

PERBANDINGAN ALGORITMA C4.5 DAN NAIVE BAYES UNTUK MENETUKAN KARYAWAN BERPRESTASI

Kevin Merico Setiawan, Febriana Santi Wahyuni, Ahmad Faisol

Program Studi Teknik Informatika S1, Fakultas Teknologi Industri
Institut Teknologi Nasional Malang, Jalan Raya Karanglo km 2 Malang, Indonesia
mericosetiawan@gmail.com

ABSTRAK

Karyawan merupakan salah satu aset terpenting sebuah perusahaan dalam usahanya menjalankan roda bisnisnya yang semakin kompetitif. Salah satu upaya dalam meningkatkan kualitas sumber daya manusia banyak jalan untuk perusahaan melakukannya, termasuk memberikan bonus bulanan seperti CV. Mitra Jaya Bersama Malang. Namun pemberian bonus tersebut masih menggunakan cara secara konvensional sehingga sangat memakan waktu terutama apabila karena hanya manager yang akan menghitung tersebut. Namun juga penentuan bonus tersebut juga terkadang dapat meleset karena satu dua kendala seperti kedekatan dengan manager atau yang lain. Algoritma yang digunakan untuk membentuk pohon keputusan. Sebuah pohon keputusan merupakan struktur yang digunakan sebagai himpunan-himpunan data yang lebih kecil. Algoritma Naive Bayes Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Dari kedua algoritma tersebut akan dicari nilai yang terbaik dengan menggunakan *Confusion Matrix* untuk mencari informasi perbandingan hasil klasifikasi yang dilakukan oleh algoritma atau model. Berdasarkan atribut karyawan CV. Mitra Jaya Bersama Malang yang telah diuji menggunakan *Confusion Matrix*, kinerja Algoritma C4.5 dan Algoritma Naive Bayes tergantung dalam data nya untuk kinerjanya. Pada data sebanyak 162 kinerja Algoritma C4.5 dan Algoritma Naive Bayes sama baiknya, sedangkan pada data sebanyak 187 Algoritma C4.5 memiliki tingkat akurasi sebesar 98,91%, presisi dan *recall* sebesar 99,17%, sedangkan untuk Algoritma Naive Bayes memiliki tingkat akurasi sebesar 96,21%, presisi 95,23% dan *recall* sebesar 99,17%.

Kata Kunci : Karyawan, Algoritma C4.5, Naive Bayes, *Confusion Matrix*.

1. PENDAHULUAN

1.1 Latar Belakang

Karyawan merupakan salah satu aset terpenting sebuah perusahaan dalam usahanya menjalankan roda bisnisnya yang semakin kompetitif. Dengan persaingan yang semakin kompetitif ini, memacu sebuah perusahaan berupaya lebih keras dalam meningkatkan kualitas perusahaannya. Salah satu upaya yang dapat dilakukan yaitu dengan meningkatkan kualitas sumber daya manusia seperti pada CV. Mitra Jaya Bersama Malang yang memberikan bonus tambahan setiap bulan apabila memenuhi target serta absen yang telah diatur sedemikian rupa.

Penentuan karyawan terbaik merupakan cara yang dapat meningkatkan kompetensi karyawan tanpa mengeluarkan dana yang besar bagi perusahaan dibandingkan dengan solusi lain. Dengan penentuan karyawan terbaik serta dapat digunakan untuk mendapatkan informasi yang *valid* dan dapat menjadi acuan perusahaan untuk melakukan keputusan penting seperti promosi jabatan, pembagian *reward* dan keputusan keputusan penting lainnya berdasarkan kriteria-kriteria seperti absensi, pengambilan cuti, pencapaian target atau bahkan yang lain bersarkan perusahaan yang terkait. Namun menentukan karyawan terbaik pada beberapa perusahaan masih dalam bentuk manual dan subjektivitas sehingga dinilai masih kurang efisien serta kurang akurat.

Berbagai penelitian tentang karyawan terbaik telah banyak dilakukan seperti Dimas Imam Baihaqi yang berjudul Perbandingan Metode Naive Bayes dan C4.5 untuk Memprediksi Mortalitas Pada Peternakan Ayam Broiler, Algoritma C4.5 lebih baik di bandingkan dengan Naive Bayes. Pada Penelitian lain Choirul Anam yang berjudul Perbandingan Kinerja Algoritma C4.5 dan Naive Bayes untuk Klasifikasi Penerimaan Beasiswa, kinerja Algoritma Naive Bayes lebih baik. Dengan latar belakang itu penulis melakukan penelitian untuk melihat kinerja Algoritma C4.5 dan Naive Bayes dalam studi kasus karyawan terbaik, menggunakan *Confusion Matrix* sebagai alat ukur untuk menghitung kinerja algoritma C4.5 dan Naive Bayes.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Menurut Yustina Retno Wahyu A. (dalam penelitiannya Implementasi Algoritma C4.5 Dalam Pembuatan Aplikasi Penunjang Keputusan Penerimaan Pegawai CV. Dinamika Ilmu, 2014:32). Hasil dari penelitian ini Algoritma C4.5 dapat di implementasikan dengan output berupa klasifikasi calon pegawai baru dengan variable tujuan diterima atau tidak terima dibentuk dengan algoritma C4.5[1].

