

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan hasil penelitian mengenai penerapan Extreme Gradient Boosting (XGBoost) dan K-Nearest Neighbors Enhanced (K-NN Enhanced) dengan dan tanpa Principal Component Analysis (PCA) untuk prediksi tingkat keparahan pasien hemodialisis, diperoleh kesimpulan sebagai berikut.

1. Kedua algoritma berhasil diterapkan untuk klasifikasi tingkat keparahan berbasis parameter klinis yang tersedia. Setelah rangkaian prapemrosesan dan penyeimbangan data, keduanya menunjukkan performa yang stabil pada data uji.
2. XGBoost menjadi model terbaik dengan akurasi 91,65 persen, precision 0,92, recall 0,92, dan F1-score rata-rata 0,92 pada test set. Keunggulan ini selaras dengan karakter XGBoost yang mampu menangkap interaksi fitur, memiliki regularisasi, serta pembelajaran bertahap berbasis boosting.
3. Dampak PCA berbeda antar algoritma. Pada K-NN Enhanced, PCA meningkatkan akurasi dari 82,84 persen menjadi 84,12 persen karena reduksi dimensi mengurangi redundansi dan efek curse of dimensionality. Pada XGBoost, PCA sedikit menurunkan akurasi dari 91,65 persen menjadi 90,47 persen karena XGBoost sendiri efektif dalam seleksi fitur.
4. PCA berperan penting pada efisiensi komputasi terutama untuk algoritma berbasis jarak seperti K-NN. Reduksi dimensi mampu menekan noise tanpa kehilangan informasi yang berarti sehingga proses prediksi lebih efisien.
5. Optimasi hiperparameter dengan GridSearchCV 10-fold berkontribusi nyata terhadap peningkatan kinerja. Kombinasi terbaik yang diperoleh antara lain pada XGBoost: `learning_rate` 0,3, `max_depth` 6, `subsample` 0,8, `colsample_bytree` 1,0; serta pada K-NN Enhanced: `n_neighbors` 3 dengan pembobotan `distance`.
6. Integrasi reduksi dimensi pada skenario yang tepat dan optimasi hiperparameter memberikan hasil signifikan untuk data medis yang berdimensi tinggi dan tidak seimbang, meningkatkan efisiensi tanpa mengorbankan akurasi.

7. Dari sisi penerapan, model terbaik diintegrasikan ke dalam antarmuka pengguna grafis sebagai demonstrasi proses prediksi pada mode inferensi. Pengguna dapat mengunggah berkas data, sistem menjalankan prapemrosesan yang sama seperti pada penelitian, lalu menampilkan label prediksi Mild, Moderate, atau Severe tanpa pelatihan ulang.
8. Hasil penelitian berpotensi menjadi dasar pengembangan sistem pendukung keputusan medis untuk membantu pemantauan kondisi pasien hemodialisis secara lebih cepat dan akurat.

## 5.2 Saran

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, maka saran yang dapat diberikan untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

### 1. Pengayaan Dataset dan Variasi Fitur:

Disarankan untuk menggunakan dataset dengan jumlah data yang lebih besar dan bervariasi, mencakup lebih banyak parameter medis seperti kadar albumin, natrium, kalium, dan indeks massa tubuh (BMI) agar model dapat mempelajari pola yang lebih kompleks dan representatif terhadap kondisi pasien nyata.

### 2. Eksplorasi Algoritma dan Kombinasi Model Lain:

Penelitian berikutnya dapat mengombinasikan metode XGBoost dan K-NN Enhanced dengan teknik lain seperti Stacking, Bagging, atau Random Forest, atau membandingkannya dengan model berbasis neural network (misalnya MLP atau LSTM) untuk menilai performa pada data sekuensial pasien.

### 3. Penerapan Feature Engineering dan Optimasi Dimensi Lanjutan:

Selain PCA, dapat diuji metode reduksi dimensi lain seperti Linear Discriminant Analysis (LDA), t-SNE, atau Autoencoder untuk melihat perbandingan kinerja dan visualisasi distribusi data hasil reduksi secara lebih mendalam.

### 4. Penelitian Lanjutan pada Ketidakseimbangan Kelas (Class Imbalance):

Meskipun teknik Random OverSampling telah membantu menyeimbangkan distribusi kelas, pendekatan lanjutan seperti SMOTE, ADASYN, atau kombinasi oversampling-undersampling adaptif dapat dieksplorasi untuk menghasilkan representasi data minoritas yang lebih realistis.

**5. Penelitian Lanjutan pada Implementasi :**

Penelitian lanjutan dapat diarahkan pada integrasi model prediksi dengan sistem database online sehingga hasil prediksi dapat tersimpan secara otomatis dan digunakan untuk keperluan monitoring serta pengambilan keputusan dalam program kesehatan gizi balita secara berkelanjutan.

*Halaman ini sengaja dikosongkan*