

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Tahapan Penelitian

4.1.1. Pengumpulan Data

Pada penelitian ini, proses pengumpulan data dilakukan dengan memanfaatkan *dataset* obesitas yang tersedia pada *ML Repository*, yaitu himpunan data publik yang secara luas digunakan dalam penelitian *machine learning* di bidang kesehatan. *Dataset* tersebut terdiri dari 2.112 individu dengan total 17 variabel, yang dikategorikan berdasarkan status obesitas. Dari keseluruhan variabel tersebut, penelitian ini menggunakan 14 variabel prediktor yang dipisahkan ke dalam tiga kategori, (1) faktor kebiasaan makan, (2) faktor kondisi fisik, dan (3) faktor karakteristik fisik serta demografis.

Kategori pertama berfokus pada faktor risiko terkait pola makan harian, yang meliputi variabel tingkat konsumsi alkohol (CALC), frekuensi minum air harian (CH2O), frekuensi konsumsi sayuran (FCVC), kebiasaan mengonsumsi makanan utama di antara waktu makan (CAEC), dan pola konsumsi makanan harian (NCP). Kategori kedua mencakup variabel yang menggambarkan aktivitas fisik, seperti durasi aktivitas fisik harian (TUE), jenis transportasi yang digunakan (MTRANS), serta kebiasaan memantau asupan kalori (SCC). Sementara itu, kategori ketiga terdiri atas variabel yang menggambarkan karakteristik fisik dan demografis individu, antara lain usia, tinggi badan, dan berat badan. Pemilihan ketiga kategori variabel tersebut didasarkan pada pertimbangan teoretis yang menunjukkan hubungan kuat terhadap risiko terjadinya obesitas.

Untuk memastikan kualitas evaluasi model serta meningkatkan kemampuan generalisasi hasil, penelitian ini menerapkan strategi pembagian data yang komprehensif dan sistematis. *Dataset* secara keseluruhan dibagi menggunakan metode *stratified random split*, dengan 80% data (1.690 sampel) digunakan untuk pelatihan dan 20% data (422 sampel) untuk pengujian. Hal ini digunakan untuk menjaga keseimbangan proporsi setiap kelas obesitas pada seluruh *subset*, sehingga mengurangi potensi bias akibat ketidakseimbangan kelas. Pada tahap berikutnya, data pelatihan diproses menggunakan metode *5-fold cross-validation* untuk keperluan

optimasi serta validasi pemilihan fitur. Pendekatan ini bertujuan untuk meminimalkan risiko *overfitting* dan memberikan estimasi performa model yang lebih stabil, terutama dalam permasalahan klasifikasi multi-kelas yang memiliki kompleksitas tinggi. Meskipun validasi eksternal merupakan pendekatan ideal, kombinasi antara *stratified sampling* dan *cross-validation* memberikan evaluasi performa yang reliabel dan dapat dipertanggungjawabkan secara statistik. Tabel 4.1 menyajikan daftar variabel prediktif yang dipakai dalam penelitian ini, yang dikelompokkan ke dalam tiga kategori utama, yaitu kebiasaan makan, kondisi fisik, serta karakteristik individu dan demografis.

Tabel 4.1. Dataset Obesitas dan Variabel Prediktor

Kategori 1: <i>Eating Habits</i>	Kategori 2: <i>Physical Conditions</i>	Kategori 3: <i>Individual Characteristics</i>
<i>Vegetable Consumption Frequency (FCVC)</i>	<i>Physical Activity Frequency (FAF)</i>	<i>Age</i>
<i>Number of Major Food Consumption Between Meals (CAEC)</i>	<i>Transportation Usage (MTRANS)</i>	<i>Weight</i>
<i>Frequency of High Caloric Food Consumption (FAVC)</i>	<i>Calories Consumption Monitoring (SCC)</i>	<i>Gender</i>
<i>Frequency of Water Intake Daily (CH20)</i>	<i>Time Use (TUE)</i>	<i>Height</i>
<i>Consumption of Alcohol (CALC)</i>	-	-
<i>Consumption of Food Between Meals (NCP)</i>	-	-

