

Implementasi Algoritma Naive Bayes Dalam Penentuan Bonus Karyawan

Wulan Dari

Program Studi Sistem Informasi Fakultas Teknik dan Ilmu Komputer
Universitas Potensi Utama

Alamat: Jl. K.L Yos Sudarso, Km. 6,5 , No. 3-a, Tj. Mulia, Kec. Medan Deli, Kota Medan,
Sumatera Utara 20241

Korespondensi penulis: ulandari2796@gmail.com

Abstract. CV. Aneka Kaca is a company engaged in the supply of the best glass materials, because CV. Aneka Kaca is a direct distributor from Japan. In a company like CV. Aneka Kaca, employees are one of the important assets in helping to improve and stabilize the company. Employees are people who work in a company or agency to carry out an operational task and expect remuneration in the form of a commission or salary. Usually employees will also get bonuses from the company. Bonuses that will be received by employees are one of the ways companies use to increase the motivation of their employees' performance. But giving bonuses to CV. Aneka Kaca still uses the conventional method, so it can take a long time if only managers do the math. The determination of bonuses can sometimes be missed due to one or two problems that can have an impact on awarding bonuses. The Naive Bayes algorithm is a simple probabilistic classifier that computes a probabilistic set by summing the frequencies and value combinations from a given dataset.

Keywords: Employees, Naive Bayes Algorithm

Abstrak. CV. Aneka Kaca merupakan perusahaan yang bergerak dalam bidang penyediaan material kaca terbaik, karena CV. Aneka Kaca merupakan distributor langsung dari negara Jepang. Dalam sebuah perusahaan seperti pada CV. Aneka Kaca, karyawan merupakan salah satu aset penting dalam membantu meningkatkan dan menstabilkan perusahaan tersebut. Karyawan adalah orang yang bekerja pada suatu perusahaan atau instansi untuk mengerjakan sebuah tugas operasional dan mengharapkan balas jasa berupa komisi atau gaji. Biasanya karyawan juga akan mendapatkan bonus dari perusahaan. Bonus yang akan diterima karyawan menjadi salah satu cara yang digunakan perusahaan untuk meningkatkan motivasi kinerja karyawannya. Namun pemberian bonus pada CV. Aneka Kaca masih menggunakan cara yang konvensional sehingga dapat memakan waktu yang lama apabila hanya manajer yang menghitung. Penentuan bonus terkadang juga dapat meleset karena satu dua kendala yang dapat berdampak dalam pemberian bonus. Algoritma Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilistik dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan.

Kata kunci: Karyawan, Algoritma Naive Bayes

LATAR BELAKANG

Dalam sebuah perusahaan sangat dibutuhkan yang namanya karyawan. Karyawan merupakan orang yang bekerja baik dalam jasa pikiran maupun tenaga dengan tujuan dibalas jasa berupa uang. Hal ini tidak jauh-jauh dengan pembahasan mengenai bonus karyawan. Bonus karyawan adalah upah tambahan yang didapatkan karyawan diluar gaji tetap. Bonus ini biasanya diberikan oleh perusahaan guna meningkatkan motivasi kinerja karyawan dalam bekerja. Hal ini juga memberikan nilai positif dari karyawan kepada perusahaan.

CV. Aneka Kaca ingin memberikan bonus kepada setiap karyawan yang layak mendapatkannya. Beberapa kriteria juga menjadi acuan perusahaan dalam memberikan bonus tersebut, diantaranya yaitu jumlah absen, kinerja karyawan dalam bekerja dan jumlah lembur yang diambil oleh karyawan. Data setiap karyawan akan direkap oleh manajer untuk dapat

melihat karyawan yang layak menerima bonus. Namun dalam merekap data, manajer masih menggunakan cara manual, sehingga dalam pengerjaan membutuhkan waktu yang cukup lama dan belum tentu akurat. Hal ini dapat terjadi karena terdapat satu dua kendala yang menghambat perekapan data. Oleh karena itu penggunaan naive bayes diperlukan untuk memudahkan manajer dalam merekap data tersebut.

