

Penerapan *Fuzzy K-Nearest Neighbor* (FK-NN) Dalam Menentukan Status Gizi Balita

Satria Dwi Nugraha¹, Rekyan Regasari Mardi Putri², Randy Cahya Wihandika³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹snugrahadwi@gmail.com, ²rekyan.rmp@ub.ac.id, ³rendicahya@ub.ac.id

Abstrak

Balita merupakan golongan yang memiliki masa penting dalam tumbuh kembang fisik anak. Balita sendiri adalah istilah umum bagi anak usia 1-3 tahun (batita) dan anak prasekolah (3-5 tahun). Masa balita sendiri sering dikatakan sebagai masa keemasan karena keberhasilan pertumbuhan seseorang di masa selanjutnya ditentukan pada saat masa balita. Pertumbuhan pada balita tidak hanya digunakan sebagai gambaran dalam bertambahnya ukuran anggota tubuh, tetapi juga digunakan sebagai gambaran mengenai kesinambungan antara asupan dan kebutuhan gizi. Salah satu indikator yang dapat mengetahui tingkat kesehatan balita sendiri adalah dengan melihat status gizi nya dengan menggunakan skala antropometri. Pada perhitungan antropometri, penentuan status gizi balita hanya didasarkan atas 4 faktor internal yaitu jenis kelamin, umur, berat badan dan tinggi badan. Oleh karena itu, antropometri dirasa masih kurang fleksibel karena hanya terpaku pada 4 faktor tersebut mengingat status gizi balita juga dipengaruhi oleh faktor lain seperti status kesehatan, pendidikan, pengetahuan, genetik, dan pendapatan. Metode klasifikasi Fuzzy K-Nearest Neighbor merupakan metode klasifikasi yang akan mampu melakukan perhitungan penentuan status gizi balita di masa yang akan datang dimana klasifikasi sendiri merupakan metode yang memakai data training sebagai bagian dalam mengambil keputusannya sehingga dapat menyesuaikan parameter-parameter penentu status gizi balita lainnya agar dapat menghasilkan hasil yang semakin akurat. Berdasarkan hasil pengujian perangkat lunak yang dibuat pada penelitian ini, akurasi yang dihasilkan sebesar 84,37% dengan menggunakan 160 data latih dengan 32 data uji dan nilai $k = 4$.

Kata Kunci: status gizi balita, klasifikasi, Fuzzy K-Nearest Neighbor

Abstract

Infants or so-called children is a group that have an important period in physical growth. Infants itself is categorized as a group of children between age 1 to 3 as a toddler group, and age 3 to 5 as a pre-school group. Some says children has a big role in order to attaining of growth success in the future for human, hence they call it as the golden age of living. Children's growth not only describing as an increasing of body dimensions but also as the continuity of intake and nutrient needs. An indicator to know the children's health is by determining their nutritional status. Based on SK Minister of Health in Indonesia, they use a method called anthropometry to determining children's nutritonal status. While this method only reviewing 4 internal factors, there're some other factors which influence of children's nutritional status itself such as genetic, disease, education, knowledge, and income. Therefore Fuzzy K-Nearest Neighbor is used in this study as a classifiaction method that can determining children's nutritional status because this method using the data training as the knowledge to clasify and would adjust other factors of nutritional status itself right in the future. From the test results of the study, this system can clasify well with maximum accuracy of 84,37% when using 160 training data with k value = 4.

Keywords: nutritional status of children, classification, Fuzzy K-Nearest Neighbor.

1. PENDAHULUAN

Balita merupakan golongan yang memiliki

masa penting dalam tumbuh kembang fisik anak. Balita sendiri adalah istilah umum bagi anak usia 1-3 tahun (batita) dan anak prasekolah (3-5 tahun) (Soetomo, 2010). Masa balita sendiri

