

**ANALISIS KOMPARATIF CLUSTER POLA GAJI-USIA DI TIGA SEKTOR KUNCI EKONOMI
(PENJUALAN, JASA, DAN PERTANIAN) MENGGUNAKAN ALGORITMA K-MEANS**

Seftia Ananta¹, Muhammad Luthfi Ramadhan², Maulana Dwi Saputro³, Naila Qobilah
Adiansyah⁴

^{1,2,3,4,5}Program Studi Sistem Informasi, Fakultas Teknik Informatika, Universitas Bina Sarana
Informatika, Jl. Raya Kaliabang Tengah No8, Bekasi, Jawa Barat, Indonesia

Email penulis pertama: seftiaanantaa@gmail.com¹, luthfiramadhan420@gmail.com²,
maulanasaputro217@gmail.com³, nailaputri26605@gmail.com⁴

ABSTRACT

This study aims to analyze and compare the relationship between wages and age in three main economic sectors, namely the Sales, Services, and Agriculture, Forestry, Hunting, and Fisheries (PKPP) sectors. The main focus of this study is the wage level and age group of workers in each sector, with the hope of identifying the characteristics of worker groups that reflect economic and demographic trends in the workforce. To analyze the data, the method used is clustering with the K-Means algorithm run through RapidMiner software. The data source is taken from the Central Statistics Agency (BPS) in 2025, namely datasets taken in February 2024 and February 2025 containing the average wages or net salaries per month for laborers, employees, and staff based on age groups and main types of work. The results show the existence of three main clusters, namely: Cluster 0: Group of workers of all ages with high salaries, Cluster 1: Group of workers of productive age with low salaries, Cluster 2: Group of young and senior workers with middle salaries. This analysis provides a comparative overview of the distribution of workers' welfare across various sectors and can serve as a basis for formulating data-based employment policies

Keywords: K-Means, Clustering, Salary, Age, Sector

ABSTRAK

Penelitian ini bertujuan untuk menganalisis dan membandingkan pola hubungan antara gaji dan usia di tiga sektor utama ekonomi, yaitu sektor Penjualan, Jasa, serta sektor (PKPP) Pertanian, Kehutanan, Perburuan, dan Perikanan. Fokus utama dari penelitian ini adalah tingkat gaji dan kelompok usia para pekerja di masing-masing sektor, dengan harapan bisa mengenali ciri-ciri kelompok pekerja yang mencerminkan tren ekonomi dan demografi tenaga kerja. Untuk menganalisis data, metode yang digunakan adalah clustering dengan algoritma K-Means yang dijalankan melalui perangkat lunak RapidMiner. Sumber datanya diambil dari Badan Pusat Statistik (BPS) tahun 2025, yaitu dataset yang diambil pada bulan februari 2024 dan febuari 2025 berisi rata-rata upah atau gaji bersih per bulan untuk buruh, karyawan, dan pegawai berdasarkan

Article History

Received: Desember 2025

Reviewed: Desember 2025

Published: Desember 2025

Plagiarism Checker No
235

Prefix DOI :

[10.8734/Kohesi.v1i2.365](https://doi.org/10.8734/Kohesi.v1i2.365)

Copyright : Author

Publish by : Kohesi



This work is licensed
under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



kelompok usia dan jenis pekerjaan utama. Hasil penelitian menunjukkan adanya tiga kelompok utama, yaitu: Cluster 0: Grup pekerja semua usia dengan gaji tinggi, Cluster 1: Grup pekerja usia produktif dengan gaji rendah, Cluster 2: Grup pekerja muda dan senior dengan gaji menengah. Analisis ini memberikan gambaran perbandingan mengenai distribusi kesejahteraan pekerja di berbagai sektor dan dapat menjadi dasar dalam penyusunan kebijakan ketenagakerjaan yang didasarkan pada data.

Kata Kunci: *K-Means, Clustering, Gaji, Usia, Sektor*

1. PENDAHULUAN

Gaji adalah uang yang diberikan oleh perusahaan kepada karyawan sebagai pertukaran atas kerja dan layanan yang mereka berikan. Secara umum, Imbalan yang harus diterima karyawan atas pekerjaan yang telah diselesaikan setiap harian, setiap bulanan, atau setiap tahun (Joiwi & Nainggolan, 2021). Sebagai negara berkembang, Indonesia menghadapi berbagai tantangan dalam upaya meningkatkan kesejahteraan rakyatnya, terutama di bidang ketenagakerjaan (Ananda et al., 2025).

Berdasarkan data dari Badan Pusat Statistik, mayoritas pekerja di Indonesia masih bekerja di sektor informal. Perbedaan gaji antara usia dan jenis pekerjaan menunjukkan adanya ketimpangan ekonomi yang cukup besar di berbagai daerah (Statistik, 2024). Kondisi ini mengindikasikan bahwa faktor usia dan tenaga kerja memiliki hubungan yang erat dengan produktivitas serta tingkat gaji yang diterima pekerja. Begitupun beberapa hal lainnya yang memengaruhi kemampuan pekerja di perusahaan, baik itu dari kebijakan yang diambil perusahaan maupun kondisi yang dialami oleh para pekerja itu sendiri (Jamaludin et al., 2024).

Tujuan utama penelitian ini adalah Menerapkan algoritma *K-Means* untuk membandingkan perbedaan tingkat gaji berdasarkan kelompok usia di tiga sektor pekerjaan di Indonesia. Dengan pendekatan ini, penelitian memperoleh hasil pola pengelompokan usia dan gaji yang mampu menunjukkan bagaimana kesejahteraan pekerja didistribusikan secara lebih jelas. Hasilnya nanti bisa menjadi acuan dalam membuat kebijakan ketenagakerjaan yang didasarkan pada data. Algoritma *K-Means Clustering* dalam *tools RapidMiner* dipilih karena digunakan untuk memproses data, sehingga bisa dianalisis hubungan antara gaji dan usia di tiga sektor utama, yaitu sektor penjualan, jasa, serta sektor (PKPP) pertanian, kehutanan, perburuan, dan perikanan.

