

KLASIFIKASI PENYAKIT DIABETES MELLITUS (DM) MENGUNAKAN ALGORITMA NAÏVE BAYES PADA KLINIK PRATAMA JAMBI

Lucy Simorangkir¹, Akenda Putra²

^{1,2}Program Studi Teknik Informatika, Universitas Nurdin Hamzah, Jambi

Email: lucy.simorangkir@yahoo.co.id, akendaputra@gmail.com

Abstract - Diabetes Mellitus (DM) is a chronic metabolic disorder characterized by sustained hyperglycemia, arising from impaired blood sugar utilization. This can lead to multi-organ complications affecting the cardiovascular, ocular, renal, and nervous systems. This research addresses the prevalent global health issue of Diabetes Mellitus (DM) by developing a classification system for early detection. Utilizing the Naïve Bayes algorithm, the system aims to assist healthcare professionals in primary clinics and pharmacies, specifically U.K Jambi, in identifying potential DM cases. A Waterfall model guides the system development process, encompassing needs analysis, design, implementation, and testing phases. Implementation is achieved through the Python programming language. Training data comprises 100 patient medical records, encompassing 16 DM symptoms. System testing employs new patient data input, including age, gender, and presenting symptoms. The Naïve Bayes algorithm processes this data to compute the probability of positive or negative DM classification. System output comprises a binary classification result (positive/negative) for DM, accompanied by performance metrics: accuracy, recall, and precision. Evaluation results demonstrate an 80% accuracy rate, with a recall of 0.82 and precision of 0.87, underscoring the efficacy of the Naïve Bayes algorithm for DM classification within this context.

Keywords: Diabetes Mellitus, Classification, Naïve Bayes, Python, Waterfall Model, Accuracy, Recall, Precision..

I. PENDAHULUAN

1.1. Latar Belakang

Kesehatan adalah hal yang sangat penting bagi semua makhluk hidup, terutama manusia. Untuk menghindari penyakit, kita harus menjaga makanan yang masuk ke dalam tubuh kita dengan mengatur pola makan dan berolahraga secara teratur. Diabetes atau Diabetes Mellitus (DM) adalah penyakit metabolik kronis yang ditandai dengan tingginya kadar gula darah dalam jangka waktu yang lama. Kondisi ini terjadi karena tubuh tidak dapat menggunakan gula darah secara maksimal. Diabetes dapat menyebabkan kerusakan pada jantung, pembuluh darah, mata, ginjal, dan saraf.

Berdasarkan penyebabnya, DM menurut *World Health Organization* (WHO) diklasifikasikan menjadi DM tipe 1, DM tipe 2, dan DM Gestasional (Irianto, 2014). DM gestasional adalah jenis diabetes yang hanya terjadi pada kehamilan karena perubahan hormonal. DM tipe 1 dan tipe 2 pada umumnya memiliki gejala yang sama, hanya saja rentang waktu kemunculan berbeda. DM tipe 1 memiliki gejala yang datang secara seketika dan berkembang dengan cepat dalam waktu beberapa minggu, sedangkan DM tipe 2 memiliki gejala yang tampak tidak jelas, tetapi akan memburuk secara perlahan. Adapun gejala umum diabetes dari berbagai tipe adalah *poliuria*, *polidipsia*, *polifagia*, penurunan berat badan mendadak, kelelahan, obesitas, penyembuhan tertunda, penglihatan kabur,

gatal, *iritabilitas*, sariawan genital, *paresis parsial*, otot kaku, *alopecia*, dan lain-lain. (Najib, Nurcahyono, & Setiawan, 2019).

Klinik Pratama Jambi adalah suatu fasilitas kesehatan publik yang didirikan untuk memberikan perawatan kepada pasien luar. Pelayanan konsultasi penyakit diabetes pada Klinik Pratama Jambi memerlukan proses yang cukup rumit, dimana harus dilakukan pengetesan gula darah seperti tes gula darah sewaktu tes gula darah puasa, tes gula darah yang dilakukan 2 jam setelah makan, dan tes HbA1c. Tentunya proses tersebut cukup memakan waktu dan biaya dalam prosesnya, padahal dengan modal gejala penderita pasien diabetes yang ada sebelumnya dapat dilakukan prediksi untuk mengetahui penyakit yang diderita pasien dengan menggunakan metode pada data mining yaitu metode Naïve Bayes dalam klasifikasi penyakit diabetes.