sering dikatakan sebagai masa golden age atau masa keemasan karena keberhasilan pertumbuhan dan perkembangan seseorang di masa selanjutnya ditentukan pada saat masa balita. Pertumbuhan pada balita tidak hanya digunakan sebagai gambaran dalam bertambahnya ukuran anggota tubuh, tetapi juga digunakan sebagai gambaran mengenai kesinambungan antara asupan dan kebutuhan gizi. Status gizi adalah keadaan tubuh sebagai akibat dari pemakaian, penyerapan dan penggunaan nutrisi dari makanan. Salah satu indikator yang dapat menentukan tingkat kesehatan balita sendiri adalah dengan melihat status gizi nya. Berdasarkan SK Menkes Nomor 1995/Menkes/SK/XII/2010 bahwa untuk menilai status gizi anak diperlukan standar antropometri yang mengacu pada Standar World Health Organization. Indeks antropometri yang digunakan sendiri adalah berat badan terhadap umur (BB/U), tinggi badan terhadap umur (TT/U), dan berat badan terhadap tinggi badan (BB/TB). Pada pengukuran status gizi balita melalui skala antropometri, parameter yang menjadi perhitungan untuk menghasilkan output status gizi berupa gizi baik, gizi kurang, gizi buruk, dan gizi lebih adalah berat badan dan umur. Status gizi balita sendiri dipengaruhi langsung oleh dua faktor utama yaitu faktor internal dan faktor eksternal. Faktor internal adalah status kesehatan, umur, jenis kelamin, dan ukuran tubuh, sedangkan faktor eksternal meliputi pendidikan, pengetahuan pendapatan. (Radiansyah, 2007). Pada perhitungan antropometri, penentuan status gizi hanya didasarkan atas 4 faktor internal yaitu jenis kelamin, umur, berat badan dan tinggi badan. Oleh karena itu, antropometri dirasa masih kurang fleksibel karena hanya terpaku pada 4 faktor tersebut. Antropometri tidak bisa mendeteksi status gizi seorang balita berdasarkan faktor penentu status gizi lainnya seperti genetik dan penyakit. Antropometri juga dirasa kurang dapat mendeteksi status gizi dalam waktu singkat karena pada perhitungan ini cukup membutuhkan ketelitian sehingga keasalahan pengukuran akan dapat mempengaruhi presisi, akurasi, dan validitas (Supariasa, 2002). Oleh apa yang telah dijelaskan diatas, maka penelitian ini dilakukan untuk memenuhi kekurangan dalam menentukan status gizi balita dengan metode klasifikasi. Metode klasifikasi pada penelitian ini akan mampu melakukan perhitungan penentuan status gizi balita di masa yang akan datang karena klasifikasi merupakan

metode yang memakai data training sebagai bagian dalam mengambil keputusan dan dapat menyesuaikan parameter-parameter penentuan status gizi balita lainnya sehingga nantinya akan menghasilkan hasil yang akan semakin baik.

2. TINJAUAN PUSTAKA

2.1. Balita

Balita terbagi menjadi dua kategori berdasarkan karakteristiknya yaitu batita (1-3 tahun), dan anak usia prasekolah. Batita dan anak pra sekolah memiliki karakteristik pada laju pertumbuhan yang berbeda. Seorang batita memiliki laju pertumbuhan yang lebih cepat daripada anak usia pra sekolah. Asupan makanan pada batita pun juga masih membutuhkan peranan orang tua. Berbeda dengan anak pra sekolah, mereka sudah bisa memilih makanan sesuai selera mereka. Pada anak pra sekolah, prilaku juga mulai berubah karena pada masa ini mereka mulai mengenal lingkungan sekitar.

2.2. Status Gizi Balita

Terpenuhi atau tidaknya asupan gizi pada balita dapat dilihat berdasarkan status gizinya. Status gizi baik atau normal adalah merupakan keadaan yang sangat diinginkan oleh semua orang. Status gizi kurang adalah keadaan kekurangan energi protein yang disebabkan oleh rendah konsumsi energi dan protein dalam makanan sehari-hari sehingga tidak memenuhi angka kecukupan gizi (Suparisa, 2002). Status gizi lebih adalah keadaan gizi seseorang dimana jumlah energi yang masuk ke dalam tubuh lebih besar dari jumlah energi yang dikeluarkan (Nix, 2005). Gizi buruk adalah keadaan kurang gizi yang disebabkan karena kekurangan asupan energi dan protein juga mikronutrien dalam jangka waktu lama. Anak disebut gizi buruk apabila berat badan dibanding umur tidak sesuai (selama 3 bulan berturut-turut tidak naik) dan tidak disertai tanda-tanda bahaya.

2.3. Antropometri

Untuk mengetahui status gizi balita sendiri digunakan satuan baku yang biasa disebut dengan reference. Ukuran baku yang dipakai di Indonesia adalah baku antropometri. Satuan antropometri sendiri merupakan standar yang dipakai di Indonesia dalam menentukan status gizi balita mengacu pada World Health Organization – National Centre for Health Statistics (WHO-NCHS). Antropometri adalah

metode penentuan status gizi berdasarkan ukuran tubuh. Antropometri sendiri berhubungan dengan berbagai macam pengukuran dimensi dan komposisi tubuh dari tingkatan umur dan tingkatan gizi. Parameter penentuan status gizi berdasarkan antropometri sendiri adalah umur, berat badan, dan tinggi badan.

