

Klasifikasi Penyakit Dengue Menggunakan Algoritma K-Nearest Neighbors Berbasis Flask

¹Isyaa Akhsanil Khakim, ²Abdul Rozaq
Universitas PGRI Madiun
Madiun, Indonesia

isak.akhsanil@gmail.com, rozaq@unipma.ac.id

*Corresponding Author

Diajukan : 01/07/2022

Diterima : 23/07/2022

Dipublikasi : 01/08/2022

ABSTRAK

Penanggulangan penyakit dengue merupakan salah satu elemen penting dalam memastikan masyarakat hidup sehat dan mendapatkan penanganan kesehatan yang baik. Penanggulangan penyakit dengue yang terlambat dapat mengakibatkan kematian. Berdasarkan permasalahan tersebut, maka diperlukan suatu sistem klasifikasi yang dapat mendiagnosis penyakit demam berdarah berdasarkan pemeriksaan kesehatan pasien. Tujuan dari penelitian ini adalah untuk menerapkan algoritma *k-nearest neighbor* pada pembangunan sistem klasifikasi dengue. Algoritma *K-Nearest Neighbors* bekerja dengan menetapkan kelas dari suatu objek berdasarkan kesamaan atribut dari data yang sebelumnya dijadikan model. Data model yang dibangun menggunakan data kasus *dengue* yang terjadi di kota madiun dengan total data sebanyak 276 baris dengan kolom data sebanyak 7 kolom. Data model tersebut nantinya akan di *pickle* menggunakan *library* yang ada pada *python*. Setelah di *pickle* kemudian data di split menjadi data *training* sebanyak 70% dan data *testing* sebanyak 30% yang nantinya akan digunakan sebagai bahan untuk melakukan prediksi data. Metode pengembangan sistem untuk penelitian ini menggunakan metode extreme programming yang terdiri dari empat tahap yaitu: (1) perencanaan, (2) perancangan, (3) pengkodean, dan (4) pengujian. Pengembangan sistem dilakukan pada aplikasi berbasis web dengan menggunakan framework Flask menggunakan bahasa pemrograman Python, dengan akurasi rata-rata 0,72, akurasi rata-rata 0,83, nilai recall rata-rata 0,59, dan nilai rata-rata f-1 0,64. Itu telah dibuat atau sebesar 64%.

Kata Kunci: Klasifikasi, *K-Nearest Neighbors*, *Extreme Programming*, *Framework Flask*, *Dengue*.

I. PENDAHULUAN

Perkembangan teknologi dan ilmu pengetahuan Indonesia dari tahun ke tahun sangat pesat khususnya dalam bidang teknologi informasi. Ada banyak indikasi bahwa hampir semua aktivitas manusia didukung oleh sistem komputer. Perkembangan teknologi ini juga akan mempengaruhi dunia kesehatan. Oleh karena itu, semua fasilitas kesehatan perlu mengikuti perkembangan teknologi saat ini. Salah satu instansi kesehatan adalah Kementerian Kesehatan.

Dinas Kesehatan Kota Madiun merupakan salah satu lembaga Kesehatan di Kota Madiun, dan salah satu kasus yang sedang ditangani oleh tenaga kesehatan saat ini adalah kasus *dengue* yang kian meningkat. Saat ini sistem pendeteksian *dengue* masih terbatas dengan tes yang dilakukan pihak rumah sakit. Hal ini membuat beberapa masyarakat yang berada di daerah yang rawan terjangkit *dengue* terutama masyarakat yang belum mengetahui indikator serta klasifikasi tingkatan *dengue* tidak dapat melakukan tes mandiri untuk segera mengetahui tingkat penyakit *dengue* yang mereka rasakan agar dapat memperoleh penanganan kesehatan berikutnya.

Salah satu teknologi yang dapat digunakan untuk mengklasifikasikan *dengue* adalah teknologi data mining algoritma K-nearest neighbor (K-NN). Metode *k-nearest neighbor* merupakan algoritma takson yang bertujuan untuk membuat prediksi berdasarkan data yang ada dengan menghitung jarak terdekat dari kumpulan data yang besar. Data ini diperoleh dari kasus *dengue* sebelumnya untuk menjadi data training menentukan klasifikasi DBD selanjutnya. Python merupakan bahasa pemrograman yang menyediakan library algoritma K-Nearest Neighbors selain itu *python* juga menyediakan *library flask* untuk pengembangan aplikasi berbasis *web* dan modul bernama *pickle* yang berfungsi untuk menyimpan serta membaca data didalam suatu file.