Beberapa penelitian sebelumnya berkaitan dengan pengembangan klasifikasi penyakit diabetes mellitus. Penelitian oleh Khasanah, Nasution, dan Amijaya (2022) menggunakan algoritma Naïve Bayes Classifier untuk klasifikasi diabetes, dengan mempertimbangkan lima atribut utama, yaitu usia, jenis kelamin, status merokok, kadar glukosa, dan tekanan darah. Selanjutnya, Widodo, Anggraeni, dan Sutabri (2021) mengembangkan sistem pakar berbasis web untuk diagnosis diabetes, yang juga memanfaatkan algoritma Naïve Bayes. Sistem ini diuji dengan dataset pasien dan divalidasi melalui uji kepakaran, serta pengujian whitebox dan blackbox. Selain itu, penelitian Ridwan (2020) menggunakan algoritma Naïve Bayes untuk klasifikasi diabetes

dengan atribut yang lebih beragam, termasuk gejala klinis seperti poliuria, polidipsia, kehilangan berat badan mendadak, kelemahan, gatal, dan lainnya, sebanyak 16 atribut. Terakhir, penelitian oleh Najib, Nurcahyono, dan Setiawan (2019) menerapkan algoritma C4.4 untuk klasifikasi diagnosis diabetes mellitus, menggunakan atribut jenis kelamin, usia, GDP, GDS, dan HbA1c.

Metode Naive Bayes adalah salah satu metode klasifikasi yang populer dan efektif dalam *machine learning*. Kelebihan metode ini adalah sederhana dan mudah dipahami karena implementasinya cukup mudah meskipun menggunakan dataset kecil yang kemungkinan ada beberapa fitur dataset yang hilang atau tidak relevan sehingga dapat dengan mudah digunakan untuk beradaptasi dengan data baru. Berdasarkan kelebihan tersebut membuat metode Naive Bayes menjadi salah satu pilihan yang cukup tepat untuk mengetahui apakah pasien tersebut didiagnosa positif atau negatif diabetes.

Aplikasi yang dibangun menggunakan pemrograman Python yang memudahkan dalam membantu pemrosesan data. Output dari model program ini bisa menjadi rekomendasi dalam melakukan diagnosa awal penyakit diabetes mellitus. Hasil yang diharapkan dengan adanya aplikasi ini agar proses diagnosa dapat dilakukan lebih cepat dan efisien, sehingga memberikan manfaat yang signifikan dalam pelayanan kesehatan terkait diabetes.

1.2. Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah : “Bagaimana membangun sistem klasifikasi penyakit diabetes mellitus menggunakan algoritma Naives Bayes pada Klinik Pratama Jambi ?”.

1.3. Tujuan Penelitian

Adapun tujuan penelitian ini adalah untuk membangun sistem klasifikasi penyakit diabetes mellitus menggunakan algoritma Naives Bayes Pada Klinik Pratama Jambi.

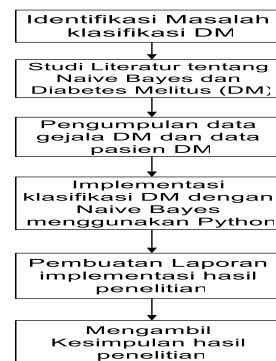
1.3 Manfaat Penelitian

Adapun manfaat penelitian ini adalah dapat membantu dalam deteksi dini penyakit diabetes berdasarkan gejala-gejala yang muncul, mendukung upaya pencegahan dan pengelolaan penyakit diabetes sejak dini.

1.4 Metode Penelitian

Adapun metode penelitian yang digunakan adalah *Waterfall*. Penerapan metode *Waterfall* mulai dari tahapan analisis, kebutuhan sistem,

perancangan, implementasi, dan pengujian dapat menghasilkan rancangan sistem aplikasi yang dibutuhkan oleh pihak klinik untuk pelayanan kepada pasien yang terkena diabetes mellitus. Berikut ini Gambar 1 menunjukkan tahapan penelitian mengikuti alur metode *Waterfall*.



Gambar 1. Tahapan Metode Penelitian

II. TINJAUAN PUSTAKA

2.1 Klasifikasi

Klasifikasi adalah sebuah proses yang menemukan properti-properti yang sama pada sebuah himpunan obyek didalam sebuah basis data dan mengklasifikasikannya ke dalam kelas-kelas yang berbeda menurut model klasifikasi yang diterapkan (Abdurrahman, 2022).

