



Analysis of Family Health with a Combination of Naive Bayes KNN Methods

Ona Maliki¹, Muis Nanja²

^{1,2}Jurusan Sistem Informatika, Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo, Gorontalo

^{1,2}Jl. Achmad Nadjamudin No 10. Kel. Dulalowo, Kec. Kota Tengah, Kota Gorontalo, Gorontalo, Indonesia

E-mail: Onamaliki08@gmail.com², muisnanja@unisan.ac.id¹

Abstrak

Kesehatan merupakan bagian yang sangat penting dalam kehidupan bangsa dan Negara kita, di dalam bangsa yang maju dan berkembang terdapat masyarakat yang sehat. Sehat dapat diartikan keadaan individu yang sejahtera baik dari segi fisik, ekonomi maupun jiwa social. Kesehatan inilah yang perlu dijaga dan dirawat untuk kelangsungan suatu bangsa dan Negara. Melihat dari pentingnya unsur kesehatan dalam suatu bangsa, pemerintah telah melaksanakan program nawa cita dengan salah satu pointnya yaitu merujuk pada kesehatan keluarga. Pelaksanaan program keluarga sehat, pemerintah telah mengupayakan pendataan dan pengontrolan keluarga sehat dengan pendekatan keluarga. Melihat dari system pendataan keluarga sehat dengan penetapan status yaitu meratakan nilai indicator untuk mendapatkan status kesehatan keluarga. Peneliti bermaksud melakukan pengembangan dari segi penentuan status kesehatan keluarga dengan menggunakan metode kombinasi metode K-Nearest Neighbor (KNN) dengan Naive Bayes dengan tujuan mengetahui hasil penerapan metode tersebut serta memperoleh hasil yang lebih baik. Berdasarkan dari hasil penelitian yang dilakukan diperoleh nilai akurasi 93,53% pada model kombinasi metode Naive Bayes KNN. Sehingga dapat disimpulkan bahwa kombinasi dari metode tersebut baik digunakan untuk melakukan penentuan status kesehatan keluarga.

Kata Kunci: Keluarga Sehat, KNN, KNN-Naive Bayes, Nive Bayes

Abstract

Health is a very important part in the life of our nation and country, in a developed and developing nation there is a healthy society. Healthy can be defined as the state of an individual who is prosperous both physically, economically and socially. Health is what needs to be maintained and cared for for the survival of a nation and state. Seeing the importance of the health element in a nation, the government has implemented the Nawacita program with one of the points referring to family health. In implementing the healthy family program, the government has sought to collect data and control healthy families with a family approach. Looking at the data collection system for healthy families with status determination, that is, averaging the indicator values to get the family's health status. The researcher intends to develop in terms of determining family health status by using a combination method of the K-Nearest Neighbor (KNN) method with Naive Bayes with the aim of knowing the results of applying the method and obtaining better results. Based on the results of the research conducted, the accuracy value of 93.53% was obtained in the combination model of the Naive Bayes KNN method. So it can be concluded that the combination of these methods is good for determining family health status.

Keywords: Healthy Family, KNN, KNN-Naive, Nive Bayes

I. PENDAHULUAN

Kemajuan suatu Negara merupakan kemajuan yang dilatar belakangi oleh masyarakat atau keluarga yang sehat, karena keluarga yang sehat akan menghasilkan sumber daya manusia yang baik dan produktif, sehingga dapat meningkatkan kualitas suatu bangsa baik dari segi ekonomi, pendidikan maupun bidang lainnya. Keluarga sehat dapat diartikan bahwa dalam satu keluarga setiap individunya merupakan individu yang sejahtera, baik dari segi ekonomi, mental maupun fisik.[1] Keluarga yang sehat akan dapat hidup

normal secara sosial dan ekonomi di tengah masyarakat lainnya. Kesehatan merupakan salah satu bagian program pemerintah yakni nawa cita yaitu peningkatan kualitas hidup dari segi kesehatan. Program Nawa Cita tidak hanya mengupayakan dari segi sarana kesehatan masyarakat akan tetapi ikut mengupayakan terbentuknya suatu keluarga sehat.[2] Upaya pengontrolan keluarga sehat, pelaksanaannya yaitu dengan melakukan pendataan keluarga yang dalam pengolahan data sebelumnya menggunakan perhitungan rata-rata. Berdasarkan dari data yang

diperoleh peneliti akan melakukan pengembangan pengolahan data untuk menentukan status keluarga sehat yaitu dengan melakukan analisa status kesehatan keluarga dengan menggunakan kombinasi metode K-Nearest Neighbor dengan Naïve Bayes dalam pengolahan datanya.

