



SISTEM PENDUKUNG KEPUTUSAN PENYAKIT DIABETES MEGGUNAKAN METODE KLASIFIKASI NAÏVE BAYES

Aa Kurniawan

Universitas Pamulang

*E-mail : dosen02361@unpam.ac.id

Kata kunci:	Abstrak
Diabetes, naïve bayes, klasifikasi	Pemanfaatan komputer pada saat ini sudah menyentuh berbagai aspek kehidupan manusia. Hampir seluruh bidang pekerjaan mulai masuk ke era komputerisasi, salah satunya termasuk dalam bidang kesehatan. Diabetes adalah penyakit menahun yang terjadi karena adanya gangguan metabolik yang ditandai dengan kadar gula darah yang melebihi batas normal. Penyakit ini tidak hanya menyebabkan kematian premature di seluruh dunia, lebih jauh lagi menjadi penyebab utama kebutaan, penyakit jantung dan ginjal. Berdasarkan Riset Kesehatan Dasar (RISKESDA) yang dilaksanakan pada tahun 2018 menyimpulkan bahwa baru sekitar 25% penderita diabetes yang mengetahui dirinya menderita diabetes. Hal ini, menjadikan status diabetes sebagai <i>silent killer</i> dari penyakit ini masih menghantui dunia. Diagnosis terhadap penyakit ini secara medis masih mengalami kesulitan dan reduksi data. Untuk mendukung hal ini perlu menggunakan teknik <i>data mining</i> berbasis komputer agar dapat menggali informasi berharga dari kumpulan informasi diabetes. Klasifikasi dan klusterisasi merupakan dua bidang garapan yang paling sering ditemui untuk mengekstrak pengetahuan. Dalam peranannya klasifikasi adalah proses untuk menemukan model yang menggambarkan dan membedakan kelas data atau konsep dengan tujuan memprediksi kelas untuk data yang tidak diketahui kelasnya. Model algoritma yang digunakan dalam klasifikasi adalah algoritma naïve bayes. Pada penelitian ini, implementasi kinerja algoritma naïve bayes digunakan untuk melakukan teknik klasifikasi. Penelitian ini mengambil sampel dari UCI data set dengan menggunakan 16 atribut independen dan satu atribut dependen. Sehingga, diharapkan tujuan dari penelitian ini kita dapat mengimplementasikan dan memanfaatkan media internet untuk mendiagnosa dan mengklasifikasi penyakit diabetes secara dini dan mandiri.

Pendahuluan

Pada saat ini penggunaan komputer sudah mempengaruhi hampir di setiap lapisan aspek kehidupan, hampir seluruh bidang pekerjaan mulai masuk ke era komputerisasi dimana komputer digunakan untuk membantu seseorang dalam pekerjaannya, bahkan termasuk dalam bidang kesehatan. Oleh karena itu, perkembangan teknologi *Artificial Intelligence* atau kecerdasan buatan menjadi bagian penting dalam mengembangkan dan membangun aplikasi dalam dunia kesehatan. Sistem pakar merupakan aplikasi berbasis

komputer yang pemanfaatannya digunakan sebagai solusi pemecahan masalah sebagaimana yang dipikirkan oleh pakar (Pratiwi, 2019).

Diabetes merupakan salah satu penyakit kronis menahun yang sering terjadi sebagai akibat adanya gangguan dalam metabolisme tubuh yang sering dijumpai dengan tingkat kandungan gula darah yang berlebih. Hal ini selain menimbulkan kematian premature tidak hanya di Indonesia maupun dunia, lebih jauh lagi sering menjadi pemicu awal kebutaan, penyakit ginjal dan jantung.

Pada 14 Mei 2020, (IDF) dalam halaman resminya menginformasikan bahwa kurang lebih 463 juta jiwa di seluruh dunia mengidap diabetes dan memiliki nilai prevalensi global hingga 9,3%. Namun, hal yang paling memprihatinkan adalah sekitar 50,1% pengidap diabetes tidak terdiagnosis. Hal ini menjadikan penyakit ini dikenal sebagai *silent killer* dan terus menghantui dunia.