Penelitian ini bertujuan untuk memberikan pemahaman mengenai penggunaan naive bayes kepada perusahaan dan juga kepada setiap pembaca yang telah membaca penelitian ini nantinya.

KAJIAN TEORITIS

1. Data Mining

Data mining adalah proses yang memanfaatkan teknik-teknik statistik, matematika, dan kecerdasan buatan untuk mengekstrak dan mengidentifikasi informasi dan knowledge selanjutnya (atau pola-pola) yang berasal dari sekumpulan data yang sangat besar. Berbagai macam pola tersebut bisa dalam bentuk aturan bisnis, kesamaan-kesamaan, korelasi, trend, atau model-model prediksi.

Data mining mempunyai 4 dasar fungsi diantaranya ialah :

- a. Prediksi : setelah menemukan suatu pola dari kumpulan data, pola tersebut nantinya digunakan dalam memprediksi hasil yang terjadi di periode berikutnya.
- b. Dekripsi : untuk memahami karakteristik dari suatu data.
- c. Asosiasi : untuk menemukan hubungan antar data, sehingga bisa tahu apakah karakteristik suatu data mempengaruhi data lain.
- d. Klasifikasi : untuk memberikan atribut tertentu pada suatu data. Agar datanya lebih mudah diinterpretasikan.

2. Algoritma Naive Bayes

Naive Bayes merupakan metode yang cocok untuk klasifikasi biner dan multiclass. Metode yang juga dikenal sebagai Naive Bayes Classifier ini menerapkan teknik supervised klasifikasi objek di masa depan dengan menetapkan label kelas ke instance/catatan menggunakan probabilitas bersyarat. Probabilitas bersyarat adalah ukuran peluang suatu peristiwa yang terjadi berdasarkan peristiwa lain yang telah (dengan asumsi, praduga, pernyataan, atau terbukti) terjadi.

Dasar dari NBC merupakan teorema bayes dimana mempunyai kemampuan untuk melakukan klasifikasi, NBC hamper sama dengan pohon keputusan serta jaringan syaraf. NBC memiliki tingkatan akurasi yang tinggi jika digunakan pada basis data yang memiliki data yang besar. Teorema Bayes memiliki persamaan umum :

$$P[H|X] = \frac{P(X|H)P(H)}{P(X)}$$

X = Data dengan class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)

P(H) = Probabilitas hipotesis H (prior prob.) P(X|H) = Probabilitas X berdasarkan kondisi tersebut

P(X) = Probabilitas dari X

3. Kelebihan dan Kekuran Metode Naive Bayes

Adapun kelebihan dari metode naive bayes yaitu :

- Klasifikasi Naive Bayes mudah diimplementasikan dan cepat.
- Ini akan menyatu lebih cepat daripada model diskriminatif seperti regresi logistik.
- Membutuhkan lebih sedikit data pelatihan.
- Sifatnya sangat skalabel, atau skalanya linier dengan jumlah prediktor dan titik data.
- Dapat membuat prediksi probabilistik dan dapat menangani kontinu beserta diskrit data.
- Algoritma klasifikasi Naive Bayes dapat digunakan untuk biner maupun multi-kelas masalah klasifikasi keduanya.

Dan adapun kekurangan dari naive bayes adalah :

- Salah satu Kelemahan terpenting dari klasifikasi Naive Bayes adalah fiturnya yang kuat kemandirian karena dalam kehidupan nyata hampir tidak mungkin untuk memiliki seperangkat fitur yang benar-benar independen satu sama lain.
- Masalah lain dengan klasifikasi Naive Bayes adalah 'frekuensi nol' yang berarti bahwa jika variabel kategori memiliki kategori tetapi tidak diamati dalam data pelatihan set, maka model Naive Bayes akan menetapkan probabilitas nol untuk itu dan itu tidak akan dapat untuk membuat prediksi.