Sementara itu, Tabel 4.2 menunjukkan distribusi jumlah sampel pada setiap kelas obesitas, yang menggambarkan variasi proporsi antar kelas dan menjadi dasar dalam penerapan *stratified sampling* serta evaluasi performa model klasifikasi. Distribusi kelas pada variabel target menunjukkan variasi proporsi yang cukup signifikan di antara kategori obesitas. Kelas dengan jumlah kasus terbanyak adalah *Obesity Type I* dengan 351 individu, diikuti oleh *Obesity Type III* sebanyak 324 individu dan *Obesity Type II* sebanyak 297 individu. Sementara itu, kategori *Overweight Grade I* dan *Overweight Grade II* masing-masing tercatat memiliki 290 kasus. Adapun kategori *Normal Weight* memiliki 287 individu, *sedangkan Insufficient Weight* merupakan kelas dengan jumlah paling sedikit, yaitu 272 individu. Informasi mengenai distribusi kelas ini penting untuk memastikan proses evaluasi model berjalan secara adil pada setiap kategori obesitas serta menghindari kecenderungan

model yang bias terhadap kelas tertentu. Selain itu, distribusi ini memberikan gambaran awal mengenai karakteristik populasi dalam *dataset*, sehingga memperkuat interpretasi hasil pada tahap analisis dan pengembangan model.

Tabel 4.2. Distribusi Variabel Target Obesitas

<i>Obesity Type</i>	<i>Number of Cases</i>
<i>Obesity Type I</i>	351
<i>Obesity Type II</i>	297
<i>Obesity Type III</i>	324
<i>Overweight Level I</i>	290
<i>Overweight Level II</i>	290
<i>Normal Weight</i>	287
<i>Insufficient Weight</i>	272

4.2. Hasil Penelitian

4.2.1. Pengujian Algoritma *Machine Learning*

Dalam penelitian ini, evaluasi performa berbagai algoritma *machine learning* dihitung dengan menggunakan metrik akurasi sebagai indikator utama untuk menilai kemampuan model dalam mengklasifikasikan tingkat obesitas secara tepat. Berdasarkan hasil pengujian, algoritma RF menghasilkan performa paling unggul dengan akurasi mencapai 94%. Tingginya akurasi ini menggambarkan bahwa RF dapat mengidentifikasi pola-pola kompleks *dataset* lebih unggul dibandingkan algoritma lainnya. Keunggulan RF terutama berasal dari mekanisme *ensemble learning* yang memadukan banyak *decision tree* sehingga mampu mengurangi variansi model dan menghasilkan prediksi yang lebih stabil dan akurat.

Di urutan berikutnya, algoritma *Decision Tree* (DT) berhasil memperoleh akurasi sebesar 92%, yang menunjukkan bahwa model ini juga cukup efektif dalam menangani data klasifikasi obesitas. Namun, performanya tetap berada sedikit di bawah RF karena DT lebih rentan terhadap *overfitting* ketika berhadapan dengan data berdimensi tinggi. Selanjutnya, algoritma k-NN mencatat akurasi 84%, yang relatif baik, meskipun performanya cukup dipengaruhi oleh distribusi data dan sensitivitas terhadap pemilihan nilai k . *Logistic Regression* (LR) menunjukkan performa menengah dengan akurasi 78%, menandakan bahwa model linear ini mampu menangkap sebagian pola dalam data, tetapi kurang ideal untuk hubungan *non-linear* yang lebih kompleks.

Sementara itu, algoritma *Naive Bayes* (NB) hanya mencapai akurasi 65%, yang mengindikasikan keterbatasannya dalam menangani variabel yang saling berkorelasi, karena NB bekerja berdasarkan asumsi independensi antar fitur. Algoritma SVM menjadi model dengan performa terendah, yaitu 62%, kemungkinan karena data bersifat multi-kelas dengan struktur *non-linear* sehingga menyulitkan SVM untuk menemukan *hyperplane* pemisah yang optimal.

