

Penerapan Data Mining Dalam Menentukan Standard Kelayakan Penerima Insentif Pegawai PT. Sinarmas Multifinance Menggunakan Algoritma K-Means

William Saputra^{1*}, Sumarno², Sundari Retno Andani³, Indra Gunawan⁴, Zulaini Masruro Nasution⁵

^{1,2,3,4,5}Teknik Informatika; STIKOM Tunas Bangsa Pematangsiantar; Jl. Jend. Sudirman Blok A No.1,2,&3 Pematangsiantar

e-mail: ¹williamsaputra77@gmail.com, ²sumarno@amiktunasbangsa.ac.id,

³sundari.ra@amiktunasbangsa.ac.id, ⁴indra@amiktunasbangsa.ac.id, ⁵zulaini@amiktunasbangsa.ac.id

Abstrak: Penerapan Data Mining digunakan untuk mengatasi permasalahan dalam penentuan insentif pegawai setiap bulan di PT Sinarmas Multifinance. PT Sinarmas Multifinance adalah sebuah perusahaan yang bergerak dalam jasa usaha pembiayaan sewa guna usaha, anjak piutang dan pembiayaan konsumen. Dalam menentukan jumlah insentif pegawai, data yang digunakan sebagai kriteria adalah kedisiplinan, integritas, tanggung jawab, komunikasi, antusiasme dalam bekerja masing-masing pegawai. Metode yang digunakan adalah penerapan algoritma K- Means dimana membutuhkan beberapa kriteria. Penerapan Data Mining ini dibuat dengan menggunakan Rapid Miner. Hasil implementasi dari Penerapan Data Mining ini dapat mempermudah pihak Manager (HRD) PT Sinarmas Multifinance dalam menentukan insentif pegawai dengan tepat dan akurat. Penilaian terhadap suatu kedisiplinan merupakan suatu tahap evaluasi kinerja yang penting bagi kelangsungan aktivitas perusahaan. Pemantauan dan penilaian terhadap suatu kedisiplinan pegawai harus dilakukan secara terus-menerus agar dapat meningkatkan kualitas sumber daya manusia. Selanjutnya data pegawai tersebut diolah menggunakan software RapidMiner.

Kata kunci: Data mining, K-means, Insentif, Karyawan, Rapidminer

Abstract: The application of Data Mining is used to overcome problems in determining employee incentives every month at PT Sinarmas Multifinance. PT Sinarmas Multifinance is a company engaged in business leasing, factoring and consumer financing services. In determining the number of employee incentives, the data used as criteria are discipline, integrity, responsibility, communication, and enthusiasm in each employee's work. The method used is the application of the K-Means algorithm which requires several criteria. This Data Mining application is made using Rapid Miner. The results of the implementation of the Data Mining Application can facilitate the Manager (HRD) of PT Sinarmas Multifinance in determining employee incentives correctly and accurately. Assessment of a discipline is an important stage of performance evaluation for the continuity of the company's activities. Monitoring and assessment of an employee's discipline must be carried out continuously in order to improve the quality of human resources. Furthermore, the employee data is processed using RapidMiner software.

Keywords: Data Mining, K-means, Incentives, Employees, Rapidminer

1. Pendahuluan

Teknik *clustering* penting digunakan untuk mempermudah pihak manajemen dalam menentukan tingkat kedisiplinan pada pegawai. *Clustering* memiliki peran sebagai salah satu metode yang digunakan pada Data Mining dimana proses ini dapat mengelompokkan dan mencari data yang mempunyai karakteristik mirip antara data satu dengan data lainnya. Data Mining merupakan bagian penting dalam mencari informasi menarik atau pola suatu data yang dapat juga dikenal dengan istilah sebagai ilmu pengetahuan di dalam suatu database dengan menggunakan teknik tertentu. Data Mining merupakan proses menggunakan kecerdasan buatan, *machine learning*, teknik *statistic* dan matematika dalam mengidentifikasi dan menemukan informasi yang bermanfaat bagi pengetahuan yang terkait dengan berbagai database berkapasitas besar. Algoritma K-

Means adalah salah satu metode atau teknik yang terdapat dalam data mining untuk mengelompokkan (*Clustering*) data kedalam beberapa kelompok berdasarkan jarak, kriteria, kondisi atau karakteristik.

