

BAB V

PENUTUP

5.1. Kesimpulan

Dengan mengacu pada hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa:

1. Algoritma *Random Forest* terbukti efektif dalam memprediksi pilihan paket Umrah berdasarkan data calon jemaah seperti usia, jenis kelamin, wilayah, waktu keberangkatan, dan metode pembayaran, yang menjadi fitur penting dalam klasifikasi. Model ini berhasil mengklasifikasikan calon jemaah ke dalam enam kategori paket yang tersedia, yaitu Paket Reguler 3, 4, dan 5 Bintang, serta Paket Plus A, B, dan C. Hasil prediksi dapat langsung digunakan dalam bentuk *dashboard* interaktif berbasis *Streamlit*, untuk mendapatkan rekomendasi paket secara otomatis.
2. Model mencapai akurasi sebesar 93,25% pada data uji. Nilai ini didukung oleh skor *F1-score macro* sebesar 0,90 dan *weighted* sebesar 0,93, menunjukkan model cukup andal dalam mengenali semua kelas secara seimbang. Meskipun masih terdapat kesalahan klasifikasi antar beberapa kelas yang mirip, performa keseluruhan model sudah tergolong sangat baik untuk kebutuhan sistem rekomendasi paket Umrah.
3. Penerapan metode SMOTE memiliki pengaruh yang signifikan terhadap peningkatan performa model. Sebelum SMOTE diterapkan, model hanya memperoleh akurasi sekitar 73,69%, dan cenderung bias terhadap kelas mayoritas. Setelah dilakukan *oversampling* dengan SMOTE pada data *training*, akurasi meningkat drastis menjadi 93,25%, serta terjadi peningkatan pada *precision* dan *recall* di hampir semua kelas, termasuk kelas minoritas. Hal ini menunjukkan bahwa penggunaan SMOTE sangat penting untuk mengatasi ketidakseimbangan data dalam proses klasifikasi, dan berkontribusi terhadap model yang lebih adil dan akurat.

5.2. Saran

Penelitian ini memberikan beberapa saran yang bertujuan untuk memperbaiki dan mengembangkan penelitian di masa depan. Beberapa saran yang disarankan oleh peneliti antara lain:

1. Menggunakan model algoritma klasifikasi lain, seperti *Support Vector Machine (SVM)*, atau *XGBoost* untuk membuktikan dan membandingkan apakah akurasi lebih baik.
2. Memanfaatkan metode *oversampling* atau *undersampling* lainnya untuk menentukan pendekatan yang paling efektif dalam mengatasi ketidakseimbangan kelas data.
3. Meningkatkan antarmuka prediksi dengan mengembangkan *dashboard* interaktif yang lebih kompleks, menambahkan fitur, dan responsif untuk memudahkan pengguna menelusuri hasil prediksi secara intuitif dan efisien.