

BAB III

HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Berdasarkan hasil penelitian yang telah dilakukan, tujuan utama adalah untuk mengevaluasi kinerja algoritma Random Forest dengan tambahan metode *SMOTE* dan *ANOVA* dalam klasifikasi penyakit Stunting di Kota Samarinda, dengan menggunakan metrik utama seperti akurasi. Akurasi memberikan wawasan mengenai efektivitas algoritma dalam mengklasifikasi Stunting. Dataset yang didapat memiliki 23 kolom dengan 22 atribut dan 1 atribut sebagai label (TB/U).

3.1.1 Data Selection

Pada Tabel 3.1 merupakan atribut yang telah dipilih secara manual terdapat 13 Atribut yang terpilih dan 1 atribut TB/U sebagai target atau kelas

Tabel 3.1 Data Selection

	Nama	JK	Berat	Tinggi	LiL A	BB/U	ZS BB/U	Tanggal Pengukuran	ZS TB/ U	BB/TB	ZS BB/T B	Naik Berat Badan	Jml Vit A	TB/U
1	DIMAS ADITYA	L	9.01		0	Kurang	-0.39	2023-01-02	-0.21	Gizi Baik	-0.39	O		Normal
2	SITI AISYAH	P	12	94	0		-2.25	2023-01-02	-2.09	Gizi Baik	-1.46	O		Pendek
3	M AL FATIH ALMAHIRA	L	8.01	69	0	Berat Badan Normal	-0.53	2023-01-02	-0.65	Gizi Baik	-0.14	O		Normal
4	AKIRA AKBAR GIUNIA	P	6.03		0	Berat Badan Normal	-0.31 0.09097	2023-01-02	0.42	Gizi Baik	-0.74	O		Normal
5	QAMELA	P	10.06		0	Risiko Lebih	2222	2023-01-02	0.23	Gizi Lebih	0.09	O		Normal
150462	ADNAN IRAGUSTI	L	3	50		Berat Badan Normal	-0.73	2023-12-30	0.06	Gizi Baik	-1.19	-		Normal
150463	SIENA AL RAISHA	P	13		0	Berat Badan Normal	-1.22	2023-12-13	-1.11	Gizi Baik	-0.85	O		Normal
150464	AFIZAH KHAIRINA	P	2.05	45		Kurang	-2.03	2023-12-09	-2.63	Gizi Baik	-0.35	-		Normal
150465	M ARSYA KHOLIF	P	3	49		Berat Badan Normal	-1.56	2023-12-19	-1.3	Gizi Baik	-1.04	-		Pendek
150466	MUHAMMAD IQBAL	L	2.09	49		Kurang	-2.97	2023-12-29	-2.48	Gizi Baik	-1.37	-		Normal

3.1.2 Data Cleaning

Pada Tabel 3.2 Merupakan tampilan dataset sebelum dilakukan pembersihan data untuk menghapus baris yang memiliki nilai hilang dan menghapus baris yang terduplikat pada dataset stunting dengan jumlah 150466 record. Proses ini memastikan data bersih dan siap untuk analisis lebih lanjut.

