

## Rancangan Sistem Klasifikasi Kekurangan Gizi Balita dengan Metode K-Nearest Neighbor

Syahrani Lonang<sup>1\*</sup>, Anton Yudhana<sup>2</sup>, Muhammad Kunta Biddinika<sup>3</sup>

<sup>1,3</sup> Magister Informatika, Fakultas TEKNOLOGI INDUSTRI, Universitas Ahmad Dahlan

<sup>2</sup> Jurusan TEKNIK ELEKTRO, Fakultas TEKNOLOGI INDUSTRI, Universitas Ahmad Dahlan

\* Email: lonangsyahrani3@gmail.com

### Abstrak

Kekurangan gizi pada balita (*malnutrisi*) merupakan masalah serius yang dihadapi negara berkembang seperti Indonesia, efek jangka panjang yang diakibatkan dapat menurunkan kecerdasan balita. Pengklasifikasian status gizi balita masih dilakukan secara konvensional di pusat kesehatan masyarakat. Algoritma K-Nearest Neighbor termasuk dalam algoritma pembelajaran mesin yang dapat dimanfaatkan untuk melakukan klasifikasi salah satunya masalah klasifikasi status gizi. K-NN digunakan sebagai algoritma penentuan kelas dari data baru yang akan diinputkan sesuai format. Penelitian ini dimulai dengan studi pustaka, lalu mengidentifikasi kebutuhan, dilanjutkan pengumpulan data yang rencananya akan digunakan dalam sistem yang akan dibangun juga menjadi acuan membuat rancangan dan tahapan akhir perancangan sistem. Penelitian ini berhasil membuat sebuah rancangan sistem dengan menggunakan Unified Model Language (UML), satu use case yang berisikan empat functional sistem di antaranya file upload dataset, menampilkan dataset, menguji akurasi terhadap dataset, memprediksi data baru dan rancangan antarmuka sistem yang nantinya mempermudah pembuatan sistem.

**Kata kunci:** *malnutrisi, k-nn, uml, rancangan*

### Abstract

Malnutrition in toddlers is a serious problem faced by developing countries like Indonesia, and the resulting long-term effects can reduce the intelligence of toddlers. The classification of the nutritional status of children under five is still carried out conventionally in community health centers. The K-Nearest Neighbor algorithm is included in a machine learning algorithm that can be used to classify one of the nutritional status classification problems. K-NN is used as a class determination algorithm for new data to be input according to the format. This research begins with a literature study, then identifies needs, followed by data collection that is planned to be used in the system to be built as well as a reference for making the design and the final stage of system design. This research succeeded in creating a system design using the Unified Model Language (UML), one use case that contains four functional systems, including uploading dataset files, displaying datasets, testing the accuracy of datasets, predicting new data, and designing system interfaces that will make system development easier..

**Keyword:** *malnutrition, k-nn, uml, design*

## PENDAHULUAN

Ada dua masalah gizi yang dihadapi dunia dalam lingkup rumah tangga maupun individu adalah kekurangan gizi dan obesitas (Lowe et al., 2021). Indonesia adalah salah satu negara yang memiliki tantangan terkait permasalahan gizi yaitu gizi kurang, bayi usia dibawah lima tahun (balita) akan mengalami dampak serius bagi kesehatannya baik dalam jangka pendek maupun panjang jika mengalami permasalahan gizi. Masalah gizi juga adalah

salah satu faktor penyebab kematian anak di negara-negara berkembang (Utami & Mubasyiroh, 2019). Efek kekurangan gizi pada balita seringkali menyebabkan infeksi kronis dan rentan peradangan, penurunan IQ (*Intelligence Quotien*) juga penurunan kepercayaan diri adalah efek jangka panjang yang diakibatkan kekurangan gizi (Husna & Izzah, 2021). Berdasarkan Pantauan Status Gizi Tahun 2017 oleh Kemenkes RI (Kementerian Kesehatan Republik Indonesia) prevalensi

masalah gizi balita Indonesia mencapai 17,8% dimana 14% adalah masalah gizi kurang. Di Indonesia pengklasifikasian status gizi balita masih dilakukan secara manual yang rentan akan ketidaktepatan juga memakan waktu lebih lama (Lubna Irkanisa, Cholissodin, & Abdurrachman Bachtiar, 2019). Status gizi balita dapat dinilai dengan mengukur tubuh yang dikenal sebagai antropometri, dengan usia, jenis kelamin, berat serta tinggi badan balita sebagai indikator pengukuran (Hafizan & Putri, 2020). Oleh karena itu diperlukan perancangan sistem guna mempermudah pembuatan sistem yang dapat mengklasifikasi/memprediksi status gizi secara cepat dan akurat guna melakukan diagnosis dini pada balita.

