



## Klasifikasi Status Stunting Pada Balita Menggunakan *K-Nearest Neighbor* Dengan *Feature Selection Backward Elimination*

Syahrani Lonang\*, Dwi Normawati

Program Studi Teknik Informatika, Fakultas Teknologi Industri, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Email: <sup>1</sup>lonangsyahrani3@email.com, <sup>2</sup>dwi.normawati@tif.uad.ac.id

Email Penulis Korespondensi: lonangsyahrani3@gmail.com

**Abstrak**—Permasalahan utama tentang gizi yang dihadapi Indonesia adalah stunting, dimana Indonesia menempati peringkat kelima dunia dengan tingkat prevalensi stunting terbesar pada tahun 2017 yaitu 29,6% dari seluruh anak Indonesia. Stunting adalah anak bayi dibawah lima tahun yang memiliki nilai *z-score* nya kurang dari -3 standar deviasi (SD). Stunting memiliki dampak negatif yaitu dapat terganggunya perkembangan fisik maupun intelektual balita di masa depan. Dalam kasus ini, pemeriksaan status stunting oleh tenaga medis masih dilaksanakan secara manual yang memakan waktu cukup lama dan rentan ketidaktelitian. Penelitian ini bertujuan untuk mengklasifikasi status stunting pada balita dengan menerapkan metode *K-Nearest Neighbor* menggunakan seleksi fitur *Backward Elimination* untuk mendapatkan hasil yang cepat dan akurat. Berdasarkan hasil dari penelitian ini menunjukkan rata-rata akurasi yang dihasilkan algoritma *K-Nearest Neighbor* pada  $k=5$  adalah 91,90% dengan 9 atribut dan rata-rata akurasi yang dihasilkan algoritma *K-Nearest Neighbor* dengan penambahan *Backward Elimination* sebesar 92,20% dengan 8 atribut. Hasil tersebut menunjukkan bahwa penerapan *Backward Elimination* mampu menaikkan nilai akurasi pada algoritma *K-Nearest Neighbor* dan juga melakukan seleksi atribut.

**Kata Kunci:** Klasifikasi; Stunting; *K-Nearest Neighbor*; *Backward Elimination*; Data Mining

**Abstract**—The main problem regarding nutrition faced by Indonesia is stunting, where Indonesia is ranked fifth in the world with the largest stunting prevalence rate in 2017, which is 29.6% of all Indonesian children. Stunting is a child under five years who has a *z-score* value of less than -3 standard deviations (SD). Stunting has a negative impact, namely it can disrupt the physical and intellectual development of toddlers in the future. In this case, the examination of stunting status by medical personnel is still carried out manually which takes a long time and is prone to inaccuracies. This study aims to classify stunting status in toddlers by applying the *K-Nearest Neighbor* method using the *Backward Elimination* feature selection to get fast and accurate results. Based on the results of this study, the average accuracy produced by the *K-Nearest Neighbor* algorithm at  $k=5$  is 91.90% with 9 attributes and the average accuracy produced by the *K-Nearest Neighbor* algorithm with the addition of *Backward Elimination* is 92.20%. with 8 attributes. These results indicate that the application of *Backward Elimination* can increase the accuracy value of the *K-Nearest Neighbor* algorithm and also perform attribute selection.

**Keywords:** Classification; Stunting; *K-Nearest Neighbor*; *Backward Elimination*; Data Mining

### 1. PENDAHULUAN

Permasalahan utama tentang gizi yang dihadapi Indonesia adalah balita stunting atau biasa disebut balita kerdil. Indonesia menempati peringkat kelima dunia dengan tingkat prevalensi stunting tertinggi sebesar 29,6% pada tahun 2017[1]. Stunting adalah kondisi gagal tumbuh pada balita karena kekurangan gizi kronis yang mengakibatkan balita sangat pendek dibandingkan dengan usianya [2]. Stunting dapat terjadi karena beberapa faktor yaitu praktek pengasuhan yang kurang baik, kurangnya pemahaman ibu terkait kesehatan dan gizi sebelum ataupun sesudah melahirkan, masih terbatasnya layanan kesehatan, minimnya akses keluarga ke makanan bergizi, dan terbatasnya akses air bersih juga sanitasi [2]. Peranan penting dipegang oleh orang tua guna membantu mengatasi masalah gizi terutama stunting, Orang tua yang memiliki pengetahuan tentang tanda-tanda, cara pencegahan dan efek yang timbul dapat menentukan tindakan juga perbuatan dalam mencegah terjadinya stunting [3]. Pemeriksaan status stunting dilakukan dengan perhitungan *z-score* yang mengacu pada antropometri yang ditetapkan oleh Kementerian Kesehatan Republik Indonesia dengan ambang batas  $<-3$  standar deviasi, dimana pemeriksaan status stunting pada balita masih membutuhkan waktu cukup lama karena dilakukan secara manual juga rentan akan ketidakakuratan. Oleh karena itu dibutuhkan sebuah sistem yang dapat melakukan klasifikasi data pemeriksaan balita guna memprediksi apakah anak tersebut masuk dalam status stunting atau tidak stunting dengan cepat dan akurat.