Tabel 3.2 Dataset sebelum dibersihkan

	Nama	JK	Berat	Tinggi	LiL A	BB/U	ZS BB/U	Tanggal Pengukuran	ZS TB/ U	BB/TB	ZS BB/T B	Naik Berat Badan	Jml Vit A	TB/U
1	DIMAS ADITYA	L	9.01		0	Kurang	-0.39	2023-01-02	-0.21	Gizi Baik	-0.39	O		Normal
2	SITI AISYAH	P	12	94	0		-2.25	2023-01-02	-2.09	Gizi Baik	-1.46	O		Pendek
3	M AL FATIH ALMAHIRA	L	8.01	69	0	Berat Badan Normal	-0.53	2023-01-02	-0.65	Gizi Baik	-0.14	O		Normal
4	AKIRA AKBAR GIUNIA	P	6.03		0	Berat Badan Normal	-0.31 0.09097	2023-01-02	0.42	Gizi Baik	-0.74	O		Normal
5	QAMELA	P	10.06		0	Risiko Lebih	2222	2023-01-02	0.23	Gizi Lebih	0.09	O		Normal
150462	ADNAN IRAGUSTI	L	3	50		Berat Badan Normal	-0.73	2023-12-30	0.06	Gizi Baik	-1.19	-		Normal
150463	SIENA AL RAISHA	P	13		0	Berat Badan Normal	-1.22	2023-12-13	-1.11	Gizi Baik	-0.85	O		Normal
150464	AFIZAH KHAIRINA	P	2.05	45		Kurang	-2.03	2023-12-09	-2.63	Gizi Baik	-0.35	-		Normal
150465	M ARSYA KHOLIF	P	3	49		Berat Badan Normal	-1.56	2023-12-19	-1.3	Gizi Baik	-1.04	-		Pendek
150466	MUHAMMAD IQBAL	L	2.09	49		Kurang	-2.97	2023-12-29	-2.48	Gizi Baik	-1.37	-		Normal

Pada tampilan data sebelum dilakukan proses pembersihan terdapat banyak data yang memiliki kesamaan nama dikarenakan pemeriksaan stunting selama periode tahun 2023 terjadi beberapa kali namun untuk penelitian ini hanya akan menggunakan tanggal pemeriksaan terbaru sehingga perlu untuk melakukan penghapusan data untuk memilih tanggal pemeriksaan terbaru setelah proses penghapusan data yang terduplikasi dilanjutkan dengan penghapusan kolom yang memuat nilai N/A (no value is available) atau data yang tidak memiliki nilai.

```
Jumlah missing values untuk setiap kolom:
Nama          6
JK            0
Berat        558
Tinggi       51222
LiLA         29075
BB/U         0
ZS BB/U     159
TB/U        12871
ZS TB/U     12264
BB/TB       12812
ZS BB/TB   12200
Naik Berat Badan  57
Jml Vit A   95334
Tanggal Pengukuran 0
dtype: int64
```

Gambar 3. 1 Jumlah Nilai Kosong Tiap Kolom Sebelum Pembersihan

Tabel 3. 3 Dataset setelah dibersihkan

	Nama	JK	Berat	Tinggi	LiLA	BB/U	ZS BB/U	Tanggal Pengukuran	ZS TB/U	BB/TB	ZS BB/T B	Naik Berat Badan	Jml Vit A	TB/U
1	A FARIS WICAKSONO	L	9.07	78.0	16	Berat Badan Normal	-1.75	2023-10-06	-2.84	Gizi Baik	-0.47	O	1	Pendek
2	A FATHAN	L	15.00	107.0	17	Berat Badan Normal	-1.08	2023-10-23	0.14	Gizi Baik	-1.83	T	1	Normal
3	A FAUJAN	L	14.00	100.0	0	Berat Badan Normal	-0.85	2023-02-07	-0.16	Gizi Baik	-1.14	O	1	Normal
4	A MISHAEL	P	15.00	103.0	0	Berat Badan Normal	0.06	2023-02-09	1.05	Gizi Baik	-0.80	O	1	Normal
5	A ZHAFIR KAMIL	L	14.00	102.0	0	Berat Badan Normal	-1.91	2023-10-14	-1.49	Gizi Baik	-1.59	O	1	Normal
...
8055	yelmi	L	17.04	111.0	0	Berat Badan Normal	-0.04	2023-02-27	0.06	Gizi Baik	-0.93	N	1	Normal
8056	yumna	P	13.07	100.0	14	Berat Badan Normal	-1.07	2023-10-14	-0.45	Gizi Baik	-1.18	O	1	Normal
8057	ziaan aqilla kamo	L	14.05	102.0	0	Berat Badan Normal	-1.71	2023-10-21	-1.62	Gizi Baik	-1.15	T	1	Normal
8058	zubair	L	17.03	107.0	0	Berat Badan Normal	-0.14	2023-02-20	-0.13	Gizi Baik	-0.13	O	1	Normal
8059	zulkifii abdi	L	15.06	102.0	0.0	Berat Badan Normal	-0.59	2023-10-03	-0.69	Gizi Baik	-0.25	N	1	Normal