Maka sesuai dengan penjelasan penulis mengajukan judul penelitian dengan judul “Klasifikasi Penyakit Dengue Menggunakan Algoritma K-Nearest Neighbors (K-NN) Berbasis Flask”. Sistem yang dibuat ini agar dapat membantu tenaga kesehatan dalam mengetahui klasifikasi dengue yang sedang marak di Kota Madiun.

II. STUDI LITERATUR

Penelitian Terdahulu

Pada penelitian ini penulis mengacu pada penelitian sebelumnya yang telah dilakukan sebagai referensi. Pada penelitian yang dilakukan oleh (Dinata & Hasdyna, 2020), Tujuan dari penelitian ini adalah untuk membantu masyarakat umum menemukan sekolah menengah/ sederajat berdasarkan delapan kriteria: akreditasi, fasilitas, fasilitas atletik, laboratorium, kegiatan ekstrakurikuler, biaya, tingkat kelas, dan masa studi. Metode yang dipakai dalam penelitian ini adalah klasifikasi menggunakan algoritma K-nearest neighbor menggunakan pendekatan jarak Eludian. Dengan menggunakan 140 data latih dan 60 data uji diperoleh akurasi dengan nilai 79,62% $k = 5$.

Penelitian selanjutnya yang dilakukan oleh (Kafil, 2019), Tujuan dari penelitian ini adalah untuk mengetahui peramalan hasil penjualan di Boutiq Dealove Bondowoso dengan metode klasifikasi menggunakan algoritma *k-nearest neighbor* dan sistem prediksi berbasis web native PHP. Hasil uji akurasi dengan menggunakan 12 data latih dan 12 data uji menghasilkan hasil akurasi sebesar 83,3% dengan nilai error sebesar 16,7%.

Selanjutnya pada penelitian yang dilakukan oleh (Siregar et al., 2019) Tujuan dari penelitian ini adalah untuk mengefisiensi pengolahan data komentar, dan analisis sentimen berlaku untuk proses menganalisis apakah suatu komentar *positif* atau *negatif* dan proses klasifikasi komentar. Metode yang dipakai dalam penelitian ini adalah *classification* menggunakan algoritma *k-nearest neighbor*. Pengujian dilakukan dengan menguji pengaruh k menggunakan 64 data uji dan 11, 20, 120 data latih. Hasil uji k mencapai akurasi rata-rata 94,23%.

Extreme Programming

Pemrograman ekstrim adalah salah satu teknik pengembangan yang paling banyak digunakan oleh tim pengembangan perangkat lunak dalam lingkungan tangkas yang ideal. Tim tangkas adalah tim lintas fungsi kecil dengan semua keterampilan yang dibutuhkan untuk memberikan peningkatan produk secara teratur dan sering. Anda mandiri dan dapat bekerja tanpa ketergantungan eksternal (Paasivaara & Kruchten, 2020:219). Pemrograman ekstrim menekankan kerja tim. Manajer, pelanggan, dan pengembang semuanya adalah mitra yang setara dalam tim kolaborasi. Nilai inti Extreme Programming adalah kekompakan dan saling menghargai. (Baumeister et al., 2017:122).

K-Nearest Neighbors

Algoritma *k-nearest neighbor* (K-NN) merupakan algoritma yang melakukan klasifikasi berdasarkan posisi data yang dekat dengan data lain (jarak) (Eko, 2012:50). Algoritma K-NN dapat didefinisikan berdasarkan namanya. Dimana K adalah banyaknya bilangan terdekat (*nearest neighbor*) yang digunakan untuk mengidentifikasi kesamaan antara objek baru dengan tetangganya. (Daqiqil, 2021:134).

