

BAB V

KESIMPULAN

5.1 KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan mulai dari tahap pendahuluan, landasan teori, metodologi, hingga analisis dan interpretasi pada Bab IV, serta mengacu pada dua tujuan penelitian maka, dapat diambil beberapa kesimpulan sebagai berikut:

1. Penerapan metode klasifikasi *Naïve Bayes*, *Decision Tree* (J48), dan *Random Forest* dengan teknik SMOTE serta seleksi fitur *Gain Ratio* berhasil diterapkan untuk menentukan kelayakan siswa SMA penerima PIP di Kota Jambi.

Seluruh tahapan mulai dari pre-processing, balancing menggunakan SMOTE, seleksi fitur *Gain Ratio*, pembentukan model, hingga pengujian telah berjalan sesuai prosedur. SMOTE berhasil menyeimbangkan distribusi kelas sehingga model dapat mengenali kelas “Ya” dengan lebih baik.

Penerapan *Gain Ratio* mampu menyaring atribut yang paling relevan, meskipun pada dataset ini penghilangan atribut tertentu membuat sebagian model mengalami penurunan akurasi.

Secara keseluruhan, ketiga algoritma dapat digunakan untuk proses klasifikasi kelayakan siswa penerima PIP dan menghasilkan model yang dapat membantu proses penentuan kelayakan secara objektif dan berbasis data.

2. Hasil komparasi menunjukkan bahwa algoritma *Random Forest* memiliki performa terbaik dibandingkan *Naïve Bayes* dan *Decision Tree* (J48), baik pada pengujian data *Training* maupun data *Testing*.

- a. *Random Forest* menghasilkan akurasi paling tinggi dan stabil pada seluruh skenario pengujian.
- b. *Naïve Bayes* dan *Decision Tree* (J48) juga mampu melakukan klasifikasi dengan baik, tetapi performanya lebih rendah dibanding *Random Forest*. Pengujian juga menunjukkan bahwa akurasi tanpa seleksi fitur pada umumnya lebih tinggi dibandingkan dengan seleksi fitur. Hal ini terjadi karena sebagian atribut yang dihapus melalui *Gain Ratio* ternyata masih mengandung informasi penting yang dibutuhkan model, sehingga penghilangan atribut menyebabkan hilangnya pola tertentu dalam data. Dengan demikian, model terbaik dalam penelitian ini adalah *Random Forest* tanpa seleksi fitur.

5.2 SARAN

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan, maka saran-saran yang dapat diberikan untuk penelitian selanjutnya maupun pihak terkait adalah sebagai berikut:

1. Penelitian selanjutnya dapat mencoba algoritma lain seperti *XGBoost*, *AdaBoost*, *SVM*, atau *Deep Learning* untuk melihat apakah terdapat peningkatan akurasi dibandingkan *Random Forest*.

2. Atribut penelitian dapat diperluas dengan menambahkan faktor-faktor lain seperti kondisi tempat tinggal, status kepemilikan rumah, jumlah tanggungan, maupun indikator sosial ekonomi lainnya agar hasil prediksi semakin kaya dan akurat.
3. Sistem klasifikasi dapat dikembangkan menjadi Sistem Pendukung Keputusan (SPK) berbasis web atau aplikasi, sehingga pihak sekolah, BPMP, dan instansi terkait dapat menggunakannya secara langsung dalam proses verifikasi kelayakan penerima PIP.
4. Perlu dilakukan pembaruan data secara berkala karena kondisi ekonomi keluarga dapat berubah dari waktu ke waktu. Model yang baik tetap memerlukan data terbaru agar performa klasifikasi tetap akurat.
5. Penerapan SMOTE sudah baik, namun dapat ditingkatkan dengan menguji variasi parameter SMOTE atau mencoba teknik balancing lain seperti ADASYN atau *Random Undersampling* untuk melihat pengaruh terhadap kualitas model.
6. Penelitian selanjutnya dapat menerapkan *ensemble model* atau *hybrid model* seperti kombinasi *Random Forest* + *Naïve Bayes* atau *stacking classifier* untuk meningkatkan stabilitas dan akurasi prediksi.