Setelah dilakukan proses pembersihan dapat terlihat pada tabel 3.3 jumlah data yang memiliki nilai kosong sudah tidak ada dan data yang tersisa setelah proses pembersihan data berjumlah 8059 record .

```
Jumlah missing values untuk setiap kolom:
Nama          0
JK            0
Berat         0
Tinggi        0
LiLA          0
BB/U          0
ZS BB/U       0
TB/U          0
ZS TB/U       0
BB/TB         0
ZS BB/TB     0
Naik Berat Badan  0
Jml Vit A     0
Tanggal Pengukuran 0
dtype: int64
```

Gambar 3. 2 Jumlah Nilai Kosong Tiap Kolom Setelah Pembersihan

3.1.3 Data Transformation

Beberapa atribut yang ditransformasi meliputi jenis kelamin, naik berat badan, berat badan menurut umur, dan berat badan menurut tinggi badan. Proses ini melibatkan teknik encoding untuk mengubah data kategori menjadi numerik dan penggantian nilai pada kolom label TB/U untuk mengasumsikan kelas menjadi dua kategori. Dengan transformasi ini, data menjadi lebih siap untuk digunakan dalam analisis dan pemodelan.

Tabel 3. 4 Dataset sebelum ditransformasi

	JK	BB/U	BB/TB	Naik Berat Badan	TB/U
1	L	Berat Badan Normal	Gizi Baik	O	Pendek
2	L	Berat Badan Normal	Gizi Baik	T	Normal
3	L	Berat Badan Normal	Gizi Baik	O	Normal
4	P	Berat Badan Normal	Gizi Baik	O	Normal
5	L	Berat Badan Normal	Gizi Baik	O	Normal
...
8055	L	Berat Badan Normal	Gizi Baik	N	Normal
8056	P	Berat Badan Normal	Gizi Baik	O	Normal
8057	L	Berat Badan Normal	Gizi Baik	T	Normal
8058	L	Berat Badan Normal	Gizi Baik	O	Normal
8059	L	Berat Badan Normal	Gizi Baik	N	Normal

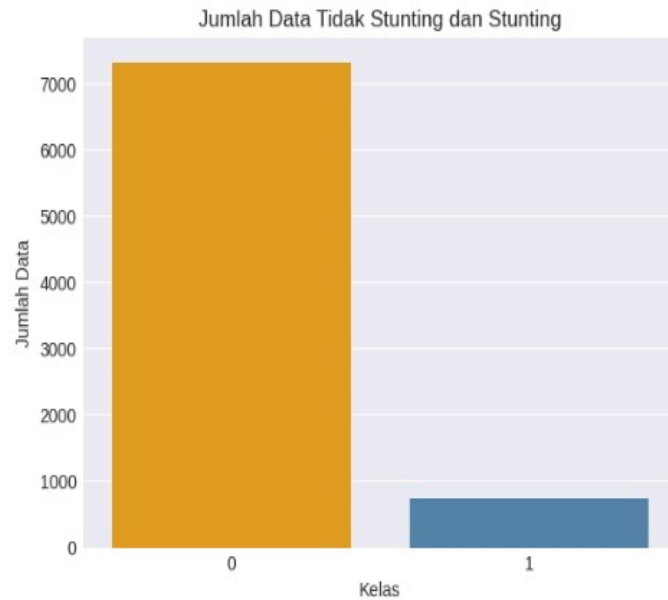
Tabel 3. 5 Dataset Setelah ditransformasi

	JK	BB/U	BB/T B	Naik Berat Badan	TB/U
1	0	0	0	2	1
2	0	0	0	3	0
3	0	0	0	2	0
4	1	0	0	2	0
5	0	0	0	2	0
...
8055	0	0	0	4	0
8056	1	0	0	2	0
8057	0	0	0	3	0
8058	0	0	0	2	0
8059	0	0	0	4	0

Pada tabel 3.5 tampilan data pada kolom atribut 'JK', 'BB/U', 'BB/TB', 'Naik Berat Badan' dan label 'TB/U' setelah dilakukan transformasi data dimana data yang sebelumnya berupa String di ubah menjadi Integer untuk memudahkan proses klasifikasi.

