

BAB V

PENUTUP

5.1 Kesimpulan

Penerapan metode Synthetic Minority Over-sampling Technique (SMOTE) berhasil digunakan untuk mengatasi ketidakseimbangan distribusi kelas pada dataset klasifikasi cuaca. Proses oversampling menghasilkan jumlah data yang lebih seimbang pada setiap kategori cuaca sehingga model memperoleh kesempatan yang sama dalam mempelajari karakteristik masing-masing kelas. Dengan demikian, penggunaan SMOTE membantu mengurangi bias model terhadap kelas mayoritas dan mendukung proses pelatihan model klasifikasi yang lebih optimal.

Algoritma Extreme Gradient Boosting (XGBoost) berhasil diterapkan untuk membangun model klasifikasi cuaca dengan memanfaatkan teknik hyperparameter tuning melalui metode Grid Search dan Random Search. Proses tuning dilakukan terhadap beberapa parameter penting, yaitu `learning_rate`, `max_depth`, `min_child_weight`, `gamma`, `subsample`, `colsample_bytree`, dan `n_estimators`, sehingga diperoleh kombinasi parameter yang mampu meningkatkan performa model dibandingkan penggunaan parameter default.

Hasil perbandingan menunjukkan bahwa kedua metode hyperparameter tuning mampu meningkatkan performa model dibandingkan model XGBoost dasar. Model hasil Grid Search memperoleh CV Score sebesar 0,9116 dengan accuracy sebesar 0,8309, precision sebesar 0,8313, recall sebesar 0,8309, dan F1-score sebesar 0,8290. Sementara itu, model hasil Random Search memperoleh CV Score sebesar 0,9091 dengan accuracy sebesar 0,8358, precision sebesar 0,8379, recall sebesar 0,8358, dan F1-score sebesar 0,8342. Berdasarkan hasil tersebut, Grid Search menghasilkan nilai validasi silang tertinggi, sedangkan Random Search memberikan performa terbaik pada data uji sehingga dapat disimpulkan bahwa Random Search merupakan metode optimasi yang paling efektif pada dataset dan ruang pencarian parameter yang digunakan dalam penelitian ini.

5.2 Saran

Berdasarkan hasil implementasi dan pengujian yang telah dilakukan, beberapa saran untuk penelitian selanjutnya adalah sebagai berikut:

1. Menggunakan dataset dengan karakteristik yang lebih beragam, baik dari segi jumlah data, jumlah kelas, distribusi kelas, maupun cakupan wilayah pengamatan.
2. Melakukan pengembangan pada tahap prapemrosesan data, seperti rekayasa fitur (*feature engineering*), seleksi fitur, maupun pembentukan fitur baru yang lebih representatif terhadap kondisi cuaca.
3. Memperluas ruang pencarian hyperparameter dengan menambahkan parameter lain atau menggunakan rentang nilai yang lebih luas.
4. Membandingkan metode optimasi hyperparameter dengan pendekatan lainnya, seperti Bayesian Optimization, Genetic Algorithm, Particle Swarm Optimization, atau metode optimasi lainnya.
5. Melakukan pengujian pada model klasifikasi lain untuk mengetahui apakah pola peningkatan performa yang diperoleh melalui optimasi hyperparameter juga terjadi pada model yang berbeda.