Menurut Victor Marudut Mulia Siregar (pada paper Sistem Pendukung Keputusan Penentuan Insentif Bulanan Pegawai dengan Menggunakan

Algoritma Naive Bayes, 2018:94). Untuk mempermudah manager Edene Sayangku Cafe & Bakery dalam menghitung insentif bulanan karyawan dengan tepat dan akurat. Hasil dari penelitian tersebut metode Naive Bayes dapat berjalan secara benar dan memberikan hasil yang akurat sehingga metode Naive Bayes dapat dijadikan alat bantu bagi pengambilan keputusan dalam memutuskan jumlah insentif yang layak untuk masing masing pegawai [2].

Menurut Dimas Imam Haihaqi (dalam papernya Perbandingan Metode Naive Bayes dan C4.5 untuk Memprediksi Moralitas pada Peternakan Ayam Broiler, 2019:389). Tujuan pada penelitian ini merupakan membandingkan kinerja antara Naive Bayes dan C4.5 pada Moralitas Ayam Broiler. Hasil yang didapatkan Algoritma C4.5 merupakan algoritma yang terbaik dengan hasil akurasi sebesar 93%, presisi sebesar 0,68% dan *recall* sebesar 0,76% serta *Tree* keputusan yang dihasilkan dari metode C4.5 menampilkan dua variabel yaitu *Body Weight* dan umur panen. Variabel tersebut adalah variabel yang paling berpengaruh untuk memprediksi mortalitas pada ayam broiler [3].

Menurut Ulfa Paujiah (dalam papernya yang berjudul Analisis Penentuan Karyawan Terbaik Menggunakan Metode Naive Bayes Studi Kasus PT. XYZ, 2017:101). Tujuan dari penelitian tersebut adalah mengukur tingkat akurasi dari kajian algoritma Naive Bayes dalam penentuan karyawan terbaik di PT. XYZ. Hasil dari penelitian tersebut disimpulkan bahwa algoritma Naive Bayes terbukti dapat dipakai dalam penentuan karyawan terbaik dengan hasil bahwa 98,6% algoritma Naive Bayes dapat membantu pengambilan keputusan pada penentuan karyawan terbaik [4].

Menurut Saefudin dan Sri Lestari (dalam penelitiannya berjudul Sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5, 2015:42). Tujuan dari penelitian tersebut membangun sebuah sistem pendukung keputusan menggunakan algoritma C4.5 untuk membantu pihak perusahaan dalam menentukan kinerja karyawan sesuai kriteria. Hasil dari penelitian tersebut adalah terciptanya sebuah sistem berbasis web yang digunakan PT. Mulya Spindo Mills dengan pengujian 10 karyawan dan menghasilkan 6 karyawan dengan keputusan istimewa [5].

2.2 Data Mining

Data Mining adalah suatu proses untuk menemukan pengetahuan baru yang berasal dari data dalam jumlah yang besar yang telah diolah. Pengetahuan baru akan digunakan sebagai acuan untuk sebuah keputusan di masa depan yang valid dari sumber data lama.

2.3 Karyawan

Karyawan merupakan orang yang bekerja pada sebuah lembaga, kantor, perusahaan dan sebagainya

dengan mendapatkan gaji atau upah. Karyawan juga disebut seorang penjual jasa dengan kompensasi atau gaji yang telah disepakati di awal.

2.4 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sebuah pohon keputusan merupakan struktur yang digunakan sebagai himpunan - himpunan data yang lebih kecil. Pada awal perhitungan algoritma C4.5 harus menentukan *root* awal dengan mencari nilai *entropy* dan *gain* pada setiap atribut.

Berikut cara mencari nilai *entropy* dan *gain* :

$$\text{Entropy (S)} = \sum_{i=1}^n - P_i * \log_2 P_i$$

Dimana :

S : Himpunan Kasus

n : Jumlah Partisi Kasus

Pi: Proporsi dari Si

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

Dimana :

S : Himpunan Kasus

A : Atribut

n : Jumlah Partisi Atribut A

|S_i| : Jumlah Kasus pada Partisi Ke I

|S| : Jumlah Kasus dalam S

2.5 Algoritma Naive Bayes

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas.

Formula untuk menghitung probabilitas pada algoritma Naive Bayes sebagai berikut :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^d P(X_i|Y)}{P(X)}$$

Dimana

P(Y|X) : Probabilitas data dengan vektor X pada kelas Y.

