

Klasifikasi Penyakit Liver dengan NAIVE BAYES

Liver Disease Classification using the NAIVE BAYES

¹Vitra Nurhalisa, ²Ika Nur Fajri*

^{1,2}Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta
^{1,2}Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa
Yogyakarta 55281

*e-mail: vitraturhalisa@students.amikom.ac.id, fajri@amikom.ac.id

(received: 9 February 2025, revised: 30 May 2025, accepted: 1 June 2025)

Abstrak

Perkembangan teknologi kecerdasan buatan membuka peluang baru dalam mendukung tenaga medis untuk mengambil keputusan klinis yang lebih cepat dan tepat. Penelitian ini menghadirkan sistem klasifikasi penyakit liver berbasis Naive Bayes, yang dirancang agar mudah diinterpretasikan oleh dokter dan tenaga kesehatan. Kami menggunakan $n = 580$ pasien dengan 11 atribut klinis—mulai dari kadar bilirubin hingga rasio albumin–globulin—yang diproses melalui tahapan pembersihan data dan normalisasi. Model Bernoulli Naive Bayes kemudian dilatih dan dievaluasi menggunakan confusion matrix serta analisis ROC-AUC. Hasil menunjukkan akurasi 67 % dengan kemampuan mendeteksi pasien berisiko penyakit liver (recall 0,82) dan menandai kesehatan normal (recall 0,28). Kecepatan pelatihan dan transparansi probabilitas prediksi Naive Bayes menjadikannya solusi praktis untuk prototipe sistem pendukung keputusan medis. Rekomendasi ke depan mencakup penambahan fitur klinis relevan serta penerapan pendekatan ensemble untuk meningkatkan sensitivitas dan spesifisitas diagnosis.

Kata kunci: klasifikasi penyakit liver, naive bayes, pra-pemrosesan data, evaluasi model, sistem pendukung keputusan medis

Abstract

The advancement of artificial intelligence technology presents new opportunities to support medical professionals in making faster and more accurate clinical decisions. This study introduces a liver disease classification system based on the Naive Bayes algorithm, designed to be easily interpretable by doctors and healthcare personnel. A dataset of 580 patients with 11 clinical attributes—ranging from bilirubin levels to albumin–globulin ratio—was used and processed through data cleaning and normalization stages. The Bernoulli Naive Bayes model was then trained and evaluated using a confusion matrix and ROC-AUC analysis. The results show an accuracy of 67%, with strong performance in identifying patients at risk of liver disease (recall of 0.82), but weaker in classifying healthy individuals (recall of 0.28). The fast training time and transparent probabilistic predictions of the Naive Bayes algorithm make it a practical solution for developing a prototype of a medical decision support system. Future recommendations include incorporating additional relevant clinical features and applying ensemble methods to improve diagnostic sensitivity and specificity.

Keywords: liver disease classification, naive bayes, data preprocessing, model evaluation, clinical decision support system

1 Pendahuluan

Kesehatan liver merupakan salah satu hal rentan yang perlu dijaga untuk kehidupan manusia[1]. Penanganan yang cepat dan diagnosis yang tepat sangatlah penting untuk mengatasi masalah ini[2]. Di Indonesia, masalah ini menjadi perhatian besar, terutama akibat infeksi hepatitis B dan C[3]. Berdasarkan data dari Kementerian Kesehatan, prevalensi hepatitis B mengalami angka penurunan dari 7,1% yang kemudian pada tahun 2013 angka tersebut mengalami penurunan yang lebih jauh menjadi 2,4% pada tahun 2023[4]. Namun, angka kematian tetap tinggi setiap tahunnya, sekitar 51.100 orang meninggal akibat hepatitis

<http://sistemasi.ftik.unisi.ac.id>

B, dan 5.942 orang akibat hepatitis C. Dalam konteks ini, teknologi berbasis kecerdasan buatan, khususnya *machine learning*, dapat memberikan solusi yang menjanjikan untuk menawarkan solusi analisis pola data klinis dan mendukung keputusan medis[5]. Salah satu metode klasifikasi yang sederhana namun efektif adalah Naive Bayes[6]. Melalui pendekatan ini, sistem dapat mengenali pola dalam data pasien dan memprediksi kemungkinan adanya penyakit liver[7]. Sistem ini juga memungkinkan untuk menjadi bagian dari *decision support system* (DSS) yang membantu profesional medis dalam mempercepat dan mempermudah pengambilan keputusan dengan tingkat ketepatan yang lebih tinggi [8].