Secara keseluruhan, hasil kinerja seluruh algoritma dirangkum dalam Tabel 4.3. Dari hasil tersebut dapat disimpulkan bahwa RF secara konsisten mengungguli algoritma lain berkat sifat *ensemble-based* yang mampu meningkatkan kemampuan generalisasi model. Di sisi lain, performa rendah pada NB, k-NN, dan terutama SVM dapat disebabkan oleh sensitivitas model terhadap ketidakseimbangan kelas, keberadaan fitur yang tidak relevan, serta kompleksitas pola hubungan antar variabel. Temuan ini menekankan bahwa ketepatan pemilihan algoritma sangat dipengaruhi oleh karakteristik *dataset*, struktur fitur, serta tingkat kompleksitas masalah klasifikasi yang ditangani.

Tabel 4.3. Hasil Evaluasi Semua Algoritma Klasifikasi

<i>Algorithms</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>Standard Deviation</i>
<i>Random Forest</i>	0,94	0,96	0,96	0,96	0,063466
<i>Naive Bayes</i>	0,65	0,63	0,65	0,63	0,046410
<i>k-Nearest Neighbors</i>	0,84	0,86	0,86	0,86	0,034830
<i>Decision Tree</i>	0,92	0,93	0,93	0,93	0,033861
<i>Support Vector Machine</i>	0,62	0,65	0,64	0,62	0,046250
<i>Logistic Regression</i>	0,78	0,80	0,80	0,80	0,033922

4.2.2. Feature Importance Analysis (Analisis Kepentingan Fitur)

Analisis pentingnya fitur dilakukan untuk memahami variabel mana yang memberikan kontribusi paling besar dalam proses prediksi tingkat obesitas. Pada penelitian ini, *dataset* yang digunakan merupakan data pemeriksaan kesehatan rutin yang dikumpulkan pada tahun 2019, dengan total 2.112 individu dan 17 variabel yang tercatat. Variabel-variabel tersebut mencakup karakteristik fisiologis, perilaku konsumsi makanan, serta kebiasaan aktivitas fisik. Karena data kesehatan bersifat dinamis dan dapat berubah mengikuti tren populasi tertentu, karakteristik *dataset*

termasuk jumlah instansi dan faktor-faktor yang direkam berpotensi mempengaruhi kinerja model *machine learning* dalam mengidentifikasi pola obesitas di masa mendatang.

Dalam penelitian ini, variabel “NObesidad” ditetapkan sebagai target fitur atau variabel yang diprediksi. Untuk meningkatkan kinerja model, dilakukan proses seleksi fitur menggunakan metode *Sparse Feature Selection and Ranking* (SpFSR). Hasil proses seleksi fitur menunjukkan bahwa pengurangan jumlah variabel dari 16 menjadi 10 fitur memberikan peningkatan kinerja model yang cukup signifikan, terutama pada algoritma *Random Forest*. Sebagaimana ditunjukkan pada Tabel 4.4, model *Random Forest* dengan 10 fitur berhasil mencapai akurasi sebesar 96.3%, lebih tinggi dibandingkan akurasi model dengan keseluruhan fitur (95.7%) [85]. Bahkan, model dengan 9 dan 8 fitur menunjukkan performa yang hampir setara, yang menegaskan bahwa pengurangan dimensi fitur tidak menurunkan kualitas prediksi secara substantif. Peningkatan ini menunjukkan bahwa proses seleksi fitur tidak hanya membantu meningkatkan performa model, tetapi juga meningkatkan interpretabilitas dengan mengurangi fitur-fitur yang redundan. SpFSR berhasil mengekstraksi fitur-fitur yang paling relevan dalam mempengaruhi tingkat obesitas, seperti riwayat obesitas dalam keluarga, tinggi badan, dan frekuensi konsumsi makanan utama, yang semuanya memiliki keterkaitan logis secara klinis maupun perilaku.