Gaji berperan penting dalam perekonomian karena menjadi sumber utama penghasilan bagi pekerja dan langsung mempengaruhi kesejahteraan masyarakat. Memberi kompensasi atau gaji yang adil tidak hanya membantu kondisi keuangan karyawan, tetapi juga membuat mereka



lebih tenang dalam hati, sehingga lebih semangat bekerja dan lebih aktif mengambil inisiatif (Denny et al., 2021). Selain itu, kebijakan menentukan gaji yang adil dan sesuai dapat mengurangi perbedaan sosial serta menyeimbangkan kebutuhan para pekerja dengan kemampuan perusahaan dalam memberikan kompensasi. Oleh karena itu, Kinerja yang baik adalah kinerja yang terbaik, yaitu kinerja yang sesuai dengan standar perusahaan dan membantu mencapai tujuan perusahaan (Sasmitasari et al., 2023).

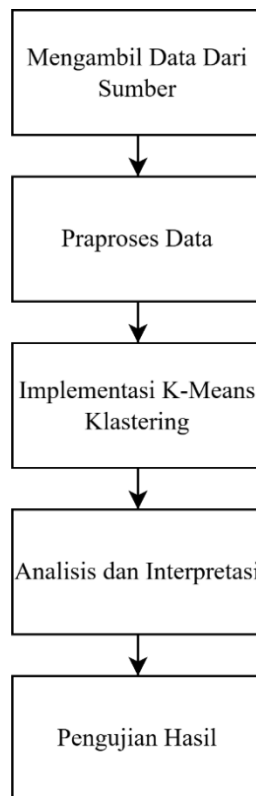
Adapun penelitian yang telah dilakukan sebelumnya yang berjudul “Analisis Clustering Penduduk Berdasarkan Kelompok Umur dengan K-Means dan Hierarchical Clustering untuk Perencanaan Demografi”. Penelitian ini dilakukan untuk mengidentifikasi pola distribusi penduduk berdasarkan kelompok umur, hasil dari penelitian tersebut menunjukkan bahwa algoritma *K-Means* mampu mengelompokkan penduduk ke dalam tiga klaster utama, di mana klaster 2 mendominasi dengan proporsi sebesar 65%, yang terdiri dari kelompok usia muda di bawah 40 tahun (Lestari, 2024).

Berdasarkan penelitian sebelumnya yang telah diuraikan, penelitian tersebut berfokus pada klasterisasi penduduk berdasarkan kelompok umur tanpa memperhitungkan faktor ekonomi seperti tingkat gaji atau sektor pekerjaan. Dengan demikian, penelitian ini bertujuan untuk mengembangkan penelitian sebelumnya dengan menerapkan metode *K-Means Clustering* untuk menganalisis hubungan antara usia dan tingkat gaji pada tiga sektor utama tenaga kerja di Indonesia. Melalui pembuatan jurnal penelitian ini, dapat memberikan hasil gambaran yang lebih komprehensif mengenai distribusi kesejahteraan tenaga kerja serta menjadi dasar bagi perumusan kebijakan ketenagakerjaan berbasis data di Indonesia.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Metode penelitian yang dilakukan dapat dilihat pada Gambar 1, yang menunjukkan alur proses dari pengambilan data hingga evaluasi hasil sebagai berikut:



Gambar 1. Tahapan penelitian

Adapun penjelasan dari gambar di atas yaitu:

1. Mengambil Data Dari Sumber

Pengumpulan data sekunder yang diambil dari situs resmi Badan Pusat Statistik (BPS), yaitu data Rata-rata upah/Gaji Bersih Sebulan Buruh/Karyawan/Pegawai menurut umur dan jenis pekerjaan Utama pada bulan februari 2024 dan februari 2025. Data diunduh dalam format tabel statistik yang kemudian dikonversi menjadi format CSV untuk diolah menggunakan *RapidMiner*.

2. Praproses Data

Tahap ini meliputi proses pembersihan data (*data cleaning*) untuk menghilangkan nilai kosong atau data duplikat, transformasi data agar seluruh variabel berada dalam satuan yang seragam, serta normalisasi data menggunakan metode *Min-Max Scaling* agar setiap atribut memiliki rentang nilai antara 0 dan 1.

3. Implementasi K-Means Clustering

Tahapan inti penelitian ini adalah penerapan algoritma *K-Means* untuk melakukan pengelompokan data. Jumlah kluster ditetapkan berdasarkan evaluasi *Davies-Bouldin Index* (DBI), dengan $K = 3$ sebagai nilai yang menghasilkan performa klusterisasi optimal. Visualisasi hasil kluster kemudian disajikan dalam bentuk *scatter plot* guna menggambarkan distribusi data pada masing-masing cluster.

4. Analisis dan Interpretasi



Hasil klasterisasi kemudian dianalisis untuk mengidentifikasi pola hubungan antara usia dan tingkat gaji pada masing-masing sektor. Setiap klaster diinterpretasikan berdasarkan karakteristik dominannya, seperti kelompok usia produktif, kelompok usia muda, dan kelompok usia senior.

5. Pengujian Hasil

Tahap akhir adalah pengujian hasil untuk mengetahui kecocokan hasil analisis, hasil analisis algoritma *K-Means* akan diuji menggunakan *tools RapidMiner*. Metode yang digunakan untuk dilakukannya pengujian hasil adalah *Davies-Bouldin Index* (DBI).

2.2 Data Mining

Data mining adalah proses mengambil pengetahuan atau informasi yang berguna dari sekumpulan data yang besar dan rumit. Tujuan utamanya adalah menemukan pola, hubungan, atau informasi yang tidak terlihat jelas secara langsung, sehingga bisa memberikan pemahaman yang lebih dalam dan bernilai (Rahayu et al., 2024).

Berikut adalah beberapa teknik analisis data mining yang umum (Wororowi et al., 2024).

- a. Klasifikasi: Teknik ini digunakan untuk mengklasifikasikan data ke dalam kategori atau kelas berdasarkan atribut-atribut tertentu.
- b. Klustering: Adalah teknik yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang memiliki karakteristik atau atribut yang serupa
- c. Regresi: Teknik regresi digunakan untuk memodelkan hubungan antara satu atau lebih variabel independen dan variabel dependen. Ini membantu dalam memprediksi nilai variabel dependen berdasarkan variabel independen yang diberikan.
- d. Asosiasi: Teknik asosiasi digunakan untuk menemukan hubungan atau asosiasi antara item-item dalam kumpulan data.