Klasifikasi merupakan suatu proses penempatan sebuah objek spesifik ke himpunan kategori yang ada, berdasarkan masing-masing property objek. Proses klasifikasi didasarkan pada empat komponen mendasar yaitu *class*, *predictor*, *training dataset* dan *testing dataset* (Becker, 2015). K-Nearest Neighbor (K-NN) merupakan salah satu metode klasifikasi dengan melakukan pengklasifikasian berdasarkan jarak terdekat antara data latih dengan objek, dalam K-NN perhitungan jarak antar data akan dilakukan menggunakan *Euclidean Distance* (Yudhana, Muslim, et al., 2020).

Dalam topik penambangan data K-NN adalah salah satu algoritma yang paling banyak diimplementasikan dalam penelitian (Koncz & Paralic, 2011). K-NN bisa memberikan nilai akurasi yang baik menggunakan jumlah dataset yang besar, hal itu merupakan kelebihan dari K-NN. Namun kekurangannya, K-NN sensitif terhadap atribut yang tidak relevan (Lonang & Normawati, 2022).

Pada penelitian sebelumnya K-NN dimanfaatkan untuk melakukan prediksi pada hasil penggergajian kayu sengon yang dilakukan oleh Anton Yudhana, Sunardi dan Agus Jaka Sri Hartanta. *Dataset* yang digunakan berjumlah 135 sebagai data latih dan sepuluh data uji. Didapatkan hasil prediksi dengan nilai akurasi 70% setelah membangun sistem berbasis *framework* laravel (Yudhana, Sunardi, & Hartanta, 2020).

K-NN juga diimplementasikan dalam penelitian untuk mengklasifikasikan kanker payudara, fitur seleksi yang digunakan berbasis *chi square*. Ada 2 *dataset* yang dipakai yaitu WDBC dan WBC *dataset* bersumber dari UCI *Repository*. Akurasi tertinggi yang didapatkan

sebesar 98.85% jika nilai *k* yang digunakan adalah 3, 99.42% jika *k* = 1 dan 94.86% jika *k* = 4. Nilai *k* paling optimal pada penelitian ini dari 1-9. Dan berdasarkan hasil penelitian *chi square* memberikan peningkatan pada akurasi dikomparasikan dengan model tanpa fitur seleksi (Mushtaq, Yaqub, Sani, & Khalid, 2020).

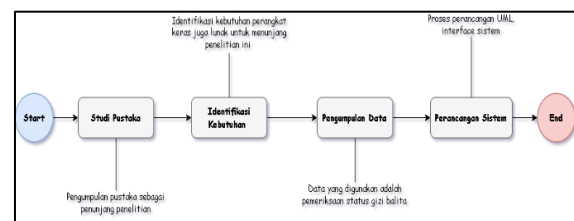
Pada penelitian lain metode K-NN dimanfaatkan untuk melakukan klasifikasi sebagai alat pendukung implantasi simulator otak pada pasien pengidap Parkinson. Penelitian ini menggunakan data neuronal electrical activity yang diperoleh dari operasi Deep Brain Stimulator (DBS). Hasilnya menunjukkan bahwa K-NN menggunakan data yang distandarisasi menghasilkan kinerja terbaik dengan akurasi 94.35% (Bellino, Schiaffino, Battisti, Guerrero, & Rosado-Muñoz, 2019).

Penelitian lain dengan mengimplementasikan seleksi fitur *forward selection* dan K-NN untuk memprediksi bedah toraks menggunakan dataset thoracic surgery yang bersumber dari UCI Repository dengan record berjumlah 470, 16 variabel satu kelas. Penelitian yang dilakukan menghasilkan nilai akurasi tertinggi sebesar 83,40% tanpa seleksi fitur dan meningkat 2,34% menjadi 85,73 % setelah dioptimasi menggunakan *forward selection* (Sanjaya & Fitriyani, 2019).

