

ANALISIS DAN IMPLEMENTASI DATA MINING UNTUK MENENTUKAN GAJI KARYAWAN TETAP SERTA HONORER MEMAKAI PROSEDUR ALGORITMA K-MEANS CLUSTERING DAN C4.5

David Ardian Darma¹⁾, Iwan Wahyudin²⁾

^{1,2)} Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional
Jalan Sawo Manila, Pasar Minggu, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12520
e-mail: davidardiandarma2018@student.unas.ac.id¹⁾, iwanwahyuddi@civitas.unas.ac.id²⁾

ABSTRAK

PT Bhara Induk Merupakan perusahaan yang berkecimpung dalam bisnis perkebunan kelapa sawit. Di dalam PT bhara induk mempunyai banyak tenaga kerja dimana dalam pekerjaannya berbeda – beda, ada yang berada di kantor pusat dan ada juga yang berada di luar lapangan. Temuan dalam kasusnya perusahaan masih menggunakan mekanisme penginputan data sistem yang masih manual dalam menentukan penggajian, PPH (Pajak Penghasilan) serta buruknya mekanisme pemerataan gaji yang tidak akurat dengan status tunjangan karyawan, gaji pokok karyawan, serta potongan - potongan PPH (Pajak Penghasilan). Temuan lainnya dalam tahap menentukan kesesuaian gaji tenaga kerja PT Bhara Induk masih belum sangat efektif dalam menentukan kesesuaian penggajian, Karena masih menggunakan mekanisme dan aplikasi sistem sederhana seperti Ms.word dan Ms.excel. Maka dalam hal ini untuk menentukan kesesuaian gaji terhadap tenaga kerja tetap dan tenaga kerja honorer. Pada PT Bhara induk dengan menghitung daftar hadir dan cuti perbulan serta status tenaga kerja agar dapat mengetahui berapa hasil dari penetapan kesesuaian gaji karyawan yang di dapatkan dan di peroleh setiap karyawan. Dalam jurnal ini akan lebih menekankan mekanisme pembahasan sistem dan pengaplikasian berdasarkan studi kasus, pada jurnal ini merupakan pengklasifikasian hasil dari penyesuaian gaji karyawan tetap dan karyawan honorer yang di dapat pada PT Bhara Induk dengan menggunakan metode penelitian Algoritma Clustering, dan untuk metode perhitungan menggunakan K-Means serta algoritma C4.5. Dimana dinilai cukup efektif dalam penentuan kesesuaian gaji karyawan honorer dan karyawan tetap. Dan dalam tahap implementasinya maka akan dibuatkan rancang bangun sistem penggajian berbasis website agar mempermudah tenaga kerja bagian keuangan untuk mendaftarkan serta menentukan dan menyesuaikan gaji karyawan berdasarkan statusnya.

Kata Kunci: : Clustering K-Means ,C4.5 ,honor penggajian.

ABSTRACT

PT Bhara Induk is a company engaged in the oil palm plantation business. In PT Bhara Utama, there are many workers who work differently, some are at the head office and some are outside the field. The findings in this case are that the company still uses a manual data entry system in determining payroll, income tax (PPH) and the poor salary distribution mechanism that is not accurate with the status of employee benefits, employee basic salaries, and PPH (Income Tax) deductions. Another finding in the stage of determining the suitability of salaries for PT Bhara Induk's workers is still not very effective in determining the suitability of salaries, because they still use simple system mechanisms and applications such as Ms.word and Ms.excel. So in this case to determine the suitability of salaries for permanent workers and temporary workers. At PT Bhara parent by calculating the attendance list and monthly leave as well as the status of the workforce in order to find out how much the results of determining the suitability of employee salaries are obtained and earned by each employee. This journal will emphasize the mechanism for discussing the system and its application based on case studies, in this journal it is a classification of the results of adjusting the salaries of permanent employees and honorary employees obtained at PT Bhara Induk using the Clustering Algorithm research method, and for the calculation method using K- Means as well as the C4.5 algorithm. Which is considered quite effective in determining the suitability of salaries of honorary employees and permanent employees. And in the implementation stage, a website-based payroll system design will be made to make it easier for the finance department workforce to record and determine and adjust employee salaries based on their status.

Keywords: Clustering K-Means ,C4.5 ,honor salary.

I. PENDAHULUAN

DATA mining ialah satuan proses pengerukan atau pengumpulan data info penting asal kesatuan data yang besar. Proses data mining selalu menggunakan metode statistika, matematika, sampai memanfaatkan teknologi kecerdasan buatan atau *artificial intelligence* [1]. PT Bhara induk merupakan perusahaan yang berkecimpung dalam bisnis perkebunan kelapa sawit. Di dalam PT Bhara Induk sangat lah minim teknologi yang digunakan sangatalah minim sehingga kurangnya efektifitas dalam menentukan honor atau gaji pada karyawan.

Maka, sistem ini akan menjadi sebuah alat bagi instansi untuk membangun sistem informasi penggajian pada PT Bhara Induk. Sistem penggajian karyawan yang dibuat adalah aplikasi berbasis web. Website dibuat menggunakan code igniter. Website penggajian karyawan ini akan dibangun dengan sistem yaitu Sistem Penggajian sebagai alat untuk menghitung gaji atau honor karyawan.

Pada jurnal ini akan memperkecil pembahasan dengan membahas sistem aplikasi penggajian PT Bhara Induk. Pada studi kasus ini, sistem yang dirancang akan memiliki satu user interface dan form penginputan dan serta di masing masing bagian form mempunyai Table laporan. Dalam Jurnal pembuatan aplikasi menggunakan algoritma clustering k-means untuk dan algoritma c4.5

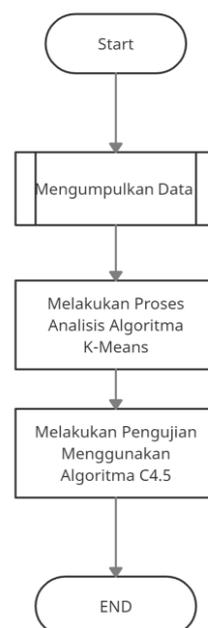
K-means clustering merupakan salah satu metode *cluster analysis* non hirarki yang berpotensi buat mempartisi objek yang terdapat kedalam satu atau lebih *cluster* atau pengelompokkan objek berdasarkan karakteristiknya. sebagai objek yang mempunyai ciri yang sama dikelompokkan dalam satu cluster dan objek yang mempunyai ciri yang tidak sama dikelompokkan kedalam cluster yang lain. [2] Algoritma C4.5 merupakan bagian algoritma yang digunakan untuk membentuk pembagian terstruktur atau pengelompokan di dataset. Dasar dari algoritma C4.5 adalah bentuk pohon keputusan (*Decision Tree*). Cabang-cabang berasal pohon keputusan ialah pertanyaan penjabaran sedangkan untuk daun-daunnya mencerminkan kelas-kelas atau kelompoknya [3].