Tantangan penelitian ini adalah bagaimana performa Bernoulli Naïve Bayes pada dataset 580 pasien lokal, serta sejauh mana model ini dapat diinterpretasikan dan potensinya. Penelitian ini diarahkan untuk mengembangkan serta menguji efektivitas algoritma Bernoulli Naive Bayes dalam menghasilkan prediksi kesehatan yang akurat dan informatif. Penilaian terhadap hasil model bertujuan untuk memperoleh pemahaman yang lebih menyeluruh mengenai implementasi *machine learning* dalam membantu diagnosis penyakit dan pengelolaan sistem kesehatan. Penelitian ini menyajikan performa evaluasi dengan Naïve bayes dengan dataset lokal untuk pemodelan serta pengembangan kualitas input model didunia medis serta analisis performa, yang berdasarkan perbandingan studi terkini sebagai acuan serta penerapan sistem dss dengan antarmuka.

Tujuan dari penelitian ini adalah untuk mengembangkan dan mengevaluasi efektivitas Bernoulli Naive Bayes dalam memberikan prediksi kesehatan yang akurat dan relevan khususnya dalam penyakit liver. Dari hasil analisis model diharapkan mampu memberikan pemahaman tentang bagaimana *machine learning* dapat diterapkan dunia kedokteran dalam mendiagnosis.

