

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian, pengujian, dan pembahasan mengenai penerapan data mining untuk klasifikasi tingkat kesehatan mental pekerja menggunakan algoritma Naive Bayes dan XGBoost, maka dapat ditarik beberapa kesimpulan sebagai berikut:

- 1 Penelitian ini berhasil mengimplementasikan enam algoritma klasifikasi untuk memprediksi kebutuhan penanganan kesehatan mental pekerja dengan memanfaatkan delapan fitur kunci yang paling relevan, yaitu *Age*, *Gender*, *family_history*, *benefits*, *care_options*, *anonymity*, *leave*, dan *work_interfere*.
- 2 Penerapan teknik SMOTE (*Synthetic Minority Over-sampling Technique*) terbukti sangat efektif dalam menangani ketidakseimbangan kelas pada dataset, yang ditunjukkan dengan peningkatan signifikan pada nilai *Recall* atau kemampuan model dalam mendeteksi individu yang benar-benar membutuhkan bantuan profesional.
- 3 Berdasarkan hasil evaluasi komprehensif, algoritma Naive Bayes dan XGBoost muncul sebagai model dengan performa terbaik. Algoritma Naive Bayes berhasil mencatatkan akurasi sebesar 0,8413 dengan skor F1 yang seimbang, sementara XGBoost menunjukkan keunggulan pada nilai *Recall* mencapai 0,9055, menjadikannya model yang paling sensitif dan reliabel untuk deteksi dini.
- 4 Perbandingan dengan penelitian terdahulu menunjukkan bahwa model yang diusulkan dalam penelitian ini mampu melampaui capaian akurasi studi sebelumnya. Hal ini mengonfirmasi bahwa integrasi antara prosedur pra-pemrosesan data yang sistematis (normalisasi dan feature selection) dengan pemilihan algoritma ensemble learning merupakan pendekatan yang sangat unggul untuk mengklasifikasikan karakteristik dataset kesehatan mental pekerja yang kompleks.

5.2 Saran

Berdasarkan hasil penelitian yang telah dicapai, terdapat beberapa saran yang dapat diajukan untuk pengembangan penelitian selanjutnya, antara lain:

1. Melakukan deployment model terbaik ke dalam sebuah platform aplikasi berbasis web menggunakan framework Streamlit agar dapat digunakan secara praktis oleh perusahaan dalam memantau kesehatan mental pekerja secara real-time.
2. Menambahkan variabel prediktor baru yang lebih spesifik bagi lingkungan profesional, seperti beban kerja, jam kerja lembur, atau data biometrik untuk meningkatkan kedalaman analisis dan akurasi prediksi.
3. Melakukan penyetelan parameter (hyperparameter tuning) yang lebih mendalam pada algoritma ensemble lainnya guna mengeksplorasi potensi peningkatan kinerja model di luar hasil yang telah dicapai.
4. Mempertimbangkan penggunaan teknologi pendukung seperti blockchain untuk menjamin keamanan dan anonimitas data medis pekerja sesuai dengan standar privasi data kesehatan yang berlaku.