Permasalahan yang terjadi pada kasus penggajian adalah : Bagaimana cara membedakan gaji honor karyawan tetap dengan karyawan honorer ? Perbedaan gaji dari PPH dan Status Karyawan tetap dan karyawan honorer, Bagaimana penerapan dari metode Clustering K-Mean dan C4.5 ? Penerapan dari Clustering K-means pada jurnal kali ini adalah perbedaan dari gaji karyawan tetap dan gaji karyawan honorer, Bagaimana tahapan pemrosesan dari Clustering K-Mean dan C4.5 ? Tahapan Pemrosesan clustering k-means dan algoritma C4.5 dibedakan menjadi 2 yakni tahap analisis dan pengujian untuk lebih lanjut pembahasan akan dibahas pada bab hasil dan pembahasan, Cara menguji proses perhitungan penggajian menggunakan C4.5 ? Pengujian C4.5 menggunakan aplikasi rapid miner dan pembahasan akan dibahas pada bab hasil dan pembahasan.

Dari penelitian ini mengharapkan beberapa hal sebagai berikut : Membantu dan mengembangkan sistem pada PT Bhara Induk agar dapat menggunakan sistem yang mempermudah dan mengefisiensikan waktu dalam pembuatan laporan penggajian, Mengadaptasikan karyawan pada PT Bhara Induk untuk mengefektif kan pekerjaan agar dapat mengerjakan perkerjaan lainnya dengan cepat dan lancar , untuk mengembangkan kemampuan merancang dan menganalisis kebutuhan yang sedang dibutuhkan oleh perusahaan berkembang ataupun perusahaan maju agar terus berimprovisasi dalam memajukan perusahaannya , Mengembangkan Teknologi yang terdapat pada Perusahaan agar dapat Bersaing dengan perusahaan lain

II. METODE PENELITIAN

A. Tahapan Penelitian



Gambar 1. Tahapan Penelitian

Pada gambar 1 alur dari tahapan penelitian diatas yang dilakukan adalah mengumpulkan data, melakukan proses analisis menggunakan algoritma k-means, melakukan pengujian menggunakan pengujian C4.5 dan selesa.

B. Algoritma Clustering K-Means

Adapun struktur kerja penelitian ini dengan menggunakan metode clustering Algoritma K-Means sebagai berikut :

Analisa Data

Dalam mengumpulkan data dilakukan dengan cara melakukam wawancara serta mengambil data kepada administrasi kantor. [7] . pada penelitian ini faktor yang digunakan adalah faktor Intern. Faktor ini nantinya akan dikelompokkan sehingga membantu dalam memahami data gaji karyawan. Penganalisaan dilakukan berdasarkan fakta dan data yang didapat di PT BHARA INDUK.

Penetapan Data.

Sebelum melakukan proses pengelompokan data Karyawan, mula – mula menentukan data yang akan diolah supaya tujuan penelitian dapat tercapai. memilih Atribut yang dikenakan pada penelitian ini faktor yang mempengaruhi honor karyawan tetap dan karyawan honorer yaitu:

1. Jabatan karyawan
2. Gaji Pokok karyawan
3. Hari kerja karyawan
4. Status karyawan
5. Total gaji karyawan

Proses Clustering Menggunakan Algoritma K-Means

Data yang sudah ditetapkan akan dilakukan pengolahan data dengan proses clustering dengan menggunakan algoritma K-Means sehingga dapat mendapatkan hasil data aga dapat diolah oleh algoritma clustering

Penentuan jumlah cluster

Penentuan jumlah cluster diberlakukan bertujuan untuk mengetahui hasil dari keseluruhan gaji karyawan. Maka dalam analisis ini jumlah cluster yang digunakan ialah sebanyak 6 cluster ($k=6$), sehingga selanjutnya akan diketahui cluster yang optimal dalam melakukan penelitian ini. Berikut merupakan Transformasi Data Sebelum di konversi pada Tabel 1, dan Data yang sudah di konversi pada Tabel II

TABEL I
DATA KARYAWAN SEBELUM DI KONVERSI

no	Nama	Jabatan	Status	Gaji Pokok	Jumlah hari Kerja	Total Gaji
1	Drs Agus Kusnadi	Komisaris	Tetap	4.905.501	20	13.695.991
2	Drs Darsian Rasyid,SH	Direktur	Tetap	13.653.057	20	19.859.073
3	Ir Oneng Noor Dewati	Kabag Ren Min	Tetap	8.791.358	20	14.298.300
4	Yani Herawati,SKom	Kasi Adum/Set	Tetap	3.776.788	20	7.404.282
5	Ahmad Higuaita,SPd	Bendahara	Tetap	3.776.788	20	7.334.889
6	M Susila Raharja,SE	Kasi Pembukaan	Tetap	3.699.749	20	7.260.861
7	Eko Suwarno	Satpam	Tetap	3.098.197	20	5.193.654
8	Syahrullah	Satpam	Tetap	2.618.872	20	4.733.067
9	Moch Yusuf	Pramu Bakti	Tetap	2.093.259	20	4.223.828
10	Eko Adi Purnomo	Staff Renmin	Tetap	4.490.200	20	4.331.332
11	Estranto Purnomo	Staff Renmin	Honoror	5.800.000	20	5.568.000
12	Dicky Noer Diansyah	Pengemudi	Honoror	2.500.000	20	2.500.000

TABEL II
DATA KARYAWAN SESUDAH DI KONVERSI

no	Nama	Jabatan	Status	Gaji Pokok	Jumlah hari Kerja	Total Gaji
1	Drs Agus Kusnadi	Komisaris	1	4	1	2
2	Drs Darsian Rasyid,SH	Direktur	1	2	1	1
3	Ir Oneng Noor Dewati	Kabag Ren Min	1	3	1	2
4	Yani Herawati,SKom	Kasi Adum/Set	1	4	1	4
5	Ahmad Higuaita,SPd	Bendahara	1	4	1	4
6	M Susila Raharja,SE	Kasi Pembukaan	1	4	1	3
7	Eko Suwarno	Satpam	1	4	1	3
8	Syahrullah	Satpam	1	4	1	4
9	Moch Yusuf	Pramu Bakti	1	4	1	4
10	Eko Adi Purnomo	Staff Renmin	1	4	1	4
11	Estranto Purnomo	Staff Renmin	2	3	1	3
12	Dicky Noer Diansyah	Pengemudi	2	4	1	4

Menentukan centroid

Pemilihan pusat awal cluster centroid ditentukan acak, dimana nilai cluster 0 diambil dari baris ke-1, nilai cluster 1 pada baris ke-2. Pada Tabel III merupakan nilai titik poin awal cluster (centroid) yang dapat ditinjau pada tabel berikut :

TABEL III
CENTEROID YANG DITENTUKAN

no	Nama	Jabatan	Status	Gaji Pokok	Jumlah hari Kerja	Total Gaji
2	Drs Darsian Rasyid,SH	Direktur	1	2	1	1
3	Ir Oneng Noor Dewati	Kabag Ren Min	1	3	1	2

Pusat Menghitung jarak antara titik centroid menggunakan titik tiap objek dengan menggunakan *Euclidian Distance*. terdapat kalkulasi centroid awal secara manual. Kalkulasinya ialah sebagai berikut: [8]

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{kn} - X_{kj})^2} \quad (1)$$

Dimana :

x = pusat cluster

y = data

dimana pada x merupakan data pusat cluster pertama dan data y merupakan data cluster yang akan diuji dengan menggunakan data pusat cluster.

