

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang diperoleh, maka peneliti dapat memberikan beberapa kesimpulan yang dihasilkan sebagai berikut:

1. Hasil analisis menunjukkan bahwa model dengan `max_depth` 24 dan `n_estimator` 50 menghasilkan akurasi dan f1-score tertinggi, yaitu sebesar 80,94% dan 81,10% secara berturut-turut. Hasil ini menandakan bahwa kombinasi nilai parameter tersebut memberikan performa yang lebih baik dalam melakukan klasifikasi data. Sebaliknya, model dengan `max_depth` 12 dan `n_estimator` 100 menunjukkan akurasi dan f1-score terendah, yaitu sebesar 78,89% dan 78,93%. Perbedaan hasil ini menegaskan bahwa nilai `n_estimator` menjadi faktor paling berpengaruh dalam meningkatkan akurasi dan f1-score. Semakin kecil nilai `n_estimator`, model cenderung lebih baik dalam melakukan prediksi. Meskipun `max_depth` juga mempengaruhi performa model, tidak selalu berarti semakin tinggi nilai `max_depth` akan meningkatkan akurasi dan f1-score. Hasil analisis menunjukkan bahwa ada titik optimal dimana `max_depth` mencapai nilai tertentu yang memberikan performa terbaik. Penetapan nilai `max_depth` yang terlalu tinggi dapat menyebabkan *overfitting*, dimana model terlalu kompleks dan hanya cocok dengan data pelatihan, namun tidak mampu melakukan prediksi dengan baik pada data baru. Oleh karena itu, penting untuk mencari titik keseimbangan yang tepat dalam memilih nilai `max_depth` yang sesuai untuk memaksimalkan kinerja model.
2. Parameter terbaik setelah dilakukan optimasi *tunning* hyperparameter menggunakan *grid search* yaitu dengan parameter `booster = gbtrees`, `learning_rate = 0.1`, `max_depth = 24`, `min_child_weight = 30`, `n_estimator = 50`, dan `n_thread = 6`. Parameter tersebut memberikan akurasi secara keseluruhan dengan nilai 80,94%. Dari perhitungan menggunakan *confusion*

matrix didapatkan nilai akurasi, precision, recall, dan F1-score masing-masing sebesar 80,94%, 80,41%, 81,80%, dan 81,10% yang berarti bahwa klasifikasi yang diperoleh cukup.

5.2 Saran

Berdasarkan hasil analisis yang telah dilakukan, ada beberapa saran yang dapat diberikan untuk meningkatkan performa model dalam klasifikasi data menggunakan parameter `max_depth` dan `n_estimator`.

Pertama, parameter `n_estimator` memiliki pengaruh signifikan dalam meningkatkan akurasi dan f1-score pada model. Oleh karena itu, disarankan untuk mengurangi nilai `n_estimator` agar model lebih baik dalam melakukan prediksi. Namun, perlu diingat bahwa penurunan nilai `n_estimator` tidak boleh terlalu drastis karena dapat menyebabkan model kehilangan kemampuan untuk menggeneralisasi data dengan baik. Cobalah untuk melakukan eksperimen dengan nilai `n_estimator` yang berbeda secara bertahap untuk mencari titik optimal yang memberikan hasil terbaik.

Kedua, perhatikan nilai `max_depth` dengan cermat. Meskipun `max_depth` juga mempengaruhi performa model, bukan berarti semakin tinggi nilai `max_depth` akan selalu meningkatkan akurasi dan f1-score. Jika `max_depth` terlalu tinggi, maka model berpotensi mengalami overfitting dan hanya cocok dengan data pelatihan saja. Cari nilai `max_depth` yang mencapai titik keseimbangan di mana performa model mencapai hasil terbaik tanpa terjadi overfitting pada data.

Selain itu, penting untuk selalu memperhatikan seimbangannya data dan nilai parameter yang digunakan. Pastikan bahwa dataset telah diseimbangkan dengan melakukan pruning untuk menghindari bias dan mencegah dominasi satu kelas atas kelas lainnya. Dengan memperhatikan faktor-faktor tersebut dan memahami interaksi antara kedua parameter, kita dapat mengoptimalkan kinerja model dan menciptakan prediksi yang lebih akurat dalam berbagai aplikasi machine learning.