

Implementasi Algoritma *k-Nearest Neighbor* (k-NN) dalam Klasifikasi Status Gizi Balita

¹Sahara Syarifatul Choeriyah, ²Riezan Syauqi Fanhas, ³Adittia Fathah, ⁴Haerul Pebriyansyah

^{1,2,4}Program Studi Informatika Sekolah Tinggi Teknologi Cipasung

³Program Studi Sistem Informasi STMIK LIKMI Bandung

Jl. Raya Singaparna-Ciawi Km.1 PO. BOX 24 Cilampunghilir Padakembang Tasikmalaya 46466

Telp./Fax : 0265-2550424

Email: sahara@sttcipasung.ac.id

riezansfanhas@sttcipasung.ac.id

milki10220013@sttcipasung.ac.id

Abstract— *Toddlers are children under the age of 5 years or 0-60 months. This age is included in a group that is at high risk of disease. Nutritional status in toddlers is an important factor that must be considered because the development of toddlers is very important for their bodies which are still very vulnerable to the name malnutrition. The purpose of this study is to group the nutritional status of toddlers by utilizing the k-NN algorithm and knowing the level of accuracy. The method used in this research is quantitative using the k-NN algorithm. While the evaluation uses Confusion Matrix The results showed an AUC value of 85.1%. Based on these results, it can be concluded that the k-NN algorithm is well used for the classification of toddler nutrition in the future.*

Keywords— *Classification; Toddler Nutrition; k-NN; Confusion Matrix.*

Abstrak— Balita adalah anak usia dibawah 5 tahun atau 0-60 bulan. Usia tersebut termasuk kedalam kelompok yang beresiko tinggi terhadap penyakit. Status Gizi pada balita adalah faktor penting yang harus diperhatikan karena masa balita perkembangan menjadi hal yang sangat penting bagi tubuh mereka yang masih sangat rentan dengan namanya kekurangan gizi. Tujuan penelitian ini adalah untuk melakukan pengelompokan status gizi balita dengan memanfaatkan algoritma k-NN serta mengetahui tingkat akurasi. Metode yang digunakan pada penelitian ini adalah kuantitatif dengan menggunakan algoritma k-NN. Sedangkan evaluasinya menggunakan *Confusion Matrix* Hasil penelitian menunjukkan nilai AUC sebesar 85,1%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma k-NN baik digunakan untuk klasifikasi gizi balita di masa yang akan datang.

Kata kunci— *Klasifikasi; Gizi Balita; K-NN; Confusion Matrix.*

I. PENDAHULUAN

Balita adalah anak usia dibawah 5 tahun atau 0-60 bulan. Usia tersebut termasuk kedalam kelompok yang beresiko tinggi terhadap penyakit. Kekurangan maupun kelebihan asupan zat gizi pada balita dapat berdampak pada status gizi dan status kesehatannya. Berdasarkan data laporan kesehatan *The World Health Organization* (WHO) dan Posyandu balita, ada beberapa masalah status gizi yang dapat mempengaruhi kesehatannya, seperti kekurangan energi protein, obesitas, kurang vitamin A, gangguan akibat kekurangan Iodium dan anemia atau kurang zat besi (FE). Status Gizi pada balita adalah faktor penting yang harus diperhatikan karena masa balita perkembangan menjadi hal yang sangat penting bagi tubuh mereka yang masih sangat rentan dengan namanya kekurangan gizi [1].

Penelitian tentang pemanfaatan data

mining untuk klasifikasi gizi bayi sudah banyak dilakukan.[2] melakukan penelitian di posyandu melati IV untuk mengetahui apakah mengalami gangguan gizi atau normal. Dataset yang digunakan berjumlah 219. [3] melakukan penelitian di Posyandu Melati di wilayah Plumbon kabupaten Cirebon untuk melakukan deteksi pada bayi yang mengalami gizi buruk. Dataset yang diperoleh sebanyak 176. [4] melakukan penelitian di posyandu Kenanga yang berada di kabupaten Simalungun dengan tujuan melakukan uji coba model untuk klasifikasi status gizi balita berdasarkan baku rujukan WHO-2005. Data yang diolah sebanyak 40 data. Sedangkan [1] melakukan penelitian di Posyandu Melati, Kelurahan Margatunggal, Kecamatan Jayaloka, Kabupaten Musi-Rawas, Sumatera Selatan dengan tujuan untuk penentuan balita yang layak sebagai penerima makanan tambahan (PMT) dan meminimalisir terjadinya kesalahan dalam memilih balita yang layak menerima

bantuan. Data yang digunakan berjumlah 207. Penelitian selanjutnya dilakukan oleh [5] di Puskesmas Padangsari, Kec. Banyumanik, Kota Semarang untuk mengetahui bagaimana pengklasifikasian status gizi balita stunting agar dapat diketahui secara mendetail dengan akurasi yang tepat. Dataset yang didapatkan sebanyak 1066 data. Penelitian juga telah dilakukan oleh [6] di Kota Blitar. Dataset yang diperoleh sebanyak 6.908.