Adapun Langkah-langkah algoritma K-NN (Daqiqil, 2021:134) adalah sebagai berikut :
Tentukan nilai K Nilai K harus bilangan asli ganjil untuk menghindari hasil yang seimbang.
Gunakan rumus Euclidean untuk menghitung jarak antara titik baru dan semua data latih

Berikut rumus umum Euclidean :

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \dots \text{Persamaan (1)}$$

Keterangan :

- pi = data latih (data training)
- qi = data uji (data testing)
- i = variable data
- n = dimensi data

Pilih titik K yang paling dekat dengan titik baru.

Saat mengklasifikasikan, jumlah titik yang berdekatan dihitung berdasarkan kategori. Poin baru masuk dalam kategori dengan jumlah poin terdekat terbanyak. Di sisi lain, dalam kasus regresi, titik baru akan menjadi nilai rata-rata di dekat K.

Confusion Matrix

Ada empat istilah tambahan yang perlu diketahui yang merupakan "blok bangunan" yang digunakan dalam menghitung banyak ukuran evaluasi. Agar memudahkan untuk memahami arti dari berbagai ukuran (Han et al., 2012). Berikut penjelasan empat istilah yang perlu diketahui :

- a. True Positives (TP) Ini mengacu pada tupel *positif* yang diberi label dengan benar oleh pengklasifikasi. Biarkan TP menjadi jumlah *positif* sejati..
- b. True Negative (TN): Ini adalah tupel *negatif* yang diberi label dengan benar oleh classifier. Biarkan TN menjadi jumlah negatif benar.
- c. False Positives (FP): Ini adalah tupel *negatif* yang salah diberi label sebagai *positif* (misalnya, tupel kelas *buys_computer* = tidak yang diprediksi oleh pengklasifikasi *buys_computer* = ya). Biarkan FP menjadi jumlah *positif* palsu.
- d. False Negatif (FN): Ini adalah tupel positif yang salah dilabeli sebagai *negatif* (misalnya, tupel kelas *buys_computer* = *yes* yang diprediksi oleh pengklasifikasi *buys_computer* = tidak). Biarkan FN menjadi jumlah negatif palsu.

Tabel kelas prediksi Confusion Matrix sebagai berikut : Tabel confusion matrix dapat dilihat pada Tabel 1. (Han et al., 2012).

Tabel 1. Confusion Matrix

		Predicted class		Total
		Yes	No	
Actualy class	Yes	TP	FN	P
	No	FP	TN	N
	Total	P	N	P+N

Sumber : (Han et al., 2012)

Data Mining

Menurut Tan (dalam Eko, 2012:2) *Data mining* sebagai proses mengambil informasi yang berguna dari gudang *database* yang besar. *Data mining* juga dapat diartikan sebagai penggalian informasi baru dari sejumlah besar data yang berguna untuk pengambilan keputusan. *Istilah data minin* juga dikenal sebagai penemuan pengetahuan. *Data mining* adalah proses berulang yang kemajuannya ditentukan oleh deteksi otomatis atau manual (Kantardzic, 2020:2).

Klasifikasi

Klasifikasi merupakan suatu proses penilaian terhadap suatu objek yang nantinya akan di kelompokkan menjadi beberapa kelas yang telah di tentukan. Dalam klasifikasi terdapat dua proses utama yang harus dikerjakan yaitu pembuatan model sebagai kerangka untuk disimpan di dalam suatu penyimpanan dan pemakaian model guna melakukan classification pada suatu objek untuk mengetahui kelas suatu objek tersebut dalam model yang telah disimpan sebelumnya. Klasifikasi adalah pendekatan learning yang terjaga, yaitu merumuskan suatu data yang telah mempunyai kelas. Data yang di pakai adalah data yang berasal dari satu set contoh training, dimana setiap contoh training merupakan sepasang data input (Aggarwal, 2015:66).

Dangue

Demam berdarah disebabkan oleh setidaknya empat virus terpisah, yang semuanya menyebar terutama *Aedes aegypti* yang berkembang biak melalui air yang tergenang (Adelman, 2016). Penyakit ini masuk dalam kategori vector(mosquito-borne disae) yang banyak di hadapi dunia. Kasus pertama demam berdarah dilaporkan di Asia, Afrika, dan Amerika Utara dari tahun 1779 hingga 1780 (Suwandono, 2019:1). Infeksi virus dengue dapat dibagi menjadi tiga tahap: demam berdarah (DD), demam berdarah dengue (DBD), dan sindrom syok dengue (SSD) (Suwandono, 2019:7).

