

## **BAB V PENUTUP**

### **5.1. Kesimpulan**

Mengacu pada analisis dan pembahasan yang telah disajikan sebelumnya, dapat ditarik sejumlah kesimpulan sebagai berikut:

1. Model *Random Forest* (RF) terbukti memiliki performa terbaik dalam memprediksi tingkat obesitas dengan akurasi sebesar 94%, lebih unggul dibandingkan algoritma lain seperti *Decision Tree* (92%), *k-Nearest Neighbor* (84%), *Logistic Regression* (78%), *Naïve Bayes* (65%), dan *Support Vector Machine* (62%). Keunggulan RF terutama berasal dari mekanisme *ensemble learning* yang menggabungkan beberapa *decision tree* sehingga mampu menangkap pola kompleks dalam data dan mengurangi variansi prediksi.
2. Proses seleksi fitur menggunakan SpFSR (*Sparse Feature Selection and Ranking*) berhasil mereduksi jumlah fitur dari 16 menjadi 10 tanpa menurunkan performa model. Bahkan, pengurangan fitur ini meningkatkan akurasi *Random Forest* menjadi 96,3%. Variabel yang paling berpengaruh meliputi riwayat keluarga obesitas, tinggi badan, usia, jumlah konsumsi makanan utama per hari, dan frekuensi konsumsi sayuran. Temuan ini menunjukkan bahwa seleksi fitur tidak hanya meningkatkan efisiensi komputasi, tetapi juga memperjelas interpretabilitas model.
3. Temuan penelitian menegaskan bahwa obesitas merupakan kondisi multifaktorial yang dipengaruhi oleh faktor genetik, demografis, dan gaya hidup. Model *machine learning*, khususnya *Random Forest* dengan integrasi SpFSR, dapat dimanfaatkan sebagai alat pendukung bagi tenaga kesehatan dan pembuat kebijakan dalam melakukan deteksi dini serta merancang intervensi pencegahan obesitas yang lebih terarah.
4. Penelitian ini menggunakan satu *dataset* tunggal dari *ML Repository*, sehingga generalisasi model ke populasi lain masih terbatas. Selain itu, beberapa kelas obesitas yang memiliki karakteristik mirip cenderung mengalami kesalahan klasifikasi, meskipun performa keseluruhan tetap tinggi

## 5.2. Saran

Mengacu pada temuan penelitian, terdapat beberapa rekomendasi yang relevan untuk pengembangan penelitian selanjutnya agar diperoleh hasil yang lebih optimal dalam klasifikasi tingkat obesitas, yaitu sebagai berikut:

1. Studi di masa mendatang disarankan menggunakan *dataset* yang lebih luas serta lebih beragam, baik dari sisi jumlah sampel maupun karakteristik responden, sehingga model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik terhadap populasi yang berbeda.
2. Penggunaan algoritma pembelajaran mesin lain yang lebih canggih atau berbasis *ensemble* dan *boosting*, seperti *XGBoost*, *LightGBM*, atau *CatBoost*, dapat dipertimbangkan sebagai alternatif untuk meningkatkan performa klasifikasi dan membandingkan efektivitasnya dengan penelitian ini.
3. Proses evaluasi model pada penelitian selanjutnya dapat diperluas dengan menambahkan metrik evaluasi lain, seperti AUC-ROC atau *Matthews Correlation Coefficient* (MCC), guna memperoleh gambaran performa model yang lebih komprehensif, khususnya dalam menghadapi distribusi kelas yang tidak seimbang.
4. Selain metode seleksi fitur SpFSR, penelitian lanjutan dapat mengeksplorasi dan membandingkan metode seleksi fitur lainnya, baik berbasis *filter*, *wrapper*, maupun *embedded*, untuk mengetahui pengaruhnya terhadap efisiensi dan akurasi model klasifikasi obesitas.