METODE PENELITIAN

Metode yang digunakan dalam penelitian ini merupakan metode Literatur studi pustaka. Dengan membaca berbagai sumber materi yang akan dijadikan sebagai referensi. Setelah melakukan literatur studi pustaka maka akan didapat berbagai data dan informasi yang berkaitan dengan Data Mining metode Naive Bayes yang dimana selanjutnya data dan informasi tersebut dapat diolah menjadi bahan untuk mengenal lebih dalam lagi serta menemukan bagaimana cara menghitung data-data tersebut.

HASIL DAN PEMBAHASAN

1. Analisa Masalah

CV. Aneka Kaca merupakan sebuah perusahaan yang bergerak dalam bidang penyediaan material aneka kaca sesuai dengan kebutuhan pelanggan. Permasalahan yang tengah dihadapi oleh perusahaan ini yaitu perhitungan bonus karyawan yang masih belum akurat, sehingga terlihat satu atau lebih kendala masalah dalam pemberian bonus karyawan.

2. Pembahasan

Hal pertama yang dilakukan yaitu mengumpulkan data-data karyawan sebelumnya untuk dijadikan sebagai perhitungan data selanjutnya. Berdasarkan dari data yang sudah dikumpulkan dan juga tahapan preprocessing atau pengelompokan data, maka dapat dibuat data baru yang sudah melalui tahapan Preprocessing tersebut, dimana data tersebut yang akan digunakan pada tahap pemrosesan dan ditarik kesimpulannya.

No	Jumlah Absen/Bulan	Kinerja Pegawai	Jumlah Lembur/Bulan	Status Penerimaan
1	26	Baik	0	Layak
2	24	Kurang	0	Tidak
3	26	Baik	14	Layak
4	25	Baik	14	Layak
5	25	Kurang	0	Tidak
6	24	Baik	14	Layak
7	26	Kurang	7	Layak
8	24	Kurang	7	Layak
9	26	kurang	0	Tidak
10	25	Baik	0	Layak
11	26	baik	7	?

Selanjutnya mencari efek yang ditimbulkan dari data testing dengan menggunakan Naïve Bayes Classifier. Dengan mencari peluang kelas.

a. Menghitung Probabilitas Class

Langkah pertama perhitungan untuk menentukan karyawan yang layak mendapatkan bonus dengan metode naive bayes adalah dengan mencari probabilitas dari masing-masing kelas. Dalam penentuan karyawan tetap akan ditentukan 2 kelas yaitu kelas "Layak" dan "Tidak". Cara perhitungannya dengan mencari berapa jumlah data yang ada pada kelas "Layak" dan berapa jumlah pada kelas "Tidak" dari total keseluruhan data training, lalu membaginya dengan total keseluruhan data. Dengan menggunakan excel, rumusnya adalah :

"=COUNTIF(Layak)/COUNTA(Seluruh Data)",

"=COUNTIF(Tidak)/COUNTA(Seluruh Data)".

Maka hasil yang akan didapat terlihat pada tabel dibawah ini.

Label	Jumlah Data	Jumlah Seluruh Data	Hasil
Layak	7	10	0,7
Tidak	3	10	0,3

b. Menghitung Probabilitas Masing-Masing Atribut

Langkah selanjutnya yaitu menghitung probabilitas masing-masing atribut. Cara mencarinya yaitu dengan membandingkan atribut dari data testing dengan atribut dari data training. Berapa jumlah atribut dengan kelas "Layak" yang berada pada data training, kemudian bagi dengan probabilitas kelas "Layak". Begitu juga dengan mencari probabilitas untuk kelas "Tidak". Jika menggunakan excel, rumus yang digunakan ialah :

"=COUNTIFS(jumlah absen "Layak")/COUNTIF(class layak)",

"=COUNTIFS(jumlah absen "Tidak")/COUNTIF(class tidak)".