Klasifikasi merupakan suatu proses pengkategorian yang digunakan untuk menentukan kelas dari data yang tidak diketahui label *class* nya [4]. Memprediksi label kelas adalah tujuan utama dari klasifikasi. Ada dua tahapan yang dimiliki proses klasifikasi, pertama *learning*: menganalisa data *training* dengan menggunakan algoritma klasifikasi. Dan yang kedua klasifikasi: memprediksi ketepatan dari klasifikasi dengan menggunakan data *testing* [5].

Algoritma *K-Nearest Neighbor* adalah algoritma *supervised learning* dimana hasil klasifikasi dipengaruhi oleh mayoritas tetangga terdekat ke- $k$ , algoritma *K-NN* memiliki prinsip kerja yaitu menghitung jarak terdekat dari data baru dengan tetangga terdekatnya dan hanya mengandalkan memori juga menggunakan mayoritas tetangga untuk klasifikasi [5].

Seleksi fitur adalah penghapusan fitur yang tidak berpengaruh atau tidak signifikan dari set data guna memberikan efektivitas [6]. Salah satu metode seleksi fitur adalah *backward elimination*, merupakan sebuah metode yang mampu mengeliminasi variabel yang tidak relevan atau tidak signifikan dalam set data [7]. *K-NN*



dapat memberikan hasil akurasi yang baik pada data yang berjumlah besar tetapi K-NN sensitif terhadap atribut yang tidak signifikan, oleh karena itu kehadiran seleksi fitur menggunakan *backward elimination* turut membantu dalam meningkatkan performanya [8].

Pada penelitian sebelumnya yang dilakukan oleh Agus Byna dan Fadhiyah Noor Anisa dengan judul penelitian “*Backward Elimination* Untuk Meningkatkan Akurasi Kejadian Stunting Dengan Analisis Algoritma *Support Vector Machine*”. Jumlah data yang digunakan sebanyak 457 data dengan jumlah atribut sebanyak 13 atribut, berdasarkan hasil pengujian didapatkan nilai akurasi sebesar 81,62% menggunakan algoritma SVM sedangkan untuk algoritma SVM dengan *backward elimination* sebesar 90,16% [7].

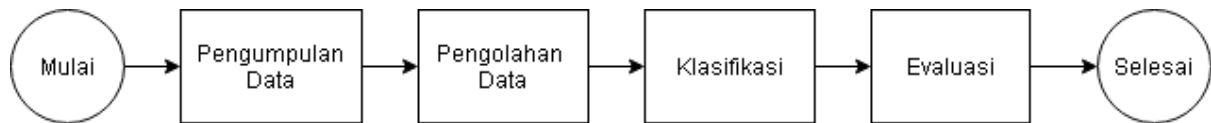
Pada penelitian yang dilakukan oleh Tio Prasetya, Irfan Ali, Cep Lukman Rohmat dan Odi Nurdiawan dengan judul penelitian “Klasifikasi Status Stunting Di Desa Slangit Menggunakan Metode *K-Nearest Neighbor*”. Menggunakan data penimbangan posyandu sebanyak 300 data dan 5 atribut. Berdasarkan hasil pengujian yang telah dilakukan, tingkat akurasi klasifikasi yang didapatkan sebesar 98.89% [9].

Penelitian lain yang dilakukan oleh Junta Zeniarja, Kiki Widia dan Ramadhan Rakhmat Sani dengan judul “Penerapan Algoritma *Naïve Bayes* dan *Forward Selection* Dalam Pengklasifikasian Status Gizi Stunting Pada Puskesmas Semarang”. Penelitian ini menggunakan set data berjumlah 300. Nilai akurasi yang didapatkan sebesar 83,33% tanpa menggunakan seleksi fitur dan meningkat menjadi 86% dengan seleksi fitur metode *forward selection* [10].

Berdasarkan penjelasan diatas maka penelitian ini akan membuat sebuah sistem yang dapat mengklasifikasi juga memeriksa status stunting pada balita dengan memanfaatkan algoritma *K-Nearest Neighbor* dengan seleksi fitur *Backward Elimination*.

