

## 1. Pendahuluan

Demam berdarah dengue (DBD) merupakan masalah kesehatan di Indonesia. Dengan jumlah penduduk 267 juta jiwa, Indonesia berisiko tertular penyakit demam berdarah. Faktor utama penyebab penyakit Demam berdarah disebabkan oleh gigitan nyamuk *Aedes Aegypti* yang membawa virus *dengue*. faktor lain dapat disebabkan oleh curah hujan yang tinggi, kelembapan, dan temperatur suhu. faktor cuaca sangat berpengaruh terhadap proses berkembang-biaknya nyamuk tersebut [1].

Menurut Kementerian Kesehatan Republik Indonesia, jumlah penderita DBD di Indonesia dari awal tahun hingga Juli 2020 sebanyak 71.633 kasus dengan 459 kematian, sedangkan pada tahun 2019 sebanyak 919 kematian dengan 138.127 kasus [2]. Wilayah di Indonesia yang memiliki jumlah kasus terbanyak adalah Jawa Barat dengan peringkat tertinggi dengan 10.772 kasus. wilayah Jawa Barat yang memiliki jumlah kasus terbanyak adalah Kota Bandung. Kota Bandung tercatat memiliki jumlah kasus terbanyak, yakni 4.424 kasus.

Untuk mengatasi banyaknya masalah jumlah kasus DBD yang terjadi di Kota Bandung, maka diberikan solusi untuk mengklasifikasi pengaruh cuaca terhadap jumlah kasus yang terjadi. klasifikasi berguna untuk mengelompokkan data berdasarkan kelas yang sudah di tetapkan [3]. salah satunya penggunaan metode klasifikasi pada algoritma machine learning. belum banyak penelitian untuk meningkatkan hasil kinerja klasifikasi. biasanya hanya berfokus pada hasil klasifikasi akurasi yang diperoleh.

Pada penelitian sebelumnya menggunakan *hybrid stacked classifier* dengan SVM, KNN, dan C5.0. hasil yang diperoleh pada klasifikasi individu SVM diperoleh akurasi terbaik dengan 88,2%. sedangkan untuk *hybrid SVM, KNN dan C5.0* menggunakan meta RF dan GLM dengan 10 cross validation masing-masing sebesar 91,2% dan 90,8%. dalam penelitian ini menyimpulkan bahwa variasi model kombinasi klasifikasi dengan pendekatan ensemble dapat meningkatkan akurasi hasil [4].

Pada penelitian serupa lainnya, penyakit DBD di Indonesia diprediksi menggunakan *Decision Tree* dengan akurasi 76% [5]. Penelitian lain menggunakan perbandingan algoritma *Support Vector Machine* dan C4.5 untuk identifikasi hama dan penyakit pada tanaman cabai, dengan akurasi masing-masing 82,33% dan 89,25% [6]. Studi lain menggunakan *K-Nearest Neighbor* untuk diagnosis kanker payudara memperoleh akurasi 77,5% [7].

Berdasarkan penelitian di atas dapat disimpulkan bahwa masih sedikit penelitian yang mengembangkan kinerja algoritma klasifikasi menggunakan ensemble learning. Dalam penelitian ini, penulis memfokuskan pada *hybrid classifier* dengan menggunakan metode *hard voting ensemble* dari tiga algoritma klasifikasi yaitu *Support Vector Machine*, *K-Nearest Neighbor* dan *Decision Tree*. Pendekatan *voting ensemble* dapat memecahkan masalah penentuan hasil prediksi yang lebih baik. dapat dilihat dari penelitian diatas bahwa setiap algoritma klasifikasi memiliki akurasi yang cukup baik. dengan menggunakan *voting ensemble* dari ketiga algoritma klasifikasi maka hasil prediksi dapat lebih optimal dan akurasi yang diperoleh akan meningkat. kontribusi penelitian ini adalah untuk memberikan informasi bahwa metode *voting ensemble* dapat meningkatkan hasil prediksi yang lebih optimal dalam proses pengklasifikasian.