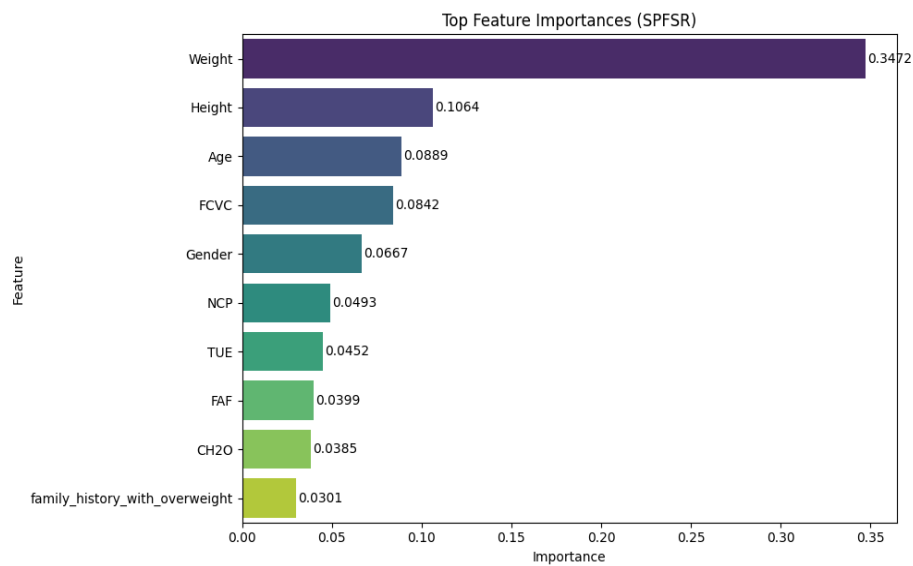
Di sisi lain, analisis tingkat kepentingan fitur dari algoritma *Random Forest* yang divisualisasikan dalam Gambar 4.1 menunjukkan bahwa berat badan merupakan prediktor paling dominan dengan kontribusi sebesar 34.72%. Disusul oleh tinggi badan (10.64%) dan usia (8.89%). Menariknya, meskipun berat badan muncul sebagai fitur paling penting dalam grafik kontribusi prediksi, fitur ini tidak terpilih sebagai fitur utama oleh metode SpFSR. Hal ini kemungkinan disebabkan oleh adanya korelasi yang sangat erat terkait berat dan tinggi yang membentuk indeks BMI, sehingga informasi tersebut dianggap redundan oleh proses seleksi fitur. Pada model tertentu, redundansi membuat suatu fitur dikesampingkan untuk meminimalkan kompleksitas tanpa mengurangi akurasi prediksi.

Secara keseluruhan, hasil analisis ini menunjukkan bahwa pengurangan fitur mampu meningkatkan efisiensi komputasi sekaligus mempertahankan tingkat prediksi yang tinggi. Temuan ini menegaskan bahwa proses seleksi fitur merupakan tahap

penting dalam pengembangan model pembelajaran mesin untuk memprediksi tingkat obesitas, terutama ketika data memiliki korelasi antar-fitur yang tinggi.

