

Sistem Deteksi Dini Pneumonia Balita Berdasarkan Rekam Medis Menggunakan Algoritma C4.5

Tashya Eka Melinda*, Mochammad Choirur Roziqin, Veronika Vestine, Muhammad Ifantara Putra

Program Studi Manajemen Informasi Kesehatan, Jurusan Kesehatan, Politeknik Negeri Jember
tashyaorsi06@gmail.com, irul@polije.ac.id, veronikavestine@polije.ac.id, m.ifantar_p@polije.ac.id

Keywords:

C4.5,
Pneumonia,
Public Health Center,
Medical Record

ABSTRACT

The detection of pneumonia cases in children under five at Jabung Public Health Center has not reached the targeted rate. From 2019 to 2022, the number of identified cases remained below the expected target of 4.45%. This study aimed to design and develop an early detection system for childhood pneumonia based on medical records using the C4.5 algorithm. The research applied the waterfall development method and utilised data collection techniques including interviews and document analysis. The subjects were program officers for childhood pneumonia and medical record staff, while the objects were medical records of children diagnosed with pneumonia and acute respiratory infections (ARI). System development involved several stages, starting with data preprocessing, including data cleaning, selection, reduction, and transformation. Data mining was conducted using the C4.5 algorithm with the help of RapidMiner software. The result was an early detection system tailored to the needs of Jabung Public Health Center. The system achieved an accuracy rate of 97.50% based on the confusion matrix. This system was expected to assist health workers in identifying pneumonia cases in children more effectively, thereby improving disease monitoring and early treatment efforts at the community healthcare level.

Kata Kunci

C4.5,
Pneumonia,
Puskesmas,
Rekam Medis

ABSTRAK

Penemuan kasus pneumonia pada balita di wilayah kerja Puskesmas Jabung belum mencapai target yang telah ditetapkan. Jumlah kasus pneumonia yang ditemukan selama tahun 2019 hingga 2022 selalu berada di bawah target, yaitu sebesar 4,45%. Penelitian ini bertujuan untuk merancang dan membangun Sistem Deteksi Dini Pneumonia Balita Berdasarkan Rekam Medis Menggunakan Algoritma C4.5 di Puskesmas Jabung. Penelitian ini menggunakan metode waterfall dan algoritma C4.5, dengan teknik pengumpulan data melalui wawancara dan studi dokumentasi. Subjek penelitian adalah penanggung jawab program pneumonia balita dan petugas rekam medis, sedangkan objek penelitian berupa data rekam medis balita dengan diagnosis pneumonia dan ISPA. Hasil dari penelitian ini adalah sistem deteksi dini pneumonia balita berbasis rekam medis menggunakan algoritma C4.5 yang dirancang khusus untuk digunakan di Puskesmas Jabung. Pembuatan sistem melalui beberapa tahapan, dimulai dari analisis data yang mencakup proses preprocessing seperti data cleaning, data selection, data reduction, dan data transformation. Selanjutnya dilakukan proses data mining menggunakan algoritma C4.5 dengan bantuan perangkat lunak RapidMiner, yang menghasilkan tingkat akurasi berdasarkan confusion matrix sebesar 97,50%.

Korespondensi Penulis:

Tashya Eka Mleinda,
Politeknik Negeri Jember,
Jl. Mastrip PO BOX 164 Jember
Telepon: +6285749581062
Email: tashyaorsi06@gmail.com

Submitted : 24-08-2024; Accepted : 25-06-2025;
Published : 29-06-2025

Copyright (c) 2025 The Author (s)

This article is distributed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0)

1. PENDAHULUAN

Infeksi Saluran Pernapasan Akut (ISPA) merupakan infeksi saluran pernapasan yang menyerang sistem pernapasan salah satunya sistem pernapasan pada bagian bawah yang menurunkan usia harapan hidup penderitanya sebesar 2,09 pada tahun 2019 [1]. Salah satu infeksi yang menyerang sistem pernapasan bagian bawah adalah Pneumonia. Pneumonia didefinisikan sebagai suatu inflamasi yang terjadi pada parenkim paru-paru yang mencakup bronkiolus respiratorius dan alveoli sehingga dapat menimbulkan konsolidasi paru dan gangguan pertukaran gas saat bernapas [2]. Infeksi ini disebabkan oleh beberapa faktor, yakni bakteri, virus, dan jamur [3], [4], [5]. Pneumonia lebih rentan menyerang usia balita. Sehingga balita merupakan usia yang berisiko [6]. Bahkan di dunia, Pneumonia merupakan penyebab kematian utama balita dibandingkan dengan penyakit lainnya, seperti AIDS, malaria, dan campak [7].

Di Indonesia, kasus Pneumonia pada balita lebih tinggi daripada kasus Pneumonia pada kelompok usia lainnya. Dengan tingginya kasus Pneumonia tersebut, pemerintah Indonesia membuat program tatalaksana Pneumonia melalui Manajemen Terpadu Balita Sakit (MTBS) yang dilakukan oleh Puskesmas yang selanjutnya di atur dalam Peraturan Menteri Kesehatan Nomor 25 Tahun 2014 Tentang Upaya Kesehatan Anak [8]. Dengan adanya program ini diyakini dapat menurunkan angka kematian balita dengan kasus Pneumonia sebesar 40%. Penurunan tingkat kematian balita terhadap Pneumonia dapat dicapai ketika kabupaten/kota yang 50% Puskesmasnya melaksanakan deteksi dini Pneumonia melalui MTBS atau program lainnya [9].