2.3 Clustering

Clustering dapat diartikan sebagai cara mengelompokkan objek-objek yang memiliki kesamaan. Dengan metode *clustering*, kita bisa lebih mudah mengenali bagian-bagian yang padat dan jarak antar objek dalam ruang. Selain itu, kita juga dapat memahami secara umum pola penyebaran dan hubungan antar atribut. Pendekatan klasifikasi juga bisa digunakan untuk membedakan kelompok atau jenis objek yang berbeda (Afifah & Nurdiyanto, 2023).

Dalam teknik *data mining* terdapat dua metode *clustering* yang digunakan untuk mengelompokkan data, yaitu *clustering* hierarkis dan *clustering* non hierarkis (Ulfah & Sri Irtwaty, 2022).



2.4 Algoritma K-Means

Algoritma *K-means* merupakan salah satu metode yang digunakan untuk mempartisi objek yang ada dalam satu atau lebih kelompok atau kelompok objek yang berdasarkan karakteristiknya, maka objek yang berkarakteristik sama akan dikelompokkan dalam sebuah klaster yang sama serta objek yang memiliki karakteristik berbeda akan dikelompokkan ke dalam klaster yang lain (Mustika et al., 2021).

2.5 Davies Bouldin-Index

Davies-Bouldin Index (DBI) mengukur rata-rata tingkat kesamaan setiap klaster dengan klaster-klaster lain yang memiliki tingkat kesamaan tertinggi dengan klaster tersebut (Raihannabil, 2024). Semakin kecil nilai DBI, semakin baik hasil pengelompokan yang diperoleh. Rumus untuk menghitung DBI adalah sebagai berikut:

$$DBI = \frac{1}{K} \sum_{k=1}^K \max_{j \neq k} \left(\frac{s_k + s_j}{d_{kj}} \right)$$

Gambar 2. Rumus *Davies Bouldin-Index*

Keterangan :

- s_i dan s_j = *intra-cluster distance* (jarak dalam klaster)
- d_{kj} = *inter-cluster distance* (jarak antarklaster).

2.6 RapidMiner

RapidMiner adalah perangkat lunak yang dikembangkan oleh Dr. Markus Hofmann dari *Institute of Technology Blanchardstown* serta *Ralf Klinkenberg* dari *rapid-i.com*, yang memiliki antar muka grafis (GUI) sehingga menyediakan kemudahan bagi pengguna dalam mengoperasikannya. Perangkat lunak ini bersifat *open source*, dikembangkan dengan menggunakan bahasa pemrograman *Java*, dan berada di bawah lisensi *GNU Public Licence*. *RapidMiner* dapat dijalankan pada berbagai jenis sistem operasi (Sudarsono et al., 2021).

3. HASIL DAN PEMBAHASAN

3.1 Proses Pembentukan Dataset

Data diambil dari sumber Badan Pusat Statistik (BPS) yang berjudul “Rata-Rata Upah/Gaji Bersih Sebulan Buruh/Karyawan/Pegawai menurut Kelompok Umur dan Jenis Pekerjaan Utama (Rupiah)” dibuat dalam bentuk file *Excel*, seperti yang ditunjukkan pada gambar berikut:



Gambar 3. Dataset penelitian pada tahun 2024

Gambar 4. Dataset penelitian pada tahun 2025

Data Pada penelitian ini menggunakan beberapa variabel untuk melakukan analisa menggunakan metode *K-Means Clustering*. Berikut ini adalah variabel yang digunakan:

Tabel 1. Variabel yang dianalisis

No	Nama Variabel
1	Kelompok Umur
2	Tenaga Kerja

Adapun penjelasan dari variabel adalah seperti berikut ini:

1. Kelompok Umur adalah grup pekerja dengan umur masing-masing
2. Tenaga Kerja memiliki bidang kerja yaitu Usaha penjualan, Jasa, PKPP (Pertanian, Kehutanan, Perburuan dan Perikanan)

Jumlah isi kolom adalah 30, yang mencakup semua variabel. Kemudian menyiapkan dataset yang terstruktur dalam bentuk tabel. Berikut ini adalah dataset yang digunakan pada penelitian ini hasil pembersihan data:

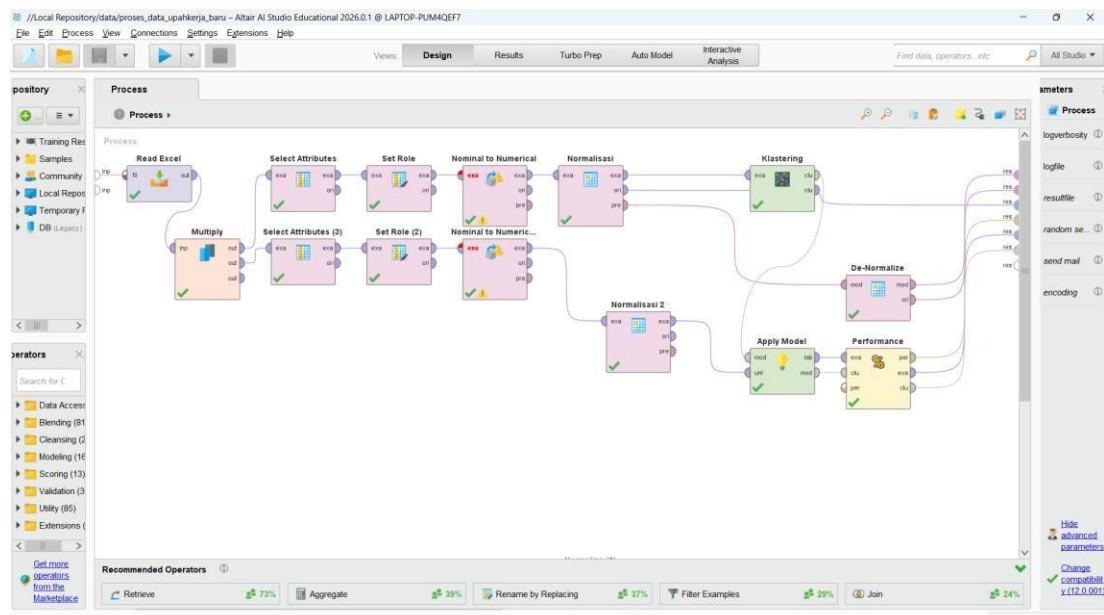
Tabel 2. Dataset yang digunakan untuk penelitian

Kelompok Umur	Tenaga Kerja	Februari 2024	Februari 2025
15-19	Tenaga usaha penjualan	Rp1.479.264	Rp1.499.624
15-19	Tenaga usaha jasa	Rp1.512.434	Rp1.582.490
15-19	Tenaga usaha pertanian, kehutanan, perburuan dan perikanan	Rp1.574.818	Rp2.082.134
....			

