

Implementasi Metode Naive Bayes pada Sistem Diagnosis Penyakit Mata (Studi Kasus Poli Mata Rumah Sakit Islam Jemursari Surabaya)

Raihan Febrianto Grahadi^{1)*}, Rokhmatul Insani²⁾, Berlian Rahmy Lidiawaty³⁾

Telkom University, Surabaya Campus^{1),2),3)}

raihanfebriantog@student.telkomuniversity.ac.id^{1)*}, rokhmatul@telkomuniversity.ac.id²⁾,
berlianerel@telkomuniversity.ac.id³⁾

Abstrak

Lamanya waktu tunggu dalam pemeriksaan dokter mata di rumah sakit dapat meningkatkan risiko terpapar infeksi virus lain. Maka dari itu diperlukannya sebuah sistem diagnosis yang cepat, akurat dan efektif untuk mengurangi lamanya waktu tunggu. Oleh karena itu, penelitian ini bertujuan untuk membuat sistem diagnosis penyakit mata berbasis web dengan menggunakan metode klasifikasi Naive Bayes. Sistem dikembangkan dengan mengambil dataset berupa data rekam medik pasien sebanyak 6074 data. Atribut pada dataset yang akan digunakan dalam sistem ini meliputi umur, jenis kelamin, keluhan, dan diagnosis. Sistem ini dibangun dengan metode Naive Bayes sebagai algoritma data mining yang menggunakan bahasa pemrograman Python sebagai tools pengolahan data. HTML, CSS, JS, dan PHP sebagai bahasa pemrograman pembangun website yang digunakan sebagai tempat visualisasi dan inputan data oleh user yang didukung MySQL sebagai database penyimpanan data. Hasil yang didapatkan dari sistem diagnosis penyakit mata berbasis web dengan perbandingan metode Gaussian Naive Bayes dan Bernoulli Naive Bayes ini mampu mendapatkan akurasi sebesar 93.42% dan 90.79%, sehingga menjadi sistem pendukung keputusan yang dapat membantu dokter dalam mengambil keputusan diagnosis penyakit mata pasien. Dengan adanya sistem ini, dapat mempermudah proses diagnosis penyakit mata serta dapat membantu dokter dalam mengambil keputusan yang tepat dalam waktu singkat.

Kata kunci: Klasifikasi Penyakit Mata, Sistem Diagnosis Penyakit Mata, Naive Bayes

Abstract

[Implementation of Naive Bayes Method on Eye Disease Diagnosis System (Ophthalmology Poly Case Study at Jemursari Islamic Hospital Surabaya)] The waiting time during an eye doctor's examination at the hospital can increase the risk of exposure to other virus infections. Therefore, a fast, accurate, and effective diagnostic system is needed to reduce waiting time. This study aims to develop a web-based eye disease diagnosis system using the Naive Bayes classification method. The system was developed using a dataset consisting of 6074 patient medical records. The attributes of the dataset used in this system include age, gender, complaints, and diagnosis. The system was built using the Naive Bayes method as a data-mining algorithm using the Python programming language for data processing. HTML, CSS, JS, and PHP are website builder programming languages used for visualization and data input by users, supported by MySQL as the data storage database. The results obtained from the web-based eye disease diagnosis system with a comparison of Gaussian Naive Bayes and Bernoulli Naive Bayes methods achieved accuracies of 93.42% and 90.79%, respectively, making it a decision support system that can assist doctors in making diagnoses of patients' eye diseases. With this system, the process of diagnosing eye diseases can be simplified and doctors can be assisted in making the right decisions in a short time.

Keywords: Classification of Eye Diseases, System for Diagnosing Eye Diseases, Naive Bayes.

1. PENDAHULUAN

Kurangnya jumlah dokter [1] [2] [3] mata di beberapa daerah, dapat menyebabkan pasien harus menempuh jarak yang cukup jauh [4] ataupun mengantri dengan waktu yang lama untuk mendapatkan pelayanan kesehatan yang sesuai. Berdasarkan Keputusan menteri kesehatan Indonesia

Nomor 129/Menkes/SK/II/2008 [5] [6] waktu tunggu pasien idealnya ≤ 60 menit. Lebih dari 60 menit tingkat kepuasan pasien dapat menurun. Lamanya waktu tunggu dalam pemeriksaan dokter mata di rumah sakit juga dapat meningkatkan risiko terpapar infeksi virus lain [7] Tidak hanya itu, jika lamanya waktu tunggu