Kabupaten Malang merupakan salah satu kabupaten yang ada di Jawa Timur yang keseluruhan Puskesmasnya telah melakukan tatalaksana Pneumonia melalui MTBS. Salah satunya adalah Puskesmas Jabung. Jumlah target penemuan Pneumonia di wilayah kerja Puskesmas Jabung adalah sebesar 4,45% dan target pemenuhan Standar Pelayanan Minimal Kesehatan Kabupaten sebesar 100% dari keseluruhan jumlah balita yang ada di di wilayah kerjanya. Akan tetapi, kenyataan di lapangan pada tahun 2019 – 2022 Puskesmas jabung belum memenuhi target yang telah ditetapkan. Ada beberapa faktor yang mempengaruhi penemuan kasus Pneumonia balita yang di bawah target, yakni Puskesmas Jabung masih melakukan deteksi secara pasif dengan menunggu balita dengan kondisi sakit untuk berobat atau saat pelaksanaan posyandu. Akan tetapi, tidak semua balita datang untuk melakukan Posyandu sehingga deteksi tidak bisa dilakukan secara menyeluruh. Selain itu, kurangnya jumlah tenaga kesehatan yang turun ke masyarakat untuk melakukan deteksi dan sistem pelaporan yang kurang maksimal.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk merancang Sistem Deteksi Dini Pneumonia Balita Berdasarkan Rekam Medis Menggunakan Algoritma C4.5 Di Puskesmas Jabung. Dalam penelitian ini sistem dibangun menggunakan metode *waterfall* dan klasifikasi data menggunakan algoritma C4.5. Metode *waterfall* dipilih karena salah satu metode pengembangan sistem yang sistematis atau dilaksanakan tahap demi tahap untuk meminimalisir kesalahan yang ada [10]. Algoritma C4.5 dipilih karena dapat menghasilkan sebuah pohon keputusan yang mudah dimengerti sebagai *rules* pengembangan sistem, memiliki tingkat akurasi yang mudah diterima, dapat menangani variabel dengan tipe diskrit dan kontinyu, dan dapat memangkas pohon keputusan [11].

Selain itu, algoritma C4.5 dipilih karena memiliki nilai akurasi yang lebih tinggi dibandingkan algoritma KNN dalam penelitian dengan judul komparasi uji performa Algoritma C4.5 dan KNN dalam memprediksi penyakit diabetes [12]. Algoritma C.45 juga memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma naïve bayes dalam penelitian yang berjudul perbandingan klasifikasi metode Naïve Bayes dan metode *Decission Tree* Algoritma J48 pada pasien penderita stroke di RSUD Abdul Wahab Sjahranie Samarinda [13].

2. METODE PENELITIAN

Penelitian ini menggunakan algoritma C4.5 dalam melakukan klasifikasi data. Proses klasifikasi data melalui beberapa tahapan yakni, pengumpulan data, *pre-processing data* yang meliputi, *data cleaning*, *data selection*, *data reduction*, dan *data transformation*. Setelah melalui tahapan *pre-processing data*, selanjutnya masuk ke dalam tahapan *data mining* menggunakan algoritma C4.5 dan *pattern evaluation* menggunakan *confusion matrix* sehingga dihasilkan sebuah *knowledge*. Penelitian ini juga menghasilkan sebuah produk sistem deteksi dini Pneumonia balita. Dalam proses pengembangan sistem digunakan metode *waterfall* [14]. Tahapan metode *waterfall* tersebut terdiri dari analisis kebutuhan, desain, pengkodean, dan pengujian.

1) Analisis kebutuhan

Pengumpulan kebutuhan *user* yang bertujuan untuk mengidentifikasi kebutuhan sistem deteksi dini Pneumonia pada balita. Analisis kebutuhan ini dilakukan dengan wawancara kepada penanggungjawab Pneumonia dan penanggungjawab rekam medis

2) Desain

Proses desain bertujuan untuk merancang atau membuat desain sistem informasi yang sesuai dengan kebutuhan *user*. Proses desain sistem deteksi dini Pneumonia pada balita akan dibuat menggunakan *flowchart*, *Data Flow Diagram* (DFD), *Entity Relationship Diagram* (ERD), dan desain *interface* sistem informasi

3) Pembuatan kode program

Tahap ini peneliti akan merepresentasikan rancangan yang telah dibuat dan melakukan pengkodean ke dalam bahasan pemrograman PHP dengan menggunakan *framework CodeIgniter 4* sebagai struktur *coding* dan *Bootstrap* sebagai *front-end-framework* untuk pengembangan antar muka.

4) Pengujian

Pada tahap akan dilakukan pengujian untuk memastikan keluaran dari sistem sesuai dengan yang diinginkan.

Penelitian ini dilakukan di Puskesmas Jabung dengan subjek penelitian dalam penelitian ini adalah penanggungjawab Pneumonia balita dan penanggungjawab rekam medis. Pengumpulan data melalui dokumentasi dan wawancara. Objek penelitiannya adalah rekam medis balita dengan diagnosis ISPA dan Pneumonia (Pneumonia dan Pneumonia Berat). Jumlah keseluruhan rekam medis yang digunakan sebanyak 200 rekam medis. Atribut yang digunakan dalam penelitian ini merupakan gejala dari penyakit ISPA dan Pneumonia yang meliputi, demam, batuk, sukar bernapas, stridor, tidak mau minum atau menyusu, tarikan dinding dada, napas cepat, kejang, letergis, dan cuping hidung. Klasifikasi dalam penelitian ini dibagi menjadi tiga, yakni Bukan Pneumonia, Pneumonia, dan Pneumonia Berat.

Penelitian ini telah lolos uji etik yang diselenggarakan oleh Komisi Etik Politeknik Negeri Jember yang dibuktikan dengan Surat Keterangan Persetujuan Etik Nomor 101/PL17.4/PG/2024.

3. HASIL DAN ANALISIS

3.1 Analisis Data

3.1.1 Pengumpulan Data

Data yang digunakan merupakan data sekunder berupa rekam medis pasien balita dengan diagnosis Pneumonia dan ISPA. Data yang digunakan sebanyak 200 rekam medis. Data yang dikumpulkan berupa nama, usia, gejala, dan klasifikasi penyakit. Gejala yang dikumpulkan berdasarkan pedoman Manajemen terpadu Balita Sakit (MTBS) yang terdiri dari, demam, batuk, sukar bernapas, stridor, tidak mau minum atau menyusu, tarikan dinding dada, napas cepat, kejang, letergis, dan cuping hidung [8]. Berikut adalah contoh data yang dikumpulkan oleh peneliti.