Python

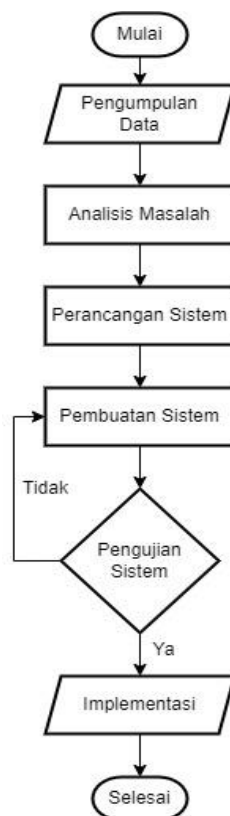
Python adalah bahasa yang ditafsirkan yang bekerja dengan membuat program membaca setiap baris yang ditulis dalam bahasa tingkat tinggi. Interpreter kemudian memproses setiap baris secara langsung dan mengeluarkan output. Jika terjadi kesalahan program selama eksekusi, Python akan menjalankan program secara otomatis. Ada dua cara untuk menjalankan juru bahasa python ini pada berikutnya yang memanggil python. Menggunakan command yang ada pada system windows. Yang kedua memakai IDLE, yang merupakan standar untuk python itu sendiri. (Wibowo et al., 2020:3). Python merupakan bahasa pemrograman yang efisien syntax program lebih sedikit dibandingkan dengan bahasa pemrograman yang lain, syntax python lebih mirip dengan algoritma pseudocode yang umum sehingga lebih mudah dipahami dan membutuhkan waktu yang cepat untuk memecahkan masalah (Wibawa et al., 2021:12).

Flask

Flask adalah kerangka kerja kecil menurut sebagian besar standar, sangat kecil sehingga disebut "kerangka mikro". Ini cukup kecil sehingga pengguna dapat membaca dan memahami semua kode sumber setelah pengguna terbiasa. Namun, menjadi kecil berarti tidak kalah dengan *framework* lainnya. Flask dirancang dari awal sebagai prototipe system yang dapat dikembangkan. Serta memiliki core yang solid dengan layanan dasar sementara ekstensi menyediakan sisanya (Grinberg, 2014:3).

III. METODE

Rancangan penelitian yang digunakan dapat dilihat pada Gambar 1, sebagai berikut:



Gambar 1. Flowchart rancangan penelitian

Keterangan:

A. Pengumpulan Data

Pada tahap ini dilakukan pengambilan data dengan metode wawancara, observasi langsung dan studi pustaka. Pelaksanaan wawancara dan observasi dilakukan di Dinas Kesehatan Kota Madiun.

B. Analisis Masalah

Data yang telah terkumpul kemudian dibuat user story untuk dijadikan dasar dalam pembangunan sistem yang ramah pada pengguna.

C. Perancangan Sistem

Pada tahap ini dilakukan perancangan alur kerja sistem, pemodelan alur kerja sistem menggunakan Unified Modelling Language (UML), perhitungan K-Nearest Neighbors serta perancangan User Interface (antarmuka) sistem.

D. Pembuatan Sistem

Pada tahap ini dilakukan penulisan kode program sesuai rancangan yang telah dilakukan. Selanjutnya, modul-modul yang telah jadi dilakukan pengujian apakah terdapat ketidaksesuaian dalam pengerjaan.

E. Pengujian Sistem

Pada tahap ini sistem diuji untuk memastikan seluruh alur kerja sistem bekerja sesuai dengan fungsinya.

F. Implementasi

Pada tahap ini sistem yang sudah dilakukan pengujian dapat diimplementasikan dengan menambah hosting supaya mendapat akses terhadap internet dan dapat diakses secara terbuka oleh masyarakat daerah terdampak dalam mengidentifikasi *dengue* pada pasien.

IV. HASIL DAN PEMBAHASAN

Data merupakan data sekunder yang diambil dari Dinas Kesehatan Kota Madiun 2021-2022. Data yang diambil yaitu kasus penyakit dengue pada periode tahun 2021-2022 bersumber dari Dinas Kesehatan Kota Madiun.