C. Algoritma C4.5

Penyeleksian data merupakan pemilihan data yang akan dipakai dalam proses algoritma klasifikasi C4.5. [9]. Tujuan dari penyeleksian data adalah memperoleh himpunan data target, pemilihan pengelompokkan data atau menetapkan data pada *subset variable* atau *sample* data dimana penemuan akan dilakukan. [10]

Sumber Data

Transformasi data ialah proses mengubah atau mengkonversi data ke dalam bentuk yang sesuai. Pada Tabel IV ialah sumber data yang akan diolah oleh algoritma C4.5 :

TABELIV
DATA KARYAWAN SESUDAH DI KONVERSI

no	Nama	Jumlah Kerja	Status	Masa kerja	Rekomendasi
1	Drs Agus Kusnadi	R	T	L	Y
2	Drs Darsian Rasyid,SH	R	T	L	Y
3	Ir Oneng Noor Dewati	R	T	L	Y
4	Yani Herawati,SKom	R	T	B	Y
5	Ahmad Higuita,SPd	R	T	B	Y
6	M Susila Raharja,SE	R	T	B	Y
7	Eko Suwarno	R	T	L	T
8	Syahrullah	R	T	L	T
9	Moch Yusuf	R	T	B	Y
10	Eko Adi Purnomo	R	T	B	Y
11	Estranto Purnomo	R	H	B	T
12	Dicky Noer Diansyah	R	H	B	T

Penentuan Information Gain dan Entropy

Perhitungan semua atribut/variable untuk mendapatkan informasi gain dan Entropy tertinggi yang nantinya akan dijadikan sebagai akar pada pembuatan pohon keputusan [11]

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (3)$$

Keterangan :

S = Himpunan Kasus.

N = jumlah partisi S.

Pi : proporsi dari S, terhadap S

dimana S adalah himpunan kasus atau entropy selanjutnya N digunakan untuk menentukan banyaknya partisi pada data yang digunakan dan pi ialah proporsi berasal S terhadap S . pada tabel V merupakan hasil perhitungan entropy berasal dari sumber data yang ada :

TABEL V
HASIL ENTROPY

	Jumlah(Si)	Ya (Si)	Tidak(Si)	Entropy
Total	12	10	2	0,650022422
Jumlah kerja				
Rajin	12	10	2	0,650022422
Cukup Rajin	12	3	2	0,930827083
Status				
Buruk	0	0	0	0
Tetap	10	10	2	0,464385619
Honorer	12	2	2	0,861654167
Masa Kerja				
Lama	5	5	2	0,528771238
Baru	7	5	2	0,863120569

Berikut adalah rumus dari Entropy :

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

Keterangan:

S = Himpunan Kasus

A = Atribut

n = Jumlah partisi atribut A

|Si| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Pada rumus diatas dijabarkan dimana S dan A merupakan Himpunan kasus data lalu S merupakan hasil Entropy yang sudah di hitung lalu dikurangkan hasil dari partisi pada atribut A dan |Si| adalah Jumlah kasus dari partisi pada data – i dan |S| merupakan jumlah kasus pada data 80S . Pada Tabel VI merupakan hasil dari perhitungan entropy dari sumber data yang ada:

TABEL VI
HASIL GAIN

		Jumlah(SI)	Ya (SI)	Tidak(SI)	Entropy	Gain
Total		12	10	2	0,650022422	
Jumlah kerja						
	Rajin	12	10	2	0,650022422	-0,93082708
	Cukup Rajin	12	3	2	0,930827083	
Status						
	Buruk	0	0	0	0	
	Tetap	10	10	2	0,464385619	-0,59861976
	Honorer	12	2	2	0,861654167	
Masa Kerja						
	Lama	5	5	2	0,528771238	-0,07378593
	Baru	7	5	2	0,863120569	

Pohon Keputusan

Pohon keputusan merupakan hasil berasal entropy dan information gain , sesudah perumusan dilakukan berulang – ulang sampai semua atribut pohon memiliki kelas serta tidak mampu lagi untuk diteruskan proses perhitungan[12]

Aturan Rule /Model

Aturan model pada naskah penjelasan yang mencerminkan sebuah pohon keputusan.

1. Jika Pekerja masuk dan memperoleh predikat Rajin, status kerja tetap, dan masa kerja Lama maka kemungkinan untuk mendapatkan rekomendasi gaji akan besar
2. Jika pekerja masuk dan memperoleh predikat cukup, status kerja tetap dan masa kerja baru maka kemungkinan untuk mendapatkan rekomendasi akan besar
3. Jika pekerja masuk dan memperoleh predikat buruk, status kerja tetap, masa kerja lama maka akan mendapatkan rekomendasi gaji kecil
4. Jika pekerja masuk dan memperoleh predikat rajin atau cukup , statu kerja honorer , dan masa kerja lama maka rekomendasi gaji akan kecil
5. Jika pekerja masuk dan memperoleh preikat buruk,status kerja honorer, masa kerja baru ,maka rekomendasi gaji akan kecil[13]

Analisa Kebutuhan Hardware

Hardware merupakan salah satu kebutuhan dalam pembuatan aplikasi pada penelitian. Berikut merupakan kebutuhan hardware yang digunakan pada Tabel VII:

TABEL VII
HARDWARE YANG DIGUNAKAN

Perrangkat	Spesifikasi
Processor	Intel ® Core™ i3-6006U CPU @2.00GHz (4 CPUs)
RAM	4GB
Harddisk	500GB

Analisa Kebutuhan Software

Selain hardware, software juga dibutuhkan dalam pembuatan website pada penelitian. Berikut merupakan kebutuhan software yang digunakan pada Tabel VIII:

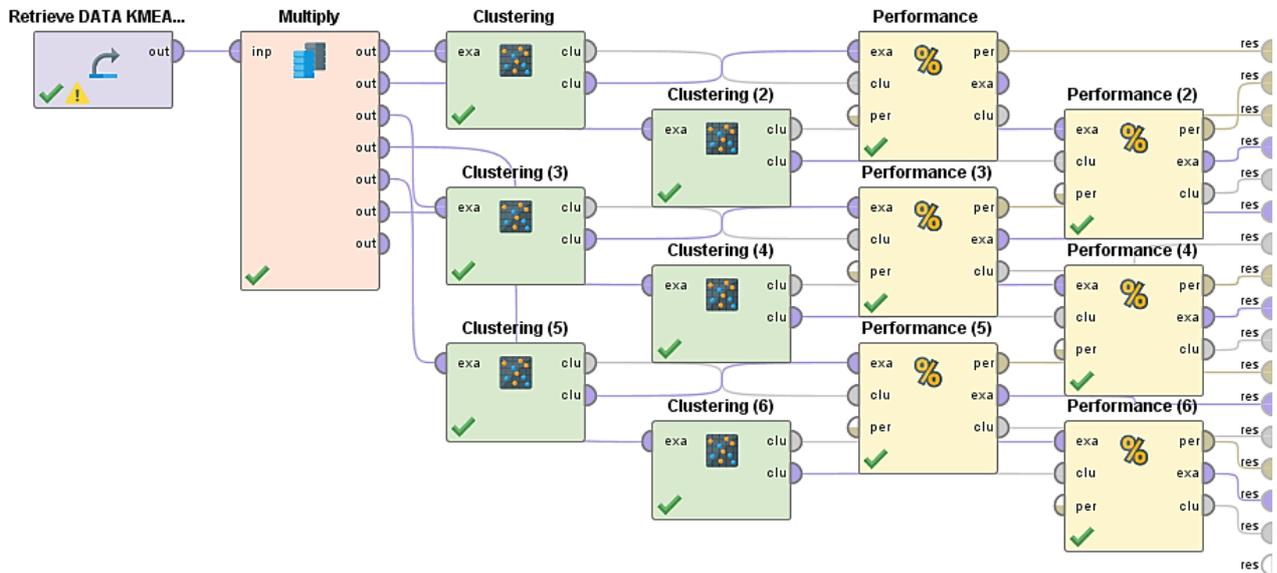
TABEL VIII
 SOFTWARE YANG DIGUNAKAN

Perangkat	Kegunaan
Sistem operasi microsoft windows 10 Professional 64-bit	Sistem operasi yang digunakan peneliti dalam membangun aplikasi ini
Microsoft Visual Studio Code,	Text Editor yang digunakan untuk menuliskan Script kode berupa HTML,CSS, dan PHP untuk membuat website
XAMPP	Sebagai webserver lokal yang digunakan untuk menyimpan database guna untuk menghubungkan antar PHP dan MYSQL
Web Hosting	Merupakan media yang digunakan untuk menghosting atau mengupload hasil dari website yang sudah dibuat agar hasilnya dapat diakses secara skala besar

III. HASIL DAN PEMBAHASAN

A. penggajian menggunakan clustering K-Means

Tahapan Clustrering 1



Gambar 1. Rangkaian Clustering K-mean

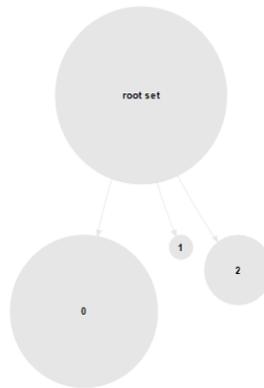
Berikut adalah hasil process dari rapid miner pada gambar diatas ada beberapa komponen yang digunakan yakni menggunakan Retrive data, Multiply, Clustering K-Means, dan Perfomance, Retrive data berfungsi untuk memuat data excel atau csv yang digunakan sebagai sumber data Multiply berfungsi sebagai pengganda data sebelum di transfer atau di buat alurnya ke dalam Clustering Clustering berfungsi sebagai perhitungan pada analisis kali ini dan akan dihasilkan berupa table , graph dan data clusternya Perfomance berfungsi sebagai penampil data yang akan digunakan untuk menampilkan hasil dari Clustrering data dari Clustering K-Means

Tahapan Clustrering 2

Cluster Model

Cluster 0: 8 items
 Cluster 1: 1 items
 Cluster 2: 3 items
 Total number of items: 12

Attribute	cluster_0	cluster_1	cluster_2
Status	1.125	1	1.333
Gaji Pokok	4	2	3.333
Jumlah Hari Kerja	1	1	1
Total Gaji	3.750	1	2.333



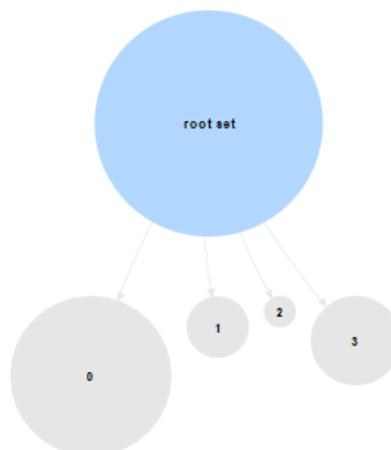
Gambar 2. Cluster Model 2, Graph Tree 2 dan hasil centroid 2

Berikut adalah graphic Tree Clustering k-mean pada cluster ke 2 diperoleh sebagai berikut dimana root set dibawa mendapatkan 3 data graphic yakni 0,1, dan 2 dan hasil dari centroid gambar ke 2 adalah dimana adalah tahapan pertama dalam pengujian clustering k-means karena data masih pengujian pertama maka akan diulangi lagi sampai data tidak berulang dan hasilnya tetap Tahapan Clustrering 3

Cluster Model

Cluster 0: 6 items
 Cluster 1: 2 items
 Cluster 2: 1 items
 Cluster 3: 3 items
 Total number of items: 12

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
Status	1.167	1	2	1
Gaji Pokok	4	2.500	3	4
Jumlah Hari Ke...	1	1	1	1
Total Gaji	4	1.500	3	2.667



Gambar 3 . Cluster Model 3 , Graph Tree Tahap 3 dan Centroid Tabel Tahap 3

Berikut hasil dari cluster 2 untuk model ke 3 dari perhitungan data klustering dapat diperoleh pada data dibawah merupakan cluster model yang diperoleh yakni ada 4 cluster yakni Berikut adalah graphic Tree Clustering k-mean pada cluster ke 3 diperoleh sebagai berikut dimana root set dibawa mendapatkan 4 data graphic yakni 0,1,2, dan 3

dan hasil dari centeroid table ke 3 adalah dimana adalah tahapan ketiga dalam pengujian clustering k-means karena data masih pengujian pertama maka akan diulangi lagi sampai data tidak berulang dan hasilnya tetap

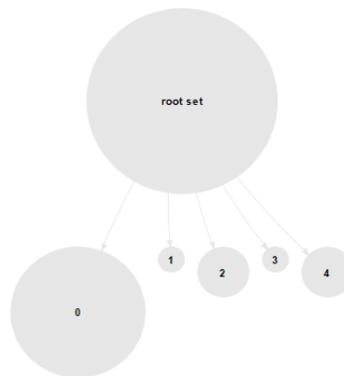
Tahapan Clustrering 4

Berikut hasil dari cluster data ke 3 model ke 4 dari perhitungan data klustering dapat diperoleh pada data dibawah merupakan cluster model yang diperoleh yakni beberapa item dibawah

Cluster Model

Cluster 0: 6 items
 Cluster 1: 1 items
 Cluster 2: 2 items
 Cluster 3: 1 items
 Cluster 4: 2 items
 Total number of items: 12

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Status	1.167	1	1	2	1
Gaji Pokok	4	2	3.500	3	4
Jumlah Hari Kerja	1	1	1	1	1
Total Gaji	4	1	2	3	3



Gambar 4. Cluster Model 4 dan Graph Tree Tahap 4 Centeroid Tabel Model 4

Berikut adalah graphic Tree Clustering k-mean pada cluster ke 4 diperoleh dimana root set dibawa mendapatkan 4 data graphic yakni 0,1,2,3,4 dan hasil dari centeroid table ke 4 adalah dimana adalah tahapan ke 4 dalam pengujian clustering k-means karena data masih pengujian pertama maka akan diulangi lagi sampai data tidak berulang dan hasilnya tetap.