No.	Tanggal	Nama Pasien	No. eRM	NK	No. RK	No. RM Lam	No. Dokum	Diri Kalam	No Telp	Alamat	RT	RW	Pekerjaan	Tanggal Pemeriksaan	Kelurahan	Tempat Lahir
1	05/12/2023 08:00		00172304					Laki-laki	00	pendang	001	000		05/12/2023 10:00	MANJARI	MANJARI
2	05/12/2023 08:10		00100962					Laki-laki			001	000	PELAKSI/MA PAGOWA	05/12/2023 10:00	MANJARI	MANJARI
3	05/12/2023 08:30		001785140			02 36 14		Perempuan			001	000	MENDURUS PUNDA TANGGA	05/12/2023 09:00	MANJARI	MANJARI
4	05/12/2023 10:00		00100971			02 08 05		Laki-laki	01		001	000	KARTIRAN DINDA	05/12/2023 10:00	MANJARI	MANJARI
5	05/12/2023 07:30		00100940			02 02 05		Laki-laki	01		001	000	KARTIRAN DINDA	05/12/2023 08:00	MANJARI	MANJARI
6	05/12/2023 07:30		00100940			02 02 05		Laki-laki	01		001	000	KARTIRAN DINDA	05/12/2023 08:00	MANJARI	MANJARI

Gambar 1. Contoh Data Kolom No-Tempat Lahir

Gambar di atas adalah contoh data yang diperoleh peneliti dari hasil export pada e-pus Puskesmas Jabung. Data yang diperoleh masih terpisah-pisah tiap bulan dalam bentuk excel yang memiliki banyak atribut atau kolom. Kemudian peneliti melakukan rekapitulasi data berdasarkan umur balita dan menyimpannya dalam bentuk excel yang baru. Data yang diambil peneliti berupa nama, usia, gejala, dan diagnosis.

Pada excel yang baru, peneliti membuat atribut atau kolom berupa nomor, nama, usia, gejala yang diamati, yakni Demam (D), Batuk (B), Sukar Napas (SN), Stridor (S), Tidak mau minum atau menyusui (TM), Muntah (M), Tarikan Dinding Dada (TDD), Napas Cepat (NC), Kejang (K), Letergis (L), Cuping Hidung (CH) yang didasarkan pada pedoman Manajemen Terpadu Balita Sakit (MTBS), serta kolom klasifikasi untuk diagnosis pasien yang dibagi menjadi 3, yakni Bukan Pneumonia, Pneumonia, dan Pneumonia Berat. Nantinya gejala yang dialami balita akan dituliskan jawaban Ya atau Tidak berdasarkan gejala yang dialami balita yang tercantum pada kolom keluhan pada excel yang diperoleh dari hasil *export* e-pus. Berikut adalah contoh data yang peneliti gunakan.

3.1.2 Pre-Processing Data

Data Cleaning

Proses membersihkan data dari duplikasi, isian yang kosong, dan data yang tidak valid [7]. Saat melakukan pengumpulan data diketahui bahwa banyak gejala yang tidak tertulis dalam rekam medis. Sehingga banyak dijumpai data kosong. Berikut adalah contoh data kosong dan data setelah dilakukan pengisian.

No	Nama	Usia	D	B	SN	S	TM	M	TDD	NC	K	L	CH	Klasifikasi
1.	Mxxxxxxx	1 Th 1 Bln 28 Hr	Ya	Ya										Bukan Pneumonia
2.	Mxxxxxxx	0 Th 6 Bln 0 Hr	Tidak	Ya										Bukan Pneumonia
3.	Axxxxxxx	1 Th 6 Bln 5 Hr	Ya	Ya										Bukan Pneumonia

Gambar 1. Data Kosong

No	Nama	Usia	D	B	SN	S	TM	M	TDD	NC	K	L	CH	Klasifikasi
1.	Mxxxxxxx	1 Th 1 Bln 28 Hr	Ya	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia
2.	Mxxxxxxx	0 Th 6 Bln 0 Hr	Tidak	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia
3.	Axxxxxxx	1 Th 6 Bln 5 Hr	Ya	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia

Gambar 2. Data Setelah Dilakukan Pengisian

Dari kedua gambar di atas dapat diketahui bahwa penelitian ini melakukan pengisian data kosong dengan mengisi secara manual gejala atau atribut yang kosong dengan jawaban “tidak”. Pengisian jawaban “tidak” tersebut berdasarkan konfirmasi dari penanggungjawab Pneumonia balita bahwasannya gejala yang tidak tertulis dalam rekam medis artinya balita tersebut tidak mengalami gejala tersebut.

Data Selection

Data selection dilakukan dengan menghilangkan atribut yang tidak ikut dalam proses *mining* [7]. Dalam penelitian ini atribut yang dihilangkan berupa nama dan umur karena peneliti tidak melakukan klasifikasi atau *mining* terhadap dua atribut tersebut. Berikut adalah data setelah dilakukan proses *selection*.

D	B	SN	S	TM	M	TDD	NC	K	L	CH	Klasifikasi
Ya	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia
Tidak	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia
Ya	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia

Gambar 3. Data Setelah Dilakukan *Selection*

Dari gambar di atas atribut nama dan umur dihilangkan. Atribut tersebut hanya atribut bantu dalam proses pengumpulan data. Sehingga atribut yang ikut dalam proses *mining* tersisa gejala dan klasifikasi penyakit.

Data Reduction

Data reduction yang dilakukan berupa penghilangan atribut yang tidak memberikan nilai terhadap klasifikasi atau proses *mining* yang dilakukan. *Data reduction* dilakukan dalam penelitian ini dengan menghapus atribut yang memiliki isian yang sama. Berikut adalah data yang telah dilakukan *reduction*.

D	B	SN	S	TM	M	TDD	NC	Klasifikasi
Ya	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia
Tidak	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia
Ya	Ya	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak	Bukan Pneumonia