A. Data Training dan Data Testing

Data training diperlukan sebagai pemodelan machine learning dengan kebutuhan data sebanyak 7 kolom yang terdiri dari usia, jenis kelamin (JK), nilai trombosit terendah (NTT), nilai Hemoglobin (NHB), nilai hematokrit terendah (NHR), nilai hematokrit tertinggi (NHT), label (Demam Dengue (DD), Demam Berdarah Dengue (DBD) dan Dengue Shock Syndrome (DSS)). Data training yang digunakan sebanyak 70% dari total keseluruhan data yaitu sebanyak 193 baris. Pembagian data training dapat dilihat pada Tabel 2 sebagai berikut

Tabel 2. Data Training

No	Usia	JK	NTT	NHB	NHR	NHT	Label
1	46	L	72	15	46	47	DD
2	11	L	47	13	37	38	DD
3	3	L	111	12	37	38	DD
4	31	P	76	14	34	40	DD
5	9	L	74	13	40	41	DBD
6	1	L	77	11	31	33	DD
7	1	L	104	12	35	38	DD
8	20	L	39	16	50	52	DD
9	2	P	99	11	29	33	DBD
10	3	L	84	11	30	31	DD
...
193	10	P	13	19	55	57	DSS

Data testing diperlukan sebagai validasi dan pengukuran peforma terhadap pemodelan machine learning yang dibuat data sebanyak 7 kolom yang terdiri dari usia, jenis kelamin (JK), nilai trombosit terendah (NTT), nilai Hemoglobin (NHB), nilai hematokrit terendah (NHR), nilai hematokrit tertinggi (NHT), label (Demam Dengue (DD), Demam Berdarah Dengue (DBD) dan Dengue Shock Syndrome (DSS)). Data testing yang digunakan sebanyak 30% dari total keseluruhan data yaitu sebanyak 83 baris. Pembagian data testing dapat dilihat pada Tabel 3 sebagai berikut.

Tabel 3. Data Testing

No	Usia	JK	NTT	NHB	NHR	NHT	Label
1	19	L	69	14	35	38	DBD
2	8	P	128	14	40	43	DD
3	6	P	94	13	36	37	DD
4	2	L	61	12	37	38	DBD
5	6	P	108	13	38	39	DD
6	3	P	95	14	36	39	DD
7	48	P	65	13	37	41	DD
8	17	L	64	15	42	47	DD
9	44	L	184	14	38	44	DBD
10	3	L	78	13	36	37	DD
...
83	10	P	13	19	55	57	DSS

Sedangkan sebagai data target yang dihasilkan dapat dilihat pada Tabel 4.9. sebagai berikut.

Tabel 4.9. Variabel target

Variabel	Keterangan
1	Demam Dengue (DD)
2	Demam Berdarah Dengue (DBD)
3	Dengue Shock Syndrome (DSS)

B. Perhitungan Metode *K-Nearest Neighbors* (K-NN)

Sebelum melakukan perhitungan data akan di normalisasi terlebih dahulu. Normalisasi data dilakukan dengan tujuan menyamakan data dalam bentuk rata-rata sehingga antara data yang satu dengan data yang lainnya tidak memiliki nilai terlalu jauh yaitu dalam range 0 sampai 1. Normalisasi data dilakukan dengan menggunakan persamaan 1.

$$\frac{Data\ Input - Min}{Range} \dots\dots (1)$$

Data yang telah dilakukan normalisasi dapat dilihat pada Tabel 4 sebagai berikut.

Tabel 4. Data normalisasi

No	Usia	JK	NTT	NHB	NHR	NHT	Label
1	1,00	0,00	0,49	0,29	0,00	0,10	DBD
2	0,35	1,00	1,00	0,29	0,25	0,33	DD
3	0,00	0,00	0,42	0,00	0,10	0,10	DBD
4	0,24	1,00	0,83	0,14	0,15	0,14	DD
5	0,29	1,00	0,14	0,00	0,00	0,00	DBD
6	0,47	1,00	0,00	1,00	1,00	1,00	DSS

Perhitungan jarak dilakukan dengan menghitung setiap jarak dari setiap data training terhadap data testing. Data yang digunakan diambil dari data nomor 1 sebagai data testing dan data nomor 2 sampai seterusnya sebagai data training.