Penelitian ini dilakukan dengan tujuan untuk mengetahui apakah kombinasi metode K-Neares Neighbor dengan Naïve Bayes dapat digunakan untuk menganalisa status keluarga sehat dan bagaimana hasil penerapannya. Penerapan kombinasi metode K-Neares Neighbor dengan Naïve Bayes digunakan karena merujuk pada penelitian yang dilakukan oleh Naisah Marito Putry dan Betha Nurina Sari dengan Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Melitus, mendapat hasil akurasi 80% untuk naïve bayes dan 75% untuk KNN,[3] penelitian yang dilakukan oleh syafitri hidayatul annur aini .dkk dengan judul seleksi fitur *information gain* untuk klasifikasi penyakit jantung menggunakan kombinasi metode *k-nearest neighbor* dan *naïve bayes* dengan hasil akurasi 92,31%[4] dan penelitian yang dilakukan oleh Mohammad Abdul Manan Sarwido, Gentur Wahyu Nyipto Wibowo dengan judul Klasifikasi Penyakit Kelamin Pada Wanita Dengan Menggunakan Kombinasi Metode K-Nearest Neighbor Dan Naïve Bayes Classifier, dengan hasil akurasi dari kombinasi metode K-Nearest Neighbor dan Naïve Bayes Classifier sebesar 97.5% dengan menggunakan average accuracy dan sebesar 99.17%. [5]

II. LANDASAN TEORI

A. Keluarga Sehat

Keluarga Sehat dapat didefinisikan bahwa dalam satu keluarga setiap individunya merupakan individu yang sejahtera, baik dari segi ekonomi, mental maupun fisik.[6] Karena kesehatan merupakan factor utama yang mencerminkan kualitas sumber daya manusia dan juga merupakan investasi bagi peningkatan kualitas sumber daya manusia, sesuai dengan Undang-Undang Nomor 36 Tahun 2009 tentang kesehatan. [7]

B. Indikator Keluarga Sehat

Terdapat 12 indikator yang telah disepakati dalam pelaksanaan program keluarga sehat .12 indikator tersebut digunakan untuk melakukan analisa status kesehatan keluarga. Indikator keluarga sehat dapat dilihat pada table 1 berikut ini :[7] Sumber(<https://www.alodokter.com/kenali-12-indikator-keluarga-sehat-menurut-kemenkes-ri>)

C. K-Neares Neighbor (KNN)

K-Nearest Neighbor merupakan algoritma yang mengklasifikasi objek data baru dengan system kerjanya mencari jarak terdekat atau nilai k dari sejumlah data latih, kemudian menyimpulkan nilai kelas berdasarkan nilai yang paling banyak muncul dari sejumlah nilai k yang ditentukan.[8] Ada banyak cara untuk mengukur jarak kedekatan antara data baru

dengan data lama (data training), diantaranya euclidean .

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

Keterangan:

x = sampel data

y = data uji

D = Jarak

Dengan D adalah jarak antara titik pada data *training* x dan titik data *testing* y yang akan diklasifikasi, dimana $x = x_1, x_2, \dots, x_i$ dan $y = y_1, y_2, \dots, y_i$ dan I merepresentasikan nilai atribut serta n merupakan dimensi atribut. Langkah-langkah untuk menghitung metode Algoritma K-Nearest Neighbor.[9]

- Tentukan jumlah K = Jumlah tetangga terdekat.
- Hitung jarak antara data baru dengan semua data *training*
- Urutkan jarak tersebut dan tetapkan tetangga terdekat berdasarkan jarak minimum ke-K
- Periksa kelas dari tetangga terdekat
- Gunakan mayoritas sederhana dari kelas tetangga terdekat sebagai nilai prediksi data baru