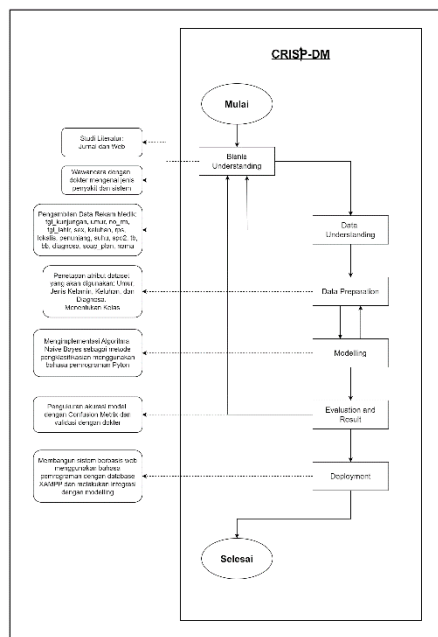
memiliki frekuensi pengulangan yang cukup tinggi, dapat menciptakan rasa frustrasi bagi pasien [8] [9].

Penelitian ini bertujuan untuk menyelesaikan permasalahan dalam mempersingkat waktu dan memudahkan proses pemeriksaan [10] mata dengan mengembangkan sistem diagnosis penyakit mata berbasis web. Dataset berupa data rekam medik pasien yang telah disetujui oleh Rumah Sakit Islam Jemursari Surabaya. Hasil yang diharapkan dari sistem diagnosis penyakit mata berbasis web dengan metode Naïve Bayes adalah mampu mendapatkan akurasi yang terbaik dan optimal dengan membandingkan jenis Naïve Bayes (Bernoulli dan Gaussian), sehingga menjadi sistem pendukung keputusan [11] yang membantu dokter dalam mendiagnosis penyakit mata pasien.

2. BAHAN DAN METODE

Bahan yang digunakan dalam penelitian ini adalah dataset rekam medik yang terdiri dari 6074 data. Dataset ini mencakup beberapa atribut yang penting dalam analisis, seperti umur, jenis kelamin, keluhan, diagnosis, dan tindak lanjut. Atribut-atribut ini akan digunakan sebagai basis untuk pengembangan sistem yang dapat mendukung diagnosa penyakit secara otomatis.

Metode penelitian ini adalah pendekatan kuantitatif data sekunder dengan mengelola data yang telah diberikan pihak Rumah Sakit Islam Jemursari Surabaya, dan melakukan observasi dan wawancara sebagai penguat data. Untuk metode dalam mengelola data, dalam penelitian ini menerapkan metode algoritma Naïve Bayes. Dalam penelitian ini mengikuti alur pedoman metode model CRISP-DM sebagai alur kerja penelitian.



Gambar 1. Gambar Diagram CRISP-DM Penelitian

Berikut ini adalah penjelasan mengenai urutan prosedur penelitian menggunakan CRISP DM [12] [13]:

1. Business Understanding

Business understanding merupakan fase di mana masalah bisnis dijelaskan dengan cara yang singkat dan akurat.

2. Data Understanding

Proses ini melibatkan pemahaman mengenai data yang dimiliki dengan data yang diperlukan.

3. Data Preparation

Data preparation merupakan tahap pengolahan data yang bertujuan untuk menghasilkan model yang berkualitas dan bermanfaat.

4. Modelling

Modeling adalah suatu tahap dimana melakukan sebuah eksekusi knowledge.

5. Evaluation and result

Evaluasi merupakan sebuah untuk memastikan kesesuaian dengan tujuan yang ingin dicapai dalam tahap awal.

6. Deployment

Deployment merupakan sebuah tahap dimana model Naïve Bayes telah di implementasikan dan di integrasikan ke dalam sistem berbasis website.

3. HASIL DAN PEMBAHASAN

Hasil implementasi Metode Naive Bayes pada Sistem Diagnosis Penyakit Mata dengan fokus pada studi kasus di Poli Mata Rumah Sakit Islam Jemursari Surabaya. Analisis mendalam terhadap data yang dikumpulkan akan disajikan, memberikan gambaran komprehensif tentang efektivitas metode tersebut.

A. Business Understanding

Seperti yang telah di paparkan dalam pendahuluan, bahwa kurangnya dokter dan lamanya waktu tunggu dapat menimbulkan dampak yang buruk. Dengan membangun sebuah sistem diagnosis yang menerapkan data mining, dapat mempermudah dan mempercepat proses mendiagnosa pasien.

B. Data Understanding

Data merupakan hasil rekam medik pasien yang berasal dari Rumah Sakit Islam Jemursari Surabaya dengan total 6074 data. Setiap entitas pasien direpresentasikan oleh 16 atribut yaitu `tgl_kunjungan`, `umur`, `no_rm`, `tgl_lahir`, `sex`, `keluhan`, `rps`, `lokalis`, `penunjang`, `suhu`, `spo2`, `tb`, `bb`, `diagnosa`, `soap_plan`, `nama`.