2.4. Data Mining

Data mining merupakan suatu proses yang menggunakan teknik statistika, matematika, kecerdasan buatan dan machine learning untuk mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar (Turban, 2005). Data mining sangat diperlukan untuk bagaimana membantu instansi atau perusahaan dalam menemukan informasi-informasi penting di dalam sebuah basis data dengan proses atau teknik-teknik statistika, kecerdasan buatan, matematika, dan machine learning.

2.5. Klasifikasi

Klasifikasi adalah suatu fungsionalitas data mining yang akan menghasilkan model untuk memprediksi kelas atau kategori dari objek-objek di dalam baris data. Klasifikasi merupakan proses yang terdiri dari dua tahap, yaitu tahap pembelajaran dan tahap pengklasifikasian (Khusnawi, 2007). Setiap algoritma klasifikasi mempunyai prinsip yang sama dimana melakukan learning dalam model sehingga bisa melakukan prediksi pada masukan ke label kelas yang benar sebagai keluarannya. Sedangkan pada tahap klasifikasi, tahap ini mengambil sebuah data yang tidak diketahui label kelasnya untuk diprediksi kemudian diuji menggunakan model yang telah dibuat pada proses pembelajaran.

2.6. Logika Fuzzy

Konsep ini diperkenalkan dan dipublikasikan pertama kali oleh Lotfi A. Zadeh, seorang profesor dari University of California di Berkeley pada tahun 1965. Logika fuzzy menggunakan ungkapan bahasa untuk menggambarkan nilai variabel. Logika fuzzy bekerja dengan menggunakan derajat keanggotaan dari sebuah nilai yang kemudian digunakan untuk menentukan hasil yang ingin dihasilkan berdasarkan atas spesifikasi yang telah ditentukan.

2.7. K-Nearest Neighbor (K-NN)

Algoritma K-NN adalah algoritma yang melakukan klasifikasi berdasarkan kedekatan lokasi (jarak) suatu data dengan data yang lain. Prinsip kerja K-Nearest Neighbor (K-NN) sendiri adalah mencari jarak terdekat antara data yang dievaluasi dengan k tetangga terdekatnya dalam data pelatihan.

Sebelum mencari jarak terdekat antara data yang dievaluasi, pada algoritma K-NN harus dilakukan preprocessing atau normalisasi terlebih dahulu. Preprocessing sendiri bertujuan untuk mendapatkan standar nilai pada semua atribut atau indikator dalam perhitungan. Pada penelitian ini, normalisasi yang digunakan adalah min-max normalization.

Proses min-max normalization ditunjukkan pada persamaan (1) :

$$V' = \frac{v(x) - \min(x)}{\text{Range}(x)} \quad (1)$$

dimana :

V' = Hasil Normalisasi yang nilainya berkisar antara 0 - 1

$V(x)$ = Nilai atribut yang akan dinormalisasikan

$\max(x)$ = Nilai maksimum dari suatu atribut $\max(x)$

$\min(x)$ = Nilai minimum dari suatu atribut

Pencarian nilai range pada perhitungan min max normalization ditunjukkan pada persamaan (2).

$$\text{Range}(x) = \text{nilai dari } (\max(x) - \min(x)) \quad (2)$$

Setelah melakukan normalisasi, maka dilakukan proses perhitungan jarak terdekat antara data latih dan data uji. Perhitungan ini bertujuan agar mengetahui jarak antara x_1 dan x_2 pada masing-masing record. Perhitungan ini menggunakan euclidean distance seperti pada persamaan (3).

$$d_i = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2} \quad (3)$$

dimana :

d_i = Jarak kedekatan

p = Jumlah atribut data

x_1 = Data latih

x_2 = Data uji

2.8. Fuzzy K-Nearest Neighbor (K-NN)

Fuzzy K-NN merupakan salah satu metode klasifikasi dengan menggabungkan teknik Fuzzy dan K-NN. Metode ini tidak seperti metode lain yang mana pada metode ini akan secara tegas memprediksi kelas yang diikuti oleh data uji berdasarkan perbandingan K terdekat. Dasar dari algoritma FK-NN adalah untuk menetapkan nilai keanggotaan sebagai fungsi jarak vektor dari K-NN dan keanggotaan tetangga mereka di kelas-kelas yang memungkinkan.

Metode ini berperan penting dalam menghilangkan ambiguitas dalam klasifikasi. Selain itu, sebuah instance akan memiliki derajat nilai keanggotaan pada setiap kelas sehingga akan lebih memberikan kekuatan atau kepercayaan suatu instance berada pada suatu kelas.