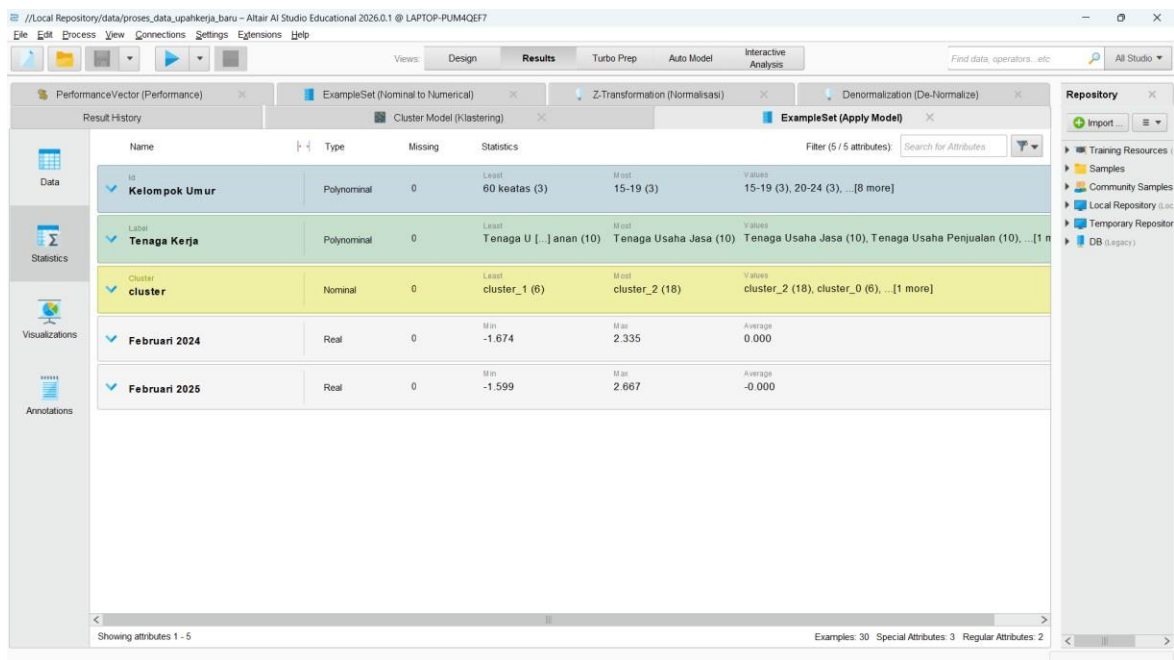
Setelah menentukan dataset, selanjutnya melakukan pengolahan data menggunakan algoritma *K- Means* dengan aplikasi *Rapidminer Studio*.

3.2 Pengolahan Data

- A. Dalam Pengujian ini, *RapidMiner* digunakan untuk menggolongkan data gaji dalam bentuk *klaster*. Berikut ini adalah proses *pre-processing*

Gambar 5. Proses *Pre-Processing*

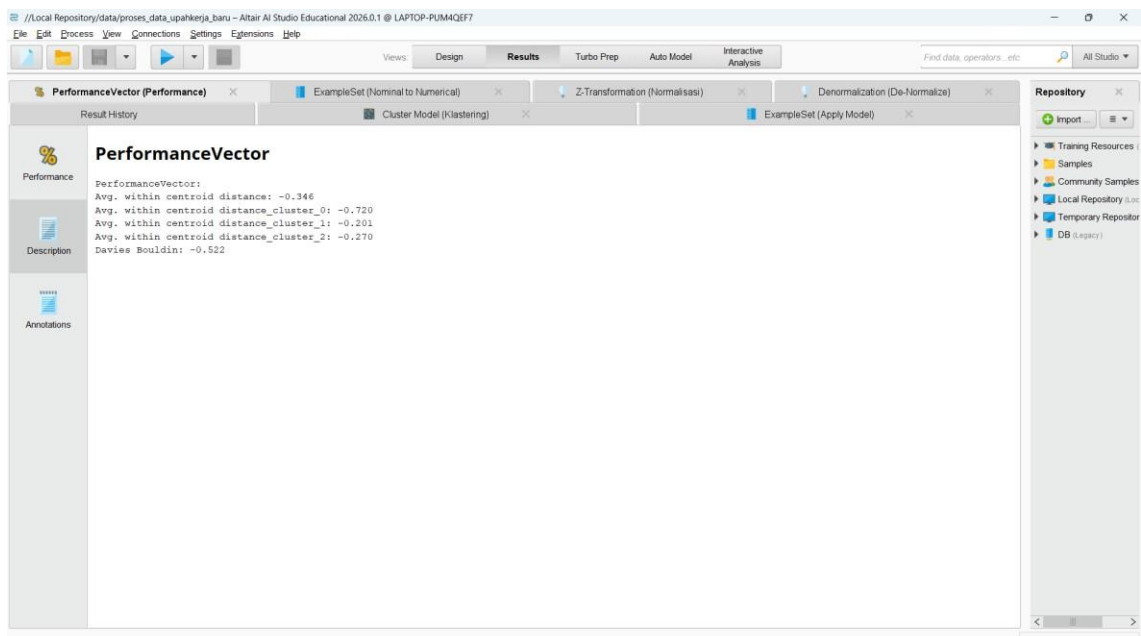
- B. Setelah semua operator terhubung dengan benar, klik tombol *Run* untuk memulai proses. *RapidMiner* menghasilkan gambaran struktur *clustering* dan *scatter plot*.