D. Naïve Bayes

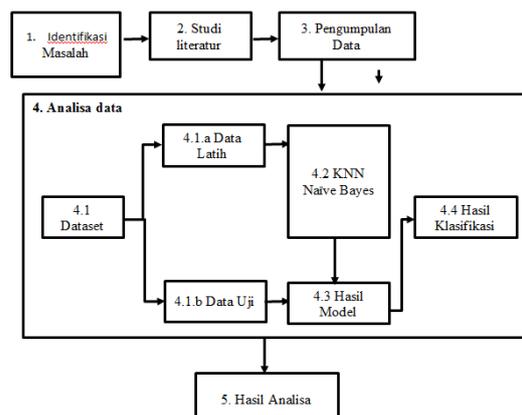
Naïve Bayes Classifier merupakan salah satu algoritma yang menerapkan perhitungan probabilitas dan statistic. Dasar dari Naïve Bayes yang dipakai dalam pemrograman adalah rumus Bayes:

$$P(A|B) = (P(B|A) * P(A)) / P(B) \dots\dots\dots(2)$$

Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, dan peluang B. Pada pengaplikasiannya nanti rumus ini berubah menjadi : $P(C_i|D) = (P(D|C_i) * P(C_i)) / P(D) \dots\dots\dots(3)[10]$

E. Metode Yang Diusulkan

Langkah awal dari penelitian adalah Identifikasi permasalahan, Setelah permasalahan berhasil diuraikan, lanjut ke tahap studi literatur /studi pustaka yang berkaitan dengan objek penelitian dan algoritma yang digunakan. Metode penelitian dapat dijabarkan pada gambar berikut:



Gambar 1. Metode Yang Diusulkan

1. Identifikasi masalah merupakan langkah awal yang dilakukan peneliti dengan Kepala rumah tangga (KK) yang menjadi objek penelitian. mengingat masih banyak masyarakat yang belum menerapkan hidup sehat
2. Studi literatu merupakan tahapan selanjutnya yang ditempuh peneliti dengan melakukan pengumpulan beberapa refrensi yang berkaitan metode dan lingkungan keluarga sehat.
3. Pengumpulan data merupakan tahapan yang dilakukan peneliti untuk memperoleh data baik data primer maupun data sekunder.
4. Analisa data, dalam melakukan analisa data, dataset yang diperoleh akan dibagi menjadi dua bagian yaitu data latih dan data uji. Data latih merupakan data yang digunakan untuk keperluan analisa model, guna mendapatkan model dari metode yang diterapkan. Data uji merupakan data yang digunakan untuk menguji suatu model untuk mendapatkan hasil klasifikasi.
5. Hasil analisa merupakan capaian yang diperoleh dari proses pengolahan data yaitu berupa hasil pemodelan, dengan tingkat akurasi yang bagus dapat disimpulkan bahawa model dari metode yang diperoleh juga bagus dan hasil klasifikasi merupakan hasil dari implementasi dari model yang diperoleh.

F. Jenis Penelitian

Metode yang digunakan dalam penelitian ini adalah metode deskriptif. Metode deskriptif; yaitu, Analisis statistik yang digunakan untuk menganalisis data dengan cara mendeskripsikan atau menggambarkan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi. [11]

G. Jenis dan Sumber Data

Dua jenis data yang diperlukan dalam penelitian ini yaitu. Data Primer, yaitu data dari desa idaman dan batu tiga. Data Sekunder, merupakan data yang didapat dari berupa jurnal, buku ataupun laporan hasil penelitian sebelumnya.[12]

