

1. Pendahuluan

Demam Berdarah Dengue merupakan salah satu penyakit yang memprihatinkan karena kemunculannya yang terus menerus dan kasusnya yang fluktuatif. Berdasarkan data tahunan Kementerian Kesehatan, pada tahun 2022 Jawa Barat menjadi provinsi dengan kasus DBD tertinggi. Jumlah kasusnya mencapai 36.608 dari total 143.266 kasus DBD di Indonesia [1]. Pada tahun 2023 Jawa Barat juga memiliki jumlah kasus DBD yang besar, yaitu mencapai 19.328 dari total 114.720 kasus DBD di Indonesia [2]. Sedangkan di Bandung, sebagai ibu kota Jawa Barat, kasus DBD-nya berfluktuasi setiap tahunnya. Berdasarkan beberapa data puskesmas di Bandung antara tahun 2017 sampai dengan tahun 2022, kasus terendah terjadi pada tahun 2017 sebanyak 1.786 kasus dan tertinggi terjadi pada tahun 2022 sebanyak 5.205 kasus [3]. Data ini menunjukkan betapa fluktuatif dan memprihatinkannya kasus DBD di Jawa Barat, begitu pula dengan Bandung sebagai ibu kotanya. Oleh karena itu, penelitian ini bertujuan untuk memprediksi dan mengklasifikasikan kejadian DBD di kecamatan-kecamatan di Kota Bandung tahun 2022, 2023, 2024 dan 2025 dalam suatu bentuk kontribusi untuk mencegah penyebaran DBD.

Latar Belakang

Meskipun banyak penelitian mengenai prediksi kejadian DBD, jarang ada penelitian yang mendalami untuk memprediksi tingkat kejadian DBD beberapa tahun ke depan menggunakan *Decision Tree* berdasarkan spasial-temporal lokasi tertentu. Mengenai prediksi masalah kesehatan, penelitian-penelitian ini [4], [5], dan [6] telah mencoba membuat prediksi menggunakan beberapa metode termasuk *Decision Tree*. Penelitian [4] meneliti prediksi kasus DBD menggunakan *Support Vector Machines* (SVM) dengan akurasi 84%, *K-Nearest Neighbor* (KNN) memperoleh akurasi 87%, dan *Decision Tree* (DT) memperoleh akurasi 79%. Dengan menerapkan pendekatan klasifikasi *hybrid* menggunakan teknik *hard voting*, ketiga metode akurasi klasifikasi dapat ditingkatkan mencapai akurasi 91%. Penelitian [5] meneliti prediksi kanker payudara menggunakan *Decision Tree* (DT) dan *Adaptive Boosting* (Ad boost), di mana dataset sangat tidak seimbang. Untuk meningkatkan kinerja *Decision Tree* dalam mengidentifikasi observasi ganas, *Adaptive Boosting* digunakan. Mendapatkan akurasi sebesar 92,53% untuk *Adaptive Boosting* dan 88,80% untuk *Decision Tree*. Studi [6] menguji prediksi probabilitas stroke dengan *Decision Tree* dan *Naive Bayes*. Studi ini memperoleh *Decision Tree* dengan menghitung koefisien Gini dari setiap fitur untuk memilih divisi. Setiap prediksi model *Decision Tree* dan *Naive Bayes* memberikan akurasi sebesar 88% dan 79%. Studi-studi ini menunjukkan bagaimana *Decision Tree* dapat memberikan akurasi yang tinggi dengan peningkatan yang tepat. Hal ini juga menunjukkan bahwa pemilihan metode terbaik untuk membangun *Decision Tree* dan menangani data ketidakseimbangan juga penting. Dalam penelitian ini, pemilihan fitur dan perluasan fitur diimplementasikan dalam upaya agar *Decision Tree* menghasilkan prediksi yang akurat.

Terkait analisis spasial-temporal dengan ekspansi fitur berbasis waktu, penelitian-penelitian [7], [8], [9] memprediksi iklim dan penyebaran penyakit berdasarkan beberapa fitur yang mempengaruhi penyebabnya. Penelitian [7] menggunakan *Naive Bayes* untuk memprediksi klasifikasi kasus demam berdarah Kota Bandung dan sebaran curah hujan bulanan Pulau Jawa dengan menggunakan ekspansi fitur yang memperoleh akurasi lebih dari 97% untuk kedua prediksi tersebut. Penelitian [8] melakukan penelitian untuk memprediksi klasifikasi penularan COVID-19 dan demam berdarah. Dengan menggunakan ekspansi fitur berbasis waktu SVM, prediksi masing-masing kelas penularan DBD dan COVID-19 memperoleh akurasi sebesar 90% dan 93%. Penelitian [9] melakukan prediksi kelas penyebaran DBD dengan mengimplementasikan *Random Forest* dan juga mengadaptasi ekspansi fitur, sehingga menghasilkan tiga model optimum yaitu tiga model dengan akurasi sebesar 97%, 93%, dan 93%. Penelitian-penelitian tersebut menunjukkan bagaimana ekspansi fitur berhasil meningkatkan akurasi suatu model dengan akurasi keseluruhan lebih dari 90%, bahkan lebih baik dibandingkan penelitian-penelitian [4], [5], [6] yang modelnya tidak menggunakan ekspansi fitur berbasis waktu.