Jumlah Absen	Layak	Tidak	P(Positif)	P(Negatif)
26	3	1	0,43	0,33
25	2	1	0,29	0,33
24	2	1	0,29	0,33
Total	7	3		

Maka didapatkan hasil seperti diatas :

$$\bullet P(\text{atribut 1}) = \frac{26|\text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{3}{7} = 0,43$$

$$\bullet P(\text{atribut 1}) = \frac{26|\text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{1}{3} = 0,33$$

$$\bullet P(\text{atribut 1}) = \frac{25|\text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{2}{7} = 0,29$$

$$\bullet P(\text{atribut 1}) = \frac{25|\text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{1}{3} = 0,33$$

$$\bullet P(\text{atribut 1}) = \frac{24|\text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{2}{7} = 0,29$$

$$\bullet P(\text{atribut 1}) = \frac{24|\text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{1}{3} = 0,33$$

Begitu juga ketika mencari atribut “Kinerja Karyawan”. Jika menggunakan excel rumus yang dipakai ialah :

“=COUNTIFS(kinerja karyawan “Layak”)/COUNTIF(class layak)”, “=COUNTIFS(kinerja karyawan “Tidak”)/COUNTIF(class tidak)”.

Kinerja Pegawai	Layak	Tidak	P(Positif)	P(Negatif)
Baik	5	0	0,71	0
Kurang	2	3	0,29	1
Total	7	3		

Tabel 5. Mencari Peluang Kinerja Karyawan

$$\bullet P(\text{atribut 2}) = \frac{\text{Baik}|\text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{5}{7} = 0,71$$

$$\bullet P(\text{atribut 2}) = \frac{\text{Baik}|\text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{0}{3} = 0$$

$$\bullet P(\text{atribut 2}) = \frac{\text{Kurang}|\text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{2}{7} = 0,29$$

$$\bullet P(\text{atribut 2}) = \frac{\text{Kurang}|\text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{3}{3} = 1$$

Hal yang sama juga dilakukan dalam mencari atribut jumlah lembur.

“=COUNTIFS(jumlah lembur “Layak”)/COUNTIF(class layak)”,

“=COUNTIFS(jumlah lembur “Tidak”)/COUNTIF(class tidak)”.

Jumlah Absen	Layak	Tidak	P(Positif)	P(Negatif)
14	3	0	0,43	0
7	2	0	0,29	0
0	2	3	0,29	1
Total	7	3		

Tabel 6. Mencari Peluang Jumlah Lembur

$$\begin{aligned} \bullet P(\text{atribut 3}) &= \frac{14 | \text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{3}{7} = 0,43 \\ \bullet P(\text{atribut 3}) &= \frac{14 | \text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{0}{3} = 0 \\ \bullet P(\text{atribut 3}) &= \frac{7 | \text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{2}{7} = 0,29 \\ \bullet P(\text{atribut 3}) &= \frac{7 | \text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{0}{3} = 0 \\ \bullet P(\text{atribut 3}) &= \frac{0 | \text{"Layak"}}{\text{Jumlah Status "Layak"}} = \frac{2}{7} = 0,29 \\ \bullet P(\text{atribut 3}) &= \frac{0 | \text{"Tidak"}}{\text{Jumlah Status "Tidak"}} = \frac{3}{3} = 1 \end{aligned}$$

c. Menghitung Hasil Akhir Data Testing

No	Jumlah Absen/Bulan	Kinerja Pegawai	Jumlah Lembur/Bulan	Status Penerimaan
11	26	Baik	7	?

Tabel 2. Data Testing

Untuk mencari hasil yang ada tanda tanya pada data karyawan diatas maka rumus yang digunakan ialah dengan cara mengalikan setiap probabilitas yang sesuai dengan kriterianya. Seperti pada tabel dibawah ini. Sehingga hasil yang didapatkan nantinya adalah dengan nilai yang paling tinggi. Seperti terlihat pada tabel dibawah :

Layak	0,43	0,71	0,29	0,09
Tidak	0,33	0	0	0

Tabel 7. Perhitungan Hasil Data Testing

$$\text{Layak} = 26 \times \text{Baik} \times 7$$

$$= 0,43 \times 0,71 \times 0,29$$

$$= 0,09$$

$$\text{Tidak} = 26 \times \text{Baik} \times 7$$

$$= 0,33 \times 0 \times 0$$

$$= 0$$

Maka hasil akhir yang diambil adalah dengan nilai yang paling tinggi yaitu "Layak | 0,09" hasilnya akan terlihat pada tabel dibawa.