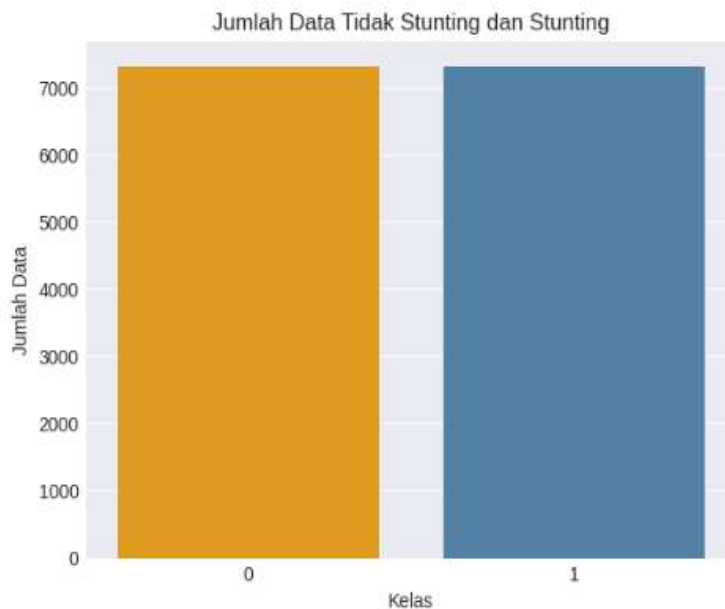
3.1.4 Data Balancing

Proses penyeimbangan data dilakukan dengan menggunakan teknik oversampling. Teknik ini bertujuan untuk menyamakan jumlah sampel antara kelas minoritas dan kelas mayoritas dalam dataset.



Gambar 3. 3 Jumlah Kelas Sebelum Data *Balancing*

Pada Gambar 3.3 terdapat perbedaan jumlah kelas dimana kategori tidak stunting (0) berjumlah 7317 data dan kategori stunting (1) berjumlah 742 data.



Gambar 3. 4 Jumlah Kelas Sesudah Data *Balancing*

Pada gambar 3.4 menunjukkan perbandingan jumlah kelas yang sudah seimbang antara kelas mayoritas dan kelas minoritas setelah proses *SMOTE* dimana jumlah data kategori tidak stunting (0) berjumlah 7317 data dan kategori stunting (1) berjumlah 7317 data.

3.1.5 Permodelan *Random forest* Tanpa *ANOVA*

Pada tahap ini, akan menjalankan eksperimen pertama dengan melakukan pengujian model menggunakan *Python*. Pengujian ini melibatkan implementasi algoritma *Random Forest* tanpa penggunaan teknik *Analysis of variances* (*ANOVA*). Tujuan dari pengujian ini adalah untuk menilai kinerja dasar model *Random Forest* dalam mengklasifikasikan data.

Hasil pengujian pada tabel 3.6 menunjukkan bahwa model *Random Forest* tanpa seleksi fitur *ANOVA* pada setiap fold menunjukkan kinerja yang sangat baik dalam memprediksi kelas pada dataset Stunting Kota Samarinda. Dalam kesepuluh *fold* yang dievaluasi, model ini mencapai akurasi yang konsisten tinggi, dengan nilai rata-rata mencapai 98.83%. Tingkat akurasi yang tinggi ini menunjukkan kemampuan model untuk secara tepat mengklasifikasikan data, dengan tingkat kesalahan yang rendah.