Tabel 4.4. Hasil Seleksi Fitur

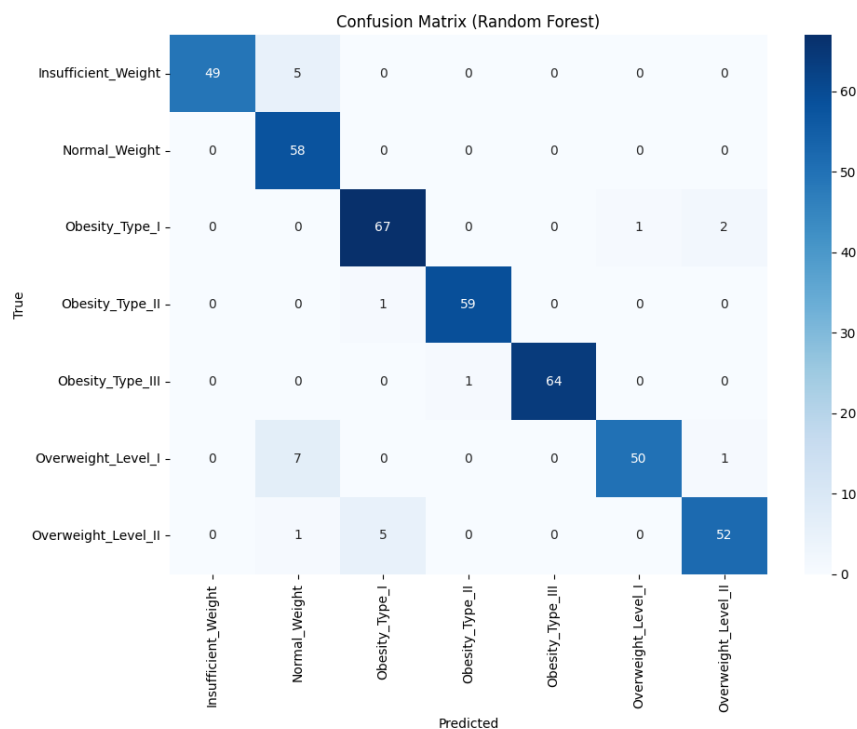
Metode	Feature	Hasil
<i>Simultaneous Perturbation Feature Selection and Ranking</i> (SpFSR)	16	0.957
	10	0.963
	9	0.962
	8	0.957



Gambar 4.1. SpFSR Feature Importance

Hasil analisis tingkat kepentingan fitur menggunakan metode *Sparse Feature Selection and Ranking* (SpFSR) menunjukkan bahwa beberapa variabel tidak memberikan kontribusi unik terhadap proses prediksi, sehingga dapat dieliminasi tanpa menurunkan performa model. Salah satu temuan penting adalah tidaknya terpilihnya variabel “Berat” dalam 10 fitur terbaik. “Berat” sebenarnya memiliki kontribusi prediksi yang sangat tinggi pada model *Random Forest*, yakni sebesar 34,72%, sebagaimana diperlihatkan pada Gambar 4.1. Namun, variabel ini tidak dipilih oleh SpFSR karena memiliki korelasi yang sangat kuat dengan variabel “Tinggi” dan karakteristik antropometrik lainnya. Dengan kata lain, “Berat” dianggap sebagai fitur yang redundan atau berlebihan, sehingga penghapusannya sejalan dengan tujuan SpFSR untuk mengurangi dimensi data dan mempertahankan hanya prediktor yang benar-benar unik.

Temuan ini memperlihatkan bagaimana teknik seleksi fitur modern dapat meningkatkan efisiensi model sekaligus mempertahankan interpretabilitas, terutama dalam konteks aplikasi medis dan epidemiologi obesitas. Hasil tingkat kepentingan fitur pada Gambar 4.1 juga menunjukkan kontribusi beberapa variabel lain yang berpengaruh terhadap status obesitas, antara lain: *family_history_with_overweight* (3%), *CH2O* atau konsumsi air harian (3,85%), *FAF* atau frekuensi aktivitas fisik (3,99%), *TUE* atau waktu penggunaan teknologi (4,52%), *NCP* atau jumlah konsumsi makanan utama per hari (4,93%), *Gender* (6,67%), *FCVC* atau frekuensi konsumsi sayur (8,42%), *Age* (8,89%), serta *Height* (10,64%). Variabel-variabel tersebut, bersama dengan pola makan, aktivitas fisik, kebiasaan transportasi, dan perilaku kesehatan lainnya, berkontribusi secara signifikan dalam membentuk prediksi tingkat obesitas individu.