Sebelum menghitung nilai keanggotaan pada Fuzzy K-NN, terlebih dahulu dilakukan proses menggunakan pada persamaan (4).

$$u_{ij} = \begin{cases} 0,51 + \left(\frac{n_j}{n}\right) * 0,49, & \text{jika } j = 1 \\ \left(\frac{n_j}{n}\right) * 0,49, & \text{jika } j \neq 1 \end{cases} \quad (4)$$

keterangan :

n_j = Jumlah anggota kelas j pada suatu data latih n

n = Jumlah data latih yang digunakan

j = Kelas data

Selanjutnya menghitung nilai keanggotaan masing-masing kelas dengan persamaan (5).

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij} (\|x - x_j\|^{-2/(m-1)})}{\sum_{j=1}^k (\|x - x_j\|^{-2/(m-1)})} \quad (5)$$

keterangan :

u_{ij} = nilai keanggotaan fuzzy pada contoh pengujian (x, xj)

k = nilai tetangga terdekat

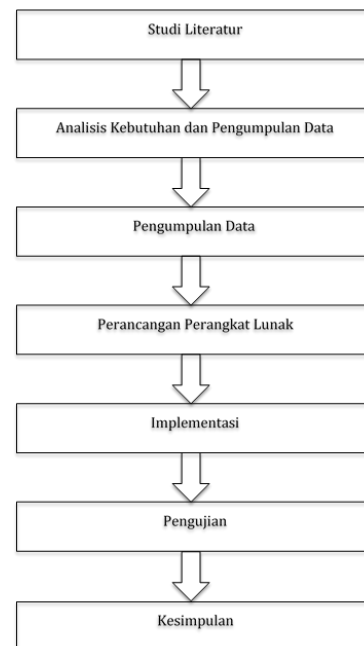
j = variabel data keanggotaan data uji

m = bobot pangkat yang besarnya $m > 1$

3. TABEL DAN GAMBAR

Metodologi menjelaskan metode mengenai alur yang digunakan dalam implementasi metode Fuzzy K-Nearest Neighbor untuk menentukan status gizi balita pada perangkat lunak. Pada bab ini, alur tersebut dimulai dari tahapan mempelajari studi literatur yang

dilakukan hingga pengambilan kesimpulan. Metode penelitian pada bab ini dapat diilustrasikan pada gambar berikut.



Gambar 3.1 Diagram alur metodologi penelitian

Penjelasan diagram alur metodologi penelitian sebagai berikut :

1. Studi literatur merupakan tahapan agar dapat mengetahui referensi lain yang berhubungan dengan penelitian agar peneliti dapat memahami permasalahan beserta pengetahuan yang menjadi landasan teori dalam menemukan solusi terkait penelitian.
2. Analisis kebutuhan merupakan tahapan untuk mengetahui kebutuhan-kebutuhan apa saja yang diperlukan dalam mendukung perangkat lunak. Kebutuhan-kebutuhan yang diperlukan meliputi kebutuhan input, kebutuhan proses, dan kebutuhan output.
3. Setelah mengetahui kebutuhan yang diperlukan, maka penulis akan mulai ke tahap pengumpulan data. Data yang diperoleh dari objek yang akan diproses pada penelitian penentuan status gizi balita sendiri didapatkan dari hasil penelitian sebelumnya pada Puskesmas dan Posyandu di Kecamatan Kertosono Kabupaten Nganjuk
4. Perancangan dibuat berdasarkan kebutuhan input dan output. Agar dapat lebih mudah memahami perancangan, maka tahapan ini akan digambarkan kedalam bentuk diagram blok.

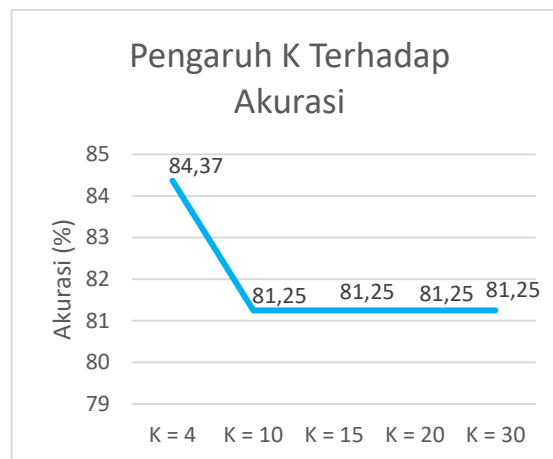
5. Proses implementasi merupakan tahapan dimana mengimplementasikan metode F-KNN ke dalam perangkat lunak yang akan dibuat dengan menggunakan bahasa pemrograman yang direncanakan yaitu menggunakan bahasa pemrograman php. Selain itu, digunakan beberapa tools yang mendukung seperti MYSQL, XAMPP, sebagai database management.
6. Pada tahapan ini dilakukan pengujian terhadap perangkat lunak dengan menggunakan data latih dan data uji pada tiap pengujiannya. Pengujian dilakukan agar pengguna dapat mengetahui tingkat akurasi terhadap beberapa skenario.
7. Tahapan ini memuat mengenai apa yang menjadi kesalahan atau kekurangan pada perangkat lunak sebagai saran bagi penulis untuk memperbaiki kesalahan tersebut. Saran yang diperoleh juga digunakan sebagai penulis untuk mencapai tujuan pada penelitian ini.