P(Y) : Probabilitas awal kelas Y

$\prod_{i=1}^d P(X_i|Y)$: Probabilitas independen kelas Y dari semua fitur dalam vektor X

2.6 Confusion Matrix

Confusion matrix juga sering disebut dengan *error matrix*. Tujuan pertama dengan adanya *confusion matrix* merupakan memberikan hasil perbandingan hasil klasifikasi yang dilakukan oleh sistem dengan

hasil klasifikasi yang sebenarnya. *Confusion matrix* berbentuk tabel matrix yang menggambarkan kinerja mode klasifikasi pada serangkaian data uji yang nilai sebenarnya sudah diketahui. *Confusion matrix* mempunyai 4 kombinasi nilai prediksi dan nilai aktual yang berbeda yang dapat dilihat pada gambar berikut :

	Positive	Negative
Positive	TRUE POSITIVE	FALSE NEGATIVE
Negative	FALSE NEGATIVE	TRUE NEGATIVE

Gambar 1 Tabel *confusion matrix*s

Berikut adalah penjelasan dari 4 istilah yang terdapat pada *Confusion Matrix*:

1. *True Positive* (TP)
merupakan data nyata dan hasil dari sistem bernilai sama namun dalam klasifikasi yang bernilai positif.
2. *True Negative* (TN)
merupakan data nyata dan hasil dari sistem bernilai sama namun dalam klasifikasi yang bernilai negatif.
3. *False Positive* (FP)
merupakan data nyata bernilai negatif namun pada data klasifikasi pada sistem bernilai positif.
4. *False Negative* (FN)
merupakan data nyata bernilai positif namun pada data klasifikasi pada sistem bernilai negatif.

Untuk mengukur performa matrix dari *Confusion matrix* berikut beberapa *performance matrix* yang digunakan :

1. Accuracy

Accuracy menggambarkan seberapa akurat model dapat mengklasifikasikan data dengan benar. Bisa dikatakan *accuracy* ini mengukur tingkat kedekatan antara nilai prediksi dan nilai asli. Berikut rumus mencari *accuracy*:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision

Precision menggambarkan tingkat keakuratan antara data yang diminta dengan prediksi yang diberikan oleh hasil klasifikasi dengan kata lain perbandingan antara rasio prediksi benar dan rasio prediksi salah. Berikut rumus mencari *precision* :

$$Precision = \frac{TP}{TP + FP}$$

3. Recall

Recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi dikatakan lain dengan rasio antara prediksi benar dan semua data yang benar. Berikut rumus mencari *recall* :

$$Recall = \frac{TP}{TP + FN}$$

3. METODE PENELITIAN

3.1. Flowchart Sistem

Flowchart sistem ini menjelaskan proses berjalannya aplikasi seperti ditunjukkan pada Gambar 2 :

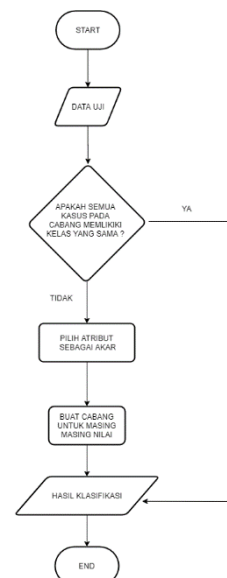


Gambar 2 Blok diagram sistem

Data uji merupakan variabel variabel yang terkait dengan karyawan, kemudian dilakukan klasifikasi oleh dua algoritma yang berbeda algoritma C4.5 dan Naive Bayes. Setelah mendapatkan hasil prediksi dari kedua metode tersebut maka dilakukan perbandingan dengan menghitung nilai akurasi, presisi dan *recall* menggunakan *Confusion Matrix*, sehingga dapat ditarik kesimpulan algoritma mana yang lebih optimal untuk melakukan penentuan karyawan terbaik.

3.2. Flowchart Algoritma C4.5

Flowchart Algoritma C4.5 ditunjukkan pada Gambar 3:

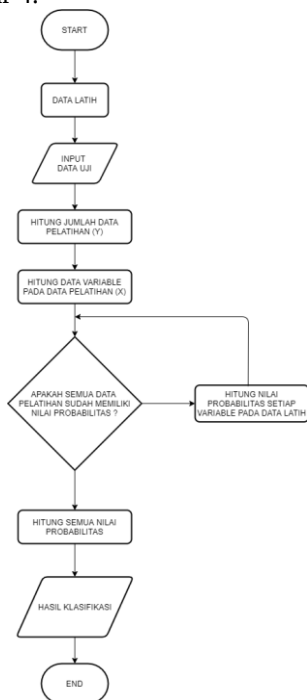


Gambar 3 Flowchart Algoritma C4.5

Berdasarkan flowchart pada Gambar 3, data uji akan mencari atribut sebagai akar pertama dengan mencari nilai *entropy*, *gain*, *split info* dan *gain ratio* pada setiap atribut. Pilih akar selanjutnya dengan cara memilih nilai terbesar dari *gain ratio* pada tiap - tiap atribut. Hingga mendapatkan atribut dengan salah satu atribut bernilai 0.