Tabel 3. 6 Hasil Pengujian model *Random Forest*

<i>Fold</i>	Akurasi
1	97.95%
2	98.84%
3	98.91%
4	98.91%
5	98.56%
6	99.18%
7	99.38%
8	98.91%
9	99.18%
10	98.50%

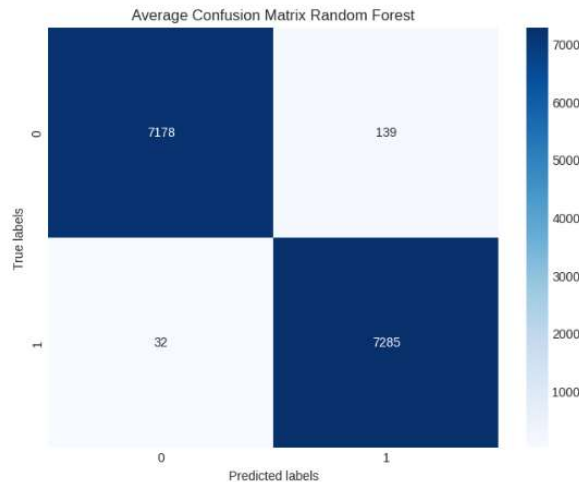
Tabel 3. 7 Rata Rata Akurasi Pengujian Model *Random Forest*

Hasil Rata Rata	
Akurasi	98.83%

Dari evaluasi model *Random Forest* tanpa *ANOVA*, menampilkan hasil yang cukup baik, dimana mendapatkan evaluasi di atas 90% pada tiap parameter. Hasil ini menunjukkan bahwa model *Random Forest* yang dikembangkan mampu melakukan klasifikasi dengan tingkat keakuratan yang sangat tinggi. Hal ini memberikan indikasi kuat bahwa fitur-fitur yang digunakan dalam model telah efektif dalam membedakan berbagai kelas dengan hasil yang diinginkan. Meskipun tanpa penerapan teknik *ANOVA* sebagai seleksi fitur, untuk membandingkan hasil evaluasi tersebut maka dilakukan perhitungan manual nilai *confusion matrix* untuk mengkonfirmasi keakuratan hasil evaluasi tersebut.

Tabel 3. 8 *Confusion Matrix*

	<i>Predicted Positive</i> (0)	<i>Predicted Negative</i> (1)
<i>Actual Positive</i> (0)	7178	139
<i>Actual Negative</i> (1)	32	7285



Gambar 3. 5 *Confusion Matrix Random Forest*

$$Accuracy = \frac{7178 + 7285}{7178 + 7285 + 139 + 32} = 0.9883 = 98.83\%$$

Setelah dilakukan proses perhitungan manual pada tiap parameter maka didapatkan hasil yang sama dengan perhitungan menggunakan *python*, hasil ini menunjukkan kesesuaian yang sangat baik dengan perhitungan manual, hal ini menunjukkan bahwa model memiliki kemampuan yang baik dalam menerapkan data latih ke data uji dan mampu mengklasifikasikan data dengan akurat.

3.1.6 Permodelan *Random Forest* Dengan *ANOVA*

Pada tahap ini, akan menjalankan eksperimen kedua dengan melakukan pengujian model menggunakan *Python*. Pengujian ini melibatkan implementasi algoritma *Random Forest* dengan tambahan penggunaan teknik *Analysis of variances* (*ANOVA*). Tujuan dari pengujian ini adalah untuk menilai kinerja model *Random Forest* setelah penerapan *ANOVA* dalam mengklasifikasikan data.

a) Seleksi Fitur

Seleksi fitur dilakukan untuk merangsang atribut yang ada, mulai dari atribut yang memiliki pengaruh besar (mendapatkan nilai F tertinggi) terhadap hasil klasifikasi hingga atribut yang memiliki pengaruh kecil (mendapatkan nilai F terendah). Dengan melakukan seleksi fitur, maka dapat dipilih subset dari atribut-atribut tersebut yang paling berpengaruh dalam memprediksi atau mengklasifikasikan target. Seleksi fitur membantu mengurangi dimensi data, mempercepat proses pembelajaran, dan mencegah overfitting (Bengnga and Ishak, 2022). Proses seleksi fitur sangat penting dalam dataset stunting di Kota Samarinda, di mana terdapat beragam atribut yang mungkin memiliki kontribusi berbeda terhadap status stunting. Dengan mengidentifikasi atribut yang paling relevan, maka dapat meningkatkan akurasi model prediksi, mengurangi waktu komputasi, dan meningkatkan interpretabilitas model. Selain itu, seleksi fitur juga membantu dalam mengeliminasi atribut yang tidak relevan, sehingga model menjadi lebih efisien dan mudah dipahami. Dengan demikian, seleksi fitur tidak hanya meningkatkan kinerja model secara keseluruhan, tetapi juga memungkinkan penggunaan sumber daya yang lebih efektif dalam analisis data.