Karyawan atau tenaga kerja sebagai sumber daya manusia pada suatu organisasi atau perusahaan adalah faktor yang paling penting dalam usaha pencapaian tujuan perusahaan. Agar dapat mencapai tujuan tersebut pemimpin perusahaan harus memberikan perhatian khusus kepada karyawan, Penilaian terhadap suatu kedisiplinan merupakan suatu hal yang penting untuk menjadi bahan evaluasi kinerja yang memiliki peran untuk kelangsungan aktivitas perusahaan. Pemantauan dan penilaian terhadap suatu kedisiplinan pegawai harus menjadi rutinitas agar dapat meningkatkan kualitas sumber daya manusia. Penelitian ini dilakukan untuk membuat standarisasi pemberian insentif berdasarkan tingkat kedisiplinan pegawai. Sistem yang berjalan saat ini pada pemberian insentif masih menggunakan Microsoft Excel dengan input data secara manual dan kebijakan dari pada pimpinan PT. Sinarmas Multifinance, dengan begini masih sering terjadi data kesalahan dalam memberikan insentif yang tidak tepat sasaran.

2. Metode Penelitian

2.1. Data Mining

Data Mining merupakan suatu rangkaian proses untuk menggali nilai tambah dari sekumpulan data yang berupa pengetahuan yang selama ini tidak diketahui secara manual (Handoko, 2016). Menurut (Maulida, 2018) Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database.

Data mining memiliki pengertian lain yaitu *knowledge discovery* ataupun *pattern recognition* merupakan suatu istilah yang digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang berukuran sangat besar. Tujuan utama data mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki (Werdiningsih et al., 2020).

Menurut (Prasetyowati, 2017) Data Mining atau yang disebut juga *Knowledge Discovery in Database* (KDD) merupakan suatu proses penggunaan data yang bertujuan untuk menemukan pola atau hubungan dari sekumpulan data yang berupa pengetahuan yang selama ini tidak diketahui.

Menurut (Saragih et al., 2021) Data mining adalah proses dalam menganalisa data dari perspektif yang berbeda dan menyimpulkannya sebuah informasi-informasi penting yang dapat dipakai untuk keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya.

2.2. Algoritma K-Means

Menurut Pralita dalam (Putra & Wadisman, 2018) K-Means merupakan algoritma clustering yang berulang-ulang. Menurut (Sibuea & Sapta, 2017) Dalam menentukan nilai centroid untuk awal iterasi, nilai awal centroid dilakukan secara acak. Sedangkan jika menentukan nilai centroid yang merupakan tahap dari iterasi, maka digunakan rumus persamaan 1:

$$d_e = \sqrt{(x_i - s_j)^2 + (y_i - t_i)^2} \dots\dots\dots (1)$$

1. Menghitung jarak antara titik centroid dengan titik tiap objek.
2. Pengelompokan object untuk menentukan anggota *cluster* adalah dengan memperhitungkan jarak minimum objek.
3. Kembali ke tahap ke-2, lakukan perulangan hingga nilai centroid yang dihasilkan tetap dan anggota *cluster* tidak berpindah ke cluster lain.

Metode K-Means *Clustering* hanya bisa mengolah data dalam bentuk angka, maka untuk data yang berbentuk nominal harus diinisialisasikan terlebih dahulu ke dalam bentuk angka (Sibuea & Sapta, 2017) Langkahnya adalah:

1. Urutkan data berdasarkan frekuensi kemunculannya.
2. Inisialisasikan data tersebut mulai dari data tertinggi dengan nilai 1. Kemudian data selanjutnya 2, 3 dan seterusnya.