Algoritma yang digunakan untuk melakukan klasifikasi juga berbeda-beda. [2], [1], [5] dan [6] memilih untuk menggunakan naïve bayes. Berbeda dengan tiga peneliti lainnya [4] menggunakan Decision Tree dan algoritma C.45. Penelitian sebelumnya sebagian besar menggunakan software Rapid Mincer untuk melakukan klasifikasi, namun ada juga yang menggunakan WEKA, yaitu [2]. Atribut yang digunakan pun hampir sama. Atribut yang dipakai oleh [2] yaitu inisial, jenis kelamin, umur dan berat badan. [3] menggunakan umur, berat badan dan tinggi badan. [4] memakai atribut berat badan, tinggi badan dan bentuk badan. [1] mengambil data jenis kelamin, berat badan, tinggi badan. [5] menambah atribut lain yaitu lingkaran lengan atas dan usia sedangkan [6] hanya mengambil atribut usia, berat badan dan tinggi badan.

Akurasi yang dihasilkan pada penelitian sebelumnya cukup besar, bahkan ada yang mencapai 100 persen. Penelitian yang dilakukan oleh [2] menghasilkan akurasi sebesar 60% dan error 40%. Hasil tersebut menunjukkan Naive Bayes bisa digunakan untuk mengklasifikasi gizi anak dengan menggunakan 5 atribut yang diteliti. Sedangkan penelitian yang lainnya menghasilkan akurasi lebih dari 90 persen. Misalnya Nilai akurasi yang diperoleh [3] dan [5] yaitu 98,86% dan 96%. Hasil penelitian menunjukkan bayi yang berumur 0 sampai 10 bulan memiliki gizi yang normal sedangkan bayi umur 11 sampai 19,5 bulan merupakan rawan akan gizi buruk dan umur lebih dari 19,5 bulan memiliki gizi yang normal. Hasil akurasi yang paling baik diperoleh pada penelitian [4]. Akurasi yang dihasilkan adalah 100%. Hal tersebut bisa terjadi karena data yang digunakan sedikit. Berbeda dengan nilai akurasi sebelumnya, hasil akurasi pada penelitian [1] dibagi menjadi 2, yaitu akurasi tertinggi pada iterasi ke-8 dengan akurasi sebesar 95,14%, sedangkan akurasi terendah berada pada iterasi ke-3 dengan akurasi sebesar 81,73%. Sedangkan pada penelitian [6] nilai akurasi diperoleh berdasarkan 3 skenario pengujian, didapatkan bahwa dengan akurasi sebesar 86% dan error rate 1,7% melalui perbandingan data testing dan training

yang digunakan adalah 70:30. Ketika perbandingan tersebut dirubah menjadi 50:50, maka akurasi yang dihasilkan akan semakin kecil yaitu sekitar 75% dan apabila data testing dan training tersebut diturunkan kembali menjadi 30:70 maka akurasi yang dihasilkan akan semakin kecil yaitu 66,8%. Berdasarkan penelitian [6] maka dapat dikatakan semakin banyak data testing yang digunakan maka semakin besar nilai akurasi yang dihasilkan.

Berdasarkan penelitian terkait yang sudah dibahas, penulis akan melakukan penelitian dengan algoritma yang berbeda yaitu k-NN. Tujuan penelitian ini adalah untuk melakukan pengelompokan status gizi balita dengan memanfaatkan algoritma K-NN serta mengetahui tingkat akurasinya.

II. LANDASAN TEORI

A. Data Mining

Data mining merupakan inti dari proses *Knowledge Discovery In Database* (KDD), meliputi dugaan algoritma yang terdiri dari eksplor data, membangun model dan menentukan pola yang belum diketahui. KDD sifatnya otomatis, yaitu bisa didefinisikan sebagai pengidentifikasian yang benar, berguna dan penemuan pola dari kumpulan data yang besar dan kompleks [7].

B. Klasifikasi

Klasifikasi adalah salah satu metode *data mining* yang bisa membantu dalam melakukan prediksi sesuai label kelas sampel yang ingin dikelompokkan. Klasifikasi merupakan proses analisa data yang menghasilkan model-model dan dapat digambarkan melalui kelas-kelas yang ada didalam data. Metode ini memiliki syarat, yaitu atribut datanya harus numerik atau nominal dan label data nominal [8]. Klasifikasi bisa dipakai pada pencarian pola dengan membangun model klasifikasi berdasarkan variabel kelas prediktor dan variabel kelas target [9].

C. k-NN

Klasifikasi *k-Nearest Neighbor* (*k-NN*) menggunakan algoritma terawasi yang mengelompokkan hasil sampel uji baru berdasarkan kelas k-NN terbanyak. Metode k-NN dapat digunakan untuk memilih input untuk model nonlinier dan juga menghasilkan informasi yang akurat untuk peramalan deret waktu. Saat digunakan untuk regresi, nilai diberikan ke properti objek sebagai rata-rata nilai yang diketahui dari k tetangga terdekatnya. Algoritma k-NN menggunakan klasifikasi tetangga sebagai prediktor pola pengujian baru dengan jarak *Euclidean* [10].