Klasifikasi adalah suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia (Putro, Vuldari, & Saptomo, 2020)

Sehingga dapat disimpulkan bahwa klasifikasi merujuk pada proses pengelompokan atau penataan objek, informasi, atau fenomena ke dalam kategori atau kelas berdasarkan karakteristik atau atribut tertentu yang dimilikinya.

2.2 Diabetes Mellitus

Diabetes adalah suatu penyakit metabolik dimana pankreas tidak cukup memproduksi insulin atau sel-sel dalam tubuh tidak dapat menggunakan insulin yang diproduksi dengan efektif. (Ridwan, A., 2020).

Diabetes mellitus adalah penyakit kronis yang ditimbulkan karena kekurangan produksi insulin (hormon yang diproduksi oleh pankreas untuk mengatur tingkat glukosa) dalam tubuh manusia. (Widodo, Anggraeini, & Sutabri, 2021).

Sehingga dapat disimpulkan bahwa diabetes mellitus adalah suatu kelompok penyakit kronis yang ditandai oleh kadar gula (glukosa) darah yang tinggi. Hal ini disebabkan oleh ketidakmampuan tubuh untuk memproduksi atau menggunakan hormon insulin dengan efisien. Insulin diperlukan

untuk mengatur kadar glukosa dalam darah, dan ketidakseimbangan ini dapat menyebabkan masalah kesehatan serius.

2.3 Naïve Bayes

Naïve bayes merupakan suatu bentuk klasifikasi data dengan menggunakan metode probabilitas dan statistik. (Waru & Astuti, 2021).

Naïve bayes adalah sebuah pengklasifikasian probabilistic sederhana yang menghitungkan sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dataset yang diberikan (Suntoro, J., 2019).

Sehingga dapat disimpulkan bahwa Naïve Bayes adalah sebuah metode klasifikasi statistik yang berdasarkan teorema Bayes. Secara singkat algoritma Naïve Bayes classification adalah pengklasifikasian kumpulan data statistika yang mana memprediksi semua probabilitas tiap anggota suatu class.

Adapun Langkah-langkah algoritma Naïve Bayes yaitu (Asfi & Fitrianiingsih, 2020, h.45)

1. Pengumpulan dan menyiapkan dataset.
2. Hitung semua kelas pada data training.
3. Hitung jumlah kasus yang sama dengan kelas yang sama.
4. Kalikan semua hasil dengan data testing yang akan dicari kelasnya dengan menggunakan persamaan, kemudian kalikan dengan hasil pada langkah kedua.
5. Bandingkan hasil perkalian nilai tertinggi ditetapkan sebagai kelas baru.

Dari langkah-langkah tersebut dapat dilihat pada rumus algoritma naïve bayes sebagai berikut :

$$P(H | X) = \frac{P(X | H)P(H)}{P(X)}$$

Rumus teorema Naïve bayes :

X = Data dengan class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

P(H | X) = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob)

P(H) = Probabilitas hipotesis (prior prob)

P(X | H) = Probabilitas X berdasarkan kondisi tersebut

P(X) = Probabilitas dari X

III. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data survey kuisioner gejala-gejala yang merujuk pada penyakit diabetes mellitus yang dialami penderita berdasarkan hasil diagnosa dokter, bagi yang sudah pernah mengalami ataupun yang belum mengalami diabetes. Data lainnya dihasilkan

dari data yang diambil melalui dataset UCI Machine Learning Repository yang nantinya digunakan sebagai sample dan disimpan kedalam satu tabel yang berjumlah 100 record. Berikut ini gambar 2 contoh data mentah hasil survey kuisioner