Gambar 4. Data Setelah Dilakukan *Reduction*

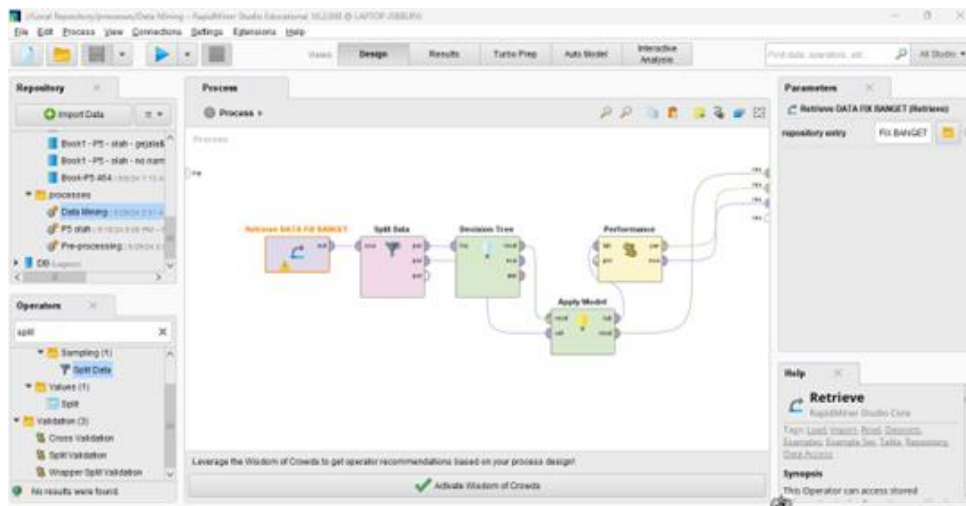
Dari gambar di atas terdapat penghapusan atribut. Terdapat tiga atribut yang memiliki isian jawaban sama, yakni berupa jawaban “tidak”. Atribut tersebut merupakan gejala yang terdiri dari kejang, letergis, dan cuping hidung. Sehingga ketiga atribut tersebut dihapus dan atribut yang dilakukan *mining*, yakni demam, batuk, sukar bernapas, stridor, tidak mau minum atau menyusui, muntah, tarikan dinding dada, dan napas cepat.

Data Transformation

Data transformation merupakan perubahan tipe data yang akan dilakukan *mining* dan format file yang didukung oleh *rapidminer*. Dalam penelitian ini, tipe data yang digunakan berupa *binary* dan nominal, serta format file yang digunakan adalah excel. Hal tersebut sudah sesuai dengan algoritma C4.5 yang dapat menangani tipe data diskrit dan kontinu serta *rapidminer* yang dapat menangani file dalam bentuk excel. Sehingga *data transformation* tidak dilakukan dalam penelitian ini.

3.1.3 Data Mining

Proses *data mining* yang dilakukan menggunakan tools *rapidminer* dengan cara melakukan *drag and drop* file data yang telah dilakukan *preprocessing* dan operator yang akan digunakan seperti gambar di bawah ini.



Gambar 5. Proses Data Mining

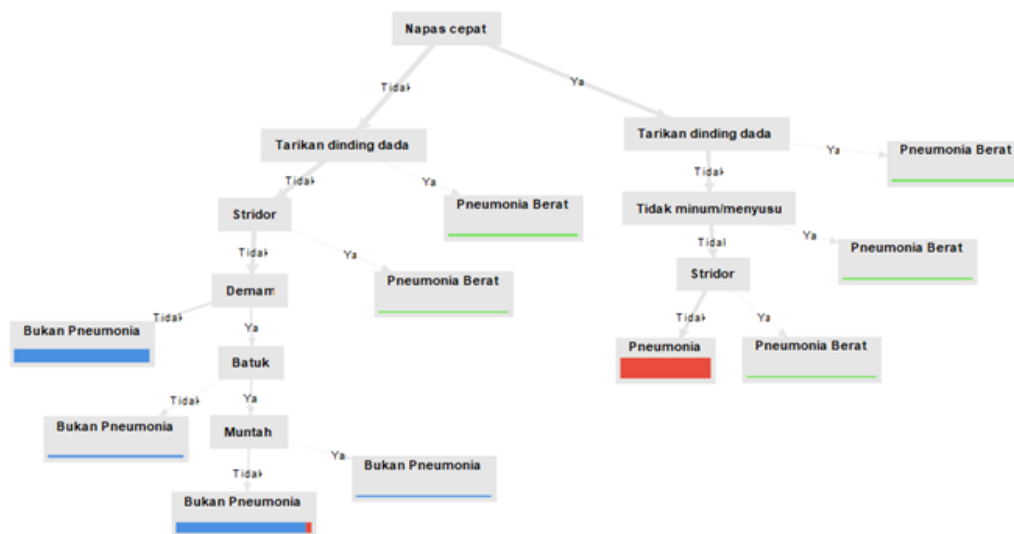
Gambar di atas menunjukkan operator-operator yang digunakan dalam *data mining* meliputi, *split data*, *decision tree*, *apply model*, dan *performance*. Masing-masing operator tersebut memiliki fungsi tersendiri. Operator *split data* nantinya akan membagi komposisi data menjadi *data training* dan *data testing*. Operator *decision tree* merupakan algoritma yang digunakan dalam proses *mining*, yakni algoritma C4.5. Operator *performance* nantinya akan mengevaluasi hasil *mining* yang telah dilakukan.

Dalam operator *split data* yang digunakan untuk membagi menjadi *data training* dan *testing* dengan beberapa komposisi. Selain itu, terdapat 3 teknik *sampling* yang diuji cobakan terhadap data untuk mengetahui hasil akurasi yang paling maksimal. Teknik *sampling* yang pertama, yakni *linier sampling* yang membagi *data training* dan *data testing* berdasarkan urutan datanya. Yang kedua, yakni *shuffled salmpling* dengan membagi *data training* dan *data testing* secara acak. Teknik *sampling* yang terakhir, yakni *stratified sampling* yang membagi *data training* dan *data testing* secara acak, tetapi tetap mempertimbangkan komposisi data yang dimiliki [8].

Komposisi data yang diujicobakan, yakni (50:50), (60:40), (63:37), dan (90:10). Uji coba yang pertama dengan komposisi data 50:50 menghasilkan nilai akurasi 0,00% dengan teknik *linear sampling*, 97,00% dengan teknik *shuffled sampling*. Tetapi pada teknik ini tidak memunculkan atribut stridor pada pohon keputusan yang merupakan gejala ciri khas Pneumonia berat, dan 98,00% dengan teknik *stratified sampling*. Uji coba yang kedua dengan komposisi data 60:40 menghasilkan akurasi sebesar 91,25% dengan teknik *linier sampling*. Akan tetapi, tetapi pada teknik *linier sampling* tidak munculnya atribut tarikan dinding dada yang merupakan gejala ciri khas Pneumonia berat. Selanjutnya diperoleh akurasi sebesar 97,50% dengan teknik *shuffled sampling* dan 97,50% dengan teknik *stratified sampling*.

