

Journal of Data Mining and Information Systems

<https://journal.yp3a.org/index.php/JDMIS>

DOI Prefix: 10.54259/jdmis.v3i1

Vol. 3, No. 1, February 2025

e-ISSN: 2986-3473

p-ISSN: 2986-5271



Diterbitkan Oleh:

Yayasan Pendidikan Penelitian Pengabdian Algero

Jl. Glugur Rimbun, Perum. Medan Hills, Cluster Eboni, Blok J No. 3. Indonesia

Website: <https://journal.yp3a.org/index.php/jdmis>

Email: jurnal.jdmis@gmail.com



Journal of Data Mining and Information Systems

<https://journal.yp3a.org/index.php/JDMIS>

DOI Prefix: 10.54259/jdmis.v3i1

Vol. 3, No. 1, January 2025

e-ISSN: 2986-3473

p-ISSN: 2986-5271

Volume 3 Nomor 1 (Februari 2025)

Klasifikasi Kualitas Udara di Jakarta Pada Bulan Agustus 2024 Menggunakan Algoritma C4.5

Ika Juni Nur Ayu, Nayla Rahmania Putri, Ridho Putra Nugraha, Ryan Febriansyah

Analisis Persepsi Publik Terhadap Pilkada Jakarta 2024 dengan Clustering dan Sentimen pada Artikel Berita

Anggi Nur Zavira, Ilma Fathiyarahmani, Khansa Nadhifa, Kinanti Anindia Putri, Widya Amanda Mulyadi, Zalfa Syawaliana

Studi Literatur Perancangan Arsitektur Data dan Aplikasi pada Perusahaan Telekomunikasi

Jason Jason, Jefferson Sutanto, Verrel Angkasa, Vicky Darmana, Ade Maulana

Color Space Influence on Photosynthetic Pigment Measurement Accuracy Using CNN in Color Constancy

Ade May Luky Harefa, Arief Muhazir Insandi

Pengembangan Aplikasi Insentif dan Komisi Salesman sebagai Strategi Peningkatan Kinerja SDM Marketing

Suwandi Suwandi, Muhammad Hatta, Turini Turini, Safitri Akbari, Limbong Aprilina Yanti

Analisis Average Waiting Time Penjadwalan CPU Menggunakan Algoritma Shortest Remaining First dan Algoritma Round Robin

Ronald Belferik, Evander Banjarnahor



Diterbitkan Oleh:

Yayasan Pendidikan Penelitian Pengabdian Algero

Jl. Glugur Rimbun, Perum. Medan Hills, Cluster Eboni, Blok J No. 3. Indonesia

Website: <https://journal.yp3a.org/index.php/jdmis>

Email: jurnal.jdmis@gmail.com

JDIMS (Journal of Data Mining and Information Systems)

Vol. 3 No. 1 (Februari 2025)

EDITOR IN CHIEF

Jefri Junifer Pangaribuan (Scopus ID: 57213221087, Universitas Pelita Harapan, Jakarta, Indonesia)

EDITORIAL BOARD

1. Mirza Ilhami (Scopus ID: 56703365400, Universitas Mikroskil, Medan, Indonesia)
2. Eddy Kurniawan (Scopus ID: 57215113944, Universitas Pesantren Tinggi Darul Ulum, Jombang, Indonesia)
3. Ali Akbar Lubis (Scopus ID: 57214068544, Universitas Negeri Medan, Indonesia)
4. Faisal Nadjar (Scopus ID: 57988165100, Universitas Pelita Harapan, Jakarta, Indonesia)
5. Erlangga (Google Scholar: nyaoWYgAAAAJ, Universitas Bandar Lampung, Indonesia)
6. Jusin (Scopus ID: 57211317974, Universitas Pelita Harapan, Jakarta, Indonesia)
7. Raissa Amanda Putri (Scopus ID: 57207730940, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia)
8. Muhamad Fahmi Nugraha (Google Scholar: UBy3dMwAAAAJ, Universitas Ma'soem, Sumedang, Indonesia)
9. Segar Napitupulu (Google Scholar: 0kVXTn4AAAAJ, Universitas Mikroskil, Medan, Indonesia)

REVIEWER BOARD

1. Prof. Dr. Dahlan Abdullah, ST, M.Kom, IPU, ASEAN Eng. (Scopus ID: 57205132023, Universitas Malikussaleh, Aceh, Indonesia)
2. Dr. Ronsen Purba, M.Sc. (Scopus ID: 57215525192, Universitas Mikroskil, Medan, Indonesia)
3. Romindo, S.Kom., M.Kom. (Scopus ID: 57222623835, Universitas Pelita Harapan, Medan, Indonesia)

4. Kelik Sussolaikah, S.Kom., M.Kom. (Scopus ID: 57209271915, Universitas PGRI, Madiun, Indonesia)
5. Ade Maulana, S.Kom., M.T.I. (Scopus ID: 57211039635, Universitas Pelita Harapan, Medan, Indonesia)
6. Darsono Nababan, S.Kom., M.Kom. (Scopus ID: 57203861998, Universitas Timor, Kefamenanu, Indonesia)
7. Yudhistira Adhitya Pratama, S.Kom., M.Kom. (Scopus ID: 57987583100, Universitas Sumatera Utara, Medan, Indonesia)
8. I Made Ardwi Pradnyana, S.T., M.T. (Scopus ID: 57202607891, Universitas Pendidikan Ganesha, Bali, Indonesia)
9. Dr. Adli Abdillah Nababan, S.Kom., M.Kom. (Scopus ID: 57202310083, Universitas Prima Indonesia, Medan, Indonesia)
10. Oman Somantri, S.Kom., M.Kom. (Scopus ID: 57208898676, Politeknik Negeri Cilacap, Indonesia)
11. Frans Mikael Sinaga, S.Kom., M.Kom. (Scopus ID: 57215527202, Universitas Mikroskil, Medan, Indonesia)

JDIMS (Journal of Data Mining and Information Systems)

Vol. 3 No. 1 (Februari 2025)

Daftar Isi

Klasifikasi Kualitas Udara di Jakarta Pada Bulan Agustus 2024 Menggunakan Algoritma C4.5 <i>Ika Juni Nur Ayu, Nayla Rahmania Putri, Ridho Putra Nugraha, Ryan Febriansyah</i>	01-08
Analisis Persepsi Publik Terhadap Pilkada Jakarta 2024 dengan Clustering dan Sentimen pada Artikel Berita <i>Anggi Nur Zavira, Ilma Fathiyarahmani, Khansa Nadhifa, Kinanti Anindia Putri, Widya Amanda Mulyadi, Zalfa Syawaliana</i>	09-17
Studi Literatur Perancangan Arsitektur Data dan Aplikasi pada Perusahaan Telekomunikasi <i>Jason Jason, Jefferson Sutanto, Verrel Angkasa, Vicky Darmana, Ade Maulana</i> ...	18-25
Color Space Influence on Photosynthetic Pigment Measurement Accuracy Using CNN in Color Constancy <i>Ade May Luky Harefa, Arief Muhazir Insandi</i>	26-35
Pengembangan Aplikasi Insentif dan Komisi Salesman sebagai Strategi Peningkatan Kinerja SDM Marketing <i>Suwandi Suwandi, Muhammad Hatta, Turini Turini, Safitri Akbari, Limbong Aprilina Yanti</i>	36-42
Analisis Average Waiting Time Penjadwalan CPU Menggunakan Algoritma Shortest Remaining First dan Algoritma Round Robin <i>Ronald Belferik, Evander Banjarnahor</i>	43-53

Klasifikasi Kualitas Udara di Jakarta Pada Bulan Agustus 2024 Menggunakan Algoritma C4.5

Ika Juni Nur Ayu¹, Nayla Rahmania Putri², Ridho Putra Nugraha³, Ryan Febriansyah⁴

^{1,2,3,4}Sistem Informasi, Universitas Bina Sarana Informatika, Indonesia

Email: ¹ikajuninurayu@gmail.com, ²naylarahmaniaputri14@gmail.com, ³ridhoputra2519@gmail.com, ⁴ryanfebriansyah556@gmail.com

ABSTRAK

Udara merupakan salah satu faktor penting dalam keberlangsungan hidup manusia selain tanah dan air. Kualitas udara yang buruk akan mengakibatkan dampak negatif terhadap kesehatan manusia, keseimbangan ekosistem dan perubahan iklim. Penelitian ini bertujuan untuk mengklasifikasi kualitas udara yang ada di Jakarta pada bulan Agustus 2024 menggunakan teknik data mining dengan algoritma C4.5. Data yang dianalisis diperoleh dari website Satu Data Jakarta yang diterbitkan pada 19 Februari 2024, mencakup beberapa parameter pengukuran, yaitu `pm_sepuluh`, `pm_duakomalima`, `sulfur_dioksida`, `karbon_monoksida`, `ozon` dan `nitrogen_dioksida`. Dalam implementasinya, penelitian ini menggunakan *tools* RapidMiner untuk memproses dan menganalisis data. Hasil klasifikasi menunjukkan bahwa kualitas udara di Jakarta pada periode tersebut dapat dikategorikan menjadi dua kelompok, yaitu tidak sehat dan sedang, dengan mayoritas pengukuran masuk dalam kategori sedang. Model klasifikasi yang dihasilkan mencapai tingkat akurasi 99,35%, menunjukkan bahwa algoritma C4.5 sangat efektif dalam mengidentifikasi dan memprediksi kualitas udara di Jakarta. Hasil ini menunjukkan bahwa sebagian besar data pengukuran kualitas udara di Jakarta masih berada dalam kategori yang memenuhi standar kualitas udara yang baik.

Kata Kunci: Kualitas Udara, Algoritma C4.5, Data Mining, Jakarta

ABSTRACT

Air is one of the important factors in human survival besides land and water. Poor air quality will have a negative impact on human health, ecosystem balance and climate change. This study aims to classify air quality in Jakarta in August 2024 using data mining techniques with the C4.5 algorithm. The data analyzed was obtained from the Satu Data Jakarta website published on February 19, 2024, including several measurement parameters, namely `pm_sepuluh`, `pm_duakomalima`, `sulfur_dioxide`, `carbon_monoxide`, `ozone` and `nitrogen_dioxide`. In its implementation, this research uses RapidMiner tools to process and analyze data. The classification results show that air quality in Jakarta during the period can be categorized into two groups, namely unhealthy and moderate, with the majority of measurements falling into the moderate category. The resulting classification model achieved an accuracy rate of 99.35%, indicating that the C4.5 algorithm is very effective in identifying and predicting air quality in Jakarta. This result shows that most of the air quality measurement data in Jakarta is still in a category that meets good air quality standards.

Keywords: Air Quality, C4.5 Algorithm, Data Mining, Jakarta

Penulis Korespondensi:

Ryan Febriansyah

Email: ryanfebriansyah556@gmail.com

Article Info

Diterima: 18 Desember 2024

Direvisi: 8 Januari 2025

Disetujui: 14 Januari 2025

This is an open access article under the [CC BY](#) license.



1. PENDAHULUAN

Udara merupakan salah satu faktor penting dalam keberlangsungan hidup manusia selain tanah dan air. Kualitas udara yang buruk akan mengakibatkan dampak negatif terhadap kesehatan manusia, keseimbangan ekosistem dan perubahan iklim. Jakarta merupakan kota besar di Indonesia yang menghadapi masalah terkait kualitas udara yang harus ditangani. *Data mining*

adalah salah satu teknik yang digunakan untuk menggali informasi yang berharga dari data yang besar dan kompleks. Teknik yang akan digunakan dalam penelitian ini adalah *data mining*.

Data mining adalah teknik analisis data otomatis untuk menemukan atau memisahkan hubungan dari sejumlah besar data yang sebelumnya tidak diketahui [1]. *Data mining* merupakan salah satu teknik untuk menggali atau “menambang” pengetahuan dari sekumpulan besar data.

Data Mining mempunyai 5 fungsi: *classification*, yaitu menyimpulkan definisi-definisi karakteristik sebuah grup. Contoh: pelanggan-pelanggan perusahaan yang telah berpindah kesainan perusahaan yang lain. *Clustering*, yaitu mengidentifikasi kelompok-kelompok dari barang-barang atau produk-produk yang mempunyai karakteristik khusus (*clustering* berbeda dengan *classification*, dimana pada *clustering* tidak terdapat definisi-definisi karakteristik awak yang di berikan pada waktu *classification*.) *Association*, yaitu mengidentifikasi hubungan antara kejadian-kejadian yang terjadi pada suatu waktu, seperti isi-isi dari keranjang belanja. *Sequencing*, hampir sama dengan *association*, *sequencing* mengidentifikasi hubungan-hubungan yang berbeda pada suatu periode waktu tertentu, seperti pelanggan-pelanggan yang mengunjungi supermarket secara berulang-ulang. *Forecasting* memperkirakan nilai pada masa yang akan datang berdasarkan pola-pola dengan sekumpulan data yang besar, seperti peramalan permintaan pasar [2].

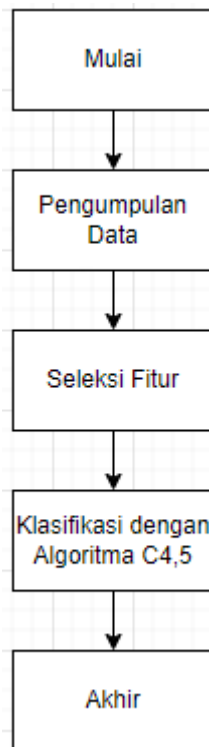
Salah satu metode *data mining* yang paling banyak digunakan adalah algoritma C4.5. Algoritma C4.5 adalah program yang menghasilkan pohon keputusan sebagai keluaran dan menyediakan sekumpulan data yang dilabelkan [3]. Pohon keputusan adalah sebuah struktur data yang terdiri dari simpul (node) dan rusuk (edge) simpul pada sebuah pohon dibedakan menjadi tiga, yaitu simpul akar (root node), simpul percabangan (branch node) dan simpul daun atau leaf node [4]. Manfaat utama menggunakan pohon keputusan adalah kemampuan untuk menyederhanakan proses pengambilan keputusan yang kompleks sehingga pengambil keputusan dapat menafsirkan solusi untuk masalah [5]. Dimana Pohon keputusan termasuk metode klasifikasi dan prediksi yang terkenal karena sangat mudah dipahami dengan bahasa alami sehingga dapat diinterpretasikan dengan cepat [6].

Dalam hal ini, penerapan algoritma C4.5 menjadi jawaban yang efektif dalam mengklasifikasikan kualitas udara di Jakarta. Dengan menganalisa data udara pada dataset yang diperoleh dari website Satu Data Jakarta algoritma C4.5 dapat menemukan pola dan hubungan pada data kualitas udara.

Oleh karena itu, tujuan penelitian ini adalah mengimplementasikan algoritma C4.5 pada data kualitas udara Jakarta untuk mengidentifikasi faktor-faktor yang mempengaruhi kualitas udara dan memberikan rekomendasi untuk meningkatkan kualitas udara di wilayah tersebut.

2. METODE PENELITIAN

Gambar di bawah ini adalah alur tahapan yang akan dilakukan.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Penelitian ini memanfaatkan *dataset* yang berasal dari website Satu data Jakarta yang diterbitkan pada 19 Februari 2024. *Dataset* berisi informasi seperti pengukuran PM10, PM2.5, SO₂, CO, O₃ dan NO₂. PM10 merupakan campuran asap yang kompleks dan heterogen, jelaga, debu, garam, asam dan logam yang mempunyai konsentrasi bervariasi, ukuran, komposisi kimia, luas permukaan dan sumber asalnya [7]. PM2,5 merupakan salah satu zat pencemar yang bebas di udara dan bertahan di udara cukup lama, ukurannya kecil hanya 2,5 mikron dapat masuk kedalam tubuh manusia lewat sistem pernafasan [8]. SO₂ merupakan gas tak berwarna yang memiliki toksisitas tinggi dan berbahaya bagi kesehatan terutama terhadap orang tua, anak-anak dan penderita yang mengalami penyakit kronis pada sistem pernafasan dan kardiovaskular [9].

Karbon monoksida (CO) merupakan gas yang tidak memiliki warna dan bau, yang dapat membahayakan apabila terhirup dengan jumlah yang besar. Sumber gas CO di udara adalah mobil, truk dan kendaraan lainnya, barang yang ada di rumah seperti pemanas ruangan dengan minyak tanah, cerobong asap, dan tungku yang bocor, kompor gas, asap rokok merupakan benda yang dapat melepaskan gas CO di dalam ruangan [10]. Ozon (O₃) merupakan suatu zat gas yang terkandung pada udara yang kita hirup. Ozon yang buruk terbentuk di dekat permukaan tanah ketika polutan (yang disebabkan oleh mobil, pembangkit listrik, industri, kilang minyak, dan pabrik kimia) bereaksi secara kimia dengan cahaya matahari [11]. NO₂ merupakan salah satu jenis polutan yang dapat menurunkan kualitas udara dan berdampak pada kesehatan manusia dan lingkungan. Gas NO₂ termasuk prekursor utama dalam pembentukan Ozon (O₃) di atmosfer [12].

2.2. Seleksi Fitur

Seleksi fitur bertujuan guna mengurangi atribut, sebab banyaknya fitur pada *dataset* yang digunakan dapat mengakibatkan *overfitting* dan perlu menentukan fitur yang dibutuhkan dalam pengolahan data. Selain itu, dilakukan normalisasi dan pemilihan atribut yang paling berpengaruh dalam klasifikasi kualitas udara [13].

2.3. Klasifikasi Dengan Algoritma C4.5

Klasifikasi merupakan salah satu teknik dalam data mining. Klasifikasi (taksonomi) merupakan proses penempatan objek atau konsep tertentu ke dalam satu set kategori berdasarkan objek yang digunakan [14]. Algoritma C4.5 dapat menangani atribut kontiniu dan diskrit, dapat menyelesaikan data tanpa nilai dari atribut., serta dapat meminimalkan pohon keputusan sesudah dibuat [15]. Dalam Penelitian ini untuk memodelkan pohon keputusan peneliti menggunakan *tools* rapidminer. RapidMiner adalah platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Ini menyediakan beragam alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. RapidMiner dirancang untuk mudah digunakan dan memungkinkan pengguna untuk dengan mudah membangun dan menguji berbagai model, bahkan tanpa pengalaman pemrograman [16]

3. HASIL DAN PEMBAHASAN

3.1. Pre-Processing

Dataset yang digunakan didapatkan dari website Satu Data Jakarta yang diterbitkan pada 19 Februari 2024 dalam bentuk format XLSX (*Microsoft Excel*). *Dataset* ini memuat pengukuran kualitas udara di Provinsi DKI Jakarta yang diterbitkan pada tahun 2024. Berikut penjelasan parameter yang terdapat di dalam *dataset*:

- periode_data: penjelasan periode data 1 bulan sekali.
- bulan: pengambilan data pengukuran udara.
- tanggal: tanggal pengambilan data pengukuran udara.
- stasiun: lokasi penempatan alat pengukuran udara.
- pm_sepuluh: partikulat dengan ukuran di bawah 10 mikron.
- pm_duakomalima: partikulat dengan ukuran di bawah 2.5 mikron.
- karbon_monoksida: pengukuran zat karbon monoksida (CO).
- ozon: pengukuran zat ozon (O₃).
- nitrogen_dioksida: pengukuran zat nitrogen dioksida (NO₂).
- max: hasil pengukuran tertinggi pada setiap parameter yang ditangkap oleh stasiun pemantau.
- parameter_pencemaran_kritis: nama parameter yang memiliki nilai tertinggi.
- kategori: kategori hasil pengukuran yang didapat dari perhitungan rata-rata setiap parameter yang kemudian diklasifikasikan berdasarkan Indeks Standar Pencemaran Udara.

Tabel 1. Tabel ISPU

No	Nilai ISPU	Kategori
1	0 – 15	Baik
2	15 – 65	Sedang
3	66 – 150	Tidak Sehat
4	151 – 250	Sangat Tidak Sehat
5	>250	Berbahaya

Secara keseluruhan, *dataset* ini memuat 155 *record* data dengan 12 atribut dan 1 kelas. Atribut yang terdapat pada *dataset* ini, yaitu *periode_data*, *bulan*, *tanggal*, *stasiun*, *pm_sepuluh*, *pm_duakomalima*, *sulfur_dioksida*, *karbon_monoksida*, *ozon*, *nitrogen_dioksida*, *max*, *parameter_pencemar_kritis* dan *kategori*.

Tabel 2. Tabel Penggabungan Data

periode_data	bulan	tanggal	stasiun	pm_sepuluh	pm_duakomalima	sulfur_dioksida	karbon_monoksida	ozon	nitrogen_dioksida	max	parameter_pencemar_kritis	kategori
202408	8	1	DKI1 Bundaran Hotel Indonesia (HI)	68	103	13	28	26	44	103	PM25	TIDAK SEHAT
202408	8	2	DKI1 Bundaran Hotel Indonesia (HI)	60	85	11	27	37	44	85	PM25	SEDANG
202408	8	3	DKI1 Bundaran Hotel Indonesia (HI)	47	64	10	28	26	45	64	PM25	SEDANG
202408	8	4	DKI1 Bundaran Hotel Indonesia (HI)	50	74	10	29	29	39	74	PM25	SEDANG
202408	8	5	DKI1 Bundaran Hotel Indonesia (HI)	53	78	10	26	23	37	78	PM25	SEDANG
202408	8	6	DKI1 Bundaran Hotel Indonesia (HI)	53	75	16	24	21	42	75	PM25	SEDANG
202408	8	7	DKI1 Bundaran Hotel Indonesia (HI)	62	91	13	26	27	42	91	PM25	SEDANG
202408	8	8	DKI1 Bundaran Hotel Indonesia (HI)	64	96	16	28	23	42	96	PM25	SEDANG
202408	8	11	DKI1 Bundaran Hotel Indonesia (HI)	56	77	12	22	28	28	77	PM25	SEDANG
202408	8	12	DKI1 Bundaran Hotel Indonesia (HI)	54	78	12	23	24	30	78	PM25	SEDANG
202408	8	13	DKI1 Bundaran Hotel Indonesia (HI)	62	93	19	26	25	44	93	PM25	SEDANG
202408	8	14	DKI1 Bundaran Hotel Indonesia (HI)	62	93	17	26	27	43	93	PM25	SEDANG
202408	8	15	DKI1 Bundaran Hotel Indonesia (HI)	63	90	14	26	38	38	90	PM25	SEDANG
202408	8	16	DKI1 Bundaran Hotel Indonesia (HI)	64	91	15	28	33	48	91	PM25	SEDANG
202408	8	17	DKI1 Bundaran Hotel Indonesia (HI)	65	92	15	25	31	44	92	PM25	SEDANG
202408	8	18	DKI1 Bundaran Hotel Indonesia (HI)	62	93	13	23	30	32	93	PM25	SEDANG
202408	8	19	DKI1 Bundaran Hotel Indonesia (HI)	60	86	13	24	32	38	86	PM25	SEDANG
202408	8	20	DKI1 Bundaran Hotel Indonesia (HI)	64	93	14	26	27	46	93	PM25	SEDANG
202408	8	21	DKI1 Bundaran Hotel Indonesia (HI)	67	96	17	27	27	47	96	PM25	SEDANG
202408	8	22	DKI1 Bundaran Hotel Indonesia (HI)	63	84	19	25	26	40	84	PM25	SEDANG
202408	8	23	DKI1 Bundaran Hotel Indonesia (HI)	56	77	12	23	28	32	77	PM25	SEDANG
202408	8	24	DKI1 Bundaran Hotel Indonesia (HI)	62	98	13	27	24	43	98	PM25	SEDANG
202408	8	25	DKI1 Bundaran Hotel Indonesia (HI)	66	90	13	25	28	37	90	PM25	SEDANG
202408	8	26	DKI1 Bundaran Hotel Indonesia (HI)	55	78	13	24	28	39	78	PM25	SEDANG
202408	8	27	DKI1 Bundaran Hotel Indonesia (HI)	69	106	14	30	36	50	106	PM25	TIDAK SEHAT
202408	8	28	DKI1 Bundaran Hotel Indonesia (HI)	62	86	14	24	33	41	86	PM25	SEDANG
202408	8	29	DKI1 Bundaran Hotel Indonesia (HI)	65	98	14	26	41	42	98	PM25	SEDANG
202408	8	30	DKI1 Bundaran Hotel Indonesia (HI)	68	98	15	28	30	46	98	PM25	SEDANG
202408	8	31	DKI1 Bundaran Hotel Indonesia (HI)	56	75	14	25	24	40	75	PM25	SEDANG

Proses selanjutnya menyeleksi data. Pada penelitian penulis mengambil 7 kolom dengan 6 atribut, yaitu *pm_sepuluh*, *pm_duakomalima*, *sulfur_dioksida*, *karbon_monoksida*, *ozon*, *nitrogen_dioksida* dan *kategori*. Kemudian masuk ke tahap pembersihan data untuk menghapus, memperbaiki dan mengidentifikasi data yang tidak akurat atau tidak valid dengan melakukan penataan ulang data yang terdapat dalam catatan, tabel atau *database*. Sampel yang perlu dilakukan pembersihan data adalah sebagai berikut:

Tabel 3. Tabel Pembersihan Data

pm_sepuluh	pm_duakomalima	sulfur_dioksida	karbon_monoksida	ozon	nitrogen_dioksida	kategori
68	103	13	28	26	44	TIDAK SEHAT
60	85	11	27	37	44	SEDANG
47	64	10	28	26	45	SEDANG
50	74	10	29	29	39	SEDANG
53	78	10	26	23	37	SEDANG
53	75	16	24	21	42	SEDANG
62	91	13	26	27	42	SEDANG
64	96	16	28	23	42	SEDANG
61	91	15	25	19	33	SEDANG
55	81	15	23	21	35	SEDANG
56	77	12	22	28	28	SEDANG
54	78	12	23	24	30	SEDANG
62	93	19	26	25	44	SEDANG
62	93	17	26	27	43	SEDANG
63	90	14	26	38	38	SEDANG
64	91	15	28	33	48	SEDANG
65	92	15	25	31	44	SEDANG
62	93	13	23	30	32	SEDANG
60	86	13	24	32	38	SEDANG
64	93	14	26	27	46	SEDANG
67	96	17	27	27	47	SEDANG
63	84	19	25	26	40	SEDANG
56	77	12	23	28	32	SEDANG
62	98	13	27	24	43	SEDANG
66	90	13	25	28	37	SEDANG
55	78	13	24	28	39	SEDANG
69	106	14	30	36	50	TIDAK SEHAT
62	86	14	24	33	41	SEDANG
65	98	14	26	41	42	SEDANG

Setelah dilakukan pembersihan data maka didapatkan data yang akan siap diolah. Data inilah yang akan dilakukan klasifikasi menggunakan *Tools Rapidminer*. Berikut tampilan data siap olah:

Tabel 4. Tabel Siap Olah

pm_sepuluh	pm_duakomalima	sulfur_dioksida	karbon_monoksida	ozon	nitrogen_dioksida	kategori
68	103	13	28	26	44	TIDAK SEHAT
60	85	11	27	37	44	SEDANG
47	64	10	28	26	45	SEDANG
50	74	10	29	29	39	SEDANG
53	78	10	26	23	37	SEDANG
53	75	16	24	21	42	SEDANG
62	91	13	26	27	42	SEDANG
64	96	16	28	23	42	SEDANG
61	91	15	25	19	33	SEDANG
55	81	15	23	21	35	SEDANG
56	77	12	22	28	28	SEDANG
54	78	12	23	24	30	SEDANG
62	93	19	26	25	44	SEDANG
62	93	17	26	27	43	SEDANG
63	90	14	26	38	38	SEDANG
64	91	15	28	33	48	SEDANG
65	92	15	25	31	44	SEDANG
62	93	13	23	30	32	SEDANG
60	86	13	24	32	38	SEDANG
64	93	14	26	27	46	SEDANG
67	96	17	27	27	47	SEDANG
63	84	19	25	26	40	SEDANG
56	77	12	23	28	32	SEDANG
62	98	13	27	24	43	SEDANG
66	90	13	25	28	37	SEDANG
55	78	13	24	28	39	SEDANG
69	106	14	30	36	50	TIDAK SEHAT
62	86	14	24	33	41	SEDANG
65	98	14	26	41	42	SEDANG

3.2. Proses Klasifikasi dan Pohon Keputusan

Algoritma C4.5 menerapkan pohon keputusan dengan langkah-langkah berikut. Proses pertama adalah memilih atribut sebagai akar pohon keputusan. Selanjutnya, membuat cabang untuk setiap nilai berdasarkan akar tersebut. Kemudian, bagi kasus-kasus tersebut ke dalam cabang. Proses ini diulang pada cabang hingga tercapai kelas yang sama. Algoritma C4.5 menggunakan metode gain untuk menentukan atribut akar. Gain dihitung untuk semua atribut yang ada, apabila atribut dengan nilai gain tertinggi maka akan dipilih sebagai atribut akar pohon keputusan. Maka untuk menghitung klasifikasi data dengan algoritma C4.5 dapat dihitung gain dan entropi untuk membuat pohon keputusan dari semua nilai yang ada pada tabel data. Berikut tampilan hasil perhitungan gain dan entropi.