Penelitian ini bertujuan untuk merancang sistem klasifikasi kekurangan gizi balita guna membantu peneliti membangun sistem yang mengacu pada rancangan yang ada dikarenakan penelitian terdahulu yang sudah disebutkan diatas belum membangun sebuah yang mengimplementasikan algoritma K-NN.

**METODE**

Adapun alur penelitian yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

**2.1 Studi Pustaka**

Studi pustaka adalah aktivitas penelitian yang dilakukan guna mengumpulkan pustaka sebagai alat penunjang penelitian yang bersumber dari buku, paper ataupun artikel yang berkaitan pada penelitian ini (Melinda & Zainil, 2020).

**2.2 Identifikasi Kebutuhan**

Identifikasi kebutuhan diperlukan untuk membantu dalam melakukan penelitian yang diperoleh dari studi literatur. Terdapat dua kebutuhan yang akan dipetakan yaitu kebutuhan perangkat keras dan kebutuhan perangkat lunak.

**2.3 Pengumpulan Data**

Data yang digunakan sebagai parameter perancangan sistem pada penelitian ini adalah data pemeriksaan gizi balita Puskesmas Ubung, Jonggat, Lombok Tengah. Ada sembilan atribut yaitu jenis kelamin, umur, berat badan, tinggi badan, berat badan menurut tinggi (bb/tb), z-score bb/tb, tinggi badan menurut umur (tb/u), z-score tb/u, z-score berat badan menurut umur (bb/u) dan kelas yaitu 0 = tidak kekurangan gizi dan 1 = kurang gizi

**2.4 Perancangan Sistem**

Pada proses ini akan dibuat suatu perancangan dari sistem klasifikasi kurang gizi balita dengan memanfaatkan Unified Modeling Language (UML), UML merupakan bahasa grafis yang berfungsi untuk merepresentasikan juga mengomunikasikan aspek yang berbeda dalam sistem (Abdelmadjid & Mimoun, 2021). Use case diagram sebagai deskripsi kebutuhan fungsional, activity diagram sebagai visualisasi alur pada sistem (Widyatmoko & Pamungkas, 2022). Use case adalah diagram dalam UML yang menggambarkan kebutuhan fungsional dari sistem dengan bentuk oval, setiap oval dalam diagram dapat diperluas lebih lanjut sebagai Activity diagram. Fungsi activity diagram adalah menguraikan secara spesifik perilaku sistem (Syarif & Nugraha, 2020).

**2.5 K-Nearest Neighbor**

K-NN merupakan algoritma sederhana berbasis jarak yang sering digunakan untuk melakukan klasifikasi, dikarenakan termasuk *lazy learning* kelebihananya K-NN tidak

membutuhkan pengetahuan dan melestarikan model yang diberikan. Oleh karena itu model dapat beradaptasi dengan perubahan cepat (Atallah, Badawy, El-Sayed, & Ghoneim, 2019). K-NN memiliki prinsip kerja yaitu menghitung jarak titik data baru dengan titik yang paling dekat dengannya. Ada beberapa fungsi jarak yang dapat digunakan seperti *Euclidean*, *Cosine*, *Manhattan* dimana nilai k adalah bilangan bulat acak (Abbad Ur Rehman, Lin, & Mushtaq, 2021). Perhitungan jarak menggunakan fungsi Euclidean yang ditunjukkan pada persamaan 1 sebagai berikut.

$$d(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (1)$$

Keterangan:

- d(a,b) = jarak
- i = variabel data
- n = dimensi data
- ai = data latih
- bi = data uji

**HASIL DAN PEMBAHASAN**

Dalam perancangan sistem ini akan dibuat analisis kebutuhan baik dari sisi perangkat keras dan perangkat lunak, UML digunakan untuk membuat use case diagram dan activity diagram untuk mendeskripsikan sistem yang akan dibuat juga perancangan desain antarmuka.