Uji coba yang ketiga dengan komposisi data 63:37 menghasilkan akurasi sebesar 90,54% dengan teknik *linier sampling*. Akan tetapi, tetapi pada teknik *linier sampling* tidak munculnya atribut tarikan dinding dada dan stridor yang merupakan gejala ciri khas Pneumonia berat. Selanjutnya diperoleh akurasi sebesar 97,50% dengan teknik *shuffled sampling*, tetapi pada teknik ini tidak munculnya gejala stridor dan 97,50% dengan teknik *stratified sampling*. Uji coba yang terakhir dengan komposisi data 90:10 menghasilkan akurasi sebesar 75% dengan teknik *linier sampling*. Akan tetapi, tetapi pada teknik *linier sampling* tidak munculnya atribut tarikan dinding dada yang merupakan gejala ciri khas Pneumonia berat. Selanjutnya diperoleh akurasi sebesar 95% dengan teknik *shuffled sampling* dan 95% dengan teknik *stratified sampling*.

Dari hasil percobaan yang telah dilakukan diketahui bahwa hasil komposisi data terbaik menunjukkan komposisi data 60 % dan 40% dengan teknik *stratified sampling*. Sehingga komposisi data dan teknik *sampling* yang terbaik dipilih untuk proses *data mining*. Selain itu, hasil dari *data mining* juga dihasilkan sebuah pohon keputusan seperti gambar di bawah ini.



Gambar 6. Pohon Keputusan

Pohon keputusan di atas dapat terbentuk karena adanya operator *decision tree* yang digunakan dalam proses *mining*. Dari pohon keputusan di atas dapat dibentuk 10 *rules*. *Rules* yang dihasilkan tersebut dapat digunakan untuk mengembangkan sistem yang akan dilakukan. Berikut adalah *rules* yang dihasilkan.

1. If napas cepat=ya AND tarikan dinding dada=ya THEN class=Pneumonia Berat
2. If napas cepat=ya AND tarikan dinding dada=tidak AND tidak minum/menyusu=ya THEN class=Pneumonia Berat
3. If napas cepat=ya AND tarikan dinding dada=tidak AND tidak minum/menyusu=tidak AND stridor=ya THEN class=Pneumonia Berat
4. If napas cepat=ya AND tarikan dinding dada=tidak AND tidak minum/menyusu=tidak AND stridor=tidak THEN class=Pneumonia
5. If napas cepat=tidak AND tarikan dinding dada=ya THEN class=Pneumonia berat
6. If napas cepat=tidak AND tarikan dinding dada=tidak AND stridor=ya THEN class=Pneumonia berat
7. If napas cepat=tidak AND tarikan dinding dada=tidak AND stridor=tidak AND demam=tidak THEN class=Bukan Pneumonia
8. If napas cepat=tidak AND tarikan dinding dada=tidak AND stridor=tidak AND demam=tidak AND batuk= tidak THEN class=Bukan Pneumonia
9. If napas cepat=tidak AND tarikan dinding dada=tidak AND stridor=tidak AND demam=tidak AND batuk= tidak AND muntah=ya THEN class=Bukan Pneumonia
10. If napas cepat=tidak AND tarikan dinding dada=tidak AND stridor=tidak AND demam=tidak AND batuk=tidak AND muntah=tidak THEN class=Bukan Pneumonia

3.1.4 Patterns Evaluation

Proses evaluasi dibutuhkan untuk identifikasi penilaian *rules* atau pola yang dihasilkan sebelum digunakan sebagai *knowledge*. Evaluasi pola memanfaatkan operator pada *rapidminer*, yakni *confusion matrix*. Berikut adalah hasil dari *confusion matrix*.

accuracy: 97.50%

	true Bukan Pneumonia	true Pneumonia Berat	true Pneumonia	class precision
pred. Bukan Pneumonia	40	0	1	97.56%
pred. Pneumonia Berat	0	5	0	100.00%
pred. Pneumonia	0	1	33	97.06%
class recall	100.00%	83.33%	97.06%	

Gambar 7. Confusion Matrix

Dari hasil perhitungan *confusion matrix* diketahui nilai akurasi yang diperoleh, yakni 97,50%. Hal tersebut juga dapat dibuktikan dengan perhitungan akurasi sebagai berikut.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100\%$$

$$Accuracy = \frac{5 + 33 + 40}{5 + 33 + 1 + 1 + 40} = \frac{78}{80} \times 100\% = 97,50\%$$

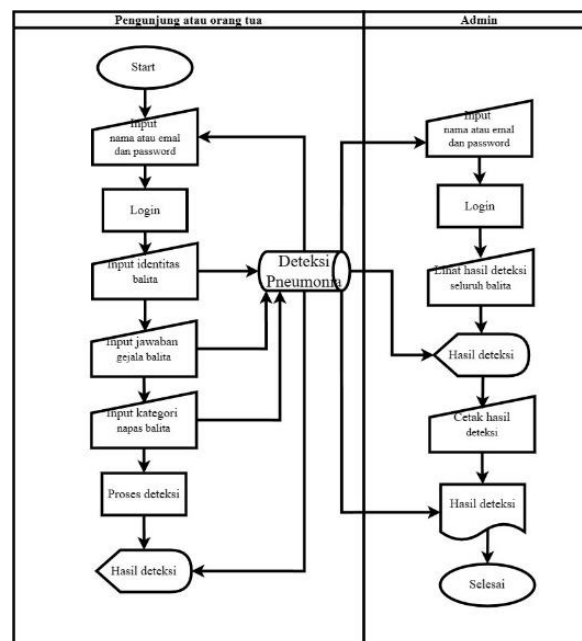
Nilai akurasi menggambarkan tingkat kebenaran model dalam meprediksi secara benar kelas dari Bukan Pnrumonia, Pneumonia, dan Pneumonia Berat. Selain nilai akurasi, nilai presisi yang dihasilkan juga tinggi. Nilai presisi yang baik dapat digunakan sebagai acuan penerapan rules untuk melakukan pengembangan sistem karena nilai presisi menunjukkan tingkat kebenaran prediksi dalam memprediksi kelas positif dari semua data positif. Hal tersebut juga didukung oleh penelitian yang dilakukan yang mengatakan bahwa nilai akurasi, presisi, dan *recall* dapat dikatakan tinggi apabila mendekati 100% [15].