Tabel 3. 9 Hasil Perengkingan *ANOVA*

Atribut	Nilai F	Ranking
ZS TB/U	3091	1
ZS BB/U	975	2
BB/U	737	3
Tinggi	441	4
Berat	290	5
LiLA	32	6
ZS BB/TB	4	7
JK	8	8
BB/TB	4	9
Naik Berat Badan	3	10
Jml Vit A	NaN	11

Pada hasil perengkingan atribut yang dapat dilihat pada tabel 3.9 , maka ditentukan atribut yang akan digunakan adalah atribut dengan ranking 1 – 6 sebagai atribut dalam pemodelan *Random Forest* karena memiliki nilai F yang tinggi sehingga atribut pada pemodelan ini hanya 6 yaitu ZS TB/U, ZS BB/U, BB/U, Tinggi, Berat, LiLA Selain itu, kolom 'Jml Vit A' perlu dihapus dari dataset karena library *pandas* membacanya sebagai nilai NaN, yang tidak dapat diproses dalam pemodelan.

b) Pemodelan

Hasil pengujian pada tabel 3.10 menunjukkan bahwa model *Random Forest* dengan seleksi fitur *ANOVA* pada kesepuluh fold menunjukkan kinerja yang sangat baik dalam memprediksi kelas pada dataset Stunting Kota Samarinda. Model ini mencapai akurasi yang tinggi, dengan nilai rata-rata mencapai 99.78%. Tingkat akurasi yang tinggi ini menunjukkan kemampuan model untuk secara tepat mengklasifikasikan data, dengan tingkat kesalahan yang sangat rendah. Penggunaan seleksi fitur *ANOVA* dalam proses pemodelan telah membantu model dalam mengidentifikasi atribut-atribut yang paling berpengaruh dalam klasifikasi stunting, sehingga meningkatkan akurasi prediksinya.

Tabel 3. 10 Hasil Pengujian Model *Random Forest*

<i>Fold</i>	Akurasi
1	99.66%
2	99.80%
3	99.86%
4	99.93%
5	99.66%
6	99.80%
7	99.80%
8	99.52%
9	99.86%
10	99.86%

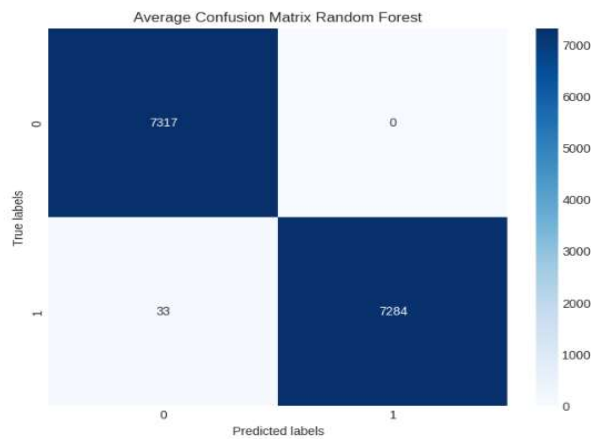
Tabel 3. 11 Rata Rata Akurasi pengujian model *Random Forest*

Hasil Rata Rata	
Akurasi	99.77%

Dari evaluasi model *Random Forest* dengan *ANOVA*, menunjukkan hasil yang cukup baik dilihat dari kenaikan akurasi setelah penerapan metode seleksi fitur. Hasil ini menunjukkan bahwa model *Random Forest* mampu memaksimalkan hasil yang didapat setelah penambahan metode tersebut. Untuk membandingkan hasil evaluasi tersebut maka dilakukan perhitungan manual nilai confusion matrix untuk mengkonfirmasi keakuratan hasil tersebut.