Name	Type	Missing	Statistics
Kelompok Umur	Polynomial	0	Least: 60 keatas (3), Most: 15-19 (3), Values: 15-19 (3), 20-24 (3), ...[8 more]
Tenaga Kerja	Polynomial	0	Least: Tenaga U [...] anan (10), Most: Tenaga Usaha Jasa (10), Tenaga Usaha Jasa (10), Tenaga Usaha Penjualan (10), ...[1 more]
cluster	Nominal	0	Least: cluster_1 (6), Most: cluster_2 (18), Values: cluster_2 (18), cluster_0 (6), ...[1 more]
Februari 2024	Real	0	Min: -1.674, Max: 2.335, Average: 0.000
Februari 2025	Real	0	Min: -1.599, Max: 2.667, Average: -0.000

Gambar 6. Hasil data *Pre-Processing*

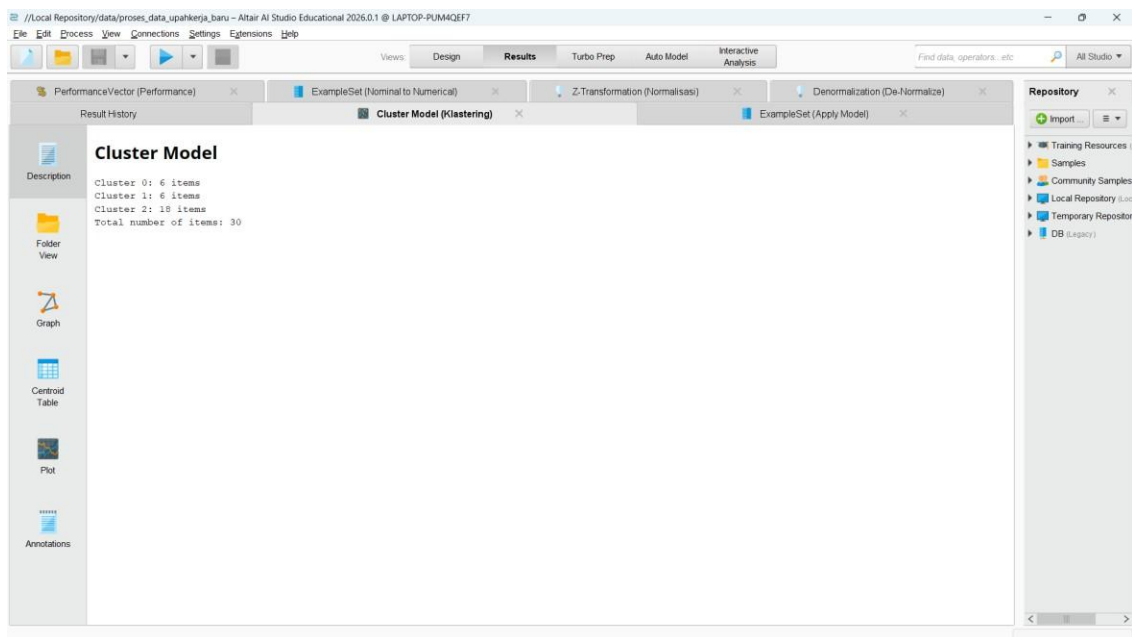
- C. Berdasarkan hasil analisis menggunakan *RapidMiner*, diperoleh nilai *Davies-Bouldin Index* (DBI) sebesar -0.522. Nilai ini menunjukkan bahwa kualitas kluster yang terbentuk cukup baik, dengan pemisahan antar kluster yang memadai dan kekompakan intrakluster yang terjaga.



Performance
PerformanceVector:
Avg. within centroid distance: -0.346
Avg. within centroid distance_cluster_0: -0.720
Avg. within centroid distance_cluster_1: -0.201
Avg. within centroid distance_cluster_2: -0.270
Davies Bouldin: -0.522

Gambar 7. Hasil analisis dari *Davies-Bouldin Index*

- D. Berdasarkan dari hasil pengolahan data di atas, diperoleh model kluster yang menghasilkan tiga kelompok.



Gambar 8. Model klaster yang dihasilkan

E. Hasil klaster dari 3 kelompok model klaster sebagai berikut:

1. *Cluster 0* terdiri dari 6 data, sebagaimana ditampilkan pada gambar berikut:

Row No.	Kelompok U...	Tenaga Kerja	cluster ↑	Februari 2024	Februari 2025
10	30-34	Tenaga Usaha Penjualan	cluster_0	1.573	0.956
13	35-39	Tenaga Usaha Penjualan	cluster_0	1.181	1.527
16	40-44	Tenaga Usaha Penjualan	cluster_0	1.919	1.098
19	45-49	Tenaga Usaha Penjualan	cluster_0	1.211	2.667
22	50-54	Tenaga Usaha Penjualan	cluster_0	2.335	2.345
25	55-59	Tenaga Usaha Penjualan	cluster_0	1.016	0.779

Gambar 9. *Cluster 0*

Pada gambar ini, *Cluster 0* didominasi oleh Tenaga kerja usaha penjualan dengan umur 30-34 hingga 55-59 dengan gaji tinggi antara 2.9-3.5 juta pada februari 2024 dan 2.9-4.0 juta pada februari 2025.

2. *Cluster 1* juga terdiri dari 6 data, ditunjukkan pada gambar berikut:

1	15-19	Tenaga Usaha Penjualan	cluster_1	-1.306	-1.211
2	15-19	Tenaga Usaha Jasa	cluster_1	-1.255	-1.087
3	15-19	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_1	-1.160	-0.339
27	55-59	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_1	-1.220	-0.847
29	60 keatas	Tenaga Usaha Jasa	cluster_1	-1.674	-1.599
30	60 keatas	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_1	-1.654	-1.321

Gambar 10. *Cluster 1*

Pada gambar ini, *Cluster 1* didominasi oleh tiga sektor usaha yaitu penjualan, jasa dan (PKPP) Pertanian, Kehutanan Perburuan dan Perikanan.

Umur antara 15-19, 55-59 dan 60 keatas dengan gaji rendah antara 1.2-1.5 juta pada februari 2024 dan 1.2-2.0 juta di februari 2025.