Tabel 5. Tabel Gain dan Entropi

Attribute		Jumlah (s)	Kepuasan			Entropy	Gain
			Baik	Tidak sehat	Sedang		
Total		155	1	32	122	0,78870534	
pm_sepuluh	<= 50	19	1	0	18	0,00000000	0,788705
	> 50	136	0	104	32	0,00000000	
pm_duakomalima	<=100	122	1	0	121	0,00000000	0,788705
	>100	33	0	33	0	0,00000000	
Sulfur dioksida	<=50	63	1	16	46	0,92832992	0,411384
	>50	92	0	15	77	0,00000000	
Karbon_monoksida	<= 25	135	1	30	104	0,82457524	0,070527
	>25	20	0	2	18	0,00000000	
Ozon	<=50	143	1	28	114	0,77137204	0,077052
	>50	12	0	6	6	0,00000000	
nitrogen_dioksida	<=50	155	1	122	32	0,78870534	0,000000
	>50	0	0	0	0	0,00000000	

Hasil Perhitungan gambar 5 dapat dijabarkan sebagai berikut:

Rumus Entropi:

$$Entropy (S) = \sum_{i=1}^n -p_i \cdot \log_2(p_i) \tag{1}$$

Keterangan:

S = Himpunan Kasus

A = Atribut

N = Jumlah partisi s

Pi = Proporsi dari Si terhadap S

Berikut perhitungan entropi dari masing-masing atribut:

entropi total:

$$Entropy (Total) = ((-1/155) * \log_2 (1/155)) + ((-32/155) * \log_2 (32/155)) + ((-122/155) * \log_2 (122/155)) = 0,78870534$$

Entropi sulfur_dioksida:

$$Entropy (sulfur_dioksida) = ((-1/63) * \log_2 (1/63)) + ((-16/63) * \log_2 (16/63)) + ((-46/63) * \log_2 (46/63)) = 0,92832992$$

Entropi karbon_monoksida:

$$Entropy (karbon_monoksida) = ((-1/135) * \log_2 (1/135)) + ((-30/135) * \log_2 (30/135)) + ((-104/135) * \log_2 (104/135)) = 0,82457524$$

Entropi ozon:

$$Entropy (ozon) = ((-1/143) * \log_2 (1/143)) + ((-28/143) * \log_2 (28/143)) + ((-114/143) * \log_2 (114/143)) = 0,77137204$$

Entropi nitrogen_dioksida:

$$Entropy (nitrogen_dioksida) = ((-1/155) * \log_2 (1/155)) + ((-122/155) * \log_2 (122/155)) + ((-3/155) * \log_2 (32/155)) = 0,78870534$$

Lalu perhitungan untuk gain:

Rumus Gain:

$$Gain (S, A) = Entropy (S) - \sum_{i=1}^n |S_i| * Entropy (S_i) \tag{2}$$

Keterangan:

S = Himpunan Kasus

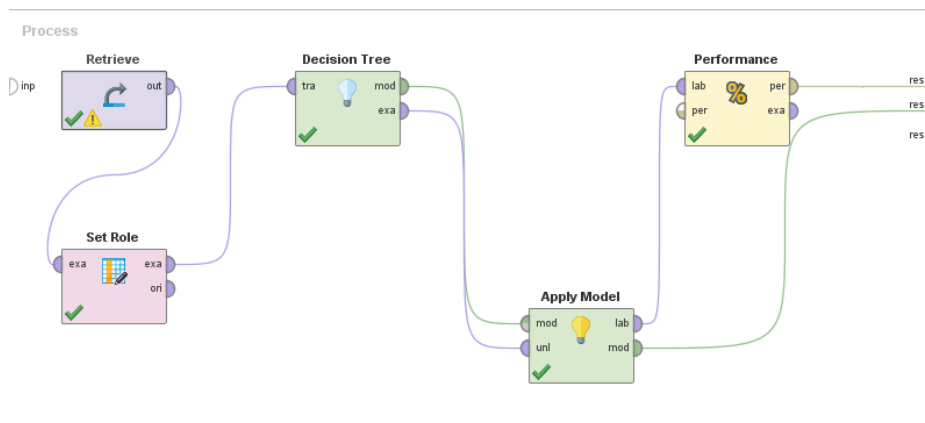
A = Atribut

N = Jumlah partisi pada atribut A

|Si| = Jumlah kasus pada partisi ke-i

$$Gain (pm_duakomalima) = 0,78870534 - ((122/155) * 0,00000000) - ((33/155) * 0,00000000) = 0,788705$$

Selanjutnya adalah klasifikasi menggunakan pohon keputusan pada *tools* rapidminer. Berikut adalah langkah-langkahnya:



Gambar 2. Proses Klasifikasi

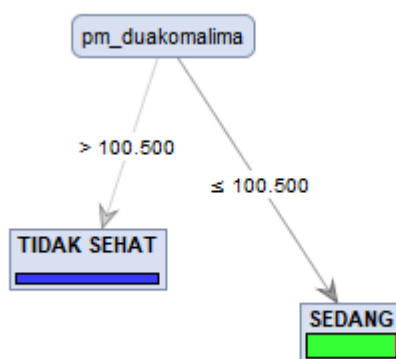
Sementara itu, pada gambar 7 di bawah ini adalah hasil dari proses klasifikasi pada gambar 6.

PerformanceVector

```
PerformanceVector:  
accuracy: 99.35%  
ConfusionMatrix:  
True:   TIDAK SEHAT   SEDANG   BAIK  
TIDAK SEHAT:   32     0     0  
SEDANG:  0     122    1  
BAIK:    0     0     0
```

Gambar 3. Output Klasifikasi

Akhirnya didapatkan pohon keputusan sebagai berikut.



Gambar 4. Pohon Keputusan

Berdasarkan pohon keputusan dari algoritma C4.5 untuk melakukan klasifikasi kualitas udara di Jakarta pada bulan Agustus, diperoleh nilai sebagai berikut:

1. Jika nilai `pm_duakomalima` lebih besar dari 100.500, maka kualitas udara masuk dalam kategori tidak sehat.
2. Jika nilai `pm_duakomalima` kurang dari atau sama dengan 100.500, maka kualitas udara masuk dalam kategori sedang.

4. KESIMPULAN

Berdasarkan pembahasan dan analisis data yang telah diuji menggunakan algoritma C4.5 untuk mengklasifikasi Kualitas udara di Jakarta pada bulan Agustus berdasarkan data yang diperoleh dari website Satu Data Jakarta, data tersebut berisi 7 parameter untuk mengukur kualitas udara, yaitu `pm_sepuluh`, `pm_duakomalima`, `sulfur_dioksida`, `karbon_monoksida`, `ozon`, `nitrogen_dioksida` dan kategori. Dengan melakukan penerapan *data mining* menggunakan algoritma C4.5 untuk melakukan klasifikasi udara di Jakarta pada bulan Agustus, diketahui bahwa kualitas udara diklasifikasikan menjadi 2 kategori, yaitu tidak sehat dan sedang. Sehingga dapat disimpulkan bahwa kualitas udara di Jakarta pada bulan Agustus masuk ke dalam kategori sedang dengan tingkat akurasi mencapai 99,35%. Hasil ini menunjukkan bahwa model klasifikasi mampu mengidentifikasi dan memprediksi dengan tingkat akurasi yang tinggi dengan menunjukkan bahwa mayoritas data pengukuran kualitas udara berada dalam kategori yang mencukupi standar untuk kualitas udara yang baik.

REFERENSI

- [1] W. W. Kristianto and C. Rudianto, "Penerapan Data Mining Pada Penjualan Produk Menggunakan Metode K- Means Clustering (Studi Kasus Toko Sepatu Kakikaki)," *J. Pendidik. Teknol. Inf.*, no. 5, pp. 90–98, 2020.
- [2] C. Zai, "IMPLEMENTASI DATA MINING SEBAGAI PENGOLAHAN DATA," *J. Portal Data*, vol. 2, no. 3, pp. 1–12, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>
- [3] M. Andreas Rayando, I. Fajri, F. Kurniawan, M. Stefanus Rafael Sihotang, and S. Suharjanti, "Analisis Tingkat Kepuasan Masyarakat Terhadap Kinerja Ketua Rw 007 Menggunakan Algoritma C4.5," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 4, pp.

- 7463–7469, 2024, doi: 10.36040/jati.v8i4.10334.
- [4] J. A. Sidette, E. Eko, and O. D. Nurhayati, “Pendekatan Metode Pohon Keputusan Menggunakan Algoritma ID3 Untuk Sistem Informasi Pengukuran Kinerja PNS,” *J. Sist. Inf. Bisnis*, vol. 4, no. 2, pp. 75–86, 2014, doi: 10.21456/vol4iss2pp75-86.
- [5] P. R. Adinda, “Pohon Keputusan C4. 5 Algoritma Untuk Klasifikasi Program Bantuan Belajar,” *J. Portal Data*, vol. 2, no. 9, pp. 1–13, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/215%0Ahttp://portaldata.org/index.php/portaldata/article/download/215/204>
- [6] S. Ucha Putri, E. Irawan, F. Rizky, S. Tunas Bangsa, P. A. -Indonesia Jln Sudirman Blok No, and S. Utara, “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5,” *Januari*, vol. 2, no. 1, pp. 39–46, 2021.
- [7] E. Nur, B. A. Seno, and R. Hidayanti, “Risk of Public Health Problems Due to PM10 Exposure in Padang City,” *J. Kesehatan Lingkungan. Indones.*, vol. 20, no. 2, pp. 97–103, 2021, [Online]. Available: <https://shorturl.asia/6u94Z>
- [8] A. E. Saputra, “Identifikasi Paparan PM2,5 Di Wilayah Kota Cimahi,” *J. Kesehat. Kartika*, vol. 16, no. 3, pp. 104–109, 2021, doi: 10.26874/jkkes.v16i3.205.
- [9] A. R. Nurfadillah and S. Petasule, “ENVIRONMENTAL HEALTH RISK ANALYSIS (SO2, NO2, CO and TSP) IN THE BONE BOLANGO AREA ROAD SEGMENT,” *J. Heal. Sci. Gorontalo J. Heal. Sci. Community*, vol. 6, no. 2, pp. 76–89, 2022, doi: 10.35971/gojhes.v5i3.13544.
- [10] M. A. Rizaldi, R. Azizah, M. T. Latif, L. Sulistyorini, and B. P. Salindra, “Literature Review: Dampak Paparan Gas Karbon Monoksida Terhadap Kesehatan Masyarakat yang Rentan dan Berisiko Tinggi,” *J. Kesehat. Lingkungan. Indones.*, vol. 21, no. 3, pp. 253–265, 2022, doi: 10.14710/jkli.21.3.253-265.
- [11] R. Wahyudi, S. Annas, and Z. Rais, “Analisis Support Vector Regression (Svr) Untuk Meramalkan Indeks Kualitas Udara Di Kota Makassar,” *Variansi J. Stat. Its Appl. Teach. Res.*, vol. 5, no. 3, pp. 104–117, 2023, doi: 10.35580/variasiunm107.
- [12] Y. Serlina, “Pengaruh Faktor Meteorologi Terhadap Konsentrasi NO2 di Udara Ambien (Studi Kasus Bundaran Hotel Indonesia DKI Jakarta),” *J. Serambi Eng.*, vol. 5, no. 3, 2020, doi: 10.32672/jse.v5i3.2146.
- [13] A. Z. D. Nur Adiya, A. F. Desvita, A. Fidela, D. Amelia, and T. Astuti, “Penerapan Data Mining Untuk Klasifikasi Kualitas Udara di Daerah Istimewa Yogyakarta Menggunakan Algoritma C4.5,” *JDMIS J. Data Min. Inf. Syst.*, vol. 2, no. 2, pp. 59–65, 2024, doi: 10.54259/jdmis.v2i2.2800.
- [14] D. Ardiansyah and W. Walim, “Algoritma c4.5 untuk klasifikasi calon peserta lomba cerdas cermat siswa smp dengan menggunakan aplikasi rapid miner,” *J. Inkofar*, vol. 1, no. 2, pp. 5–12, 2018, [Online]. Available: <https://politeknikmeta.ac.id/meta/ojs/index.php/inkofar/article/view/29/45>
- [15] Y. Hendra Kusuma, S. Suprapto, and Y. Setiawan, “Analisis Kepuasan Penumpang pada Maskapai Penerbangan Menggunakan Algoritma C4.5 dan Naïve Bayes,” *SENTIMAS Semin. Nas. Penelit. dan Pengabd. Masy.*, pp. 162–171, 2022, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas/article/view/320/125>
- [16] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, “Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.

Analisis Persepsi Publik Terhadap Pilkada Jakarta 2024 dengan Clustering dan Sentimen pada Artikel Berita

Anggi Nur Zavira¹, Ilma Fathiyarahmani², Khansa Nadhifa³, Kinanti Anindia Putri⁴, Widya Amanda Mulyadi⁵, Zalfa Syawaliana⁶

^{1,2,3,4,5,6} Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta, Indonesia

Email: ¹anggi_1314622009@mhs.unj.ac.id, ²ilma_1314622023@mhs.unj.ac.id, ³khansa_1314622032@mhs.unj.ac.id, ⁴kinanti_1314622024@mhs.unj.ac.id, ⁵widya_1314622025@mhs.unj.ac.id, ⁶zalfa_1314622027@mhs.unj.ac.id

ABSTRAK

Penelitian ini menganalisis pola pemberitaan media massa terkait Pilkada Jakarta 2024 menggunakan pendekatan *text mining* dengan algoritma K-Means Clustering dan analisis sentimen berbasis Lexicon. Data diperoleh melalui *web scraping* dari RSS Feed Google News, menghasilkan 100 artikel yang dianalisis setelah melalui proses *preprocessing* berupa *tokenizing*, *filtering*, dan *stemming*. Algoritma K-Means digunakan untuk mengelompokkan artikel ke dalam delapan *cluster* berdasarkan tema dominan, seperti dukungan politik, kegagalan pencalonan, dan isu strategis terkait Pilkada. Algoritma ini bekerja dengan menghitung jarak antara data dan centroid, yang terus diperbarui hingga mencapai *cluster* optimal berdasarkan metode *Silhouette Score*. Analisis sentimen menunjukkan bahwa sebagian besar artikel memiliki sentimen netral, mencerminkan objektivitas media, meskipun beberapa *cluster* memperlihatkan adanya sentimen positif dan negatif, yang mengindikasikan potensi bias dalam pemberitaan. Temuan ini memberikan wawasan mengenai peran media dalam membentuk opini publik serta pengaruh pola pemberitaan terhadap persepsi masyarakat dalam proses demokrasi. Penelitian ini diharapkan dapat meningkatkan literasi politik masyarakat dan mendorong partisipasi yang lebih kritis, sekaligus membuka peluang pengembangan metode analisis lebih lanjut untuk memahami bias media secara lebih mendalam.

Kata Kunci: Text Mining, K-Means Clustering, Sentiment Analysis

ABSTRACT

This study analyzes the pattern of mass media coverage related to the 2024 Jakarta gubernatorial election using a text mining approach with the K-Means Clustering algorithm and Lexicon-based sentiment analysis. Data were obtained through web scraping from Google News RSS Feeds, resulting in 100 articles that were analyzed after undergoing preprocessing processes such as tokenizing, filtering, and stemming. The K-Means algorithm was used to cluster the articles into eight clusters based on dominant themes, such as political support, candidacy failures, and strategic issues related to the election. This algorithm works by calculating the distance between data points and centroids, which are continuously updated until an optimal cluster is achieved based on the Silhouette Score method. Sentiment analysis revealed that most articles had a neutral sentiment, reflecting media objectivity, although some clusters showed positive and negative sentiments, indicating potential bias in the coverage. These findings provide insights into the role of media in shaping public opinion and the influence of news coverage on public perceptions in the democratic process. This study is expected to enhance political literacy among the public and encourage more critical participation, while also opening opportunities for further development of analysis methods to better understand media bias.

Keywords: Text Mining, K-Means Clustering, Sentiment Analysis

Penulis Korespondensi:

Widya Amanda Mulyadi

Email: widya_1314622025@mhs.unj.ac.id

Article Info

Diterima: 26 Desember 2024

Direvisi: 3 Januari 2025

Disetujui: 12 Februari 2025

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.



1. PENDAHULUAN

Pilkada merupakan momen yang sangat penting dalam sistem demokrasi di Indonesia terutama pada Pilkada Jakarta dimana memiliki dampak signifikan terhadap dinamika politik nasional. Salah satu sumber informasi bagi masyarakat untuk mengetahui informasi terbaru terkait perkembangan isu politik, kompetisi para calon, serta strategi kampanye yang telah diterapkan oleh partai politik ataupun individu adalah media massa terutama media massa *online* yang sangat *online* sangat memudahkan masyarakat untuk mendapatkan informasi terkait Pilkada Jakarta 2024.

Media massa memiliki peran penting dalam pesta demokrasi di Indonesia dengan menyalurkan berbagai informasi politik, media massa dapat memengaruhi masyarakat untuk pro aktif dalam mendukung terciptanya sistem demokrasi melalui pemilihan umum yang sesuai dengan amanat undang-undang yakni langsung, umum, bebas, rahasia, jujur dan adil (LUBER JURDIL). Namun, netralitas media massa sering kali menjadi isu yang diperdebatkan, terutama dalam konteks pemberitaan politik. Menurut penelitian pada Pilkada Manado 2020 yang dilakukan oleh Claudia Rahmim, dkk [1] terdapat dua media massa yang tidak netral dalam penyajian data yaitu, media lokal Sulut Online dan Berita Manado, ketidaknetralan ini terlihat dari porsi pemberitaan yang tidak seimbang dan isi berita yang sering memihak pada kandidat tertentu.

Indonesia Corruption Watch (ICW) membuka diskusi publik mengenai kritisi kampanye Pilkada Jakarta 2024 yang dianggap minim gagasan konkret juga kurang solutif terhadap isu-isu masyarakat. Para paslon dinilai terlalu kota-sentris sehingga mengabaikan permasalahan mendesak seperti perampasan ruang hidup, krisis iklim, kerentanan ekonomi kelas menengah, dan maraknya praktik korupsi pada pengelolaan anggaran daerah. Perubahan status Jakarta menjadi pusat ekonomi nasional juga mempengaruhi jalannya Pilkada Jakarta. Ternormalisasinya politik uang, lemahnya visi antikorupsi, serta kegagalan memahami kebutuhan masyarakat semakin memperburuk legitimasi proses demokrasi.

Pilkada Jakarta tidak hanya berdampak pada dinamika politik, tetapi juga pada aspek-aspek sosial dalam masyarakat. Pada proses pemilihan sering kali memicu diskusi publik yang cukup intens sehingga memengaruhi kohesi sosial dan membentuk kesadaran politik di kalangan masyarakat. Dalam hal ini, peran media massa sangatlah penting untuk membentuk opini-opini publik maka dari itu media massa haruslah menyuguhkan pemberitaan yang bersifat netral agar tidak memicu bias atau konflik lebih lanjut, dengan memahami pola pemberitaan dan sentimen media massa sangat penting untuk menangkap fenomena sosial yang berkembang selama Pilkada Jakarta. Tingginya volume informasi yang tersedia menyebabkan tantangan tersendiri, baik bagi masyarakat umum maupun bagi peneliti yang ingin menganalisis tren pemberitaan. Adanya data berita yang sangat beragam sering kali sulit dipahami tanpa adanya sistem pengelompokan atau analisis yang terstruktur.

Untuk mengatasi tantangan tersebut, teknologi seperti *text mining* dan algoritma *K-Means Clustering* menjadi solusi yang efektif untuk mengelompokkan artikel-artikel berita berdasarkan kesamaan tema, sehingga dapat membantu dalam memahami pola pemberitaan secara lebih jelas. Selain pengelompokan berita, penelitian ini akan menambahkan analisis sentimen pada setiap kluster. Analisis sentimen akan dilakukan menggunakan *text mining* untuk mengklasifikasikan sentimen menjadi tiga kategori utama, yaitu sentimen positif, di mana artikel memberikan pandangan atau opini yang mendukung calon tertentu; sentimen negatif, di mana artikel berisi kritik, isu kontroversial, atau sentimen buruk terhadap calon atau kebijakan tertentu; dan sentimen netral, di mana artikel bersifat informatif dan tidak memihak.

Penelitian yang dilakukan oleh Yepin Andrianus, dkk [2] membahas implementasi algoritma *K-Means Clustering* untuk menganalisis opini masyarakat terkait pemilu 2024 berdasarkan data dari Twitter. Hasil yang didapat menunjukkan bahwa mayoritas opini masyarakat bersifat netral dengan persentase 92,50%, sementara sentimen positif hanya 4,30% dan sentimen negatif sebesar 3,20%. Tingkat akurasi pengelompokan diuji menggunakan metode *Silhouette Coefficient*, yang menunjukkan hasil kuat dengan nilai 0,981. Penelitian ini membuktikan bahwa *K-Means* efektif dalam mengelompokkan opini publik dan memberikan gambaran tentang sentimen masyarakat terhadap pemilu.

Selain itu, penelitian terdahulu oleh Atika Rahmawati, dkk [3] membandingkan algoritma *Support Vector Machine* (SVM) dan *K-Means* untuk analisis sentimen publik terhadap pelaksanaan Pilkada serentak. Hasil yang didapatkan melalui metode SVM mendapatkan akurasi tertinggi sebesar 91% sedangkan *K-Means* hanya mencatat 83%. Penelitian ini menunjukkan efektivitas algoritma SVM lebih unggul dalam klasifikasi sentimen, namun *K-Means* tetap relevan untuk pengelompokan tema dalam *dataset* yang besar dan heterogen.

Tujuan penelitian ini tidak hanya memberikan kontribusi bagi para peneliti dalam memahami pola pemberitaan dan sentimen media terkait Pilkada Jakarta 2024, tetapi juga memberikan manfaat nyata bagi masyarakat umum. Melalui pengelompokan artikel berita menggunakan algoritma *K-Means Clustering* dan analisis sentimen, dengan memanfaatkan data berita yang diambil melalui teknik *web scraping* dari RSS Feed Google News, penelitian ini diharapkan masyarakat dapat memperoleh informasi terkait isu-isu yang paling banyak dibicarakan dari Pilkada Jakarta 2024 pada media massa *online* yaitu pada portal Google News. Dengan demikian masyarakat dapat memahami lebih jelas bagaimana media massa *online* membentuk opini publik terhadap Pilkada Jakarta 2024. Hasil penelitian ini diharapkan dapat meningkatkan literasi politik masyarakat, membantu mereka dalam mengidentifikasi pemberitaan yang relevan, dan mendorong partisipasi aktif serta kritis dalam proses demokrasi yang lebih berkualitas.

2. METODE PENELITIAN

2.1 Teknik Pengambilan Data

Data dalam penelitian ini diperoleh dengan menggunakan metode *web scraping* untuk mengumpulkan artikel berita dari Google News RSS Feed terkait topik “Pilkada Jakarta 2024”. Asfandyer et. al. (2023) dalam penelitian *Automatic Scientific Literature Gathering and Analysis Using Web Scraping and Locality Sensitive Hashing* menyebutkan bahwa metode *web scraping* terbukti efektif dalam pengumpulan data teks terstruktur untuk kepentingan analisis ilmiah. Sejalan dengan Papadopoulus et. al. (2022) menilai bahwa *web scraping* mendukung analisis berbasis data secara komprehensif dan akurat dalam penelitian *Big Data Analysis and Web Mining Techniques for News Articles*. Proses pengambilan data menggunakan metode *web scraping* dilakukan dengan memanfaatkan fungsi *feedparser* dalam Python yang secara otomatis mengambil, menguraikan, hingga memecah data menjadi bagian-bagian yang lebih kecil dalam XML atau JSON dari sumber RSS Feed.

Data yang dikumpulkan meliputi judul artikel, tautan berita, tanggal publikasi, serta ringkasan isi artikel dalam format terstruktur seperti CSV atau DataFrame untuk dianalisis lebih lanjut. Tahapan pengambilan data dimulai dengan mengakses RSS Feed, lalu menentukan kata kunci pencarian yang relevan, yaitu “Pilkada Jakarta 2024” melalui tautan <https://news.google.com/rss/search?q=pilkada%20jakarta%202024&hl=en-ID&gl=ID&ceid=ID%3Aen>. Selanjutnya, *feedparser* akan membaca dan mengekstrak konten dari setiap entri RSS Feed untuk memastikan konsistensi data sehingga dinyatakan layak untuk proses analisis lebih lanjut. Teknik ini memungkinkan peneliti untuk mengumpulkan data skala besar secara otomatis, sistematis, dan meminimalisir waktu dan upaya manual pengumpulan data berita dari berbagai sumber berbeda.

2.2 Analisis Clustering dengan K-Means

Analisis *clustering* dilakukan dengan menggunakan algoritma *K-Means* untuk mengelompokkan artikel berita menjadi beberapa kluster berdasarkan kemiripan tema agar memudahkan pemahaman mengenai berbagai topik yang dibahas dalam artikel berita itu sendiri. Tahapan analisis *clustering* dengan *K-Means* didahului dengan *pre-processing* data mengikuti prosedur berikut.

1. Proses *Tokenizing*.

Proses *Tokenizing* adalah proses pemotongan sebuah kalimat menjadi unit-unit kata yang lebih kecil sehingga setiap kata memiliki makna yang berbeda. Anggap kalimat “Pilkada Jakarta 2024” akan ditokenisasi menjadi kata “Pilkada”, “Jakarta”, dan “2024”.

2. Proses *Filtering*.

Proses *Filtering* merupakan proses penghapusan kata yang tidak memiliki nilai informasi seperti karakter-karakter non-alfabet lain meliputi angka, tanda baca, atau simbol. Daftar kata umum atau *stopwords* seperti kata “atau”, “dan”, atau “yang” akan turut dihapus pada proses ini.

3. Proses *Stemming*

Proses *Stemming* merupakan proses perubahan kata menjadi bentuk dasarnya mengikuti kaidah KBBI. Sebagai contoh, kata “Pemilihan” diubah menjadi kata dasar “Pilih”. Oleh karena itu, proses dapat dikatakan juga sebagai lanjutan dari proses *Filtering*.

4. Proses Ekstraksi Fitur

Metode TF-IDF atau *Term Frequency-Inverse Document Frequency* akan digunakan dalam proses ekstraksi fitur untuk menampilkan teks dalam bentuk numerik berdasarkan bobot kepentingan kata dalam korpus. Pada akhirnya, nilai tersebut digunakan dalam menentukan relevansi suatu kata terdapat prediksi sentimen.

Setelah *pre-processing* dilakukan, metode *Elbow* akan diterapkan guna menentukan jumlah optimal kluster dengan menghitung varians intra-kluster atau inersia pada berbagai nilai k . Inersia mengukur seberapa dekat titik data dalam suatu kluster. Nilai inersia menunjukkan kepadatan kluster. Apabila nilai inersia semakin kecil, maka anggota-anggota kluster memiliki kesamaan yang lebih tinggi dengan jarak antar titik-titik data lebih kecil. Nilai k yang optimal ditemukan pada titik *elbow*, yaitu titik di mana penurunan inersia mulai melambat sekaligus menunjukkan jumlah kluster terbaik untuk menggambarkan data secara efektif. *Silhouette Score* juga dilakukan sebagai evaluasi tambahan untuk mengukur seberapa baik artikel-artikel dikelompokkan dalam kluster dengan rentang nilai dari -1 sampai 1. Semakin tinggi *Silhouette Score* menunjukkan pengelompokan yang lebih baik.

Mengikuti perolehan nilai k optimal, algoritma *K-Means* dijalankan untuk pengelompokan data. Hasil pengelompokan dianalisis lebih lanjut dengan mengidentifikasi kata kunci atau *topwords* dalam setiap kluster berdasarkan nilai TF-IDF tertinggi. Penggunaan alat visualisasi *Word Cloud* turut digunakan untuk menunjang visualisasi kluster yang lebih intuitif. Kata-kata yang memiliki frekuensi muncul lebih banyak akan ditampilkan dalam ukuran yang lebih besar.

2.3 Analisis Sentimen dengan Lexicon

Analisis sentimen diterapkan untuk memahami distribusi nada pemberitaan dari setiap kluster dengan pendekatan Lexicon dan didukung oleh *TextBlob* untuk memudahkan analisis teks dalam berbagai aplikasi, seperti klasifikasi sentimen

atau ekstraksi informasi. Nilai sentimen atau *sentiment score* dihitung berdasarkan nilai polaritas atau *polarity score* dengan cara menjumlahkan keseluruhan bobot kata yang terdeteksi sistem sehingga dapat diklasifikasikan ke dalam jenis sentimen berikut ini.

Jika nilai polaritas > 0	termasuk dalam sentimen positif.
Jika nilai polaritas = 0	termasuk dalam sentimen netral.
Jika nilai polaritas < 0	termasuk dalam sentimen negatif.

Klasifikasi sentimen ini memberikan pemahaman persebaran sentimen untuk masing-masing tema sekaligus perkiraan nada pemberitaan dalam berbagai topik yang muncul dari hasil *clustering*. Melalui penelitian *Sentiment Analysis Using Lexicon-Based Methods for News Data*, Medhat et. al. (2022) beranggapan pendekatan berbasis Lexicon sangat efektif untuk menganalisis sentimen teks dalam skala besar karena sederhana, kredibel, dan tidak memerlukan pelatihan data. Tidak berbeda dengan Thelwall (2023) dalam penelitian *Lexicon-Based Sentiment Analysis for Big Data Applications* menekankan keunggulan metode Lexicon dalam analisis sentimen teks yang bersifat informal, salah satunya adalah artikel berita.