Tabel 3. 12 *Confusion Matrix*

	<i>Predicted Positive (0)</i>	<i>Predicted Negative (1)</i>
<i>Actual Positive (0)</i>	7317	0
<i>Actual Negative(1)</i>	33	7284



Gambar 3. 6 *Confusion Matrix Random Forest*

$$Accuracy = \frac{7284 + 7317}{7284 + 7317 + 0 + 33} = 0.9977 = 99.77\%$$

Setelah dilakukan proses perhitungan manual pada tiap parameter, didapatkan hasil yang sama dengan perhitungan menggunakan *Python*. Hasil ini menunjukkan kesesuaian yang sangat baik dengan perhitungan manual, mengindikasikan bahwa model memiliki kemampuan yang kuat dalam menerapkan data latih ke data uji dan mampu mengklasifikasikan data dengan akurat. Hal ini menegaskan validitas metode yang digunakan dalam penelitian ini, serta memberikan kepercayaan tambahan bahwa algoritma yang diterapkan dapat diandalkan. Kesesuaian antara hasil perhitungan manual dan hasil dari implementasi *Python* juga menunjukkan bahwa tidak ada kesalahan signifikan dalam proses pemrograman atau implementasi algoritma, serta pemilihan fitur telah dilakukan dengan tepat.

3.1.7 Perbandingan Hasil

Dalam melakukan perbandingan antara dua model *Random Forest*, di mana satu menggunakan seleksi fitur *ANOVA* dan yang lainnya tidak, fokus pada perbedaan akurasi memberikan pandangan yang penting dalam mengevaluasi efektivitas seleksi fitur dalam meningkatkan kinerja model. Hasil yang signifikan dapat mengindikasikan bahwa seleksi fitur mampu mengidentifikasi subset fitur yang lebih

relevan, sehingga memungkinkan model untuk mempelajari pola yang lebih penting dalam data. Hal ini dapat menghasilkan prediksi yang lebih akurat dan efisien.

Tabel 3. 13 Perbandingan Hasil Akurasi Pengujian *Random Forest*

<i>Fold</i>	<i>Random Forest</i> (Tanpa <i>ANOVA</i>)	<i>Random Forest</i> (Dengan <i>ANOVA</i>)	Status
1	97.95%	99.66%	Naik
2	98.84%	99.80%	Naik
3	98.91%	99.86%	Naik
4	98.91%	99.93%	Naik
5	98.56%	99.66%	Naik
6	99.18%	99.80%	Naik
7	99.38%	99.80%	Naik
8	98.91%	99.52%	Naik
9	99.18%	99.86%	Naik
10	98.50%	99.86%	Naik

Tabel 3. 14 Perbandingan Hasil Akurasi Rata-Rata *Random Forest*

<i>Random Forest</i> (Tanpa <i>ANOVA</i>)	<i>Random Forest</i> (Dengan <i>ANOVA</i>)
98.83%	99.77%

Berdasarkan tabel perbandingan akurasi antara model *Random Forest* (RF) dengan dan tanpa seleksi fitur *ANOVA* pada tabel 3.13, terlihat bahwa model dengan seleksi fitur *ANOVA* cenderung memberikan akurasi yang lebih tinggi dalam mengklasifikasi. Dari 10 lipatan (folds) pada *cross validation*, model RF dengan seleksi fitur *ANOVA* konsisten menunjukkan peningkatan akurasi dibandingkan dengan model RF tanpa seleksi fitur *ANOVA*. Peningkatan akurasi yang dicapai sebesar 0.94% menunjukkan bahwa seleksi fitur *ANOVA* efektif dalam meningkatkan kinerja model Random Forest.

3.2 Pembahasan

Penelitian ini menggunakan data Stunting Kota Samarinda periode tahun 2023, data yang didapat akan melalalu beberapa tahapan pengolahan mulai dari data *selection*, data *cleaning*, data *integration* dan data *balancing*. Tahapan tersebut dilakukan agar data yang akan digunakan dalam klasifikasi memiliki kualitas terbaik karena data yang diperoleh memiliki ketidakseimbangan kelas maka dilakukanlah teknik *oversampling* menggunakan *SMOTE* sedangkan untuk membagi datanya menjadi data latih dan data uji dilakukan dengan menggunakan *K-fold Cross Validation* dengan nilai K sebesar sepuluh.