3.3. Flowchart Algoritma Naive Bayes

Flowchart Algoritma Naive Bayes ditunjukkan pada Gambar 4:



Gambar 4 Flowchart Algoritma Naive Bayes

Berdasarkan flowchart pada Gambar 4, akan di pisahkan terlebih dahulu antara data numerik dan tidak. Kemudian akan dicari nilai setiap probabilitas setiap atribut yang terdapat pada data training. Setelah ditemukan hasil probabilitas setiap atribut maka data set akan dihitung dengan mencari probabilitas ya dan tidak pada setiap data set. Setelah diketahui setiap probabilitas ya dan tidak akan dibandingkan mana yang terbesar maka itulah hasil dari presiksi algoritma Naive Bayes.

4. HASIL DAN PEMBAHASAN

4.1. Perhitungan Matematis Algoritma C4.5

Algoritma C4.5 merupakan salah satu klasifikasi yang dihasilkan dari pembuatan sebuah pohon keputusan. Pada penelitian ini menggunakan 162 data karyawan dari bulan Juni hingga November 2020 yang dapat dilihat pada tabel 4.1 berikut :

Tabel 4.1 Data Training

N o	Nama	Absen	Lembur	Target	Bonus
1.	Slamet Basori	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
2.	Heri Romdoni	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
3.	Heru Kurniawan	Sangat Baik	Tidak Lembur	Baik	Ya
4.	Reivant Dhea Prasadana	Baik	Tidak Lembur	Kurang Baik	Tidak
5.	Sugeng Waluyo	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
6.	Sunardi	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
7.	Endri Agus Subianto	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
8.	Sulkhan	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
9.	Abadi Kusuma	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
..					
..					
..					
16 2.	Moch Aga Yudistira	Sangat Baik	Cukup Baik	Sangat Baik	Ya

Dari tabel 4.1 terdapat 4 atribut yang akan digunakan sebagai acuan perhitungan pada proses klasifikasi. Setelah ditetapkan atribut mana saja yang akan digunakan, perhitungan maka dicari akar pohon keputusan yang di tampilkan pada tabel 4.2 berikut :

Tabel 4.2 Pencarian akar pohon keputusan

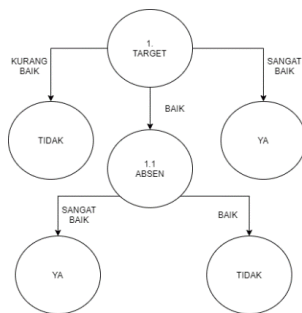
	Entropy	Gain	Split Info	Gain Ratio
Total	0.88			
Absen		0.63	0.86	0.73
Sangat Baik	0.75			
Baik	0.99			
Lembur		0.44	1.20	0.37
Tidak Lembur	0.94			
Cukup Baik	0.78			
Baik	0.30			
Target		1.08	1.40	0.77
Kurang Baik	0			
Baik	0.98			
Sangat Baik	0			

Setelah diketahui pada tabel 4.2 *gain ration* yang terbesar adalah atribut target. Maka dari itu atribut target menjadi cabang selanjutnya untuk mencari pohon keputusan dengan perhitungan yang sama dan syarat atribut yang telah di pilih tidak dapat dimasukkan pada proses selanjutnya. Perhitungan cabang pohon keputusan dapat dilihat pada table 4.3 sebagai berikut :

Tabel 4.3 Pencarian cabang pohon keputusan

	Entropy	Gain	Split Info	Gain Ratio
Total	0.98			
Absen		0.80	0.96	0.84
Sangat Baik	0.28			
	Entropy	Gain	Split Info	Gain Ratio
Baik	0			
Lembur		0.26	0.68	0.39
Tidak Lembur	0.87			
Cukup Baik	0			
Baik	0			

Pada tabel 4.3 atribut absen mempunyai nilai *gain ratio* yang paling besar, maka dari itu absen sebagai akar selanjutnya dari pohon keputusan. Dikarenakan pada atribut absen terdapat nilai 0 maka perhitungan selesai dan pohon keputusan telah terbentuk seperti pada gambar 5 berikut :



Gambar 5 Pohon Keputusan Algoritma C4.5

Setelah pohon keputusan telah terbentuk maka proses klasifikasi dapat dilakukan. Hasil klasifikasi disajikan pada tabel 4.4 berikut :

Tabel 4.4 Hasil klasifikasi algoritma C4.5

N o	Nama	Absen	Lembur	Target	Bonus	Prediksi
1.	Slam et Ba sori	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
2.	Heri Romd oni	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
3.	Heru Kurniawan	Sang at Baik	Tidak Lembur	Baik	Ya	Ya
4.	Reiva nt Dhea Prasa dana	Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
5.	Suge ng	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak

	Waluyo					
6.	Sunar di	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
7.	Endri Agus S.	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
N o	Nama	Absen	Lembur	Target	Bonus	Prediksi
8.	Sulkh an	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
9.	Abadi Kusuma	Sang at Baik	Tidak Lembur	Kura ng Baik	Tida k	Tidak
..						
..						
..						
16 2.	Moch Aga Yudis tira	Sang at Baik	Cuku p Baik	Sang at Baik	Ya	Ya

Dari hasil perhitungan algoritma C4.5 yang terdapat pada tabel 4.4, dari 162 data karyawan yang menjadi data uji hanya terdapat 1 prediksi yang tidak tepat pada nomor 87.