Gambar 4.2. Representasi *Confusion Matrix* per Kelas dari Kinerja Model *Random Forest* dalam Klasifikasi Obesitas Multi-Kelas

Selanjutnya, kinerja model dievaluasi menggunakan *confusion matrix* yang ditampilkan pada Gambar 4.2. Hal ini memberikan gambaran menyeluruh mengenai kemampuan model *Random Forest* dalam membedakan tujuh kategori obesitas. Secara umum, model menghasilkan performa sangat unggul pada kelas-kelas dengan karakteristik yang jelas, seperti *Obesity_Type_I* dengan 67 prediksi benar dari 70

kasus, serta kategori *Normal_Weight* yang berhasil diprediksi dengan benar untuk seluruh 58 kasus. Tingginya akurasi pada kedua kategori ini menunjukkan bahwa model mampu menangkap pola antropometrik dan perilaku kesehatan yang khas pada individu dengan kondisi tersebut.

Namun demikian, beberapa pola kesalahan diperhatikan pada kelas yang memiliki karakteristik berdekatan. Misalnya, terdapat tujuh kasus *Overweight_Level_I* yang salah diklasifikasikan sebagai *Normal_Weight*. Kesalahan ini disebabkan oleh adanya irisan atau *overlap* pada fitur-fitur prediktif seperti pola makan, aktivitas fisik, maupun indikator antropometri yang hampir serupa antara kedua kelompok tersebut. Fenomena ini juga sering ditemukan pada penelitian serupa, di mana model *machine learning* cenderung mengalami kesulitan saat membedakan kategori yang berada pada batas transisi (*borderline classes*).

Secara keseluruhan, temuan ini menekankan pada kombinasi SpFSR dan *Random Forest* yang tidak hanya mampu meningkatkan performa prediksi melalui reduksi fitur yang optimal, tetapi juga memberikan pemahaman lebih dalam mengenai variabel-variabel penting yang memengaruhi status obesitas. Analisis *confusion matrix* serta *feature importance* memperkuat interpretasi bahwa pendekatan ini efektif untuk mendukung analisis epidemiologis maupun implementasi sistem pendukung keputusan di bidang kesehatan.

4.3. Analisis Kesalahan Prediksi Model

Berdasarkan confusion matrix pada Gambar 4.2, dapat dilihat bahwa sebagian besar data berhasil diklasifikasikan dengan benar oleh model *Random Forest*. Namun demikian, masih terdapat beberapa kasus kesalahan prediksi pada beberapa kelas tingkat obesitas.

Kesalahan prediksi umumnya terjadi pada kelas yang memiliki karakteristik yang relatif mirip, terutama pada kategori *Normal Weight*, *Overweight Level I*, dan *Overweight Level II*. Pada beberapa kasus, data yang sebenarnya termasuk dalam kategori *Overweight Level I* diprediksi sebagai *Normal Weight*. Hal ini dapat terjadi karena nilai pada beberapa variabel seperti pola konsumsi makanan, aktivitas fisik, dan

kebiasaan hidup tidak menunjukkan perbedaan yang terlalu signifikan antara kedua kategori tersebut.

Selain itu, kesalahan klasifikasi juga dapat dipengaruhi oleh kemiripan distribusi data antar kelas. Individu yang berada pada batas antara dua kategori tingkat obesitas sering memiliki karakteristik yang hampir sama sehingga model machine learning mengalami kesulitan dalam membedakan kelas tersebut secara tepat.

Faktor lain yang dapat memengaruhi kesalahan prediksi adalah proses seleksi fitur menggunakan metode SpFSR. Metode ini bertujuan untuk memilih fitur yang paling relevan dan mengurangi dimensi data. Namun, dalam beberapa kasus, pengurangan jumlah fitur dapat menyebabkan sebagian informasi yang berpotensi membantu proses klasifikasi tidak sepenuhnya digunakan oleh model.

Meskipun masih terdapat beberapa kesalahan klasifikasi, jumlah prediksi yang benar masih jauh lebih besar dibandingkan dengan prediksi yang salah. Hal ini menunjukkan bahwa model *Random Forest* tetap memiliki performa yang baik dalam mengidentifikasi pola data dan memprediksi tingkat obesitas pada dataset yang digunakan dalam penelitian ini.