3.2 Analisis Kebutuhan

Analisis kebutuhan dilakukan dengan wawancara terhadap penanggungjawab Pneumonia balita dan penanggungjawab rekam medis. Dari hasil analisis kebutuhan diketahui bahwa nantinya sistem memiliki menu informasi yang berisi tata cara penggunaan sistem, menu deteksi, menu hasil deteksi, dan menu hasil deteksi seluruh balita yang dapat diakses oleh admin. Proses deteksi yang diinginkan dalam sistem ini dimulai dari pengisian identitas balita, pemilihan gejala, dan pemilihan kategori napas yang dialami balita. Identitas yang diinginkan untuk input di dalam sistem berupa NIK, jenis kelamin, tanggal lahir, umur, dan alamat. Output sistem yang diinginkan berupa hasil deteksi yang telah dilakukan.

3.3 System Flowchart

System flowchart merupakan bagan alir yang menunjukkan pekerjaan sistem secara keseluruhan [16]. Berikut adalah *system flowchart* dalam penelitian ini.

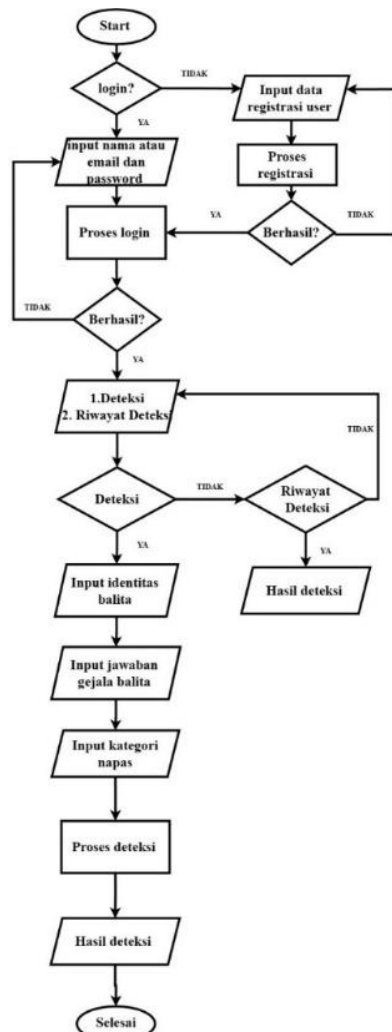


Gambar 8. *System Flowchart*

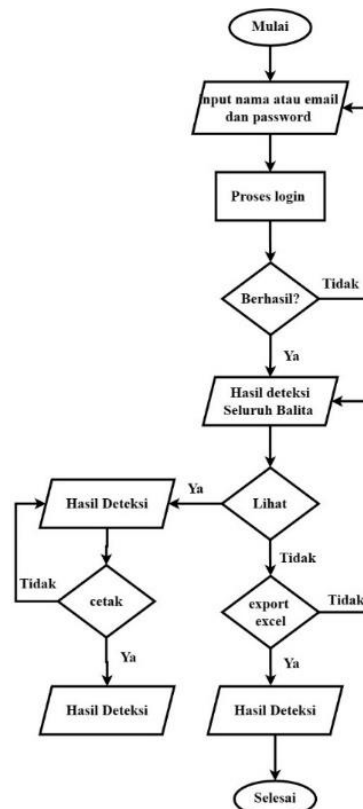
Flowchart di atas menjelaskan urutan pekerjaan sistem secara keseluruhan. Urutan pekerjaan di atas diawali dengan melakukan *login* ke dalam sistem dengan melakukan input *username* dan *password* yang sudah terdaftar. Kemudian untuk pengunjung atau orang tua balita dapat melakukan deteksi dengan melakukan input identitas, gejala yang dialami balita, dan kategori napas yang dialami. Kemudian pengunjung atau orang tua dapat melihat hasil deteksi yang telah dilakukan. Sedangkan untuk admin dapat melihat seluruh hasil deteksi yang telah dilakukan oleh balita dan dapat melakukan cetak hasil deteksi.

3.4 Program Flowchart

Program flowchart menggambarkan langkah-langkah program komputer secara logika. Berikut adalah *program flowchart* dalam penelitian ini.



Gambar 9. Program Flowchart Pengunjung



Gambar 10. Program Flowchart Admin

Flowchart di atas merupakan *program flowchart* yang menggambarkan urutan pekerjaan sistem secara logika. Perbedaannya dengan *flowchart system*, yakni terdapat pemilihan kondisi atau *decision*, yakni dalam *login*, dalam pemilihan menu, dan aksi yang akan dipilih.

3.5 Context Diagram

Context diagram menggambarkan entitas, input, dan output sistem secara keseluruhan [17]. Berikut adalah *context diagram* dari penelitian ini.

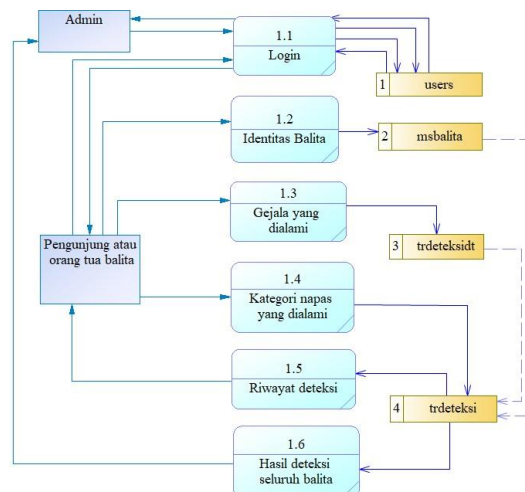


Gambar 11. Context Diagram

Dari *Context Diagram* di atas terdiri dari dua entitas, yakni pengunjung atau orang tua dan admin. Entitas pengunjung atau orang tua dapat melakukan *login*, input identitas balita, input gejala yang dialami, dan input kategori napas yang dialami. Sehingga diperoleh output berupa riwayat deteksi yang menampilkan hasil deteksi yang sudah dilakukan sebelumnya. Sedangkan admin dapat melakukan *login* dan memiliki output hasil deteksi seluruh balita.

3.6 Data Flow Diagram

Data Flow Diagram (DFD) merupakan bagan yang menggambarkan entitas, input, output, serta aliran data yang ada di dalam sebuah sistem [17]. Berikut adalah *Data Flow Diagram* dari sistem deteksi dini Pneumonia balita yang dibuat.