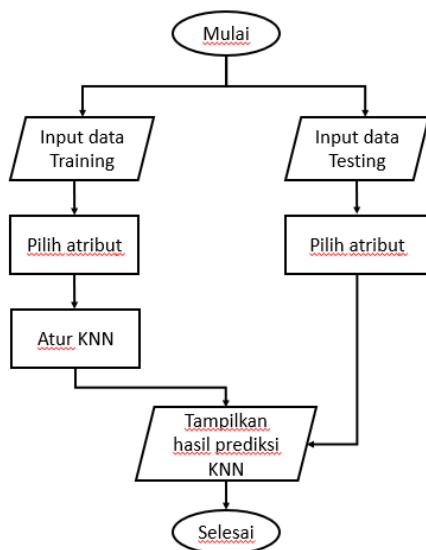
D. Confusion Matrix

Confusion matriks adalah salah satu pengujian akurasi yang digunakan dalam klasifikasi. *Confusion matrix* dapat digunakan untuk mengetahui tingkat AUC (*Area Under Curve*), *recall*, *precision* dan *F1 score*. AUC digunakan untuk menghitung perbedaan performansi metode yang digunakan [11]. Precision bisa digunakan untuk mengevaluasi kemampuan sistem untuk meningkatkan peringkat yang paling relevan. Recall bertujuan untuk menghitung kemampuan sistem dalam menemukan yang relevan terhadap data. *Accuracy* dapat dipakai untuk menghitung data yang diduga benar dengan jumlah seluruh data, sedangkan *error rate* digunakan untuk menghitung data yang diduga salah dengan membandingkan seluruh jumlah data [10].

III. METODE PENELITIAN

Penelitian menggunakan metode kuantitatif. Data yang digunakan adalah usia, tinggi badan dan berat badan. Dataset diperoleh dari website Kaggle [12]. Data yang diperoleh selanjutnya akan dianalisis berdasarkan studi pustaka dari berbagai sumber, seperti buku, jurnal dan internet yang berkaitan dengan objek penelitian [13]. Jumlah dataset yang diteliti sebanyak 200 data yang terdiri dari 160 data training dan 40 data testing.

Pengolahan data dilakukan dengan menerapkan algoritma K-NN. Software yang digunakan penulis adalah *Orange Data Mining*. Data dibagi menjadi 4 klasifikasi, yaitu gizi buruk, gizi kurang, gizi normal serta gizi lebih. Setelah diolah, data diuji akurasinya dengan menggunakan *Confusion Matriks*. Tahapan klasifikasi menggunakan algoritma K-NN dapat dilihat pada Gambar 1.



Gambar 1. Skema K-NN

Terlihat pada skema Gambar 1, data *Training* yang sudah diklasifikasi dengan K-NN kemudian dihubungkan dengan data *Testing*. Hal tersebut bertujuan untuk melihat prediksi klasifikasi data berdasarkan pola yang dipelajari oleh algoritma K-NN pada *data testing*.

IV. HASIL PENELITIAN

Dataset yang didapatkan dari Kaggle [12] dapat dibagi menjadi dua, yaitu data *training* dan *data testing*. Data dibagi dengan perbandingan 80:20. *Data training* dapat dilihat pada Tabel 1.

TABEL I
DATA TRAINING

No	Umur (bulan)	Berat (kg)	Tinggi (cm)	Status Gizi
1	24	5,8	65	GIZI_KURANG
2	24	5,5	59	GIZI_NORMAL
3	28	6,7	71,5	GIZI_KURANG
4	30	8,1	72,5	GIZI_NORMAL
5	28	6,9	73	GIZI_KURANG
6	25	5,8	77	GIZI_BURUK
7	23	4,7	52,5	GIZI_NORMAL
8	28	6,8	54	GIZI_NORMAL
9	21	3,5	55	GIZI_KURANG
10	27	5,8	53	GIZI_NORMAL
11	28	7,1	62	GIZI_NORMAL
12	32	9,4	61	GIZI_LEBIH
13	24	8,2	65	GIZI_NORMAL
14	22	5	68	GIZI_BURUK
15	23	6	67	GIZI_KURANG
16	31	10,2	50	GIZI_LEBIH
17	31	12,8	72	GIZI_NORMAL
18	28	10,1	95	GIZI_KURANG
19	24	9,4	75	GIZI_NORMAL
20	22	10	88	GIZI_KURANG
21	21	10	87	GIZI_KURANG
22	23	12	98	GIZI_KURANG
23	32	5,8	54	GIZI_NORMAL
24	19	4,7	67	GIZI_BURUK
25	19	4,7	67	GIZI_BURUK
26	23	5,7	77	GIZI_NORMAL
27	41	8,9	75	GIZI_KURANG
28	23	9	78	GIZI_LEBIH
29	32	10	56	GIZI_LEBIH
30	13	12	67	GIZI_NORMAL