**3.1 Identifikasi Kebutuhan**

Tabel 1. Kebutuhan Perangkat Keras

Hardware	Detail
Processor	Intel Core i5 Ge 8
RAM	8 GB
HDD	1 TB

Dari Tabel 1 dapat dilihat daftar perangkat keras yang digunakan untuk menunjang penelitian ini dengan laptop yang digunakan adalah Acer aspire E14 E5-476G dengan RAM 8 GB dan media penyimpanan HDD sebesar 1 TB.

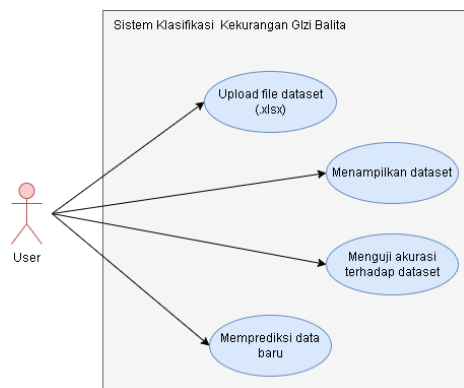
Tabel 2. Kebutuhan Perangkat Lunak

Software
Windows 10 Home Single Language
Microsoft Office 365
Canva
Draw io

Dari Tabel 2 terdapat beberapa perangkat lunak yang mendukung penelitian ini diantaranya *microsoft office* digunakan sebagai aplikasi pengolah kata, *canva* untuk melakukan desain antar muka sistem yang akan dibuat, dan *draw io* untuk pembuatan *use case diagram*, *activity diagram*.

### 3.2 Use Case Diagram

Pada tahapan ini akan dibuat *use case diagram* seperti yang ditunjukkan dalam Gambar 2 sebagai deskripsi *functional requirements*.

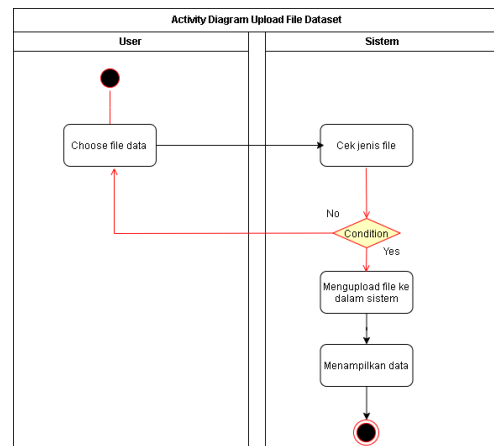


Gambar 2. Use Case Diagram

*Use case diagram* digunakan untuk merepresentasikan fungsionalitas dari sistem yang akan dibangun dimana terdapat empat proses diantaranya, upload file yang berfungsi untuk mengunggah *dataset* yang akan digunakan dengan ekstensi file *xlsx*. Melihat *dataset* proses ini menampilkan data yang sudah ada pada *database* berupa tabel. Menguji akurasi, dalam proses ini sistem akan melakukan split data menjadi data latih dan data uji untuk diklasifikasi menggunakan algoritma K-NN lalu menampilkan hasil *accuracy*, *precision*, *recall*, *f1-score* serta hasil validasi dengan *cross validation*. Dan yang terakhir adalah prediksi data, data baru akan diklasifikasikan berdasarkan jarak data satu dengan yang lainnya.

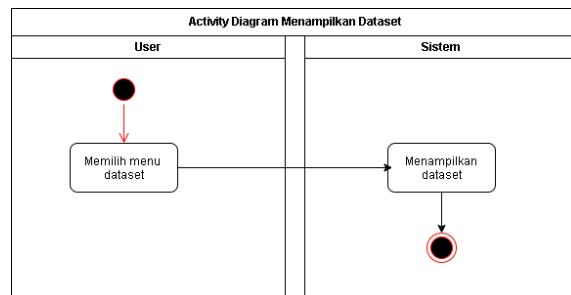
### 3.3 Activity Diagram

*Activity diagram* digunakan sebagai penggambaran aktivitas pada sistem yang akan dibuat seperti pada Gambar 3 menjelaskan skema diantara aktor dengan sistem saat melakukan proses pemilihan file *dataset*, jenis file yang dapat di unggah pada sistem ini hanya file yang berekstensi *xlsx*.