RESPONDEN	DATA KRITERIA																TARGET
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	
R1	23	L	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Negative
R2	25	L	Iya	Iya	Tidak	Iya	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Iya	Iya	Tidak	Tidak	Positive
R3	54	P	Iya	Tidak	Iya	Iya	Tidak	Tidak	Iya	Iya	Iya	Iya	Iya	Iya	Iya	Tidak	Positive
R4	63	P	Tidak	Iya	Tidak	Iya	Iya	Tidak	Iya	Iya	Iya	Iya	Iya	Iya	Iya	Tidak	Positive
R5	23	L	Iya	Iya	Iya	Iya	Tidak	Iya	Tidak	Tidak	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Positive
R6	34	L	Tidak	Iya	Iya	Iya	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Positive
R7	44	P	Tidak	Tidak	Tidak	Iya	Tidak	Tidak	Iya	Tidak	Tidak	Tidak	Iya	Iya	Tidak	Positive	
R8	29	P	Iya	Tidak	Tidak	Iya	Iya	Tidak	Tidak	Iya	Tidak	Iya	Iya	Iya	Tidak	Positive	
R9	68	L	Tidak	Tidak	Iya	Iya	Iya	Iya	Tidak	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Tidak	Positive
R10	31	P	Tidak	Tidak	Tidak	Iya	Tidak	Tidak	Tidak	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Negative
R11	34	L	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Iya	Positive
R12	39	P	Iya	Iya	Tidak	Iya	Iya	Iya	Tidak	Tidak	Tidak	Iya	Tidak	Iya	Tidak	Tidak	Positive
R13	54	P	Iya	Iya	Iya	Iya	Tidak	Tidak	Iya	Iya	Tidak	Iya	Iya	Tidak	Tidak	Tidak	Positive
R14	50	L	Iya	Iya	Tidak	Iya	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Positive
R15	46	P	Tidak	Iya	Iya	Iya	Iya	Tidak	Iya	Iya	Tidak	Tidak	Tidak	Tidak	Iya	Tidak	Positive
R16	54	L	Iya	Iya	Tidak	Iya	Tidak	Iya	Tidak	Iya	Iya	Iya	Iya	Tidak	Iya	Tidak	Positive
R17	68	L	Tidak	Iya	Tidak	Tidak	Iya	Iya	Tidak	Iya	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Positive
R18	49	P	Iya	Tidak	Tidak	Tidak	Iya	Iya	Tidak	Tidak	Iya	Iya	Tidak	Tidak	Iya	Tidak	Positive
R19	40	L	Tidak	Tidak	Tidak	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Iya	Iya	Tidak	Tidak	Tidak	Negative
R20	48	P	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Iya	Tidak	Iya	Tidak	Iya	Tidak	Tidak	Positive
R21	24	L	Tidak	Tidak	Tidak	Iya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Positive

Gambar 2. Data Mentah

Langkah selanjutnya adalah melakukan pemilihan data training, yang meliputi kriteria gejala dan target atribut yang diperlukan saat proses klasifikasi. Adapun data kriteria gejala yang diambil dapat dilihat pada tabel 1 berikut :

Tabel 1. Data Kriteria

Kriteria	Keterangan
Umur (Thn) (X1)	≤ 25 = 1, 26-45 = 2, >45 = 3
Jenis kelamin (X2)	Laki- Laki = 1 Perempuan = 2
Sering buang air (X3)	Iya = 1, Tidak = 2
Sering haus (X4)	Iya = 1, Tidak = 2
Penurunan berat badan (X5)	Iya = 1, Tidak = 2
Lemas (X6)	Iya = 1, Tidak = 2
Nafsu makan tinggi (X7)	Iya = 1, Tidak = 2
Jamur pada area sensitif (X8)	Iya = 1, Tidak = 2
Penglihatan kabur (X9)	Iya = 1, Tidak = 2
Gatal (X10)	Iya = 1, Tidak = 2
Emosi tidak stabil (X11)	Iya = 1, Tidak = 2
Luka sulit sembuh (X12)	Iya = 1, Tidak = 2
Nyeri atau kesemutan (X13)	Iya = 1, Tidak = 2
Kram (X14)	Iya = 1, Tidak = 2
Kerontokan (X15)	Iya = 1, Tidak = 2
Obesitas (X16)	Iya = 1, Tidak = 2
Target	(+) = 1, (-) = 0

Setelah melakukan pemilihan data, lakukan pembersihan data yang bertujuan untuk pra pemrosesan data yang dilakukan sebelum melakukan *mining* data. Pembersihan data berisi beberapa kegiatan yang bertujuan untuk melakukan pengenalan dan perbaikan pada data yang akan diteliti. Perbaikan pada data perlu dilakukan karena data mentah cenderung tidak siap untuk di-*mining*.

3.2 Perhitungan Klasifikasi Naïve Bayes

Setelah seluruh data yang diperlukan dikumpulkan, selanjutnya dapat dilakukan proses perhitungan dengan Algoritma Naïve Bayes dengan menggunakan Data pada Gambar 2 yang berjumlah 100 data dengan kriteria pada Tabel 1 sebelumnya. Adapun langkah perhitungan algoritma Naïve Bayes

IV. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil pembahasan di atas, maka dapat disimpulkan sebagai berikut:

1. Aplikasi Website Kelurahan Solok Sipin Berbasis Web ini, dibangun menggunakan bahasa pemrograman web dan MySQL sebagai databasenya.
2. Aplikasi Website Kelurahan Solok Sipin berbasis web ini, merupakan suatu cara untuk memudahkan pihak kelurahan menginput data penduduk yang berada di wilayah kelurahan Solok Sipin.
3. Aplikasi Website Kelurahan Solok Sipin berbasis web ini, dapat menyajikan informasi data penduduk, informasi data kartu keluarga, informasi data mutasi warga pendatang ataupun keluar, dan informasi lainnya.
4. Aplikasi Website Kelurahan Solok Sipin ini berbasis web ini, dapat menampilkan laporan yang terdiri dari laporan penduduk, laporan kartu keluarga, laporan mutasi warga yang pendatang dan pindah dan laporan lainnya.