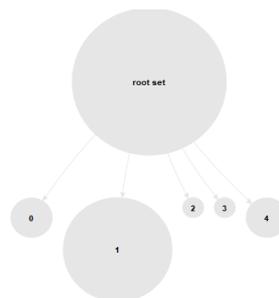
Tahapan Clustrering 5

Berikut hasil dari cluster data ke 4 model ke 5 dari perhitungan data klustering dapat diperoleh pada data dibawah merupakan cluster model yang diperoleh yakni beberapa item dibawah

Cluster Model

Cluster 0: 2 items
 Cluster 1: 6 items
 Cluster 2: 1 items
 Cluster 3: 1 items
 Cluster 4: 2 items
 Total number of items: 12

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Status	1	1.167	1	2	1
Gaji Pokok	3.500	4	2	3	4
Jumlah Hari Kerja	1	1	1	1	1
Total Gaji	2	4	1	3	3



Gambar 5 . Cluster Model 5, Graph Tree Tahap 5 dan Centeroid Tabel 5

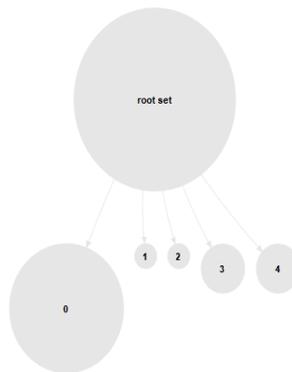
Berikut adalah graphic Tree Clustering k-mean pada cluster ke 5 diperoleh dimana root set dibawa mendapatkan 5 data graphic yakni 0,1,2,3,4 dan hasil dari centeroid table ke 5 adalah dimana adalah tahapan ke 5 dalam pengujian clustering k-means karena data masih pengujian pertama maka akan diulangi lagi sampai data tidak berulang dan hasilnya tetap

Tahapan Clustrering 6

Cluster Model

Cluster 0: 6 items
 Cluster 1: 1 items
 Cluster 2: 1 items
 Cluster 3: 2 items
 Cluster 4: 2 items
 Total number of items: 12

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Status	1.167	1	2	1	1
Gaji Pokok	4	2	3	3.500	4
Jumlah Hari Kerja	1	1	1	1	1
Total Gaji	4	1	3	2	3



Gambar 6. Cluster Model 6 , Graph Tree Tahap 6, dan hasil centeroid 6

Berikut hasil dari cluster data ke 5 model ke 6 dari perhitungan data klustering dapat diperoleh pada data dibawah merupakan cluster model yang diperoleh yakni beberapa item dibawah dan hasil graphic Tree Clustering k-mean pada cluster ke 6 diperoleh dimana root set dibawa mendapatkan 5 data graphic yakni 0,1,2,3,4 dan hasil dari centeroid table ke 6 adalah tahap terakhir pengujian dan data yang didapat sudah terakhir dan tidak berubah lagi.

B. Rancangan Design

1. Data Penggajian

Data Penggajian Karyawan

Copy CSV Excel PDF Print

Search:

Nomor	Nama Karyawan	Jabatan	Gaji Pokok	Honor Tunjangan	Potongan IURAN JHT	Total Gaji Tunjangan	Jumlah Gaji JHT	PPH21	Honor Diterima	Action
PER1494769662	Ade Muhammad Nur Fauzi	Komisaris	Rp 100.000.000	Rp 10.000.000	Rp 2.310.000	Rp 115.500.000	Rp 113.190.000	Rp 5.659.500	Rp 107.530.500	Edit
PER1629463237	Amir Simalakama	Manager juga	Rp 10.000.000	Rp 5.000.000	Rp 309.000	Rp 15.450.000	Rp 15.141.000	Rp 757.050	Rp 14.383.950	Edit
PER373296506	David Ardian Darma	direktur	Rp 0	Rp 0	Rp 0	Rp 0	Rp 0	Rp 0	Rp 0	Edit
PER797314576	David Ardian	Manager	Rp 5.000.000	Rp 500.000	Rp 113.300	Rp 5.665.000	Rp 5.551.700	Rp 277.585	Rp 5.274.115	Edit

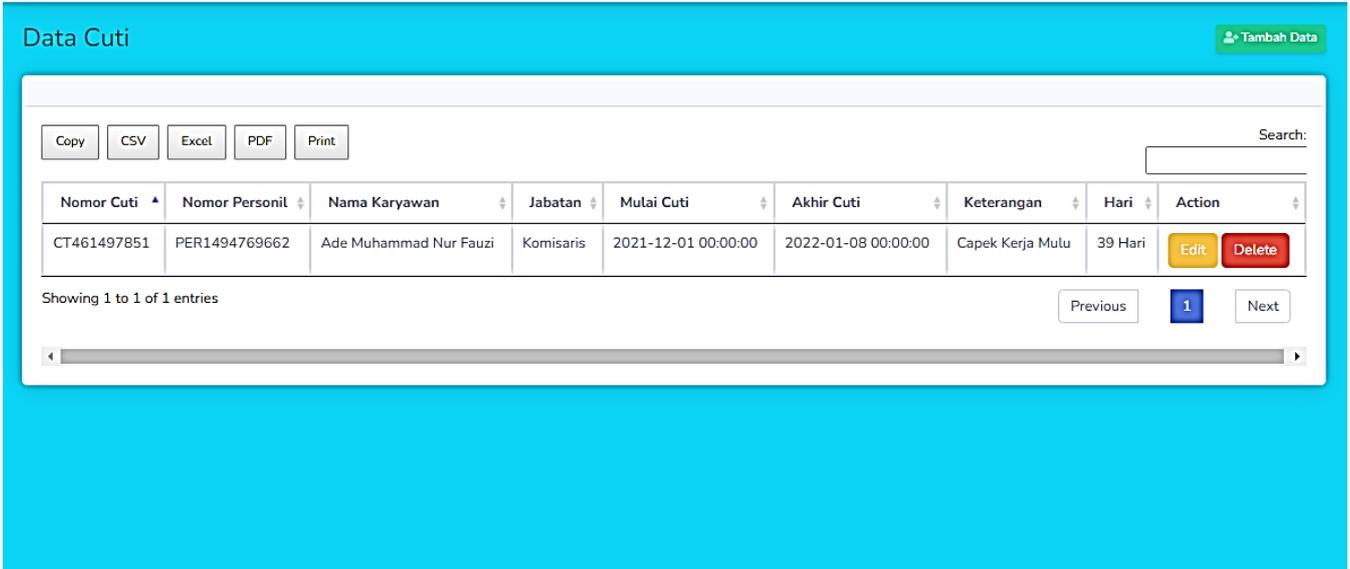
Showing 1 to 4 of 4 entries

Previous 1 Next

Gambar 7 .Hasil Data Penggajian

Dalam data penggajian akan dilakukan penginputan semua data gaji setiap karyawan dan akan dilakukan perhitungan data karyawan tetap dan honorer serta hasil diatas merupakan hasil data gaji yang sudah diinput dimana menampilkan nomor yang berupa nomor identitas karyawan, nama karyawan, jabatan, gaji pokok, honor tunjangan, potongan iuran jht, jumlah gaji jht, PPH 21, dan Honor yang diterima oleh karyawan