4.2. Perhitungan Matematis Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan sebuah algoritma yang menghitung jumlah probabilitas pada setiap atribut. Untuk data *training* yang digunakan adalah data karyawan CV. Mitra Jaya Bersama di Kota Malang sebanyak 162 data seperti pada tabel 4.5 berikut :

Tabel 4.5 Data Training

N o	Nama	Absen	Lembur	Target	Bonus
1.	Slamet Basori	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
2.	Heri Romdoni	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
3.	Heru Kurniawan	Sangat Baik	Tidak Lembur	Baik	Ya
4.	Reivant Dhea Prasadana	Baik	Tidak Lembur	Kurang Baik	Tidak
5.	Sugeng Waluyo	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
6.	Sunardi	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
7.	Endri Agus Subianto	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
8.	Sulkhana	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
9.	Abadi Kusuma	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak
..					
..					

..					
16 2.	Moch Agast Yudistira	Sangat Baik	Cukup Baik	Sangat Baik	Ya

Dari tabel 4.5 terdapat 4 atribut yang akan digunakan sebagai acuan perhitungan pada proses klasifikasi. Setelah ditetapkan atribut mana saja yang akan digunakan, perhitungan selanjutnya adalah mencari nilai probabilitas pada setiap atribut yang dapat dilihat pada tabel 4.6 berikut :

Tabel 4.6 Nilai probabilitas atribut

Probabilitas	Ya	Tidak
Bonus	0.70	0.30
Absen Sangat Baik	0.80	0.51
Absen Baik	0.20	0.49
Lembur Tidak Lembur	0.62	0.82
Lembur Cukup Baik	0.23	0.16
Probabilitas	Ya	Tidak
Lembur Baik	0.15	0.02
Target Kurang Baik	0.00	0.72
Target Baik	0.17	0.29
Target Sangat Baik	0.83	0

Setelah diketahui nilai probabilitas pada setiap atribut nyam aka proses selanjutnya adalah melakukan probabilitas pada setiap nilai setiap data dengan nilai probabilitas pada setiap atribut pada tabel 4.6. Untuk hasil klasifikasi dapat dilihat pada tabel 4.7 berikut :

Tabel 4.7 Hasil klasifikasi algoritma Naïve Bayes

No	Nama	Absen	Lembur	Target	Bonus	Prediksi
1.	Slamet Basori	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
2.	Heri Romdoni	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
3.	Heru Kurniawan	Sangat Baik	Tidak Lembur	Baik	Ya	Ya
4.	Reivandhe Prasada	Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
5.	Sugeng Waluyo	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
6.	Sunardi	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
7.	Endri Agus Subianto	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
8.	Sulkhan	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
9.	Abadi Kusuma	Sangat Baik	Tidak Lembur	Kurang Baik	Tidak	Tidak
..						

..						
..						
16 2.	Moch Agast Y.	Sangat Baik	Cukup Baik	Sangat Baik	Ya	Ya

Dari hasil perhitungan algoritma Naïve Bayes yang terdapat pada tabel 4.7, dari 162 data karyawan yang menjadi data uji hanya terdapat 1 prediksi yang tidak tepat pada nomor 87.

4.3. Perhitungan Confusion Matrix

Setelah diketahui hasil klasifikasi dari algoritma C4.5 dan Naïve Bayes untuk mengukur kinerja antara 2 buah algoritma tersebut digunakan *Confusion Matrix* dengan mencari nilai kombinasi terlebih dahulu sebelum mendapatkan nilai akurasi, presisi, dan *recall*. Untuk nilai kombinasi dari tiap tiap prediksi dapat dilihat pada tabel 4.8 berikut :

Tabel 4.8 Nilai Kombinasi

	True Positif (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)
Naïve Bayes	113	48	1	0
C4.5	113	48	1	0

Setelah mendapatkan nilai kombinasi seperti pada tabel 4.8, maka proses untuk mencari nilai akurasi, presisi, dan *recall* dapat dikerjakan. Untuk nilai akurasi, presisi dan *recall* dari algoritma C4.5 dan Naïve Bayes dapat dilihat pada tabel 4.9 berikut :

Tabel 4.9 Perbandingan kinerja algoritma C4.5 dan Naïve Bayes

	Akurasi	Presisi	Recall
Naïve Bayes	99.38 %	99.12%	100%
C4.5	99.38 %	99.12%	100%

Berdasarkan tabel 4.9 kinerja algoritma C4.5 dan Naïve Bayes sama baiknya dengan nilai presentasi untuk akurasi, presisi, dan *recall* berturut-turut 99,38%, 99,12%, dan 100%.

