

## **KLASIFIKASI STATUS GIZI PADA PERTUMBUHAN BALITA MENGGUNAKAN K-NEAREST NEIGHBOR (KNN)**

Dewi Fitrianiingsih<sup>1</sup>, Martaleli Bettiza<sup>2</sup>, Alena Uperiati<sup>3</sup>  
dewi130199@gmail.com

Program studi Teknik Informatika Fakultas Teknik, Universitas Maritim Raja Ali Haji

### **Abstract**

*The nutritional status for babies is something that is very important for parents to know, because there are still many cases of under-five malnutrition in Indonesia that never go away. Especially in the Kijang area at the Sei Lekop Health Center, there are still many toddlers who have unbalanced nutrition, due to lack of parental knowledge about toddler nutrition. In this study, the K-Nearest Neighbor method was used to classify the nutritional status of toddlers and can also create a Nutritional Status Classification System for Toddler Growth using the KNN (K-Nearest Neighbor) Method. Toddler nutrition data uses 4 classifications, namely: More, Good, Less, Bad. The amount of data used for this study is 170 data, of which 80% are training data totaling 136 data, and 20% are test data totaling 34 data. The level of accuracy of the results of this test is to produce an accuracy rate of 73.53% using the K-Nearest Neighbor algorithm.*

*Keywords: K-Nearest Neighbor, Classification, Nutrition for Toddlers*

### **I. Pendahuluan**

Status gizi untuk balita adalah sesuatu hal yang sangat penting untuk diketahui oleh orang tua, karena masih banyak ditemukan kasus kekurangan gizi balita di Indonesia yang tak kunjung hilang. Anak yang mengalami masalah gizi pada usia dini akan mengalami masalah gangguan pada tumbuh kembang, dan mengalami gangguan kesakitan bahkan mengalami kematian. Oleh karena itu untuk mendapatkan kesehatan balita yang baik, salah satunya adalah dengan cara mengkonsumsi gizi yang cukup sesuai dengan kebutuhan pada tubuh balita, selain itu untuk mengetahui tingkat kesehatan balita dapat dilihat melalui status gizinya. Penilaian status gizi balita dapat ditentukan melalui pengukuran tubuh balita. Parameter yang digunakan diantaranya adalah Jenis Kelamin (JK), Umur (U), Berat Badan (BB), Tinggi Badan (TB), Lingkar Kepala (LK). Untuk menghilangkan kasus gizi buruk di daerah Kijang perlu adanya pendeteksi atau suatu sistem yang bisa membuat status gizi balita dapat terpantau dengan baik. Pada puskesmas Sei Lekop, parameter yang umum digunakan dalam penentuan status gizi balita adalah berdasarkan Jenis Kelamin, Umur, Berat Badan, Tinggi badan, Lingkar Kepala (BB/U). Maka dari itu perlu adanya pemantauan dari orangtua mengenai status gizi balita agar dapat terus memantau pertumbuhan dan kebutuhan zat gizi yang harus terpenuhi hingga mencapai status gizi optimal pada balita. Gizi pada balita masih sangat perlu pengawasan dari orangtua dan pihak terkait karena dapat mempengaruhi tumbuh kembang seorang balita. Metode yang digunakan dalam penentuan status gizi balita ini adalah metode KNN (*K-Nearest Neighbor*). Metode ini sering digunakan untuk kasus pengklasifikasian data. Sehingga metode ini cocok untuk diterapkan dalam pengklasifikasian status gizi balita.

### **II. Metode Penelitian**

#### **2.1 Data Mining**

Secara sederhana data *mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar (Davies, 2004). Data *mining* juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data (Pramudiono, 2007).

## 2.2 Klasifikasi

Klasifikasi adalah proses untuk menemukan atau menentukan sebuah model atau fungsi yang menjelaskan atau membedakan kelas data dengan tujuan untuk memperkirakan kelas yang tidak diketahui dari suatu objek. Klasifikasi dokumen adalah pemberian kategori yang telah didefinisikan kepada dokumen yang belum memiliki kategori. Mengklasifikasi dokumen merupakan salah satu cara untuk mengorganisasikan dokumen.

## 2.3 K-Nearest Neighbor (KNN)

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan obyek. Prinsip kerja dari *K-Nearest Neighbor* (KNN) adalah untuk mencari jarak terdekat antara data yang akan dievaluasi dengan  $k$  tetangga (*neighbor*) terdekatnya dalam data pelatihan. Ketepatan algoritma K-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur agar performa klasifikasi menjadi lebih baik. Sesuai dengan prinsip kerja *K-Nearest Neighbor* yaitu mencari jarak terdekat antara data yang akan dievaluasi dengan  $k$  tetangga (*neighbor*) terdekatnya dalam data pelatihan.