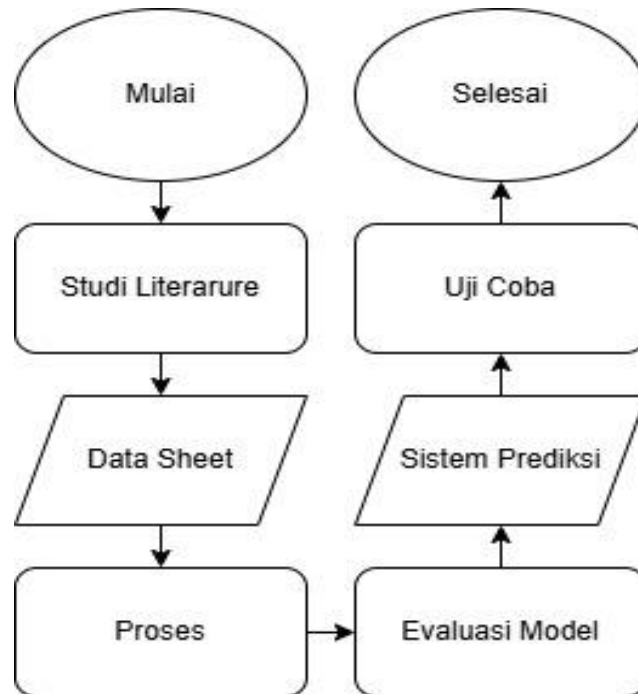
2 Tinjauan Literatur

Machine learning memiliki peran sentral dalam meningkatkan akurasi dan efisiensi dalam diagnosis klinis berdasarkan rekam medis elektronik, yang seringkali menghasilkan analisis menyamai tenaga medis manusia[9][10]. Implementasi nyata aplikasi Machine learning dalam praktik klinis termasuk faktor kunci kesuksesan seperti integrasi antarmuka pengguna dan pelatihan staf yang menunjukkan kesiapan teknologi ini untuk diadopsi skala luas dalam dunia klinis[11]. Beberapa penelitian mengenai penyakit hati mengindikasikan bahwa algoritma Naive Bayes (NB) memiliki performa yang kompetitif jika dibandingkan dengan algoritma lain. Sebuah studi yang memanfaatkan *Indian Liver Patient Dataset* (ILPD) menunjukkan bahwa NB mampu mencapai akurasi sebesar 82,65%, hanya sedikit lebih rendah dari algoritma Random Forest yang mencatat akurasi 87,48% [12]. Dalam penelitian yang lebih terkini, penerapan NB pada data pasien sirosis hati menunjukkan hasil akurasi mendekati 80%, yang menegaskan kapabilitas NB meskipun menghadapi data dengan tingkat variabilitas klinis yang tinggi [7]. Meskipun algoritma seperti Random Forest dan Support Vector Machine (SVM) umumnya unggul dalam hal akurasi, NB tetap menjadi alternatif menarik karena proses pelatihannya yang cepat dan sifatnya yang mudah dipahami [13]. Dalam studi komparatif pada dataset penyakit liver, *Logistic Regression* unggul dengan akurasi ~85%, namun NB tetap dipandang sebagai pilihan praktis ketika sumber daya komputasi terbatas dan hasil perlu dijelaskan secara transparan kepada tenaga medis[14]. Demikian pula, model NB yang dirancang untuk data cirrhosis menunjukkan performa mendekati K-Nearest Neighbors, dengan keunggulan signifikan pada waktu inferens[15]. Keandalan dan kesederhanaan Naive Bayes, berpadu dengan kemampuannya menyediakan probabilitas prediktif yang mudah ditafsirkan, menjadikannya pondasi yang ideal untuk sistem pendukung keputusan medis (DSS). Dengan demikian, meski algoritma ensemble atau kernel-based lain mungkin sedikit lebih akurat, NB memberikan kombinasi optimal antara performa, interpretabilitas, dan efisiensi komputasi.

3 Metode Penelitian

Metode yang dijadikan sebagai acuan pendekatan penelitian yang digunakan dalam studi ini bersifat kuantitatif dan dianalisis secara numerik. Sehingga pada tahap awal dilakukan studi

literatur dengan permasalahan dan tujuan penelitian ini. Kemudian adalah pemilihan dataset, yang digunakan adalah data dari catatan 580 pasien yang memiliki 11 atribut penyakit, dalam penyiapan dataset nilai yang hilang diganti dengan nilai mean per atribut dengan rentang 0-1[16]. Pada tahap pemrosesan, pembagian data validasi dibagi menjadi 70% training data dan 30% untuk test dengan validasi silang 5-fold untuk tuning[17]. Selanjutnya dilakukan analisis model dengan menguji performa. Di Tahap selanjutnya adalah pembuatan sistem prediksi dengan dasar data model yang sebelumnya kemudian dilakukan Uji Coba sistem prediksi. Alur penelitian pada Gambar 1 menunjukkan metode yang digunakan.



Gambar 1. Alur penelitian

Dataset yang dipilih akan diproses untuk pembuatan model menggunakan Python dengan *library scikit-learn*, *pandas*, dan *pickle* [18]. Pada tahap pembuatan sistem menggunakan Python dengan *library flask* untuk menciptakan antarmuka[19]. Sedangkan acuan pembuatan sistem menggunakan rumus dasar Naive Bayes berdasarkan prinsip Teorema Bayes, yang dirumuskan pada Persamaan (1) yang terlampir pada perumusan berikut:

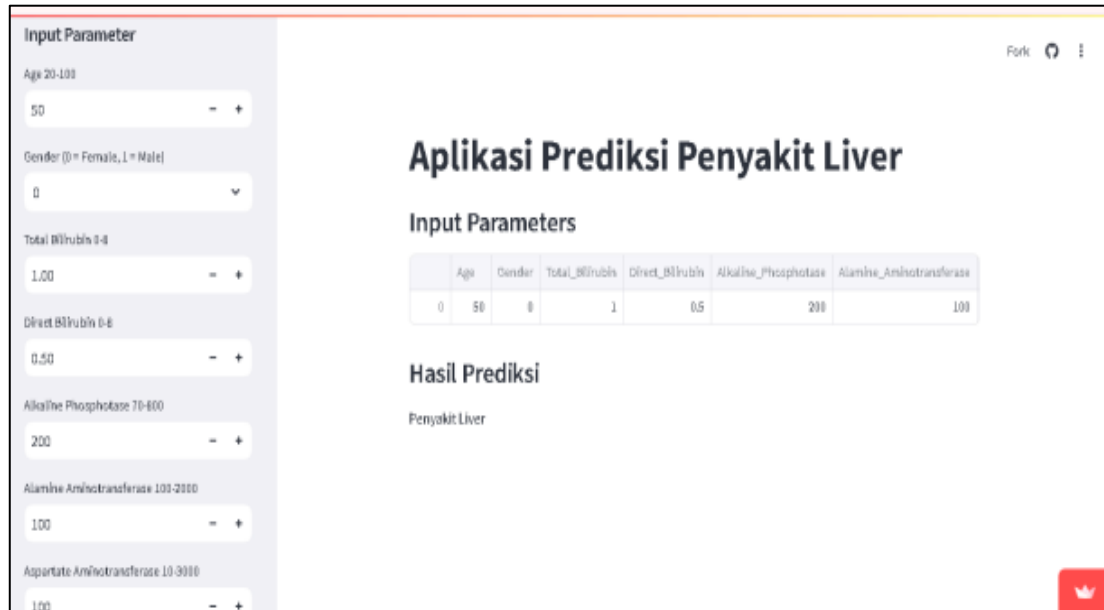
$$\frac{P(E) = P(H).P(H)}{P(E)} \quad (1)$$