3. *Cluster 2* memiliki jumlah data terbanyak, yaitu 18 data, seperti terlihat pada gambar berikut:

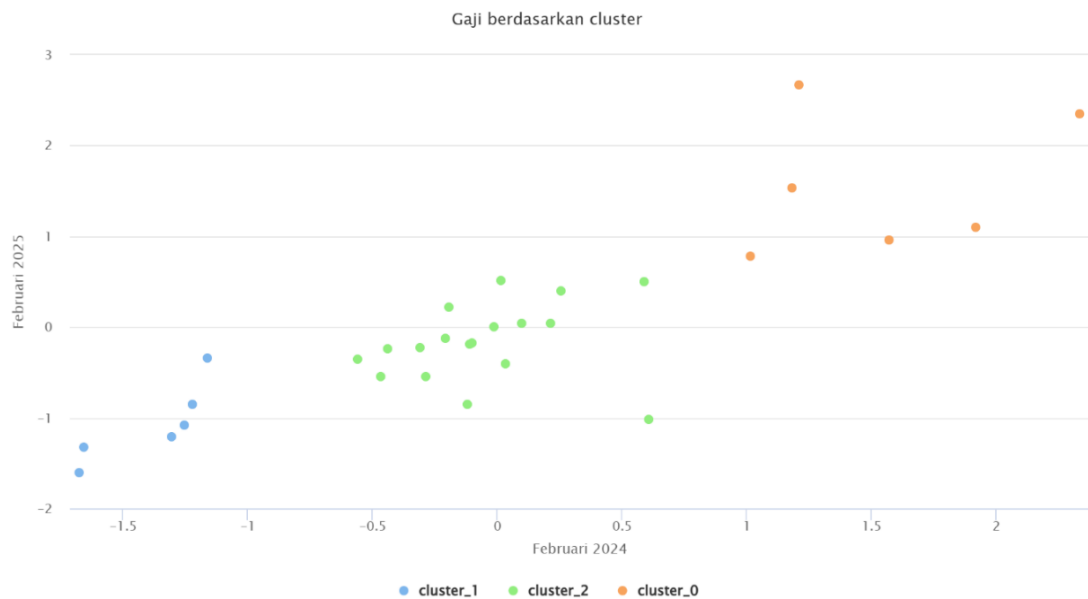
4	20-24	Tenaga Usaha Penjualan	cluster_2	-0.437	-0.234
5	20-24	Tenaga Usaha Jasa	cluster_2	-0.557	-0.351
6	20-24	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	-0.467	-0.552
7	25-29	Tenaga Usaha Penjualan	cluster_2	0.591	0.497
8	25-29	Tenaga Usaha Jasa	cluster_2	0.013	0.509
9	25-29	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	0.255	0.393
11	30-34	Tenaga Usaha Jasa	cluster_2	-0.191	0.225
12	30-34	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	0.213	0.037
14	35-39	Tenaga Usaha Jasa	cluster_2	-0.101	-0.170
15	35-39	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	0.099	0.040
17	40-44	Tenaga Usaha Jasa	cluster_2	-0.205	-0.129
18	40-44	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	-0.107	-0.193
20	45-49	Tenaga Usaha Jasa	cluster_2	-0.283	-0.541
21	45-49	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	-0.307	-0.233
23	50-54	Tenaga Usaha Jasa	cluster_2	-0.010	0.006
24	50-54	Tenaga Usaha Pertanian, Kehutanan, Perburuan dan Perikanan	cluster_2	0.036	-0.408
26	55-59	Tenaga Usaha Jasa	cluster_2	-0.119	-0.846
28	60 keatas	Tenaga Usaha Penjualan	cluster_2	0.610	-1.017

Gambar 11. *Cluster 2*

Pada gambar ini, *Cluster 2* didominasi oleh tiga sektor usaha yaitu penjualan, jasa dan (PKPP) Pertanian, Kehutanan Perburuan dan Perikanan. Umur antara 20-29, 30-39, 40-49, 50-59 dan 60 keatas dengan gaji menengah antara 1.9-2.7 juta pada februari 2024 dan 1.6-2.6 juta pada februari 2025.

F. Selanjutnya, untuk memvisualisasikan distribusi data pada masing-masing klaster, disajikan *scatter plot* yang menggambarkan pola pengelompokan tersebut.

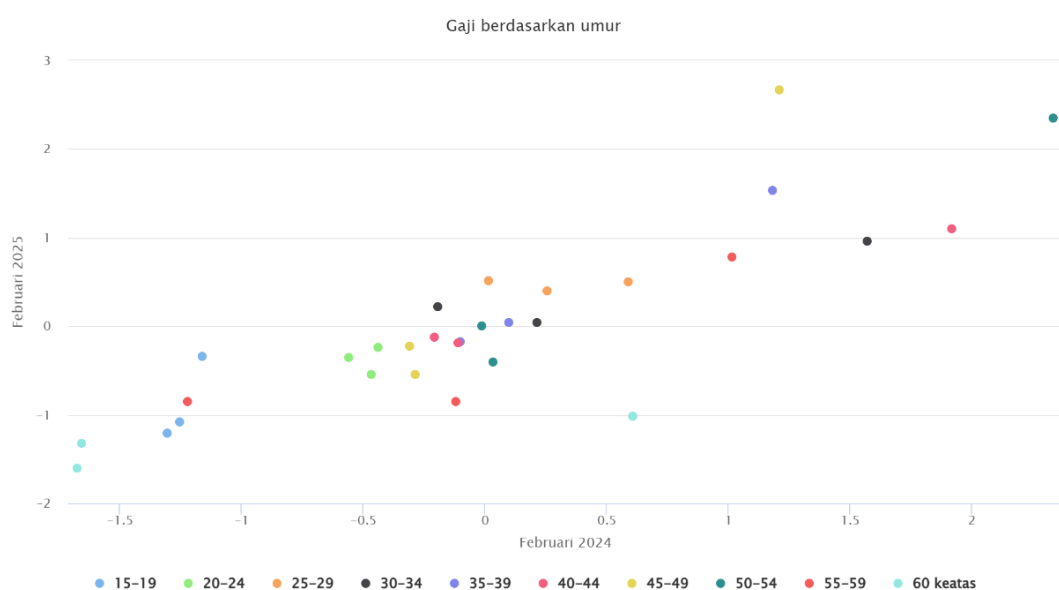
1. Gaji Berdasarkan *Cluster*

Gambar 12. Grafik *Scatter Plot* berdasarkan *Cluster*

Berdasarkan *scatter plot* gaji berdasarkan tenaga kerja pada gambar ini, masing-masing kelompok menunjukkan pola distribusi pendapatan yang berbeda.