3. HASIL DAN PEMBAHASAN

3.1 Web Scraping

Proses *web scraping* pada penelitian ini mengumpulkan 100 artikel berita terkait Pilkada Jakarta 2024. Tahap pengumpulan tautan menggunakan RSS *feed* Google News yang menghasilkan daftar tautan berita yang relevan. Kumpulan tautan tersebut disimpan dalam file "links.txt". Selanjutnya, pada tahap pengunduhan konten, digunakan *library* Python seperti "BeautifulSoup4" untuk mengekstrak isi artikel dari masing-masing tautan. Data tersebut kemudian digunakan untuk menganalisis lebih lanjut, termasuk *clustering* dan analisis sentimen.

3.2 Preprocessing Data

Tahapan *preprocessing* teks melibatkan tiga langkah utama: *tokenizing*, *filtering*, dan *stemming*. *Tokenizing* adalah proses memecah kalimat menjadi unit kata-kata kecil agar setiap kata dapat memiliki makna yang berbeda, misalnya kalimat "PAN membuka peluang untuk mendukung Anies Baswedan di Pilgub Jakarta 2024" menjadi ['PAN', 'membuka', 'peluang', 'untuk', 'mendukung', 'Anies', 'Baswedan', 'di', 'Pilgub', 'Jakarta', '2024']. *Filtering* dilakukan untuk menghapus elemen yang tidak relevan seperti angka, tanda baca, dan kata-kata umum (*stopwords*) sehingga hasilnya menjadi ['PAN', 'membuka', 'peluang', 'mendukung', 'Anies', 'Baswedan', 'Pilgub', 'Jakarta']. *Stemming* adalah tahap akhir di mana kata-kata diubah ke bentuk dasar, contohnya "membuka" menjadi "buka" dan "mendukung" menjadi "dukung", menghasilkan ['pan', 'buka', 'peluang', 'dukung', 'anies', 'baswedan', 'pilgub', 'jakarta']. Proses ini menyederhanakan teks untuk memudahkan analisis lebih lanjut.

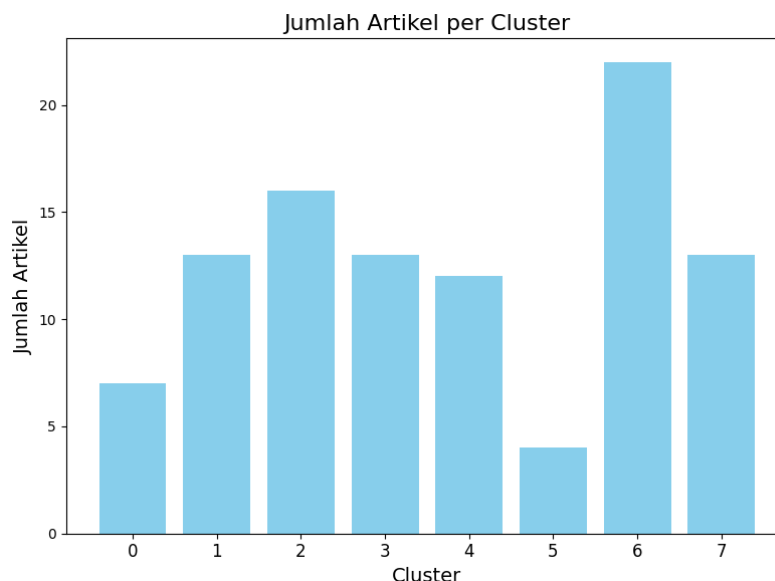
3.3 Hasil Analisis Clustering

Berikan pernyataan bahwa apa yang diharapkan, seperti yang dinyatakan pada bab "Pendahuluan" pada akhirnya dapat menghasilkan bab "Hasil dan Analisis ", sehingga ada kesesuaian. Selain itu juga dapat ditambahkan prospek pengembangan hasil penelitian dan penerapan prospek studi lanjut ke selanjutnya (berdasarkan Penentuan jumlah kluster dilakukan dengan menggunakan *Silhouette Score* yang dapat mengukur seberapa baik setiap titik data dikelompokkan, dengan nilai lebih tinggi menunjukkan pemisahan kluster yang lebih baik. Berdasarkan tabel 4 jumlah kluster dari 1 hingga 10 yang menunjukkan optimal adalah 8 kluster dengan nilai *Silhouette Score* tertinggi yaitu 0.01612.

Tabel 1. Nilai *Silhouette Score* Jumlah Kluster

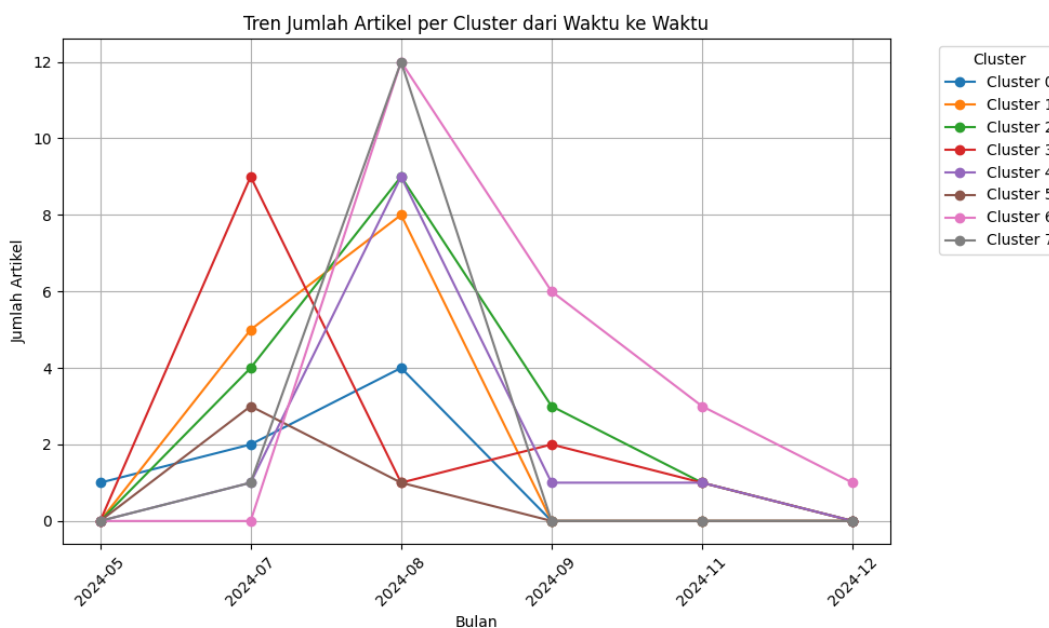
Jumlah Cluster	Silhouette Score
2 Cluster	0.00602
3 Cluster	0.00134
4 Cluster	0.01026
5 Cluster	0.01234
6 Cluster	0.01206
7 Cluster	0.01305
8 Cluster	0.01612
9 Cluster	0.01611
10 Cluster	0.01348

Dari jumlah 8 kluster tersebut, kluster dengan jumlah artikel terbanyak adalah kluster 6 dengan 22 artikel diikuti dengan kluster ke 2 yaitu 16 artikel. Sementara itu jumlah artikel sedikit ada pada kluster 0 sebanyak 7 artikel dan kluster ke 5 sebanyak 4 artikel. Kemudian untuk kluster 1, 3 dan 4 memiliki jumlah artikel yang sama yaitu sebanyak 13 artikel.



Gambar 1. Histogram Jumlah Artikel per Kluster

Pada gambar 2 menunjukkan tren jumlah artikel yang dikelompokkan berdasarkan kluster (Kluster 0 hingga 7) dari waktu ke waktu, dengan sumbu horizontal mewakili bulan dari Mei hingga Desember 2024, dan sumbu vertikal menunjukkan jumlah artikel.



Gambar 2. Grafik Tren Jumlah Artikel per Kluster dari Waktu ke Waktu

Terlihat bahwa jumlah artikel untuk hampir semua kluster mencapai puncaknya pada bulan Agustus 2024, sebelum mengalami penurunan tajam pada bulan-bulan berikutnya. Karena pada bulan Agustus menjadi bulan dimana isu-isu dalam Pilkada Jakarta 2024 mulai ramai dibicarakan mulai terkait pencalonan, permasalahan antar pencalonan dan partai politik. Hasil analisis menunjukkan bahwa setiap kluster memiliki topik pembahasan spesifik terkait Pilkada Jakarta 2024, yang ditentukan dengan mengidentifikasi kata kunci sebagai kata yang paling sering muncul dalam artikel di setiap kluster.

Kluster 0, misalnya, berfokus pada dukungan dan kandidat Pilkada dengan kata kunci seperti "dukungan kandidat" dan "pilkada," yang mencerminkan pembahasan tentang kandidat dari partai politik seperti PDIP dan PKB, serta kandidat independen.



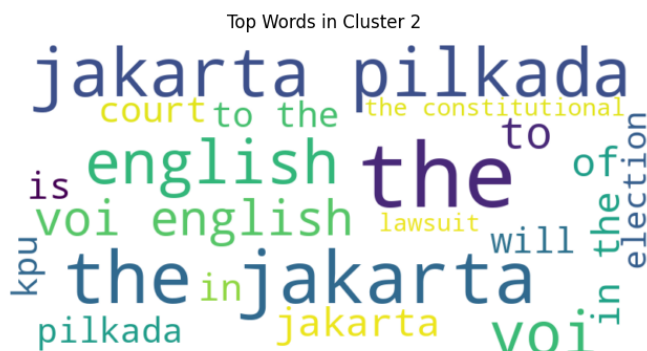
Gambar 3. *Word Clouds* dalam Kluster 0

Kluster 1 menonjol dengan kata-kata seperti "nasdem" dan "anies," yang menunjukkan fokus pada pencalonan Anies Baswedan yang secara resmi didukung Partai Nasdem.



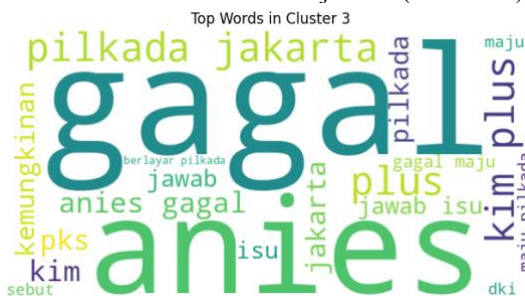
Gambar 4. *Word Clouds* dalam Kluster 1

Kluster 2, dengan kata-kata seperti "English" dan "voi," mengelompokkan artikel berbahasa Inggris dari media VOI yang terkait dengan Pilkada Jakarta.



Gambar 5. *Word Clouds* dalam Kluster 2

Sementara itu, kluster 3 membahas kegagalan Anies Baswedan dalam pencalonan yang tercermin dari kata "gagal" dan "kim" yang merujuk pada dinamika politik terkait Koalisi Indonesia Maju Plus (KIM Plus).



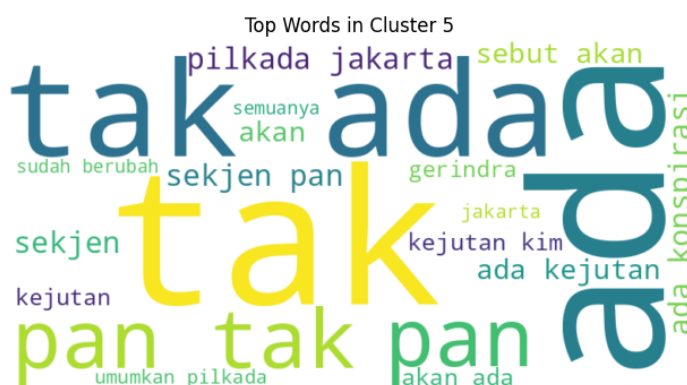
Gambar 6. *Word Clouds* dalam Kluster 3

Kluster 4 dengan kata kunci "anies" dan "jokowi" mengangkat isu dukungan politik Anies, ambang batas pencalonan, dan peran Jokowi dalam Pilgub DKI Jakarta dan Jawa Tengah.



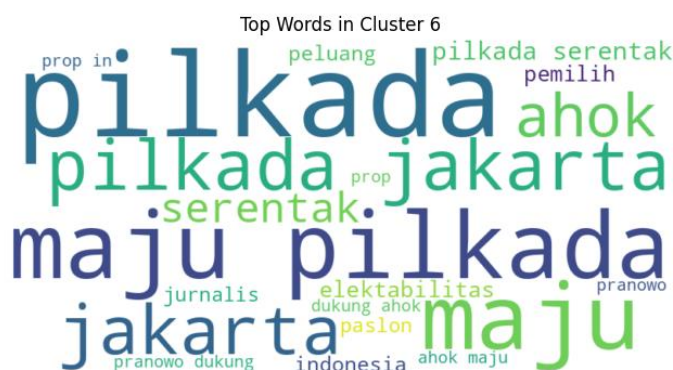
Gambar 7. *Word Clouds* dalam Kluster 4

Kluster 5 menyoroti isu ketidakpastian dan teori konspirasi dalam pilgub yang diwakili oleh kata "tak ada" dan "konspirasi" yang menegaskan klarifikasi PAN bahwa tidak ada upaya untuk menggagalkan Anies maupun Koalisi Indonesia Maju (KIM).



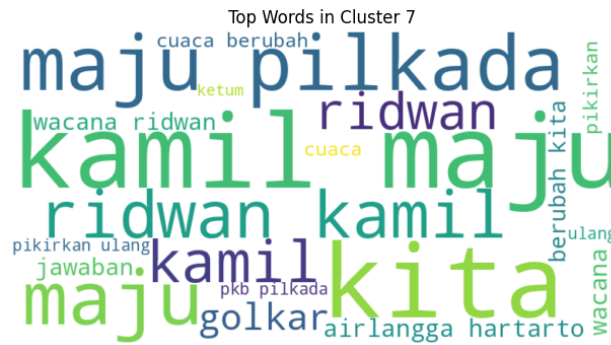
Gambar 8. *Word Clouds* dalam Kluster 5

Selanjutnya, Kluster 6 membahas Pilkada Serentak dan dinamika politik di Jakarta, dengan kata-kata seperti "pilkada" dan "serentak," yang mencakup isu persiapan kandidat, peluang bagi kandidat seperti Anies, serta dampak Pilkada terhadap ekonomi Jakarta.



Gambar 9. *Word Clouds* dalam Kluster 6

Terakhir, kluster 7 menyoroti Ridwan Kamil sebagai kandidat potensial dengan dukungan kuat dari Golkar dan Koalisi Indonesia Maju (KIM), tercermin dari kata-kata seperti "ridwan kamil" dan "golkar," serta rencana pengumuman pasangannya.



Gambar 10. *Word Clouds* dalam Kluster 7

Tren pembahasan artikel mencapai puncaknya pada Agustus 2024, dengan topik yang hangat diperbincangkan meliputi isu pencalonan, konflik politik, dan dukungan partai. Hal ini mencerminkan tingginya tingkat perhatian publik terhadap pemilihan gubernur Jakarta, sebagaimana terlihat dari pola kata yang dominan di setiap kluster.

3.4 Hasil Analisis Sentimen tiap Kluster

Hasil *clustering* yang terdiri dari delapan *cluster* berdasarkan topik utama kemudian dinilai apakah kumpulan berita tersebut bernilai positif, negatif, atau netral menggunakan pendekatan Lexicon. Mayoritas artikel dalam tiap *cluster* memiliki sentimen netral yang menunjukkan bahwa media massa umumnya menjaga objektivitas. Akan tetapi, pada *cluster* 2 (Media berita VOI dengan bahasa Inggris) terdapat indikasi sentimen positif sebanyak 5 artikel, negatif sebanyak 1 artikel, dan netral sebanyak 10 artikel. Hal ini bisa disebabkan oleh penggunaan kata atau bahasa tertentu dalam artikel.

Tabel 2. Hasil Analisis Sentimen tiap Cluster

Kluster	Topik Kluster	Netral	Positif	Negatif
0	Dukungan dan Kandidat Pilkada	7	0	0
1	Partai Nasdem mengusung Anies untuk Pilkada Jakarta	13	0	0
2	Media berita VOI dengan bahasa Inggris mengenai Pilkada Jakarta	10	5	1
3	Kegagalan Anies dalam pencalonan Pilkada Jakarta	13	0	0
4	Anies Baswedan dan Dukungan Politik	12	0	0
5	Tak Ada Dukungan dan Konspirasi dalam Pilkada	4	0	0
6	Pemilihan Serentak dan Dukungan Pilkada Jakarta 2024	22	0	0
7	Ridwan Kamil dan Koalisi Politik dalam Pilkada Jakarta 2024	13	0	0

Hasil tersebut menunjukkan bahwa sebagian besar artikel bersifat netral sesuai dengan peran media sebagai penyampai informasi yang objektif. Akan tetapi, indikasi bias sentimen pada salah satu *cluster* dapat digarisbawahi untuk analisis lebih lanjut agar mengetahui faktor-faktor yang memengaruhi bias tersebut.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa melalui algoritma *K-Means*, pola pemberitaan media dari 100 artikel mengenai Pilkada Jakarta 2024 dapat dikelompokkan menjadi delapan tema utama dengan fokus pada isu dukungan politik, calon kandidat, hingga dinamika politik yang terjadi selama Pilkada Jakarta 2024 berlangsung. Sebagian besar artikel-artikel di setiap kluster menunjukkan distribusi sentimen yang bersifat netral, meskipun ditemukan adanya sentimen positif dan negatif yang mengindikasikan adanya bias dalam artikel-artikel pada kluster 2 mengenai topik Media Berita VOI dengan bahasa Inggris mengenai Pilkada Jakarta.

Pendekatan *text mining* yang melibatkan analisis *clustering* dengan algoritma *K-Means* dan analisis sentimen berbasis Lexicon terbukti efektif dalam mengidentifikasi pola pemberitaan dan sentimen media massa terkait Pilkada Jakarta 2024 sekaligus memberikan pemahaman yang lebih mendalam tentang perkembangan dinamika politik. Hasil ini diharapkan dapat menjadi acuan bagi pemerintah, media, dan masyarakat dalam meningkatkan pemahaman tentang persepsi publik, menjaga netralitas pemberitaan, dan mendukung proses demokrasi yang berkualitas.

REFERENSI

- [1] F. V. Sari dan A. Wibowo, "ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 10, no. 2, hlm. 681–686, Nov 2019, Diakses: 26 Desember 2024. [Daring]. Tersedia pada: <https://jurnal.umk.ac.id/index.php/simet/article/view/3487>
- [2] A. Goyal, V. Gupta, dan M. Kumar, "Recent Named Entity Recognition and Classification techniques: A systematic review," *Comput Sci Rev*, vol. 29, hlm. 21–43, Agu 2018, doi: 10.1016/J.COSREV.2018.06.001.

- [3] T. ALASALI dan Y. ORTAKCI, "Clustering Techniques in Data Mining: A Survey of Methods, Challenges, and Applications," *Computer Science*, Mar 2024, doi: 10.53070/BBD.1421527.
- [4] D. M. Thomas dan S. Mathur, "Data Analysis by Web Scraping using Python," *Proceedings of the 3rd International Conference on Electronics and Communication and Aerospace Technology, ICECA 2019*, hlm. 450–454, Jun 2019, doi: 10.1109/ICECA.2019.8822022.
- [5] N. Syamsu Wais Al Qorni, "PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA ULASAN PENGGUNA MERDEKA MENGAJAR DI PLAY STORE PROGRAM STUDI TEKNIK INFORMATIKA SEKOLAH TINGGI MANAJEMEN INFORMATIKA dan KOMPUTER WIDYA CIPTA DHARMA 2024".
- [6] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, dan M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, hlm. 267–307, 2011, doi: 10.1162/COLI_A_00049.
- [7] "View of IMPLEMENTASI ALGORITMA K-MEANS TERHADAP OPINI MASYARAKAT MENGENAI PERKIRAAN PEMILU 2024 PADA TWITTER." Diakses: 26 Desember 2024. [Daring]. Tersedia pada: <https://ejournal.caturasaki.ac.id/index.php/simtek/article/view/271/233>
- [8] A. Rahmawati, A. Marjuni, J. Zeniarja, J. Informatika, U. Dian, dan N. Semarang, "ANALISIS SENTIMEN PUBLIK PADA MEDIA SOSIAL TWITTER TERHADAP PELAKSANAAN PILKADA SERENTAK MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE PUBLIC SENTIMENT ANALYSIS ON TWITTER SOCIAL MEDIA TO PILKADA SERENTAK EVENT USING SUPPORT VECTOR MACHINE ALGORITHM," vol. 10, no. 2, 2017.
- [9] "Kampanye Tiga Pasangan Calon Pilkada Jakarta 2024: Masalah Krusial Warga Tak Terjamah, Minim Solusi Nyata | ICW." Diakses: 26 Desember 2024. [Daring]. Tersedia pada: <https://antikorupsi.org/id/kampanye-tiga-pasangan-calon-pilkada-jakarta-2024-masalah-krusial-warga-tak-terjamah-minim-solusi>
- [10] O. R. Romli dan E. Nurlia, "Menimbang Isu Politik Identitas Pada Pilkada Dki Jakarta Tahun 2024," *Jurnal Ilmu Politik dan Pemerintahan*, vol. 9, no. 2, Des 2023, doi: 10.37058/jipp.v9i2.8944.
- [11] Oleh, "PROGRAM STUDI SARJANA ILMU HUKUM FAKULTAS HUKUM UNIVERSITAS HASANUDDIN MAKASSAR 2024 SKRIPSI TINJAUAN YURIDIS TERHADAP PENGATURAN HUKUM PILKADA SERENTAK TAHUN 2024 DAN IMPLIKASINYA JUDICIAL REVIEW OF LEGAL REGULATIONS OF SIMULTANEOUS REGIONAL ELECTIONS IN 2024 AND THE IMPLICATIONS".
- [12] C. Rahim, B. Niode, A. B. Pati, D. M. Liando, dan M. Mamentu, "NETRALITAS MEDIA MASSA BERBASIS ONLINE PADA PEMILIHAN WALI KOTA DAN WAKIL WALI KOTA MANADO TAHUN 2020," *Jurnal Review Pendidikan dan Pengajaran (JRPP)*, vol. 7, no. 1, hlm. 776–792, Jan 2024, doi: 10.31004/JRPP.V7I1.24692.
- [13] "MEDIA MONITORING ISU PILKADA PROVINSI DKI JAKARTA TAHUN 2024".
- [14] J. Homepage *dkk.*, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, hlm. 24–33, Mar 2021, doi: 10.57152/MALCOM.V1I1.20.
- [15] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, dan W. Gata, "ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI," *Jurnal Teknoinfo*, vol. 14, no. 2, hlm. 115, Jul 2020, doi: 10.33365/JTI.V14I2.679.
- [16] "(PDF) McQuail's Mass Communication Theory | Khalsa Alrahbi - Academia.edu." Diakses: 26 Desember 2024. [Daring]. Tersedia pada: https://www.academia.edu/36441695/McQuails_Mass_Communication_Theory
- [17] R. P. Bastoni, "Analisis Sentimen Opini Publik Terhadap Calon Presiden Indonesia 2024 Menggunakan Metode Naive Bayes Dan Lexicon Based Pada Platform YouTube Skripsi Oleh".
- [18] E. Susanto, V. C. Mawardi, dan M. D. Lauro, "APLIKASI CLUSTERING BERITA DENGAN METODE K MEANS DAN PERINGKAS BERITA DENGAN METODE MAXIMUM MARGINAL RELEVANCE".
- [19] "View of PENCARIAN DATA QUICK COUNT PILPRES DENGAN TEKNIK WEB SCRAPING." Diakses: 26 Desember 2024. [Daring]. Tersedia pada: <https://bajangjournal.com/index.php/JIRK/article/view/6695/5147>
- [20] "ANALISIS SENTIMEN UNTUK MEMPREDIKSI HASIL CALON PEMILU PRESIDEN MENGGUNAKAN LEXICON BASED DAN RANDOM FOREST." Diakses: 26 Desember 2024. [Daring]. Tersedia pada: <https://ejournal.upbatam.ac.id/index.php/jif/article/view/7987/3319>

Studi Literatur Perancangan Arsitektur Data dan Aplikasi pada Perusahaan Telekomunikasi

Jason¹, Jefferson Sutanto², Verrel Angkasa³, Vicky Darmana⁴, Ade Maulana⁵

Sistem Informasi, Universitas Pelita Harapan, Medan, Indonesia

Email: ¹03081220018@student.uph.edu, ²03081220030@student.uph.edu, ³03081220031@student.uph.edu,

⁴03081220038@student.uph.edu, ⁵ade.maulana@lecturer.uph.edu

ABSTRAK

Arsitektur Perusahaan merupakan sebuah perancangan proses bisnis dan teknologi yang terdapat pada sebuah organisasi yang kemudian akan diintegrasikan untuk mencapai tujuan dari organisasi tersebut. Pada dasarnya, arsitektur perusahaan adalah konsep-konsep dasar yang bertujuan untuk membantu sebuah organisasi untuk melakukan integrasi terhadap setiap unit-unit proses bisnis organisasi agar menjadi lebih efektif sehingga mendapatkan hasil yang maksimal. Dalam sebuah organisasi, alangkah baiknya ada arsitektur yang baik agar dapat mencapai kesuksesan bisnis. Penelitian ini akan membahas analisis penggunaan model perancangan Arsitektur Data dan aplikasi. Tidak dapat dihindari bahwa perusahaan telekomunikasi juga membutuhkan arsitektur perusahaan untuk mengolah proses-proses bisnis dan data-data perusahaan agar dapat saling terintegrasi dan selaras dengan misi dari perusahaan tersebut. Metode penelitian yang digunakan adalah studi literatur dengan pendekatan kualitatif. Hasil dari penelitian ini adalah untuk mengetahui model arsitektur paling banyak digunakan dalam perancangan *Enterprise Architecture* sehingga dapat bermanfaat bagi perusahaan telekomunikasi dalam melakukan perancangan model arsitektur yang lebih baik lagi di masa yang akan datang.

Kata Kunci: Arsitektur Perusahaan, Studi Literatur, Perusahaan Telekomunikasi

ABSTRACT

Enterprise Architecture (EA) serves as a blueprint for aligning business processes and IT systems to achieve organizational objectives. It provides a structured framework for integrating and optimizing various business units, enhancing overall effectiveness and maximizing outcomes. A well-defined EA is crucial for achieving business success. This research delves into the analysis of data and application modeling approaches in EA frameworks. Telecommunication companies, like other organizations, rely on EA to integrate and align their business processes and data with their corporate mission. Employing a qualitative literature review methodology, this study aims to identify the most prevalent data and application modeling approaches in EA frameworks, providing valuable insights for telecommunication companies seeking to refine their EA models in the future.

Keywords: Enterprise Architecture, Literature Review, Telecommunication Company

Penulis Korespondensi:

Verrel Angkasa

Email: 03081220031@student.uph.edu

Article Info

Diterima: 26 Juli 2024

Direvisi: 26 Januari 2025

Disetujui: 12 Februari 2025

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.



1. PENDAHULUAN

Pada zaman sekarang, era digitalisasi berkembang semakin pesat. Oleh karena itu, para perusahaan atau organisasi tertentu perlu mengadakan perubahan ke arah yang lebih baik. Dengan menerapkan konsep *Enterprise Architecture*, bisnis perusahaan dapat berjalan, digambarkan serta didefinisikan dengan baik sehingga dapat membuat perencanaan yang strategis dan dalam perencanaan yang ditetapkan [1]. *Enterprise Architecture* juga dapat memberikan manfaat yang baik kepada perusahaan seperti peningkatan kemampuan bisnis perusahaan, pengubahan visi misi perusahaan untuk mencapai tujuan tertentu dalam perusahaan dan masih banyak lagi [2]. *Enterprise Architecture* (EA) merupakan sebuah aspek penting dalam mengintegrasikan sistem informasi di sebuah perusahaan, EA adalah sebuah kerangka konseptual yang mendeskripsikan bagaimana perusahaan tersebut dibangun dengan mendefinisikan komponen-komponen utama dan hubungan antar komponen-komponen [3]. Konsep EA pertama kali diperkenalkan oleh seorang pengusaha bisnis asal Amerika Serikat bernama John Zachman pada tahun 1987 dengan sebuah manuskrip yang berjudul '*A Framework for Information Systems Architecture*' [4].

Terdapat banyak sekali model atau kerangka kerja dari *Enterprise Architecture*. Setiap model dan kerangka kerja itu sendiri memiliki kelebihan dan kekurangannya masing-masing, maka diperlukan pemahaman yang mendalam agar perusahaan dapat memilih metode yang paling tepat. Diantara semua kerangka kerja yang ada di *Enterprise Architecture*, TOGAF (*The Open Group Architecture Framework*) adalah salah satu yang paling populer di kalangan perusahaan [5]. TOGAF digunakan untuk membentuk arsitektur melalui deretan berikut: pengenalan, visi perusahaan, arsitektur bisnis, arsitektur sistem informasi, dan arsitektur teknologi, terdapat TOGAF 9.2 yang secara keseluruhan melengkapi versi sebelumnya [6]. Sementara TOGAF itu sendiri menganggap setiap tahapan dimulai dari visi dan misi perusahaan, hingga metode implementasi dan tata kelola, serta tahapan untuk ditinjau kembali apakah perubahan tersebut perlu ditingkatkan lebih lanjut atau tidak. Pada penelitian ini, jurnal yang menggunakan *framework* TOGAF akan dibahas dan tidak tertinggal dengan jurnal yang menggunakan *framework* lain seperti *Enterprise Architecture Planning* [7].