2.3. Insentif

Menurut (Iskandar & Pandoyo, 2020) Insentif merupakan perangsang atau pendorong yang diberikan kepada karyawan sehingga dapat meraih prestasi di lingkungan kerjanya. Hal ini tidak dapat

lepas dalam rangka menunjang pemenuhan kebutuhan hidup sehari-hari. Insentif adalah tambahan balas jasa yang diberikan kepada karyawan tertentu yang prestasinya di atas prestasi standar (Ayu & Sinaulan, 2019).

Menurut (Ardian, 2019) Kebutuhan karyawan sebagai individu dapat berupa materi dan non materil, masalah kebutuhan ini dapat menjadi pendorong manusia untuk bekerja atau dapat menyebabkan karyawan lebih bersemangat dalam melakukan pekerjaan dengan mengharapkan, memperoleh imbalan balas jasa dari perusahaan tempat ia bekerja untuk memenuhi kebutuhannya tersebut. Salah satu balas jasa yang biasanya diberikan adalah insentif.

2.4. Analisa Data

Tahap proses data merupakan tahap seleksi data yang bertujuan untuk mendapatkan data yang bersih dan siap untuk digunakan dalam penelitian. Dalam seleksi data harus memperhatikan konsistensi data, *missing value* dan *redundant* data. Lalu data dilakukan pengolahan menggunakan *Rapid Miner* yaitu *Performance* yang berguna untuk validasi dan reabilitas data untuk menemukan keakuratan data.

Dalam penelitian ini data yang digunakan merupakan data karyawan yang diperoleh dari PT Sinarmas *Multifinance* dengan lima variabel sebagai standart kelayakan penerima insentif. Berikut ini merupakan data karyawan yang diperoleh dari PT. Sinarmas *Multifinance* pada tahun 2020 yang dapat dilihat pada Tabel 1.

Tabel 1: Data Pegawai Yang Akan Mendapat Insentif Sesuai Standart

No	Nama Pegawai	Kedisipilanan	Integritas	Tanggung Jawab	Komunikasi	Antusiasme dalam Bekerja
1	Viky Ferdinan Saragih	Sangat Baik	Sangat Baik	Sangat Baik	Baik	Sangat Baik
2	Muhammad Safii	Sangat Baik	Cukup	Sangat Baik	Sangat Baik	Baik
3	Julvan Franata Sinaga	Baik	Baik	Baik	Baik	Baik
4	Firman Simanjuntak	Sangat Baik	Cukup	Baik	Sangat Baik	Baik
5	Daniel W A Manurung	Cukup	Baik	Baik	Cukup	Cukup
6	Anwar Pardede	Baik	Sangat Baik	Cukup	Baik	Sangat Baik
7	Romauli Saragih	Sangat Baik	Cukup	Sangat Baik	Sangat Baik	Baik
8	Irwan Syahputra Sitio	Sangat Baik	Sangat Baik	Cukup	Sangat Baik	Baik
9	Rosmiyanti Simbolon	Sangat Baik	Sangat Baik	Sangat Baik	Baik	Baik
10	Dony M F Siregar	Baik	Baik	Baik	Cukup	Cukup
11	Justro Marbun	Sangat Baik	Cukup	Baik	Baik	Baik
12	Brian Aryanto Tambunan	Cukup	Baik	Baik	Sangat Baik	Sangat Baik
13	Sahrul Parlindungan	Baik	Sangat Baik	Cukup	Cukup	Cukup
14	Andi Samuel Sinaga	Sangat Baik	Cukup	Sangat Baik	Sangat Baik	Sangat Baik