Prediksi terhadap penyakit diabetes dapat dihasilkan dengan cara mengumpulkan informasi penderita diabetes yang disimpan dalam basis data, setelah itu datanya diolah dan hasilnya akan membentuk suatu pola tertentu, dimana hasil akhirnya bisa digunakan dalam mendiagnosa penyakit diabetes (Arsi & Somantri, 2018).

Klasifikasi terhadap penyakit diabete secara medis masih menghadapi kesulitan serta sering mengalami reduksi data. Banyak data medis yang dimana fitur yang digunakan tidak relevan, dan mengalami *redundant* sehingga dapat mempengaruhi terhadap hasil kualitas dari diagnosa penyakit (N. Chu, J. Li, & Y. Zhou, 2010). Untuk mendukung hal ini dibutuhkan sebuah metode teknik data mining berbasis computer sehingga dapat menggali informasi berharga dari kumpulan data penderita diabetes (Lesmana, 2012).

Metode klasifikasi merupakan salah satu teknik metode yang digunakan oleh beberapa peneliti sebelumnya untuk dapat memberikan klasifikasi diagnosis penyakit diabetes, seperti penelitian yang berkaitan dengan teknik optimasi Naïve Bayes menggunakan teknik particle swarm (Susilawati & Riana, 2021). Komparasi algoritma Naïve Bayes dengan Decision Tree C4.5 (Luki, Rosyid & Mohammad, 2019), serta implementasi system pakar untuk mendiagnosa penyakit diabetes menggunakan teknik metode Naïve Bayes dan Certainty Factor (Ilham, Alamsyah & Trisnawan, 2021) (Insani, Alamsyah, & Putra, 2021).

Berdasarkan beberapa penelitian di atas dapat kita simpulkan bahwa Naïve Bayes cukup populer dan mampu menghasilkan klasifikasi dengan baik.

Metode

Dalam sebuah analisa sistem berjalan, kita sudah mengetahui bahwa dalam proses untuk mendiagnosa sebuah penyakit diabetes memerlukan seorang pakar. Sehingga sistem usulan haruslah dianalisa supaya dapat dirancang dan dibuat dengan baik untuk menghasilkan suatu sistem yang mampu bekerja layaknya seorang pakar yang dapat melakukan klasifikasi dengan baik.

Di lain sisi, agar perangkat lunak yang akan dibangun dapat *user friendly* dan perangkat kerasnya dapat mendukung secara maksimal terhadap kinerja sistem pakar. Maka dibutuhkan juga tahap analisa kebutuhan dan analisa pengetahuan.

Pada penelitian ini terdapat 16 atribut independen serta 1 atribut dependen, seperti yang disajikan dalam table berikut:

Tabel 2. 1 Variabel Diabetes

No.	Variabel	Kriteria Diabetes	Value
1.	Umur	Usia pasien	1. 35-45, 2. 45-55, 3. 55-65
2.	Jenis kelamin	Pria/wanita	1. Male 2. Female
3.	<i>Polyura</i>	Mengeluarkan urin lebih dari 3-5 liter perhari	1. Yes 2. No
4.	<i>polydipsia</i>	Sering merasa haus	1. Yes 2. No