2. Data Cuti



Data Cuti + Tambah Data

Copy CSV Excel PDF Print Search:

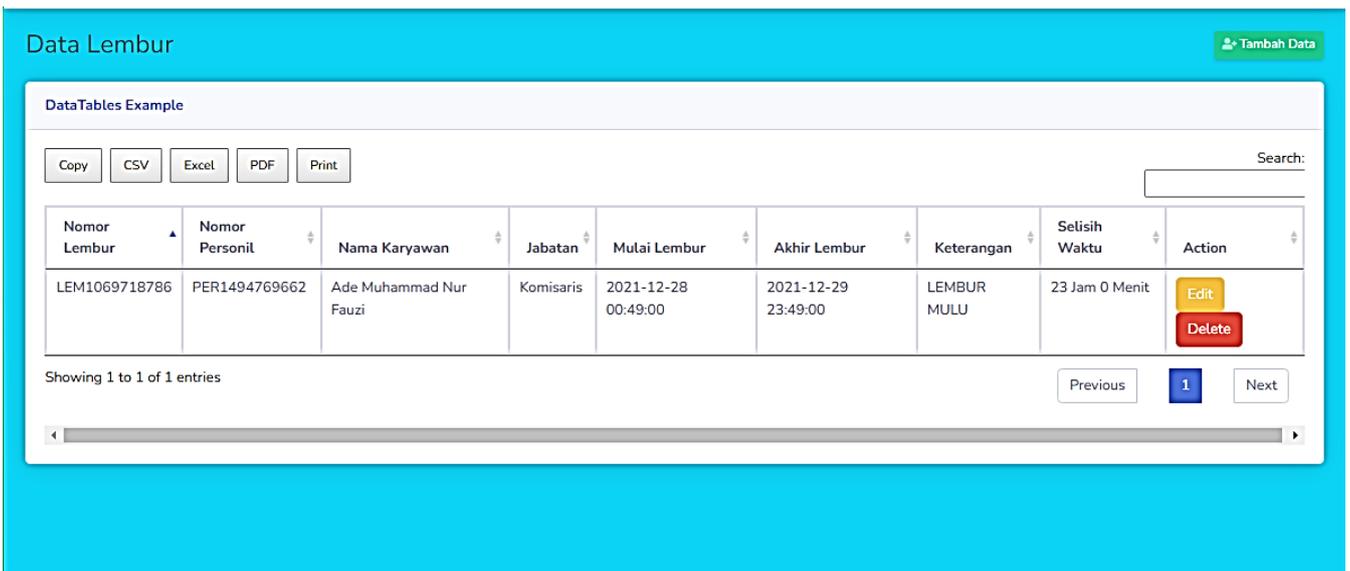
Nomor Cuti	Nomor Personil	Nama Karyawan	Jabatan	Mulai Cuti	Akhir Cuti	Keterangan	Hari	Action
CT461497851	PER1494769662	Ade Muhammad Nur Fauzi	Komisaris	2021-12-01 00:00:00	2022-01-08 00:00:00	Capek Kerja Mulu	39 Hari	Edit Delete

Showing 1 to 1 of 1 entries Previous **1** Next

Gambar 8. Hasil Data Cuti

Pada gambar diatas merupakan hasil data cuti yang digunakan untuk melihat dan menginput data cuti yang didalamnya terdiri dari nomor cuti, nomor personil digunakan untuk memanggil data personil atau karyawan, nama karyawannya, jabatan, mulai cuti, akhir cuti, keterangan yang merupakan kenapa karyawan melakukan cuti, hari yang merupakan berapa hari cuti karyawan tersebut.

3. Data Lembur



Data Lembur + Tambah Data

DataTables Example

Copy CSV Excel PDF Print Search:

Nomor Lembur	Nomor Personil	Nama Karyawan	Jabatan	Mulai Lembur	Akhir Lembur	Keterangan	Setisih Waktu	Action
LEM1069718786	PER1494769662	Ade Muhammad Nur Fauzi	Komisaris	2021-12-28 00:49:00	2021-12-29 23:49:00	LEMBUR MULU	23 Jam 0 Menit	Edit Delete

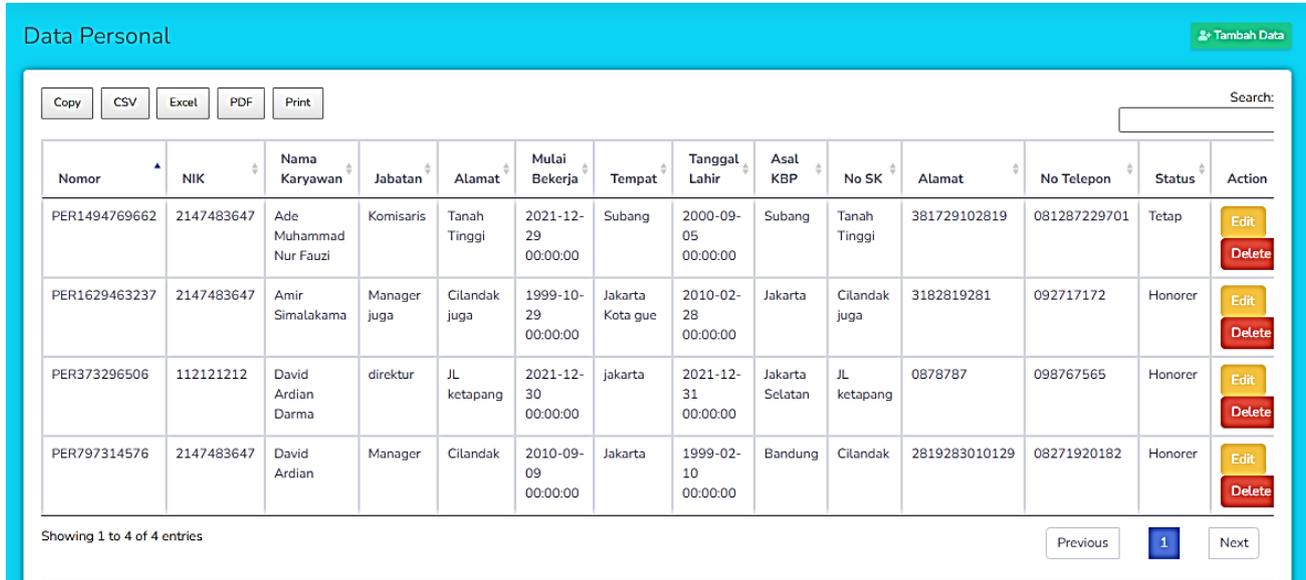
Showing 1 to 1 of 1 entries Previous **1** Next

