagi menjadi:

1. Ringan: Gejala minimal, tidak memerlukan perawatan intensif.
2. Sedang: Gejala moderat, memerlukan pengawasan medis.
3. Berat: Gejala parah yang membutuhkan penanganan segera dan intensif.

Machine Learning

Machine learning adalah cabang kecerdasan buatan yang memungkinkan komputer mempelajari pola dari data untuk membuat prediksi atau keputusan. Dalam bidang kesehatan, *machine learning* digunakan untuk diagnosis penyakit, prediksi komplikasi, dan klasifikasi tingkat keparahan. Keunggulan *machine learning* terletak pada kemampuannya mengolah data berukuran besar dengan variabel kompleks dan saling terkait.

Random Forest

Random Forest adalah metode *ensemble learning* berbasis *decision tree* yang menggabungkan prediksi dari banyak pohon keputusan untuk meningkatkan akurasi dan mengurangi *overfitting*. Keunggulan algoritma ini meliputi:

- a. Mampu menangani data numerik dan kategorikal.
- b. Robust terhadap *outlier* dan *noise*.
- c. Memiliki kemampuan mengukur *feature importance*.
Pada penelitian ini, *Random Forest* digunakan untuk mengklasifikasikan tingkat keparahan penyakit pasien menjadi tiga kelas (*Ringan, Sedang, Berat*)(Reza & Rohman, 2024).

Evaluasi Model

Evaluasi performa model klasifikasi dilakukan menggunakan beberapa metrik, di antaranya: Precision: Mengukur ketepatan prediksi positif dibandingkan dengan seluruh prediksi positif yang dibuat model. Recall (*Sensitivity*): Mengukur kemampuan model dalam menemukan seluruh data positif yang sebenarnya ada. F1-Score: Rata-rata harmonis antara precision dan recall, berguna ketika data tidak seimbang. Accuracy: Persentase prediksi benar dibandingkan dengan keseluruhan data uji. Macro Average: Rata-rata metrik yang dihitung secara setara untuk setiap kelas. Weighted Average: Rata-rata metrik yang dihitung dengan mempertimbangkan jumlah data pada setiap kelas.

Evaluasi model dilakukan menggunakan beberapa metrik, yaitu:

- a. Accuracy: proporsi prediksi benar terhadap seluruh data uji. Pada penelitian ini, akurasi tercatat sebesar 74,77%.
- b. Precision: ketepatan prediksi positif terhadap semua prediksi positif. Kelas *Berat* memiliki precision tertinggi (1,00).
- c. Recall: kemampuan model menangkap semua data positif sebenarnya. Kelas *Ringan* memiliki recall tertinggi (0,95), sementara kelas *Berat* rendah (0,12).
- d. F1-Score: rata-rata harmonis antara *precision* dan *recall*.

- e. Macro Average: rata-rata metrik tanpa mempertimbangkan jumlah data per kelas (precision 0,83; recall 0,60; f1-score 0,59).
- f. Weighted Average: rata-rata metrik dengan bobot sesuai jumlah data per kelas (precision 0,78; recall 0,75; f1-score 0,71).

Confusion Matrix

Confusion matrix adalah tabel yang menunjukkan distribusi prediksi model dibandingkan dengan label sebenarnya. Pada penelitian ini, kelas *Ringan* memiliki jumlah prediksi benar paling tinggi, yaitu 104 dari 109 data. Sementara itu, kelas *Berat* sering salah diklasifikasikan sebagai kelas *Ringan* atau *Sedang*, dan kelas *Sedang* juga menunjukkan kesalahan prediksi yang cukup sering, terutama dialihkan ke kelas *Ringan*. Analisis *confusion matrix* ini membantu mengidentifikasi kelemahan model, khususnya pada kemampuan mendeteksi kelas dengan tingkat keparahan yang lebih tinggi, sehingga dapat menjadi dasar untuk melakukan perbaikan dan optimasi model pada penelitian selanjutnya.

3. Metode Penelitian

Penelitian ini menggunakan desain eksperimen dengan pendekatan kuantitatif. Pendekatan kuantitatif dipilih karena fokus penelitian adalah pada pengolahan data numerik dari rekam medis pasien dan pengukuran kinerja model secara statistik. Desain eksperimen digunakan untuk menguji efektivitas algoritma Random Forest dalam mengklasifikasikan tingkat keparahan penyakit ke dalam tiga kategori, yaitu Ringan, Sedang, dan Berat.