C. Data Preparation

Adapun tahapan dalam penelitian ini sebagai berikut:

- Data Cleaning

Dalam tahap Data Cleaning, proses pertama dilakukan dengan mengidentifikasi dan menghapus data yang kosong untuk meningkatkan kualitas dataset.

Proses kedua menitikberatkan pada penghapusan data yang memiliki lebih dari satu diagnosis guna meningkatkan spesifikasi analisis. Terakhir, proses ketiga fokus pada penghapusan data di dalam kelas yang memiliki sampel kurang dari 20 data, memastikan bahwa setiap kategori memiliki representasi yang memadai. Melalui langkah-langkah ini, data menjadi lebih bersih, lebih spesifik, dan lebih seimbang, membentuk dasar yang kuat untuk analisis data yang akurat [14], [15] dan informatif.

- **Data Transformation**

Dalam tahap transformasi data, fokus utama adalah mengubah data umur dari bentuk angka menjadi kategorikal guna memudahkan analisis [16]. Proses ini memetakan rentang umur ke dalam empat kategori spesifik: 0-19 tahun untuk Kategori 1 (Balita, Anak-anak, dan remaja), 20-44 tahun untuk Kategori 2 (Dewasa), 45-59 tahun untuk Kategori 3 (Paruh Baya), dan 60 tahun ke atas untuk Kategori 4 (Lansia).
- **Penguraian Kolom**

Pada tahap Penguraian Kolom, berfokus pada perubahan kolom keluhan menjadi atribut yang efektif untuk model Naive Bayes dengan mempertimbangkan probabilitas munculnya keluhan tertentu terkait klasifikasi kelas yang diinginkan. Semula data keluhan menjadi satu dalam kolom keluhan kemudian data dipecah menjadi atribut
- **Normalisasi Bahasa**

Normalisasi bahasa daerah atau bahasa non-formal menjadi bahasa Indonesia atau bahasa medis memiliki peran penting dalam memudahkan pemilihan keluhan yang tepat untuk diagnosis medis. Langkah pertama dalam proses ini adalah mengidentifikasi ekspresi atau istilah dalam bahasa non-formal yang sering digunakan oleh pasien. Selanjutnya, dilakukan terjemahan atau normalisasi ke dalam bahasa Indonesia yang umumnya digunakan dalam konteks medis.
- **Dataset Akhir**

Dataset terakhir setelah tahap persiapan data mencakup berbagai atribut yang relevan untuk analisis kesehatan mata. Atribut-atribut ini

mencakup rentang usia, jenis kelamin (sex), dan sejumlah gejala atau kondisi mata yang diobservasi. Dataset terakhir yang digunakan terdiri dari 251 data.

D. Modelling

Penerapan pemodelan menggunakan dua jenis Naïve Bayes, yaitu Gaussian Naive Bayes dan Bernoulli Naive Bayes, dengan menggunakan library Python. Gaussian Naïve Bayes merupakan sebuah algoritma data mining yang bertujuan untuk melakukan klasifikasi [17]. Dalam penelitian ini Gaussian Naïve Bayes menunjukkan hasil yang baik dengan menampilkan hasil akurasi yang tinggi sebesar 93.42105263157895 %, dengan test size = 0.3 dan random state = 60. Bernoulli Naïve Bayes merupakan sebuah algoritma data mining yang bertujuan untuk melakukan klasifikasi dalam bentuk data biner [18] [19]. Dalam penelitian ini Bernoulli Naïve Bayes menunjukkan hasil yang baik dengan menampilkan hasil akurasi yang tinggi sebesar 90.79%, dengan test size = 0.3 dan random state = 6.

