



Literature Review: Klasifikasi Penyakit Demam Berdarah Memakai Algoritma Random Forest

Dimas Kurnia Putra^{1*}, Fakhri Kamaluddin², Fasqila Sakti³, Novy Rahdiansyah⁴, Perani Rosyani⁵

^{1,2,3,4,5}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email : ^{1*}dimasmas610@gmail.com, ²fakhrikamaluddin@gmail.com, ³fasqilasakti84@gmail.com,
⁴novyrahdiansyah@gmail.com, ⁵dosen00837@unpam.ac.id

(* : coressponding author)

Abstrak - Demam Berdarah Dengue(DBD) ialah penyakit meluas yang diakibatkan oleh virus demam berdarah lewat gigitan nyamuk *Aedes aegypti*. Penyakit ini masih jadi permasalahan sungguh- sungguh untuk kesehatan di bermacam negeri tropis, tercantum Indonesia, sebab mempunyai tingkatan penyebaran yang besar serta bisa berdampak parah. Klasifikasi penyakit DBD memakai tata cara yang efisien sangat diperlukan buat penaksiran yang kilat serta pas. Algoritma Random Forest merupakan salah satu tata cara pendidikan mesin yang kerap digunakan dalam klasifikasi penyakit sebab keakuratannya yang besar serta keahlian buat menanggulangi informasi dengan variabel yang lingkungan. Postingan ini bertujuan buat mereview sebagian riset yang memakai algoritma Random Forest dalam klasifikasi penyakit DBD, dan menganalisis kelebihan, kekurangan, serta akurasi tata cara ini bersumber pada literatur yang terdapat.

Kata Kunci: Demam Berdarah Dengue, Klasifikasi Penyakit, Random Forest, Pendidikan Mesin, Informasi Mining.

Abstract - *Dengue Hemorrhagic Fever(DHF) is an infectious disease caused by the dengue fever virus through the bite of the Aedes aegypti mosquito. This disease is still a serious health problem in various tropical countries, including Indonesia, because it has a high rate of spread and can be parah. Classification of dengue fever using effective methods is needed for fast and correct penaksiran. The Random Forest algorithm is a machine learning method that is often used in disease classification because of its high accuracy and ability to handle informasi with complex variables. This article aims to review several studies that use the Random Forest algorithm in the classification of dengue fever, as well as analyzing the advantages, disadvantages and accuracy of this method based on existing literature.*

Keywords: *Dengue Fever, Disease Classification, Random Forest, Machine Learning, Information Mining.*

1. PENDAHULUAN

Penyakit Demam Berdarah Dengue(DBD) ialah salah satu penyakit yang diakibatkan oleh peradangan virus demam berdarah yang menyebar lewat gigitan nyamuk *Aedes aegypti*. Penyakit ini mempunyai indikasi yang mirip dengan flu pada biasanya, semacam demam besar, perih sendi, serta ruam, sehingga kerap susah dibedakan dari penyakit lain tanpa pengecekan lebih lanjut. Oleh sebab itu, pengembangan model klasifikasi buat penaksiran DBD sangat berarti buat tingkatkan akurasi serta efisiensi penaksiran.

Pemakaian machine learning(ML) buat klasifikasi penyakit sudah jadi tren dalam sebagian tahun terakhir. Di antara bermacam algoritma ML, Random Forest kerap diseleksi sebab kemampuannya dalam menanggulangi informasi dengan bermacam tipe fitur, membagikan hasil klasifikasi yang lebih normal, serta kurangi resiko overfitting. Oleh sebab itu, postingan ini hendak mangulas sebagian riset yang memakai algoritma Random Forest buat klasifikasi penyakit DBD.

2. METODE

Riset ini dicoba memakai pendekatan Systematic Literature Review(SLR) buat mengenali, menyebarkan, serta menginterpretasikan riset yang relevan menimpa klasifikasi penyakit Demam Berdarah(DBD) memakai algoritma Random Forest, Certainty Factor, serta sistem ahli. SLR diseleksi buat membenarkan keterandalan hasil yang diperoleh, dengan menjajaki panduan yang sistematis dalam memilah serta menyebarkan literatur terpaut.



