

BAB V PENUTUP

5.1. Kesimpulan

Kesimpulan yang didapatkan dari hasil penelitian yang dilakukan, yaitu

1. Setelah dilakukan penerapan kombinasi metode *oversampling* dan *undersampling*. Terdapat peningkatan dari performa model KNN. Hal tersebut dibuktikan dengan adanya peningkatan nilai *recall* dan *f1-score*. Pada penerapan KNN dengan metode *oversampling* SMOTE dan TomekLink menghasilkan nilai akurasi sebesar 79,9%, *recall* 46,6%, dan *f1-score* 18,3%. Terdapat penurunan akurasi sebanyak 0,1%, penurunan 0,1% pada *recall*, dan penurunan 0,2% dari *f1-score* jika dibandingkan dengan penerapan teknik SMOTE. Sedangkan pada penerapan metode *oversampling* SMOTE dan ENN menghasilkan nilai akurasi sebesar 76,8%, nilai *recall* 59,6%, dan *f1-score* sebesar 20,1%. Terdapat penurunan akurasi sebesar 3,2%, peningkatan *recall* sebesar 13% dan peningkatan *f1-score* sebesar 1,4%. Hasil tersebut membuktikan bahwa penggunaan kombinasi metode *oversampling* dan *undersampling* tidak selalu meningkatkan performa. Pada Tomek Link, performa model menurun. Namun pada penerapan ENN dapat meningkatkan performa model sebagaimana hasil evaluasi yang disebutkan sebelumnya. Berdasarkan hasil evaluasi model. Metode kombinasi *resampling* yang terbaik untuk meningkatkan performa model pada penelitian ini adalah SMOTEEN.
2. Berdasarkan hasil evaluasi model, *information gain* berhasil meningkatkan performa model. Hal tersebut dibuktikan dengan adanya peningkatan nilai *recall* dan *f1-score* pada SMOTEENN. Sebelum dilakukan *resampling* nilai *recall* yang diperoleh sebesar 45,3% dan *f1-score* sebesar 16,8% meningkat menjadi *recall* sebesar 59,6% dan *f1-score* sebesar 20,1%. Pada SMOTETomek juga dibuktikan adanya peningkatan nilai *recall* dan *f1-score*. Sebelum dilakukan *resampling* nilai *recall* yang diperoleh sebesar 25,1% dan

f1-score sebesar 12,5% meningkat menjadi *recall* sebesar 46,6% dan *f1-score* sebesar 18,3%.

3. Sistem untuk memprediksi penyakit stroke telah dibangun dengan menggunakan *cloud function* sebagai *backend* dan Nextjs sebagai *frontend*. *Cloud function* digunakan untuk mengeksekusi model K-NN dengan metode kombinasi *resampling* dan fitur terbaik dari *information gain* yang telah diuji validasi dan disimpan pada *supabase storage*. Ketiga komponen sistem saling berkomunikasi dengan baik, sehingga sistem dapat menampilkan hasil dari prediksi penyakit stroke kedalam tampilan pengguna.
4. Penerapan teknik *resampling* dapat meningkatkan hasil evaluasi jika dibandingkan dengan pembangunan model menggunakan data yang *imbalance*. Hal tersebut dibuktikan dengan penerapan KNN dengan data yang *imbalance data* menghasilkan akurasi yang tinggi sebesar 94,1%. Namun, nilai evaluasi yang lain tergolong rendah yaitu *recall* sebesar 5% dan *f1-score* sebesar 13%.

5.2. Saran

Berdasarkan hasil analisis dan pembahasan pada penelitian yang telah dilaksanakan, diberikan saran diantara lain.

1. Penelitian selanjutnya dapat menerapkan teknik seleksi fitur yang lain seperti *Particle Swarm Optimization*, *Random Forest Feature Importance*, atau Algoritma Genetika.
2. Penelitian selanjutnya dapat melakukan percobaan untuk mengatasi *imbalanced data* dengan menerapkan metode lainnya seperti *Borderline-SMOTE*, *Borderline-SMOTE SVM*, *Adaptive Synthetic Sampling* (ADASYN).