## 2. METODOLOGI PENELITIAN

Penelitian ini diawali dengan pengumpulan data, lalu pengolahan data yaitu proses *preprocessing*, transformasi dan seleksi fitur menggunakan *backward elimination*, klasifikasi, dan terakhir melakukan evaluasi dengan *k-fold cross validation* dan *confusion matrix*. Gambar 1 memperlihatkan langkah-langkah penelitian yang dilakukan.



**Gambar 1.** Diagram Penelitian

### 2.1 Pengumpulan Data

Data yang dipakai adalah data pemeriksaan status gizi balita yang diambil dari Puskesmas Ubung, Kecamatan Jonggat, Kabupaten Lombok Tengah. Berupa file dengan ekstensi .xlsx dengan total data 1000 dengan 243 jumlah balita stunting dan 757 balita tidak stunting. Data ini diambil pada bulan Mei 2021. Data tersebut memiliki 9 atribut yaitu jenis kelamin, umur, berat, tinggi, berat badan dibandingkan dengan umur (bb/u), *z-score* bb/u, berat badan dibandingkan dengan tinggi badan (bb/tb), *z-score* bb/tb, *z-score* tb/u., dan 1 kelas *binary* yaitu 0 = tidak stunting dan 1 = stunting.

**Tabel 1.** Atribut dataset

Nama	Deskripsi
Jenis kelamin	0: Laki-laki, 1: Perempuan
Umur	Umur balita (0-59bulan)
Berat	Berat balita (kg)
Tinggi	Tinggi badan balita (cm)
BB/U	Berat badan dibandingkan umur (0: sangat kurang, 1: kurang, 2: normal, 3: risiko lebih)
Z-Score BB/U	Nilai <i>z-score</i> berat dibandingkan dengan umur
BB/TB	Berat badan dibandingkan tinggi (0: gizi buruk, 1: gizi kurang, 2: gizi baik, 3: risiko gizi lebih)
Z-Score BB/TB	Nilai <i>z-score</i> berat dibandingkan dengan tinggi badan
Z-Score TB/U	Nilai <i>z-score</i> tinggi dibandingkan dengan umur

### 2.2 Preprocessing (Data Cleaning)

Preprocessing adalah langkah penting dalam proses penemuan pengetahuan karena data yang didapatkan tidak lengkap, memiliki *noise* ataupun tidak konsisten [11]. Data pemeriksaan status gizi balita yang digunakan sudah lengkap, tidak memiliki *noise* dan sudah konsisten oleh karena itu dapat melanjutkan ke tahap selanjutnya.



**2.3 Transformasi Data**

Proses transformasi bertujuan merubah data dengan nilai kategori menjadi data bernilai angka sehingga data dapat diimplementasikan dengan mudah [12]. Terdapat 3 atribut yang akan ditransformasi yaitu jenis kelamin, bb/u dan bb/tb. Dimana jenis kelamin laki-laki akan ditransformasikan dengan nilai 0 dan 1 bagi balita perempuan. Data berat badan dibandingkan dengan umur (bb/u) memiliki 4 kategori yaitu sangat kurang, kurang, normal, risiko lebih. Data tersebut akan ditransformasikan menjadi nilai 0,1,2,3. Begitu juga dengan data berat badan berbanding dengan tinggi badan ditransformasikan menjadi 0,1,2,3 yang memiliki nilai kategori gizi buruk, gizi kurang, gizi baik, risiko gizi lebih. Hasil transformasi dapat anda lihat pada Tabel 2 dibawah.

**Tabel 2.** Transformasi Data

JK	Umur	Berat	Tinggi	BB/U	ZS BB/U	BB/TB	ZS BB/TB	ZS TB/U
0	44	6.7	81	0	-5.8	0	-5.02	-4.71
1	59	15	103.2	2	-1.32	2	-0.84	-1.24
1	58	15.8	102.4	2	-1.02	2	-0.19	-1.48
0	29	11.8	87.7	2	-0.91	2	-0.55	-1.05
1	58	17.1	102.8	2	-0.32	2	0.67	-1.22