Satu penelitian yang membahas tentang sebuah desain *Enterprise Architecture* menggunakan kerangka kerja TOGAF ADM 9.2 untuk PT. Industri Telekomunikasi Indonesia (INTI). Kegunaan dari IT di PT. Industri Telekomunikasi Indonesia telah memainkan peran penting dalam mendukung tujuan dari suatu perusahaan serta visi dan misi yang mampu menghasilkan sebuah sinergi dan membangun integritas bisnis. Dengan adanya *Enterprise Architecture*, mampu membantu dalam mendeskripsikan perusahaan dan mendeskripsikan status perusahaan untuk memajukan kondisi terkini dari perusahaan ke arah yang lebih baik serta mendukung manajemen berbasis teknologi informasi terkait pengembangan kinerja perusahaan terintegrasi di seluruh sumber daya teknologi, aliran informasi, proses bisnis, dan orientasi strategis [8]. Pada akhirnya, penelitian ini bertujuan untuk membantu perancangan *Enterprise Architecture* bagi perusahaan telekomunikasi khususnya dalam bagian Arsitektur Data dan Arsitektur Aplikasi.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah studi literatur dengan pendekatan kualitatif. Metode ini melibatkan pengumpulan artikel jurnal yang relevan untuk mengidentifikasi referensi akurat terkait topik penelitian. Setelah memperoleh artikel jurnal yang jelas dan akurat, peneliti menganalisisnya dan menarik kesimpulan berdasarkan temuan-temuan tersebut [9].

Studi literatur merupakan pondasi penelitian di berbagai disiplin ilmu. Metode ini berperan sebagai landasan bagi peneliti, menyediakan pengetahuan dasar yang penting untuk menghasilkan karya inovatif. Dengan mendalami penelitian sebelumnya, peneliti dapat menghindari redundansi dan membangun pengetahuan yang ada. Namun, manfaatnya tidak hanya sebatas akumulasi pengetahuan. Studi literatur juga dapat memicu penemuan-penemuan revolusioner dengan menantang pemahaman saat ini dan mendorong perspektif baru [10].

Studi literatur telah memperoleh perhatian yang signifikan dalam komunitas penelitian, terutama dalam studi yang mengeksplorasi Arsitektur Perusahaan (*Enterprise Architecture/EA*) dalam organisasi. Pelaksanaan yang cermat pada setiap tahap studi literatur sangat penting untuk menghasilkan hasil penelitian berkualitas tinggi. Penelitian ini menggali tahap-tahap kunci yang terlibat dalam pelaksanaan studi literatur [11]:

2.1. Penentuan Ruang Lingkup

Penelitian ini menggunakan sumber artikel jurnal yang diterbitkan antara tahun 2019 hingga 2024 sebagai referensi. Pembatasan rentang waktu literatur ini dimaksudkan untuk memastikan relevansi kajian literatur dengan perancangan *Enterprise Architecture* bagi perusahaan telekomunikasi.

2.2. Pengumpulan Data

Enterprise Architecture (EA) telah menjadi komponen integral dalam berbagai aktivitas operasional lintas proses bisnis organisasi. Dalam pencarian sumber artikel jurnal, penggunaan kata kunci tersebut memudahkan identifikasi literatur relevan terkait topik penelitian yang sedang dikaji. Berikut kriteria inklusi dan eksklusi dalam pencarian jurnal artikel dapat dijabarkan sebagai berikut:

Tabel 1. Kriteria Inklusi dan Eksklusi

Kriteria Inklusi	Kriteria Eksklusi
Artikel dalam bahasa Indonesia atau Inggris	Abstrak, editorial, opini, dan surat pembaca tidak dimasukkan.
Artikel yang secara eksplisit membahas <i>Enterprise Architecture</i> (EA), <i>TOGAF</i> (<i>The Open Group Architecture Framework</i>) ADM, dan telekomunikasi.	Artikel yang tidak relevan dengan topik EA, <i>TOGAF</i> ADM, dan telekomunikasi tidak dimasukkan.
Artikel yang menggunakan metodologi penelitian yang kredibel, seperti studi kasus, penelitian kuantitatif, atau penelitian kualitatif.	Artikel yang menggunakan metodologi penelitian yang tidak kredibel tidak dimasukkan.
Artikel yang diterbitkan dalam jurnal ilmiah ternama dengan proses <i>peer-review</i> yang ketat	Artikel yang diterbitkan dalam jurnal ilmiah yang tidak ternama atau tidak memiliki proses <i>peer-review</i> yang ketat tidak dimasukkan.
Artikel yang dapat diakses secara penuh secara <i>online</i>	Artikel yang tidak dapat diakses secara penuh secara <i>online</i>

2.3. Analisis Data

Setelah mengumpulkan referensi yang diperlukan untuk penelitian ini, artikel jurnal yang diperoleh disimpan dan diorganisir menggunakan *reference manager*. *Reference manager* memudahkan penyimpanan sumber referensi untuk keperluan penelitian dan memperlancar proses sitasi, sehingga meningkatkan efisiensi. Mengingat fokus penelitian pada *Enterprise Architecture* di sektor telekomunikasi, sebagian besar artikel jurnal yang digunakan sebagai referensi membahas analisis dan perancangan *Enterprise Architecture* pada perusahaan telekomunikasi.

2.4. Pembahasan

Bagian selanjutnya adalah pembahasan yang di mana poin-poin pokok pembahasan penulisan studi literatur akan dijabarkan dalam bagian pembahasan setelah melakukan penarikan data dari sumber-sumber artikel jurnal yang telah ditelaah.

2.5. Kesimpulan

Selanjutnya, bagian terakhir dari studi literatur adalah melakukan penarikan kesimpulan. Data, informasi, dan hasil-hasil dari studi literatur akan dituangkan ke dalam bagian kesimpulan yang berisi saran untuk pembaca mengenai hal-hal yang dapat dikembangkan kedepannya dan hasil penelitian studi literatur yang telah dilakukan.



Gambar 1. Alur Penelitian

3. HASIL DAN PEMBAHASAN

Pada bagian ini, akan dijelaskan dengan rinci dan komprehensif mengenai hasil-hasil dari penarikan data yang telah dilakukan dari sumber jurnal artikel yang telah ditelaah.

3.1. Penelitian Terdahulu

Pada penelitian yang membahas tentang analisis dan perancangan *Enterprise Architecture* pada PT. XYZ di bagian divisi operasional layanan menemukan permasalahan bahwa setelah menerapkan EA masih terdapat beberapa masalah, seperti proses bisnis yang berjalan masih tidak optimal, kemampuan sumber daya manusia yang masih kurang mumpuni, dan tidak terintegrasinya sistem-sistem yang ada di perusahaan. Maka dari itu, perancangan EA yang baru dilakukan untuk memperbaiki masalah-masalah yang ada, salah satunya di arsitektur sistem informasi perusahaan. Pada Arsitektur Data dihasilkan artefak EA yang dapat membantu mengelola data perusahaan yaitu *Data Dissemination Diagram Existing* dan *Data Dissemination Diagram Targeting*. Sedangkan pada Arsitektur Aplikasi yang dirancang, penelitian tersebut menggunakan *Application Portfolio Catalog* yang berguna untuk mendeskripsikan aplikasi apa yang cocok untuk digunakan dalam mengolah data dalam perusahaan tersebut dan *Application Communication Diagram* yang dimana dapat membantu perusahaan dalam menggambarkan bagaimana hubungan interaksi antar aplikasi yang akan digunakan dalam kegiatan operasionalnya [12].

Dalam penelitian [13], diungkapkan bahwa sistem yang digunakan oleh PT. Bestari Kalimantan untuk mengolah data-data perusahaan masih sama sekali belum terkomputerisasi dan masih menggunakan cara manual sehingga ada perancangan EA yang telah dibuat dalam penelitian tersebut. Arsitektur Data dan Aplikasi yang dirancang memunculkan kebutuhan untuk menggunakan *value chain* sebagai pemodelan bisnis dan berhasil mendapatkan 9 entitas bisnis. Entitas bisnis ini kemudian digunakan sebagai representasi dari alur sistem. Tabel matrik entitas bisnis dan data membahas *jobdesk* spesifik dari masing-masing entitas, seperti contohnya Entitas Pengelolaan SDM, yang memiliki *jobdesk* perekrutan karyawan, training karyawan, dan lain lain. Ada juga kolom atribut sebagai penjelas dari tabel ini, yaitu seperti contohnya *id_pelamar*, *nama_pelamar*, dan *umur_pelamar*. Selain itu, berdasarkan ke-9 entitas tersebut, dilaksanakan sebuah perancangan Arsitektur Aplikasi sebagai solusi untuk pemrosesan data. Rancangan Arsitektur Aplikasi ini adalah penentuan jenis aplikasi seperti apa yang cocok untuk pemrosesan data. Adanya evaluasi terhadap ke-9 sistem ini dan rancangan implementasi target teknologi baru sebagai upaya dan solusi dalam menjalankan proses bisnis yang lebih efektif dan efisien. Seperti contohnya, untuk mendefinisikan pengelolaan SDM, diperlukan aplikasi rekrutmen karyawan dan aplikasi absensi karyawan. Aplikasi rekrutmen karyawan diperlukan untuk menggantikan SDM dan aplikasi absensi karyawan digunakan untuk pendataan SDM. Penelitian ini merencanakan usulan perancangan implementasi yang memakan waktu lebih kurang 5 tahun, terutama pada aplikasi yang berbasis *android* dan *web* yang dirancang untuk mempermudah proses bisnis, meningkatkan pelayanan terhadap pelanggan, dan memberikan nilai tambah bagi perusahaan.

Selanjutnya, pada penelitian yang melakukan analisis dan identifikasi data yang dibutuhkan untuk pengembangan *Enterprise Architecture*, yang kemudian digambarkan menjadi diagram konseptual menggunakan metode *Entity Relationship Diagram* (ERD). ERD adalah diagram dengan notasi grafis yang digunakan dalam pembuatan *database* untuk menghubungkan berbagai data yang dapat berfungsi sebagai alat bantu dalam merancang *database* dan memberikan gambaran tentang cara kerja *database* yang akan dibuat [14]. Dari entitas data tersebut, hubungan antara entitas data pada objek penelitian digambarkan dalam bentuk *Data Dissemination Diagram*. Adanya diagram yang lebih detail membuat pemetaan menjadi lebih terstruktur, contohnya seperti bagian *fraud*, *risk register*, dan laporan perubahan. Pada fase Arsitektur Aplikasi, dilakukan perancangan aplikasi yang *existing* dan target yang dipetakan dengan fungsi bisnis di objek penelitian. Proses ini mencakup pemetaan hubungan antara aplikasi yang *existing* dan target dengan fungsi yang ada di objek penelitian. Penulis menyimpulkan bahwa adanya perbaikan pada proses bisnis seperti, pengelolaan *fraud*, penyusunan risiko, dan pengelolaan perubahan mampu meningkatkan kualitas proses bisnis. Penelitian tersebut juga menghasilkan *blueprint Enterprise Architecture* untuk fungsi *Risk and Quality Management* sebagai solusi integratif terhadap penggunaan aplikasi yang sebelumnya terpisah-pisah, dengan tujuan meningkatkan kinerja, efektivitas pengelolaan data, dan nilai fungsi tersebut dalam mendukung perusahaan serta menyelesaikan permasalahan yang ada [15].

Kemudian, dalam penelitian [16], dilakukannya desain 2 buah model, yaitu pemodelan dari data dan aplikasi. Kedua hal ini merupakan bagian dari Arsitektur Sistem Informasi. Pemodelan data melibatkan implementasi dari beberapa entitas, seperti MariaDB, Apache Cassandra dan NoSQL Database, yang bisa memudahkan pengelolaan data. Dalam bagian pemodelan data, peneliti terdahulu mencantumkan banyak contoh implementasi dalam *data layer*, *business layer*, dan *presentation layer*. Setelah itu, *System Building Block Application Landscape Design Model* dirancang sesuai dengan porsi *jobdesk* masing-masing dari pemodelan yang sebelumnya. Model ini menghasilkan sebuah struktur pekerjaan secara keseluruhan, seperti contohnya bagian *Enterprise Resource Planning*, *Finance and Accounting Service*, dan *Help Desk*. Ada pula perancangan *Entity Relationship Diagram* yang berfungsi untuk melakukan *mapping* pada *database*. *Database* yang dirancang dapat digunakan diintegrasikan dengan *database* lainnya. Selain itu, *blueprint* ini dihasilkan sebagai rancangan sistem informasi operasional arsitektur perusahaan yang bertujuan untuk integrasi dan mempermudah proses bisnis. *Blueprint* ini juga memanfaatkan *cloud computing* agar jaminan infrastruktur, keamanan informasi, dan kecepatan akses yang dimiliki oleh penyedia layanan menjadi lebih efisien.

Dan pada penelitian [17], dilakukan analisis untuk EA agar mengikuti perkembangan zaman dikarenakan PT. Time Excelindo masih menggunakan cara manual dalam proses pengerjaannya. Oleh karena itu, dengan memanfaatkan dan mengimplementasikan *Enterprise Architecture* untuk mempunyai sistem informasi billing yang lebih canggih, cepat dan akurat serta dapat dimengerti oleh para pengguna dan memilih kerangka kerja TOGAF ADM untuk membuat *blueprint*. Pemodelan Arsitektur Data dibuat bertujuan untuk mendefinisikan tipe dan dari mana sumber data yang dibutuhkan untuk mendukung bisnis dan dapat dimengerti oleh semua pihak yang terlibat. Pada Arsitektur Data berfokus pada seperti apa data dapat digunakan untuk proses bisnis, kebutuhan bisnis, dan pelayanan. Pada pemodelan Arsitektur Aplikasi lebih menekankan pada kebutuhan aplikasi yang direncanakan dengan menggunakan *Application Portfolio Catalog* dan berfokus pada model aplikasi yang dirancang. Pada

perusahaan ini juga memanfaatkan *Value Chain* Proses Bisnis yang bertujuan untuk mengidentifikasi beberapa bagian yang terdapat pada bagian aktivitas pendukung yang ada di perusahaan dan bertujuan untuk meningkatkan nilai pada perusahaan.

3.2. Analisis Arsitektur Data

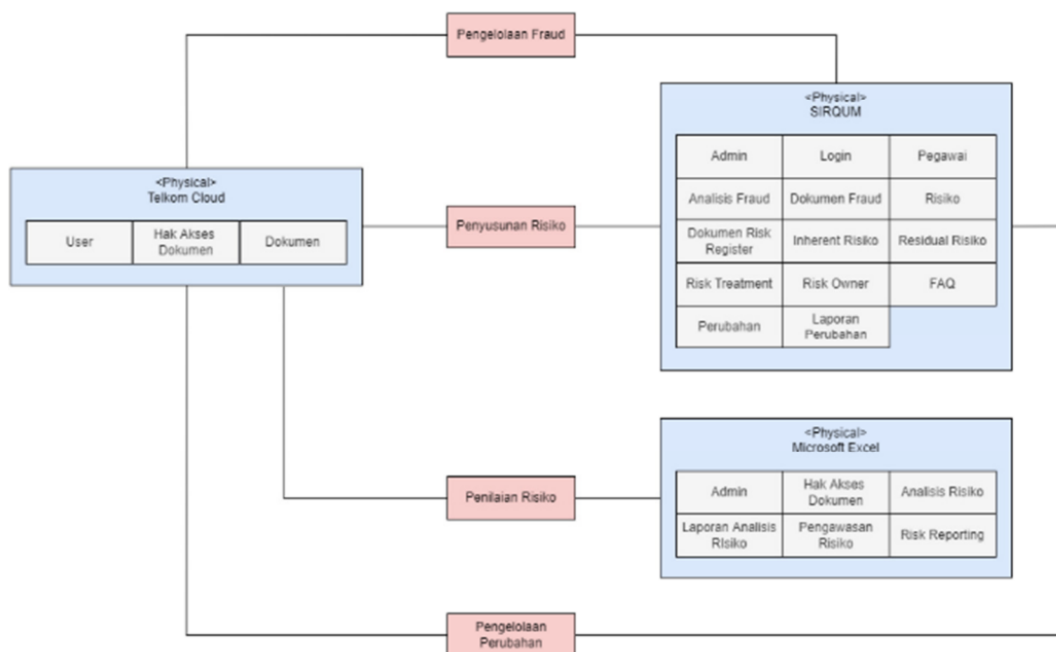
Setelah melakukan studi literatur terhadap perancangan Arsitektur Data yang terdapat pada perusahaan-perusahaan telekomunikasi, telah ditemukannya banyak jenis model yang dirancang di dalamnya. Model-model perancangan Arsitektur Data yang digunakan dapat dikelompokkan dalam Tabel 2.

Tabel 2. Model Arsitektur Data

Model Arsitektur Data	Data Dissemination Diagram	Matriks Entitas	Database	Entity Relationship Diagram (ERD)
Penelitian				
[12]	√			
[13]		√		
[15]	√			√
[16]			√	

Dalam jurnal [17], tidak terdapat model perancangan Arsitektur Data dikarenakan pada penelitian tersebut proses perancangan arsitektur masih sampai dengan tahap visi arsitektur perusahaan. Pada Tabel 2, dilakukan pengelompokkan dari perancangan model Arsitektur Data yang digunakan, tanda √ diberikan kepada jurnal yang menggunakan model Arsitektur Data, seperti *Data Dissemination Diagram*, Matriks Entitas, *Database*, *Entity Relationship Diagram (ERD)*. Dari hasil pengelompokkan, dapat dilihat pada Tabel 2, model Arsitektur Data yang paling banyak digunakan ialah *Data Dissemination Diagram*.

Alasan dibalik banyaknya penggunaan *Data Dissemination Diagram* karena diagram ini memiliki kemampuan untuk menggambarkan hubungan dari setiap komponen data-data yang digunakan dalam suatu perusahaan. Diagram ini juga direkomendasikan untuk digunakan sebagai panduan dalam melakukan model Arsitektur Data pada saat merancang sebuah *Enterprise Architecture* suatu perusahaan dikarenakan dengan kemampuannya dalam memvisualisasikan hubungan antar data sehingga dapat membantu perancangan Arsitektur Data suatu perusahaan menjadi semakin efektif karena hubungan setiap penggunaan data dijabarkan dengan jelas. Pada Gambar 2, terdapat contoh *Data Dissemination Diagram* yang dirancang dalam [15].



Gambar 2. Contoh Model *Data Dissemination Diagram*

3.3. Analisis Arsitektur Aplikasi

Bagian selanjutnya ialah Arsitektur Aplikasi. Setelah melakukan studi literatur terhadap perancangan Arsitektur Aplikasi yang terdapat pada perusahaan-perusahaan telekomunikasi, telah ditemukannya banyak jenis model yang dirancang di dalamnya. Model-model perancangan Arsitektur Aplikasi yang digunakan dapat dikelompokkan dalam Tabel 3.

Tabel 3. Model Arsitektur Aplikasi

Model Arsitektur Aplikasi	<i>Application Portfolio Catalog</i>	<i>Application Communication Diagram</i>	<i>Application/Function Matrix</i>	<i>System Building Block Application Landscape Design</i>	<i>Application Database Design</i>	<i>Gap Analysis</i>
Penelitian						
[12]	√	√				
[13]						√
[15]		√	√			
[16]				√	√	
[17]	√	√				

Pada Tabel 3, dilakukan pengelompokan terhadap penggunaan model dari perancangan Arsitektur Aplikasi, ada berbagai model yang digunakan pada jurnal penelitian tersebut namun 2 (dua) perancangan model Arsitektur Aplikasi yang paling sering digunakan di dalamnya adalah *Application Portfolio Catalog* dan *Application Communication Diagram*.

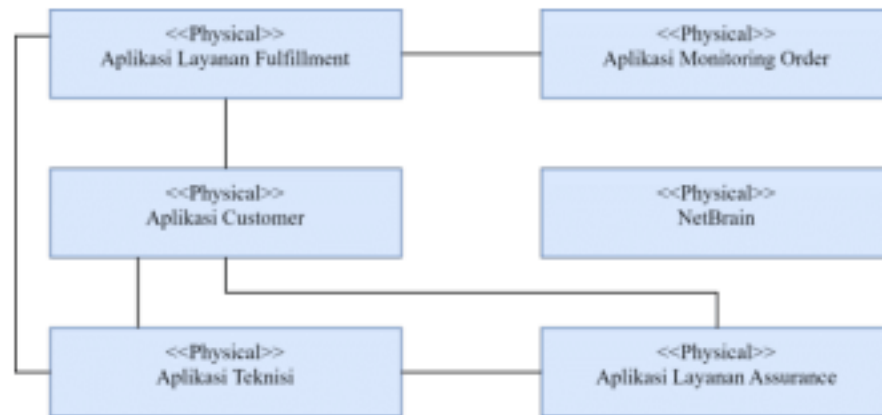
Application Portfolio Catalog banyak digunakan sebagai model perancangan Arsitektur Data adalah karena pemodelannya yang mengharuskan adanya deskripsi untuk *physical application* yang dirancang. Model ini juga berfungsi untuk mendokumentasi setiap aplikasi dalam perusahaan. Dengan ini, delegasi pekerjaan akan menjadi lebih mudah. *Application Portfolio Catalog* juga menyederhanakan manajemen dan pengelolaan aset aplikasi dalam organisasi, mendukung pengambilan keputusan terkait investasi, pengembangan, atau penghentian aplikasi, serta meningkatkan efisiensi operasional dengan mengidentifikasi redundansi dan peluang untuk konsolidasi. Selain itu, katalog ini memastikan bahwa semua aplikasi memenuhi standar keamanan dan kepatuhan yang dibutuhkan, serta memfasilitasi perencanaan pengembangan dan integrasi aplikasi baru dengan yang sudah ada dalam suatu perusahaan. Salah satu contoh model dari *Application Portfolio Catalog* yang dirancang pada [12] dapat dilihat pada Gambar 3 yang tertera di bawah ini.

APPLICATION PORTFOLIO CATALOG		
No.	Physical Application	Deskripsi
1.	Aplikasi Layanan Fulfillment	Aplikasi Layanan Fulfillment merupakan aplikasi berbasis website yang berguna untuk memproses pemesanan layanan melalui unit CFU dan melakukan aktivasi layanan telekomunikasi kepada pelanggan dengan efektif dan efisien.
2.	Aplikasi Layanan Assurance	Aplikasi Layanan Assurance adalah aplikasi berbasis website yang berguna untuk mengelola dan menyelesaikan masalah layanan telekomunikasi melalui sistem tiket, dengan pendekatan berbasis logika untuk memecahkan masalah.
3.	Aplikasi Customer	Aplikasi Customer adalah sebuah aplikasi berbasis website yang digunakan oleh Customer Facing Unit (CFU) untuk membuat tiket order dan tiket gangguan bagi pelanggan, yang selanjutnya akan diserahkan kepada Functional Unit (FU) untuk dikerjakan dan diselesaikan sesuai dengan tiket yang diberikan

Gambar 3. Contoh Model *Application Portfolio Catalog*

Sedangkan penggunaan *Application Communication Diagram* banyak digunakan adalah karena model ini memvisualisasikan dan mendokumentasikan interaksi antara berbagai aplikasi dalam suatu lingkungan TI. Dengan interaksi yang telah dipetakan, penggunaan *Application Communication Diagram* dapat memberikan banyak manfaat kepada perusahaan. Diagram ini memudahkan pemahaman tim TI dan bisnis mengenai aliran informasi dan dependensi antar aplikasi. Selain itu, diagram ini mendukung pengambilan keputusan dengan menyediakan informasi yang diperlukan untuk pengembangan, integrasi, dan pemeliharaan aplikasi secara tepat. Dalam hal pemecahan masalah, diagram ini mempermudah identifikasi dan

solusi terhadap masalah yang muncul dalam interaksi antar aplikasi. Diagram ini juga sangat berguna untuk perencanaan integrasi aplikasi baru, memastikan bahwa semua jalur komunikasi dan dependensi telah dipertimbangkan dengan baik. Terakhir, *Application Communication Diagram* membantu memastikan bahwa semua komunikasi antar aplikasi aman dan sesuai dengan regulasi yang berlaku, meningkatkan keamanan dan kepatuhan dalam lingkungan TI. Contoh model dari *Application Communication Diagram* yang dirancang pada jurnal [12] dapat dilihat pada Gambar 4.



Gambar 4. Contoh Model *Application Communication Diagram*

4. KESIMPULAN

Enterprise Architecture sangat penting dalam perusahaan telekomunikasi karena memungkinkan pengelolaan yang lebih terstruktur dan efisien terhadap sistem dan teknologi yang kompleks. Dengan *Enterprise Architecture*, perusahaan dapat menyelaraskan strategi bisnis dengan teknologi, memastikan integrasi yang lancar antara berbagai sistem, dan meningkatkan efisiensi operasional. Ini juga membantu dalam pengambilan keputusan strategis, mengidentifikasi dan mengurangi redundansi, serta memastikan kepatuhan terhadap standar industri dan regulasi. Selain itu, *Enterprise Architecture* memungkinkan perusahaan untuk beradaptasi dengan cepat terhadap perubahan teknologi dan kebutuhan pasar, menjaga daya saing, dan meningkatkan kualitas layanan kepada pelanggan.

Tahapan ini diisi oleh perancangan dari Arsitektur Data dan Arsitektur Aplikasi. Dalam tahapan Arsitektur Data, penelitian ini menunjukkan bahwa *Data Dissemination Diagram* merupakan model Arsitektur Data yang paling banyak digunakan karena kemampuannya untuk menggambarkan hubungan komponen data secara jelas, memudahkan perancangan Arsitektur Data yang efektif. Selain itu, dalam Arsitektur Aplikasi, *Application Portfolio Catalog* dan *Application Communication Diagram* adalah dua model Arsitektur Aplikasi yang paling sering digunakan. *Application Portfolio Catalog* membantu dalam mengatur manajemen aset aplikasi, mendukung pengambilan keputusan, dan memastikan kepatuhan keamanan. Sementara itu, *Application Communication Diagram* memudahkan visualisasi interaksi antar aplikasi, mendukung perencanaan integrasi, serta meningkatkan keamanan dan kepatuhan komunikasi antar aplikasi. Kedua model ini memberikan manfaat signifikan dalam pengelolaan dan pengembangan aplikasi dalam organisasi.

Saran dari penulis adalah untuk mencari sumber rujukan dan bacaan yang lebih komprehensif, memperbanyak riset khususnya pada bagian pembahasan dan metodologi untuk mencari informasi yang lebih valid dan terpercaya. Dalam judul riset penulis yang bersinggungan dengan perusahaan telekomunikasi, terkadang ada beberapa jurnal yang membawakan satu topik yang sama. Topik ini harus dikaji dengan lebih teliti lagi, khususnya pada bagian pembahasan, karena penulis takut adanya manipulasi data seperti parafrase dan pemalsuan data. Saran lainnya adalah untuk membaca sebanyak mungkin jurnal untuk menambah pengetahuan dan referensi.

REFERENSI

- [1] A. M. Karim, I. Santosa, and L. Ramadani, "PERANCANGAN ARSITEKTUR UMUM PERUSAHAAN TELEKOMUNIKASI PADA PROSES PRODUCT MANAGEMENT MENGGUNAKAN ETOM, SID DAN TAM," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 2, pp. 340–353, May 2022, doi: 10.29100/jipi.v7i2.2695.
- [2] B. Setyawan, "ENTERPRISE ARCHITECTURE UNTUK INDUSTRI KONSTRUKSI INDONESIA," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 9, no. 2, pp. 1175–1186, Jun. 2022, doi: 10.35957/jatisi.v9i2.2027.
- [3] R. Suriانشa and H. Sasongko, "Enterprise Architecture Management and its Role in Corporate Strategic Management," *Journal of Economics and Business UBS*, vol. 11, no. 1, pp. 72–78, Jun. 2022, doi: 10.52644/joeb.v11i1.83.
- [4] N. Wayan Rena Mariani, I. Made Sudjana, and M. Arya Astina, "PEMANFAATAN ZACHMAN FRAMEWORK DALAM PENGEMBANGAN BLUEPRINT SISTEM INFORMASI LEMBAGA SERTIFIKASI USAHA PARIWISATA," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 2, pp. 637–650, Jan. 2023, doi: 10.54443/sibatik.v2i2.635.
- [5] H. Supriyadi and E. Amalia, "Development of Enterprise Architecture in Senior High School Using TOGAF as Framework," *Universal Journal of Educational Research*, vol. 7, no. 4A, pp. 8–14, Apr. 2019, doi: 10.13189/ujer.2019.071402.