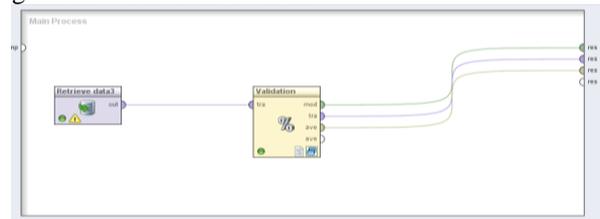
III. HASIL DAN PEMBAHASAN

A. Hasil Eksperimen Dan Pengujian Model

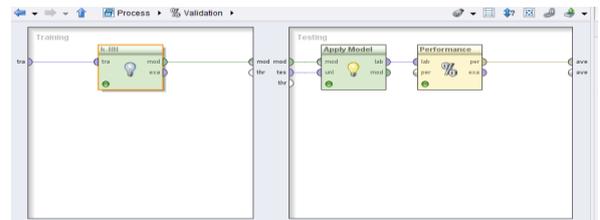
Berdasarkan dari hasil eksperimen yang telah dilakukan oleh peneliti. Peneliti melakukan beberapa tahap untuk memperoleh suatu hasil penelitian yaitu mulai dari tahap pencarian model terbaik yaitu dengan menggunakan beberapa metode atau algoritma yaitu KNN, Naïve Bayes, kombinasi KNN Naïve Bayes dan kombinasi Naïve Bayes KNN. Model yang dianggap terbaik adalah model yang memiliki tingkat accuracy yang paling besar. Model yang memiliki nilai accuracy yang paling besar akan digunakan untuk mengklasifikasikan data uji.

Parameter KNN

Berdasarkan dari tahapan eksperimen yang dilakukan oleh peneliti diperoleh hasil pemodelan dari parameter KNN dengan nilai ketetanggan 3,5,7 dan 9 dengan nilai cross validation 10 dapat dilihat pada gambar 2, gambar 3 dan tabel 2 berikut :



Gambar 2 Model validasi



Gambar 3 Model KNN

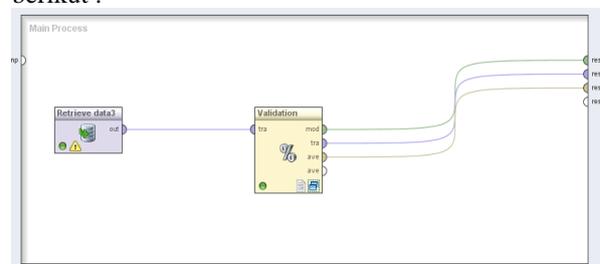
Tabel 2 Hasil Pengujian Metode KNN

K	Number of validation	Accuracy
3	10	91,24%
5	10	90,62%
7	10	92,39%
9	10	88.30%

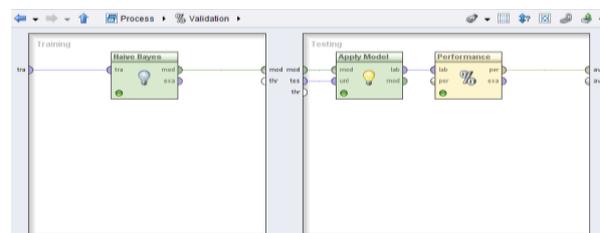
Berdasarkan pada Tabel 2 yaitu eksperimen penentuan model dengan nilai ketetanggan (k) 3,5,7 dan 9 dengan cross validation 10 diperoleh nilai accuracy yang paling baik, terdapat pada model KNN dengan nilai k =7 yaitu dengan nilai accuracy 92,39%.

Parameter Niave Bayes

Berdasarkan dari tahapan eksperimen yang dilakukan oleh peneliti diperoleh hasil pemodelan dari parameter Naïve bayes dengan nilai cross validation 10 dapat dilihat pada gambar 4, gambar 5 dan tabel 3 berikut :



Gambar 4 Model Validasi



Gambar 5 Model Naïve Bayes

Tabel 3 Hasil Pengujian Metode Naïve Bayes

Number of validation	Accuracy
10	91,83%

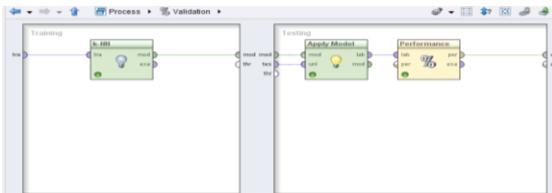
Berdasarkan pada Tabel 3 yaitu eksperiment penentuan model naïve bayes dengan cross validation 10 diperoleh nilai accuracy sebesar 91.83%