Perhitungan jarak data training menggunakan euclidean distance dengan menggunakan persamaan 2.

$$euc = \sqrt{(\sum_{i=1}^n (p_i - q_i)^2)} \dots(2)$$

Perhitungan jarak data ke-2 terhadap data testing.

$$d = \sqrt{(0,35 - 1)^2 + (1 - 0)^2 + (1 - 0,49)^2 + (0,29 - 0,29)^2 + (0,25 - 0)^2 + (0,33 - 0,10)^2} = 1,34205$$

Perhitungan jarak data ke-3 terhadap data testing.

$$d = \sqrt{(0 - 1)^2 + (0 - 0)^2 + (0,42 - 0,49)^2 + (0 - 0,29)^2 + (0,10 - 0)^2 + (0,10 - 0,10)^2} = 1,04713$$

Perhitungan jarak data ke-4 terhadap data testing.

$$d = \sqrt{(0,24 - 1)^2 + (1 - 0)^2 + (0,83 - 0,49)^2 + (0,14 - 0,29)^2 + (0,15 - 0)^2 + (0,14 - 0,10)^2} = 1,32097$$

Perhitungan jarak data ke-5 terhadap data testing.

$$d = \sqrt{(0,29 - 1)^2 + (1 - 0)^2 + (0,14 - 0,49)^2 + (0 - 0,29)^2 + (0 - 0)^2 + (0 - 0,10)^2} = 1,30765$$

Perhitungan jarak data ke-6 terhadap data testing.

$$d = \sqrt{(0,47 - 1)^2 + (1 - 0)^2 + (0 - 0,49)^2 + (1 - 0,29)^2 + (1 - 0)^2 + (1 - 0,10)^2} = 1,96117$$

Berdasarkan perhitungan jarak data training sebanyak 5 data terhadap data testing menggunakan euclidean distance didapat hasil jarak terdekat terhadap data testing ditemukan pada data ke-3 yang berlabel “DBD” dengan jarak sebesar 1,04713.

C. Perhitungan *Confusion Matrix*

Evaluasi performa dilakukan menggunakan teknik Confusion Matrix untuk mengukur performa model data yang dibuat. Evaluasi yang dilakukan menggunakan teknik confusion matrix meliputi akurasi, presisi recall dan f-1 score. Hasil Evaluasi menggunakan confusion matrix dapat dilihat pada Tabel 5.

Tabel 5. Hasil Evaluasi Confusion Matrix

Actual Class	Classified Class		
	1	2	3
1	39	15	0
2	16	20	0
3	2	0	1

Pada tabel 5 terdapat actual class (kelas sebenarnya) dan classified class (kelas klasifikasi). Actual class merupakan kelas yang sebelumnya sudah ditentukan kelas/labelnya, sedangkan Classified class merupakan kelas hasil klasifikasi menggunakan metode K-Nearest Neighbors (K-NN).

Hasil dari perhitungan confusion matrix pada tabel menghasilkan label 1 (DD) terdapat 39 data uji yang terklasifikasi secara benar oleh sistem dan 15 data yang terklasifikasi salah. label 2 (DBD) terdapat 20 data uji yang terklasifikasi secara benar oleh sistem dan 16 data yang terklasifikasi salah. label 3 (DSS) terdapat 1 data uji yang terklasifikasi secara benar oleh sistem dan 2 data yang terklasifikasi salah.

Akurasi, presisi, recall dan f-1 score masing-masing dihitung dengan modul confusion matrix pada library sklearn menggunakan jupyter notebook. Perhitungan nilai akurasi, presisi, recall dan f-1 score dapat dilihat pada Tabel 6.