Populasi penelitian adalah seluruh data rekam medis pasien yang tercatat pada sistem informasi rumah sakit mitra penelitian. Sampel penelitian diambil dengan teknik purposive sampling, yaitu pemilihan data berdasarkan kriteria kelengkapan variabel yang dibutuhkan, seperti hasil pemeriksaan laboratorium, tekanan darah, suhu tubuh, dan parameter klinis lainnya. Jumlah total sampel yang digunakan adalah 214 data pasien yang telah melalui proses seleksi.

Data penelitian diperoleh melalui teknik dokumentasi, yaitu pengambilan data yang sudah tersedia dalam arsip digital rumah sakit dengan izin resmi dan setelah dilakukan proses anonimisasi untuk menjaga kerahasiaan identitas pasien. Data yang diperoleh bersifat terstruktur dalam bentuk tabel digital yang memuat variabel numerik dan kategorikal. Instrumen penelitian berupa perangkat lunak berbasis Python dengan pustaka pendukung seperti pandas, scikit-learn, numpy, matplotlib, dan seaborn yang digunakan untuk melakukan pra-pemrosesan data, membangun model Random Forest, dan mengevaluasi kinerja model.

Tahapan penelitian meliputi pengumpulan data dari sumber resmi, pra-pemrosesan data yang mencakup pembersihan data (data cleaning), penanganan nilai kosong (*missing values*), transformasi variabel kategorikal menjadi numerik, dan normalisasi skala data numerik. Setelah itu, dataset dibagi menjadi 80% data latih (*training set*) dan 20% data uji (*testing set*). Model Random Forest kemudian dilatih pada data latih untuk mempelajari pola hubungan antara variabel masukan dan label tingkat keparahan penyakit. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan f1-score, serta analisis confusion matrix untuk mengidentifikasi pola kesalahan prediksi.

Analisis data dilakukan secara statistik untuk menilai performa model berdasarkan hasil evaluasi. Nilai akurasi digunakan untuk mengukur ketepatan keseluruhan prediksi,

sedangkan precision, recall, dan f1-score digunakan untuk menilai ketepatan, kelengkapan, dan keseimbangan prediksi pada masing-masing kelas. Analisis confusion matrix memberikan gambaran distribusi prediksi model terhadap data sebenarnya sehingga kelemahan model dapat diidentifikasi dan digunakan sebagai dasar rekomendasi perbaikan pada penelitian selanjutnya.

3. Hasil dan Pembahasan

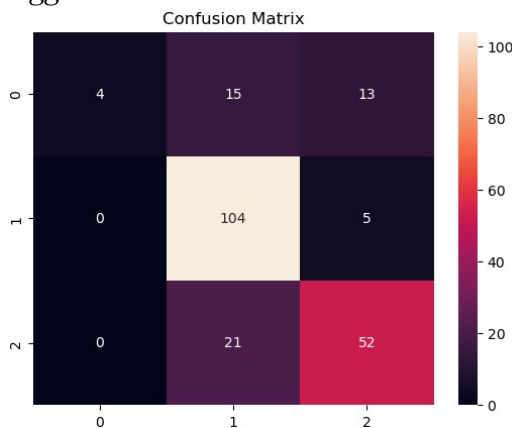
Hasil Penelitian

Penelitian ini menghasilkan model klasifikasi tingkat keparahan penyakit pasien menggunakan algoritma *Random Forest* dengan tiga kategori kelas, yaitu *Ringan*, *Sedang*, dan *Berat*. Model dibangun dari data rekam medis pasien yang telah melalui proses pra-pemrosesan dan dibagi menjadi 80% data latih serta 20% data uji. Evaluasi model dilakukan menggunakan metrik akurasi, *precision*, *recall*, *f1-score*, serta analisis *confusion matrix*.

	precision	recall	f1-score	support
Berat	1.00	0.12	0.22	32
Ringan	0.74	0.95	0.84	109
Sedang	0.74	0.71	0.73	73
accuracy			0.75	214
macro avg	0.83	0.60	0.59	214
weighted avg	0.78	0.75	0.71	214

Accuracy : 0.7476635514018691

Gambar 1. Hasil dari tingkat keparahan Pasien pengujian dengan menggunakan Metode Random Forest



Gambar 2. Hasil dari pengujian Tingkat Keparahannya dengan menggunakan Metode Random Forest

Berdasarkan hasil pengujian, model *Random Forest* memperoleh tingkat akurasi sebesar 74,77%. Nilai *precision* tertinggi diperoleh pada kelas *Berat* (1,00), menunjukkan bahwa seluruh prediksi kelas ini benar meskipun jumlah prediksinya terbatas. Nilai *recall* tertinggi dimiliki kelas *Ringan* (0,95), mengindikasikan bahwa hampir seluruh data *Ringan* berhasil dikenali dengan benar oleh model. Sementara itu, *recall* terendah terdapat pada

kelas *Berat* (0,12), menunjukkan bahwa sebagian besar kasus *Berat* tidak berhasil diidentifikasi dengan tepat. Analisis *confusion matrix* menunjukkan bahwa kelas *Ringan* memiliki jumlah prediksi benar tertinggi, yaitu 104 dari 109 data. Sebaliknya, kelas *Berat* sering salah diklasifikasikan sebagai *Ringan* atau *Sedang*, dan kelas *Sedang* cenderung diprediksi sebagai *Ringan*. Pola kesalahan ini mengindikasikan bahwa model memiliki kecenderungan bias terhadap kelas mayoritas.