15	Esbon Yosanda Purba	Sangat Baik	Cukup	Sangat Baik	Sangat Baik	Sangat Baik
16	Meisiska	Sangat Baik	Sangat Baik	Sangat Baik	Baik	Sangat Baik
17	Jefri Anto Damanik	Baik	Baik	Baik	Baik	Baik
18	Rizki Rahmayadi	Sangat Baik	Cukup	Baik	Baik	Sangat Baik
19	Abdullah Saleh	Cukup	Baik	Baik	Cukup	Cukup
20	Afrizal Ariadi	Baik	Sangat Baik	Cukup	Sangat Baik	Baik

3. Hasil dan Pembahasan

3.1. Hasil

Dalam proses penelitian penulis menggunakan data pegawai PT. Sinarmas *Multifinance* yang akan mendapat insentif sesuai standart sebagai bahan untuk menyelesaikan penelitian yang diambil dari PT. Sinarmas *Multifinance* pada tahun 2020. Untuk mendukung penelitian penulis menggunakan bantuan *software RapidMiner* 8.1 dalam pengimplementasian data agar mendapatkan hasil yang sesuai dari perhitungan manual dengan aplikasi. data pegawai PT. Sinarmas *Multifinance* akan di ubah ke dalam format data *excel* 2010 kemudian data dimasukkan kedalam rumus algoritma *K-Means* untuk mendapatkan *cluster* tertinggi, Sedang dan terendah.

Data yang digunakan dalam penelitian adalah data pegawai PT. Sinarmas *Multifinance* yang akan mendapatkan insentif tahun 2020. Berikut Tabel 2 adalah data yang akan digunakan untuk pengolahan *K-Means*.

Tabel 2 : Konversi Nilai Data Pegawai Yang Akan Mendapat Insentif Sesuai Standart

No	Nama Pegawai	Kedisiplinan	Integritas	Tanggung Jawab	Komunikasi	Antusiasme dalam Bekerja	Jumlah
1	Vicky Ferdinand Saragih	4	4	4	3	4	19
2	Muhammad Safii	4	2	4	4	3	17
3	Julivan Franata Sinaga	3	3	3	3	3	15
4	Firman SimanJunitak	4	2	3	4	3	16
5	Daniel W A Manullang	2	3	3	2	2	12
6	Anwar Pardede	3	4	2	3	4	16
7	Romauli Saragih	4	2	4	4	3	17
8	Irwan Syahputra Sitio	4	4	2	4	3	17
9	Rosmiyanti Simbolon	4	4	4	3	3	18
10	Donny M F	3	3	3	2	2	13

	Siregar						
11	Justro Maretbun	4	2	3	3	3	15
12	Brian Aryanto Tambunan	2	3	3	4	4	16
13	Sahrul Parlindungan	3	4	2	2	2	13
14	Andi Samuel Sinaga	4	2	4	4	4	18
15	Esbon Yosanda Purba	4	2	4	4	4	18
16	Meisiska	4	4	4	3	4	19
17	Jefri Anto Damanik	3	3	3	3	3	15
18	Rizki Rahmayadi	4	2	3	3	4	16
19	Abdullah Saleh	2	3	3	2	2	12
20	Afrizal Ariadi	3	4	2	4	3	16

Dalam penelitian Jumlah *Cluster* yang digunakan pada data akta kelahiran sebanyak 3 *Cluster*. *Cluster* tersebut diantaranya Ditolak dari Target (*C1*), Diterima dengan pertimbangan (*C2*) dan Diterima (*C3*) berdasarkan data insentif pegawai PT. Sinarmas *Multifinance*.