### Jarak Euclidean

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2}$$

Dengan keterangan :

$d$  = Jarak

$x_1$  = Data Latih

$x_2$  = Data Uji

$P$  = Dimensi data

$i$  = Variabel data

## 2.4 Status Gizi

Status gizi merupakan keadaan yang diakibatkan oleh status keseimbangan antara jumlah asupan (*intake*) zat gizi dan jumlah yang dibutuhkan (*requirement*) oleh tubuh untuk berbagai fungsi biologis (pertumbuhan fisik, perkembangan, aktivitas, pemeliharaan kesehatan, dan lainnya). Status gizi dapat pula diartikan sebagai gambaran kondisi fisik seseorang sebagai refleksi dari keseimbangan energi yang masuk dan yang dikeluarkan oleh tubuh.

## 2.5 Confusion Matrix

*Confusion matrix* merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya *confusion matrix* mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya.

Selanjutnya dari matriks dapat ditarik kesimpulan dari hasil klasifikasi berupa nilai *accuracy* dan *error*. *Accuracy* adalah rasio dari jumlah data ketepatan prediksi benar dari tiap kelas terhadap jumlah total dari semua data. Rumus *accuracy* ditunjukkan pada rumus dibawah berikut ini:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

*Error* adalah rasio dari jumlah data kesalahan prediksi yang salah dari tiap kelas terhadap jumlah total semua data. Rumus *error* ditunjukkan pada rumus dibawah berikut ini:

$$Error = \frac{FP + FN}{TP + FP + FN + TN} \times 100\%$$

Dengan keterangan:

TP : True Positive

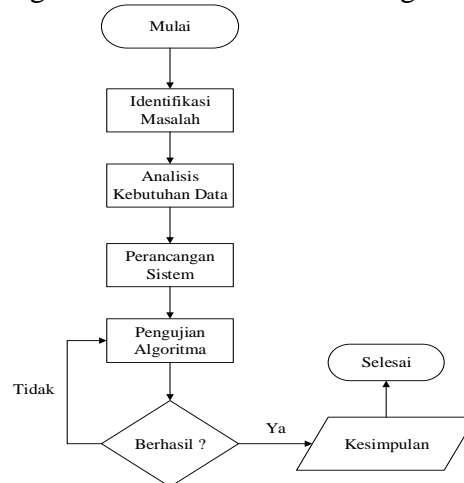
TN : True Negative

FP : False Positive

FN : False Negative

## 2.4 Kerangka Pikir Penelitian

Metode penelitian merupakan tahap-tahap yang dilakukan pada penelitian. Berikut ini merupakan tahapan penelitian yang digambarkan dalam bentuk diagram alir pada penelitian ini.



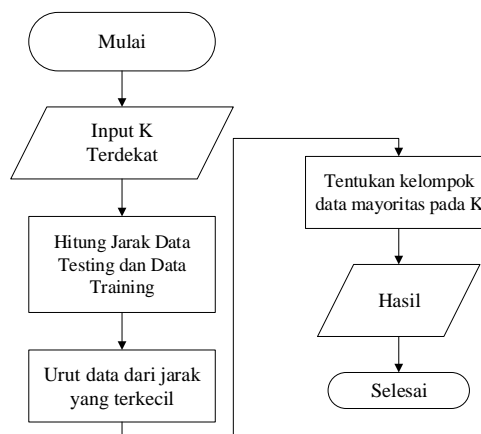
Gambar 1. Kerangka Pikir Penelitian

Deskripsi gambar 1:

1. Mulai
2. Identifikasi masalah yang mungkin akan terjadi.
3. Analisa kebutuhan data yang diperlukan untuk menentukan status gizi balita di Puskesmas Sei Lekop Kijang.
4. Perancangan sistem, merencanakan sistem yang akan dibuat.
5. Pengujian algoritma, melakukan pengujian terhadap algoritma bertujuan untuk melihat apakah algoritma cocok dengan data yang digunakan.
6. Kesimpulan, mengambil hasil dari kesimpulan.
7. Selesai.