31	23	13	78	GIZI_NORMAL	75	40	7,5	68	GIZI_NORMAL
32	23	9,4	76	GIZI_KURANG	76	34	8,9	77	GIZI_KURANG
33	32	9,7	88	GIZI_NORMAL	77	23	8,8	79	GIZI_KURANG
34	32	9,8	78	GIZI_NORMAL	78	45	9,7	85	GIZI_NORMAL
35	22	9,9	81	GIZI_BURUK	79	34	8,8	74	GIZI_NORMAL
36	21	5,8	82	GIZI_KURANG	80	23	9,5	76	GIZI_NORMAL
37	22	8,8	83	GIZI_KURANG	81	32	9	65	GIZI_LEBIH
38	32	9	84	GIZI_KURANG	82	34	12	67	GIZI_NORMAL
39	32	10	85	GIZI_KURANG	83	33	10	78	GIZI_KURANG
40	11	11	86	GIZI_NORMAL	84	12	10	87	GIZI_NORMAL
41	14	12	87	GIZI_NORMAL	85	32	9	67	GIZI_NORMAL
42	23	13	88	GIZI_BURUK	86	43	12	77	GIZI_NORMAL
43	43	7,5	89	GIZI_BURUK	87	43	11	79	GIZI_NORMAL
44	23	6,5	77	GIZI_BURUK	88	23	12	87	GIZI_NORMAL
45	27	5,4	76	GIZI_BURUK	89	43	8,9	74	GIZI_KURANG
46	28	3,4	75	GIZI_KURANG	90	32	9,5	88	GIZI_BURUK
47	37	7,9	82	GIZI_KURANG	91	34	5,8	76	GIZI_KURANG
48	34	8	81	GIZI_KURANG	92	35	7,8	75	GIZI_NORMAL
49	45	9	80	GIZI_BURUK	93	37	9,8	64	GIZI_NORMAL
50	43	7	82	GIZI_BURUK	94	45	9	67	GIZI_KURANG
51	23	5,7	75	GIZI_KURANG	95	42	10	84	GIZI_LEBIH
52	27	6,8	76	GIZI_KURANG	96	36	12	65	GIZI_NORMAL
53	34	7,8	73	GIZI_KURANG	97	37	10	65	GIZI_LEBIH
54	35	7,9	84	GIZI_KURANG	98	38	9	55	GIZI_LEBIH
55	32	8,9	82	GIZI_KURANG	99	26	9,6	53	GIZI_LEBIH
56	24	9	89	GIZI_NORMAL	100	37	9,7	57	GIZI_NORMAL
57	43	9,7	77	GIZI_LEBIH	101	45	8,9	76	GIZI_KURANG
58	35	14	73	GIZI_NORMAL	102	24	8,8	78	GIZI_BURUK
59	34	13	75	GIZI_NORMAL	103	54	7,6	93	GIZI_KURANG
60	34	12	89	GIZI_KURANG	104	34	7,7	78	GIZI_KURANG
61	23	11	91	GIZI_KURANG	105	23	8,7	77	GIZI_NORMAL
62	32	10	95	GIZI_BURUK	106	45	9	75	GIZI_NORMAL
63	36	5,4	75	GIZI_BURUK	107	43	10	77	GIZI_LEBIH
64	45	5,5	77	GIZI_KURANG	108	4	13	67	GIZI_LEBIH
65	38	7,8	73	GIZI_NORMAL	109	45	12	58	GIZI_LEBIH
66	43	8,9	72	GIZI_NORMAL	110	47	10	53	GIZI_NORMAL
67	43	9,8	67	GIZI_NORMAL	111	45	9,8	64	GIZI_NORMAL
68	44	8,8	75	GIZI_BURUK	112	44	8	65	GIZI_KURANG
69	23	6,9	94	GIZI_KURANG	113	43	7,9	78	GIZI_LEBIH
70	45	9	88	GIZI_KURANG	114	23	8,8	57	GIZI_NORMAL
71	43	8	76	GIZI_BURUK	115	24	9,8	57	GIZI_LEBIH
72	23	8,7	91	GIZI_NORMAL	116	45	7,8	76	GIZI_KURANG
73	19	7,8	69	GIZI_NORMAL	117	37	8,6	68	GIZI_NORMAL
74	29	6,6	54	GIZI_NORMAL	118	42	6,7	78	GIZI_KURANG