Selanjutnya, dilakukan perhitungan pusat *cluster* awal dan hasil perhitungan iterasi 1 dapat dilihat pada Tabel 3 dan jarak *centroid* iterasi ke-1 dapat dilihat pada Tabel 4 :

Tabel 3 : *Centroid*

<i>Centroid</i>	Kedisiplinan	Integritas	Tanggung Jawab	Komunikasi	Antusiasme dalam Bekerja
Di ambil data ke-1 sebagai pusat cluster ke-1	4	4	4	3	4
Di ambil data ke-17 sebagai pusat cluster ke-2	3	3	3	3	3
Di ambil data ke-19 sebagai pusat cluster ke-3	2	3	3	2	2

Tabel 4 : Jarak *Centroid* Iterasi Ke-1

C1	C2	C3	Jarak <i>Centroid</i>
0,0000	2,7321	6,6458	C1
3,2361	2,0000	4,1623	C2
2,7321	0,0000	2,4142	C2
3,4495	1,7321	4,0000	C2
6,6458	2,4142	0,0000	C3
2,2361	2,4142	6,0000	C1
3,2361	2,0000	4,1623	C2
3,2361	2,0000	4,1623	C2

1,0000	1,7321	3,6458	C1
6,0000	2,0000	1,0000	C3
3,2361	1,4142	3,4495	C2
2,6458	2,4142	6,0000	C2
6,4495	2,7321	1,7321	C3
2,2361	3,0000	7,1623	C1
2,2361	3,0000	7,1623	C1
0,0000	2,7321	6,6458	C1
2,7321	0,0000	2,4142	C2
2,2361	2,4142	6,4495	C1
6,6458	2,4142	0,0000	C3
3,4495	1,7321	3,6458	C2

Dalam metode *K-Means*, perhitungan berhenti apabila *Cluster* pada iterasi yang dihasilkan sama pada iterasi sebelumnya. Pada penelitian ini, proses *K-Means* berhenti di iterasi ke-4, berikut Tabel 5 adalah hasil perhitungan iterasi ke-4 :

Tabel 5 : Jarak *Centroid* Iterasi Ke-4

C1	C2	C3	Jarak <i>Centroid</i>
1,3015	2,5635	6,3184	C1
2,3553	1,3744	4,0619	C2
1,8330	0,8819	2,1726	C2
2,3015	1,0541	3,8062	C2
6,1665	3,2852	0,6124	C3
1,8844	2,7951	5,5411	C1
2,3553	1,3744	4,0619	C2
2,8844	1,6667	3,7157	C2
2,3015	1,5635	3,3184	C2
5,5972	2,6997	0,6124	C3
2,2454	1,1055	3,2079	C2
1,7261	2,7321	6,0917	C1
6,3255	3,3094	1,1726	C3
1,3553	2,3744	7,0619	C1
1,3553	2,3744	7,0619	C1
1,3015	2,5635	6,3184	C1
1,8330	0,8819	2,1726	C2
1,2454	2,1055	6,2079	C1
6,1665	3,2852	0,6124	C3
2,9219	1,7638	3,3184	C2

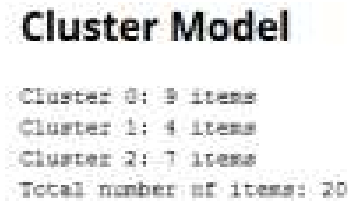
Proses perhitungan iterasi ke-3 dan iterasi ke-4 memiliki nilai *Cluster* yang sama atau tidak berubah, maka perhitungan dihentikan dan hasil yang diperoleh yaitu :

1. *Cluster 2 (C1)* memiliki 7 data yang diartikan bahwa kelompok pertama adalah kategori yang diterima sebagai penerima insentif pada tahun 2020.
2. *Cluster 0 (C2)* memiliki 9 data yang diartikan bahwa kelompok kedua adalah kategori yang dapat pertimbangan untuk mendapat atau tidaknya insentif pada tahun 2020.
3. *Cluster 1 (C3)* memiliki 4 data yang diartikan bahwa kelompok ketiga adalah kategori yang ditolak untuk mendapat insentif pada tahun 2020.

Untuk proses mengelompokkan data penulis menggunakan rumus pada metode *K-Means*. Kemudian data pegawai penerima insentif tersebut akan diproses menggunakan aplikasi *RapidMiner* 8.1 untuk mendapatkan hasil yang akurat. Data yang akan diolah adalah sample data 20 pegawai PT Sinarmas *Multifinance*.