- [6] R. R. H. R. Yusriyahti, A. A. Nur Fajrillah, and W. A. Nurtrisha, "ENTERPRISE ARCHITECTURE: STRATEGY OF SMART VILLAGE DEVELOPMENT (VILLAGE SERVICES) USING TOGAF 9.2," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 10, no. 1, pp. 19–28, Dec. 2023, doi: 10.33330/jurteksi.v10i1.2542.
- [7] N. Legowo and I. Kaharmies, "Enterprise Architecture Application and Business Process Improvement: A Case Study of Bus Terminal in Indonesia," *Journal of System and Management Sciences*, vol. 13, no. 5, Sep. 2023, doi: 10.33168/JSMS.2023.0524.
- [8] Heri Aji Setiawan, Andi Naila Quin Azisah Alisyahbana, Arnes Yuli Vandika, Bambang Prihantoro Nugroho, and Ikhwanto Asri, "Enterprise Architecture Design Using TOGAF Adm at PT. Industri Telekomunikasi Indonesia," *Indonesian Journal of Enterprise Architecture*, vol. 1, no. 2, pp. 77–87, Feb. 2024, doi: 10.61220/ijea.v1i2.0239.
- [9] N. F. Hadi and N. K. Afandi, "Literature Review is A Part of Research," *Sultra Educational Journal*, vol. 1, no. 3, pp. 64–71, Dec. 2021, doi: 10.54297/seduj.v1i3.203.
- [10] M. K. Linnenluecke, M. Marrone, and A. K. Singh, "Conducting systematic literature reviews and bibliometric analyses," *Australian Journal of Management*, vol. 45, no. 2, pp. 175–194, May 2020, doi: 10.1177/0312896219877678.
- [11] Y. Gong and M. Janssen, "The value of and myths about enterprise architecture," *Int J Inf Manage*, vol. 46, pp. 1–9, Jun. 2019, doi: 10.1016/j.ijinfomgt.2018.11.006.
- [12] J. Jimmy, D. Praditya, and A. F. Santoso, "ANALISIS DAN PERANCANGAN ARSITEKTUR PERUSAHAAN MENGGUNAKAN KERANGKA KERJA TOGAF 10 DI DIVISI OPERASIONAL LAYANAN PADA PT.XYZ," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 2, pp. 654–668, May 2024, doi: 10.29100/jipi.v9i2.4610.
- [13] M. Zulfani and A. W. R. Emanuel, "Perancangan Model Enterprise Architecture Planning pada PT. Bestari Kalimantan," *Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 8, no. 2, pp. 341–350, Apr. 2024, doi: 10.35870/jtik.v8i2.1795.
- [14] K. 'Afiifah, Z. F. Azzahra, and A. D. Anggoro, "Analisis Teknik Entity-Relationship Diagram dalam Perancangan Database Sebuah Literature Review," *INTECH*, vol. 3, no. 1, pp. 8–11, Apr. 2022, doi: 10.54895/intech.v3i1.1261.
- [15] M. A. Pangestu, R. Fauzi, and I. Santosa, "Analisis Perancangan Enterprise Architecture Fungsi Risk And Quality Management Dengan Togaf ADM (Studi Kasus: Telkom Corporate University Center)," *Jurnal Teknik Informatika dan Sistem Informasi ISSN*, vol. 2407, p. 4322, 2022.
- [16] D. Kwek, D. Maulana, E. R. Kaburuan, and N. Legowo, "Enterprise architecture planning information system based on cloud computing using togaf (case study: Pandi. Id registry)," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 8, no. 09, 2019.
- [17] Y. S. Hudha, E. Utami, and E. T. Luthfi, "Perancangan Enterprise Arsitektur Sistem Informasi Billing Menggunakan Metode TOGAF ADM Pada PT. Time Excelindo," *Creative Information Technology Journal*, vol. 5, no. 1, p. 40, May 2019, doi: 10.24076/citec.2017v5i1.125.

Color Space Influence on Photosynthetic Pigment Measurement Accuracy Using CNN in Color Constancy

Ade May Luky Harefa¹, Arief Muhazir Insandi²

¹ Ilmu Komputer, Institut Bisnis dan Komputer Indonesia, Indonesia

² Agribisnis, Institut Bisnis dan Komputer Indonesia, Indonesia

Email: ¹ademayluky@gmail.com, ²ariefmuhazir@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk merancang sistem pengukuran pigmen tanaman menggunakan citra digital dan deep learning, dengan mengintegrasikan berbagai ruang warna seperti RGB, HSV, LAB, dan YCbCr. Metode yang diusulkan merupakan alternatif yang lebih cepat, lebih hemat biaya, dan lebih akurat dibandingkan metode tradisional seperti analisis spektrofotometri dan HPLC. Hasil eksperimen menunjukkan bahwa pemilihan ruang warna dan pengaturan preprocessing inpaint secara signifikan mempengaruhi akurasi model CNN P3Net. Kombinasi RGB+YCbCr dengan inpaint serta RGB+LAB tanpa inpaint menghasilkan nilai MAE validasi terendah. Studi ini juga membuktikan bahwa fenomena color constancy memengaruhi akurasi model, di mana ruang warna yang mempertimbangkan fenomena ini, seperti RGB+YCbCr dengan inpaint, memberikan akurasi lebih baik dibandingkan yang tidak.

Kata Kunci: CNN, Ruang Warna, Citra Digital, Pengukuran Pigmen

ABSTRACT

This study aims to design a plant pigment measurement system using digital images and deep learning, incorporating various color spaces including RGB, HSV, LAB, and YCbCr. The proposed method serves as a faster, more cost-effective, and accurate alternative to traditional methods such as spectrophotometric analysis and HPLC. Experimental results indicate that the choice of color space and inpaint preprocessing settings significantly impacts the accuracy of the CNN P3Net model. The combination of RGB+YCbCr with inpaint and RGB+LAB without inpaint yielded the lowest validation MAE values. The study also demonstrates that color constancy phenomena influence model accuracy, with color spaces that account for this phenomenon, such as RGB+YCbCr with inpaint, providing better accuracy than those that do not.

Keywords: CNN, Color space, Digital image, Pigment measurement

Corresponding Author:

Ade May Luky Harefa

Email: ademayluky@gmail.com

Article Info

Received: 14 Februari 2025

Revised: 15 Februari 2025

Accepted: 17 Februari 2025

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.



1. INTRODUCTION

Plant metabolism occurs through a chemical process known as photosynthesis [1]. The mechanism of photosynthesis involves triggering a chemical reaction using light energy to convert carbon dioxide and water into glucose and oxygen as a byproduct [2]. The produced glucose can then be utilized as energy for plant metabolism or transformed into starch, which is stored and used when needed [3]. In the photosynthesis process, there are pigment molecules responsible for absorbing light energy for the chemical reaction, specifically chlorophyll and carotenoids [4]. Essentially, pigment molecules capture and reflect light energy within a specific spectrum [5]. The absorbed light energy is used to support the photosynthesis process, while the reflected light energy causes the color effect in plants, limited to the wavelength of the reflected light energy [6]. Chlorophyll pigments are green pigments, meaning they reflect the green spectrum of light. Carotenoid pigments, on the other hand, reflect light within the red, orange, and yellow spectrums [7].

Pigment content in plants is generally measured through laboratory extraction using spectrophotometric analysis, which measures the absorbance and reflectance of the leaf pigment solution, or through High-Performance Liquid Chromatography (HPLC) [8]. Both methods provide accurate results, but they differ in their measurement outcomes. Spectrophotometric analysis can only separate the results for chlorophyll a and chlorophyll b pigments, but not for other pigments like carotenoids or anthocyanins [9]. In contrast, HPLC analysis can provide detailed information on individual pigments [10]. The drawbacks of both measurement methods are that they are relatively time-consuming, costly, and destructive, as the leaves must be crushed to extract the pigments [11]. Therefore, there is a need for an easy, fast, cost-effective, and accurate method for measuring plant pigment content [12].

The proposed method is to measure pigments using digital images taken with a smartphone camera, which are then processed using deep learning methods to estimate the pigment content in the leaves. Previous research [13] discussed measuring pigment content (chlorophyll, carotenoids, anthocyanins) in plants using digital images and deep learning, but the study was limited to only onecolor space, RGB (Red, Green, Blue). Therefore, this study aims to design a plant pigment measurement system using digital images and deep learning, incorporating multiple different color spaces, including RGB, HSV (Hue, Saturation, Value), LAB, and YCbCr as inputs.

2. RESEARCH METHOD

2.1. Method of Data Collection

The data collected were 212 leaf image data along with the content of photosynthetic pigments from each leaf object in the image. The plants used in the collection of leaf image data were piper betle, jasminum, syzigium oleina, and graphtophyllum pictum [14]. The method used for data acquisition is divided into 2 steps [15]. The first step is to acquire leaf image data using a smartphone digital camera. The second step is the measurement of photosynthetic pigment levels. The measurement of photosynthetic pigment levels goes through a preparatory stage before being inserted into a UV-vis spectrophotometer [16]. The preparation stage is the division of the leaf into 2 parts which are then crushed into small parts with a weight of 0.05 grams each [17]. The first part is for the measurement of chlorophyll and carotenoid levels, while the second part is for the measurement of anthocyanin levels. The solution used for pigment extraction is distinguished by the type of pigment. For chlorophyll and anthocyanins, the solvent contains 100% acetone, while for anthocyanins the solvent is a mixture of methanol, concentrated hydrochloric acid, and filtered water [18]. A solvent of 1.5 mL is then added calcium carbonate and sodium ascorbate and put into the tube [19]. The solution and leaf samples were then mixed using *vortex* for 1 minute, and cooled for 1 minute for 3 times for each tube, then centrifuged at 14,000 rpm for 2 minutes and then cooled [20]. After the preparation stage was completed, the sample was measured for absorption using a spectrophotometer and the pigment level was calculated using the [21] and [22] method.



Figure 1. Leaf sample image data

Table 1 presents the data on photosynthetic pigment content, including anthocyanins, carotenoids, and chlorophyll, measured in different sample images. The values indicate the concentration of each pigment in the respective samples, which can provide insights into the physiological and biochemical characteristics of the plants. Anthocyanins are expressed in Abs/g, while carotenoids and chlorophyll are measured in $\mu\text{g/g}$. The variations in pigment content among the samples highlight differences in pigmentation, which may be influenced by factors such as species, environmental conditions, and developmental stages.

Image code	Anthocyanins (Abs/g)	Carotenoids ($\mu\text{g/g}$)	Chlorophyll ($\mu\text{g/g}$)
G01	25.8532	0.0000	461.9024
J01	0.0000	0.0000	1011.7251
PB01	0.7897	11.6864	102.6544
PM01	0.0296	8.8892	357.2901

2.2. The Creation of Datasets

To facilitate comparison and further analysis, the photosynthetic pigment content data were normalized. Normalization ensures that the values are scaled within a specific range, allowing for better interpretation of differences across samples. Table 2 presents the normalized pigment levels for anthocyanins, carotenoids, and chlorophyll in the analyzed images. These normalized values provide a standardized representation of pigment concentrations, aiding in identifying patterns and relationships among the samples.

Table 2. Normalized photosynthetic pigment level data

Image code	Anthocyanins (Abs/g)	Carotenoids ($\mu\text{g/g}$)	Chlorophyll ($\mu\text{g/g}$)
G01	0.2881	0.0000	0.4565
J01	0.0000	0.0000	1.0000
PB01	0.0088	0.2195	0.1015
PM01	0.0003	0.1670	0.3531

The creation of datasets was carried out on two types of data, namely photosynthetic pigment level measurement data and leaf image raw data. The measured data of photosynthetic pigment levels obtained was changed by performing a normalization operation to reduce the magnitude of the data [23]. The normalized pigment level data is shown in Table 2. The raw data of the images obtained from the data collection stage still requires the preprocessing stage to become a suitable dataset [24]. The data preprocessing process is divided into 3 types, namely, the elimination of white light reflections on the surface, the creation of color space channels of RGB, HSV, LAB, and YCbCr, and then the data augmentation process which then becomes the input dataset that is entered into the CNN [25]. The use of datasets in the training process of the CNN model is 80% data for training data and 20% for validation data. The explanation of the stage of creating a dataset is as follows [26]:

- a. Removal of the effect of white light reflection on the surface
 White light reflection is minimized using the algorithms of Alexandru Telea and Navier Stokes. Both algorithms are available in the form of an inpaint function in the OpenCv library. The inpaint operation is carried out accompanied by a mask element so that the inpaint operation is properly carried out on the white light reflection part of the leaf image. The surgical steps for the elimination of white light reflection on the leaf surface are as follows:
 1. Access leaf image data.
 2. Segmentation of leaf objects.
 3. Making a mask for the target position of the white light reflection on the leaf image.
 4. Inpaint operation.
- b. Creation of channels for each color space
 The creation of color space channels is achieved by converting RGB color spaces into HSV, LAB, and YCbCr color spaces. After the color space conversion operation is performed, the color space components are separated and stored individually as independent components. From the process carried out, a dataset of components from various color spaces is formed that can be combined as a dataset that is included in the CNN model. The use of color space channels as input data refers to research conducted by Prilianti et al. which used 10 multispectral color channels as input data to CNNs. The steps to create a color space channel are as follows.
 1. Access leaf image data.
 2. Color space conversion.
 3. Channel separation.
 4. Channel file storage.
- c. Data Augmentation
 The data augmentation process was carried out to overcome the possibility of low CNN performance due to the lack of available datasets. The data augmentation carried out is an addition to data variation by rotating the data with various rotation angles, zooming in, and applying standardization. The data generated from the previous 2 processes is applied data augmentation operations.

This study focuses on analyzing the effects resulting from the use of different types of color spaces on the accuracy of CNN, therefore the architecture used is the CNN architecture. The following is a description of the architectural design used with Table 3. and illustrated with Figure 2.

Table 3. CNN Model Architecture

Layer	Feature Map	Size	Kernel Size	Stride	Activation	Dropout	
Input	Image						
1	Convolution	32	Input	3x3	2x2	ReLU	-
2	Convolution	32		3x3	1x1	ReLU	-
3	Pooling	-		3x3	2x2		-
4	Flattening	-		-	-		-
5	Dense	-	90	-	-	Sigmoid	0.03
6	Dense	-	3	-	-	LeakyRelu (alpha=0.3)	-

The CNN architecture in this study uses mean squared error as the loss function and Nesterov Adam as the optimization function. Technically the CNN model is built using the Hard library and the Python language. The CNN input is leaf image data from various color spaces and the output produced is an estimate of the content of photosynthetic pigments, namely chlorophyll, carotenoids, and anthocyanins.

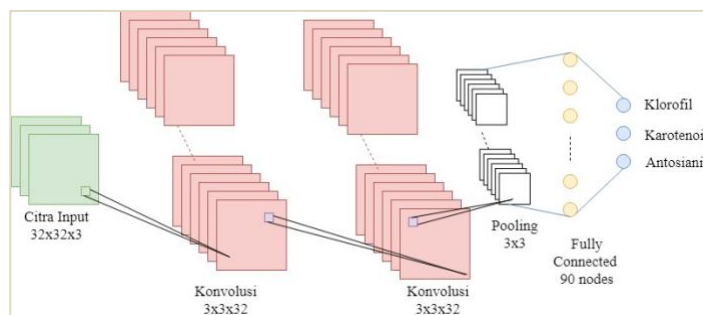


Figure 2. CNN model architecture

3. RESULTS AND DISCUSSIONS

3.1 Single Color Space with The Best Accuracy Without Inpainting

Figure 3 illustrates the sequence of accuracy values for a single-color space model without inpainting. This graph provides insight into the model's performance over different iterations, highlighting variations in accuracy across the dataset. Analyzing this trend helps evaluate the model's consistency and effectiveness in processing images without inpainting techniques.

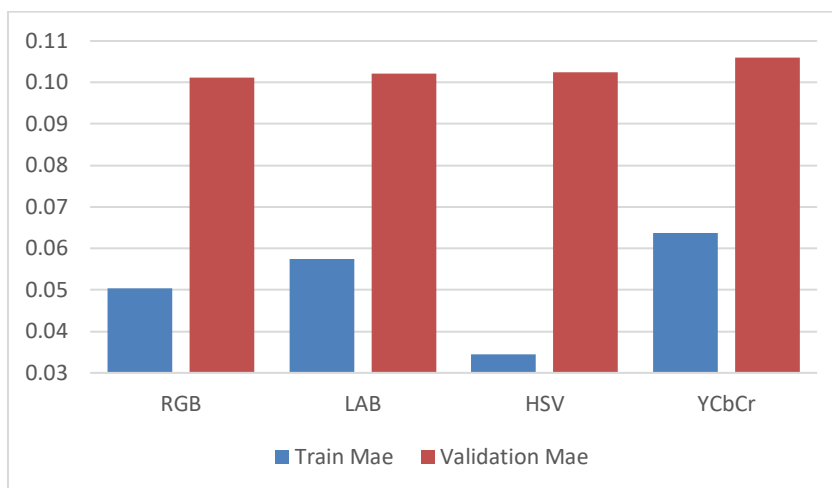


Figure 3. Graph Sequence Of Accuracy Values of a Singlecolor Space Model Without Inpaint

Table 4 presents the sequence of accuracy values for a single-color space model without inpainting, measured using the Mean Absolute Error (MAE) for both training and validation. The table compares the performance of different color spaces, including RGB, LAB, and YCbCr, by evaluating their respective MAE scores. Lower MAE values indicate better model accuracy, with differences in scores reflecting the impact of each color space on model performance.

Table 4. Sequence of accuracy values of a singlecolor space model without *inpaint*

No	Color space	Average MAE training score	Average MAE Validation score
1	RGB	0.05031	0.10115
2	LAB	0.05738	0.10221
3	YCbCr	0.06376	0.10603

The results of the experiment carried out by conducting a training process on models that use a singlecolor space without the inpaint process as input data are recorded. Then the singlecolor space without the inpaint process is sorted based on the average validation MAE value from the smallest to the largest as shown in Figure 3. The color space compared in the MAE value table is the color space that does not give rise to symptoms of overfitting. Table 3. shows the sequence of the results of the training experiment on a singlecolor space without the inpaint process.

From the data obtained, it is shown that the tw color spaces with the smallest MAE values for model validation are RGB and LAB with MAE values of 0.10115 and 0.10221. Therefore, the singl color space without the inpaint process that is recommended for input data in the use of CNN as a method of estimating pigment content is the RGB and LAB color space which is the color space with the smallest MAE value result of model validation.

3.2 Combination Color Space with The Best Accuracy Without Inpainting

Figure 4 illustrates the sequence of accuracy values for a combination color space model without inpainting. This graph compares the model's performance when multiple color spaces are utilized together, providing insights into how different combinations affect accuracy. By analyzing this trend, it is possible to determine the effectiveness of integrating multiple color spaces in improving model performance.

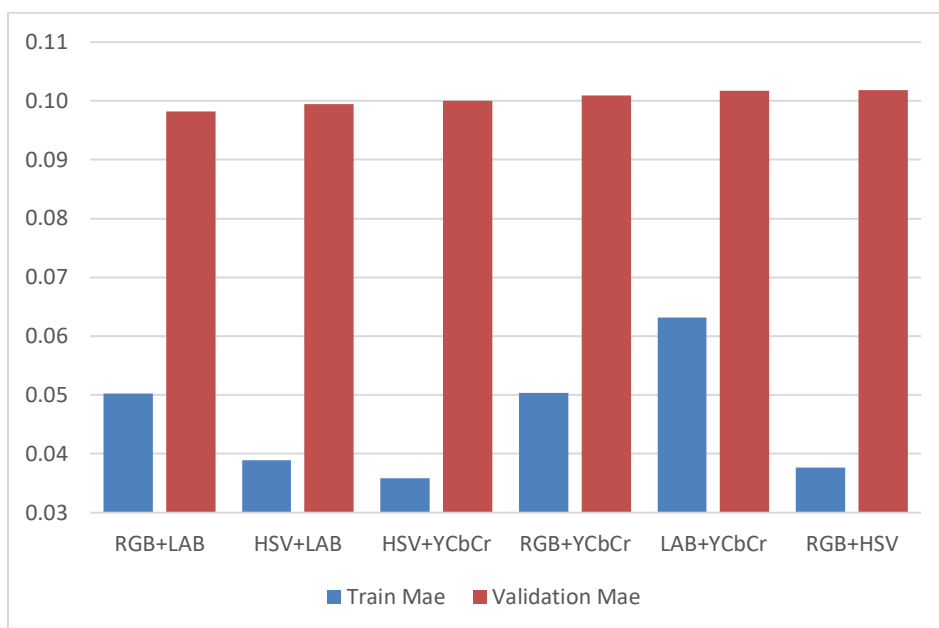


Figure 4. Graph Sequence of Accuracy Values of a Combination Color Space Model Without Inpaint

The results of the experiment carried out by conducting a training process on models that use a combination color space without the inpaint process as input data are recorded. Then the combination color space without the inpaint process is sorted based on the average validation MAE value from the smallest to the largest as shown in Figure 4. Table 5 shows the sequence of the results of the training experiment on a singlecolor space without inpaint. From the data obtained, it is shown that the twocolor spaces of the model with the smallest validation MAE values are RGB+LAB and RGB+YCbCr. Therefore, the recommended color space for input data in the use of CNN as a method of estimating pigment content is the RGB+LAB and RGB+YCbCr color space. The validation MAE value of the RGB+LAB color space is 0.09820, while for the RGB+YCbCr color space is 0.10089.

Table 5. Sequence of Accuracy Values of Combination Color Space Model Without Inpaint

No	Color space	Average MAE training score	Average MAE Validation score
1	RGB+LAB	0.05020	0.09820
2	RGB+YCbCr	0.05033	0.10089
3	LAB+YCbCr	0.06322	0.10171

3.3 Comparison of Accuracy of Single and Combined Color Spaces Without Inpaint Process

Figure 5 presents the sequence of accuracy values for both the single-color space model and the combination color space model without inpainting. This graph allows for a direct comparison of the performance between the two approaches, highlighting how the integration of multiple color spaces influences accuracy in contrast to using a singlecolor space. The analysis of this sequence provides valuable insights into the advantages and limitations of each model configuration.

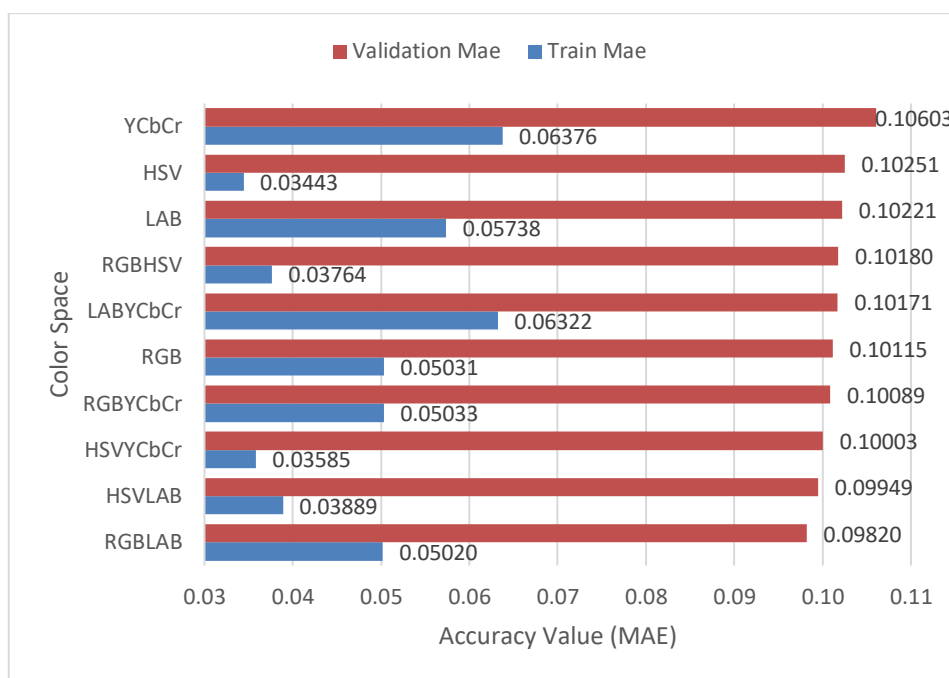


Figure 5. Graph Sequence of Accuracy Values of Singlecolor Space Model and Combination Without Inpaint

The type of color space without using the recommended inpaint process is obtained by comparing the average value of the validation MAE of the color space as a whole, both single color space and combination color space. The average value data of single and combined color space MAE is shown in Figure 5. Table 6 shows the color space data that has been sorted based on the average value of the MAE validation and the model is not overfitted. From the experiments carried out, the three-color spaces of the model with the smallest MAE values of model validation are RGB+LAB, RGB+YCbCr, and RGB color spaces with values of 0.09820, 0.10089, and 0.10115. Therefore, the three-color spaces without the inpaint process that are recommended as CNN input data as a method of estimating leaf pigment levels are RGB+LAB, RGB+YCbCr, and RGB.

Table 6. Order of Accuracy Values of All Color Space Models Without Inpaint

No	Color space	Average MAE training score	Average MAE Validation score
1	RGB+LAB	0.05020	0.09820
2	RGB+YCbCr	0.05033	0.10089
3	RGB	0.05031	0.10115
4	LAB+YCbCr	0.06322	0.10171
5	LAB	0.05738	0.10221
6	YCbCr	0.06376	0.10603

3.4 Single Color Space with The Best Accuracy with Inpaint Process

Figure 6 displays the sequence graph of accuracy values for a single-color space model with inpainting. This graph illustrates how the inpainting technique affects the model's performance over different iterations. By comparing this sequence with models without inpainting, we can assess the impact of inpainting on improving the accuracy of the single-color space model.

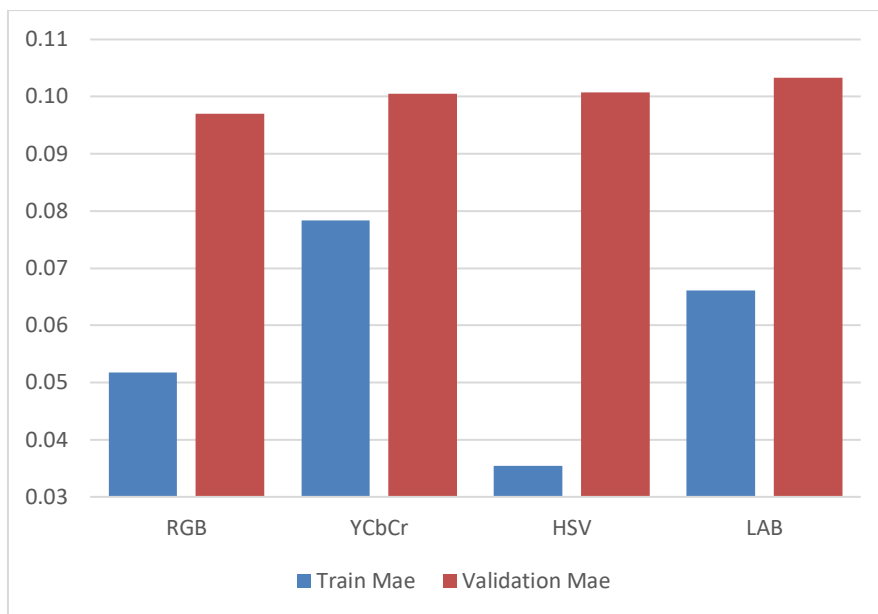


Figure 6. Sequence Graph of Accuracy Values of a SingleColor Space Model with Inpaint

The results of the experiment in the form of the average value of MAE validation from the color space with the inpaint process were obtained by conducting a training process on the CNN models of each color space that had previously gone through the inpaint process. Figure 7 shows the accuracy results of a CNN model with a single color space with inpaint as the input data. Table 6. shows the order of MAE average values of singlecolor space validation with the inpaint process as input data on the CNN model from smallest to largest.

Table 7. Sequence of accuracy values of a singlecolor space model with inpaint

No	Color space	Average MAE training score	Average MAE Validation score
1	RGB	0.05173	0.09696
2	LAB	0.06606	0.10333
3	YCbCr	0.06489	0.11199

The average value of MAE validation of a singlecolor space with *the inpaint process* shows that the twocolor spaces with the smallest MAE validation values of the model are RGB and LAB color spaces with average MAE validation values of 0.09696 and 0.10033. Therefore, the recommended singlecolor space with *the inpaint process* is RGB and LAB.

3.5 Combination color space with the best accuracy with the inpaint process

The results of the experiment carried out by training the CNN model with a combination color space obtained from images that have gone through the inpainting process are shown in Figure 7. From the results of the experiment of training the CNN model using a combination color space with the *inpaint process*, the twocolour spaces of the model with the smallest MAE validation values are RGB+YCbCr and LAB+YCbCr color spaces with average MAE validation values of 0.09486 and 0.10239. Therefore, the recommended color space for the color space input data combined with *inpaint* in the CNN model is RGB+YCbCr and LAB+YCbCr color spaces.

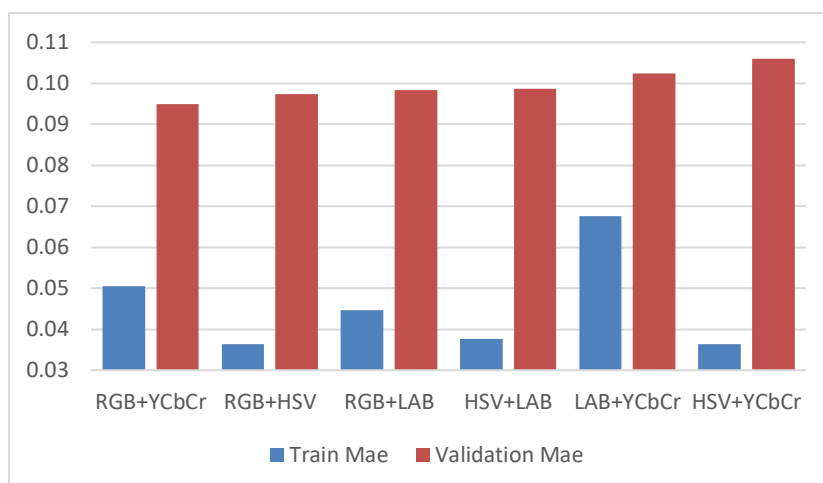


Figure 7. Graph Sequence of Accuracy Values of Color Space Model Combination with Inpaint

Table 8 presents the order of accuracy values for the color space model combination with inpainting. This table compares the performance of different combinations of color spaces, with inpainting applied, to determine which configuration provides the highest accuracy. The

accuracy values provide insight into how the integration of multiple color spaces, combined with inpainting techniques, influences the model's effectiveness.