## 2.5. Perancangan Algoritma

Pada perancangan algoritma *K-Nearest Neighbor*, menjelaskan terkait dengan proses data mining untuk mencari jarak terdekat dan merupakan nilai mayoritas untuk dijadikan sebagai hasil akhir. Digambarkan dalam bentuk flowchart untuk tahapan dalam perancangan algoritma KNN, dalam *flowchart* berikut ini berisi langkah – langkahnya dalam proses algoritma KNN, yaitu adalah:

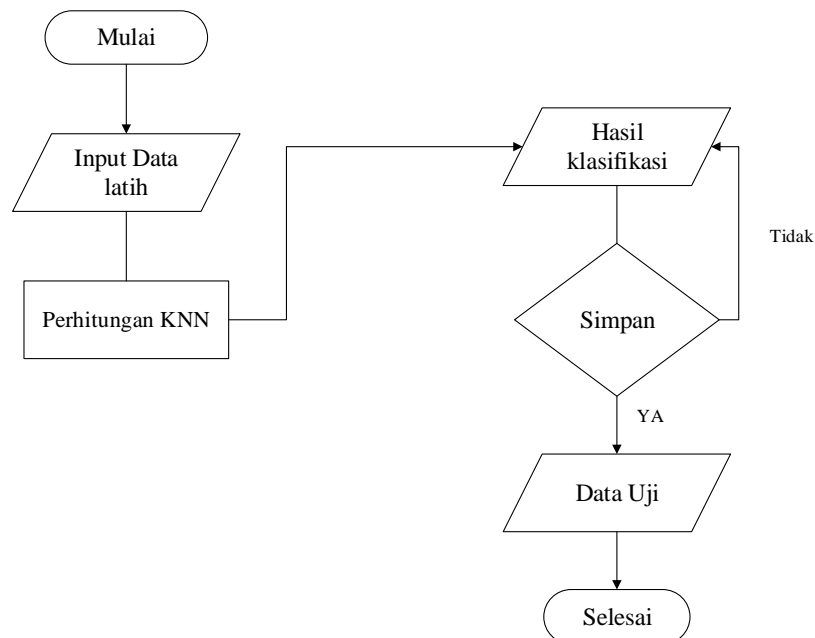


Gambar 2. Flowchart Algoritma *K-Nearest Neighbor*

Adapun penjelasan dari gambar 2 adalah sebagai berikut:

- a. *Input k Terdekat*  
K adalah jumlah tetangga terdekat yang letaknya paling dekat dengan objek data *testing* yang diuji. Berdasarkan nilai K inilah klasifikasi objek baru pada data *testing* dapat ditentukan dan akan masuk ke kelas yang telah ada nantinya.
- b. *Hitung Jarak Data Testing dan Data Training*  
Menghitung jarak Euclidean antara Data Testing dan Data Training untuk mencari peringkat dalam k yang telah ditentukan dan berfungsi untuk menguji ukuran yang dapat digunakan sebagai interpretasi kedekatan antara dua objek.
- c. *Urut Data Dari yang Terkecil*  
Hasil perhitungan yang diurutkan adalah hasil dari perhitungan jarak *Euclidean*. Pengurutan ini dilakukan dari yang terkecil hingga yang paling besar nilainya.
- d. *Tentukan Kelompok Data Mayoritas pada k*  
Menentukan kelompok data mayoritas pada k untuk menentukan kelas objek berdasarkan k yang ditentukan.

## 2.6 Perancangan Alur Sistem



Gambar 3. Alur Sistem

Adapun penjelasan dari gambar 3 adalah sebagai berikut:

- a. Mulai.
- b. Input data latih, memasukan data latih.
- c. Perhitungan dengan algoritma KNN.
- d. Hasil klasifikasi dari perhitungan yang dilakukan.
- e. Simpan hasil klasifikasi
- f. Data uji, hasil data yang telah diuji.
- g. Selesai

## III. Hasil dan Pembahasan

### 3.1 Hasil Penelitian

Hasil pengujian klasifikasi status gizi balita dengan menggunakan data sebanyak 170 data, dengan 80% sebagai data training berjumlah 136 data dan 20% sebagai data testing berjumlah 34 data, dengan menggunakan nilai  $k = 5$  maka didapatkan hasil akurasi adalah sebesar 73.53% dan hasil error adalah sebesar 26.47%.

Tabel 1. Hasil Akurasi

No	Status Gizi	Status Gizi Sistem	Hasil
1	Lebih	Lebih	Sama
2	Buruk	Buruk	Sama
3	Baik	Baik	Sama
4	Lebih	Lebih	Sama
5	Kurang	Kurang	Sama
..	...	...	...
..	...	...	...
32	Kurang	Buruk	Tidak sama
33	Kurang	Kurang	Sama
34	Kurang	Baik	Tidak sama

Ini adalah tabel perbandingan hasil akurasi antara k yang sudah diuji akurasi pada tabel-tabel diatas sebagai berikut:

Tabel 2. Tabel Perbandingan Hasil Akurasi

k	Akurasi
k=3	73,53%
k=5	73,53%
k=7	67,65%
k=9	67,65%
k=11	64,71%
k=13	55,88%
k=15	64,71%
k=17	67,65%

#### IV. Kesimpulan

Kesimpulan yang dapat diambil dari hasil penelitian tentang Klasifikasi Status Gizi Pada Pertumbuhan Balita Menggunakan *K-Nearest Neighbor* adalah yaitu dari hasil pengujian klasifikasi status gizi balita dengan menggunakan data sebanyak 170 data, dengan 136 sebagai data training dan 34 sebagai data testing lalu nilai  $K = 5$  didapatkan hasil akurasi dari data yang diuji adalah sebesar 73.53% dan hasil error dari data yang diuji adalah sebesar 26.47%. Setelah membandingkan tingkat akurasi antara  $k=3$ ,  $k=5$ ,  $k=7$ ,  $k=9$ ,  $k=11$ ,  $k=13$ ,  $k=15$  dan  $k=17$ , maka diketahui bahwa tingkat akurasi tertinggi diperoleh dari  $k=3$  dan  $k=5$  yaitu sebesar 73.53%.

#### V. Daftar Pustaka

- Kaesmetan, Y. R., & Johannis, J. A. (2017). Klasifikasi Status Gizi Balita Di Kelurahan Oesapa Barat Menggunakan Metode K-Nearest Neighbor. *Multitek Indonesia*, 11(1), 42. <https://doi.org/10.24269/mtkind.v11i1.506>
- Kandasamy, J. P., & Balamurali, S. (2015). Performance analysis of classifier models to predict diabetes mellitus. *Procedia Computer Science*, 47(C), 45–51. <https://doi.org/10.1016/j.procs.2015.03.182>

- Putri, N. E. A., Syauqy, D., & Hanafi, M. H. (2017). Sistem Klasifikasi Status Gizi Bayi dengan Metode K- Nearest Neighbor Berbasis Sistem Embedded. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(9), 933–939
- Saleh, H., Faisal, M., & Musa, R. I. (2019). Klasifikasi Status Gizi Balita Menggunakan Metode K- Nearest Neighbor. *Jurnal Sistem Informasi Dan Teknik Komputer*, Vol. 4(No. 2), Hal. 120-126.
- Hasan, S. A., Suhada, S., & Hadjaratie, L. (2015). *Klasifikasi Status Gizi Menggunakan K-Nearest Neighbor*. 3(1), 6–11.
- Kibanov, M., Becker, M., Mueller, J., Atzmueller, M., Hotho, A., & Stumme, G. (2018). Adaptive kNN using expected accuracy for classification of geo-spatial data. *Proceedings of the ACM Symposium on Applied Computing*, 857–865. <https://doi.org/10.1145/3167132.3167226>
- Mardalena, I., & Suyani, E. (2016). Keperawatan Ilmu Gizi. *Kementeriann Kesehatan Republik Indonesia*, 182.
- Fibrianda, M. F., & Bhawiyuga, A. (2018). Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(9), 3112–3123.
- Purnamasari, I., Deny, F., & Amijaya, T. (2019). *KLASIFIKASI STATUS PEMBAYARAN PREMI MENGGUNAKAN ALGORITMA NEIGHBOR WEIGHTED K-NEAREST NEIGHBOR ( NWKNN ) ( STUDI KASUS : PT . BUMIPUTERA KOTA SAMARINDA ) Classification of Premium Payment Status Using Neighbor Weighted K-Nearest Neighbor ( NWKNN ) Algorithm ( Case Study : PT . Bumiputera Kota Samarinda )*. 1, 56–63.

## Ucapan Terimakasih

Penulis mengucapkan terima kasih yang sebesar-besarnya dan penghargaan yang mulia kepada :

1. Allah SWT yang memberikan Anugerah yang luar biasa kepada Penulis.
2. Kedua orang tua Bapak Subiyanto dan Ibu Siti Halimah, Kakak Eka Andriani dan Keluarga Besar.
3. Ibu Nurul Hayaty, S.T., M.Cs. selaku ketua Prodi Jurusan Teknik Informatika dan seluruh dosen pengajar Jurusan Teknik Informatika Umrah.
4. Ibu Martaleli Bettiza, S.Si., M.Sc. selaku dosen pembimbing I yang telah memberikan arahan, dukungan dan semangat kepada Penulis.
5. Ibu Alena Uperiati, S.T., M.Cs. selaku dosen pembimbing II yang telah memberikan arahan, dukungan dan semangat kepada Penulis.
6. “Para Pejuang Skripsi Angkatan 2016” (Girls Basecamp)