4.2. Saran

Adapun saran-saran yang dapat dikemukakan adalah:

1. Mengadakan pelatihan kepada staff yang bertugas untuk mengolah data website Kelurahan Solok Sipin.
2. Diharapkan website ini dapat membantu dalam mengolah data penduduk, data kartu keluarga, data mutasi pendatang, mutasi keluar, data peristiwa kelahiran dan peristiwa kematian.
3. Diharapkan *website* ini memberikan kemudahan dalam memperoleh informasi, karena data tersimpan dalam *database* sehingga semua data dan informasi terdokumentasi dengan baik.
4. Diharapkan website Kelurahan Solok Sipin ini dapat dikembangkan lagi oleh peneliti selanjutnya.

DAFTAR REFERENSI

- Abdullah, R, 7 In 1 Pemrograman Web Untuk Pemula, 2018, In PT. Elex Media Komputindo (1 ed), Jakarta Pusat.
- Ahmad Azhar Kadim, I Ketut Sutriana, Irham H. Masir, Perancangan Sistem Aplikasi Layanan Kelurahan Berbasis Web, Vol. 4, No. 1, April 2022. <http://ejurnal.ung.ac.id/index.php/jji>.
- Al Hasri, M. V., & Sudarmilah, E, Sistem Informasi Pelayanan Administrasi Kependudukan Berbasis Website Kelurahan Banaran. *MATRIK: Jurnal Manajemen,*

Teknik Informatika dan Rekayasa Komputer, Vol 20 No 2, hh. 249-260, 2021.

<<https://doi.org/10.30812/matrik.v20i2.1056>>.

Anhar, S, Panduan Menguasai PHP & MySql Secara Otodidak, 2010, Transmedia Vol. 1999, No. December, PP. 1-6.

Ardhana, Kusuma YM, Project PHP & MySQL Membuat Website Buku Digital, 2014. Jasakom.

Arif Djunaidy, Wiwik Anggraeni, Faizal Mahananto, & Retno Aulia Vinarti, Pengembangan Website Kelurahan Gebang Putih Sukolilo Sebagai Penunjang Sarana Komunikasi Berbasis Teknologi Informasi, Vol 5, No 2, 2021.

Haswan, F, Perancangan Sistem Informasi Pendataan Penduduk Kelurahan Sungai Jering Berbasis Web Dengan Object Oriented Programming. *Jurnal Teknologi Dan Open Source*, Vol 1, No 2, hh 92-100, 2018.

<<https://www.ejournal.unlks.ac.id/Index.php/JTOS/artIcle/download/23/22>>.

Jogiyanto, HM, Analisis dan Perancangan Sistem Informasi, 2014, Wedatama Widya Sastra, DKI Jakarta.

Raharjo, M. M. I, Tata Kelola Pemerintahan Desa, 2021, Bumi Aksara.

Sari, A. R, Efektivitas Peran Kelurahan dalam Pelayanan Administrasi Kependudukan, 2021, Penerbit NEM.

Sutabri, Tata, Analisis Sistem Informasi, 2012, CV. Andi Offset, Yogyakarta.

IDENTITAS PENULIS

Nama : Lucy Simorangkir, M.Kom
 NIDN/NIK : 1028097801 / 78.11.2.0037
 TTL : Kuala Tungka/8 September 1978
 Gol/Pangkat : III D
 Jab. Fungsional : Lektor
 Alamat Rumah : Jl. Sermak Ishak Ahmad No. 24 RT 07 Mayang Jambi
 Telp. : 081366009242
 Email : lucy.simorangkir@yahoo.co.id

Nama : Akenda Putra
 TTL : Tanjung Raja, 11 November 2001
 Alamat Rumah : Jl. Ks. Tubun Lorong Pokat Telanai Pura Kota Jambi
 Telp. : 082186641357
 Email : akendaputra@gmail.com