Table 8. Order of Accuracy Values of Color Space Model Combination with Inpaint

No	Color space	Average MAE training score	Average MAE Validation score
1	RGB+YCbCr	0.05051	0.09486
2	LAB+YCbCr	0.06761	0.10239

3.6 Comparison of the Accuracy of Single and Combined Color Spaces with The Inpaint Process

The type of color space with the recommended inpaint process can be obtained by obtaining the color space with the smallest average validation MAE value of the entire color space, both single and combined color spaces. Figure 8. shows the accuracy value of the model in the form of the average MAE of each color space both for training and validation that has been determined based on the average value of the MAE validation, while for the sequence of color spaces with *the* inpaint process that is not overfitting and sorted based on the average value of the MAE validation is shown in Table 8.

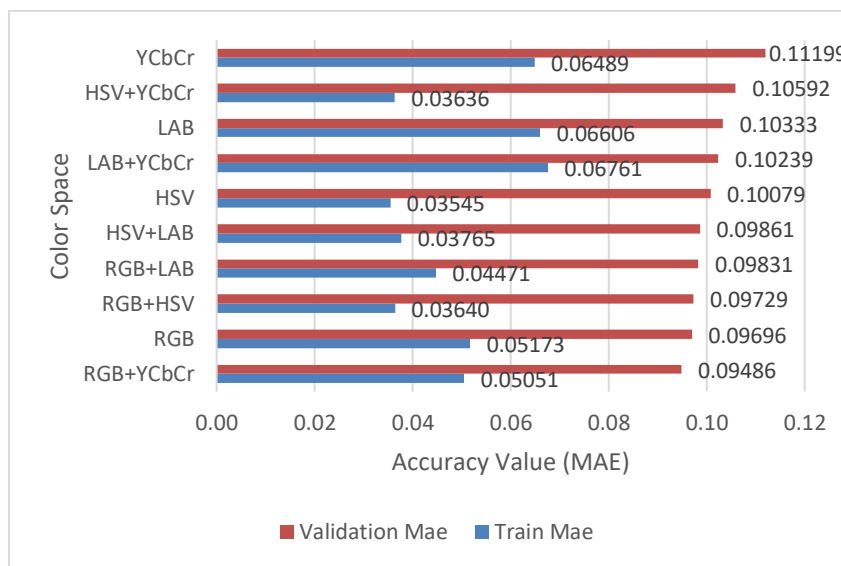


Figure 8. Graph Sequence of Accuracy Values of a Singlecolor Space Model and Combination with Inpaint

Table 9 presents the order of accuracy values for both the single-color space model and the combination color space model with inpainting. This table allows for a direct comparison between the performance of single-color space models and combination models, both utilizing inpainting. By examining these values, we can determine which approach yields the highest accuracy and evaluate the benefits of combining multiple color spaces with inpainting.

Table 9. Order of accuracy values of singlecolor space model and combination with inpaint

No	Color space	Average MAE training score	Average MAE Validation score
1	RGB+YCbCr	0.05051	0.09486
2	RGB	0.05173	0.09696
3	LAB+YCbCr	0.06761	0.10239
4	LAB	0.06606	0.10333
5	YCbCr	0.06489	0.11199

From the results of the experiment, it can be seen that the threecolor spaces with the smallest MAE accuracy values of the color space with the smallest inpaint process are RGB+YCbCr, RGB, and LAB+YCbCr with MAE accuracy values of 0.09486, 0.09696, and 0.10239. Therefore, the recommended color spaces with the inpaint process are RGB+YCbCr, RGB, and LAB+YCbCr.

3.7 Best Color Space from A Comparison of All Inpainted and Non-Inpainted Color Spaces

Table 10 presents the order of the best accuracy values for all color space models with inpainting preprocessing. This table ranks the performance of each color space model, incorporating inpainting as a preprocessing technique, to identify the configurations that achieve the highest accuracy. Analyzing these results helps determine the most effective color space models and preprocessing methods for optimizing model performance.

Table 10. Order of Best Accuracy Values for All Color Space Models and Inpaint Preprocessing

No	Color space	Inpaint	Average value of MAE trainer	Average MAE validation score
1	RGB+YCbCr	Yes	0.05051	0.09486
2	RGB	Yes	0.05173	0.09696
3	RGB+LAB	It	0.05020	0.09820
4	RGB+YCbCr	It	0.05033	0.10089

5	RGB	It	0.05031	0.10115
6	LAB+YCbCr	It	0.06322	0.10171
7	LAB	It	0.05738	0.10221
8	LAB+YCbCr	Yes	0.06761	0.10239
9	LAB	Yes	0.06606	0.10333
10	YCbCr	It	0.06376	0.10603
11	YCbCr	Yes	0.06489	0.11199

The experimental results showed that the 3 types of CNN model input data for pigment content estimation with the smallest average value of MAE validation were RGB+YCbCr color space with inpaint with an average value of 0.09486 MAE validation, RGB with inpaint with an average value of 0.09696 MAE validation, and RGB+LAB without inpaint with an average value of 0.09820 MAE validation. From the results obtained, the recommended color spaces are RGB+YCbCr with inpaint, RGB with inpaint, and RGB+LAB without inpaint. From the results obtained, color spaces with characteristics that consider lighting elements are included in the color spaces that produce the best accuracy values. This can be explained by the previous hypothesis regarding the relationship between how the type of color space gets the color value of an object, namely the separation of the lighting element from other elements such as "L" (luminance) in LAB or "Y" in YCbCr. In color space models such as LAB or YCbCr, color elements other than lighting are always consistent under different lighting, so that it can reduce the influence of the color constancy phenomenon. Table 9. shows the results of experiments on the entire singlecolor space and combination with or without the inpainting process. Table 9. shows the sequence of color space models and types of preprocesses that are sorted based on the average value of the MAE validation and do not experience overfitting.

Regarding the effect of inpainting on the accuracy of the CNN model, by looking at the data of the training MAE value and validation shown in Table 9. it can be concluded that the preprocessing of inpaint on the training image data does not provide a large decrease in the accuracy value. This conclusion was obtained from the difference from the MAE value of the training and validation between the models with image data that went through the inpainting process and did not go through the inpainting process. However, there is another conclusion related to inpaint preprocessing, namely the accuracy value of the model with the RGB color space that goes through the inpaint process is included in the group of threecolor spaces with the smallest validation MAE value. The twocolor spaces from the group of threecolor spaces with the smallest MAE validation values are RGB+YCbCr and RGB+LAB color spaces which consider lighting elements that can minimize the effect of the color constancy phenomenon, while RGB does not consider lighting elements. However, from the results of the research, the application of the inpainting process to RGB training image data is able to provide accuracy results that rival color spaces that consider lighting elements.

4. CONCLUSION

Experiments conducted by training the CNN model using different color space types and preprocessing settings showed differences in training accuracy values and model validation for input data types. By sorting the color space type and preprocessing settings based on the average value of the validation MAE, the recommended color space type for the CNN P3Net model input data of the estimated pigment content can be determined. For input data without inpaint process from a single or combination of color spaces, the recommended color spaces are RGB+LAB, RGB+YCbCr, and RGB. However, for input data without an inpaint process with a singlecolor space, the recommended color spaces are RGB and LAB. As for input data without an inpaint process with a combination color space, the recommended ones are RGB+LAB and RGB+YCbCr.

The results of the CNN P3Net model training experiment using the inpaint process on the image data resulted in the conclusion that of the entire color space, either single or combined color space, the recommended ones are RGB+YCbCr, RGB, and LAB+YCbCr. The recommended singlecolor space for input data with the inpaint process is RGB and LAB. Meanwhile, the recommended combination color space for input data with the inpaint process is RGB+YCbCr and LAB+YCbCr.

The results of the comparison of all color spaces and inpaint preprocessing settings show that the recommended color spaces are RGB+YCbCr with inpaint, RGB with inpaint, and RGB+LAB without inpaint. The experimental results showed that the type of color space and inpaint preprocessing settings affected the accuracy of the CNN P3Net model. The entry of RGB+YCbCr color space with inpaint and RGB+LAB without inpaint is the color space by considering the lighting element as model input data CNN P3Net which produces the smallest validation MAE accuracy value and the consistency of the type of color space that considers lighting elements is included in the smallest result in the comparison of the validation MAE value both with inpaint and not, proving that the type of color space by considering the phenomenon of color constancy can affect the accuracy of the CNN P3Net model and provides a better accuracy value when compared to the type of color space that does not consider the phenomenon color constancy (RGB), in this case the color space that considers the color constancy phenomenon that provides the best accuracy value is RGB+YCbCr with inpaint. The influence of the inpainting process obtained from the results of the study is twofold, namely, the inpainting process on the training image data does not provide a large decrease in MAE values and the inpainting process in the RGB color space which has the characteristic of not considering lighting elements can produce an accuracy value that occupies the second position in the group of the three smallest color spaces, so that it can compete with models that use color spaces that consider lighting elements.

REFERENSI

- [1] S. C. Bhatla and M. A. Lal, *Plant physiology, development and metabolism*. Springer Nature, 2023.
- [2] S. Ray, J. Abraham, N. Jordan, M. Lindsay, and N. Chauhan, "Synthetic, photosynthetic, and chemical strategies to enhance carbon dioxide fixation," *C*, vol. 8, no. 1, p. 18, 2022.
- [3] M. N. Nakrani, R. H. Wineland, and F. Anjum, "Physiology, glucose metabolism," 2020.
- [4] S. Zulfiqar, S. Sharif, M. Saeed, and A. Tahir, "Role of carotenoids in photosynthesis," *Carotenoids Struct. Funct. Hum. body*, pp. 147–187, 2021.

- [5] P. M. Glibert, "Capturing Light and the Diversity of Pigments," *Phytoplankt. Whispering An Introd. to Physiol. Ecol. Microalgae*, pp. 85–103, 2024.
- [6] J. Liu and M. W. Van Iersel, "Photosynthetic physiology of blue, green, and red light: Light intensity effects and underlying mechanisms," *Front. Plant Sci.*, vol. 12, p. 619987, 2021.
- [7] E. Narbona, J. C. del Valle, and J. B. Whittall, "Painting the green canvas: how pigments produce flower colours," *Biochem. (Lond.)*, vol. 43, no. 3, pp. 6–12, 2021.
- [8] F. Petibon and G. L. B. Wiesenberg, "Characterization of complex photosynthetic pigment profiles in European deciduous tree leaves by sequential extraction and reversed-phase high-performance liquid chromatography," *Front. Plant Sci.*, vol. 13, p. 957606, 2022.
- [9] M. Manninen, V.-M. Vesterinen, and J.-P. Salminen, "Chemistry of autumn colors: quantitative spectrophotometric analysis of anthocyanins and carotenoids and qualitative analysis of anthocyanins by ultra-performance liquid chromatography–tandem mass spectrometry," *J. Chem. Educ.*, vol. 97, no. 3, pp. 772–777, 2020.
- [10] S. Khani, J. B. Ghasemi, and Z. Piravi-vanak, "Development of a computer vision system for the classification of olive oil samples with different harvesting years and estimation of chlorophyll and carotenoid contents: A comparison of the proposed method's efficiency with UV-Vis spectroscopy," *J. Food Compos. Anal.*, vol. 129, p. 106078, 2024.
- [11] M. Islam, S. Bijjahalli, T. Fahey, A. Gardi, R. Sabatini, and D. W. Lamb, "Destructive and non-destructive measurement approaches and the application of AI models in precision agriculture: a review," *Precis. Agric.*, vol. 25, no. 3, pp. 1127–1180, 2024.
- [12] Z. Zlatev, V. Stoykova, G. Shivacheva, and M. Vasilev, "Design and Implementation of a Measuring Device to Determine the Content of Pigments in Plant Leaves," *Appl. Syst. Innov.*, vol. 6, no. 4, p. 64, 2023.
- [13] J. Chaki and N. Dey, *Image color feature extraction techniques: fundamentals and applications*. Springer Nature, 2020.
- [14] K. R. Prilianti, S. Anam, T. H. P. Brotosudarmo, and A. Suryanto, "Real-time assessment of plant photosynthetic pigment contents with an artificial intelligence approach in a mobile application," *J. Agric. Eng.*, vol. 51, no. 4, pp. 220–228, 2020.
- [15] J. Liu, Y. Sun, W. Liu, Z. Tan, J. Jiang, and Y. Li, "Association of spectroscopically determined leaf nutrition related traits and breeding selection in *Sassafras tzumu*," *Plant Methods*, vol. 17, pp. 1–10, 2021.
- [16] B. Fernández-Marín, J. I. García-Plazaola, A. Hernández, and R. Esteban, "Plant photosynthetic pigments: methods and tricks for correct quantification and identification," *Adv. plant Ecophysiol. Tech.*, pp. 29–50, 2018.
- [17] C. N. Waluchio, "Nutrient and antinutrient content in leaves of selected Coastal Kenya cassava varieties as affected by maturity stage, leafage and preparation method." university of Nairobi, 2016.
- [18] S. A. N. Che Musa, R. M. Taha, U. N. A. Abdul Razak, N. Anuar, and A. K. Arof, "The effects of different solvent extraction and pH on the stability of carotenoids and chlorophyll in *Cucumis melo* L. for potential coating technology," *Pigment Resin Technol.*, vol. 47, no. 6, pp. 511–516, 2018.
- [19] T. Okazaki, W. Wang, H. Kuramitz, N. Hata, and S. Taguchi, "Molybdenum blue spectrophotometry for trace arsenic in ground water using a soluble membrane filter and calcium carbonate column," *Anal. Sci.*, vol. 29, no. 1, pp. 67–72, 2013.
- [20] L. Z. Drábková, "DNA extraction from herbarium specimens," *Mol. Plant Taxon. Methods Protoc.*, pp. 69–84, 2014.
- [21] L. Zhan, J. Hu, Z. Ai, L. Pang, Y. Li, and M. Zhu, "Light exposure during storage preserving soluble sugar and L-ascorbic acid content of minimally processed romaine lettuce (*Lactuca sativa* L. var. *longifolia*)," *Food Chem.*, vol. 136, no. 1, pp. 273–278, 2013.
- [22] F. Madzimur, "Spectral transformation of image Data: Vegetation remote sensing for forest monitoring," *Int. J. Phys. Soc. Sci.*, vol. 7, no. 5, pp. 1–5, 2017.
- [23] J. A. Gamon, O. Kovalchuck, C. Y. S. Wong, A. Harris, and S. R. Garrity, "Monitoring seasonal and diurnal changes in photosynthetic pigments with automated PRI and NDVI sensors," *Biogeosciences*, vol. 12, no. 13, pp. 4149–4159, 2015.
- [24] S. Wang *et al.*, "Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects," *Inf. Fusion*, vol. 76, pp. 376–421, 2021.
- [25] A. F. M. Raffei, H. Asmuni, R. Hassan, and R. M. Othman, "Fusing the line intensity profile and support vector machine for removing reflections in frontal RGB color eye images," *Inf. Sci. (Ny)*, vol. 276, pp. 104–122, 2014.
- [26] D. M. Montserrat, Q. Lin, J. Allebach, and E. J. Delp, "Training object detection and recognition CNN models using data augmentation," *Electron. Imaging*, vol. 29, pp. 27–36, 2017.

Pengembangan Aplikasi Insentif dan Komisi Salesman sebagai Strategi Peningkatan Kinerja SDM Marketing

Suwandi¹, Muhammad Hatta², Turini³, Safitri Akbari⁴, Limbong Aprilina Yanti⁵

^{1,3,5}Komputerisasi Akuntansi, Universitas Catur Insan Cendekia, Indonesia

²Sistem Informasi, Universitas Catur Insan Cendekia, Indonesia

⁴Akuntansi, Universitas Catur Insan Cendekia, Indonesia

Email: ¹suwandi@cic.ac.id, ²muhammad.hatta@cic.ac.id, ³turini@cic.ac.id, ⁴safitri.akbari@cic.ac.id, ⁵limbong.yanti@cic.ac.id

ABSTRAK

Kinerja tenaga penjual (*salesman*) sangat dipengaruhi oleh sistem insentif dan komisi yang diterapkan oleh perusahaan. Sistem yang tidak transparan dan tidak terstruktur dapat menurunkan motivasi serta produktivitas salesman. Penelitian ini bertujuan untuk mengembangkan aplikasi insentif dan komisi *salesman* sebagai strategi peningkatan kinerja SDM marketing. Metode yang digunakan dalam penelitian ini meliputi analisis kebutuhan, perancangan sistem, implementasi, dan pengujian aplikasi. Aplikasi ini dirancang berbasis web untuk memudahkan akses serta meningkatkan transparansi dalam perhitungan insentif dan komisi. Hasil pengujian menunjukkan bahwa implementasi aplikasi ini mampu meningkatkan kepuasan dan motivasi tenaga penjual melalui sistem perhitungan komisi yang lebih adil dan *real-time*. Selain itu, adanya fitur pelacakan target penjualan dan *reward* otomatis berkontribusi dalam meningkatkan kinerja SDM *marketing*. Kesimpulan dari penelitian ini adalah bahwa penggunaan aplikasi berbasis digital dapat menjadi solusi efektif dalam meningkatkan transparansi, kepercayaan, serta motivasi salesman, yang pada akhirnya berdampak positif terhadap produktivitas penjualan.

Kata Kunci: Insentif, Komisi Salesman, Kinerja SDM, Aplikasi, Strategi.

ABSTRACT

The performance of salespeople is greatly influenced by the incentive and commission system implemented by the company. A system that is not transparent and structured can reduce salesman motivation and productivity. This research aims to develop an application for salesman incentives and commissions as a strategy to improve marketing HR performance. The methods used in this research include requirements analysis, system design, implementation and application testing. This application is designed to be web-based to make it easier to access and increase transparency in calculating incentives and commissions. The test results show that the implementation of this application is able to increase sales force satisfaction and motivation through a fairer and real-time commission calculation system. Apart from that, the sales target tracking feature and automatic rewards contribute to improving marketing HR performance. The conclusion of this research is that the use of digital-based applications can be an effective solution in increasing transparency, trust and salesman motivation, which ultimately has a positive impact on sales productivity.

Keywords: Incentives, Salesman Commission, HR Performance, Application, Strategy.

Penulis Korespondensi:

Suwandi

Email: suwandi@cic.ac.id

Article Info

Diterima: 15 Februari 2025

Direvisi: 17 Februari 2025

Disetujui: 18 Februari 2025

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.



1. PENDAHULUAN

PT. Sinar Gunung Jati Cirebon merupakan perusahaan distributor *furniture* dan *spring bed* yang mengandalkan tenaga penjual (*salesman*) sebagai ujung tombak dalam meningkatkan volume penjualan [1]. Dalam industri distribusi, sistem insentif dan komisi berperan penting dalam menjaga semangat kerja serta mendorong kinerja tenaga penjual. Namun, sistem yang kurang transparan dan masih bersifat manual dapat menimbulkan ketidakpuasan di kalangan salesman, yang pada akhirnya berdampak pada produktivitas dan loyalitas mereka terhadap perusahaan [2]. Oleh karena itu, diperlukan suatu inovasi dalam sistem manajemen insentif dan komisi yang lebih efektif dan efisien. Saat ini, PT. Sinar Gunung Jati Cirebon masih menggunakan sistem perhitungan komisi secara konvensional, yang sering kali memunculkan berbagai kendala. Beberapa di antaranya adalah ketidaktepatan dalam perhitungan komisi, keterlambatan dalam pembayaran, serta kurangnya transparansi dalam mekanisme insentif yang diterapkan. Kondisi ini dapat menyebabkan ketidakpercayaan tenaga penjual terhadap sistem yang ada, sehingga menurunkan motivasi mereka dalam mencapai target penjualan yang telah ditetapkan perusahaan [3].

Beberapa penelitian terdahulu menunjukkan bahwa sistem insentif dan komisi yang terstruktur dengan baik dapat meningkatkan motivasi dan produktivitas tenaga penjual [4]. Digitalisasi sistem insentif dan komisi juga telah terbukti mampu meningkatkan efisiensi dan transparansi dalam berbagai sektor bisnis. Berdasarkan permasalahan yang dihadapi oleh PT. Sinar Gunung Jati Cirebon serta referensi dari penelitian sebelumnya, penelitian ini mengusulkan pengembangan aplikasi insentif dan komisi salesman berbasis digital sebagai solusi yang inovatif. Aplikasi ini akan dirancang untuk mempermudah proses perhitungan komisi secara otomatis, mengurangi kesalahan dalam perhitungan, serta meningkatkan transparansi dan keterjangkauan informasi bagi tenaga penjual [5]. Dengan adanya sistem berbasis digital ini, diharapkan tenaga penjual dapat lebih termotivasi karena dapat mengakses informasi komisi dan insentif mereka secara *real-time*.

Pendekatan yang digunakan dalam penelitian ini mencakup analisis kebutuhan, perancangan sistem, pengembangan aplikasi berbasis *web* dan *mobile*, serta pengujian untuk memastikan sistem berjalan dengan optimal [6]. Aplikasi ini akan terintegrasi dengan sistem penjualan PT. Sinar Gunung Jati Cirebon, sehingga setiap transaksi yang dilakukan oleh tenaga penjual dapat tercatat secara otomatis dan langsung dikonversi menjadi komisi sesuai dengan skema yang telah ditetapkan perusahaan [7]. Selain itu, sistem ini akan dilengkapi dengan fitur pelacakan target penjualan dan penghargaan berbasis kinerja untuk meningkatkan motivasi tenaga penjual [8].

Nilai kebaruan dari penelitian ini terletak pada implementasi sistem digital yang tidak hanya mengotomatisasi perhitungan komisi, tetapi juga menghadirkan sistem pelaporan yang lebih akurat dan transparan bagi tenaga penjual [9]. Dengan adanya fitur pelacakan target penjualan dan penghargaan berbasis gamifikasi, aplikasi ini tidak hanya berfungsi sebagai alat administrasi, tetapi juga sebagai instrumen strategis dalam meningkatkan motivasi dan produktivitas tenaga penjual [10]. Penelitian ini diharapkan dapat memberikan manfaat bagi PT. Sinar Gunung Jati Cirebon dalam mengoptimalkan sistem insentif dan komisi, serta menjadi referensi bagi perusahaan distributor lainnya dalam mengembangkan strategi peningkatan kinerja tenaga penjual berbasis teknologi [11]. Dengan sistem yang lebih modern dan efisien, perusahaan dapat membangun lingkungan kerja yang lebih kompetitif dan mendukung pertumbuhan bisnis secara berkelanjutan.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini bertujuan untuk merancang dan mengembangkan aplikasi insentif dan komisi salesman berbasis digital guna meningkatkan transparansi dan efektivitas sistem penghargaan di PT. Sinar Gunung Jati Cirebon. Penelitian ini mengadopsi metode [12] *Research and Development* (R&D) dengan pendekatan *System Development Life Cycle* (SDLC) model *Waterfall*, yang meliputi tahap analisis kebutuhan, perancangan sistem, implementasi, pengujian, serta evaluasi. Berikut adalah penjelasan secara detail mengenai metode penelitian yang digunakan:

2.1. Jenis dan Pendekatan Penelitian

Penelitian ini merupakan penelitian kualitatif dan kuantitatif dengan metode *Research and Development* (R&D), yang bertujuan untuk mengembangkan solusi berbasis teknologi yang dapat menyelesaikan permasalahan sistem insentif dan komisi tenaga penjual. Pendekatan yang digunakan adalah *System Development Life Cycle* (SDLC) model *Waterfall*, yang dilakukan secara bertahap mulai dari analisis hingga implementasi [13].

2.2. Tahapan Penelitian (SDLC Model Waterfall) [14]

a. Analisis Kebutuhan (*Requirement Analysis*) [15]

Tahap ini bertujuan untuk mengidentifikasi permasalahan yang terjadi di PT. Sinar Gunung Jati Cirebon terkait sistem perhitungan insentif dan komisi salesman.

b. Perancangan Sistem (*System Design*)

Perancangan sistem (*System Design*) dalam penelitian ini bertujuan untuk mengembangkan Aplikasi Insentif dan Komisi Salesman yang dapat meningkatkan transparansi, akurasi, dan efisiensi dalam sistem penghargaan tenaga penjual di PT. Sinar Gunung Jati Cirebon.

c. Implementasi (*Development & Implementation*)

Tahap ini bertujuan untuk merealisasikan rancangan sistem ke dalam bentuk perangkat lunak yang dapat digunakan secara efektif oleh tenaga penjual dan manajemen perusahaan. Implementasi dilakukan setelah tahap perancangan sistem selesai, dan bertujuan untuk memastikan bahwa aplikasi berfungsi sesuai dengan kebutuhan bisnis.

d. Pengujian Sistem (*Testing*) [16]

Setelah aplikasi dikembangkan, dilakukan pengujian dengan metode *Black Box Testing* untuk memastikan sistem berjalan sesuai fungsinya.

3. HASIL DAN PEMBAHASAN

PT. Sinar Gunung Jati merupakan perusahaan manufaktur yang bergerak di bidang industri mebel atau *furniture*. PT. Sinar Gunung Jati didirikan pada tahun 1999, namun pada awal berdirinya PT. Sinar Gunung Jati didirikan dengan nama PT. Cahaya Buana Intitama yang merupakan anak perusahaan PT. Cahaya Buana Intitama Bogor. Pada awal berdirinya PT. Sinar Gunung Jati bergerak di bidang penjualan atau pemasaran *furniture* plastik, panel, kasur springbed dan kasur busa. Beberapa produk yang dihasilkan adalah kasur busa, kasur twin, dan springbed yang terdiri dari sandaran, matras, dan divan. Diantara beberapa jenis produk yang dijelaskan di atas terdapat tiga merek, yaitu Bigland, Bigdream dan Marcell/buana.

3.1. Syarat Pemberian Komisi

1. Pemberian komisi di PT. Sinar Gunung Jati hanya diberikan kepada *salesman*.
2. *Salesman* melakukan transaksi penjualan mencapai atau melebihi target bulanan yang dibuat oleh divisi marketing.
3. *Salesman* melakukan penagihan atas konsumen tokonya paling lama 60 hari.
4. Sesuai kebijakan perusahaan jumlah komisi yang diterima tidak bisa melebihi 1% dari realisasi jumlah nominal penjualan barang.
5. Terdapat perbedaan skala untuk menghitung komisi *salesman* yang didapat sesuai target penjualan dan target uang yang cair dari hasil penagihan sebagai berikut:

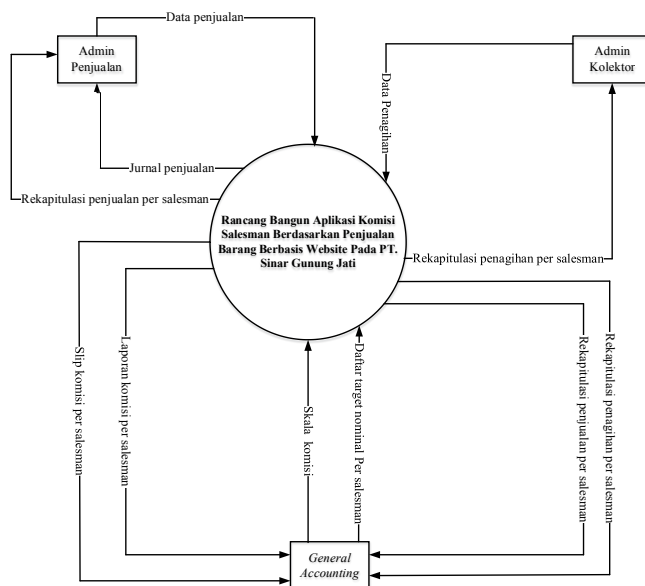
Tabel 1. Skala Komisi Penjualan dan Penagihan

Keterangan	Pencapaian			
	>= 80%	>= 90%	>= 100%	>= 120%
Pencapaian penjualan per kategori produk				
Plastik	0,09	0,12	0,15	0,20
Springbed	0,10	0,13	0,16	0,20
Busa/Panel	0,12	0,16	0,20	0,25
Umum	0,02	0,03	0,04	0,05
Komisi Penagihan	31-60 HR	22-30 HR	15-21 HR	01-14 HR
Lama Pembayaran	0,15	0,20	0,25	0,30

3.2. Diagram Konteks [17]

Naratif *diagram konteks*:

- a. Admin Penjualan memasukkan data penjualan ke dalam sistem dan dikelola datanya lalu menghasilkan keluaran berupa rekapitulasi penjualan per *salesman*.
- b. Admin Kolektor memasukkan data penagihan ke dalam sistem dan dikelola datanya untuk menghasilkan rekapitulasi penagihan per *salesman* dan Jurnal Penjualan.
- c. *General Accounting* memasukkan daftar target penjualan per *salesman* ke sistem lalu dari sistem tarik data rekapitulasi penjualan per *salesman* dan rekapitulasi penagihan per *salesman*, selanjutnya data yang sudah di-*input*-kan menghasilkan keluaran berupa laporan komisi per *salesman* dan slip komisi per *salesman*.



Gambar 1. Diagram Konteks Rancang Bangun Aplikasi Perhitungan Komisi Salesman

3.3. Tampilan Program

a. Tampilan form login

Form login merupakan form utama yang digunakan user atau pengguna untuk dapat mengakses aplikasi sesuai dengan akses yang dimiliki setiap user



Gambar 2. Form Login User

b. Tampilan form Data Penjualan

Form di bawah ini digunakan untuk menampilkan data penjualan yang sebelumnya telah di-input oleh Admin Penjualan.