Pembahasan

Hasil yang diperoleh menunjukkan bahwa algoritma *Random Forest* mampu memberikan performa prediksi yang cukup baik pada data rekam medis pasien, khususnya pada kelas mayoritas (*Ringan*). Hal ini sejalan dengan teori yang menyatakan bahwa *Random Forest* memiliki kemampuan tinggi dalam mengolah data berukuran besar dan mengurangi risiko *overfitting* melalui penggunaan beberapa *decision tree*. Namun, kinerja yang rendah pada kelas *Berat* mengindikasikan adanya masalah ketidakseimbangan data (*imbalanced data*), di mana jumlah sampel pada kelas minoritas jauh lebih sedikit dibandingkan kelas mayoritas. Temuan ini konsisten dengan penelitian sebelumnya di bidang klasifikasi medis, di mana model cenderung memiliki sensitivitas rendah terhadap kelas dengan jumlah sampel yang terbatas. Faktor lain yang memengaruhi hasil ini adalah kemiripan karakteristik variabel klinis antara kelas *Sedang* dan *Berat*, sehingga model kesulitan membedakan keduanya.

Meskipun demikian, hasil penelitian ini tetap menunjukkan potensi penerapan *Random Forest* sebagai sistem pendukung keputusan di bidang medis. Dengan akurasi mendekati 75% dan performa tinggi pada kelas *Ringan*, model ini dapat membantu tenaga medis dalam pengelompokan awal tingkat keparahan penyakit. Namun, agar sistem ini optimal, diperlukan langkah lanjutan seperti penerapan teknik *oversampling* (misalnya SMOTE) atau *class weight adjustment* untuk meningkatkan *recall* pada kelas *Berat*. Keterbatasan penelitian ini terletak pada ukuran sampel yang relatif kecil dan distribusi kelas yang tidak seimbang, yang berpengaruh pada kemampuan model dalam mendeteksi kasus pada kelas minoritas. Penelitian lanjutan disarankan untuk menggunakan dataset yang lebih besar dan seimbang, serta membandingkan *Random Forest* dengan algoritma lain seperti *XGBoost* atau *Support Vector Machine* untuk melihat perbedaan kinerja. Secara keseluruhan, penelitian ini membuktikan bahwa *Random Forest* dapat digunakan untuk klasifikasi tingkat keparahan penyakit berbasis data rekam medis dengan hasil yang cukup memuaskan, meskipun masih memerlukan optimasi lebih lanjut pada kelas minoritas agar penerapannya di dunia nyata lebih akurat dan andal.

5. Kesimpulan

Penelitian ini bertujuan untuk membangun model klasifikasi tingkat keparahan penyakit pasien berbasis data rekam medis menggunakan algoritma *Random Forest*. Model dikembangkan untuk membedakan tiga kategori tingkat keparahan, yaitu *Ringan*, *Sedang*, dan *Berat*, dengan harapan dapat membantu tenaga medis dalam pengambilan keputusan yang lebih cepat dan akurat. Hasil penelitian menunjukkan bahwa *Random Forest* mampu memberikan akurasi sebesar 74,77%, dengan performa terbaik pada kelas *Ringan* yang memiliki *recall* 0,95 dan *precision* 0,80. Meskipun kelas *Berat* mencapai *precision* sempurna (1,00), *recall*-nya rendah (0,12), yang mengindikasikan kesulitan model dalam mendeteksi sebagian besar kasus pada kelas ini. Temuan ini menjawab permasalahan penelitian bahwa

metode *machine learning*, khususnya *Random Forest*, dapat mengklasifikasikan tingkat keparahan penyakit dengan cukup baik, namun sensitifitas terhadap kelas minoritas masih menjadi tantangan. Kontribusi penelitian ini terletak pada penerapan *Random Forest* dalam klasifikasi tingkat keparahan penyakit berbasis rekam medis pasien di lingkungan rumah sakit, yang dapat menjadi dasar pengembangan sistem pendukung keputusan medis. Secara praktis, hasil ini berpotensi meningkatkan efisiensi layanan kesehatan melalui pengelompokan pasien berdasarkan tingkat keparahan secara otomatis. Namun, keterbatasan penelitian ini meliputi ukuran dataset yang relatif kecil dan distribusi kelas yang tidak seimbang, sehingga memengaruhi kemampuan model dalam mengenali kelas minoritas. Untuk penelitian lanjutan, disarankan penggunaan dataset yang lebih besar dan seimbang, penerapan teknik penyeimbangan data seperti SMOTE atau penyesuaian *class weight*, serta perbandingan kinerja *Random Forest* dengan algoritma lain seperti *XGBoost* atau *Support Vector Machine*.