Langkah- langkah metodologi riset merupakan selaku berikut:

1) Perumusan Research Question(RQ)

Riset ini dimulai dengan formulasi sebagian Research Questions(RQ) yang bertujuan buat membagikan arah dalam proses pengumpulan serta analisis informasi. Sebagian persoalan riset yang hendak dijawab lewat riset ini meliputi:

- RQ1: Gimana algoritma Random Forest diterapkan dalam klasifikasi penyakit DBD, serta seberapa akurat hasil yang diperoleh?
- RQ2: Gimana Kepastian Aspek diterapkan dalam tingkatkan kepastian penaksiran dalam sistem ahli buat klasifikasi DBD?
- RQ3: Apa saja kelebihan serta kekurangan Random Forest serta Certainty Factor dalam membangun sistem ahli buat klasifikasi DBD?
- RQ4: Faktor- faktor apa yang sangat signifikan dalam memastikan penaksiran penyakit DBD bagi model Random Forest serta Certainty Factor?

2) Pengumpulan Data

Literatur yang relevan dikumpulkan dari bermacam sumber semacam database harian ilmiah (*IEEE Xplore, ScienceDirect, PubMed, Springer, dll.*), konferensi, dan prosiding. Kata kunci yang digunakan dalam pencarian literatur meliputi: "Demam Berdarah Dengue", "Klasifikasi Penyakit", "Random Forest", "Pendidikan Mesin", "Informasi Mining".

3) Kriteria Inklusi serta Eksklusi

Kriteria inklusi serta eksklusi digunakan buat membenarkan literatur yang diambil relevan dengan topik riset. Kriteria inklusi:

- Artikel yang mangulas algoritma Random Forest ataupun Certainty Factor dalam klasifikasi DBD.
- Artikel yang memakai sistem ahli dalam klasifikasi ataupun penaksiran DBD.
- Artikel yang ditulis dalam bahasa Inggris ataupun Indonesia.
- Kriteria eksklusi:
- Artikel yang tidak mencakup tata cara Random Forest ataupun Certainty Factor secara khusus.
- Artikel yang tidak membagikan hasil penilaian ataupun analisis yang terukur.

4) Ekstraksi serta Analisis Data

Informasi dari riset yang terpilih diekstraksi serta dianalisis buat menanggapi Persoalan Riset. Informasi yang diekstrak meliputi:

- Metode yang digunakan dalam klasifikasi penyakit.
- Model akurasi serta penilaian kinerja yang dicapai oleh Random Forest serta Certainty Factor.
- Interpretasi hasil serta faktor- faktor yang berfungsi dalam penaksiran penyakit.

5) Sistem Ahli serta Kepastian Faktor

Sistem ahli digunakan selaku pendekatan buat mengimplementasikan pengetahuan yang diperoleh dari literatur ke dalam sistem yang sanggup membagikan keputusan penaksiran. Kepastian Aspek diintegrasikan ke dalam sistem ahli buat tingkatkan akurasi serta kepastian dalam penaksiran DBD. Kepastian Aspek menolong sistem dalam membagikan probabilitas kepastian sesuatu penaksiran bersumber pada indikasi yang timbul.



6) **Analisis serta Ulasan Hasil**

Sehabis informasi dianalisis, hasil riset dipaparkan buat mangulas kelebihan, kekurangan, serta kedudukan tiap- tiap tata cara (Random Forest serta Kepastian Aspek) dalam sistem ahli penaksiran DBD. Hasil ini pula membagikan pengetahuan tentang faktor- faktor yang sangat pengaruhi penaksiran DBD.

7) **Search Process**

Dengan pendekatan pencarian yang sistematis ini, diharapkan bisa terkumpul literatur yang secara merata mencakup pelaksanaan algoritma Random Forest dalam klasifikasi penyakit DBD. Hasil literatur ini hendak dianalisis buat memperoleh pengetahuan tentang daya guna, kelebihan, kekurangan, serta hasil klasifikasi DBD memakai Random Forest, dan membagikan donasi untuk pengembangan model penaksiran penyakit berbasis pendidikan mesin yang lebih efektif serta akurat.