Gambar 3. Tampilan Form Tampil Data Penjualan

c. Tampilan form Rekap Penjualan

Tampilan Rekapitulasi Penjualan dalam Aplikasi Insentif dan Komisi *Salesman* di PT. Sinar Gunung Jati Cirebon memiliki peran penting dalam menyajikan data penjualan secara terstruktur dan mudah dipahami. Fitur ini dirancang untuk membantu tenaga penjual, manajemen, dan tim keuangan dalam memantau, mengevaluasi, dan mengelola informasi penjualan serta komisi secara transparan dan akurat.



Gambar 4. Form Tampilan Rekapitulasi Penjualan

d. Tampilan form Komisi Penjualan

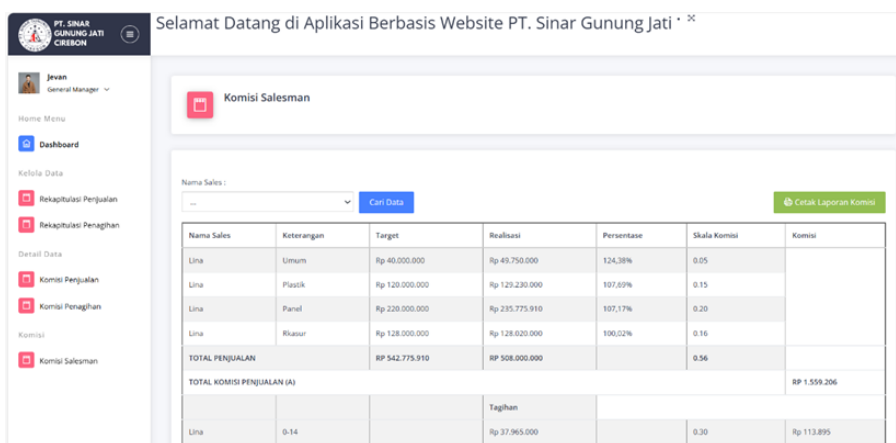
Tampilan Data Komisi Penjualan dalam Aplikasi Insentif dan Komisi *Salesman* di PT. Sinar Gunung Jati Cirebon memiliki peran penting dalam menyediakan informasi komisi yang akurat, transparan, dan mudah diakses bagi *salesman*, manajemen, dan tim keuangan. Fitur ini dirancang untuk memastikan bahwa setiap tenaga penjual dapat mengetahui secara jelas besaran komisi yang mereka terima berdasarkan kinerja penjualan mereka.



Gambar 5. Tampilan Data Komisi Penjualan

e. Tampilan Laporan Komisi Penjualan

Aplikasi Laporan Komisi *Salesman* di PT. Sinar Gunung Jati Cirebon dikembangkan untuk mengelola dan menyajikan informasi terkait perhitungan komisi tenaga penjual secara otomatis, akurat, dan transparan. Sistem ini tidak hanya mempermudah perhitungan komisi tetapi juga meningkatkan efisiensi dan akurasi dalam pengelolaan insentif tenaga penjual.



Gambar 6. Tampilan Laporan Komisi Salesman

f. Output Laporan Komisi

Output ini dibuat untuk memastikan bahwa setiap *salesman* mengetahui komisi yang diperoleh berdasarkan kinerja masing-masing. Selain itu juga untuk mencegah kesalahpahaman atau sengketa terkait perhitungan insentif, karena data komisi ditampilkan secara jelas dan sistematis. Dengan demikian, manajemen dapat melihat performa seluruh tim penjualan dalam satu periode, baik individu maupun secara kelompok. Hal ini dapat membantu dalam identifikasi *salesman* terbaik dan yang memerlukan pelatihan lebih lanjut.

PT.SINAR GUNUNG JATI								
Jalan Jend. Sudirman No. 238, Hargamak, Kec. Hargamak, Kota Cirebon, Jawa Barat 45143								
LAPORAN KOMISI								
Nama Sales	Keterangan	Realisasi	Target	Persentase	Skala Komisi	Komisi		
Lina	Lunas	RP 49.750.000	RP 40.000.000	124,38%	0,05			
Lina	Pasrah	RP 129.230.000	RP 120.000.000	107,69%	0,13			
Lina	Pasrah	RP 235.775.910	RP 220.000.000	107,17%	0,20			
Lina	Rikara	RP 128.020.000	RP 128.000.000	100,02%	0,16			
TOTAL PENJUALAN		RP 542.775.910	RP 508.000.000		0,56			
TOTAL KOMISI PENJUALAN (A)							RP 1.559.206	
Lina	0-14		Tagihan					
Lina	0-14		RP 37.863.000		0,30	RP 113.895		
Lina	15-21		RP 32.710.870		0,25	RP 81.777		
Lina	22-30		RP 6.990.600		0,30	RP 11.981		
Lina	31-60		RP 200.761.230		0,13	RP 301.145		
TOTAL PENAGIHAN							RP 278.429.720	
TOTAL KOMISI PENAGIHAN (B)							RP 516.798	
JUMLAH KOMISI (A-B)							RP 2.076.005	
DEPOSITO SALESMAN (20%)							RP 414.001	
JUMLAH KOMISI							RP 1.662.004	
INSENTIF GUDANG (5%)							RP 82.800	
JUMLAH KOMISI YANG DITERIMA							RP 1.579.204	

Cirebon, 14 September 2024
 General Accounting.

 Devi Yenny Suningsih, S.Kom

Gambar 7. Tampilan Output Komisi Salesman

4. KESIMPULAN

Penelitian ini telah berhasil mengembangkan aplikasi insentif dan komisi salesman sebagai strategi peningkatan kinerja SDM marketing di PT. Sinar Gunung Jati Cirebon. Sistem ini dirancang untuk mengotomatiskan perhitungan komisi berbasis penjualan, sehingga memberikan transparansi yang lebih tinggi, mengurangi kesalahan administrasi, dan meningkatkan efisiensi dalam pengelolaan insentif tenaga penjual. Dengan penerapan sistem ini, perusahaan dapat mengoptimalkan strategi pemberian insentif guna meningkatkan motivasi dan produktivitas salesman. Penggunaan aplikasi ini menjadikan perhitungan komisi secara otomatis, yang menghilangkan risiko kesalahan dalam pencatatan manual. Laporan komisi yang dihasilkan bersifat *real-time* dan terstruktur, sehingga memudahkan manajemen dalam melakukan evaluasi terhadap kinerja individu maupun tim sales. Dengan adanya sistem ini, tenaga penjual dapat melihat langsung perolehan komisi mereka, sehingga meningkatkan transparansi dan mendorong persaingan sehat dalam tim.

Dari hasil implementasi, dapat disimpulkan bahwa aplikasi ini mampu meningkatkan efisiensi operasional perusahaan dalam pengelolaan insentif. Selain itu, sistem ini juga memberikan dampak positif terhadap motivasi dan kepuasan tenaga penjual, karena mereka mendapatkan penghargaan yang sesuai dengan pencapaian kinerjanya. Dengan demikian, aplikasi ini berkontribusi pada peningkatan efektivitas strategi pemasaran perusahaan. Secara keseluruhan, penelitian ini menunjukkan bahwa digitalisasi dalam pengelolaan komisi dapat meningkatkan transparansi, akurasi, dan efisiensi dalam sistem remunerasi tenaga penjual. Implementasi aplikasi ini tidak hanya menguntungkan tenaga penjual, tetapi juga memberikan manfaat bagi perusahaan dalam mengoptimalkan kinerja bisnis dan pengambilan keputusan berbasis data. Oleh karena itu, aplikasi ini dapat dijadikan model dalam pengelolaan insentif tenaga penjual di industri sejenis.

Penelitian lebih lanjut dapat dilakukan untuk menganalisis dampak penerapan sistem ini terhadap retensi dan loyalitas tenaga penjual dalam jangka panjang. Selain itu, studi lanjutan bisa mengeksplorasi model insentif berbasis performa dan kepuasan pelanggan, serta meneliti bagaimana digitalisasi pengelolaan komisi berdampak pada produktivitas tenaga penjual di berbagai sektor industri.

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih yang sebesar-besarnya kepada Universitas Catur Insan Cendekia (UCIC) Cirebon, khususnya kepada para dosen dan civitas akademika yang telah memberikan dukungan, bimbingan, serta fasilitas dalam pelaksanaan penelitian ini. Penghargaan yang setinggi-tingginya juga kami sampaikan kepada PT. Sinar Gunung Jati Cirebon sebagai objek penelitian, yang telah memberikan kesempatan, data, serta informasi yang diperlukan dalam pengembangan sistem ini. Terima kasih atas kerja sama, kepercayaan, dan kontribusi yang sangat berarti dalam mendukung keberhasilan penelitian ini.

REFERENSI

- [1] F. Khasanah, M. Hatta, and S. Suwandi, "ANALISIS SENTIMEN PADA ULASAN PEMBELIAN PRODUK DENGAN METODE SUPPORT VECTOR MACHINE," *JSR: Jaringan Sistem Informasi Robotik*, vol. 8, no. 2, pp. 198–203, 2024.
- [2] S. Suwandi, A. H. A. N. Karsa, and M. Asfi, "Rancang Bangun Aplikasi Manajemen Penjualan Konsyinesi pada Perusahaan Dagang," *Jurnal Syntax Admiration*, vol. 5, no. 4, pp. 1240–1248, 2024.
- [3] R. C. IYADI and S. ISIKADI, "Salesmanship and Sales Performance of Selected Organizations," *World*, vol. 2579, p. 0544, 2023.
- [4] N. W. F. A. Lestari, S. Suwandi, T. Turini, and B. Dewansyah, "EFEKTIFITAS IMPLEMENTASI PENGEMBANGAN SDM PADA PRODUKTIVITAS KARYAWAN BAGIAN PENJUALAN," *Jurnal EBI*, vol. 6, no. 2, pp. 24–29, 2024.
- [5] F. Hasna, S. Zulaihati, and A. Sumiati, "The Effect of Capital Intensity Ratio and Sales Growth to Tax Avoidance with Independent Commissioners as Moderating Variable (Empirical Study on Consumer Goods Industry Companies on the Indonesia Stock Exchange in 2021-2022): English," *Nexus Synergy: A Business Perspective*, vol. 1, no. 2, pp. 72–83, 2023.
- [6] P. Xiao, "Design and Implementation of HTML5-based Mobile Sales Management System," in *2023 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, IEEE, 2023, pp. 400–405.
- [7] S. G. Thomson, R. McMorran, and J. P. G. Glendinning, "Rural Land Market Data Report: Analysis of land sales data and proposals for improving future reporting of land market transactions A report to the Scottish Land Commission," 2022.
- [8] M. S. Kim, A. Haslam, and V. Prasad, "Trend of sales revenue by year for top selling cancer drugs in the US and the effect of loss of market exclusivity," *J Cancer Policy*, vol. 43, p. 100533, 2025.
- [9] J. Liu, "Joint Decision of Price Discount, Sales Commission, and LS Effort Based on Numerical Simulation," in *Proceedings of the 2023 4th International Conference on Computer Science and Management Technology*, 2023, pp. 834–839.
- [10] M. M. Lastner *et al.*, "Salesperson motivation, compensation, training and deployment within the sales ecosystem," *Eur J Mark*, vol. 58, no. 3, pp. 812–841, 2024.
- [11] P. Gautier, A. Siegmann, and A. van Vuuren, "Real-estate agent commission structure and sales performance," *J Empir Finance*, vol. 72, pp. 163–187, 2023.
- [12] R. Darmayanti, "Research Design: RnD Research Method," *Available at SSRN 5012967*, 2024.
- [13] A. K. Samaratova, "FEATURES OF USING WATERFALL METHODOLOGIES IN PROJECT MANAGEMENT," *Вестник науки*, vol. 4, no. 10 (79), pp. 39–48, 2024.
- [14] D. B. Aniley, E. A. Jalew, and G. A. Agegnehu, "Selection of Software Development Life Cycle Models using Machine Learning Approach," *Int J Comput Appl*, vol. 186, pp. 36–43, 2024.
- [15] W. D. Febrian and B. Y. Geni, "PENERAPAN TEKNOLOGI BIG DATA DALAM ANALISIS KEBUTUHAN TENAGA KERJA DAN PERENCANAAN SUKSESI ORGANISASI," *EDUSAINTEK: Jurnal Pendidikan, Sains Dan Teknologi*, vol. 11, no. 3, pp. 1309–1319, 2024.
- [16] M. Jibril, "PENGUJIAN SISTEM INFORMASI E-MODUL PADA SMPN 1 TEMPULING MENGGUNAKAN BLACK BOX TESTING," *Jurnal Perangkat Lunak*, vol. 6, no. 2, pp. 327–332, 2024.
- [17] Y. Salamor, J. M. L. Wakim, J. C. Wattimena, D. R. Tisera, and G. I. Behuku, "Mengoptimalkan Arus Keuangan: Pendekatan Diagram Konteks dan DFD level 0 Pada Desain Siklus Pendapatan," *Jurnal Tagalaya Pengabdian Kepada Masyarakat*, vol. 1, no. 2, pp. 119–124, 2024.

Analisis *Average Waiting Time* Penjadwalan CPU Menggunakan Algoritma *Shortest Remaining First* dan Algoritma Round Robin

Ronald Belferik¹, Evander Banjarnahor²,

¹ Informatika, Fakultas Teknologi Informasi, Universitas Pelita Harapan, Indonesia

² Matematika, Fakultas Sains dan Teknologi, Universitas Pelita Harapan, Indonesia

Email: ¹ronald.belferik@gmail.com, ²evander.banjarnahor@gmail.com,

ABSTRAK

Dalam sistem operasi, penjadwalan proses merupakan aspek kritis untuk menentukan urutan eksekusi proses oleh CPU. Penelitian ini melakukan perbandingan waktu tunggu rata – rata atau *average waiting time* (AWT) pada algoritma *Shortest Remaining First* (SRF) dan algoritma *Round Robin* (RR) dimana masalah yang hendak diselesaikan adalah penjadwalan pada CPU. Tujuan dari penelitian ini adalah mendapatkan algoritma yang memiliki rata – rata waktu tunggu yang singkat. Hasil pengujian diperoleh bahwa algoritma SRF memiliki waktu tunggu rata – rata yang sangat singkat dengan nilai 29.85 ms dibanding algoritma RR yang mendapatkan hasil AWT 65.6 ms.

Kata Kunci: *Shortest Remaining First, Round Robin, Average Waiting Time*

ABSTRACT

In operating systems, process scheduling is a critical aspect to determine the order of process execution by the CPU. This research compares the average waiting time (AWT) of *Shortest Remaining First* (SRF) algorithm and *Round Robin* (RR) algorithm where the problem to be solved is CPU scheduling. The purpose of this research is to get an algorithm that has a short average waiting time. The test results obtained that the SRF algorithm has a very short average waiting time with a value of 29.85 ms compared to the RR algorithm which gets an AWT result of 65.6 ms.

Keywords: *Shortest Remaining First, Round Robin, Average Waiting Time*

Penulis Korespondensi:

Ronald Belferik

Email: ronald.belferik@uph.edu

Article Info

Diterima: 19 Februari 2025

Direvisi: 23 Februari 2025

Disetujui: 25 Februari 2025

This is an open access article under the [CC BY](#) license.



1. PENDAHULUAN

Di era komputasi modern, penggunaan sumber daya komputasi yang efisien menjadi sangat penting untuk memastikan kinerja sistem yang optimal. Salah satu sumber daya yang penting di dalam sistem komputer adalah *Central Processing Unit* (CPU), yang bertanggung jawab untuk mengeksekusi instruksi-instruksi program. Dalam sistem komputer yang menjalankan banyak tugas (*multitasking*), penjadwalan CPU memainkan peran penting dalam menentukan urutan proses-proses yang bersaing untuk mendapatkan akses ke CPU untuk dieksekusi [1]-[9].

Ketika prosesor tidak menjalankan tugas (kondisi *idle*), sistem operasi memilih tugas dari memori utama untuk dijalankan dan menugaskan prosesor untuk menjalankan tugas tersebut. Kebutuhan untuk memprogram prosesor untuk menjalankan tugas di memori utama untuk memastikan kelangsungan sistem operasi adalah masalah yang perlu ditangani dan diselesaikan. Ini adalah hasil dari tidak mengoptimalkan kinerja CPU. Tujuan dari *multiprogramming* adalah untuk meminimalkan waktu tidak aktif (*idle*) CPU saat menjalankan beberapa proses [10]. Untuk memaksimalkan penggunaan CPU, beberapa proses disimpan dalam memori dan CPU menggunakan waktu yang dialokasikan untuk menjalankan satu proses [11]. Dengan cara ini, waktu menganggur dapat dikurangi dan sistem operasi dapat meningkatkan produktivitas komputer dengan menjadwalkan alokasi

waktu untuk setiap proses dalam sistem [8]. Dalam sebuah sistem operasi, terdapat banyak proses yang perlu dieksekusi, sehingga timbul masalah mengenai proses mana yang harus dieksekusi dalam sistem. Oleh karena itu, perlu adanya suatu pedoman dan mekanisme mengenai urutan kerja sistem komputer di dalam sistem operasi. Untuk mengatasi masalah ini, telah dikembangkan algoritma penjadwalan yang menentukan proses mana yang dieksekusi CPU terlebih dahulu. Proses yang belum menerima alokasi dari CPU ditempatkan dalam antrian. Proses tersebut kemudian dieksekusi.

Dalam konteks ini, penulis ingin melakukan penelitian yang berjudul “Analisis *Average Waiting Time* (AWT) Penjadwalan CPU Menggunakan Algoritma *Shortest Remaining First* dan Algoritma *Round Robin*”. Penelitian ini menganalisis dan membandingkan kinerja algoritma prioritas waktu tunggu terpendek dan algoritma *Round Robin* dalam hal waktu tunggu rata-rata (AWT) di bawah beban CPU yang berbeda. Dengan membandingkan kedua algoritma ini, diharapkan dapat memberikan wawasan tentang *trade-off* antara efisiensi dan *fairness* dalam penjadwalan CPU.

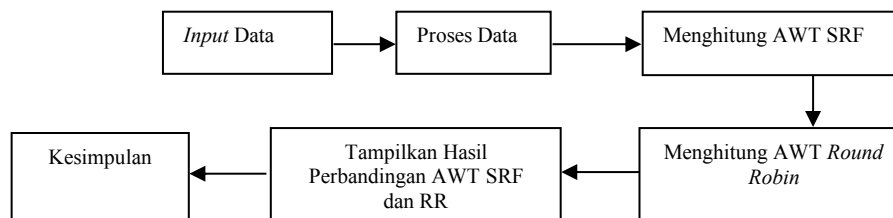
Penelitian ini akan melibatkan implementasi dan simulasi kedua algoritma penjadwalan CPU menggunakan data beban kerja sintetis. Hasil simulasi akan dianalisis untuk mengidentifikasi faktor-faktor yang mempengaruhi AWT pada masing-masing algoritma.

2. METODE PENELITIAN

Pada penelitian ini dilakukan perbandingan *Average Waiting Time* (AWT) Pada algoritma *Shortest Remaining First* dan algoritma *Round Robin*. Dalam Algoritma *Shortest Remaining First*, proses dengan sisa waktu paling sedikit untuk diselesaikan dipilih untuk dijalankan. Proses yang sedang berjalan akan terus berjalan hingga selesai atau proses baru dengan sisa waktu yang lebih pendek tiba. Dengan cara ini, proses yang dapat menyelesaikan paling cepat akan selalu diprioritaskan [12]. Algoritma *Round Robin* merupakan penjadwalan yang paling tua, sederhana, adil, banyak digunakan algoritmanya dan mudah diimplementasikan. Penjadwalan ini bukan dijalankan oleh proses lain tetapi oleh penjadwalan berdasarkan lama waktu berjalannya proses (*preempt by time*). Penjadwalan tanpa prioritas berasumsi bahwa semua proses memiliki kepentingan yang sama, sehingga tidak ada prioritas tertentu. Semua proses dianggap penting sehingga diberi sejumlah waktu oleh pemroses yang disebut kuantum (*quantum*) atau *time slice* dimana proses itu berjalan. Jika proses masih *running* sampai akhir *quantum*, maka CPU akan mem-*preempt* proses itu dan memberikannya ke proses lain [13]. Pada penelitian ini metode penelitian yang digunakan ialah *library search* dan *experiment*. Pada *library search*, maka akan dicari penelitian terkait untuk dikembangkan dan *experiment* yaitu menguji coba algoritma penjadwalan dengan mencari AWT pada tiap algoritma dan lalu membandingkannya.

2.1. Bagan Penelitian

Pada bagan penelitian, *input* data berupa contoh data penjadwalan yang akan diuji. Selanjutnya data tersebut diolah dengan menggunakan algoritma SRF dan RR. Hasil proses adalah AWT masing-masing algoritma. Setelah itu diambil kesimpulan AWT yang terbaik dari algoritma yang digunakan.



Gambar 1. Bagan penelitian

2.2. Alur Penelitian

Alur penelitian yang dilakukan adalah mengumpulkan *literature*. Melakukan uji coba pada algoritma tersebut untuk menghasilkan *average waiting time* pada algoritma penjadwalan *Shortest Remaining First* dan *Round Robin*. Jumlah proses data antrian dilakukan secara dinamis yang akan diproses oleh kedua algoritma tersebut, data antrian yang akan diproses oleh masing-masing kedua algoritma tersebut ditentukan berdasarkan *waiting time* yang telah ditentukan, sehingga proses data antrian diproses secara dinamis yang akan diharapkan mendapatkan nilai *average waiting time* antar kedua algoritma tersebut.

2.3. Penentuan Proses Data Antrian

Proses data antrian akan dilakukan secara berurutan oleh kedua algoritma penjadwalan tersebut yaitu algoritma *Shortest Remaining First* dan *Round Robin*. Berikut dapat dijelaskan proses data antrian yang dilakukan oleh kedua algoritma penjadwalan tersebut:

1. Proses Data *Shortest Remaining First* (SRF)

Pada algoritma ini seluruh proses yang berada pada *ready queue* akan diproses berdasarkan *burst time* yang terkecil. Hal ini menyebabkan setiap proses akan memiliki *waiting time* yang singkat. Dengan setiap proses yang

memiliki *waiting time* yang singkat, maka *average waiting time* juga akan menjadi semakin pendek. Oleh karena itu, algoritma ini dapat dikatakan sebagai algoritma yang optimal [10].

2. Proses Data Antrean *Round Robin* (RR)

Proses data antrean pada algoritma penjadwalan *Round Robin* merupakan algoritma yang menggunakan sistem *time sharing*, dimana setiap proses data yang mengantre pada algoritma ini akan mendapatkan bagian di proses berdasarkan *quantum time* yaitu jika proses data antrean melewati dari *quantum time* maka proses tersebut akan melakukan antrean lagi dari awal [14].

3. Menghitung *Average Waiting Time* (AWT)

Pada proses perhitungan *waiting time* ini, setiap proses data antrean yang diproses oleh ketiga algoritma tersebut akan dihitung masing-masing *waiting time* dan total *waiting time*-nya yang dipengaruhi oleh *arrival time* dan *burst time* pada suatu data antrean tersebut. Proses dalam mencari *waiting time* pada suatu data antrean adalah waktu eksekusi dikurangi dengan waktu tiba, sedangkan dalam mencari *average waiting time* pada suatu data antrean adalah dengan cara menjumlahkan seluruh *waiting time* yang ada pada seluruh proses data antrean dan dibagi dengan seluruh jumlah proses data antrean.

2.4. Pseudocode Algoritma SRF

Langkah – langkah *source code* dalam mencari AWT pada algoritma SRF yaitu:

- * *Traverse until all process gets completely executed.*
 - > *Find process with minimum remaining time at every single time lap.*
 - > *Reduce its time by 1.*
 - > *Check if its remaining time becomes 0*
 - > *Increment the counter of process completion.*
 - > *Completion time of current process = current_time + 1;*
 - > *Calculate waiting time for each completed process.*
$$wt[i] = \text{Completion time} - \text{arrival_time} - \text{burst_time}$$
 - > *Increment time lap by one.*
- * *Find turnaround time (waiting_time + burst_time).*

2.5. Tampilan Program Algoritma SRF

Pada tampilan program Algoritma RR menampilkan data 20 proses, yaitu dari P1 hingga P20, beserta waktu kedatangan (*Arrival Time*) dan waktu burst (*Burst Time*) masing-masing proses. Waktu kedatangan menunjukkan kapan proses masuk ke dalam sistem, sedangkan waktu *burst* menunjukkan durasi waktu yang dibutuhkan proses untuk dieksekusi oleh CPU. Setelah data proses diolah menggunakan suatu algoritma SRF, diperoleh hasil perhitungan *Average Waiting Time* (AWT) atau rata-rata waktu tunggu: 35.7 satuan waktu. AWT adalah rata-rata waktu yang dihabiskan setiap proses dalam antrean sebelum dieksekusi. Semakin kecil nilai AWT, semakin efisien algoritma penjadwalan dalam meminimalkan waktu tunggu proses. Kemudian juga diperoleh hasil *Average Turnaround Time* (ATT): 43.1 satuan waktu. ATT adalah rata-rata total waktu yang dibutuhkan setiap proses, mulai dari saat kedatangan hingga selesai dieksekusi. AAT mencakup waktu tunggu dan waktu *burst*. Semakin kecil nilai AAT, semakin efisien algoritma penjadwalan dalam menyelesaikan proses secara keseluruhan.

```
Enter no. of processes : 20
PROCESSES      ARRIVAL TIME      BURST TIME
P1              6                  3
P2              15                 17
P3              15                 13
P4              12                  6
P5              1                  9
P6              7                   2
P7              19                 10
P8              6                   3
P9              6                   0
P10             16                 12
P11             0                  11
P12             9                   7
P13             10                  2
P14             3                   2
P15             15                  7
P16             2                   9
P17             18                  2
P18             7                   9
P19             16                 13
P20             2                   11
Average waiting time : 35.7
Average turn around time : 43.1
```

Gambar 2. Tampilan program algoritma SRF

2.6. Pseudocode Algoritma RR

Langkah-langkah *source code* dalam mencari AWT pada algoritma RR yaitu:

- CPU scheduler picks the process from the circular/ready queue, set a timer to interrupt it after 1 time slice / quantum and dispatches it .*

- * If process has burst time less than 1 time slice/quantum
 - > Process will leave the CPU after the completion
 - > CPU will proceed with the next process in the ready queue / circular queue .
- else If process has burst time longer than 1 time slice/quantum
 - > Timer will be stopped. It cause interruption to the OS .
 - > Executed process is then placed at the tail of the circular / ready queue by applying the context switch.
 - > CPU scheduler then proceeds by selecting the next process in the ready queue.

2.7. Tampilan Program Algoritma RR

Pada tampilan program Algoritma RR menampilkan data 20 proses, yaitu dari P1 hingga P20, beserta waktu kedatangan (*Arrival Time*) dan waktu burst (*Burst Time*) masing-masing proses. Waktu kedatangan menunjukkan kapan proses masuk ke dalam sistem, sedangkan waktu *burst* menunjukkan durasi waktu yang dibutuhkan proses untuk dieksekusi oleh CPU. Setelah data proses diolah menggunakan algoritma RR, diperoleh hasil perhitungan *Average Waiting Time* (AWT) atau rata-rata waktu tunggu: 112.55 satuan waktu. AWT adalah rata-rata waktu yang dihabiskan setiap proses dalam antrian sebelum dieksekusi. Semakin kecil nilai AWT, semakin efisien algoritma penjadwalan dalam meminimalkan waktu tunggu proses. Kemudian juga diperoleh hasil *Average Turnaround Time* (ATT): 122.2 satuan waktu. ATT adalah rata-rata total waktu yang dibutuhkan setiap proses, mulai dari saat kedatangan hingga selesai dieksekusi. AAT mencakup waktu tunggu dan waktu *burst*. Semakin kecil nilai AAT, semakin efisien algoritma penjadwalan dalam menyelesaikan proses secara keseluruhan.

```

Enter Number Of Process: 20
Quantum: 3
Processes  Burst time  Waiting time  Turn around time
1           6           53           59
2           15          151          166
3           15          154          169
4           12          135          147
5           1           12           13
6           7           107          114
7           19          174          193
8           6           71           77
9           6           74           80
10          16          172          188
11          8           114          122
12          9           116          125
13          10          144          154
14          3           37           40
15          15          163          178
16          2           43           45
17          18          173          191
18          7           128          135
19          16          176          192
20          2           54           56
Average waiting time = 112.55
Average turn around time = 122.2
    
```

Gambar 3. Tampilan program algoritma RR

3. HASIL DAN PEMBAHASAN

Pada penelitian ini yang menjadi *input* adalah jumlah proses, *arrival time* dan *burst time* dari tiap proses. Cara *input* proses adalah dengan melakukan pemanggilan *file* data antrian yang sudah ditentukan proses_id, *arrival time* dan *burst time* yang telah di simpan dalam *storage*. Jika *file* proses telah di *input*, maka proses-proses tersebut akan dikerjakan dengan menggunakan algoritma *Shortest Remaining First* dan didapatkan *waiting time* dari setiap proses yang dikerjakan. Setelah itu data tersebut digunakan kembali dengan menggunakan algoritma *Round Robin* untuk menghitung *waiting time* dari setiap proses. Hal ini dilakukan sampai semua proses benar-benar selesai dilayani oleh CPU. Setelah semua proses selesai dilayani oleh CPU, maka akan ditampilkan proses data-data antrian yang dikerjakan dari kedua algoritma penjadwalan tersebut, *waiting time* pada setiap proses dan AWT seluruh proses.