Persamaan (1) mendefinisikan $P(H|E)$ sebagai probabilitas posterior hipotesis H ketika bukti E ada. $P(E|H)$ adalah probabilitas bukti E jika hipotesis H benar. $P(H)$ adalah probabilitas awal hipotesis, sedangkan $P(E)$ adalah probabilitas bukti secara keseluruhan. Rumus tersebut menjadi dasar pembuatan program yang mempermudah analisis gejala guna penerapan naive bayes dalam mendeteksi penyakit hati. Evaluasi dilakukan dengan *confusion matrix* dan metrik *precision*, *recall*, F1-score, serta ROC-AUC [20].

4 Hasil dan Pembahasan

Hasil studi literatur memberikan pondasi kuat dalam pembangunan sistem ini untuk mencapai tujuan penelitian. Setelah pemilihan dataset kemudian diproses dengan *scikit-learn*, *pandas* dan *pickle* untuk menghasilkan model yang akan digunakan dalam penelitian ini. Tahapan penelitian dilakukan secara sistematis untuk mendapatkan model, sehingga dapat digunakan dalam sistem prediksi penyakit liver. Hasil model yang diperoleh kemudian dilakukan evaluasi yang bertujuan untuk

mengetahui kesesuaian model sehingga relevan. Setelah tahap ini, dilanjutkan pada tahap pembangunan sistem menggunakan Tools Stream Cloud dan github dengan menggunakan library flask untuk menghasilkan antarmuka yang akan siap digunakan untuk diagnosa, sehingga dapat digunakan lebih mudah. Desain antarmuka dari sistem prediksi yang dibuat diperlihatkan pada Gambar 2.



Gambar 2. Tampilan sistem prediksi

Setelah melalui tahapan pengembangan sistem dan antarmuka, sistem melalui tahapan uji coba yang dimana sistem akan diuji untuk memastikan kinerjanya dalam mengklasifikasikan penyakit liver. Hasil dari pengujian ini disajikan dalam Tabel 1.