5.	<i>Sudden Weight Loss</i>	Kondisi turunnya berat badan dalam waktu yang cepat	1. Yes 2. No
6.	<i>Weakness</i>	Kondisi badan cepat lelah atau lemah	1. Yes 2. No
7.	<i>Polyphagia</i>	Rasa lapar yang berlebihan atau intens	1. Yes 2. No
8.	<i>Genital Thrush</i>	Infeksi di area kelamin	1. Yes 2. No
9.	<i>Visual Blurring</i>	Kemampuan mata yang mengabur	1. Yes 2. No
10.	<i>Itching</i>	Kondisi kulit gampang gatal	1. Yes 2. No
11.	<i>Irritability</i>	Gampang emosi atau frustrasi	1. Yes 2. No
12.	<i>delayed healing</i>	Kondisi luka yang sulit sembuh	1. Yes 2. No
13.	<i>partial paresis</i>	Mengalami kelumpuhan parsial	1. Yes 2. No
14.	<i>muscle stiffness</i>	Sering mengalami kekakuan otot	1. Yes 2. No
15.	<i>Alopecia</i>	Mengalami kerontokan rambut	1. Yes 2. No
16.	<i>Obesity</i>	Mengalami berat badan berlebih	1. Yes 2. No

Sumber :UCI Data set

Hasil dan Pembahasan

1. Perhitungan Metode

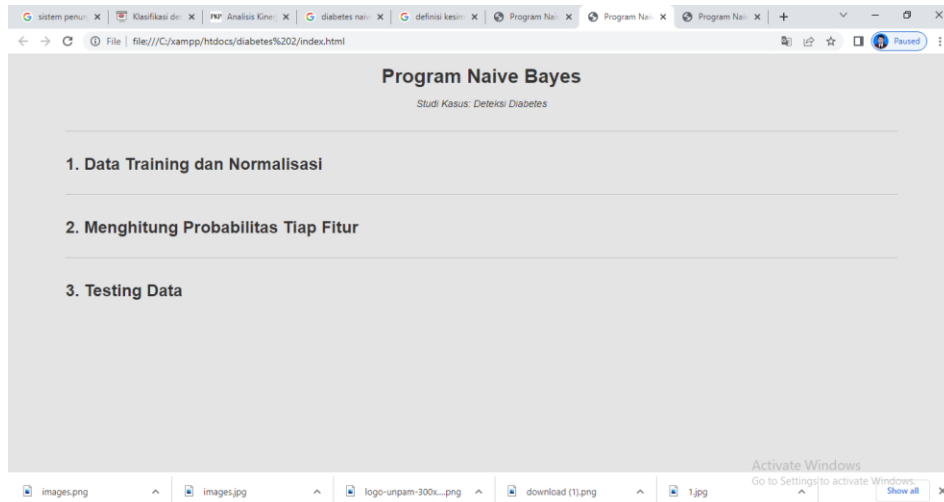
Dalam melakukan klasifikasi, naïve bayes bersandar pada sebuah nilai probabilitas dari setiap variabel yang ada. Selanjutnya untuk mengantisipasi kelemahan dari naïve bayes, yaitu jika ada fitur data yang bernilai 0(nol) maka digunakan juga smoothing terhadap probabilitas dengan menggunakan metode Laplacian.

2. Implementasi

Implementasi dapat didefinisikan sebagai keadaan dimana sistem siap untuk digunakan pada keadaan yang sebenarnya, sehingga akan diketahui apakah sistem yang dibuat telah berfungsi sesuai dengan tujuan yang diinginkan.

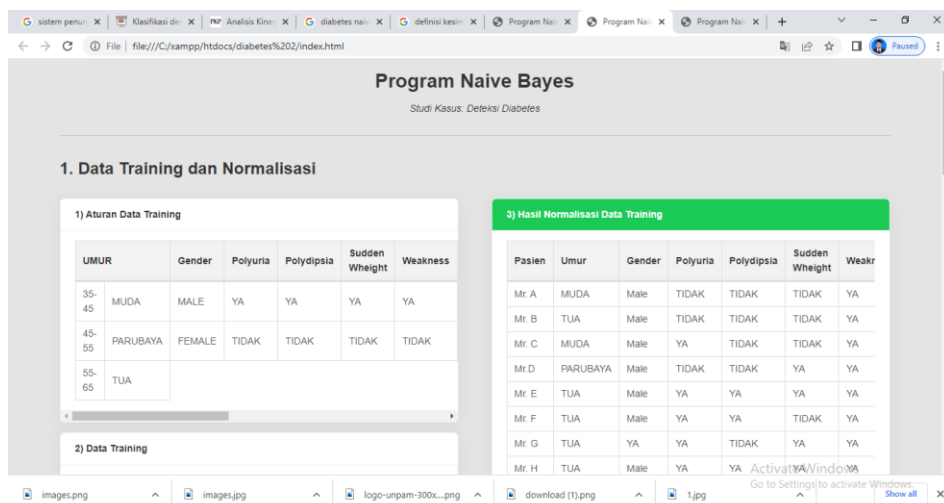
Berikut implementasi antar muka (user interface) dari sistem yang aplikasi yang dibangun