3.1. Uji Coba Dengan 20 Data

Pada pengujian yang dilakukan penulis dengan parameter *input* adalah jumlah proses sebanyak 20 proses, *arrival time* dan *burst time* dari tiap proses dengan menggunakan proses antrian *single queue* yang dikerjakan oleh kedua algoritma penjadwalan. Proses yang akan dieksekusi oleh CPU adalah seperti pada Tabel 1 berikut ini :

Tabel 1. Data antrian dengan jumlah 20 proses

Proses_ID	Arrival Time	Burst Time
P_1	0	8
P_2	2	7
P_3	5	10
P_4	11	6

Proses_ID	Arrival Time	Burst Time
P_5	16	4
P_6	19	9
P_7	25	5
P_8	30	9
P_9	30	7
P_10	30	5
P_11	30	4
P_12	34	9
P_13	36	11
P_14	39	5
P_15	43	8
P_16	51	9
P_17	58	6
P_18	58	8
P_19	58	9
P_20	58	12

Dari tabel di atas, akan dilakukan proses data antrean untuk mencari *average waiting time* pada setiap masing-masing atau ketiga algoritma penjadwalan tersebut dengan menggunakan *single queue*. Disini dapat dijelaskan beberapa proses data antrean, dimana *arrival time* adalah waktu ketika proses berada di memori utama sebelum proses tersebut dikerjakan oleh CPU. *Burst Time* mengacu pada alokasi lamanya waktu eksekusi yang telah dialokasikan kepada masing-masing proses sejak proses itu dibuat. Dalam sistem komputer, setiap proses memerlukan waktu untuk dieksekusi oleh CPU. *Burst Time* mengukur lamanya waktu tersebut [15]. Setelah dari penjelasan di atas, maka proses data antrean dari masing-masing kedua algoritma penjadwalan tersebut.

3.2. Hasil Pengujian 20 Proses Dengan Algoritma *Shortest Remaining First*

Pada algoritma SRF, seluruh proses yang berada pada *ready queue* akan diselesaikan berdasarkan proses yang memiliki *burst time* terkecil. Dikarenakan algoritma SRF menyelesaikan suatu proses berdasarkan proses yang memiliki *burst time* terkecil, maka hal ini akan berdampak pada *waiting time* yang singkat pada seluruh prosesnya. Dengan *waiting time* yang singkat dari setiap proses, maka *average waiting time* pada algoritma SRF juga menjadi singkat. Hal ini membuat algoritma SRF dapat dikatakan salah satu algoritma yang optimal.

Hal yang pertama yang harus dilakukan adalah menggambarkan kronologi eksekusi proses pada Table 1 di atas dengan *Gantt Chart*. Mengenai hasil *Gantt Chart* dari Tabel 1 di atas dapat dijelaskan proses data antrean satu per satu. Pada saat CPU tidak mengerjakan sesuatu atau dalam posisi 0 datang sebuah proses P1 yang membutuhkan waktu penyelesaian atau disebut *burst time* yang berjumlah 8, karena algoritma penjadwalan SRF merupakan algoritma yang melakukan proses berdasarkan *burst time* terkecil, maka proses selanjutnya ialah mengecek *burst time* P2, P3, P4 dan seterusnya jika sudah tiba. Jika *burst time* pada proses P2, P3, P4 dan seterusnya lebih kecil dari *burst time* P1 yang sedang diproses, maka proses yang memiliki *burst time* terkecil harus diproses terlebih dahulu dan P1 masuk ke status menunggu sampai dilayani. Jika semua proses telah selesai dikerjakan, maka langkah selanjutnya ialah menghitung waktu *average waiting time* dari seluruh proses yang telah dikerjakan.

1. Gantt Chart

P_1	P_2	P_4	P_5	P_4	P_7	P_11
0	8	15	16	20	25	30
P_10	P_14	P_9	P_15	P_17	P_18	P_6
34	39	44	51	59	65	73
P_8	P_12	P_16	P_19	P_3	P_13	P_20
82	91	100	109	118	128	139
						151

Gambar 4. *Gantt chart* 20 data dengan algoritma *shortest remaining first*

2. Average Waiting Time

Hal yang dilakukan untuk menghitung *average waiting time* adalah dengan menghitung berapa lama waktu tunggu atau *waiting time* yang dihasilkan oleh proses data antrean tersebut dan dibagi dari jumlah seluruh proses. Untuk menghitung *waiting time* dapat dilakukan dengan cara mengurangi waktu eksekusi dengan waktu datang. Hasil *waiting time* dari setiap proses data antrean pada Gambar 4 adalah seperti tabel di bawah ini:

Tabel 2. *Waiting time* dengan *shortest remaining first*

Proses	Waiting Time(ms)
P_1	0
P_2	6
P_3	113
P_4	8
P_5	0
P_6	54
P_7	0
P_8	52
P_9	14
P_10	4
P_11	0
P_12	57
P_13	92
P_14	0
P_15	8
P_16	49
P_17	1
P_18	7
P_19	51
P_20	81

Dari tabel 2. di atas maka *average waiting time* (AWT) dapat dihitung yaitu :

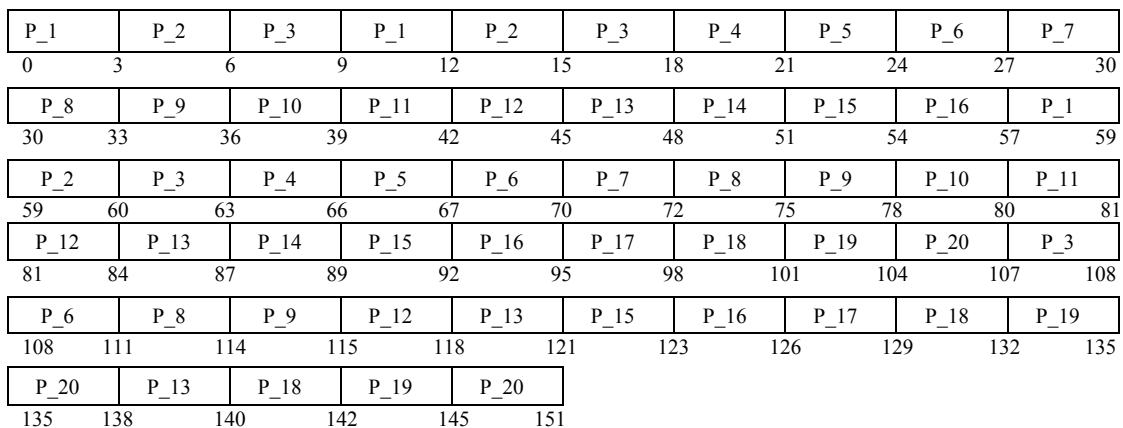
$$\begin{aligned}
 AWT &= (WT P_1 + WT P_2 + WT P_3 + WT P_4 + WT P_5 + WT P_6 + WT P_7 + WT P_8 + WT P_9 + WT P_{10} + \\
 &WT P_{11} + WT P_{12} + WT P_{13} + WT P_{14} + WT P_{15} + WT P_{16} + WT P_{17} + WT P_{18} + WT P_{19} + WT P_{20}) \\
 &/ 20 \\
 &= (0 + 6 + 113 + 8 + 0 + 54 + 0 + 52 + 14 + 4 + 0 + 57 + 92 + 0 + 8 + 49 + 1 + 7 + 51 + 81) / 20 \\
 &= \mathbf{29.85ms}
 \end{aligned}$$

Pada hasil waktu tunggu rata-rata atau *average waiting time* di atas didapat dengan cara menjumlahkan seluruh *waiting time* pada proses P1 sampai proses P20 dan dibagi dengan jumlah seluruh proses data antrean.

3.3. Hasil Pengujian 20 Proses Dengan Algoritma RR

Pada metode algoritma *Round Robin* melakukan proses antrean yaitu metode yang setiap prosesnya memperoleh alokasi waktu CPU dalam *quantum time* yang telah diketahui. Pada pengujian algoritma penjadwalan *Round Robin* ini *quantum time* yang digunakan sebesar 3 ms. Hal yang pertama yang harus dilakukan adalah menggambarkan kronologi eksekusi proses pada Tabel 2 di atas dengan *Gantt Chart*. Berikut adalah hasil pengujian dari parameter pengujian *Round Robin* dengan *quantum time* yang telah diketahui 3 ms berdasarkan Tabel 2 di atas:

1. Gantt Chart



Gambar 5. *Gantt chart* 20 data dengan algoritma *Round Robin*

Mengenai hasil *gantt chart* di atas dapat dijelaskan proses data antrean satu per satu. Dari *Gantt Chart* di atas terlihat bahwa setiap proses dikerjakan berdasarkan *quantum time* yang telah ditentukan sebesar 3 ms. Pertama P1 akan dikerjakan sebanyak 3 langkah, kemudian P2 sebanyak 3 langkah dan begitu selanjutnya sampai P20. Proses data antrean yang sudah diproses menurut waktu yang diberikan, maka akan kembali menunggu dan berada pada proses data antrean paling belakang. Pada P1 *start time* adalah 0 dan *arrival time* 0, setelah itu dikerjakan proses data antrean P1 dengan nilai *quantum time* sebesar 3 ms dan *burst time* 8. Selanjutnya, *burst time* pada proses data antrean P1 dikurangkan sebesar 3 ms berdasarkan *quantum time* yang telah ditentukan dan *burst time* pada P1 menjadi sebesar 5, maka proses P1 berhenti pada posisi 3. Selanjutnya, akan dikerjakan proses data antrean berikutnya berdasarkan nilai *arrival time* harus lebih kecil atau sama dengan nilai posisi P1,

maka proses data antrean yang dikerjakan berikutnya adalah P2 dengan *start time* 3, *arrival time* 2 dan *burst time* 7. Setelah itu *burst time* pada P2 dikurangkan sebesar 3 ms dan *burst time* pada P2 menjadi sebesar 4 ms, maka proses P2 berhenti pada posisi 6 dan begitu seterusnya proses yang akan dikerjakan pada P3 sampai P20 selesai di proses. Selanjutnya, hal yang dilakukan adalah menghitung *average waiting time*.

2. Average Waiting Time(AWT)

Hal yang dilakukan untuk menghitung *average waiting time* adalah dengan menghitung berapa lama waktu tunggu atau *waiting time* yang dihasilkan oleh proses data antrian tersebut dan dibagi dari jumlah seluruh proses. Untuk menghitung *waiting time* dapat dilakukan dengan cara mengurangi waktu eksekusi dengan waktu datang. Hasil *waiting time* dari setiap proses data antrian pada Gambar 5 adalah seperti tabel berikut ini :

Tabel 3. *Waiting time* dengan *round robin*

Proses	Waiting Time(ms)
P_1	51
P_2	51
P_3	93
P_4	49
P_5	47
P_6	83
P_7	42
P_8	75
P_9	78
P_10	45
P_11	47
P_12	75
P_13	93
P_14	45
P_15	72
P_16	66
P_17	65
P_18	76
P_19	78
P_20	81

Dari Tabel 3 di atas maka *average waiting time* (AWT) dapat dihitung yaitu :

$$\begin{aligned}
 AWT &= (WT_{P_1} + WT_{P_2} + WT_{P_3} + WT_{P_4} + WT_{P_5} + WT_{P_6} + WT_{P_7} + WT_{P_8} + WT_{P_9} + WT_{P_{10}} + \\
 &WT_{P_{11}} + WT_{P_{12}} + WT_{P_{13}} + WT_{P_{14}} + WT_{P_{15}} + WT_{P_{16}} + WT_{P_{17}} + WT_{P_{18}} + WT_{P_{19}} + WT_{P_{20}}) \\
 &/ 20 \\
 &= (51 + 51 + 93 + 49 + 47 + 83 + 42 + 75 + 78 + 45 + 47 + 75 + 93 + 45 + 72 + 66 + 65 + 76 + 78 + 81) / 20 \\
 &= 1312 / 20 \\
 &= \mathbf{65.6ms}
 \end{aligned}$$

Pada hasil waktu tunggu rata-rata atau *average waiting time* di atas didapat dengan cara menjumlahkan seluruh *waiting time* pada proses P1 sampai proses P20 dan dibagikan dengan jumlah seluruh proses data antrean.

3.4. Hasil Pengujian 35 Proses

Pada pengujian berikutnya peneliti melakukan pengujian kembali dengan jumlah proses data antrean 35 proses. Pada pengujian yang dilakukan penulis dengan parameter *input* adalah jumlah proses sebanyak 35 proses, *arrival time* dan *burst time* dari tiap proses dengan menggunakan proses antrian *single queue* yang dikerjakan oleh ketiga algoritma penjadwalan. Proses yang akan dieksekusi oleh CPU adalah seperti pada Tabel 4 berikut ini :

Tabel 4. Uji coba 35 data antrean

Proses ID	Arrival Time	BurstTime
P_1	0	5
P_2	2	7
P_3	4	3
P_4	6	8
P_5	7	4
P_6	9	3
P_7	12	6
P_8	15	8
P_9	17	5
P_10	19	7
P_11	21	10
P_12	25	5
P_13	27	9
P_14	29	6
P_15	32	4
P_16	33	5
P_17	38	7

Proses ID	Arrival Time	BurstTime
P_18	40	12
P_19	40	8
P_20	40	5
P_21	40	4
P_22	44	11
P_23	55	3
P_24	58	8
P_25	64	4
P_26	68	3
P_27	68	6
P_28	68	8
P_29	75	5
P_30	80	7
P_31	81	2
P_32	83	3
P_33	84	6
P_34	95	8
P_35	100	5

3.5. Hasil Pengujian 35 Proses Dengan Algoritma *Shortest Remaining First*

Pada algoritma SRF, seluruh proses yang berada pada *ready queue* akan diselesaikan berdasarkan proses yang memiliki *burst time* terkecil. Dikarenakan setiap proses diselesaikan berdasarkan proses yang memiliki *burst time* terkecil, maka *waiting time* setiap proses akan menjadi singkat dan hal ini juga menyebabkan *average waiting time* menjadi singkat. Oleh karena itu, algoritma ini dapat dikatakan salah satu algoritma yang optimal. Hal yang pertama yang harus dilakukan adalah menggambarkan kronologi eksekusi proses pada Tabel 4 di atas dengan *Gantt Chart*. Berikut adalah hasil pengujian dari parameter pengujian berdasarkan Tabel 4 di atas :

1. *Gantt Chart*

P 1	P 3	P 5	P 6	P 7	P 9	P 12
0	5	8	12	15	21	26
31	32	36	41	45	50	55
P 14	P 15	P 14	P 21	P 16	P 20	P 23
58	65	69	72	78	83	85
P 2	P 25	P 26	P 27	P 29	P 31	P 32
88	94	101	106	113	120	128
P 33	P 10	P 35	P 17	P 30	P 4	P 8
136	144	152	160	168	177	187
P 19	P 24	P 28	P 34	P 13	P 11	P 22
198	210					
P 18						
198	210					

Gambar 6. *Gantt chart* 35 data dengan algoritma *shortest remaining first*

Mengenai hasil *gant chart* di atas dapat dijelaskan proses data antrean satu per satu. Pada saat CPU tidak mengerjakan sesuatu atau dalam posisi 0 datang sebuah proses P1 yang membutuhkan waktu penyelesaian atau disebut *burst time* yang berjumlah 8, karena algoritma penjadwalan *SRF* merupakan algoritma yang melakukan proses berdasarkan *burst time* terkecil, maka proses selanjutnya ialah mengecek *burst time* P2, P3, P4 dan seterusnya jika sudah tiba. Jika *burst time* pada proses P2, P3, P4 dan seterusnya lebih kecil dari *burst time* P1 yang sedang diproses, maka proses yang memiliki *burst time* terkecil harus diproses terlebih dahulu dan P1 masuk ke status menunggu sampai dilayani. Jika semua proses telah selesai dikerjakan, maka langkah selanjutnya ialah menghitung waktu *average waiting time* dari seluruh proses yang telah dikerjakan.

2. *Average Waiting Time*

Hal yang dilakukan untuk menghitung *average waiting time* adalah dengan menghitung berapa lama waktu tunggu atau *waiting time* yang dihasilkan oleh proses data antrean tersebut dan dibagi dari jumlah seluruh proses. Untuk menghitung *waiting time* dapat dilakukan dengan cara mengurangi waktu eksekusi dengan waktu datang. Hasil *waiting time* dari setiap proses data antrean pada gambar 6 adalah seperti tabel di bawah ini :

Tabel 5. *Waiting time* dengan *shortest remaining first*

Proses	<i>Waiting Time</i> (ms)
P_1	0
P_2	56
P_3	1
P_4	114
P_5	1
P_6	3
P_7	3
P_8	113
P_9	4
P_10	75

Proses	Waiting Time(ms)
P_11	156
P_12	1
P_13	141
P_14	6
P_15	0
P_16	12
P_17	68
P_18	158
P_19	96
P_20	10
P_21	1
P_22	1
P_23	4
P_24	84
P_25	3
P_26	33
P_27	2
P_28	2
P_29	3
P_30	33
P_31	2
P_32	2
P_33	4
P_34	65
P_35	1

Dari tabel 4.5 diatas maka *average waiting time* (AWT) dapat dihitung yaitu :

$$\begin{aligned}
 AWT &= (WT P_1 + WT P_2 + WT P_3 + WT P_4 + WT P_5 + WT P_6 + WT P_7 + WT P_8 + WT P_9 + WT P_{10} + \\
 &WT P_{11} + WT P_{12} + WT P_{13} + WT P_{14} + WT P_{15} + WT P_{16} + WT P_{17} + WT P_{18} + WT P_{19} + WT P_{20}) \\
 &/ 20 \\
 &= (0 + 56 + 1 + 114 + 1 + 3 + 3 + 113 + 4 + 75 + 156 + 1 + 141 + 6 + 0 + 49 + 12 + 68 + 158 + 96 + 10 + 1 + 143 \\
 &+ 0 + 86 + 1 + 1 + 4 + 84 + 3 + 33 + 2 + 2 + 4 + 65 + 1) / 35 \\
 &= \mathbf{41.3714 \text{ ms}}
 \end{aligned}$$

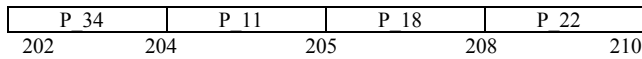
Pada hasil waktu tunggu rata-rata atau *average waiting time* di atas didapat dengan cara menjumlahkan seluruh *waiting time* pada proses P1 sampai proses P35 dan dibagi dengan jumlah seluruh proses data antrean.

3.6 Hasil Pengujian 35 Proses dengan algoritma Round Robin

Pada metode algoritma *Round Robin* melakukan proses antrean yaitu metode yang setiap prosesnya memperoleh alokasi waktu CPU dalam *quantum time* yang telah diketahui. Pada pengujian algoritma penjadwalan *Round Robin* ini *quantum time* yang digunakan sebesar 3 ms. Hal yang pertama yang harus dilakukan adalah menggambarkan kronologi eksekusi proses pada Tabel 5 di atas dengan *Gant Chart*. Berikut adalah hasil pengujian dari parameter pengujian *Round Robin* dengan *quantum time* yang telah diketahui 3 ms berdasarkan Tabel 5 di atas :

1. Gant Chart

P 1	P 2	P 3	P 4	P 5	P 6	P 7	
0	3	6	9	12	15	18	21
P 8	P 9	P 10	P 11	P 12	P 13	P 14	
21	24	27	30	33	36	39	42
P 15	P 16	P 17	P 18	P 19	P 20	P 21	
42	45	48	51	54	57	60	63
P 22	P 23	P 24	P 25	P 26	P 27	P 28	
63	66	69	72	75	78	81	84
P 29	P 30	P 31	P 32	P 33	P 34	P 35	
84	87	90	92	95	98	101	104
P 1	P 2	P 4	P 5	P 7	P 8	P 9	
104	106	109	112	113	116	119	121
P 10	P 11	P 12	P 13	P 14	P 15	P 16	
121	124	127	129	132	135	136	138
P 17	P 18	P 19	P 20	P 21	P 22	P 24	
138	141	144	147	149	150	153	156
P 25	P 27	P 28	P 29	P 30	P 33	P 34	
156	157	160	163	165	168	171	174
P 35	P 2	P 4	P 8	P 10	P 11	P 13	
174	176	177	179	181	182	185	188
P 17	P 18	P 19	P 22	P 24	P 28	P 30	
188	189	192	194	197	199	201	202



Gambar 7. Gant chart 35 data dengan algoritma round robin

Mengenai hasil *Gant Chart* diatas dapat dijelaskan proses data antrian satu per satu. Dari *Gant chart* diatas terlihat bahwa setiap proses dikerjakan berdasarkan *quantum time* yang telah ditentukan sebesar 3 ms. Pertama P_1 akan dikerjakan sebanyak 3 langkah, kemudian P_2 sebanyak 3 langkah dan begitu selanjutnya sampai P_20. Proses data antrian yang sudah diproses menurut waktu yang diberikan, maka akan kembali menunggu dan berada pada proses data antrian paling belakang. Pada P_1 *start time* adalah 0 dan *arrival time* 0, setelah itu dikerjakan proses data antrian P_1 dengan nilai *quantum time* sebesar 3 ms dan *burst time* 5 dan selanjutnya *burst time* pada proses data antrian P_1 dikurangkan sebesar 3 ms berdasarkan *quantum time* yang telah ditentukan dan *burst time* pada P_1 menjadi sebesar 2, maka proses P_1 berhenti pada posisi 3. Selanjutnya, akan dikerjakan proses data antrian berikutnya berdasarkan nilai *arrival time* harus lebih kecil atau sama dengan nilai posisi P_1, maka proses data antrian yang dikerjakan berikutnya adalah P_2 dengan *start time* 3, *arrival time* 2 dan *burst time* 7. Setelah itu *burst time* pada P_2 dikurangkan sebesar 3 ms dan *burst time* pada P_2 menjadi sebesar 4 ms, maka proses P_2 berhenti pada posisi 6 dan begitu seterusnya proses yang akan dikerjakan pada P_3 sampai P_35 selesai di proses. Selanjutnya, hal yang dilakukan adalah menghitung *average waiting time*.

2. *Average Waiting Time (AWT)*

Hal yang dilakukan untuk menghitung *average waiting time* adalah dengan menghitung berapa lama waktu tunggu atau *waiting time* yang dihasilkan oleh proses data antrean tersebut dan dibagi dari jumlah seluruh proses. Untuk menghitung *waiting time* dapat dilakukan dengan cara mengurangi waktu eksekusi dengan waktu datang. Hasil *waiting time* dari setiap proses data antrean pada Gambar 7 adalah seperti tabel berikut ini :

Tabel 6. *Waiting time* dengan round robin

Proses	<i>Waiting Time (ms)</i>
P_1	101
P_2	106
P_3	2
P_4	165
P_5	102
P_6	6
P_7	98
P_8	158
P_9	99
P_10	156
P_11	174
P_12	99
P_13	152
P_14	100
P_15	100
P_16	100
P_17	144
P_18	156
P_19	146
P_20	104
P_21	106
P_22	155
P_23	11
P_24	133
P_25	89
P_26	7
P_27	86
P_28	125
P_29	85
P_30	115
P_31	9
P_32	9
P_33	81
P_34	101
P_35	71

Dari tabel 1.6 diatas maka *average waiting time (AWT)* dapat dihitung yaitu :

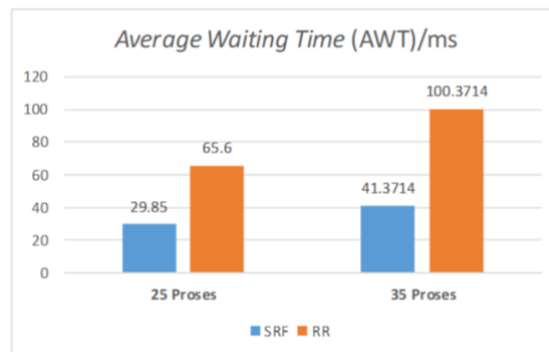
$$\begin{aligned}
 AWT &= (WT P_1 + WT P_2 + WT P_3 + WT P_4 + WT P_5 + WT P_6 + WT P_7 + WT P_8 + WT P_9 + WT P_10 + \\
 &WT P_11 + WT P_12 + WT P_13 + WT P_14 + WT P_15 + WT P_16 + WT P_17 + WT P_18 + WT P_19 + WT P_20 + WT P_21 + \\
 &WT P_22 + WT P_23 + WT P_24 + WT P_25 + WT P_26 + WT P_27 + WT P_28 + WT P_29 + WT P_30 + \\
 &WT P_31 + WT P_32 + WT P_33 + WT P_34 + WT P_35) / 35 \\
 &= (101 + 168 + 2 + 165 + 102 + 6 + 98 + 158 + 99 + 156 + 174 + 99 + 152 + 100 + 100 + 100 + 144 + 156 + 146 + 104 \\
 &+ 106 + 155 + 11 + 133 + 89 + 7
 \end{aligned}$$

$$\begin{aligned} &+ 86 + 125 + 85 + 115 + 9 + 9 + 81 + 101 + 71) / 35 \\ &= 3513 / 35 \\ &= \mathbf{100.3714 \text{ ms}} \end{aligned}$$

Pada hasil waktu tunggu rata-rata atau *average waiting time* diatas didapat dengan cara menjumlahkan seluruh *waiting time* pada proses P_1 sampai proses P_35 dan dibagi dengan jumlah seluruh proses data antrian.

4. KESIMPULAN

Dari hasil percobaan menunjukkan bahwa AWT pada algoritma SRF lebih baik dibandingkan algoritma RR. Hal ini dibuktikan dengan hasil percobaan dimana algoritma SRF mendapatkan hasil AWT 29.85 ms dan algoritma RR mendapatkan hasil AWT 65.6 ms.



Gambar 8. Hasil AWT Pada Algoritma SRF dan RR

REFERENSI

- [1] P. Tri Dharma, P. Rakhmat, "Analisis Algoritma Round Robin pada Penjadwalan CPU", Jurnal Ilmiah Teknologi Informasi Asia, Vol. 15, No. 2 Tahun 2021.
- [2] N. Nuraini and I. Ahmad, "Sistem Informasi Manajemen Kepegawaian Menggunakan Metode Key Performance Indicator Untuk Rekomendasi Kenaikan Jabatan (Studi Kasus: Kejaksaan Tinggi Lampung)", J. Teknol. dan Sist. Inf., vol. 2, no. 3, p. 81, 2021.
- [3] I. Tasya Fitria., "Analisis Performa Algoritma Penjadwalan CPU Dalam Sistem Operasi Komputer Untuk Pengoptimalan Responsivitas", Jurnal Teknologi Pintar, Vol. 3, November 2023.
- [4] O. Hoger K, J. Kamal H, H. Shalau F, Comperative, "Comparative Analysis of The Essential CPU Scheduling Algorithms", Journal Bulletin of Electrical Engineering and Informatics, Vol. 10, No. 5, PP. 2742-2750, October 2021.
- [5] P. Rakhmat, P. Tri Dharma, "Median-Average Round Robin (MARR) Algorithm for Optimal CPU Task Scheduling", Jurnal dan Penelitian Teknik Informatika (Sinkron), Vol. 9, No. 1, January 2025.
- [6] S. Siti, S. M Yogie, Rismayanti, "Penerapan Algoritma Shortest Job First (SJF) dan Priority Scheduling (PS) pada Maintenance Mesin ATM, Jurnal UINSU, Vol. 07, No. 1, April 2023.
- [7] M. Mr. Rohit, Operating System, EduGorilla Prep Experts.
- [8] H. Richki, Arsitektur dan Organisasi Komputer, Ponorogo : Uwais Inspirasi Indonesia, 2021.
- [9] A. Kande, dkk, Introduction to Operating Systems, India : Nitya Publications, 2020.
- [10] A. Mahdi S, A. Firas Sabah, M. Tariq A, "Reducing Waiting and Idle Time for A Group of Jobs in The Grid Computing", Journal Bulletin of Electrical Engineering and Informatics, Vol. 12, No. 5, PP. 3115-3123, October 2023.
- [11] N. Ahmad Mursyidun, dkk Kernel System, Malang : Ahlimedia Press, 2021.
- [12] J. Sandeep, "Shortest Remaining Time First (Preemptive SJF) Scheduling Algorithm", 03 Februari 2025, [Online]. Tersedia : <https://www.geeksforgeeks.org/shortest-remaining-time-first-preemptive-sjf-scheduling-algorithm> [Diakses : 25 Februari 2025].
- [13] S. Agung, J. Erfian, "Sistem Penjadwalan Produksi Makanan Sei Menggunakan Algoritma Round Robin di CV Gyumbox, eProsiding Teknik Informatika (PROTEKTIF), Juni 2021, Vol. 2, No. 1.
- [14] T. Raj Gaurang, A. Ambuj Kumar, K. Vikas, S. Durgesh, Design Concepts of Operating System, Delhi : Namy Press.Com, 2021.
- [15] W. Wisnu, M. Desinta, A. Kezia Jazzlyn, "Implementasi Algoritma Round Robin dan Priority Pada Sistem Antrian Rumah Sakit", Jurnal Fasilkom, Vol 14, No. 2, PP. 507-513, Agustus 2024.