

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian, beberapa temuan penting terkait penanganan ketidakseimbangan kelas pada dataset diabetes dapat disimpulkan sebagai berikut:

1. Ketidakseimbangan kelas pada dataset diabetes memengaruhi performa model Random Forest dan XGBoost, sehingga perlu diterapkan metode sampling.
2. Metode sampling terbukti dapat meningkatkan kemampuan model mendeteksi kelas minoritas berdasarkan dari kenaikan F1-score dan akurasi.
3. Dalam penelitian ini, Tomek Links menunjukkan efektivitas terbaik dalam meningkatkan performa model Random Forest dan XGBoost.
4. XGBoost dengan Tomek Links memberikan performa terbaik secara keseluruhan.
5. Random Forest dengan Tomek Links menunjukkan performa yang stabil setelah penerapan sampling.
6. Pemilihan metode sampling dan algoritma harus disesuaikan dengan karakteristik data dan tujuan penelitian.

5.2 Saran

Adapun beberapa saran yang dapat dijadikan saran untuk penelitian selanjutnya antara lain:

1. Peneliti berikutnya dapat mengeksplorasi metode sampling atau algoritma lain yang lebih kompleks untuk meningkatkan performa model pada kelas minoritas.

2. Penggunaan kombinasi beberapa Teknik sampling dapat dieksplorasi lebih lanjut untuk melihat apakah performa model Random Forest dan XGBoost bisa ditingkatkan
3. Dataset yang lebih besar, lebih beragam, atau dari sumber yang berbeda dapat digunakan agar model lebih general dan akurat dalam prediksi diabetes.
4. Analisis tambahan dengan matrik evaluasi lain dapat dilakukan untuk memberikan gambaran lebih mendalam tentang kemampuan model dalam mengenali kelas minoritas.
5. Proses preprocessing dan EDA dapat ditingkatkan untuk memperoleh informasi yang lebih banyak mengenai data agar performnya bisa meningkat.