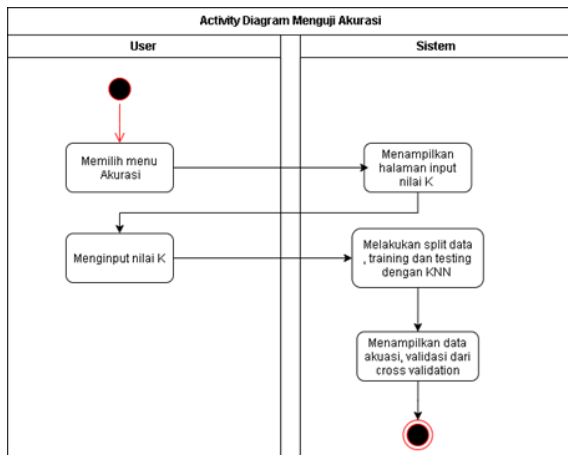
Gambar 3. Activity Diagram Upload File

Gambar 4 mendeskripsikan proses jika user ingin menampilkan *dataset* yang ada dalam sistem dengan cara user memilih menu *dataset* dan sistem akan menampilkannya dalam bentuk tabel pada layar *user*.



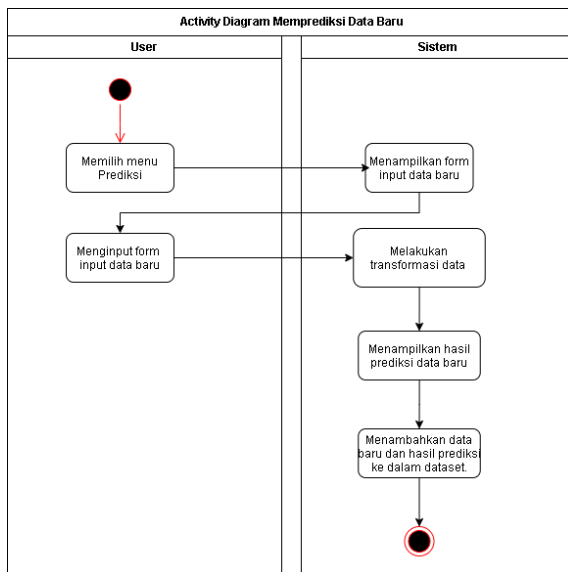
Gambar 4. Activity Diagram Menampilkan Dataset

Perancangan halaman pengujian akurasi dapat dilakukan *user* dengan memilih menu akurasi, lalu sistem akan meminta *user* memasukkan nilai bilangan bulat untuk menjadi nilai *k*. *Dataset* yang ada akan dilatih menggunakan algoritma K-NN. seperti yang ditunjukkan *activity diagram* pada Gambar 5.



Gambar 5. Activity Diagram Menguji Akurasi

User juga dapat melakukan prediksi kelas data baru seperti yang digambarkan pada activity diagram Gambar 6 dengan menginputkan nilai atribut yang ada. Jika berhasil sistem akan menampilkan prediksi kelas data baru yang ada.



Gambar 6. Activity Diagram Memprediksi Data Baru

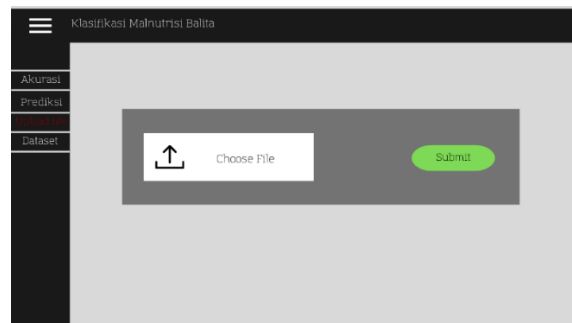
### 3.4 Desain User Interface

Pada Gambar 7 terdapat rancangan halaman yang akan menampilkan dataset serta klasifikasi dari data tersebut. Jika user melakukan prediksi data baru, maka data yang diinputkan tersimpan dalam database dan dapat dilihat pada halaman ini.