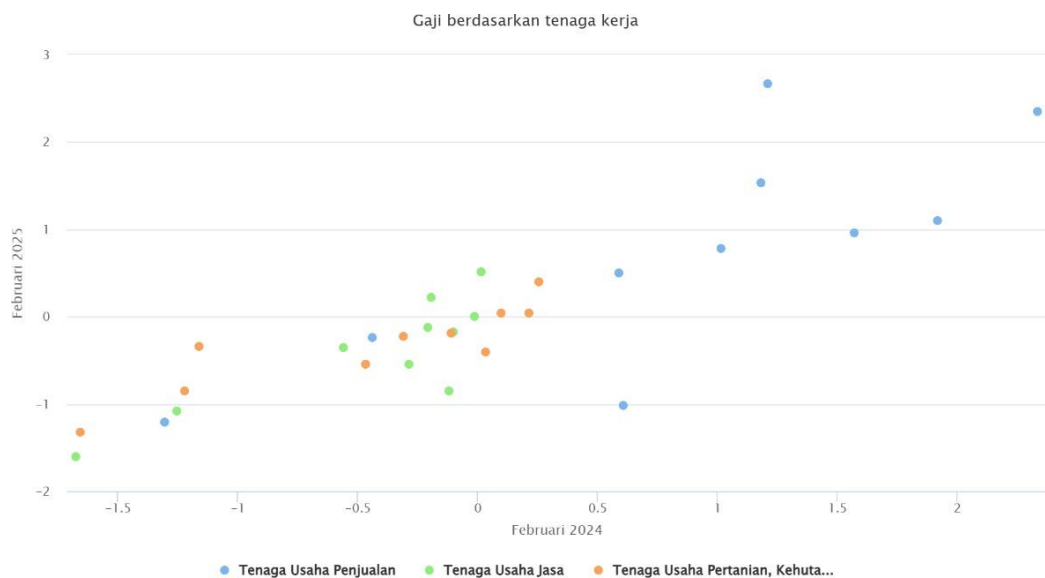
- a. *Cluster* 0 (oren) berada pada rentang nilai tertinggi dan menggambarkan pekerjaan dengan gaji besar serta pertumbuhan pendapatan yang lebih signifikan dibandingkan kelompok lainnya.
- b. *Cluster* 1 (biru) berada pada rentang nilai terendah, menunjukkan pekerja dengan gaji rendah dan perubahan pendapatan yang minimal, cenderung berada pada nilai negatif.
- c. *Cluster* 2 (hijau) menempati rentang nilai menengah dan merepresentasikan pekerja dengan gaji sedang, dengan perubahan pendapatan yang stabil dan tidak ekstrem.

2. Gaji Berdasarkan umur

Gambar 13. Grafik *Scatter Plot* gaji berdasarkan umur

Berdasarkan *scatter plot* pada gambar ini menggambarkan adanya kecenderungan peningkatan gaji dari 2024 ke 2025 pada berbagai kelompok usia. Pekerja berusia 15-24 tahun umumnya berada pada kategori pendapatan rendah, sesuai dengan fase awal pengalaman kerja. Kelompok usia 25-44 tahun memperlihatkan pendapatan pada tingkat menengah dengan pola yang relatif stabil. Sementara itu, usia 45-54 tahun menunjukkan nilai tertinggi, menandakan periode puncak pendapatan. Pada kelompok 55 tahun ke atas, sebaran pendapatan tampak lebih beragam, mencerminkan perbedaan kondisi pekerjaan pada usia lanjut. Secara keseluruhan, grafik menunjukkan bahwa pendapatan cenderung meningkat hingga usia pertengahan sebelum menjadi lebih bervariasi pada usia senior.

3. Gaji Berdasarkan tenaga kerja



Gambar 14. Grafik *Scatter Plot* gaji berdasarkan tenaga kerja

Berdasarkan *scatter plot* pada gambar ini, Tenaga Usaha Penjualan (biru) berada pada rentang nilai tertinggi dan menggambarkan pekerja dengan tingkat pendapatan yang besar serta pertumbuhan gaji yang lebih signifikan dibandingkan dua kelompok lainnya. Titik-titik pada sektor ini menunjukkan peningkatan yang konsisten dari 2024 ke 2025, sehingga mencerminkan potensi pendapatan yang lebih kuat.

Tenaga Usaha Jasa (hijau) menempati rentang nilai menengah dan merepresentasikan pekerja dengan tingkat pendapatan sedang. Perubahan gaji pada kelompok ini tampak stabil, dengan sebaran nilai yang tidak ekstrem. Pola tersebut menunjukkan konsistensi pertumbuhan pendapatan di sektor jasa.



Sementara itu, Tenaga Usaha Pertanian, Kehutanan, dan Perikanan (oren) berada pada rentang nilai terendah, menunjukkan pekerja dengan tingkat pendapatan rendah serta perubahan gaji yang cenderung terbatas. Sebaran titik yang lebih rendah mengindikasikan bahwa peningkatan pendapatan pada kelompok ini tidak sebesar kelompok lainnya.

Kemudian, perbandingan ketiga kelompok tersebut secara keseluruhan memperlihatkan adanya perbedaan yang jelas dalam struktur pendapatan antar sektor tenaga kerja, di mana sektor penjualan tampak paling dominan dalam peningkatan gaji dari tahun ke tahun.

KESIMPULAN

Dalam Penelitian ini menunjukkan bahwa *Cluster 0* adalah kelompok pekerja dengan penghasilan tertinggi, terutama terdiri dari pekerja di sektor penjualan yang berusia antara 30 hingga 59 tahun. Kelompok ini menunjukkan pekerja yang sudah cukup berpengalaman, berada dalam usia produktif, dan bekerja di sektor yang memiliki persaingan ekonomi tinggi.