Parameter Kombinasi KNN Naïve Bayes

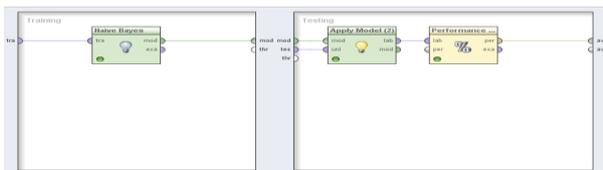
Berdasarkan dari tahapan eksperiment yang dilakukan oleh peneliti diperoleh hasil pemodelan dari parameter kombinasi KNN Naïve Bayes dengan nilai ketetanggan 3,5,7 dan 9 dengan nilai cross validation 10 dapat dilihat pada gambar 6, gambar 7, gambar 8 dan tabel 4 hasil berikut :



Gambar 6 Model Validasi



Gambar 7 Model KNN



Gambar 8 Model Naïve Bayes

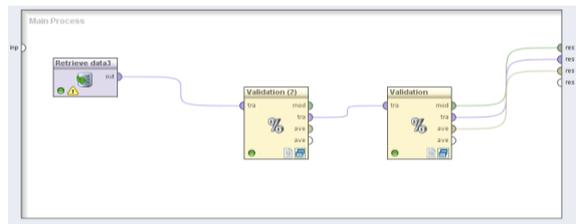
Tabel 4 Hasil Pengujian Metode KNN

KNN(k)_NBY	Number of validation	Accuracy
3	10	91,24%
5	10	91,24%
7	10	91,24%
9	10	91,24%

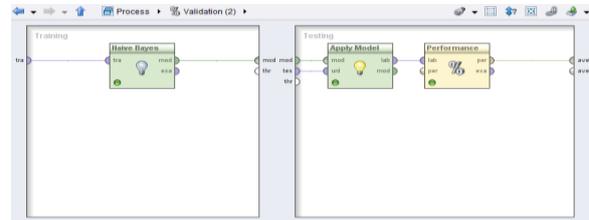
Berdasarkan pada Tabel 4 yaitu eksperiment penentuan model dengan kombinasi parameter KNN nilai ketetanggan (k) 3,5,7 dan 9 dan Naïve Bayes dengan cross validation 10 diperoleh nilai accuracy yang konstan yaitu sebesar 91,24%.

Parameter Kombinasi Naïve Bayes KNN

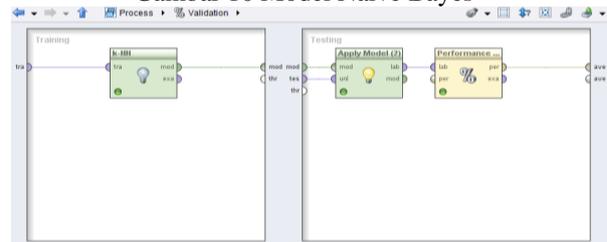
Berdasarkan dari tahapan eksperiment yang dilakukan oleh peneliti diperoleh hasil pemodelan dari parameter kombinasi Naïve Bayes KNN dengan nilai ketetanggan 3,5,7 dan 9 dengan nilai cross validation 10 dapat dilihat pada gambar 9, gambar 10, gambar 11 dan tabel 5 berikut :



Gambar 9 Model Validasi



Gambar 10 Model Naïve Bayes



Gambar 11 Model KNN

Tabel 5 Hasil Pengujian Metode KNN

NBY_KNN(k)	Number of validation	Accuracy
3	10	93,53%
5	10	91,80%
7	10	92,39%
9	10	90,65%

Berdasarkan pada Tabel 5 yaitu eksperiment penentuan model dengan kombinasi parameter Naïve Bayes KNN nilai ketetanggan (k) 3,5,7 dan 9 dan dengan cross validation 10 diperoleh nilai accuracy yang paling besar yaitu terdapat pada kombinasi metode Naïve Bayes KNN pada nilai K=3 yaitu sebesar 93,53%.