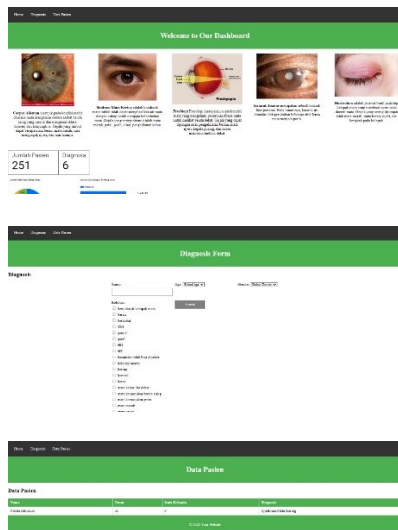
E. Evaluation and result

Hasil dari pengembangan modelling dengan menggunakan dua model yaitu Gaussian Naïve Bayes dan Bernoulli Naïve Bayes menunjukkan bahwa kedua model tersebut memiliki titik perbedaan dalam hal akurasi. Ada beberapa faktor yang dapat mempengaruhi perbedaan tersebut, seperti: pembagian data latih dan coba, random state, dan jenis data. Adapun uji dengan test size yang sama, pada test size 0.2 dan 0.5 Gaussian Naïve Bayes dan Bernoulli Naïve Bayes memiliki akurasi yang sama namun, random state yang digunakan berbeda, menunjukkan bahwa Bernoulli memerlukan pengacakan angka yang lebih tinggi dibanding Gaussian. Apabila di uji satu banding satu dengan test size yang sama dan random state yang sama hingga 200 kali menunjukkan akurasi Bernoulli bisa lebih tinggi, namun tidak pernah melebihi akurasi tertinggi yang dicapai oleh Gaussian Naïve Bayes. Meskipun dalam Bernoulli memiliki akurasi sebesar 100%, namun pada nyatanya ketika di uji dengan data baru menunjukkan terdapat kesalahan dalam diagnosis dan dapat memungkinkan terjadinya overfitting yaitu dimana data yang telah diolah mendapatkan akurasi yang baik namun mengalami ketidakcocokan dengan data True dan Prediksi [20].

F. Deployment

Dalam pembangunan website ini digunakan bahasa pemrograman HTML, CSS, PHP, Flask Python. Dalam website ini terbagi menjadi tiga page yaitu home, Diagnosis, dan

Data Pasien. Dalam home berisikan sekilas konten mengenai penyakit mata, dimana ada lima penyakit mata, serta terdapat dashboard visualisasi mengenai dataset akhir yang digunakan dalam modelling. Untuk halaman diagnosis berisikan form yang diisikan data pasien mulai dari nama, umur, gender dan keluhan-keluhan. Dalam halaman ini akan disambungkan dengan modelling menggunakan flask sebagai penghubung antara kodingan website dengan kodingan modelling yang menggunakan bahasa pemrograman yang berbeda. Hal ini dilakukan supaya data baru pasien dapat diolah untuk mengetahui diagnosa penyakit yang diderita oleh pasien. Untuk halaman data pasien merupakan halaman yang berisikan visualisasi dari data pasien yang telah tersimpan dalam database.



Gambar 2. Gambar Implementasi Website

4. KESIMPULAN

Berdasarkan penelitian ini, dapat disimpulkan menjadi beberapa point sebagai berikut :

1. Dalam penelitian ini model Gaussian dipilih menjadi model dikarenakan memiliki akurasi sebesar 93.4% dengan Test size sebesar 0.3 dan Random State sebesar 60.
2. Model Bernoulli Naïve Bayes memiliki akurasi yang bagus. Saat diuji 1 banding 1 dengan Gaussian Naive Bayes, akurasi Bernoulli bisa lebih tinggi, namun tidak pernah melebihi akurasi tertinggi yang dicapai oleh Gaussian Naïve Bayes. Adapun akurasi tertinggi Gaussian Naïve Bayes dari test size=0.1 – 0.6 sebagai berikut, 96.15, 92.16, 93.42, 91.09, 88.10, dan 86.09 . Sedangkan Bernoulli Naïve Bayes sebagai berikut 100, 92.16, 90.79, 89.11, 88.10, dan 84.11 .

3. Tiga faktor yang mempengaruhi perbandingan antara kedua model ini adalah Test Size, Random State, dan Jenis Data
4. Sistem diagnosis dengan metode Naïve Bayes telah berhasil di visualisasikan melalui website.