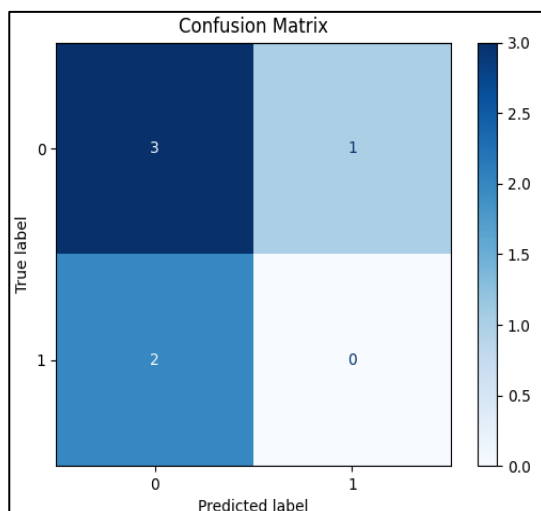
Tabel 1. Uji coba sistem prediksi

No	Age	Gender	Tot. Bilirubin	Dir. Bilirubin	Albumin			Total Proteins	Albumin D	Hasil	
					A	B	C				
1	63	female	2.89	2.27	354	322	878	4.12	2.32	0.34	Sakit
2	57	female	6.00	4.50	341	145	2254	8.00	2.53	2.52	Sakit
3	49	female	7.75	6.99	354	1055	2269	09.08	2.10	0.27	Tidak Sakit
4	20	female	1.00	4.50	78	100	10	7.00	6.00	2.00	Sakit
5	63	male	7.63	1.42	282	1311	502	6.84	5.52	0.88	Sakit
6	67	male	7.47	3.79	163	1587	454	04.08	5.11	2.29	Sakit
7	44	female	3.10	0.60	196	1382	1819	8.17	3.36	1.44	Sakit
8	57	male	5.33	02.02	682	1745	2104	9.16	02.06	2.80	Tidak Sakit
9	17	female	0.91	4.76	716	641	2074	2.98	05.09	2.49	Sakit
10	69	female	7.77	3.88	350	1385	2238	6.99	3.52	1.81	Tidak Sakit
11	44	female	1.33	1.52	107	530	2939	7.25	4.61	1.59	Sakit
12	51	female	4.51	6.31	793	1800	2367	2.72	2.46	2.31	Tidak Sakit
13	55	female	4.62	2.46	118	520	2439	7.31	5.79	0.18	Sakit

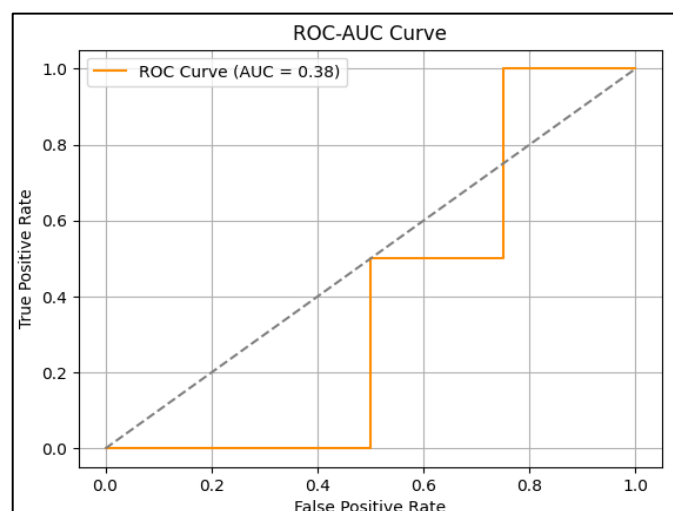
14	22	male	3.91	2.20	362	180	160	9.23	4.97	2.16	Sakit
15	44	male	2.33	6.53	86	610	323	08.08	2.75	0.88	Tidak Sakit
16	43	male	0.74	5.49	330	871	1411	2.62	3.70	2.38	Tidak Sakit
17	57	female	0.64	1.26	165	1383	2249	4.82	02.01	2.22	Sakit
18	55	male	4.76	5.11	169	708	281	6.50	4.15	2.60	Tidak Sakit
19	39	male	0.33	2.61	679	1248	36	8.54	2.57	0.07	Sakit
20	30	male	1.29	1.60	522	1817	261	9.97	4.99	0.91	Tidak Sakit

Tabel 2. Analisis data percobaan

Metrik	Penyakit	Sehat	Overall accuracy
Accuracy	—	—	0,67
Precision	0,66	0,48	—
Recall	0,82	0,28	—
F1-score	0,73	0,36	—



Gambar 3. Analisis confusion matrix



Gambar 4. Analisis ROC-AUC

Parameter A mengacu pada Alkaline Phosphatase , parameter B mengacu pada Alamine Aminotransferase, parameter C mengacu pada Aspartate Aminotransferase, dan parameter D mengacu pada Albumin and Globulin Ratio. Hasil pengujian menunjukkan bahwa model Naive Bayes efektif dalam memprediksi gejala penyakit liver. Evaluasi lebih lanjut memperlihatkan bahwa algoritma ini dapat diaplikasikan untuk diagnosis awal penyakit liver, meskipun performanya pada kelompok pasien yang sudah menderita penyakit masih perlu ditingkatkan. Analisis mendalam terhadap parameter medis juga menunjukkan bahwa fitur seperti total bilirubin dan kadar enzim hati memiliki kontribusi signifikan dalam prediksi. Analisis data percobaan pada Tabel 2 menunjukkan tingginya recall untuk penyakit (0,82) sebagai model sensitive dalam mendeteksi kasus liver, meski banyak false positives (0,66) untuk precision dan mengklasifikasikan pasien sehat sebagai sakit. *Overall accuracy* menunjukkan 67 yang mengartikan bahwa model ini memiliki keakuratan keseluruhan yang mencapai 67%. Hal ini kemungkinan didasari akibat variabilitas lokal dalam fitur klinis yang berbeda-beda. Implikasi klinis mendapatkan nilai $< 0,01$ s yang memungkinkan untuk mendukung akses real-time dan memudahkan tenaga medis, hasil Confusion Matrix dapat dilihat pada Gambar 3 dan analisis ROC-AUC pada Gambar 4.