**2.4 Seleksi Fitur**

Seleksi fitur berpengaruh langsung terhadap hasil klasifikasi [13]. Tujuan utamanya adalah memilih fitur terbaik serta mengecilkan dimensi yang tinggi sehingga mampu meningkatkan kinerja metode klasifikasi [11]. Salah satu metode untuk seleksi fitur adalah *backward elimination*. *Backward elimination* merupakan metode untuk menyeleksi fitur dengan melakukan pengujian kepada semua fitur terlebih dahulu, lalu secara bertahap mengurangi fitur yang tidak signifikan berdasarkan perbandingan evaluasi hasil uji yang didapatkan [14]. Menghilangkan sebuah *model regressor* satu per satu berdasarkan penurunan didalam sebuah model yang bagus adalah proses *backward elimination* konvensional [15]. Langkah pertama yang dilakukan saat melakukan *backward elimination* adalah menentukan nilai signifikan (*significance level*) sebesar 5% pada umumnya yang berarti *p-value* menjadi 0.05, Kedua, memasukkan semua fitur dalam model. Langkah ketiga, cari nilai fitur *p-value* terbesar dan bandingkan dengan nilai signifikan. Langkah keempat, jika *p-value* terbesar yang ada dalam model > 0.05 maka fitur tersebut akan dihapus. Langkah selanjutnya kembali ke langkah kedua sampai nilai *p-value* tertinggi dalam model < 0.05[16].

**2.5 Klasifikasi**

Klasifikasi merupakan suatu proses pengkategorian yang digunakan untuk menentukan kelas dari data yang tidak diketahui label kelasnya [4]. Ada dua tahapan proses dalam klasifikasi data. Pertama *learning*, proses menganalisa data training dengan menggunakan algoritma klasifikasi. Dan kedua klasifikasi, proses memprediksi ketepatan dari klasifikasi dengan menggunakan *data testing* [5].

**2.6 K-Nearest Neighbor**

*K-Nearest Neighbor* adalah metode klasifikasi terhadap data yang berlandaskan pada jarak terdekat pada data tersebut [17]. Nilai k sangat berpengaruh pada pengklasifikasian K-NN sebagai pembatas jarak terdekat yang akan digunakan, K-NN juga mempunyai kelebihan seperti tahan terhadap *dataset* yang besar dan juga mudah diimplementasikan [18]. Rumus algoritma K-NN adalah sebagai berikut:

$$D = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \tag{1}$$

Keterangan:

- x = sampel data
- y = data uji
- D = jarak

Ada banyak metode untuk menghitung jarak kedekatan data baru dengan *data training*, salah satu *euclidean distance*, rumus untuk mengukur jarak kedekatan dapat dilihat pada persamaan 2 dibawah:

$$D = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{2}$$

**2.6 K-Fold Cross Validation**

*K-fold cross validation* adalah salah satu teknik dari validasi silang. Tujuan penggunaan *k-fold cross validation* adalah untuk menghilangkan bias pada data. Pelatihan dan pengujian dilakukan sebanyak k [19]. Untuk k = 10, dataset akan dibagi menjadi 10 kali lipat berukuran serupa untuk tiap-tiap lipatan. Untuk masing-masing 10 subset, pengujian *cross validation* akan menggunakan 9 lipat untuk *training* dan 1 lipat untuk *testing* [20]. *10-fold cross validation* dapat anda liat pada Tabel 3 dibawah.



**Tabel 3.** 10-fold cross validation

fold	Dataset									
1	1	2	3	4	5	6	7	8	9	10
2	1	2	3	4	5	6	7	8	9	10
3	1	2	3	4	5	6	7	8	9	10
4	1	2	3	4	5	6	7	8	9	10
5	1	2	3	4	5	6	7	8	9	10
6	1	2	3	4	5	6	7	8	9	10
7	1	2	3	4	5	6	7	8	9	10
8	1	2	3	4	5	6	7	8	9	10
9	1	2	3	4	5	6	7	8	9	10
10	1	2	3	4	5	6	7	8	9	10

Keterangan:

Data training
Data testing

**2.6 Confusion Matrix**

Confusion matrix adalah metode yang digunakan untuk menganalisis seberapa akurat classifier dapat mengenali tuples dari kelas yang berbeda. Ada empat istilah yang perlu diketahui merupakan “building blocks” yang digunakan dalam menghitung ukuran evaluasi yaitu True Positives (TP), True Negatives (TN), False Positives (FP), False Negatives (FN). TP adalah jumlah data positif yang terklasifikasi benar oleh sistem, TN adalah jumlah data negatif terklasifikasi dengan benar oleh sistem. FP adalah jumlah data positif namun terklasifikasi salah oleh sistem. Dan FN jumlah data negatif terklasifikasi salah oleh sistem [21]. Pembagian nilai dapat anda lihat pada Tabel 4.

**Tabel 5.** Confusion Matrix

	Hasil prediksi		
		True	False
Hasil sebenarnya	True	TP	FN
	False	FP	TN

Confusion matrix menghasilkan keluaran yaitu akurasi, recall, precision, f1-score yang diformulasikan pada persamaan 3,4,5,6.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$F1 - score = 2 * \frac{Precision*Recall}{Precision+Recall} \tag{6}$$

Akurasi dihitung dengan membandingkan jumlah data yang benar dan tepat terklasifikasi dengan jumlah keseluruhan data, dapat juga dikatakan akurasi untuk menggambarkan seberapa akurat model dalam mengklasifikasi dengan benar. Precision diartikan sebagai ukuran ketepatan antara data yang diminta dengan hasil prediksi oleh model. Sedangkan recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi [12].