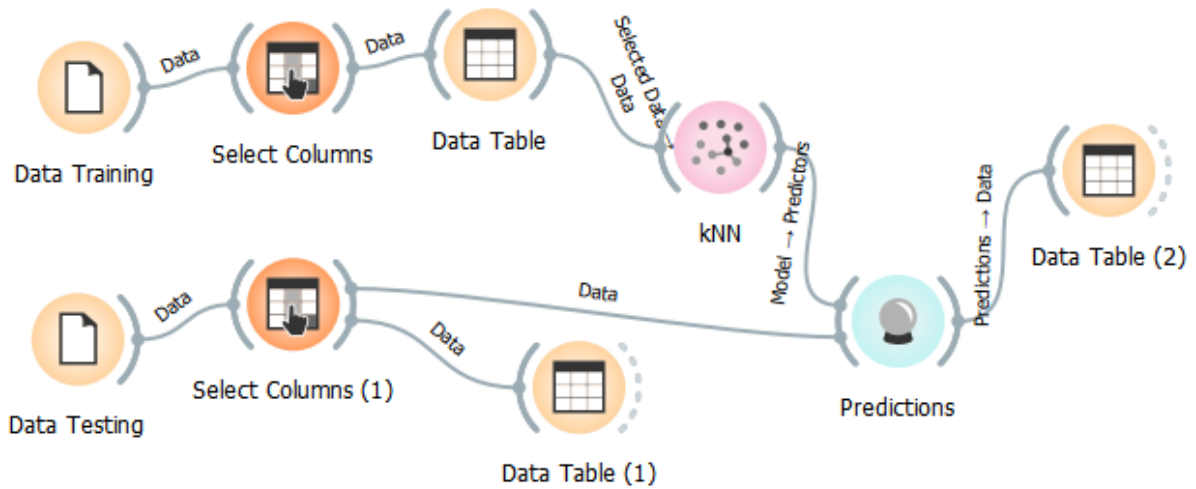
119	36	7,5	78	GIZI_KURANG
120	35	6,6	87	GIZI_BURUK
121	35	9,8	97	GIZI_BURUK
122	43	7,9	75	GIZI_KURANG
123	23	8	68	GIZI_NORMAL
124	36	9	76	GIZI_NORMAL
125	27	13	67	GIZI_LEBIH
126	36	10	85	GIZI_KURANG
127	43	9,6	75	GIZI_NORMAL
128	33	9,9	68	GIZI_NORMAL
129	23	8	87	GIZI_BURUK
130	32	7,6	87	GIZI_BURUK
131	25	4,9	88	GIZI_BURUK
132	37	5,9	67	GIZI_KURANG
133	38	8,9	87	GIZI_KURANG
134	36	9,8	66	GIZI_NORMAL
135	36	8,8	76	GIZI_NORMAL
136	35	7,9	86	GIZI_BURUK
137	39	9,4	86	GIZI_KURANG
138	42	5,7	67	GIZI_KURANG
139	22	6,7	55	GIZI_NORMAL
140	43	10	91	GIZI_KURANG
141	34	12	76	GIZI_NORMAL
142	23	11	78	GIZI_NORMAL
143	45	13	66	GIZI_LEBIH
144	38	3,5	65	GIZI_BURUK
145	12	4,7	68	GIZI_BURUK
146	23	5,8	89	GIZI_BURUK
147	42	7,9	57	GIZI_NORMAL
148	33	8,9	78	GIZI_KURANG
149	27	9	56	GIZI_LEBIH
150	45	10	56	GIZI_LEBIH
151	32	11	68	GIZI_NORMAL
152	23	12	77	GIZI_NORMAL
153	34	4,7	87	GIZI_BURUK
154	34	7,8	76	GIZI_KURANG
155	23	9,8	75	GIZI_NORMAL
156	45	8,6	65	GIZI_NORMAL
157	34	9	65	GIZI_NORMAL
158	23	10	68	GIZI_NORMAL
159	34	11	76	GIZI_NORMAL
160	37	13	89	GIZI_NORMAL

TABEL 2
DATA TESTING

No	Umur (bulan)	Berat (kg)	Tinggi (cm)
1	43	12	65
2	44	9,7	64
3	36	9,8	56
4	34	7,8	54
5	32	8,8	65
6	43	9,6	76
7	4	8,9	78
8	23	9,6	56
9	33	6,7	67
10	44	8,7	76
11	32	9,8	67
12	32	12	56
13	23	13	78
14	45	10	76
15	31	9,8	56
16	32	7,8	77
17	31	5,7	68
18	23	8,7	76
19	34	7,7	65
20	32	7,8	57
21	34	7,9	78
22	37	9,5	75
23	43	9	66
24	24	10	76
25	44	1	76
26	23	12	89
27	34	9,5	75
28	43	9,3	75
29	35	6,8	49
30	33	4,8	66
31	23	7,3	61
32	33	5,6	62
33	25	5,5	62
34	34	7,8	63
35	35	9,8	63
36	36	6,6	64
37	38	9	76
38	42	12	75
39	44	11	74
40	47	10	73

Data pada Gambar 1 dan 2 selanjutnya diolah dengan aplikasi Orange. Alur pengolahan pada aplikasi Orange dapat dilihat pada Gambar 2

Data yang digunakan untuk melakukan uji coba berjumlah 40 data. *Data Testing* dapat dilihat pada Tabel 2.