4. HASIL PENELITIAN

Pada pengujian perangkat lunak ini, data yang digunakan sebanyak 160 data latih dengan 32 data uji yang diperoleh dari Puskesmas Kertosono Kabupaten Nganjuk dan sudah terverifikasi oleh pakar yaitu dokter pada poli anak Puskesmas Dinoyo Malang. Terdapat lima pengujian yang dilakukan, yaitu :

4.1. Pengujian Pengaruh Nilai K Terhadap Akurasi

Pada pengujian ini digunakan nilai k yang berbeda yaitu k yang berbeda yaitu $k = 5$, $k = 10$, $k = 15$, $k = 20$. Jumlah data latih yang digunakan adalah sebanyak 160 dengan sebaran yang sama rata dan data uji sebanyak 32 data. Tujuan pengujian ini adalah untuk mengetahui pengaruh banyaknya tetangga terhadap akurasi dan untuk mengetahui nilai k yang paling optimal pada penentuan status gizi balita menggunakan metode fuzzy k-nearest neighbor. Hasil percobaan pengujian pengaruh nilai K terhadap akurasi dapat dilihat pada Gambar 4.1.



Gambar 4.1 Hasil pengujian pengaruh nilai k terhadap akurasi

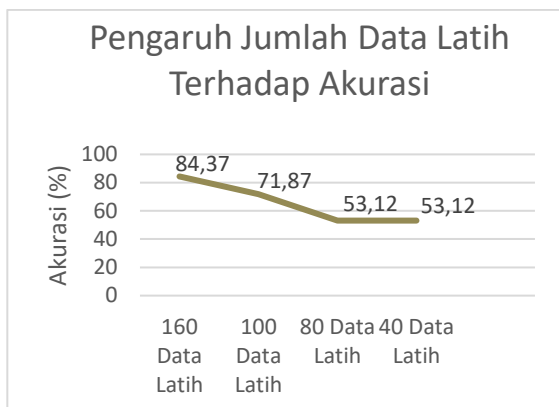
Hasil pengujian pengaruh nilai k terhadap akurasi dengan menggunakan 160 data latih dan data uji yang sama yaitu sejumlah 32 data uji. Pengujian dengan nilai $k = 4$ menghasilkan akurasi sebesar 84,37%. Pengujian dengan nilai $k = 10$ menghasilkan akurasi sebesar 81,25%, begitu pula dengan nilai $k = 15$, $k = 20$, $k = 30$ yang menghasilkan akurasi yang sama. Total rata-rata akurasi pada pengujian ini sendiri sebesar 81,87%.

Pada hasil pengujian pengaruh nilai K terhadap akurasi yang ditampilkan pada gambar 1, dapat diketahui bahwa akurasi terbesar dihasilkan pada jumlah $K = 4$ yaitu sebesar 84.37%. Sementara itu, nilai K yang berbeda menghasilkan akurasi yang berbeda pula sehingga dapat disimpulkan bahwa nilai K berpengaruh terhadap akurasi yang dihasilkan meskipun akurasi yang berubah tidak terlalu signifikan. Selain itu, dapat dilihat pada grafik tersebut bahwa semakin besar nilai K maka, akurasi yang dihasilkan akan menurun. Hal ini dikarenakan semakin besar nilai K, semakin banyak data yang diambil dalam menghasilkan klasifikasi sehingga semakin banyak pula data yang tidak relevan dan mempengaruhi kesalahan dalam prediksi.

4.2. Pengujian Pengaruh Jumlah Data Latih Terhadap Akurasi

Pada pengujian ini, data latih yang digunakan memiliki jumlah yang berbeda yaitu sebanyak 160, 100, 80, dan 40 data latih. Data uji yang digunakan adalah sebanyak 32 data uji. Pengujian ini menggunakan nilai k optimal yang diketahui pada pengujian sebelumnya yaitu $k = 4$ dengan akurasi sebesar 84,27%. Hasil

pengujian pengaruh jumlah data latih terhadap akurasi ditunjukkan pada Gambar 4.2.