**3. HASIL DAN PEMBAHASAN**

Pengujian pertama dilakukan dengan algoritma k-nn tanpa menggunakan seleksi fitur. Pengujian tersebut memakai data sejumlah 1000 data dengan 9 atribut yang telah ditransformasi menjadi data numerik yaitu jenis kelamin, umur, berat, tinggi, bb/u, z-score bb/u, bb/tb, z-score bb/tb, z-score tb/u. Nilai tetangga yang digunakan adalah k =5 karena memiliki tingkat akurasi yang paling besar diantara nilai k yang lain. Sedangkan untuk cross validation digunakan 10-fold cross validation.

**Tabel 6.** Hasil Pengujian Menggunakan K-NN

Iterasi ke	Akurasi	Presisi	Recall	F1-Score
1	95.0	100.0	64.0	78.0
2	95.0	93.0	78.0	85.0



Iterasi ke	Akurasi	Presisi	Recall	F1-Score
3	98.0	100.0	71.0	83.0
4	98.0	100.0	82.0	90.0
5	92.0	90.0	56.0	69.0
6	95.0	100.0	29.0	44.0
7	98.0	100.0	80.0	89.0
8	95.0	100.0	50.0	67.0
9	91.0	89.0	100.0	94.0
10	62.0	62.0	100.0	77.0
<b>Rata-rata</b>	<b>91.9</b>	<b>93.4</b>	<b>71.0</b>	<b>77.6</b>

Dari Tabel 6 dengan nilai  $k=5$  dan *10-fold cross validation* didapatkan hasil rata akurasi sebesar 91,9%, presisi sebesar 93,4%, *recall* sebesar 71.0% dan *f1-score* sebesar 77.6%. Setelah itu melakukan seleksi fitur dengan *backward elimination* dengan memasukkan semua atribut dalam *model regressor 1*. Tabel 7 memperlihatkan nilai *p-value* dari seluruh atribut.

**Tabel 7.** *P-value model regressor 1*

	<i>P-value</i>
<i>Intercept</i>	0.0000
JK	0.0086
Umur	0.0026
Berat	0.0106
Tinggi	0.0000
BB/U	0.0000
ZS BB/U	0.0000
BB/TB	0.2907
ZS BB/TB	0.0000
ZS TB/U	0.0037

Setelah dilakukan *backward elimination*, atribut yang memiliki nilai *p-value* terbesar adalah bb/tb, dimana akan dibandingkan apakah *p-value* bb/tb ( $0.290 > 0.05$ ) dan hasilnya benar lebih besar. Oleh karena itu bb/tb akan dieliminasi dari model. Dan dilanjutkan seleksi pada atribut yang lain tanpa menyertakan bb/tb dalam *model regressor 2*. Tabel memperlihatkan nilai *p-value* dari 8 atribut yang tersisa.

**Tabel 8.** *P-value model regressor 2*

	<i>P-value</i>
<i>Intercept</i>	0.0000
JK	0.0069
Umur	0.0021
Berat	0.0170
Tinggi	0.0000
BB/U	0.0000
ZS BB/U	0.0000
ZS BB/TB	0.0000
ZS TB/U	0.0057

Pada *model regressor 2* tidak ada lagi atribut yang memiliki nilai *p-value*  $>$  *significance level*. Oleh karena itu penyeleksian fitur dihentikan. Dan menggunakan *model regressor 2* untuk melakukan pengujian. Pada pengujian kedua yaitu *k-nn* dengan seleksi fitur *backward elimination*, atribut yang terseleksi adalah bb/tb. Oleh karena itu percobaan kedua dilakukan dengan 8 atribut yaitu jenis kelamin, umur, berat, tinggi, bb/u, *z-score* bb/u, *z-score* bb/tb, *z-score* tb/u. Percobaan kedua dilakukan sama dengan nilai  $k=5$  dan *10-fold cross validation*.