5. DAFTAR PUSTAKA

- [1] Y. Yuliyana and A. S. R. M. Sinaga, “Sistem Pakar Diagnosa Penyakit Gigi Menggunakan Metode Naive Bayes,” *Fountain of Informatics Journal*, vol. 4, no. 1, p. 19, May 2019, doi: 10.21111/fij.v4i1.3019.
- [2] B. N. Riyanto and O. Suria, “Sistem Pakar Diagnosa Penyakit Pencernaan Menggunakan Metode Teorema Bayes 7,” *Jurnal Multimedia & Artificial Intelligence*, vol. 2, no. 1, pp. 7–12, 2018.
- [3] E. Niati and S. Sitohang, “Sistem Pakar Diagnosa penyakit Mata Glaukoma Dengan Metode Teorema Bayes,” *Jurnal Comaise*, vol. 7, no. 5, pp. 117–124, 2022, [Online]. Available: <http://ejournal.upbatam.ac.id/index.php/comasiejournal>
- [4] N. Ratnaningsih and S. F. Boesoerie, *Korelasi antara Jarak Tempuh ke Rumah Sakit dengan Proporsi Gangguan Penglihatan dalam Survei Rapid Assessment of Avoidable Blindness di Jawa Barat*. 2017. [Online]. Available: perpustakaanrsmrsmcicendo.com
- [5] Kementerian Kesehatan, “Menteri Kesehatan Republik Indonesia Nomor : 129/Menkes/SK/II/2008 Tentang Standar pelayanan Minimal Rumah Sakit,” 2008.
- [6] R. J. R. Walakandou, G. A. E. Ratag, and G. E. C. Korompis, “Faktor-Faktor yang Berpengaruh pada Waktu Tunggu Pasien dalam Masa Pandemi Covid 19 di Unit Rawat Jalan Rumah Sakit,” *Journal of Public Health and Community Medicine*, vol. 2, no. 3, pp. 1–12, 2021.
- [7] Meilinda and F. Gustini, “Analisis Fasilitas Sanitasi Dalam Mencegah Penularan Covid-19 di Rumah sakit,” *Jurnal Education and development*, vol. 9, no. 4, pp. 81–85, 2021.
- [8] R. S. Kuryadinata, M. Rohmah, and Z. M. Septimar, “Hubungan Waktu Tanggap Pelayanan Kegawatdaruratan Dengan Tingkat Kepuasan Keluarga Pasien di Unit Gawat Darurat,” *Jurnal Kesehatan Panrita Husada*, vol. 7, no. 1, pp. 16–27, Mar. 2022, doi: 10.37362/jkph.v7i1.708.
- [9] A. Pandit, L. Varma, and A. P., “Impact of OPD Waiting Time on Patient Satisfaction,” *International Education & Research Journal (IERJ)*, vol. 2, no. 8, pp. 86–90, 2016, [Online]. Available:

- <https://www.researchgate.net/publication/324594239>
- [10] P. W. Ramadhan, *Sistem Pakar Diagnosa Penyakit Umum Menggunakan Metode Forward Chaining Berbasis Web*. 2019.
- [11] M. R. Handoko, "Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web," *Jurnal Teknologi dan Sistem Informasi (JTISI)*, vol. 2, no. 1, pp. 50–58, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTISI>
- [12] Y. A. Singgalen, "Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM," *Journal of Information System Research (JOSH)*, vol. 4, no. 2, pp. 614–625, Jan. 2023, doi: 10.47065/josh.v4i2.2999.
- [13] N. Alfiah, "Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan Menggunakan Metode Naive Bayes," *Jurnal Teknologi Informasi*, vol. 16, no. 1, pp. 32–40, 2021.
- [14] Scikit-learn, "Naive Bayes," [scikit-learn.org](https://scikit-learn.org/stable/modules/naive_bayes.html). Accessed: Dec. 03, 2023. [Online]. Available: https://scikit-learn.org/stable/modules/naive_bayes.html
- [15] scikit-learn, "sklearn.naive_bayes.BernoulliNB," [scikit-learn.org](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.BernoulliNB.html). Accessed: Dec. 03, 2023. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.BernoulliNB.html
- [16] G. Gumelar and H. Al-Fatta, "Kombinasi Algoritma Klasifikasi Dengan Algoritma Oversampling Untuk Menangani Ketidakseimbangan Kelas Pada Level Data," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 10, no. 2, pp. 29–39, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>
- [17] M. V. Anand, B. Kiranbala, S. R. Srividhya, K. C., M. Younus, and M. H. Rahman, "Gaussian Naïve Bayes Algorithm: A Reliable Technique Involved in the Assortment of the Segregation in Cancer," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/2436946.
- [18] E. Hasibuan and E. A. Heriyanto, "Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping di Google Play Store Menggunakan Naive Bayes Classifier," *Jurnal Teknik dan Science (JTS)*, vol. 1, no. 3, pp. 13–24, 2022.
- [19] O. V. Yamin, A. Tenriawaru, O. La Saidi, and A. G. Rahman, "Penerapan Naïve Bayes Classifier dengan Algoritma Nazief dan Adriani Untuk Deteksi Hoaks," *Seminar Nasional Pemanfaatan Sains dan Teknologi Informasi*, vol. 1, no. 1, pp. 335–344, 2023.
- [20] P. A. Nugroho, I. Fenriana, and R. Arijanto, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Pada Ekspresi Manusia," *JURNAL ALGOR*, vol. 2, no. 1, 2020, [Online]. Available: <https://jurnal.buddhidharma.ac.id/index.php/algor/index>