3.2. Pemodelan RapidMiner

Hasil percobaan pada penelitian ini menggunakan *software Rapidminer 5.3* diperoleh model aturan yang dapat dilihat pada Gambar 1.



Gambar 1 Hasil Data *Clustering*

Sehingga berdasarkan data hasil *clustering* pada Gambar 1 didapatkan *cluster* hasil dari rapidminer 8.1 berikut ini :

1. *Cluster* 0 (C2) memiliki 9 data yang diartikan bahwa kelompok pertama adalah kategori yang dapat pertimbangan untuk mendapat atau tidaknya insentif pada tahun 2020.
2. *Cluster* 1 (C3) memiliki 4 data yang diartikan bahwa kelompok kedua adalah kategori yang ditolak untuk mendapat insentif pada tahun 2020.
3. *Cluster* 2 (C1) memiliki 7 data yang diartikan bahwa kelompok ketiga adalah kategori yang diterima sebagai penerima insentif pada tahun 2020.

4. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, maka dapat disimpulkan bahwa penerapan *data mining* dengan menggunakan algoritma *K-Means* dapat dilakukan dalam *clustering* data pegawai dalam menerima insentif untuk mengetahui layak atau tidaknya pegawai dalam menerima insentif.

Hasil yang tertinggi adalah *cluster* 2 pada *rapidminer* atau (C1) pada perhitungan konvensional menggunakan *excel* dengan 7 data pegawai dengan hasil yang memuaskan dalam kinerja. Hasil hitungan *excel* menggunakan algoritma dengan aplikasi menggunakan *RapidMiner* bernilai sama

Daftar Referensi

- Ardian, N. (2019). Pengaruh insentif berbasis kinerja , motivasi kerja , dan kemampuan kerja terhadap prestasi kerja pegawai UNPAB. *Jurnal*, 4(2), 119–132.
- Ayu, D. K., & Sinaulan, J. H. (2019). 濟無No Title No Title No Title. 1, 105–112.
- Handoko, K. (2016). Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K-Means Clustering (Studi Kasus Di Program Studi Tkj Akademi Komunitas Solok Selatan). *Jurnal Teknologi Dan Sistem Informasi*, 02(03), 31–40. <http://teknosi.fti.unand.id/index.php/teknosi/article/view/70>
- Iskandar, D. S., & Pandoyo. (2020). Pengaruh Insentif Terhadap Prestasi Kerja Karyawan Pada PT. Applikanusa Lintasarta Jakarta Selatan. *Jurnal Ilmu Administrasi Publik Dan Bisnis*, 1(September), 32–38.
- Maulida, L. (2018). Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 2(3), 167. <https://doi.org/10.14421/jiska.2018.23-06>
- Prasetyowati, E. (2017). DATA MINING Pengelompokan Data untuk Informasi dan Evaluasi. In *Duta Media Publishing* (pp. 97–98).
- Putra, R. R., & Wadisman, C. (2018). IMPLEMENTASI DATA MINING PEMILIHAN PELANGGAN POTENSIAL MENGGUNAKAN ALGORITMA K-MEANS. *Journal of Information Technology and*

Computer Science, 2(1), 1–13.

Saragih, D. A. R., Safii, M., & Suhendro, D. (2021). Penerapan Data Mining Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Sistem Informasi di Program Studi Sistem Informasi. *Journal of Information System Research (JOSH)*, 2(2), 173–177.

Sibuea, F. L., & Sapta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *Jurteksti*, 4(1), 85–92. <https://doi.org/10.33330/jurteksti.v4i1.28>

Werdiningsih, I., Nuqoba, B., & Muhammadun. (2020). *Data Mining Menggunakan Android, Weka, Dan Spss* (p. 209). Airlangga University Press.