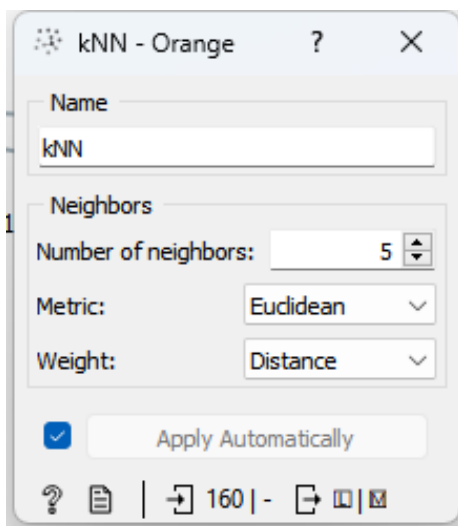
Gambar 2. Proses Klasifikasi Algoritma K-NN

Gambar 2 menunjukkan saat *data training* dan *testing* dimasukan, sebelum masuk tahap K-NN harus melewati *select colums*. Tujuannya untuk menentukan atribut apa saja yang dipakai. Penelitian ini menggunakan 3 atribut, yaitu umur, berat badan dan tinggi badan.

Tahap selanjutnya adalah mengatur algoritma K-NN pada aplikasi *Orange*. Jumlah K diatur ke 5, karena jumlah klasifikasi pada penelitian ini ada 4 kategori. Pengaturan lainnya dapat dilihat pada Gambar 3.

TABEL 3. HASIL PREDIKSI KLASIFIKASI MENGGUNAKAN ALGORITMA K-NN

No	Usia (bulan)	Berat badan (Kg)	Tinggi Badan (cm)	Status Gizi
1	43	12	65	GIZI_NORMAL
2	44	9,7	64	GIZI_NORMAL
3	36	9,8	56	GIZI_LEBIH
4	34	7,8	54	GIZI_LEBIH
5	32	8,8	65	GIZI_NORMAL
6	43	9,6	76	GIZI_NORMAL
7	4	8,9	78	GIZI_KURANG
8	23	9,6	56	GIZI_LEBIH
9	33	6,7	67	GIZI_NORMAL
10	44	8,7	76	GIZI_NORMAL
11	32	9,8	67	GIZI_NORMAL
12	32	12	56	GIZI_LEBIH
13	23	13	78	GIZI_NORMAL
14	45	10	76	GIZI_NORMAL
15	31	9,8	56	GIZI_LEBIH
16	32	7,8	77	GIZI_KURANG
17	31	5,7	68	GIZI_NORMAL
18	23	8,7	76	GIZI_NORMAL
19	34	7,7	65	GIZI_NORMAL
20	32	7,8	57	GIZI_NORMAL
21	34	7,9	78	GIZI_KURANG
22	37	9,5	75	GIZI_NORMAL
23	43	9	66	GIZI_NORMAL
24	24	10	76	GIZI_NORMAL
25	44	1	76	GIZI_KURANG
26	23	12	89	GIZI_NORMAL

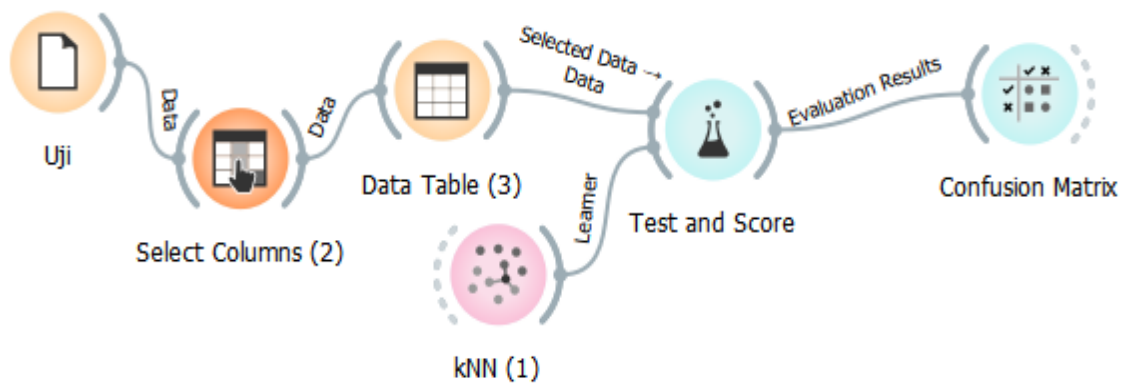


Gambar 3. Pengaturan algoritma KNN

Data testing yang telah diuji dengan algoritma K-NN, hasil prediksi klasifikasinya dapat dilihat pada Tabel 3.