Evaluasi

Dari beberapa eksperiment model yang dilakukan sebelumnya diperoleh hasil accuracy rata-rata sangat baik. Berdasarkan hasil eksperiment yang telah dilakukan, maka dapat ditentukan model yang akan diimplementasikan untuk melakukan pengelompokan atau kalsifikasi keluarga sehat yaitu model yang tingkat akurasinya paling besar, nilai accuracy model tersebut dapat dilihat pada tabel 6 berikut :

Tabel 6 Hasil Pengujian Metode

Algoritma	Accuracy
KNN (k=7)	92,39%
Naïve Byes	91,83%
KNN (k=3) Naïve Bayes	91,24%
Naïve Bayes KNN (k=3)	93,53%

Implementasi

Setelah memperoleh model terbaik, peneliti melakukan pengujian model dengan pengimplementasikan data uji ke dalam model. Dari

hasil implementasi data uji ke dalam model diperoleh hasil klasifikasi yang dapat dilihat dari gambar 12 sebagai berikut :

Row	Y	confidence	P	confidence	confidence	prediction(Y)	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
1	?	1.000	0.000	0	PS	Y	N	Y	T	N	N	N	T	T	Y	Y	Y	N
2	?	0.969	0.000	0.031	PS	Y	N	N	N	N	N	N	T	Y	Y	Y	Y	N
3	?	0.886	0.000	0.114	PS	N	N	N	N	N	N	N	T	Y	Y	Y	Y	N
4	?	0.969	0.000	0.031	PS	Y	N	N	N	N	N	N	T	Y	Y	Y	Y	N
5	?	1.000	0.000	0.000	PS	N	N	Y	T	Y	N	N	T	T	Y	Y	Y	N

Gambar 12 Hasil klasifikasi keluarga sehat dengan 5 record data uji

Setelah melakukan berbagai tahapan atau beberapa rangkaian uji coba atau eksperimen dengan menerapkan metode KNN, metode Naïve Bayes dan kombinasi metode antara metode KNN dengan metode Naïve Bayes, hasil eksperimen menunjukkan bahwa hasil analisa dari metode KNN dengan menggunakan nilai ketetangaan yang dimulai dari nilai k3, k5, k7 dan k9 dengan nilai validasi 10 masing-masing memperoleh nilai akurasi sebesar KNN(k3)= 91,24%, KNN(k5)= 90,62%, KNN(k7)= 92,39% dan KNN(k9)= 88,30%, dari beberapa eksperimen yang dilakukan dari metode KNN, nilai yang paling baik diperoleh dari KNN(k7) dengan nilai akurasi sebesar 92,39%, metode Naïve Bayes dengan nilai validasi 10 memperoleh nilai akurasi sebesar 91,83%, kemudian eksperimen dengan kombinasi dua metode yaitu metode KNN Naïve Bayes dengan nilai k3, k5, k7 dan k9 memperoleh nilai secara konstan atau tetap yaitu sebesar 91,24%, dilanjutkan dengan kombinasi Naïve Bayes KNN masing-masing memperoleh nilai, Naïve Bayes KNN(k3) =93,53%, Naïve Bayes KNN(k5) =91,80%, Naïve Bayes KNN(k7) =92,390%, Naïve Bayes KNN(k9) =90,65%. Melihat dari hasil eksperimen pengujian metode KNN, Naïve Bayes, KNN Naïve Bayes dan Naïve Bayes KNN dapat diketahui bahwa hasil kombinasi metode yang paling baik adalah kombinasi antara Naïve Bayes dengan KNN (3) dengan nilai akurasi sebesar 93,53%, sehingga diantara metode yang sudah diketahui hasilnya maka metode yang paling baik untuk diterapkan dalam menganalisa keluarga sehat adalah kombinasi metode Naïve Bayes KNN.