3. ANALISIS DAN PEMBAHASAN

Ulasan menimpa hasil riset dari 5 harian yang dianalisis terpaut pemakaian algoritma Random Forest buat klasifikasi penyakit DBD. Tabel berikut merangkum data terpaut tata cara, tujuan, serta hasil yang diperoleh dari tiap- tiap riset.

NO	NAMA PENELITI DAN TAHUN	METODE YANG DIBAHAS	TUJUAN PENELITIANNYA	HASIL YANG DIDAPAT
1	Setyawan, A. F., Ariyanto, A. D. P., & Fikriah, F. K. (2024)	Random Forest, CRISP-DM (Cross-Industry Standard Process for Data Mining)	Mengembangkan model klasifikasi untuk mendiagnosis demam berdarah dengue (DHF) dengan menggunakan algoritma Random Forest.	Model Random Forest memiliki akurasi 88.5%, presisi 88.2%, recall 65.2%, F1-score 74.9%, dan ROC AUC 0.810. Analisis fitur menunjukkan Gender_Male dan Body_Pain sebagai faktor yang paling signifikan.
2	Nisa, L. N., Muflikhah, L., Yudistira, N. (2024)	Algoritma Random Forest, SMOTE, PCA	Mendeteksi Dengue Shock Syndrome (DSS) pada pasien DBD menggunakan algoritma Random Forest dan menangani dataset yang imbalanced dengan SMOTE	Akurasi tertinggi mencapai 0,911 menggunakan PCA dan random oversampling
3	Rahmat Azul Mizan, Prima Widayani, Nur Mohammad Farda (2020)	Support Vector Machine (SVM) dan Random Forest (RF)	Memetakan dan mendeskripsikan distribusi kerentanan penyakit demam berdarah di Kota Baubau	Kecamatan Batupuaro dan Murhum didominasi zona rentan sebesar 92,54% dan 41,74% dari luas total wilayah masing-masing.



NO	NAMA PENELITI DAN TAHUN	METODE YANG DIBAHAS	TUJUAN PENELITIANNYA	HASIL YANG DIDAPAT
4	Lisa N. Nisa ¹ , Lailil Muflikhah, Novanto Yudistira. (2024)	Dalam penelitian ini, model random forest telah berhasil dalam memprediksi dengue syock syndrome pada pasien demam berdarah.	hyperparameter dapat mempengaruhi evaluasi model random forest.Extraction fitur memiliki pengaruh terhadap performa random forest, information gain memiliki performa lebih baik daripada PCA pada setiap jumlah fitur Data imbalance yang telah dilakukan oversampling dapat meningkatkan performa evaluasi random forest dibandingkan data imbalance tanpa oversampling.	Oversample Rata-rata Accuracy Oversampling 0,911342 SMOTE0,850228 Tanpa oversampling 0,808
5	Ryan Erlangga Ardiansyah.(2024)	Uji fungsionalitas sistem dilakukan dengan metode blackboxtestingatau sering disebut behavior testing.	Metode Modelling Random Forestdipilih untuk menganalisis hasil diagnosis karena menawarkan akurasi yang lebih tinggi dibandingkan metode lain yaitu Logistic Regression danDecision Tree. Sistem ini diimplementasikan menggunakan pemrograman berbasis web, sehingga bisadiakses dan digunakan oleh pengguna atau pasien.	hasil akurasi klasifikasi penyakit akibat gigitan nyamuk menggunakan modelling Random Forestdengan menggunakan tools google colabyaitu 90,00%.