Gambar 4.2 Hasil pengujian pengaruh jumlah data latih terhadap akurasi

Pada hasil pengujian pengaruh nilai jumlah data latih terhadap akurasi didapatkan bahwa pengujian dengan menggunakan 160 data latih dengan data uji sebanyak 32, akurasi yang dihasilkan sebesar 84,37%. Skenario kedua dengan memakai 100 data latih dan 32 data uji menghasilkan akurasi sebesar 71,87%. Kemudian, skenario dengan menggunakan 80 data latih dan 32 data uji dapat menghasilkan akurasi sebesar 53,12%. Sedangkan skenario pengujian pengaruh jumlah data latih terhadap akurasi terakhir dengan menggunakan 40 data latih dan 32 data uji, akurasi yang dihasilkan sebesar 53,12%.

Dari analisis tersebut, dapat diketahui bahwa jumlah data latih memang mempengaruhi akurasi yang dihasilkan dengan akurasi terbesar yaitu 84,37% pada skenario menggunakan 160 data latih. Jumlah data latih yang berpengaruh terhadap akurasi dikarenakan adanya komposisi data yang berbeda pada tiap data latih yang digunakan sehingga bisa menyebabkan kesalahan prediksi pada pengujian.

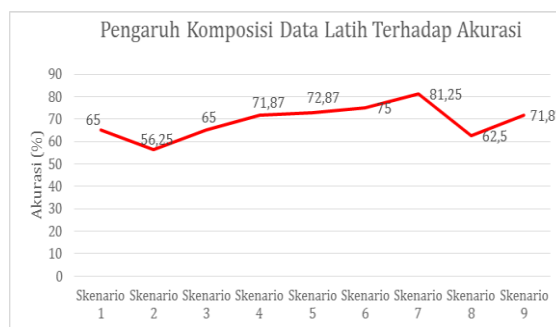
4.3. Pengujian Pengaruh Komposisi Data Latih Terhadap Akurasi Pengujian

Pada pengujian pengaruh komposisi data latih terhadap akurasi, data latih yang digunakan adalah sebanyak 100 data latih dengan komposisi yang berbeda-beda. Pengujian ini bertujuan untuk mengetahui apakah dengan komposisi data latih yang berbeda-beda, akurasi yang dihasilkan pada penelitian ini akan berpengaruh atau tidak. Komposisi data latih yang berbeda pada 100 data latih adalah rasio jumlah data latih pada empat kelas yang berbeda pada setiap skenario. Skenario pengujian ketiga

adalah sebagai berikut.

- 40 kelas gizi baik, 20 kelas gizi kurang, 20 kelas gizi buruk, 20 kelas gizi lebih.
- 20 kelas gizi baik, 40 kelas gizi kurang, 20 kelas gizi buruk, 20 kelas gizi lebih.
- 20 kelas gizi baik, 20 kelas gizi kurang, 40 kelas gizi buruk, 20 kelas gizi lebih.
- 20 kelas gizi baik, 20 kelas gizi kurang, 20 kelas gizi buruk, 40 kelas gizi lebih.
- 10 kelas gizi baik, 30 kelas gizi kurang, 30 kelas gizi buruk, 30 kelas gizi lebih.
- 30 kelas gizi baik, 10 kelas gizi kurang, 30 kelas gizi buruk, 30 kelas gizi lebih.
- 30 kelas gizi baik, 30 kelas gizi kurang, 10 kelas gizi buruk, 30 kelas gizi lebih.
- 30 kelas gizi baik, 30 kelas gizi kurang, 30 kelas gizi buruk, 10 kelas gizi lebih.
- 25 kelas gizi baik, 25 kelas gizi kurang, 25 kelas gizi buruk, 25 kelas gizi lebih.

Hasil pengujian pengaruh komposisi data latih terhadap akurasi ditunjukkan pada Gambar 4.3.

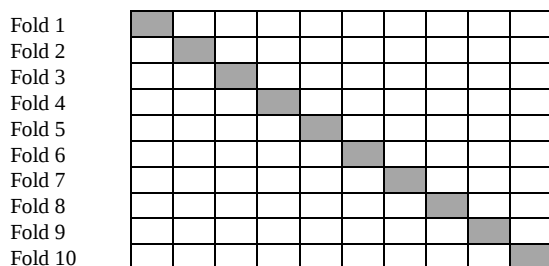


Gambar 4.3 Hasil pengujian pengaruh komposisi data latih terhadap akurasi

Berdasarkan Gambar 4.3, akurasi yang dihasilkan oleh komposisi data latih yang hanya memiliki satu kelas dominan tergolong rendah. Sedangkan komposisi data latih yang memiliki tiga kelas dominan yang ditunjukkan pada skenario ke-5 hingga skenario ke-8, akurasi yang dihasilkan cenderung tinggi. Sedangkan komposisi data latih yang memiliki beberapa kelas dengan data latih yang jumlahnya sedikit seperti pada skenario pertama hingga keempat, akurasi yang dihasilkan cenderung kecil. Dari pengujian ketiga ini bisa disimpulkan bahwa akurasi yang berubah-ubah memang dikarenakan adanya komposisi data latih yang berbeda-beda pada beberapa skenario.