- a) Seleksi fitur *ANOVA* pada dataset Stunting Kota Samarinda mengidentifikasi fitur ZS TB/U sebagai yang paling berpengaruh terhadap hasil klasifikasi dengan nilai F sebesar 3091, diikuti oleh ZS BB/U dengan nilai F 975, dan BB/U dengan nilai F 737. Fitur-fitur ini memiliki pengaruh signifikan dalam membedakan berbagai kelas dalam dataset stunting. Sementara itu, fitur-fitur seperti Tinggi, Berat, dan LiLA juga menunjukkan pengaruh yang cukup besar, menduduki peringkat 4 hingga 6. Penelitian oleh K. M. Rajabi dkk. (2023) yang menggunakan metode *Relief* juga menunjukkan

pentingnya fitur Tinggi, yang teridentifikasi sebagai salah satu dari dua fitur terbaik yang meningkatkan akurasi *k-nearest neighbor* (KNN) menjadi 98,16%. Selanjutnya, Yunus dkk (2023) menggunakan metode *Backward Elimination* (BE) berhasil mengidentifikasi 2 fitur terbaik, yaitu Tinggi dan Berat. Fitur-fitur ini berkontribusi meningkatkan akurasi model *Naive Bayes* (NB) dari 53,50% menjadi 92,54%. Dengan demikian, penggunaan seleksi fitur *ANOVA* dalam studi ini tidak hanya mengonfirmasi keefektifan atribut Tinggi dan Berat yang telah diakui dalam penelitian lain, tetapi juga meningkatkan performa model *Random Forest* dalam klasifikasi stunting.

Tabel 3. 15 Persamaan Hasil Seleksi Fitur Dengan Penelitian Lain

Hasil Seleksi Fitur <i>ANOVA</i> (<i>RF – ANOVA</i>)	Penelitian K. M. Rajabi dkk. (2023) (KNN – <i>Relief</i>)	Penelitian Yunus dkk. (2023) (NB – <i>BE</i>)
ZS TB/U		
ZS BB/U		
BB/U		
Tinggi	✓	✓
Berat		✓
LiLA		

- b) Penggunaan model *Random Forest* yang diintegrasikan dengan teknik oversampling *SMOTE* berhasil mencapai akurasi rata-rata 98.83%, tanpa menggunakan seleksi fitur. Teknik *SMOTE* diimplementasikan untuk mengatasi ketidakseimbangan kelas yang signifikan dalam dataset stunting, yang memungkinkan model untuk belajar dari distribusi kelas yang lebih seimbang. Kemudian, dengan menambahkan seleksi fitur menggunakan metode *ANOVA* terjadi peningkatan akurasi yang signifikan hingga mencapai 99.77%. Peningkatan sebesar 0.94% ini menunjukkan bahwa integrasi seleksi fitur *ANOVA* dapat efektif meningkatkan efisiensi dan efektivitas model *Random Forest* dalam klasifikasi stunting, begitupun pada penelitian terkait stunting yang dilakukan oleh Nugroho et al., (2022) yang menggabungkan *ANOVA* dengan algoritma KNN dan *Decision tree* dimana *ANOVA* terbukti mampu menaikkan akurasi sebesar 20% untuk KNN dan 39% untuk *Decision Tree* dan pada penelitian Mogarampalli et al., (2022) yang menggunakan algoritma klasifikasi seperti SVM, KNN, *Random Forest*, *Naive Bayes*, *logistic regression*, *Decision Tree* terbukti mampu meraih akurasi diatas 90% ketika dikombinasikan dengan seleksi fitur *ANOVA*. Seleksi fitur membantu model mengidentifikasi dan memprioritaskan variabel yang paling berpengaruh, sehingga meningkatkan akurasi klasifikasi meskipun data tetap diolah dengan teknik *SMOTE* untuk memastikan keseimbangan antar kelas.