5 Kesimpulan

Analisis menggunakan algoritma Naive Bayes pada dataset penyakit liver yang terdiri dari 580 data dengan 11 atribut dan 1 label kelas menghasilkan akurasi model sebesar 67%. Ini menunjukkan bahwa model memiliki tingkat keberhasilan moderat dalam memprediksi ada atau tidaknya penyakit liver pada pasien. Dari sisi precision, model lebih akurat dalam memprediksi pasien yang tidak memiliki penyakit liver (0.66) dibandingkan dengan pasien yang memiliki penyakit liver (0.48). Dari sisi recall, model lebih baik dalam mengidentifikasi pasien yang sebenarnya tidak memiliki penyakit liver (0.82) namun kurang baik dalam mendeteksi pasien yang sebenarnya memiliki penyakit liver (0.28). Hal ini berarti model cenderung lebih sering salah dalam mendeteksi pasien dengan penyakit liver, meskipun cukup baik dalam mengidentifikasi pasien yang sehat. Secara keseluruhan, model Naive Bayes ini dapat menjadi dasar awal yang cukup baik untuk sistem pendukung keputusan, terutama dalam mengidentifikasi pasien yang tidak menderita penyakit liver. Namun, perlu diingat bahwa ada keterbatasan signifikan dalam kemampuannya untuk mendeteksi pasien yang benar-benar menderita penyakit liver. Meskipun demikian, terdapat keterbatasan signifikan dalam kemampuannya untuk secara akurat mendeteksi pasien yang benar-benar menderita penyakit liver, yang ditunjukkan oleh rendahnya recall untuk kelas tersebut. Oleh karena itu, evaluasi lebih lanjut dan peningkatan model sangat diperlukan untuk meningkatkan kinerja prediksi secara keseluruhan, khususnya untuk kelas dengan penyakit liver.

Berdasarkan temuan dan keterbatasan tersebut, Disarankan untuk memperbaiki kualitas data input, mencoba model yang lebih kompleks seperti Random Forest atau SVM, dan menambahkan fitur umpan balik dokter pada sistem pendukung keputusan (DSS). Dengan perbaikan ini, DSS diharapkan lebih akurat dan bermanfaat dalam diagnosis penyakit liver.

Referensi

- [1] H. Devarbhavi, S. K. Asrani, J. P. Arab, Y. A. Nartey, E. Pose, and P. S. Kamath, “*Global Burden of Liver Disease: 2023 Update*,” Aug. 01, 2023, *Elsevier B.V.* doi: 10.1016/j.jhep.2023.03.017.
- [2] C. Gan *et al.*, “*Liver Diseases: Epidemiology, Causes, Trends and Predictions*,” Dec. 01, 2025. doi: 10.1038/s41392-024-02072-z.
- [3] A. Trickey *et al.*, “*Hepatitis C Virus Elimination in Indonesia: Epidemiological, Cost, and Cost-Effectiveness Modelling to Advance Advocacy and Strategic Planning*,” *Liver International*, Vol. 40, May 2019, doi: 10.1111/liv.14232.
- [4] kemenkes, “indo liver.”
- [5] B. S. Sheena *et al.*, “*Global, Regional, and National Burden of Hepatitis B, 1990–2019: A Systematic Analysis for the Global Burden of Disease Study 2019*,” *Lancet Gastroenterol Hepatol*, Vol. 7, No. 9, pp. 796–829, Sep. 2022, doi: 10.1016/S2468-1253(22)00124-8.
- [6] A. Aravind, A. G. Bahirvani, R. Quiambao, and T. Gonzalo, “*Machine Learning Technology for Evaluation of Liver Fibrosis, Inflammation Activity and Steatosis (LIVERFAST™)*,” *Journal of Intelligent Learning Systems and Applications*, Vol. 12, No. 02, pp. 31–49, 2020, doi: 10.4236/jilsa.2020.122003.
- [7] H. Hartatik, M. Tamam, and A. Setyanto, “*Prediction for Diagnosing Liver Disease in Patients using KNN and Naïve Bayes Algorithms*,” May 2020, pp. 1–5. doi: 10.1109/ICORIS50180.2020.9320797.
- [8] M. Elhaddad and S. Hamam, “*AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential*,” *Cureus*, Apr. 2024, doi: 10.7759/cureus.57728.
- [9] K. Kolasa, B. Admassu, M. Hołownia-Voloskova, K. J. Kędzior, J. E. Poirrier, and S. Perni, “*Systematic Reviews of Machine Learning in Healthcare: A Literature Review*,” 2024, *Taylor and Francis Ltd.* doi: 10.1080/14737167.2023.2279107.
- [10] K. H. Lee *et al.*, “*Machine Learning-based Clinical Decision Support System for Treatment Recommendation and Overall Survival Prediction of Hepatocellular Carcinoma: A Multi-Center Study*,” *NPJ Digit Med*, Vol. 7, No. 1, Jan. 2024, doi: 10.1038/s41746-023-00976-8.