Akurasi	Jenis Kelamin	Umur (Bulan)	Berat Badan (kg)	Tinggi Badan (cm)	Laki	Perempuan	BA	z-score BA	BMI	Malnutrisi (Kategori)
Prediksi	Laki-laki	33	12	667	2	-129	2	-059	-066	0
Upload file	Perempuan	44	67	61	0	-502	1	-471	-580	1

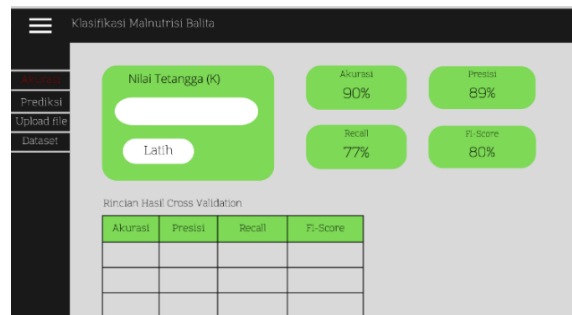
Gambar 7. Halaman Dataset

Gambar 8 merupakan rancangan halaman *upload* yang berfungsi untuk mengunggah data ke dalam database yang akan digunakan untuk *dataset* sistem ini. File yang diunggah diharuskan berekstensi *.xlsx* dengan nama dan jumlah atribut yang ditentukan.



Gambar 8. Halaman Upload Dataset

Gambar 9 merupakan rancangan halaman yang digunakan user untuk melakukan training dan evaluasi terhadap sistem klasifikasi menggunakan algoritma K-NN. User akan diminta untuk menginputkan nilai k bilangan bulat acak. Sistem akan menampilkan hasil akurasi, presisi, *recall*, *f1-score*.



Gambar 9. Halaman Akurasi

Gambar 10 merupakan rancangan antarmuka halaman prediksi yang berfungsi untuk melakukan prediksi kelas data baru yang diinputkan oleh *user* sesuai variabel yang ada dan sesuai dengan format yang sudah ditentukan. Jika berhasil, sistem akan menampilkan hasil prediksi.



Gambar 10. Halaman Prediksi

### SIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka diperoleh kesimpulan bahwa penelitian ini telah berhasil membuat rancangan sistem klasifikasi kekurangan gizi pada balita guna mempermudah dan sebagai acuan dalam pembuatan sistem. Rancangan sistem dibuat berdasarkan *unified modelling language* yang terdiri dari satu *use case diagram* berisikan empat fungsi terdiri dari file *upload dataset*, menampilkan *dataset*, menguji akurasi terhadap *dataset*, memprediksi data baru dan *activity diagram*. Terdapat juga rancangan antarmuka untuk membantu dalam membangun sistem klasifikasi kekurangan gizi pada balita. Sistem ini dirancang untuk dibuat dalam bentuk *platform web*, sehingga tampilan antarmuka masih dapat dikembangkan lebih lanjut. Setelah mendapatkan acuan rancangan, peneliti akan melanjutkan penelitian dengan mencoba mengimplementasikan hasil rancangan ini untuk membuat website klasifikasi dengan menggunakan python sebagai programming language, dan mysql sebagai database yang digunakan.

### DAFTAR PUSTAKA

Abbad Ur Rehman, H., Lin, C. Y., & Mushtaq, Z. (2021). Effective K-Nearest Neighbor Algorithms Performance Analysis of Thyroid Disease. *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers, Series A*, 44(1), 77–87.

<https://doi.org/10.1080/02533839.2020.1831967>

Abdelmajid, L., & Mimoun, M. (2021). Uncertain Decision-Making Requirements Formalizing with Complement Fuzzy UML Model. *Procedia Computer Science*, 198, 317–322. <https://doi.org/10.1016/j.procs.2021.12.247>

Atallah, D. M., Badawy, M., El-Sayed, A., & Ghoneim, M. A. (2019). Predicting kidney transplantation outcome based on hybrid feature selection and KNN classifier. *Multimedia Tools and Applications*, 78(14), 20383–20407. <https://doi.org/10.1007/s11042-019-7370-5>

Becker, F. G. (2015). *Data Mining Concepts, Models and Techniques*.