4.4. Analisis Komparatif dengan Penelitian Terdahulu

Beberapa penelitian sebelumnya menunjukkan bahwa obesitas merupakan kondisi multifaktorial yang dipengaruhi oleh faktor genetik maupun lingkungan. Lowe *et al.* [76] menemukan bahwa predisposisi keluarga merupakan faktor utama dalam timbulnya obesitas pada anak-anak serta berdampak pada kesehatan jangka panjang. Riwayat keluarga dengan kelebihan berat badan atau penyakit kardiometabolik juga terbukti terkait dengan tingkat keparahan obesitas pada anak-anak [77]. Penelitian oleh Eisenberg *et al.* [78] menekankan bahwa intervensi berbasis keluarga dalam manajemen berat badan dapat memberikan manfaat signifikan, menunjukkan pentingnya mempertimbangkan faktor keluarga dalam penanganan obesitas. Selain itu, Shammala *et al.* [79] menemukan bahwa variabel demografis seperti lokasi tempat tinggal, jenis kelamin, dan tipe obesitas memengaruhi prevalensi obesitas pada mahasiswa, menunjukkan bahwa faktor lingkungan dan gaya hidup turut berperan.

Dalam penelitian ini, algoritma *Random Forest* (RF) menunjukkan akurasi prediksi tertinggi dibandingkan model *machine learning* lain, yaitu sebesar 94%. Jika

dibandingkan dengan studi Jindal *et al.* [29] yang melaporkan akurasi RF sebesar 89,68%, peningkatan akurasi pada penelitian ini kemungkinan disebabkan oleh penerapan SpFSR (*Stochastic Permutation Feature Selection and Ranking*) sebagai metode seleksi fitur. Selain meningkatkan akurasi, pendekatan ini juga memberikan peningkatan pada efisiensi komputasi.

Analisis pentingnya fitur (*feature importance*) dalam model RF merupakan metode yang kuat untuk mengidentifikasi prediktor utama. Namun, perlu diperhatikan potensi bias pada atribut yang memiliki banyak nilai [80]. Teknik SpFSR secara stokastik memberi peringkat dan seleksi fitur, sehingga fitur paling signifikan dapat dimasukkan ke dalam model [81]. Dengan demikian, temuan penelitian ini sejalan dengan literatur sebelumnya mengenai pengaruh faktor genetik, demografis, dan gaya hidup terhadap obesitas, sekaligus menunjukkan bahwa integrasi metode seleksi fitur canggih dapat meningkatkan performa model *machine learning*.

4.5. Implikasi bagi Kesehatan Masyarakat dan Penelitian Selanjutnya

Integrasi RF dengan SpFSR dalam penelitian ini memberikan pemahaman penting untuk deteksi dini dan pencegahan obesitas. Dengan menyoroti variabel paling berpengaruh, metode SpFSR meningkatkan interpretabilitas model, sehingga dapat menjadi alat bantu bagi tenaga kesehatan dan pengambil kebijakan dalam menyusun intervensi yang lebih baik.

Keterbatasan utama penelitian ini adalah penggunaan satu *dataset* tunggal dari ML *Repository*, yang dapat membatasi generalisasi model pada populasi lain. Meskipun penelitian ini menggunakan *train-test split* terstratifikasi dan *cross-validation 5-fold* untuk mengurangi *overfitting*, penelitian selanjutnya disarankan melakukan validasi pada dataset eksternal, seperti NHANES atau RISKESDAS, agar hasil lebih dapat diterapkan secara luas. Selain itu, penerapan teknik *domain adaptation* atau *transfer learning* dapat meningkatkan adaptabilitas model dalam konteks kesehatan masyarakat nyata. Perkembangan metode *machine learning* yang berkelanjutan dan akses ke *dataset* yang lebih representatif sangat penting untuk meningkatkan akurasi prediksi dan mendukung intervensi kesehatan berbasis data.