Tabel 6. Perhitungan Nilai Akurasi, Presisi, Recall dan F-1 score

Label	precision	recall	f1-score
1	0.68	0.89	0.77
2	0.80	0.56	0.66
3	1.00	0.33	0.50
accuracy			0.72
macro avg	0.83	0.59	0.64

Pada Tabel 6 diperoleh hasil nilai rata-rata akurasi sebesar 0,72 atau sebesar 72%, nilai rata-rata presisi sebesar 0,83 atau sebesar 83%, nilai rata-rata recall sebesar 0,59 atau sebesar 59% dan nilai rata-rata f-1 score sebesar 0,64 atau sebesar 64%.

D. Implementasi Sistem

Halaman prediksi merupakan halaman inti dari sistem klasifikasi dengue. Pada halaman prediksi pengguna diarahkan untuk mengisi form data pasien sebanyak 6 form kemudian selanjutnya dilakukan proses klasifikasi terhadap data model yang sebelumnya telah dibuat. Proses prediksi berjalan dengan mengambil data dari input data pengguna yang kemudian diklasifikasi untuk menentukan posisi data pengguna tergolong dalam label terdekat berdasarkan kemiripan data. Label data terdiri sejumlah 3 jenis yang masing-masing adalah jenis dengue. Implementasi halaman prediksi dapat dilihat pada Gambar 3.

Sistem Klasifikasi Dengue

Beranda Prediksi Panduan

Apa Hasil Pemeriksaan Lab Anda ?

Umur:

Jenis Kelamin:

Nilai Trombosit Terendah:

Nilai HB (Hemoglobin):

Nilai Hematokrit Terendah:

Nilai Hematokrit Tertinggi:

Isyaaq Akhsani Khakim | Teknik Informatika | Skripsi 2022

Gambar 2. Halaman form prediksi

Sistem Klasifikasi Dengue

Beranda Prediksi Panduan

Apa Hasil Pemeriksaan Lab Anda ?

Umur:

Jenis Kelamin:

Nilai Trombosit Terendah:

Nilai HB (Hemoglobin):

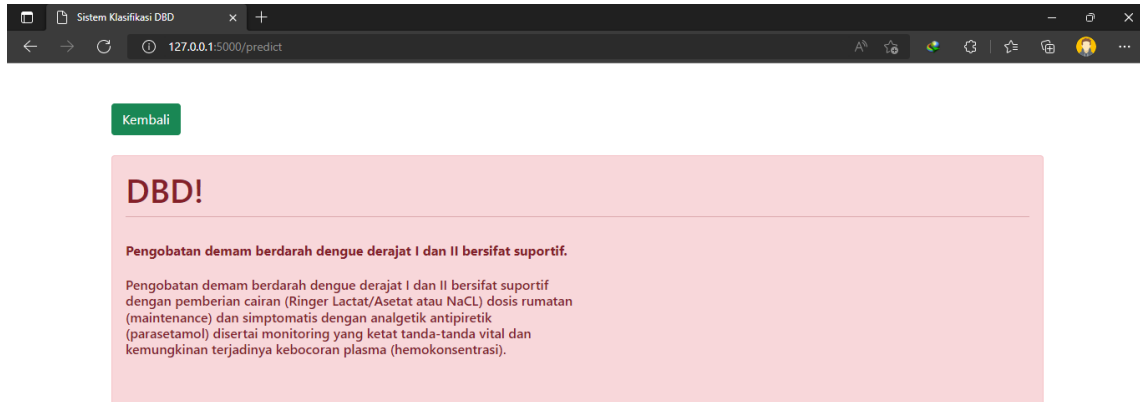
Nilai Hematokrit Terendah:

Nilai Hematokrit Tertinggi:

Isyaaq Akhsani Khakim | Teknik Informatika | Skripsi 2022

Gambar 3. Form prediksi ketika pengisian tidak terisi semuanya

Halaman hasil prediksi menampilkan data pasien yang di-input oleh pengguna dan telah mengalami proses klasifikasi data. Label data dengue akan tampil pada layar pengguna berikut dengan solusi atau langkah yang tepat yang dapat diambil oleh pengguna berdasarkan dengue pada data pasien. Implementasi halaman beranda dapat dilihat pada Gambar 4.