Bellino, G. M., Schiaffino, L., Battisti, M., Guerrero, J., & Rosado-Muñoz, A. (2019). Optimization of the KNN supervised classification algorithm as a support tool for the implantation of deep brain stimulators in patients with Parkinson'S Disease. *Entropy*, 21(4). <https://doi.org/10.3390/e21040346>

Hafizan, H., & Putri, A. N. (2020). Penerapan Metode Klasifikasi Decision Tree Pada Status Gizi Balita Di Kabupaten Simalungun. *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, 1(2), 68–72. <https://doi.org/10.30645/kesatria.v1i2.23>

Husna, L. N., & Izzah, N. (2021). Gambaran Status Gizi Pada Balita: Literature Review. *Prosiding Seminar Nasional Kesehatan*, 1, 385–392. <https://doi.org/10.48144/prosiding.v1i1.689>

Koncz, P., & Paralic, J. (2011). An approach to feature selection for sentiment analysis. *INES 2011 - 15th International Conference on Intelligent Engineering Systems, Proceedings*, 357–362. <https://doi.org/10.1109/INES.2011.5954773>

Lonang, S., & Normawati, D. (2022). Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination. *Jurnal Media Informatika Budidarma*, 6(1), 49–56.

Lowe, C., Kelly, M., Sarma, H., Richardson, A., Kurscheid, J. M., Laksono, B., ...

- Gray, D. J. (2021). The double burden of malnutrition and dietary patterns in rural Central Java, Indonesia. *The Lancet Regional Health - Western Pacific*, 14. <https://doi.org/10.1016/j.lanwpc.2021.100205>
- Lubna Irkanisa, N., Cholissodin, I., & Abdurrachman Bachtiar, F. (2019). Klasifikasi Status Gizi pada Balita Menggunakan Metode Extreme Learning Machine dan Algoritme Genetika. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(4), 3640–3646.
- Melinda, V., & Zainil, M. (2020). Penerapan model project based learning untuk meningkatkan kemampuan komunikasi matematis siswa sekolah dasar (studi literatur). *Jurnal Pendidikan Tambusai*, 4(2), 1526–1539.
- Mushtaq, Z., Yaqub, A., Sani, S., & Khalid, A. (2020). Effective K-nearest neighbor classifications for Wisconsin breast cancer data sets. *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers, Series A*, 43(1), 80–92. <https://doi.org/10.1080/02533839.2019.1676658>
- Sanjaya, R., & Fitriyani, F. (2019). Prediksi Bedah Toraks Menggunakan Seleksi Fitur Forward Selection dan K-Nearest Neighbor. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(3), 316. <https://doi.org/10.26418/jp.v5i3.35324>
- Syarif, M., & Nugraha, W. (2020). Pemodelan Diagram UML Sistem Pembayaran Tunai Pada Transaksi E-Commerce. *Jurnal Teknik Informatika Kaputama (JTIK)*, 4(1), 70 halaman. Retrieved from <http://jurnal.kaputama.ac.id/index.php/JTIK/article/view/240>
- Utami, H. N., & Mubasyiroh, R. (2019). Masalah Gizi Balita Dan Hubungannya Dengan Indeks Pembangunan Kesehatan Masyarakat (Nutritional Problems Among Underfive Children and It'S Relationship With Public Health Development Index). *Jurnal Penelitian Gizi Dan Makanan*, 42(1), 10.
- Widyatmoko, W., & Pamungkas, N. (2022). Pemodelan Unified Modeling Language pada Sistem Aplikasi Pariwisata (SiAP). *Jurnal Bumigora Information Technology (BITE)*, 4(1), 73–84. <https://doi.org/10.30812/bite.v4i1.1871>
- Yudhana, A., Muslim, A., Wati, D. E., Puspitasari, I., Azhari, A., & Mardhia, M. M. (2020). Human emotion recognition based on EEG signal using fast fourier transform and K-Nearest neighbor. *Advances in Science, Technology and Engineering Systems*, 5(6), 1082–1088. <https://doi.org/10.25046/aj0506131>
- Yudhana, A., Sunardi, S., & Hartanta, A. J. S. (2020). Algoritma K-Nn Dengan Euclidean Distance Untuk Prediksi Hasil Penggergajian Kayu Sengon. *Transmisi*, 22(4), 123–129. <https://doi.org/10.14710/transmisi.22.4.123-129>