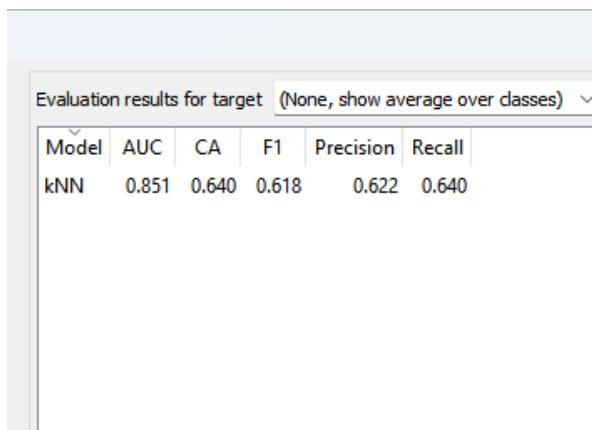
27	34	9,5	75	GIZI_NORMAL
28	43	9,3	75	GIZI_NORMAL
29	35	6,8	49	GIZI_LEBIH
30	33	4,8	66	GIZI_NORMAL
31	23	7,3	61	GIZI_NORMAL
32	33	5,6	62	GIZI_NORMAL
33	25	5,5	62	GIZI_NORMAL
34	34	7,8	63	GIZI_NORMAL
35	35	9,8	63	GIZI_NORMAL
36	36	6,6	64	GIZI_NORMAL
37	38	9	76	GIZI_NORMAL
38	42	12	75	GIZI_NORMAL
39	44	11	74	GIZI_NORMAL
40	47	10	73	GIZI_NORMAL

Data yang sudah diklasifikasi harus diperiksa dievaluasi akurasi. Metode yang digunakan adalah *Confusion Matrix*. Pemeriksaan tingkat akurasi dilakukan dengan cara menggabungkan *data training* dengan *data testing* yang sudah diklasifikasikan. Alur evaluasi *Confussion Matrix* pada aplikasi *Orange* dapat dilihat pada Gambar 4.



Gambar 4. Skema Confusion Matrix

Fitur *Test and Score* digunakan untuk melihat nilai *AUC*, *Recall*, *Precision* dan *F1 score*. Pengaturan K-NN sama seperti langkah sebelumnya. Hasil rata-rata keseluruhan dari *Test dan Score* dapat dilihat pada Gambar 5.



Model	AUC	CA	F1	Precision	Recall
kNN	0.851	0.640	0.618	0.622	0.640

Gambar 5. Hasil Test dan Score

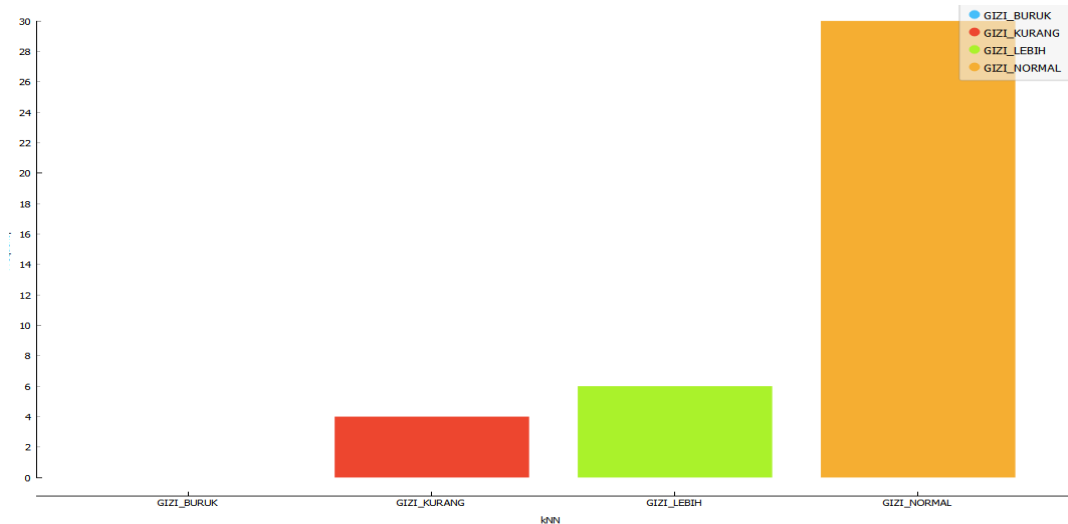
Sedangkan sebaran hasil evaluasi yang diterapkan pada *Confusion Matrix* dapat dilihat pada Gambar 6.

		Predicted				
		GIZI_BURUK	GIZI_KURANG	GIZI_LEBIH	GIZI_NORMAL	Σ
Actual	GIZI_BURUK	7	15	0	5	27
	GIZI_KURANG	6	27	0	21	54
	GIZI_LEBIH	1	0	14	10	25
	GIZI_NORMAL	0	8	4	82	94
Σ		14	50	18	118	200

Gambar 6. Confusion Matrix

IV. PEMBAHASAN

Berdasarkan hasil penelitian yang didapat, 40 dataset yang digunakan sebagai data *Testing* memiliki kelompok prediksi yang berbeda. Prediksi balita gizi normal memiliki frekuensi paling banyak. Sedangkan prediksi balita yang mengalami gizi buruk tidak ada. Distribusi gizi balita dalam setiap klasifikasi dapat dilihat pada Gambar 7.