- [11] L. M. Preti, V. Ardito, A. Compagni, F. Petracca, and G. Cappellaro, “Implementation of Machine Learning Applications in Health Care Organizations: Systematic Review of Empirical Studies (Preprint),” Dec. 30, 2023. doi: 10.2196/preprints.55897.
- [12] S. R. Velu, V. Ravi, and K. Tabianan, “Data mining In Predicting Liver Patients using Classification Model,” *Health Technol (Berl)*, Vol. 12, No. 6, pp. 1211–1235, Nov. 2022, doi: 10.1007/s12553-022-00713-3.
- [13] N. A. Aziz, A. Manzoor, M. D. M. Qureshi, M. A. Qureshi, and W. Rashwan, “Unveiling Explainable AI in Healthcare: Current Trends, Challenges, and Future Directions,” Aug. 10, 2024. doi: 10.1101/2024.08.10.24311735.
- [14] F. Rahman, D. Das, A. Sami, P. Podder, and D. L. Michael, “Liver Cirrhosis Prediction using Logistic Regression, Naïve Bayes and Knn,” *International Journal of Science and Research Archive*, Vol. 12, No. 1, pp. 2411–2420, Jun. 2024, doi: 10.30574/ijrsra.2024.12.1.1030.
- [15] Nithyashri, H. Goel, and M. Hada, “Intelligent Classification of Liver Diseases using Ensemble Machine Learning Techniques,” May 2024, pp. 1183–1188. doi: 10.1109/ICoICI62503.2024.10696789.
- [16] Y. Ramakrishnaiah, N. Macesic, G. I. Webb, A. Y. Peleg, and S. Tyagi, “EHR-ML: A Data-Driven Framework for Designing Machine Learning Applications with Electronic Health Records,” *Int J Med Inform*, Vol. 196, Apr. 2025, doi: 10.1016/j.ijmedinf.2025.105816.
- [17] J. Kruta et al., “Machine Learning for Precision Diagnostics of Autoimmunity,” *Sci Rep*, Vol. 14, No. 1, Dec. 2024, doi: 10.1038/s41598-024-76093-7.
- [18] A. Nash, “Ndata A Framework for Automated Data Preprocessing using Machine Learning Techniques.” [Online]. Available: <https://www.researchgate.net/publication/384812869>
- [19] E. C. de Andrade et al., “Towards Machine Learning Algorithms in Predicting the Clinical Evolution of Patients Diagnosed with COVID-19,” *Applied Sciences (Switzerland)*, Vol. 12, No. 18, Sep. 2022, doi: 10.3390/app12188939.
- [20] S. M. Ganie, P. K. D. Pramanik, and Z. Zhao, “Improved Liver Disease Prediction from Clinical Data Through an Evaluation of Ensemble Learning Approaches,” *BMC Med Inform Decis Mak*, Vol. 24, No. 1, Dec. 2024, doi: 10.1186/s12911-024-02550-y.