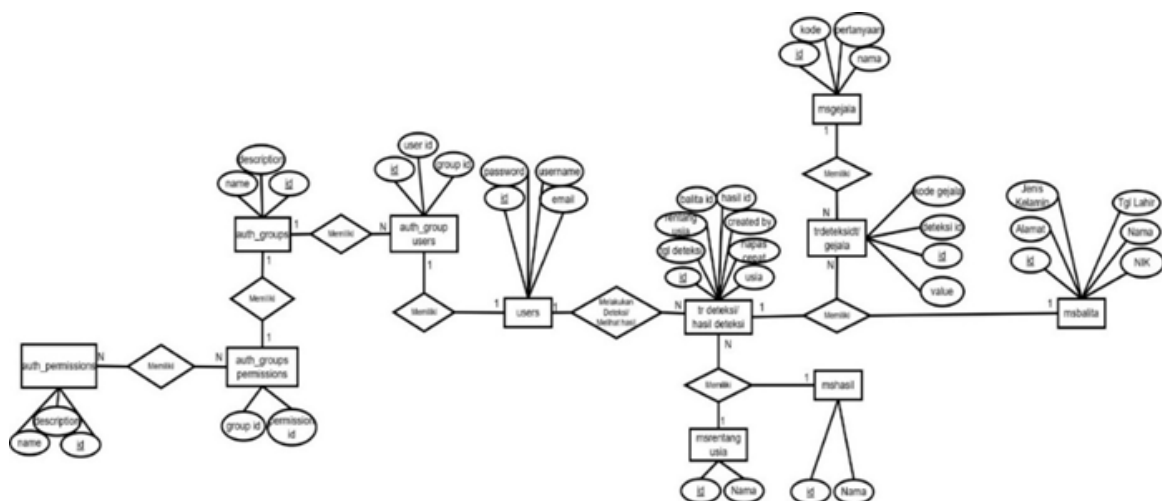


Gambar 12. *Data Flow Diagram*

DFD merupakan pengembangan dari *Context Diagram*. DFD di atas memberikan secara detail proses yang dilakukan dalam sistem. Dalam penelitian ini terdapat enam proses, yakni proses *login*, proses input identitas balita, proses input gejala yang dialami, proses input kategori napas, proses melihat riwayat deteksi, dan proses melihat hasil deteksi seluruh balita. Selain itu, dalam DFD ini juga menggambarkan *database* yang ada di dalam sistem, yakni *database users*, *msbalita*, *trdeteksidt*, dan *terdeteksi*.

3.7 Entity Relationship Diagram

Entity Relationship Diagram (ERD) merupakan bagan yang menggambarkan struktur data dan hubungan antar data yang ada di dalamnya [18]. Berikut adalah ERD yang dibuat oleh peneliti.



Gambar 13. *Entity Relationship Diagram*

ERD di atas memiliki 11 entitas, yakni *users*, *auth group*, *auth permission*, *auth group permissions*, *auth groups users*, *msbalita*, *msgejala*, *mshasil*, *rentang usia*, *terdetksidt*, dan *trdetksi*. Selain itu, ERD memiliki 10 kardinilatas antar entitas yang dimiliki.

3.8 Desain Interface

Desain *Interface* merupakan perancangan antarmuka yang akan dibuat. Desain *interface* yang dibuat sudah disesuaikan dengan kebutuhan pengguna sesuai dengan analisis kebutuhan yang telah dilakukan. Desain *interface* sistem deteksi dini Pneumonia balita dapat diakses melalui link di Berikut: <https://www.figma.com/design/rzyZbyyYN5XvoD65xpyQDC/DESAIN-INTERFACE-SISTEM-DETEKSI-DINI-PNEUMONIA-BALITA?node-id=0-1&t=9CfzY8jEaSwHzlVi-1>.

3.9 Pengkodean

Pengkodean merupakan tahapan transalasi desain ke dalam program dengan penggunaan bahasa pemrograman tertentu [14]. Dalam penelitian ini, pengkodean dilakukan menggunakan bahasa pemrograman PHP. Selain itu, untuk mempermudah dan mempercepat dalam melakukan pengkodean, peneliti menggunakan *framework CodeIgniter 4*, *Bootsrap*, serta *library Myth/Auth*.

3.10 Pengujian

Pengujian yang dilakukan menggunakan *blacbox testing* untuk memastikan fungsional sistem dapat berjalan sesuai dengan perencana [19]. Dari hasil pengujian terhadap menu menu yang dibuat, tombol, input, dan output sistem diperoleh hasil sesuai dengan perencanaan atau berhasil dilakukan. Berikut adalah hasil dari pengujian yang dilakukan.

Tabel 1. Hasil Uji Fungsionalitas

No.	Fungsionalitas	Skenario Uji	Hasilyang diharapkan	Kesimpulan
1.	Melakukan registrasi	<i>Input</i> email, nama, dan <i>password</i>	Tampilan halaman <i>login</i> apabila berhasil dan apabila gagal akan muncul peringatan	Berhasil
2.	Melakukan <i>login</i>	Menginputkan email atau nama dan kata sandi	Tampilan <i>dashboard</i> apabila <i>login</i> berhasil	Berhasil
3.	Pengunjung/orang tua dapat mengakses menu <i>dashboard</i>	Mengkases menu <i>dashboard</i>	Jika berhasil akan muncul halaman <i>dashboard</i>	Berhasil
4.	Menu deteksi	Melakukan deteksi Pneumonia pada menu deteksi	Dapat melakukan deteksi dengan cara melakukan input identitas dan gejala yang dialami balita	Berhasil
5.	Menu riwayat deteksi	Mengakses menu riwayat deteksi	Dapat melihat hasil deteksi yang telah dilakukan sebelumnya	Berhasil

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan terkait Sistem Deteksi Dini Pneumonia Balita Berdasarkan Rekam Medis Menggunakan Algoritma C4.5 Di Puskesmas Jabung dapat disimpulkan bahwa analisis data dilakukan terhadap 200 berkas rekam medis. Analisis data dilakukan melalui beberapa tahapan, yakni *pre-processing*, *data mining*, *patterns evaluation*, dan *knowledge*. Hasil dari evaluasi *data mining* yang telah dilakukan menunjukkan akurasi sebesar 98,91%. Hasil dari analisis data menghasilkan sebuah *rules* dari pohon Keputusan yang dibentuk. *Rules* tersebut digunakan dalam pengembangan sistem menggunakan metode *waterfall* yang terdiri dari analisis kebutuhan, desain sistem yang terdiri dari *flowchart*, DFD, ERD, dan desain *interface*, pengkodean menggunakan bahasa PHP, dan pengujian menggunakan *blackbox testing* dengan kesimpulan berhasil. Puskesmas disarankan untuk menambahkan fitur hapus pada menu hasil deteksi seluruh balita yang dapat diakses oleh admin. Sehingga ketika ada kesalahan input atau duplikasi deteksi yang dilakukan oleh pengunjung atau orang tua, admin dapat menghapusnya dengan mudah. Diharapkan peneliti selanjutnya dapat mengembangkan sistem dengan menu yang lebih kompleks.