IV. KESIMPULAN

Berdasarkan dari hasil *eksperimen* yang dilakukan dengan menggunakan beberapa metode dan kombinasi metode, seperti metode *K Nearest Neighbor (KNN)*, metode *Naive Bayes*, kombinasi *KNN* dengan *Naive Bayes* dan *Naive Bayes* dengan *KNN*. Masing-masing memperoleh nilai akurasi yang cukup baik seperti metode *Knearest Neighbor (KNN)* memperoleh nilai akurasi terbaiknya yaitu pada nilai $K=7$ yaitu sebesar 92,39%. Metode *Naive Bayes* yaitu sebesar 91,83%, kombinasi *KNN Naive Bayes* memperoleh nilai akurasi yang konsisten dari nilai $K=3,5,7$ dan 9 memperoleh nilai akurasi sebesar 91,24 % dan kombinasi *KNN Naive Bayes* memperoleh nilai akurasi terbaiknya pada nilai $K=3$ yaitu sebesar 93,53%. Dengan demikian dari beberapa *eksperimen* nilai

akurasi yang terbaik diperoleh dari kombinasi metode *Naive Bayes KNN*.

DAFTAR PUSTAKA

- [1] Emilia Khairani Majid, Mapeaty Nyorong, Nuraini, Evaluasi Implementasi Supervisi Keperawatan di Ruang Rawat Inap Terpadu Rumah Sakit H. Adam Malik Medan, JUMANTIK Volume 7 No.2 Mei 2022
- [2] M. R. Fadli, “Memahami desain metode penelitian kualitatif,” vol. 21, no. 1, pp. 33–54, 2021.
- [3] N. M. Putry and B. N. Sari, “Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Melitus,” *Evolusi J. Sains dan Manaj.*, vol. 10, no. 1, pp. 45–57, 2022.
- [4] A. A. Syafitri Hidayatul AA, Yuita Arum S, “Seleksi Fitur Information Gain untuk Klasifikasi Penyakit Jantung Menggunakan Kombinasi Metode K-Nearest Neighbor dan Naïve Bayes,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 2546–2554, 2018.
- [5] D. A. Nazaruddin, F. A. Bachtiar, and R. K. Dewi, “Klasifikasi Penyakit Kelamin Pada Wanita Dengan Menggunakan Kombinasi Metode K-Nearest Neighbor Dan Naïve Bayes Classifier,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 4, pp. 3266–3274, 2019.
- [6] N. Salamung et al., *Keperawatan Keluarga (Family Nursing)*. 2021.
- [7] E. K. Majid and M. Nyorong, “Evaluasi Implementasi Supervisi Keperawatan di Ruang Rawat Inap Terpadu Rumah Sakit H . Adam Malik Medan,” vol. 7, no. 2, pp. 103–116, 2022.
- [8] Y. Melia, “KLASIFIKASI MENGGUNAKAN NAÏVE BAYES DAN K-NEAREST NEIGHBOR,” no. July, 2020.
- [9] Z. Arifin, “Penerapan Metode Knn (K-Nearest Neighbor) Dalam Sistem Pendukung Keputusan Penerimaan Kip (Kartu Indonesia Pintar) Di Desa Pandean Berbasis Web Dan Mysql,” *NJCA (Nusantara J. Comput. Its Appl.*, vol. 4, no. 1, 2019.
- [10] P. Merupakan, T. Tidak, P. S. Informatika, F. Sains, D. A. N. Teknologi, and U. S. Dharma, “PENERAPAN METODE NAIVE BAYES Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh Gelar Sarjana Komputer Program Studi Informatika Disusun Oleh : Katarina Haryaningtyas THE IMPLEMENTATION OF

NAIVE BAYES METHOD TO CLASSIFY OBESITY LEVELS BASD ON EATING HAB," 2022.

- [11] M. Kompetensi, K. Karyawan, D. I. Pt, J. Nugraha, and A. Amalia, "Jurnal Ekonomi dan Bisnis , Vol . 9 No . 2 Maret 2022 P - ISSN : 2503-4413 IMPLEMENTASI PENGEMBANGAN SUMBER DAYA MANUSIA DALAM UPAYA EKAKURIR COUNTER AGEN PARK ROYAL SIDOARJO Oleh ;," vol. 9, no. 2, 2022.
- [12] J. Lexi and M. M.A., Metodologi Penelitian Kualitatif. In Metodologi Penelitian Kualitatif, no. March. 2010.