4.4. Pengujian Konsistensi

Setelah diketahui kinerja Algoritma C4.5 dan Naïve Bayes, untuk menguji tingkat konsistensi dari kedua algoritma tersebut penulis menggunakan data yang sedikit berbeda dengan menggunakan data karyawan pada bulan Juni hingga Desember 2020. Berdasarkan hasil pengujian terdapat sedikit perbedaan yang terjadi pada hasil perbandingan antara algoritma C4.5 dan Naïve Bayes seperti pada tabel 4.10 berikut :

Tabel 4.10 Pengujian konsistensi

No	Data Asli	Prediksi			
		Juni – November		Juni - Desember	
		C4.5	Naïve Bayes	C4.5	Naïve Bayes
1.	Tidak	Tidak	Tidak	Tidak	Tidak

2.	Tidak	Tidak	Tidak	Tidak	Tidak
3.	Ya	Ya	Ya	Ya	Ya
4.	Tidak	Tidak	Tidak	Tidak	Tidak
5.	Tidak	Tidak	Tidak	Tidak	Tidak
6.	Tidak	Tidak	Tidak	Tidak	Tidak
7.	Tidak	Tidak	Tidak	Tidak	Tidak
8.	Tidak	Tidak	Tidak	Tidak	Tidak
9.	Tidak	Tidak	Tidak	Tidak	Tidak
...					
...					
...					
189.	Tidak	-	-	Tidak	Tidak

Dari hasil pengujian pada tabel 4.10 dari data sebanyak 189 algoritma C4.5 memiliki 3 data dengan prediksi yang tidak tepat. Sedangkan algoritma Naïve Bayes memiliki 8 data dengan prediksi yang tidak tepat..

Setelah diketahui hasil klasifikasi pada kedua algoritma maka dilakukan pengujian kembali pada tabel *Confusion Matrix* seperti yang telah di paparkan pada tabel 4.11 :

Tabel 4.11 Perbandingan kinerja algoritma C4.5 dan Naïve Bayes

	Akurasi	Presisi	Recall
Naïve Bayes	95.76 %	95.27%	98.37%
C4.5	98.41 %	99.18%	99.37%

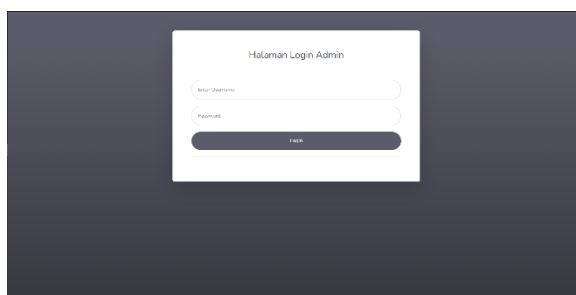
Berdasarkan tabel 4.11 terdapat perubahan kinerja pada algoritma C4.5 dan Naïve Bayes. Namun setelah menggunakan data Juni hingga Desember kinerja algoritma C4.5 lebih baik dibanding kan dengan algoritma Naïve bayes dengan nilai akurasi 98,41% dan nilai presisi sebesar 99,18%.

4.5. Implementasi Sistem

Implementasi sistem merupakan proses perancangan sistem yang telah dibuat untuk dijalankan menjadi suatu aplikasi..

4.6. Tampilan Halaman Login

Pada halaman login diberikan hak akses admin untuk login. Admin harus menginputkan username dan password untuk dapat masuk ke halaman utama admin. Tampilan halaman login dapat ditunjukkan pada Gambar 7 sebagai berikut :

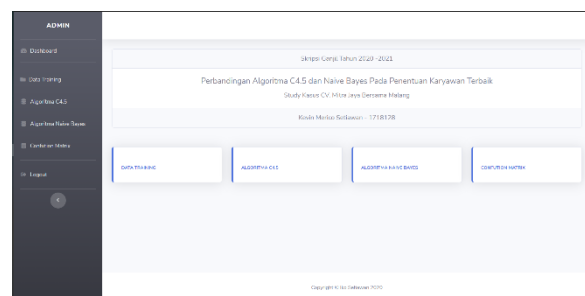


Gambar 7 Halaman Login

Gambar 7 merupakan tangkapan layar dari proses autentikasi pihak pengguna agar tidak semua pengguna dapat mengakses data data dari perusahaan. Dan jika pengguna tidak melalui autentikasi dan memaksa untuk mengakses halaman *dashboard* maka pengguna akan dilempar pada halaman autentikasi kembali

4.7. Tampilan Halaman Dashboard

Pada halaman *dashboard* terdapat pilihan menu pada sidebar dan tombol untuk mengakses ke halaman data training, perhitungan algoritma C4.5, perhitungan algoritma Naive Bayes, *Confusion Matrix* serta logout. Tampilan halaman *dashboard* dapat ditunjukkan pada gambar Gambar 8:

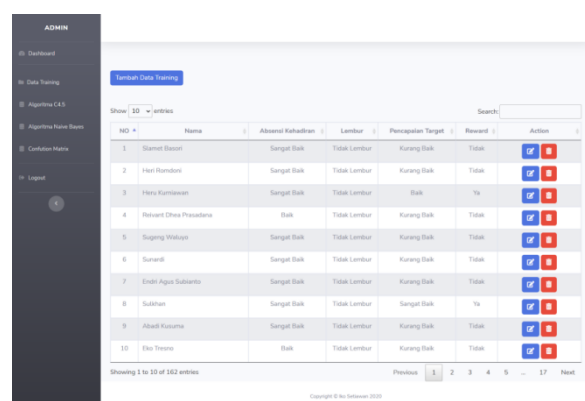


Gambar 8 Halaman Dashboard

Gambar 8 menjelaskan dari keberhasilan autentikasi pengguna, pengguna akan disuguhkan dengan 6 menu di bagian *sidebar* dan 4 menu utama di bagian *main board*.