**Tabel 7.** Hasil Pengujian Menggunakan K-NN *backward elimination*

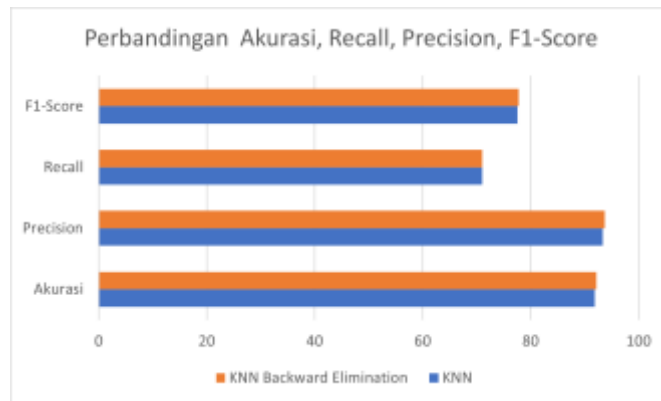
Iterasi ke	Akurasi	Presisi	<i>Recall</i>	<i>F1-Score</i>
1	95.0	100.0	64.0	78.0
2	95.0	93.0	78.0	85.0
3	98.0	100.0	71.0	83.0
4	98.0	100.0	82.0	90.0
5	92.0	90.0	56.0	69.0
6	95.0	100.0	29.0	44.0
7	98.0	100.0	80.0	89.0
8	95.0	100.0	50.0	67.0



Iterasi ke	Akurasi	Presisi	Recall	F1-Score
9	91.0	89.0	100.0	94.0
10	65.0	65.0	100.0	79.0
<b>Rata-rata</b>	<b>92.2</b>	<b>93.7</b>	<b>71.0</b>	<b>77.8</b>

Dari tabel diatas dapat dilihat hasil dari pengujian k-nn dengan *backward elimination* yang dilakukan dengan jupyter notebook didapatkan hasil rata-rata akurasi sebesar 92,2%, presisi sebesar 93,7%, recall sebesar 71.0% dan f1-score sebesar 77,8%. Melihat perbandingan dari kedua pengujian diatas, peneliti memperoleh kesimpulan bahwa rata-rata nilai akurasi tertinggi sebesar 92.2% dicapai pada saat dilakukan pengujian dengan nilai k=5 dan *10-fold cross validation* menggunakan algoritma k-nn dengan fitur seleksi *backward elimination* dan 91,9% jika tidak menggunakan fitur seleksi.

Berdasarkan nilai rata-rata akurasi yang didapatkan bisa dikatakan klasifikasi tersebut masuk dalam kategori *good classification*. Penggunaan *backward elimination* sebagai seleksi fitur pada algoritma k-nn dapat dikatakan berhasil untuk meningkatkan nilai akurasi sebesar 0,30% dari hasil akurasi pengujian sebelumnya. Untuk lebih jelasnya perbandingan dari hasil percobaan diatas dapat dilihat pada Gambar 2 dibawah.