Gambar 4. Halaman hasil prediksi

V. KESIMPULAN

Berdasarkan hasil dari perancangan yang dibuat, penulis dapat menyimpulkan dari hasil yang dilakukan yaitu:

1. Sistem klasifikasi yang dibangun menggunakan bahasa pemrograman python dengan basis web-based application menggunakan framework flask. Metode klasifikasi yang digunakan adalah algoritma k-nearest neighbors (KNN) sebagai model dalam menentukan klasifikasi dengue
2. Luaran daripada penelitian ini adalah sebuah sistem klasifikasi identifikasi dengue yang dapat digunakan untuk membantu para tenaga medis dalam mengidentifikasi dengue pada masyarakat terdampak dengue berdasarkan pemeriksaan pada kesehatan pasien.
3. Perhitungan performa model data klasifikasi k-nearest neighbors yang dibuat dihitung menggunakan confusion matrix menghasilkan nilai akurasi sebesar 72%, nilai presisi sebesar 83% dan nilai recall sebesar 59%.

Saran

Berdasarkan kesimpulan yang telah diuraikan, terdapat beberapa saran yang dapat disampaikan yaitu:

1. Pada peneliti selanjutnya dapat diperbanyak pada dataset yang digunakan, semakin banyak data yang digunakan, model yang dibuat dapat menghasilkan akurasi klasifikasi yang lebih tinggi. Sehingga klasifikasi yang didapat akan semakin akurat.
2. Pada peneliti selanjutnya dapat dikembangkan dengan menambah fitur yang akan membuat hasil dari prediksi bisa digunakan lebih maksimal
3. Pada peneliti selanjutnya sistem dapat menggunakan algoritma yang lain sehingga mendapatkan perbandingan yang lebih presisi.

VI. REFERENSI

- Adelman, Z. N. (2016). *Genetic Control of Malaria and Dengue* (Z. N. Adelman (ed.)). Sara Tenney.
- Aggarwal, C. C. (2015). *Data Classification Algorithms and Applications* (C. C. Aggarwal (ed.)). CRC Press.
- Baumeister, H., Lichter, H., & Riebisch, M. (2017). *Agile Processes in Software Engineering and Extreme Programming* (H. Baumeister, H. Lichter, & M. Riebisch (eds.)). Springer Nature.
- Daqiqil, I. (2021). *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python* (I). UR PRESS.
- Dinata, R. K., & Hasdyna, N. (2020). *Wilayah Bireuen Menggunakan Algoritma K-Nearest Neighbors Berbasis Web*. 5(1), 33–37.
- Eko, P. (2012). *Data Mining - Konsep Dan Aplikasi Menggunakan Matlab* (W. Nikodemus (ed.)).

CV. Andi Offset.

- Grinberg, M. (2014). *Flask Web Development* (M. B. and R. Roumeliotis (ed.); Edisi Pert). O'Reilly Media, Inc.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (Edisi Keti). Elsevier Inc.
- Kafil, M. (2019). *Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso*. 3(2), 59–66.
- Kantardzic, M. (2020). Data Mining Concepts, Models, Methods, and Algorithms. In *Gastronomía ecuatoriana y turismo local*. (Third Edit, Vol. 1, Issue 69). John Wiley & Sons, Inc.
- Paasivaara, M., & Kruchten, P. (2020). Agile Processes in Software Engineering and Extreme Programming – Workshops. In M. Paasivaara & P. Kruchten (Eds.), *XP 2020 Workshops Copenhagen, Denmark, June 8–12, 2020 Revised Selected Papers*. Springer Nature Switzerland AG.
- Siregar, Z. U., Siregar, R. R. A., & Arianto, R. (2019). *Klasifikasi Sentiment Analysis Pada Komentar Peserta Diklat Menggunakan Metode K-Nearest Neighbor*. 8(1), 81–92.
- Suwandono, A. (2019). *Dengue Update Menilik Perjalanan Dengue di Jawa Barat* (A. Suwandono (ed.); Edisi Pert). LIPI Press.
- Wibawa, A. P., Andika Dwiyanto, F., Widyaningtyas, T., & Mahmudy, W. F. (2021). *Struktur Data dengan Python* (Y. Umayu (ed.); Pertama). Ahlimedia Press.
- Wibowo, W., Ulama, B. S. S., & Azies, H. Al. (2020). *Belajar Pemrograman Bahasa Python* (Edisi Pert). ITS Press.