4.8. Tampilan Halaman Data Training

Pada halaman data *training* terdapat data karyawan serta menu pencarian, tambah, update dan hapus data karyawan. Tampilan halaman *dashboard* dapat ditunjukkan pada gambar 4.9 :



Gambar 9 Halaman Data Training

Gambar 4.9 menjelaskan dari seluruh tampilan total data training yang akan digunakan sebagai bahan pengujian perbandingan 2 algoritma.

4.9. Tampilan Halaman Tambah Data Training

Pada halaman tambah data *training* admin dapat menambahkan data karyawan beserta atribut atribut lainnya. Tampilan halaman tambah data *training* dapat ditunjukkan pada gambar 4.10:

Gambar 10 Halaman Tambah Data Training

Gambar 4.10 menjelaskan dari fitur tambah data training dengan ditampilkan beberapa form sesuai atribut yang di butuhkan, dengan tombol tambah sebagai aksi untuk menambahkan data yang telah terisi.

4.10. Tampilan Halaman Update Data Training

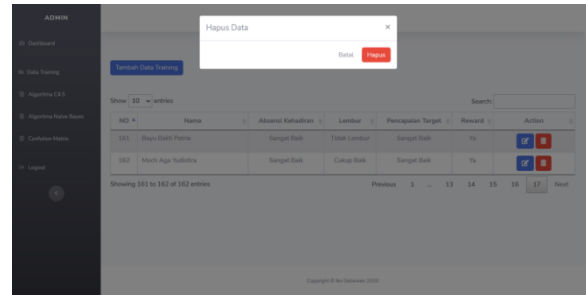
Pada halaman *update* data *training* admin dapat memperbarui data karyawan beserta atribut atribut lainnya. Tampilan halaman *update* data *training* dapat ditunjukkan pada gambar 4.11

Gambar 11 Halaman Update Data Training

Gambar 4.11 menjelaskan dari fitur *update* data training dengan menampilkan data sesuai dengan data mana yang akan dilakukan perubahan, dengan tombol perbarui sebagai aksi untuk memperbarui data yang telah diperbarui

4.2.1 Tampilan Halaman Hapus Data Training

Melakukan hapus data dapat langsung menggunakan konfirmasi notifikasi seperti gambar 4.12 berikut :



Gambar 4.12 Halaman Hapus Data Training

Gambar 4.12 menjelaskan dari fitur hapus data training dengan menekan tombol yang mempunyai *icon* sampah yang berwarna merah pada halaman data training. Yang kemudian memunculkan sebuah pesan untuk mengkonfirmasi apakah data akan dihapus atau tidak.

4.2.2 Tampilan Halaman Algoritma Naïve Bayes

Pada halaman Algoritma Naïve Bayes menampilkan hasil perhitungan setiap probabilitas atribut. Serta hasil dari prediksi pada klasifikasi algoritma Naïve Bayes. Tampilan halaman Algoritma Naïve Bayes ditunjukkan pada gambar 4.13

Label	Absen			Lembed			Target		
	Sangat Baik	Baik	Tidak Baik	Cukup Baik	Baik	Sangat Baik	Baik	Sangat Baik	Reward
Ya	0.85	0.30	0.62	0.23	0.25	0.00	0.80	0.27	0.70
Tidak	0.15	0.70	0.38	0.77	0.75	0.00	0.20	0.73	0.30

No	Nama	Absensi Kehadiran	Lembed	Persepsi Target	Reward	Probabilitas Ya	Probabilitas Tidak	Prediksi
1	Stanley Besson	Sangat Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
2	Hani Romdhoni	Sangat Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
3	Hani Kurniawan	Sangat Baik	Tidak Lembed	Baik	Ya	0.06	0.04	Ya
4	Reward Ocha Prasediana	Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
5	Begang Waluya	Sangat Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
6	Sanardi	Sangat Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
7	Endri Agus Subianto	Sangat Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
8	Sulthan	Sangat Baik	Tidak Lembed	Sangat Baik	Ya	0.29	0.00	Ya
9	Ahadi Kusuma	Sangat Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak
10	Chai Treasa	Baik	Tidak Lembed	Kurang Baik	Tidak	0.00	0.00	Tidak

Gambar 4.13 Halaman Algoritma Naïve Bayes

Gambar 4.13 menjelaskan dari perhitungan Algoritma Naïve Bayes dengan menampilkan setiap probabilitas setiap atribut dengan hasil klasifikasi yang dibandingkan dengan data aslinya.