a. Halaman Utama



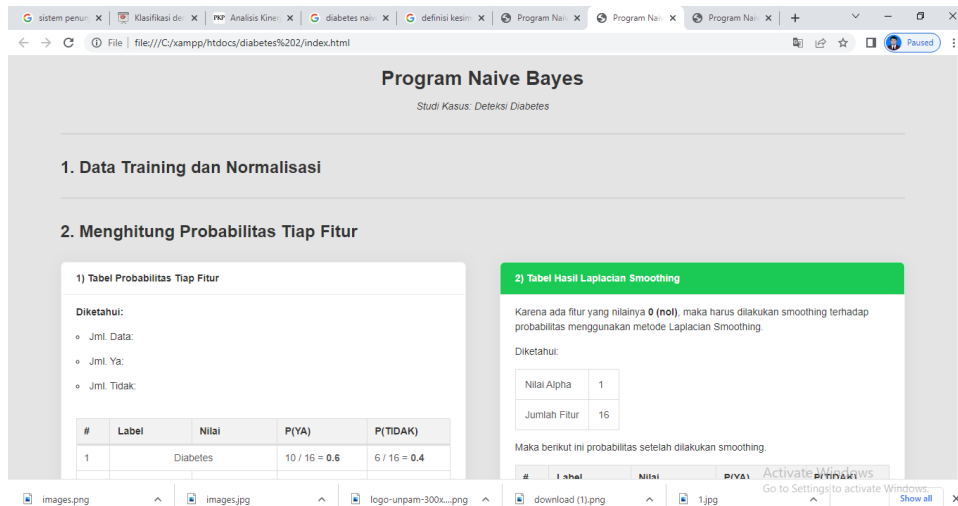
gambar 3. 1 Halaman Utama

b. Halaman Data Training



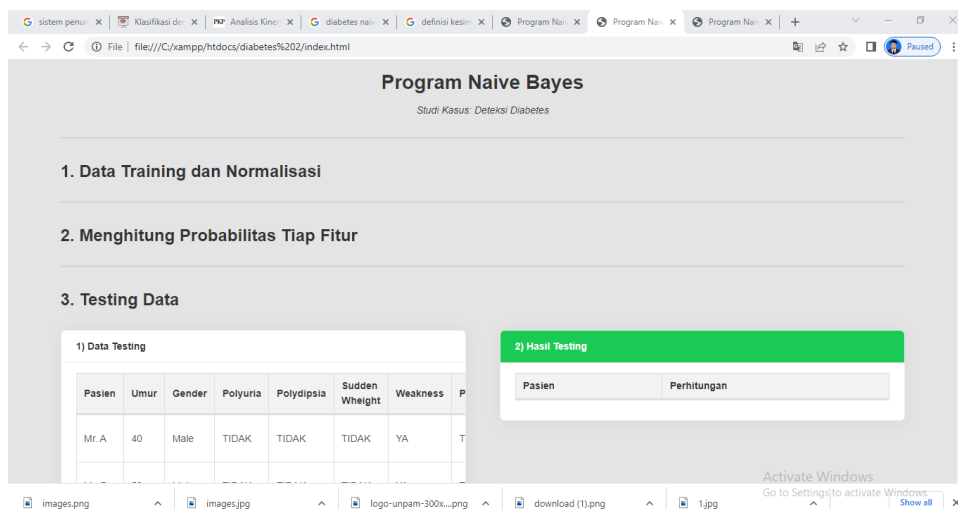
gambar 3. 2 Halaman Data Training

c. Halaman Probabilitas



gambar 3. 3 Halaman Probabilitas

d. Halaman Testing Klasifikasi



gambar 3. 4 Halaman Testing Klasifikasi

3. Pengujian

Sebelum aplikasi diterapkan, maka program harus terbebas dahulu dari kesalahan-kesalahan, oleh sebab itu program harus diuji untuk menemukan kesalahan-kesalahan yang mungkin dapat terjadi. Pengujian ini menggunakan pengujian *blackbox*. Berikut disajikan table rencana pengujian system beserta hasilnya:

Tabel 3. 1 Rencana Pengujian Sistem

No	Item Uji	Detail Pengujian	Jenis Pengujian
1	Halaman Utama	Menampilkan beranda	<i>Blackbox</i>
2	Halaman Data <i>Training</i>	Menampilkan aturan data <i>training</i> , menampilkan <i>master data training</i> dan tambah data <i>training</i>	<i>Blackbox</i>
3	Halaman Probabilitas	Menampilkan hasil probabilitas dan normalisasi berdasarkan data master	<i>Blackbox</i>
4	Halaman Data Testing	Menampilkan hasil data <i>testing</i> sebelumnya dan menambahkan data uji <i>testing</i>	<i>Blackbox</i>

Tabel 3. 2 Hasil Pengujian Sistem

Input / Event	Proses	Output	Hasil
Load Halaman Utama	Load halaman pertama kali saat <i>web</i> dibuka dan menampilkan informasi tentang aturan data <i>training</i>	Tidak ada	Sesuai
Tambah Data Training	Melakukan proses penambahan data training di <i>data master</i>	Data Penyakit Disimpan	Sesuai
Hapus Data Training	Melakukan proses hapus data training di <i>data master</i>	Data Penyakit Dihapus	Sesuai