Cluster 1 terdiri dari pekerja yang sangat muda (usia 15-19 tahun) dan usia lanjut (di atas 55 tahun) dari berbagai sektor, dengan penghasilan paling rendah. Hal ini menunjukkan bahwa usia dan pengalaman kerja memiliki pengaruh besar terhadap produktivitas, sehingga memengaruhi tingkat gaji. Sementara itu, *Cluster 2* adalah kelompok dengan penghasilan menengah yang mencakup berbagai usia produktif dari ketiga sektor, menunjukkan tingkat penghasilan yang lebih stabil dibandingkan *Cluster 1*, meskipun tidak sebesar *Cluster 0*.

Berdasarkan hasil pengelompokan, metode algoritma *K-Means* yang digunakan divalidasi menggunakan perhitungan *Davies-Bouldin Index* (DBI) dengan nilai 0,522. Nilai tersebut menunjukkan bahwa model klaster memiliki kemampuan memisahkan data yang baik, dengan jarak antar klaster yang optimal dan tingkat kekompakan dalam setiap klaster yang memadai. Nilai ini membuktikan bahwa tiga klaster yang terbentuk mampu merepresentasikan dengan tepat pola distribusi data.

DAFTAR PUSTAKA

- Afifah, I. A. N., & Nurdiyanto, H. (2023). Data Mining Clustering Dalam Pengelompokan Buku Perpustakaan Menggunakan Algoritma K-Means. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 8(3), 802-814. <https://doi.org/10.29100/jipi.v8i3.3891>
- Ananda, S. P., Muthoharoh, L., & Pratama, S. D. (2025). Analisis k-means clustering angkatan kerja berdasarkan jenis kelamin, usia, pendidikan, dan pendapatan pada data satuan kerja angkatan nasional tahun 2021. *Prosiding Seminar Nasional Sains Dan Teknologi Seri III*, 2(1), 942-951. <https://conference.ut.ac.id/index.php/saintek/article/view/5130>
- Denny, Goh, T. S., & Margery, E. (2021). Pengaruh Gaji Dan Motivasi Kerja Terhadap Cahaya Abadi Medan. *Jurnal Bisnis Kolega*, 7(2). <https://doi.org/https://doi.org/10.57249/jbk.v7i2.63>
- Jamaludin, A., Widiarto, T., Sutina, S., & Jumaeroh, S. (2024). Pengaruh Usia dan Masa Kerja terhadap Produksi Tenaga Kerja di PT. Galva Kami Industry Cikarang. *Sosio E-Kons*, 16(2),



147. <https://doi.org/10.30998/sosioekons.v16i2.22285>
- Joiwi, & Nainggolan, N. P. (2021). PENGARUH GAJI DAN STATUS KARYAWAN TERHADAP LOYALITAS KARYAWAN PADA PT CITRA BUANA BATAM INDUSTRI. *Retrieved October, 18, 2020.*
https://ejournal.upbatam.ac.id/index.php/scientia_journal/article/view/2913/1882
- Lestari, L. S. (2024). Analisis Clustering Penduduk berdasarkan Kelompok Umur dengan K-Means dan Hierarchical Clustering untuk Perencanaan Demografi. *Jurnal Komputer Dan Informatika*, 19(1), 22-29.
<https://journal.untar.ac.id/index.php/JKI/article/view/34573>
- Mustika, Ardila, Y., Manuhutu, A., Ahmad, N., Hasbi, I., Guntoro, Manuhutu, M. A., Ridwan, M., Hozairi, Wardhani, A. K., Alim, S., Romli, I., Religia, Y., Octafian, D. T., Sufandi, U. U., & Ernawati, I. (2021). Data Mining dan Aplikasinya. In N. Rismawati (Ed.), *Penerbit Widina. WIDINA BHAKTI PERSADA BANDUNG.*
- Rahayu, P. W., Sudipa, I. G. I., Suryani, Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, M. N., Harlina, S., & Maysanjaya, I. M. D. (2024). *Buku ajar data mining* (Efitra (ed.); Issue January). PT. Sonpedia Publishing Indonesia.
https://books.google.co.id/books?hl=id&lr=&id=vCruEAAQBAJ&oi=fnd&pg=PA83&dq=Buku+AJAR+DATA+MINING&ots=TpLL6QpOqB&sig=L-U7CXxD6yVCNGdgAHQio6w89-E&redir_esc=y#v=onepage&q=BUKU AJAR DATA MINING&f=false
- Raihannabil, S. D. (2024). Penerapan Metode Hierarchical Clustering untuk Klasterisasi Provinsi di Indonesia berdasarkan Indikator Status Gizi Anak Baduta (Bawah Dua Tahun) Tahun 2023. *Emerging Statistics and Data Science Journal*, 2(3), 424-436.
<https://doi.org/10.20885/esds.vol2.iss.3.art32>
- Sasmitasari, E., Hidayat, M., & Irwan. (2023). Peran motivasi dalam memediasi gaji dan insentif dalam peningkatan kualitas kerja karyawan 1,2,3. *JURNAL CASH FLOW*, 2, 360-370.
<https://e-jurnal.nobel.ac.id/index.php/jcf/article/view/4102/2175>
- Statistik, B. P. (2024). *KEADAAN PEKERJA DI INDONESIA*. 1(1), 1689-1699.
- Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner Analysis Data Mining Netflix Data Using The Rapid Miner Application. *Journal of Business and Audit Information Systems*, 4(1), 13-21. <https://doi.org/http://dx.doi.org/10.30813/jbase.v4i1.2729> Hasil
- Ulfah, M., & Sri Irtwaty, A. (2022). Penerapan Data Mining Clustering Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan. *FIDELITY: Jurnal Teknik Elektro*, 4(3), 62-68.
<https://doi.org/10.52005/fidelity.v4i3.126>
- Wororowi, J. K., Reba, F., Mandowen, S. A., Sroyer, A. M., Manurung, H. E., Koibur, M. E., Pujiarini, E. H., Edy, M. R., Yuliana, H., Yusuf, M., & Dinata, R. M. (2024). *Data Mining Memahami Data Di Balik Angka* (N. R. Jayanti (ed.)). CV.EUREKA MEDIA AKSARA. PT. Sonpedia Publishing Indonesia