Gambar 9. Hasil Data Lembur

Pada gambar diatas menjelaskan hasil data lembur yang digunakan untuk menginput data lembur dan melihat hasil data lembur serta dimana diperoleh beberapa data dalam data lembur ini yakni adalah nomor lembur , nomor personil

merupakan data unik yang digunakan untuk mengambil data karyawan, nama karyawan, jabatan, mulai lembur, akhir lembur, keterangan dari karyawan yang sedang melakukan lembur, selisih waktu berapa lama karyawan lakukan

4. Data Personil



Nomor	NIK	Nama Karyawan	Jabatan	Alamat	Mulai Bekerja	Tempat	Tanggal Lahir	Asal KBP	No SK	Alamat	No Telepon	Status	Action
PER1494769662	2147483647	Ade Muhammad Nur Fauzi	Komisaris	Tanah Tinggi	2021-12-29 00:00:00	Subang	2000-09-05 00:00:00	Subang	Tanah Tinggi	381729102819	081287229701	Tetap	Edit Delete
PER1629463237	2147483647	Amir Simalakama	Manager juga	Cilandak juga	1999-10-29 00:00:00	Jakarta Kota gue	2010-02-28 00:00:00	Jakarta	Cilandak juga	3182819281	092717172	Honoror	Edit Delete
PER373296506	112121212	David Ardian Darma	direktur	JL ketapang	2021-12-30 00:00:00	jakarta	2021-12-31 00:00:00	Jakarta Selatan	JL ketapang	0878787	098767565	Honoror	Edit Delete
PER797314576	2147483647	David Ardian	Manager	Cilandak	2010-09-09 00:00:00	Jakarta	1999-02-10 00:00:00	Bandung	Cilandak	2819283010129	08271920182	Honoror	Edit Delete

Gambar 10. Hasil Data Personil

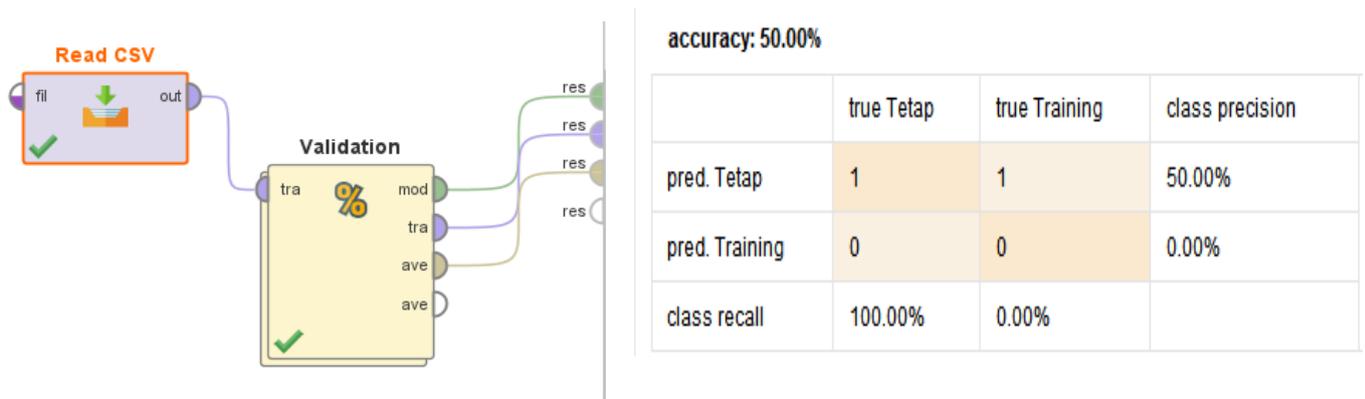
Pada gambar diatas merupakan gambar hasil data personil berguna untuk mendata karyawan dan dalam hasil diatas diperoleh beberapa data yakni nomor sebagai unik data, NIK merupakan Nomor Induk Karyawan, Nama karyawan, Jabatan yang diperoleh, alamat tinggal, mulai bekerja di PT Bhara Induk, tanggal lahir, asal kbp, no telepon , status.

C. Tahap Pengujian

Berikut tahapan dari pengujian menggunakan rapid miner. Menggunakan data yang sudah ada seperti pada Pada Tabel IX data untuk pengujian penggajian data bhara induk menggunakan klasifikasi C4.5. diperoleh data nama, jumlah kerja dan status, masa kerja, dan rekomendasi

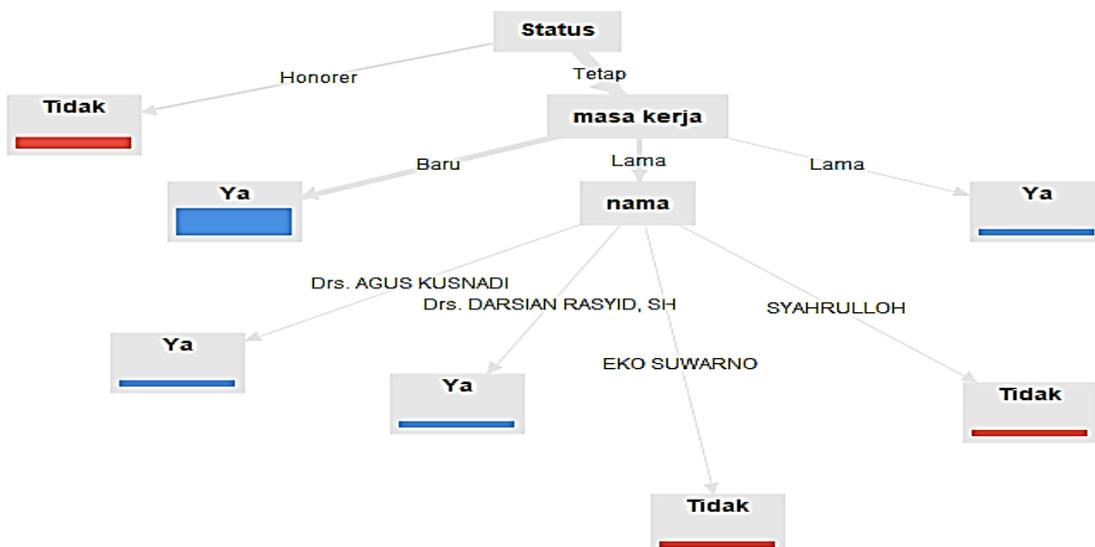
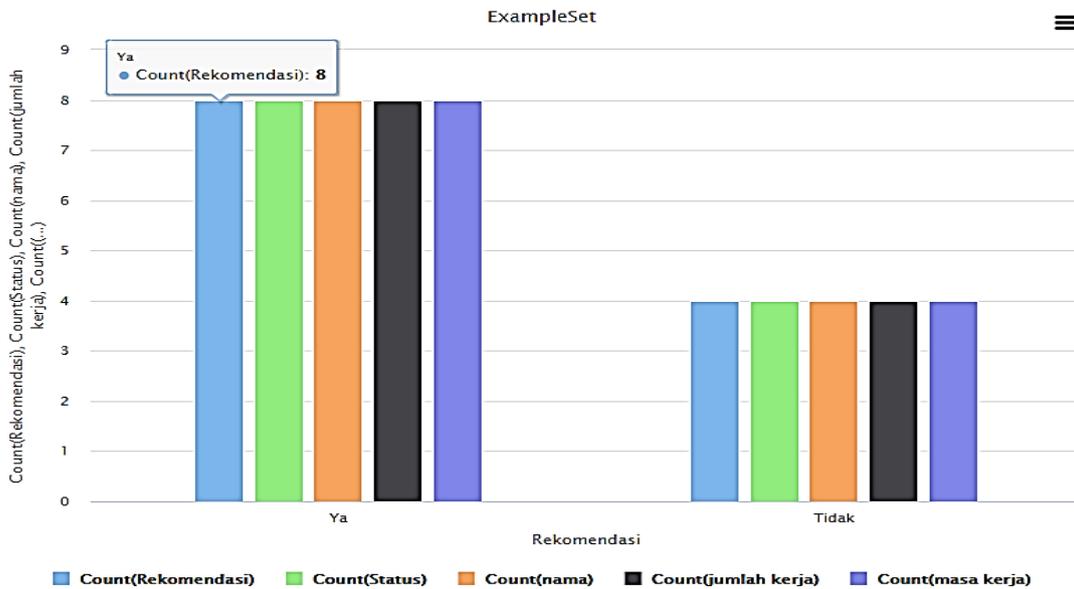
TABEL IX
DATA KARYAWANA YANG DIUJI