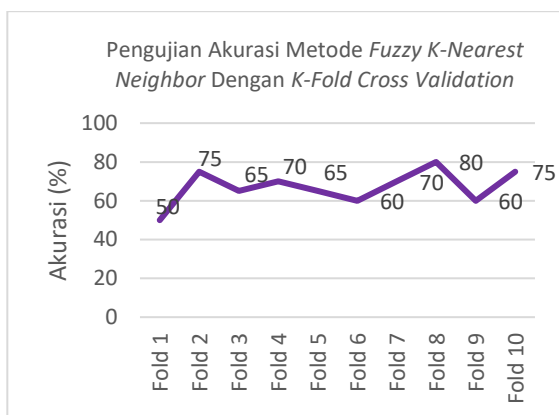
4.4. Pengujian Akurasi Metode Fuzzy K-Nearest Neighbor dengan K-Fold Cross Validation

Pada proses pengujian ini, dataset akan dibagi kedalam jumlah k subset dan selanjutnya dilakukan proses training dan testing sebanyak k kali. Subset dengan warna yang lebih gelap akan digunakan sebagai data testing, sisanya akan digunakan sebagai data training. Proses ini sebagaimana ditunjukkan pada Gambar 4.4.



Gambar 4.4 Pembagian dataset K-Fold Cross Validation

Hasil pengujian akurasi metode Fuzzy K-Nearest Neighbor dengan menggunakan metode K-Fold Cross Validation ditunjukkan pada Gambar 4.5.

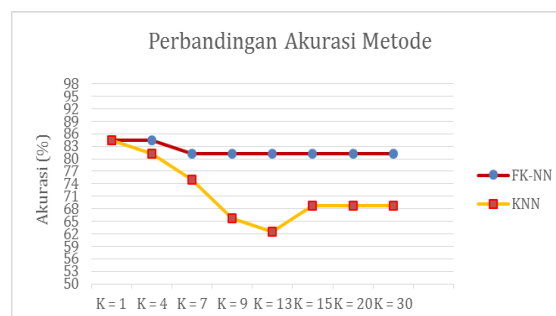


Gambar 4.5. Hasil pengujian akurasi metode F-KNN dengan metode K-Fold Cross Validation

Analisis yang kedua pada pengujian ini adalah perbedaan akurasi yang dihasilkan pada fold pertama dan fold kedelapan dikarenakan pada pengujian fold pertama, data latih yang digunakan cenderung memiliki karakteristik yang cukup berbeda dengan data uji sehingga terjadi kesalahan pada prediksi, sedangkan pada fold kedelapan dengan akurasi yang cukup tinggi dikarenakan adanya kedekatan karakteristik data latih yang digunakan dengan data uji sehingga hasil prediksi cenderung lebih baik.

4.5. Pengujian Perbandingan Akurasi Metode Fuzzy K-Nearest Neighbor dengan K-Nearest Neighbor

Skenario pengujian perbandingan akurasi metode K-Nearest Neighbor dengan Fuzzy K-Nearest Neighbor adalah melakukan beberapa skenario pengujian dengan menggunakan nilai K yang berbeda-beda yaitu $k = 1$, $k = 4$, $k = 7$, $k = 9$, $k = 13$, $k = 15$, $k = 20$, dan $k = 30$ dengan data latih dan data uji yang sama yaitu 160 data latih dan 32 data uji. Hasil pengujian pengaruh perbandingan akurasi metode K-NN dengan FK-NN ditunjukkan pada Gambar 4.6.



Gambar 4.6. Hasil pengujian akurasi metode F-KNN dan metode KNN

Pada grafik yang ditunjukkan pada Gambar 4.6, dapat diketahui bahwa masing-masing metode FK-NN dan K-NN memiliki akurasi yang sama pada jumlah $k = 1$ yaitu sebesar 84,37%.

Pada metode K-NN, akurasi yang dihasilkan cenderung mengalami penurunan yang signifikan seiring bertambahnya nilai K. Hal ini disebabkan, semakin banyak nilai K yang digunakan pada metode K-NN, semakin banyak pula rentang data yang digunakan sehingga akan banyak pula data yang tidak relevan dan mempengaruhi kesalahan dalam prediksi.