Load Halaman Normalisasi	Load halaman table normalisasi	Tampil table normalisasi Data Training	Sesuai

Load Halaman Probabilitas	Load halaman probabilitas dan informasi dalam bentuk table tingkat probabilitas tiap variabel dan <i>table Laplacian</i>	Tampil informasi probabilitas beserta table dan Table Laplacian	Sesuai
Load Halaman Data Testing/ Klasifikasi	Load halaman table dari hasil uji klasifikasi sebelumnya.	Tampil table halaman uji data testing sebelumnya	Sesuai
Tambah Data Testing	Menambahkan data uji ke dalam system untuk diklasifikasikan menggunakan metode naïve bayes	Data uji testing diklasifikasi dan disimpan	sesuai
Hapus Data Testing	Menghapus data uji testing yang sudah diinput sebelumnya	Data uji di hapus	sesuai

Berdasarkan pembahasan metode, implementasi serta pengujian yang ada, dapat kita nyatakan bahwa naïve bayes dapat dengan baik melakukan klasifikasi terhadap penyakit diabetes. Terkait dengan fitur yang memiliki nilai 0 (nol) pun dapat dilakukan *smoothing* probabilitas dengan metode Laplacian. Hasilnya aplikasi yang dibangun sudah sesuai dengan teori serta tujuan dari penelitian ini.

Lebih lanjut lagi, kesesuaian ini terjadi karena adanya landasan teori yang cukup melimpah dan pelaksanaan penelitian yang dilakukan secara sistematis dan baik.

Kesimpulan

Berdasarkan beberapa pemaparan sebelumnya, dapat kita lihat bahwa tujuan dari penelitian ini sudah tercapai dan sesuai dengan apa yang diharapkan.

Kedepannya, diharapkan ada pengembangan lebih lanjut lagi dari aplikasi ini, tidak hanya dari segi *user interface* tetapi juga fitur yang dibutuhkan serta dapat menambah nilai aplikasi ini, sehingga dapat lebih dapat dirasakan manfaatnya oleh *user*.