Meskipun kasus DBD terus meningkat dari waktu ke waktu di Jawa Barat khususnya di Bandung, ada peluang untuk mencegah penyebaran DBD dengan menerapkan prediksi kelas. Seperti yang telah ditunjukkan sebelumnya, penelitian [4], [5], [6], [7], [8] dan [9] berhasil melakukan klasifikasi mengenai masalah kesehatan dengan hasil yang memuaskan. Penelitian tersebut juga menunjukkan bahwa *Decision Tree* merupakan salah satu metode yang sering digunakan untuk menangani klasifikasi dengan hasil yang baik, sedangkan penggunaan berbasis waktu mampu meningkatkan akurasi model. Dalam penelitian ini, *Decision Tree* dipilih berdasarkan seberapa sering metode ini digunakan untuk tugas klasifikasi dalam prediksi terkait kesehatan dengan hasil yang dapat diandalkan. Analisis spasial-temporal dengan perluasan fitur berbasis waktu dipilih karena menangkap interaksi antara ruang dan waktu, memungkinkan pemahaman yang lebih komprehensif tentang pola penyebaran DBD di lokasi dan periode waktu tertentu sekaligus meningkatkan akurasi model *Decision Tree*. Meskipun penggunaan *Decision Tree* dan *time-based feature expansion* untuk prediksi kelas sudah dilakukan dalam beberapa penelitian, tidak banyak penelitian yang menggabungkan kedua metode ini untuk memberikan analisis yang terperinci mengenai prediksi kelas yang dilengkapi dengan visualisasi lokasi dan waktu tertentu. Kombinasi metode ini juga memunculkan kemungkinan untuk menciptakan hasil yang lebih akurat dan tepat dibandingkan dengan penelitian lainnya.

Topik dan Batasannya

Penelitian ini melakukan prediksi kelas penyebaran DBD menggunakan *Decision Tree* di 30 kecamatan di Kota Bandung untuk tahun 2022, 2023, 2024 dan 2025. Menggunakan dataset dari 4 tahun sebelumnya, *time-based feature expansion* diimplementasikan untuk meningkatkan akurasi. Hasil dari penelitian ini adalah fitur-fitur yang paling memengaruhi penyebaran DBD, akurasi setiap model sebagai perbandingan untuk memilih model terbaik, hasil prediksi kelas setiap kecamatan dan visualisasi berdasarkan prediksi kelas untuk memberikan pemahaman yang lebih baik. Untuk mendapatkan hasil prediksi terbaik menggunakan *Decision Tree*, penelitian ini menguji fitur terbaik untuk digunakan berdasarkan akurasinya dengan teknik pemilihan fitur dan perluasan fitur dimana dataset ini adalah fitur-fitur dari tahun 2017 sampai dengan 2021, yang dipisahkan menjadi beberapa model. Prediksi klasifikasi angka kejadian DBD tahun 2022, 2023, 2024 dan 2025 dilakukan dengan menggunakan model terbaik dengan akurasi yang luar biasa untuk memberikan prediksi yang akurat.

Tujuan

Penelitian ini bertujuan untuk memprediksi dan mengklasifikasikan kejadian DBD di kecamatan-kecamatan di Kota Bandung tahun 2022, 2023, 2024 dan 2025 dalam suatu bentuk kontribusi untuk mencegah penyebaran DBD. Penelitian ini akan bermanfaat bagi sektor kesehatan dan pemangku kepentingan lainnya dengan memberikan solusi preemtif untuk meminimalkan kejadian DBD dengan menginformasikan fitur-fitur yang memiliki pengaruh besar terhadap angka DBD dan prediksi kelas yang menunjukkan tingkat keparahan kasus DBD di kecamatan-kecamatan tersebut. Dengan adanya penelitian ini diharapkan dapat memberikan solusi secara langsung untuk memberikan informasi kepada masyarakat dan instansi kesehatan mengenai penyebaran penyakit DBD di masa yang akan datang serta faktor yang berpengaruh sehingga pemerintah dapat memberikan kebijakan sedini mungkin untuk mencegah penyebaran penyakit DBD.

Organisasi tulisan

Penelitian ini terdiri atas beberapa bagian yaitu studi terkait, sistem yang dibangun, evaluasi dan kesimpulan. Studi terkait merupakan kumpulan studi yang menjadi bahan pembelajaran dan refleksi bagi penelitian ini sehingga dapat diketahui kekurangan, kelebihan dan hal yang perlu diperbaiki atau bisa dikembangkan terhadap penelitian. Pada bagian sistem yang dibangun akan dibahas bagaimana penelitian ini menyelesaikan permasalahan untuk memprediksi kelas penyebaran DBD dengan metode yang dipilih yaitu *Decision Tree* dengan *time-based feature expansion*. Pada evaluasi akan dibahas hasil dari prediksi menggunakan metode yang telah dibangun beserta visualisasinya. Pada akhirnya dapat diambil kesimpulan mengenai bagaimana penelitian ini dapat melakukan prediksi dengan akurasi yang optimal.