Sedangkan pada metode FK-NN sendiri, penurunan akurasi yang tidak terlalu signifikan seiring bertambahnya jumlah K dikarenakan metode FK-NN tidak hanya mengambil keputusan berdasarkan jumlah kelas mayoritas yang muncul, tetapi metode FK-NN akan mencari kelas prediksi berdasarkan data yang memiliki jarak terdekat dengan data uji meskipun terdapat jumlah kelas mayoritas yang salah pada jumlah K. Metode FK-NN selanjutnya akan menghitung nilai derajat keanggotaan pada tiap kelas yang berperan pada prediksi kelas. Nilai derajat keanggotaan yang paling tinggi merupakan kelas hasil prediksinya. Semakin dekat jarak suatu data

terhadap data uji, semakin besar pula nilai derajat keanggotaan kelas dari data tersebut.

5. KESIMPULAN

Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa

1. Metode *Fuzzy K-Nearest Neighbor* digunakan dalam menentukan status gizi balita melalui tahapan melakukan normalisasi parameter data latih dan data uji, mencari nilai *euclidean distance* nya, menentukan nilai *k*, dan mencari nilai *membership* pada masing-masing kelas.
2. Metode *Fuzzy K-Nearest Neighbor* mampu melakukan penentuan status gizi balita dengan akurasi tertinggi sebesar 84,37% dengan jumlah data latih sebanyak 160 dan nilai *k* = 4.
3. Pengujian pengaruh jumlah data latih terhadap akurasi menghasilkan akurasi sebesar 84,37% pada jumlah data latih sebanyak 160 data latih.
4. Pengujian pengaruh komposisi data latih terhadap akurasi menghasilkan akurasi tertinggi sebesar 81,25% pada skenario yang memiliki komposisi data latih yang dominan.
5. Pengujian akurasi *Fuzzy K-Nearest Neighbor* menggunakan metode *K-Fold Cross Validation* menghasilkan rata-rata akurasi sebesar 67% dengan jumlah *fold* sebanyak 10.
6. Perbandingan metode F-KNN dengan KNN menghasilkan rata-rata akurasi masing-masing sebesar 82% dan 71% sehingga dapat dipastikan metode FK-NN lebih baik dibandingkan K-NN, hal ini dikarenakan FK-NN tidak hanya mengambil keputusan berdasarkan jumlah kelas mayoritas yang muncul seperti pada metode K-NN, tetapi metode FK-NN akan menghitung nilai derajat keanggotaan pada tiap kelas yang berperan pada prediksi kelas. Nilai derajat keanggotaan yang paling tinggi merupakan kelas hasil prediksinya.

6. SARAN

Saran yang dapat diambil berdasarkan penelitian ini adalah :

1. Untuk meningkatkan akurasi pada sistem, diharapkan pada penelitian selanjutnya menggunakan data latih yang lebih banyak sehingga semakin banyak karakteristik dari

data latih maka akan meningkatkan kebenaran pada prediksi.

2. Pencarian nilai *K* pada FK-NN sangat krusial sehingga diharapkan dapat mencari nilai *K* optimal dengan analisis metode pencarian nilai *K* seperti *k fold cross validation*.
3. Walaupun akurasi yang dihasilkan penelitian ini cukup baik, diharapkan melakukan penelitian serupa dengan menggunakan metode klasifikasi lain atau menggunakan kombinasi metode yang berbeda.

DAFTAR PUSTAKA

- NIX, S. 2005. *William's Basic Nutrition & Diet Therapy*, Twelfth Edition. Elsevier Mosby Inc, USA.
- RADIANSYAH, E., 2007. *Penanggulangan Gizi Buruk*. Tersedia di: <http://www.dinkespurworejo.go.id> [Diakses 26 Maret 2016].
- SUPARIASA. 2002. *Penilaian Status Gizi*. Jakarta : Buku Kedokteran EGC.
- SUTOMO, B & ANGGRAINI, D. Y. 2010. *Makanan Sehat Pendamping ASI*. Jakarta : Demedia.
- TURBAN. 2005. *Decision Support Systems and Intelligent Systems (Sistem Pendukung Keputusan dan Sistem Cerdas)* jilid 1. Yogyakarta : Andi Offset.
- KHUSNAWI. 2007. *Pengantar Solusi Data Mining (Online)*. STMIK AMIKOM. Yogyakarta. Tersedia di: <http://p3m.amikom.ac.id/p3m/56%20%20PENGANTAR%20SOLUSI%20DATA%20MINING.pdf> [Diakses 23 Maret 2016].