Namun ada yang perlu diperhatikan bahwa, Variabel dan data latih yang digunakan bersumber pada UCI data set saja, sehingga perlu diuji juga apakah variabel yang digunakan sudah baik atau masih bisa disempurnakan kembali oleh pengembang selanjutnya.

Daftar Pustaka

- Ardiansyah, M., Sunyoto, A., & Luthfi, E. T. (2021). Analisis Perbandingan Akurasi Algoritma Naïve Bayes Dan C4.5 untuk Klasifikasi Diabetes. *Edumatic*.
- Arsi, P., & Somantri, O. (2018). Deteksi Dini Penyakit Diabetes menggunakan Algoritma Neural Network Berbasis Algoritma Genetika. *J.Inform.J.Pengemb.IT*, 290-294.

- germas, k. (2020). *infodatin*. Dipetik 2021, dari <https://www.kemkes.go.id/downloads/resources/download/pusdatin/infodatin/Infodatin%202020%20Diabetes%20Melitus.pdf>
- Han, J., & Kamber, M. (2006). *Data Mining Concept and Techniques*.
- Huang, f., Zhu, Q., Zhou, J., Tao, J., Zhou, X., Jin, D., et al. (2017). Research on the Parallelization of the DBSCAN Clustering Algorithm for Spatial Data Mining Based on the Spark Platform. *MDPI*.
- idf. (t.thn.). Dipetik 2022, dari International Diabetes Federation: <https://idf.org/>
- IDF. (t.thn.). *IDF*. Diambil kembali dari International Diabetes Federation: <https://idf.org/>
- Insani, M. I., Alamsyah, & Putra, A. T. (2021). Implementasi Sistem Pakar Untuk Mendoagnosa Penyakit Diabetes Menggunakan Metode Naive Bayes dan Certainty Factor. *Scientific Journal of Informatics*.
- Larose, D. (2005). *Disxcovering Knowledge in Data* .
- Lesmana, I. (2012). *Perbandingan kinerja Decsion Tree J48 dan ID3 Dalam Pengklasifikasian Diagnosis Penyakit Diabetes Mellitus*.
- N. Chu, L., J. Li, P., & Y. Zhou. (2010). Rough set based feature selection for improved differentiation of traditional Chinese medical data. *seventh International Conference on Fuzzy Systems and Knowledge Discovery*, (hal. 2667-2672). Ynatai, China.
- Noviandi. (2018). Implementasi Algoritma Decision Tree C4.5 Untuk Prediksi Penyakit Diabetes. *Inohim*.
- P2PTM, D. (t.thn.). *Apa yang dimaksud diabetes melitus?* Jakarta: Kementerian Kesehatan Republik Indonesia.
- Perkeni. (2021). *Pengelolaan dan Pencegahan Diabetes Mellitus Tipe 2 Dewasa di Indonesia*. Dipetik 2022, dari Perkeni: <https://pbperkeni.or.id/wp-content/uploads/2021/11/22-10-21-Website-Pedoman-Pengelolaan-dan-Pencegahan-DMT2-Ebook.pdf>
- Pratiwi, H. (2019). *Sistem pakar*. Kuningan: Goresan Pena.
- Surbakti, J. (2022). mplementasi Metode NaïveBayes Untuk Diagnosa Penyakit Hati. *Jikoms*, 34-40.
- Susilawati, D., & Riana, D. (2021). Optimization the Naive bayes Classifier Method to diagnose Diabetes Mellitus. *IAIC Trans. Sustain.Digit. Innov. ITS*DI, 78-86.
- Vijayan, V., & Anjali, C. (2015). Prediction and diagnosis of diabetes mellitus-A machine learning approach. *IEEE*, 122-127.
- Zhang, H., & Wang, Z. (2011). A Normal Distribution-Based Over-Sampling Approach to Imbalanced Data Classification. 83-96.