6. Daftar Pustaka

- Araf, I., Idri, A., & Chairi, I. (2024). Cost-sensitive learning for imbalanced medical data: a review. *Artificial Intelligence Review*, 57(4). <https://doi.org/10.1007/s10462-023-10652-8>
- Reza, A. A. R., & Rohman, M. S. (2024). Prediction Stunting Analysis Using Random Forest Algorithm and Random Search Optimization. *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, 7(2), 534–544. <https://doi.org/10.31289/jite.v7i2.10628>
- Sanggamele, C., Kolibu, F. K., & Maramis, F. R. R. (2018). Analisis Pengelolaan Rekam Medis Di Rumah Sakit Umum Pancaran Kasih Manado. *Jurnal KESMAS*, 7(4), 1–11.
- Siregar, A. P., Purba, D. P., & Pasaribu, J. P. (2023). Implementasi Algoritma Random Forest Dalam Klasifikasi Diagnosis Penyakit Stroke. 2(4).
- Atika Sari, C. & Hari Rachmawanto, E. (2022). Sentiment Analyst on Twitter Using the K-Nearest Neighbors (KNN) Algorithm Against Covid-19 Vaccination. *Journal of Applied Intelligent System*, 7(2), 135–145. <https://doi.org/10.33633/jais.v7i2.6734>
- Chiu, C. C., Wu, C. M., Chien, T. N., Kao, L. J., Li, C. & Jiang, H. L. (2022). Applying an Improved Stacking Ensemble Model to Predict the Mortality of ICU Patients with Heart Failure. *Journal of Clinical Medicine*, 11(21). <https://doi.org/10.3390/jcm11216460>.
- Datta, D., Mallick, P. K., Reddy, A. V. N., Mohammed, M. A., Jaber, M. M., Alghawli, A. S. & Al-qaness, M. A. A. (2022). A Hybrid Classification of Imbalanced Hyperspectral Images Using ADASYN and Enhanced Deep Subsampled Multi-Grained Cascaded Forest. *Remote Sensing*, 14(19). <https://doi.org/10.3390/rs14194853>.
- Hairani, H., Anggrawan, A. & Priyanto, D. (2023). Improvement Performance of the Random Forest Method on Unbalanced Diabetes Data Classification Using Smote-Tomek Link. *International Journal on Informatics Visualization*, 1(7), 258–264. <https://doi.org/10.30630/joiv.7.1.1069>.
- Hamdaoui, H. El, Boujraf, S., Chaoui, N. E. H., Alami, B. & Maaroufi, M. (2021). Improving Heart Disease Prediction Using Random Forest and AdaBoost Algorithms. *International Journal of Online and Biomedical Engineering*, 17(11), 60–75. <https://doi.org/10.3991/ijoe.v17i11.24781>.
- Husain, G., Nasef, D., Jose, R., Mayer, J., Bekbolatova, M., Devine, T. & Toma, M. (2025). SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models. *Algorithms*, 18(1), 1–16. <https://doi.org/10.3390/a18010037>.

- Nugroho H, Y. D., Zakiyabarsi, F. & Paramita, A. J. (2025). Implementasi SMOTE-ENN dan Borderline SMOTE Terhadap Performa LightGBM Pada Imbalanced Class. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 10(1), 51–59. <https://doi.org/10.36341/rabit.v10i1.5436>.
- Sarra, R. R., Gorial, I. I., Manea, R. R., Korial, A. E., Mohammed, M. & Ahmed, Y. (2024). Enhanced Stacked Ensemble-Based Heart Disease Prediction with Chi-Square Feature Selection Method. *Journal of Robotics and Control (JRC)*, 5(6), 1753–1763. <https://doi.org/10.18196/jrc.v5i6.23191>.
- Sharma, H., Pangaonkar, S., Gunjan, R. & Rokade, P. (2023). Sentimental Analysis of Movie Reviews Using Machine Learning. *International Conference on Data Science and Intelligent Applications*, 53, 1–9. <https://doi.org/10.1051/itmconf/20235302006>