REFERENSI

- [1] J. P. Goyal *et al.*, “Risk Factors for the Development of Pneumonia and Severe Pneumonia in Children,” *Indian Pediatr.*, vol. 58, no. November, pp. 1036–1039, 2021.
- [2] R. L. Abdul and S. Herlina, “Asuhan Keperawatan Pada Pasien Dewasa Dengan Pneumonia: Study Kasus,” *Indones. J. Heal. Dev.*, vol. 2, no. 2, pp. 102–107, 2020, doi: <https://doi.org/10.52021/ijhd.v2i2.40>.
- [3] S. N. Budihardjo and I. W. B. Suryawan, “Faktor-faktor resiko kejadian pneumonia pada pasien pneumonia usia 12-59 bulan di RSUD Wangaya,” *Intisari Sains Medis*, vol. 11, no. 1, pp. 398–404, 2020.
- [4] N. F. Muhafidzah *et al.*, “Faktor Faktor Risiko terjadinya Pneumonia pada Stroke Akut di RSUP Dr Hasan Sadikin Bandung,” *J. Neuroanestesi Indones.*, vol. 10, no. 3, pp. 151–161, 2021.
- [5] F. Faisal, Irwandi, R. Aprilia, Suharni, and Efriza, “Tinjauan Literatur: Faktor Risiko dan Epidemiologi Pneumonia pada Balita,” *Sci. J.*, vol. 3, no. 3, pp. 166–173, May 2024, doi: 10.56260/sciena.v3i3.144.
- [6] A. Zolanda, M. Raharjo, and O. Setiani, “Faktor Risiko Kejadian Infeksi Saluran Pernapasan Akut Pada Balita Di Indonesia,” *LINK*, vol. 17, no. 1, pp. 73–80, May 2021, doi: 10.31983/link.v17i1.6828.
- [7] W. Sando, K. Kiswanto, and A. Alamsyah, “Pelaksanaan Program Pengendalian Penyakit Infeksi Saluran Pernafasan Akut (P2 ISPA) di Puskesmas Sungai Pakning Kabupaten Bengkalis,” *J. Kesehat. Komunitas*, vol. 4, no. 3, pp. 102–111, Feb. 2019, doi: 10.25311/keskom.Vol4.Iss3.269.
- [8] Kementerian Kesehatan Republik Indonesia, *Peraturan Menteri Kesehatan Nomor 25 Tahun 2014 tentang Upaya Kesehatan Anak*. Jakarta: Kementerian Kesehatan Republik Indonesia, 2014.
- [9] Dinas Kesehatan Provinsi Jawa Timur, *Profil Kesehatan Provinsi Jawa Timur*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2021.
- [10] A. A. Wahid, “Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi,” *J. Ilmu-ilmu Inform. dan Manaj. STMIK*, vol. 1, no. 1, pp. 1–6, 2020.
- [11] M. A. Muslim, *Data Mining Algoritma C4.5 : Disertai Contoh Kasus dan Penerapannya Dengan Program Komputer*. Semarang: Universitas Negeri Semarang, 2019.
- [12] A. Ardiyansyah, R. Saadah, L. Lisnawanty, and D. Purwaningtias, “Peningkatan Akurasi Metode C4.5 Untuk Memprediksi Kelayakan Kredit Berbasis Stratified Sampling Dan Optimize Selection,” *Kumpul. J. Ilmu Komput.*, vol. 10, no. 2, pp. 239–249, 2023, doi: <https://dx.doi.org/10.20527/klik.v10i2.636>.
- [13] A. Tristiyanto and Sriyanto, “Komparasi Uji Peformas Algoritma C4.5 DAN K-Nearest Neighbor Dalam Memprediksi Penyakit Diabetes,” in *Seminar Nasional Teknologi dan Multidisiplin Ilmu (SEMNASTEKMU)*, Dec. 2022, pp. 379–389. doi: 10.51903/semnastekmu.v2i1.191.
- [14] I. Lishania, R. Goejantoro, and Y. N. Nasution, “Perbandingan Klasifikasi Metode Naive Bayes dan Metode Decision Tree Algoritma (J48) pada Pasien Penderita Penyakit Stroke di RSUD Abdul Wahab Sjahranie Samarinda,” *J. Eksponensial*, vol. 10, no. 2, pp. 135–142, 2019.
- [15] D. Y. Aziz, “Perancangan Dan Pembuatan Sistem Informasi Pendaftaran Pasien Rawat Jalan Secara Online di Puskesmas Kencong Kabupaten Jember,” *Politeknik Negeri Jember*, 2022.
- [16] R. I. Borman and M. Wati, “Penerapan Data Maining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes,” *J. Ilm. Fak. Ilmu Komput.*, vol. 9, no. 1, pp. 25–34, 2020.
- [17] I. Budiman, S. Saori, R. N. Anwar, F. Fitriani, and M. Y. Pangestu, “Analisis Pengendalian Mutu Di Bidang Industri Makanan (Studi Kasus: UMKM Mochi Kaswari Lampion Kota Sukabumi),” *J. Inov. Penelit.*, vol. 1, no. 10, pp. 2185–2190, 2021, doi: <https://doi.org/10.47492/jip.v1i10.419>.
- [18] F. Soufitri, “Perancangan Data Flow Diagram Untuk Informasi Sekolah (Studi Kasus Pada SMP Plus Terpadu),” *Reg. Dev. Ind. Heal. Sci. Technol. Art Life*, vol. 2, no. 1, pp. 240–246, 2019.
- [19] M. S. Novendri, A. Saputra, and C. E. Firman, “Aplikasi Inventaris Barang Pada MTS Nurul Islam Dumai Menggunakan PHP dan MySql,” *Lentera Dumai*, vol. 10, no. 2, 2019.