no	Nama	Jumlah Kerja	Status	Masa kerja	Rekomendasi
1	Drs Agus Kusnadi	R	T	L	Y
2	Drs Darsian Rasyid,SH	R	T	L	Y
3	Ir Oneng Noor Dewati	R	T	L	Y
4	Yani Herawati,SKom	R	T	B	Y
5	Ahmad Higuaita,SPd	CR	T	B	Y
6	M Susila Raharja,SE	CR	T	B	Y
7	Eko Suwarno	R	T	L	T
8	Syahrullah	R	T	L	T
9	Moch Yusuf	R	T	B	Y
10	Eko Adi Purnomo	CR	T	B	Y
11	Estranto Purnomo	CR	H	B	T
12	Dicky Noer Diansyah	CR	H	B	T



Gambar 11. Pengujian Menggunakan Rapid miner dan Hasil Performance Vector dari Pengujian C4.5

Pada gambar diatas menunjukkan alur dari pengujian algoritma c4.5 menggunakan rapid miner dengan tahapan yakni mengambil data yang akan diuji berupa csv lalu akan diproses untuk mendapatkan hasil pohon keputusan dan hasil dari Performance Vector dari hasil analisa menggunakan rapid miner pada data Bhara Induk.



Gambar 20. Rekomendasi dari C4.5 dan Pohon Keputusan C4.5

Berikut hasil graphic dari rekomendasi yang diperoleh yakni 8 yang mendapatkan rekomendasi gaji tetap dan 4 mendapatkan gaji honorer dan hasil dari pohon keputusan yang sudah diuji menggunakan aplikasi rapid miner dengan algoritma C4.5 dari hasil yang diharapkan maka diperoleh masa kerja dan status kerja sangat diperlukan untuk membedakan honorer dan tetap [14]

IV. KESIMPULAN

Kesimpulan dari jurnal dan pembahasan kali ini ialah data mining teknik clusterin k-means dan pengujian data menggunakan algoritma C4.5 efektif dalam menentukan gaji pada karyawan tetap dan honorer . serta dalam pengimplementasiannya menggunakan website sebagai media penginputan dan seleksi data penggajian berdasarkan status sangatlah efektif pada perusahaan sehingga dapat meningkatkan efektifitas kinerja pada PT Bhara Indah.

DAFTAR PUSTAKA

- [1] M. R. Adani, "Mengetahui Konsep Data Mining Beserta Contoh Implementasinya," 2021. <https://www.sekawanmedia.co.id/data-mining/>.
- [2] Syafnidawaty, "K-MEANS CLUSTERING," 2020, [Online]. Available: <https://raharja.ac.id/2020/04/19/k-means-clustering/>.
- [3] M. I. MUBAROK, "ALGORITMA C4.5," 2018. <https://muhammadilhammubarak.wordpress.com/2018/08/14/algoritma-c4-5/>.
- [4] F. Johan and S. A. Armono, "Pemanfaatan Algoritma C4.5 Untuk Menentukan Bonus Tahunan Pada Toko Umamo Vero," vol. 02, pp. 1–5, 2020.
- [5] W. Purba, W. Siawin, and . H., "Implementasi Data Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi Phk Dengan Algoritma K-Means Clustering," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 85–90, 2019, doi: 10.34012/jusikom.v2i2.429.
- [6] Y. Irawan, "Penerapan data mining untuk evaluasi data penjualan menggunakan metode clustering dan algoritma hirarki divisive," *Jtiilm*, vol. 3, no. 1, pp. 13–20, 2019.
- [7] N. Y. S. Munti, G. W. Nurcahyo, and J. Santony, "Analisis Dan Penerapan Data Mining Untuk Menentukan Gaji Karyawan Tetap Dan Karyawan Kontrak Menggunakan Algoritma K-Means Clustering (Studi Kasus Di Pt Indomex Dwijaya Lestari)," *JITI, Vol. 1, No. 1, Maret 2018*, 2018.
- [8] E. J. Salam, "K-Means Clustering Algorithm," 2017. <https://ekojunaidisalam.com/2017/02/09/k-means-clustering-algorithm/>.
- [9] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," *J. Ultim.*, vol. 6, no. 1, pp. 15–20, 2014, doi: 10.31937/ti.v6i1.327.
- [10] R. Novita, "Teknik Data Mining : Algoritma C4.5," *Ilmu Komputer.com*, pp. 1–12, 2016, [Online]. Available: <https://ilmukomputer.org/wp-content/uploads/2019/06/Rina-Algoritma-C45.pdf>.
- [11] ilmuskripsi, "Algoritma C4.5." <https://www.ilmuskripsi.com/2016/07/algoritma-c45.html>.
- [12] Cahya , "Contoh implementasi Data Mining dengan Algoritma C4.5," *February 9, 2017*, 2017. <http://cahyadsn.phpindonesia.id/extra/c45.php>.
- [13] M. F. Arifin and D. Fitriana, "Penerapan Algoritma Klasifikasi C4.5 Dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada," *InComTech*, vol. 8, no. 2, pp. 87–102, 2018, doi: 10.22441/incomtech.v8i1.2198.
- [14] R. M. Hijriana, Nadiya, "Penerapan Metode Decision Tree Algoritma C4. 5 Untuk Klasifikasi Mahasiswa Berprestasi," *Al Ulum J. Sains Dan Teknol.*, vol. 2, no. 1, pp. 39–43, 2016.