Gambar 7. Grafik Klasifikasi K-NN

Evaluasi menunjukkan nilai AUC menghasilkan 85,1%, F1 Score 61,8%, Precision 62,2% dan Recall 64%. Untuk mengetahui akurasi dari klasifikasi dengan K-NN dapat dilihat dari nilai AUC nya. Berdasarkan tabel skala yang ada pada penelitian yang dilakukan oleh [11] pada tahun 2022, score yang didapat algoritma K-NN pada penelitian ini termasuk dalam kategori baik.

V. KESIMPULAN

Klasifikasi dengan menggunakan Algoritma K-NN dapat menghasilkan nilai akurasi yang baik. Artinya, K-NN dapat dimanfaatkan untuk melakukan klasifikasi pada gizi balita di masa yang akan datang, sehingga kondisi gizi pada balita bisa diantisipasi lebih awal. Hal tersebut dibuktikan dengan evaluasi menggunakan *Confusion Matrix* yang menghasilkan nilai tinggi pada score AUC.

DAFTAR PUSTAKA

- [1] D. Janner Lubis and G. Karunia Gusti, "Teknois : Jurnal Ilmiah Teknologi Informasi dan Sains [58] Penerapan Algoritma Naïve Bayes Untuk Penentuan Balita Penerima Makanan Tambahan (PMT) Berdasarkan Status Gizi Di Pos Pelayanan Terpadu (POSYANDU)," vol. 13, no. 1, pp. 58–66, 2023, doi: 10.36350/jbs.v13i1.
- [2] N. Rahmawati and Y. Novianto, "Klasifikasi Kondisi Gizi Balita Menggunakan Metode Naive Bayes (Studi Kasus Posyandu Melati IV)," 2020.
- [3] A. Z. Zami, O. Nurdiawan, and G. Dwilestari, "Klasifikasi Kondisi Gizi Bayi Bawah Lima Tahun Pada Posyandu Melati Dengan Menggunakan Algoritma Decision Tree," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, no. 3, p. 305, Mar. 2022, doi: 10.30865/json.v3i3.3892.
- [4] H. Hafizan and A. N. Putri, "Penerapan Metode Klasifikasi Decision Tree Pada Status Gizi Balita Di Kabupaten Simalungun," 2020.
- [5] R. Rizqi Robbi Arisandi, B. Warsito, and A. Rachman Hakim, "APLIKASI NAÏVE BAYES CLASSIFIER (NBC) PADA KLASIFIKASI STATUS GIZI BALITA STUNTING DENGAN PENGUJIAN K-FOLD CROSS VALIDATION," vol. 11, no. 1, pp. 130–139, 2022, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [6] R. Darma Rusdiyan Yusron and I. Machfud, "Klasifikasi dan Monitoring Status Gizi Balita Melalui Penerapan Metode Naïve Bayes Classification Berbasis GIS Classification and Monitoring of Toddler Nutrition Status Through Application of GIS-Based Naïve Bayes Classification Method," *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, vol. 4, no. 02, pp. 161–168, 2022.
- [7] A. Yoseva Simanjuntak and I. Septian Salomo Simatupang, "IMPLEMENTASI DATA MINING MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER UNTUK DATA KENAikan PANGKAT DINAS KETENAGAKERJAAN KOTA MEDAN," 2022. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [8] M. Yudhi Putra and D. Ismiyana Putri, "Pemanfaatan Algoritma Naïve Bayes dan K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Kelas XI.
- [9] G. Abdurrahman, "Jurnal Sistem dan Teknologi Informasi Klasifikasi Penyakit Diabetes Melitus Menggunakan Adaboost Classifier," vol. 7, no. 1, 2022, [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JUSTINDO>
- [10] M. Ula, R. Maulana, and V. Ilhadi, "Terbit online pada laman web jurnal: <https://ejurnalunsam.id/index.php/jicom/> Penerapan KNN Penentuan Pelanggan Baru PDAM dan Clustering K-Means Berdasarkan Wilayah", [Online]. Available: <https://ejurnalunsam.id/index.php/jicom/>



- [11] A. MUHARIYA, "Pengelompokan Komentar Pada Media Sosial Instagram Menggunakan Metode K-Means Clustering Untuk Identifikasi Awal Cyberbullying," 2022.
- [12] "Baby Nutrition Classification | Kaggle." <https://www.kaggle.com/datasets/mjalaluddinassuyuti/baby-nutrition-classification> (accessed Apr. 11, 2023).
- [13] J. Astri, J. Karman, and N. K. Daulay, "Prediksi Kelulusan Mahasiswa Menggunakan Metode K-Nearest Neighbor (KNN) pada Fakultas Ilmu Teknik, Univeritas Bina Insan," vol. 8, pp. 169–173, [Online]. Available: <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>