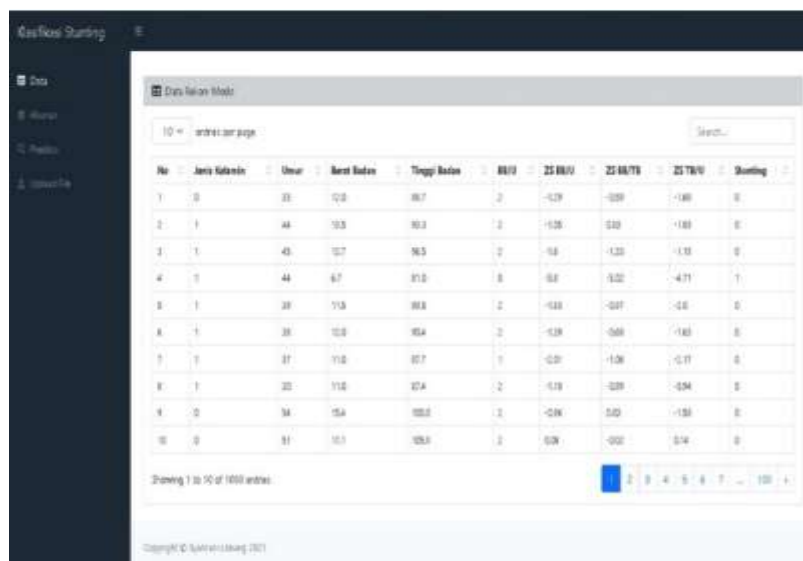


Gambar 2. Grafik Perbandingan

### 3.1 Implementasi Program

#### a. Tampilan awal sistem

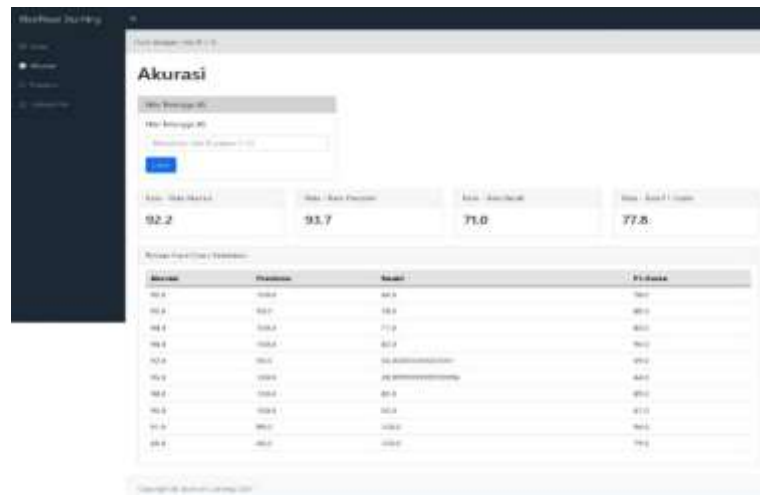
Halaman awal sistem *user* dapat melihat data-data seperti data jenis kelamim, berat, tinggi badan, bb/u, z-score bb/u, z-score bb/tb, z-score tb/u serta *class* nya apakah balita tersebut positif stunting (1) ataupun tidak stunting (0). Disamping kiri juga terdapat *side bar* yang berisikan fitur menu untuk membuka data, akurasi, prediksi maupun upload data.



Gambar 3. Tampilan awal

#### b. Halaman akurasi

Pada halaman ini *user* dapat memasukkan nilai k untuk melakukan perhitungan akurasi, presisi, recall, dan *f1-score*, *user* harus memasukkan nilai k 1 sampai dengan 10. Jika *user* tidak memasukkan nilai k, maka halaman tersebut tidak akan dapat menampilkan hasil akurasi, presisi, *recall* dan *f1-score*.



**Gambar 4.** Halaman Akurasi

c. Halaman prediksi

Pada halaman ini *user* dapat memasukkan data baru dan memeriksa apakah data balita yang baru diinputkan masuk dalam kategori stunting atau tidak. *User* harus memasukkan semua data yang ada pada form dibawah yaitu data jenis kelamin, berat, tinggi badan, *bb/u*, *z-score bb/u*, *z-score bb/tb*, *z-score tb/u*. Jika salah satu data tersebut tidak dimasukkan maka hasil tidak dapat ditampilkan dan meminta untuk diisi pada *form* yang masih kosong. Jika data balita yang dimasukkan masuk dalam kategori stunting maka akan tampil alert berwarna merah bertuliskan “Stunting” diatas kata Prediksi, jika tidak maka akan tampil *alert* berwarna hijau bertuliskan “Tidak Stunting”.



**Gambar 5.** Halaman Prediksi

d. Halaman Upload File

Pada halaman ini *user* dapat menambahkan data baru yang berfungsi untuk menambahkan *data training*, *upload* data hanya dapat dilakukan dengan ekstensi file *xlsx*.



**Gambar 5.** Halaman Upload File



#### 4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dengan judul klasifikasi status stunting pada balita menggunakan *k-nearest neighbor* dengan fitur seleksi *backward elimination*, dapat ditarik beberapa kesimpulan sebagai berikut peneliti dapat menerapkan dan membuat sistem klasifikasi status stunting pada balita dengan menerapkan metode *k-nearest neighbor* dengan seleksi fitur *backward elimination*. Tingkat akurasi yang didapatkan *k-nearest neighbor* dengan fitur seleksi *backward elimination* mencapai 92,2% meningkat 0,30%. Hal tersebut menunjukkan bahwa seleksi fitur yakni *backward elimination* meningkatkan performa yang lebih baik dan mempengaruhi hasil pencapaian akurasi yang didapatkan. Adapun saran, dapat dilakukan uji perbandingan hasil akurasi menggunakan algoritma klasifikasi selain algoritma k-nn maupun metode seleksi fitur selain *backward elimination* yang mungkin dapat menghasilkan nilai akurasi yang lebih baik untuk penelitian selanjutnya.

#### UCAPAN TERIMAKASIH

Terima kasih peneliti ucapkan kepada Tuhan Yang Maha Esa, terima kasih juga peneliti ucapkan kepada pihak Puskesmas Ubung Kec. Jongga Kab. Lombok Tengah yang telah memberikan izin penggunaan data pemeriksaan status gizi balita Puskesmas Ubung untuk penyelesaian penelitian ini. Terima kasih juga diucapkan kepada kedua orang tua, dosen pembimbing akademik juga dosen pembimbing skripsi. Terima kasih.