NO	JUMLAH ABSEN/BULAN	KINERJA PEGAWAI	JUMLAH LEMBUR/BULAN	STATUS PENERIMAAN
1	26	Baik	0	Layak
2	24	Kurang	0	Tidak
3	26	Baik	14	Layak
4	25	Baik	14	Layak
5	25	Kurang	0	Tidak
6	24	Baik	14	Layak
7	26	Kurang	7	Layak
8	24	Kurang	7	Layak
9	26	kurang	0	Tidak
10	25	Baik	0	Layak
11	26	Baik	7	Layak

KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan dapat ditarik kesimpulan bahwasannya dengan menerapkan algoritma Naive Bayes pada perusahaan datanya lebih akurat ketika mencari perhitungan karyawan yang layak mendapatkan bonus dibandingkan dengan menghitung manual. Dengan adanya data testing perusahaan dapat mengetahui hasil apa yang akan keluar sehingga perusahaan dapat memprediksi kemungkinan dikemudian hari.

UCAPAN TERIMA KASIH

Ucapan terima kasih dan puji syukur atas rahmat Allah Subhanahu wa ta'ala karena telah memberikan kesehatan, keluangan waktu dan terutama nikmat iman dan islam di dalam diri peneliti. Tidak lupa peneliti ucapkan terima kasih kepada kedua orang tua yang telah merawat dan mendidik peneliti dari kecil hingga saat ini, dan juga terhadap para mahasiswa yang membantu dalam penelitian ini. Terima kasih juga peneliti ucapkan yang membantu dalam memberikan informasi dalam penelitian ini.

DAFTAR REFERENSI

- Abdullah, M. R. (2014). Manajemen dan evaluasi kinerja karyawan.
- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah*, 10(2), 160-165.
- Information Technology Journal of UMUS, 3(02), 149-159.\
- Pauziah, U. (2017). Analisis Penentuan Karyawan Terbaik Menggunakan Metode Algoritma Naive Bayes (Studi Kasus PT. XYZ).
- Peling, I. B. A., Arnawan, I. N., Arthawan, I. P. A., & Janardana, I. G. N. (2017).

Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm. *Int. J. Eng. Emerg. Technol.*, 2(1), 53.

- Prameswari, Y. A. (2020). Rancang Bangun Sistem Informasi Penilaian Prestasi Kinerja Karyawan dengan Menggunakan Metode Naive Bayes (Doctoral dissertation, Universitas 17 Agustus 1945 Surabaya).
- Sari, V. N., Astri, L. Y., & Rasywir, E. (2020). Analisis dan Penerapan Algoritma Naive Bayes untuk Evaluasi Kinerja Karyawan pada PT. Pelita Wira
Sejahtera. Jurnal Ilmiah Mahasiswa Teknik Informatika, 2(1), 53-68.
- Silaen, N. R., Syamsuriansyah, S., Chairunnisah, R., Sari, M. R., Mahriani, E., Tanjung, R., ... & Putra, S. (2021). Kinerja Karyawan.
- Siregar, V. M. M. (2018). Sistem Pendukung Keputusan Penentuan Insentif Bulanan Pegawai Dengan Menggunakan Metode Naive Bayes. *Sistemasi: Jurnal Sistem Informasi*, 7(2), 78-86.
- Utomo, D. P., & Purba, B. (2019, September). Penerapan datamining pada data gempa bumi terhadap potensi tsunami di Indonesia. In *Prosiding Seminar Nasional Riset Information Science (SENARIS)* (Vol. 1, pp. 846-853).
- Watratan, A. F., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7-14.
- Widyassari, A. P., & Suryani, P. E. (2021). Komparasi Metode Naive Bayes dan SAW untuk Pemilihan Penerimaan Insentif Karyawan. *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, 3(02), 149-159.
- Widyassari, A. P., & Suryani, P. E. (2021). Komparasi Metode Naive Bayes dan SAW untuk Pemilihan Penerimaan Insentif Karyawan. *Jurnal Ilmiah Intech:*