4. KESIMPULAN

Kesimpulan dari kajian literatur mengenai klasifikasi penyakit Demam Berdarah Dengue (DBD) menggunakan algoritma Random Forest menunjukkan bahwa algoritma ini memiliki tingkat



akurasi yang tinggi dan keandalan dalam menangani data yang kompleks, menjadikannya pilihan yang efektif dalam diagnosis penyakit DBD. Beberapa studi menunjukkan bahwa kombinasi Random Forest dengan metode penanganan data seperti SMOTE dan PCA dapat meningkatkan performa, terutama dalam mengatasi masalah data imbalance yang sering terjadi pada data medis. Selain itu, metode ini mampu mengidentifikasi variabel signifikan yang berperan dalam diagnosis, seperti gejala nyeri tubuh dan jenis kelamin. Di sisi lain, terdapat tantangan dalam penerapan model ini, seperti perlunya penyesuaian hyperparameter untuk meningkatkan akurasi dan keakuratan sistem. Hasil penelitian ini diharapkan dapat mendukung pengembangan sistem pakar dan aplikasi berbasis web untuk diagnosis dini DBD yang lebih efisien, terutama di wilayah tropis yang rentan terhadap penyebaran penyakit ini.

REFERENCE

- [1] A. F. Setyawan, A. D. P. Ariyanto, and F. K. Fikriah, "Classification of Dengue Fever Disease Using a Machine Learning- Based Random Forest Algorithm," *Journal of Information Systems and Technology*, vol. 5, nomor. 2, pp. 87- 94, 2024.
- [2] L. N. Nisa, L. Muflikhah, and N. Yulistira, "Deteksi Dengue Shock Syndrome pada Penderita Demam Berdarah Memakai Algoritma Random Forest," *Kohesi: Harian Multidisiplin Saintek*, vol. 4, nomor. 7, pp. 1- 10, Sep. 2024.
- [3] R. A. Mizan, P. Widayani, and N. Meter. Farda, "Pemetaan Distribusi Kerentanan Penyakit Demam Berdarah di Kota Baubau Memakai Algoritma Machine Learning," *JAGAT*, vol. 4, nomor. 2, pp. 46- 57, 2020.
- [4] Lisa N. Nisa, Lailil Muflikhah2, Novanto Yulistira, "Deteksi Dengue Shock Syndrome Pada Pasien Demam Berdarah Menggunakan Algoritma Random Forest", *Kohesi: Jurnal Volume 4 No 7 Tahun 2024*, E-ISSN : 2988-1986
- [5] Ryan Erlangga Ardiansyah, "sistem pakar deteksi dini penyakit akibat gigitan nyamuk menggunakan machine learning", *Kohesi: Jurnal Multidisiplin 3 No 10 Tahun 2024*, E-ISSN : 3025-1311
- [6] Perani Rosyani, Saprudin Saprudin, Resti Amalia "Klasifikasi Citra Menggunakan Metode Random Forest dan Sequential Minimal Optimization (SMO)", *Kohesi: Jurnal Volume 9 No 2 Tahun 2021*, E-ISSN : 2620-8989
- [7] Perani Rosyani, Muhammad Faisal, Woro Agus Nurtiyanto "Implementasi Bahasa Pemrograman Python pada Metode-Metode Sistem Penunjang Keputusan", *Kohesi: Jurnal Volume 1 No 2 Tahun 2024*, ISBN : 978-623-120-128-7
- [8] N. L. Hanun and A. U. Zailani, "Penerapan algoritma klasifikasi Random Forest untuk penentuan kelayakan pemberian kredit di Koperasi Mitra Sejahtera," *Journal of Technology Information*, vol. 6, no. 1, p. 7, Juni 2020. P-ISSN: 2460-2108, E-ISSN: 2620-5181.
- [9] A. Fauzi and A. H. Yunial, "Optimasi Algoritma Klasifikasi Naive Bayes, Decision Tree, K-Nearest Neighbor, dan Random Forest menggunakan Algoritma Particle Swarm Optimization pada Diabetes Dataset," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 3, p. 470, Desember 2022. ISSN(p): 2460-0741, ISSN(e): 2548-9364.
- [10] A. Saifudin, U. U. Nabillah, Yulianti, and T. Desyani, "Bagging Technique to Reduce Misclassification in Coronary Heart Disease Prediction Based on Random Forest," in *Journal of Physics: Conference Series*, vol. 1477, no. 3, p. 032009, 2020. IOP Publishing. doi: 10.1088/1742-6596/1477/3/032009.