#### REFERENCES

- [1] Kementerian Kesehatan Republik Indonesia, "Buletin Jendela Data dan Informasi Kesehatan: Situasi Balita Pendek (Stunting) di Indonesia," *Kementeri. Kesehat. RI*, vol. 301, no. 5, pp. 1163–1178, 2018.
- [2] TNP2K, *100 Kabupaten/Kota Prioritas untuk Intervensi Anak Kerdil (Stunting): Tim Nasional Percepatan Penanggulangan Kemiskinan*, vol. 2, no. c. 2017.
- [3] A. Rahmawati, T. Nurmawati, and L. Permata Sari, "Faktor yang Berhubungan dengan Pengetahuan Orang Tua tentang Stunting pada Balita," *J. Ners dan Kebidanan (Journal Ners Midwifery)*, vol. 6, no. 3, pp. 389–395, 2019.
- [4] L. Andiani, S. Sukemi, and D. P. Rini, "Analisis Penyakit Jantung Menggunakan Metode KNN Dan Random Forest," in *Annual Research Seminar (ARS)*, 2020, vol. 5, no. 1, pp. 165–169.
- [5] A. D. Ghani, N. Salman, and Mustikasari, "Algoritma k-Nearest Neighbor Berbasis Backward Elimination Pada Client Telemarketing," in *Prosiding Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 2019, vol. 8, no. 2, pp. 141–150.
- [6] A. Bode, "K-Nearest Neighbor Dengan Feature Selection Menggunakan Backward Elimination Untuk Prediksi Harga Komoditi Kopi Arabika," *Ilk. J. Ilm.*, vol. 9, no. 2, pp. 188–195, 2017.
- [7] A. Byna and N. Anisa, "Backward Elimination Untuk Meningkatkan Akurasi Kejadian Stunting Dengan Analisis Algoritma Support Vector Machine," *Din. Kesehat. J. Kebidanan Dan Keperawatan*, vol. 9, no. 2, pp. 217–225, 2018.
- [8] H. Harafani and H. A. Al-Kautsar, "Meningkatkan Kinerja K-Nn Untuk Klasifikasi Kanker Payudara Dengan Forward Selection," *J. Pendidik. Teknol. dan Kejur.*, vol. 18, no. 1, p. 99, 2021.
- [9] T. Prasetya, I. Ali, C. L. Rohmat, and O. Nurdiawan, "Klasifikasi Status Stunting Balita Di Desa Slangit Menggunakan Metode K-Nearest Neighbor," *Informatics Educ. Prof.*, vol. 4, no. 2, pp. 93–104, 2020.
- [10] J. Zeniarja, K. Widia, and R. R. Sani, "Penerapan Algoritma Naive Bayes dan Forward Selection dalam Pengklasifikasian Status Gizi Stunting pada Puskesmas Pandanaran Semarang," *JOINS (Journal Inf. Syst.)*, vol. 5, no. 1, pp. 1–9, 2020.
- [11] Y. A. PRATAMA, "ANALISIS METODE SELEKSI FITUR UNTUK MENINGKATKAN AKURASI PADA VARIANT METODE KLASIFIKASI K-NEAREST NEIGHBOR (kNN)," pp. 7–37, 2019.
- [12] I. A. Nikmatun and I. Waspada, "Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 10, no. 2, pp. 421–432, 2019.
- [13] X. Zhang, Z. Shi, X. Liu, and X. Li, "A hybrid feature selection algorithm for classification unbalanced data processing," *Proc. - 2018 IEEE Int. Conf. Smart Internet Things, SmartIoT 2018*, pp. 269–275, 2018.
- [14] P. Haryandi, "Implementasi Seleksi Fitur dengan Backward Elimination untuk Klasifikasi Prediksi Perceraian," *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, 2021.
- [15] A. Bode, "Seleksi Fitur Untuk Prediksi Rating Film Hollywood Menggunakan Model K-Nearest Neighbor," *Jupiter*, vol. 5, no. 1, pp. 1–8, 2019.
- [16] S. Srinidhi, "Backward Elimination for Feature Selection in Machine Learning," *Towar. Data Sci.*, 2019.
- [17] S. R. Raysyah, Veri Arinal, and Dadang Iskandar Mulyana, "Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode Knn Dan Pca," *JSII (Jurnal Sist. Informasi)*, vol. 8, no. 2, pp. 88–95, 2021.
- [18] A. Ariani and S. Samsuryadi, "Klasifikasi Penyakit Ginjal Kronis menggunakan K-Nearest Neighbor," in *Annual Research Seminar (ARS)*, 2020, vol. 5, no. 1, pp. 148–151.
- [19] F. Tempola, M. Muhammad, and A. Khairan, "Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, p. 577, 2018.
- [20] D. Normawati and D. P. Ismi, "K-Fold Cross Validation for Selection of Cardiovascular Disease Diagnosis Features by Applying Rule-Based Datamining," *Signal Image Process. Lett.*, vol. 1, no. 2, pp. 23–35, 2019.
- [21] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012.