4.2.3 Tampilan Halaman Algoritma C4.5

Pada halaman Algoritma C4.5 menampilkan hasil perhitungan *entropy*, *gain*, *split info* dan *gain ratio* dapat setiap akar dan cabang pohon keputusan. Tampilan halaman Algoritma C4.5 ditunjukkan pada gambar 4.14

The screenshot shows a web application interface for the C4.5 algorithm. It displays two decision tree visualizations. The first, 'Rasio 1.1', has a root node 'Total' with an entropy of 0.99. It splits into 'Alasan' (0.63) and 'Tidak Alasan' (0.36). 'Alasan' further splits into 'Sangat Baik' (0.75) and 'Buruk' (0.25). 'Sangat Baik' splits into 'Lumbar' (0.94) and 'Tidak Lumbar' (0.06). 'Buruk' splits into 'Target' (1.00) and 'Tidak Target' (0.00). The second, 'Cabang 1.1', has a root node 'Total' with an entropy of 0.99. It splits into 'Alasan' (0.63) and 'Tidak Alasan' (0.36). 'Alasan' further splits into 'Sangat Baik' (0.75) and 'Buruk' (0.25). 'Sangat Baik' splits into 'Lumbar' (0.94) and 'Tidak Lumbar' (0.06). 'Buruk' splits into 'Target' (1.00) and 'Tidak Target' (0.00).

Gambar 4.14 Halaman Algoritma C4.5

Gambar 4.14 menjelaskan dari perhitungan Algoritma C4.5 dengan menampilkan setiap perhitungan pada cabang pohon keputusan dan hasil klasifikasi yang dibandingkan dengan data aslinya.

4.2.4 Tampilan Halaman Confusion Matrix

Pada halaman *Confusion Matrix* menampilkan hasil kinerja dari algoritma Naïve Bayes dan C4.5. Tampilan halaman *Confusion Matrix* ditunjukkan pada gambar 4.15

The screenshot shows a web application interface for the Confusion Matrix. It displays two tables. The first, 'Kinerja Algoritma Naïve Bayes', shows True Positive (TP) as 113, True Negative (TN) as 48, False Positive (FP) as 1, and False Negative (FN) as 0. The second, 'Kinerja Algoritma C4.5', shows True Positive (TP) as 113, True Negative (TN) as 48, False Positive (FP) as 0, and False Negative (FN) as 1. Both tables also show Accuracy (Akurasi) as 0.99383, Precision (Presisi) as 0.99123, and Recall (Recall) as 1.00000.

Gambar 4.15 Halaman Confusion Matrix

Gambar 4.15 menjelaskan dari perhitungan *Confusion Matrix* dengan menampilkan setiap hasil akurasi, presisi dan *recall* dari masing-masing algoritma.

4.11. Pengujian Sistem dengan Black Box

Pengujian sistem merupakan proses menampilkan sistem dengan maksud untuk menemukan adanya kesalahan atau tidak pada fitur-fitur aplikasi. Hasil pengujian sistem menggunakan metode *black box* ditunjukkan pada Tabel 4.12

Tabel 4.12 Pengujian Sistem dengan Black Box

No.	Pengujian	Nama Event	Hasil yang Diharapkan	Kesimpulan
1.	Halaman Login	Klik tombol login	Tampil halaman <i>dashboard</i>	Berhasil
2.	Sidebar	Navigasi <i>dashboard</i>	Tampil halaman <i>dashboard</i>	Berhasil
3.	Sidebar	Navigasi algoritma C4.5	Tampil halaman algoritma C4.5	Berhasil
4.	Sidebar	Navigasi Algoritma Naïve Bayes	Tampil halaman Algoritma Naïve Bayes	Berhasil
5.	Sidebar	Tombol <i>logout</i>	Tampil halaman login	Berhasil
6.	Halaman Data Training	Tambah Data Training	Menampilkan halaman tambah data <i>training</i>	Berhasil
7.	Halaman Data Training	Update Data Training	Menampilkan halaman <i>update</i> data <i>training</i> berdasarkan id data yang dipilih	Berhasil
8.	Halaman Data Training	Pencarian Data	Mencari data berdasarkan kata kunci	Berhasil
9.	Halaman Algoritma C4.5	Menampilkan Hasil Perhitungan dan prediksi Algoritma C4.5	Menampilkan dan mencari data Hasil Perhitungan dan prediksi Algoritma C4.5 berdasarkan kata kunci tertentu	Berhasil
10.	Halaman Algoritma Naïve Bayes	Menampilkan Hasil Perhitungan dan prediksi Algoritma Naïve Bayes	Menampilkan dan mencari data Hasil Perhitungan dan prediksi Algoritma Naïve Bayes berdasarkan kata kunci tertentu	Berhasil
11.	Halaman Confusion Matrix	Menampilkan hasil <i>Confusion Matrix</i>	Menampilkan hasil presisi, akurasi dan <i>recall</i> dari kedua metode yang dibandingkan	Berhasil

4.12. Pengujian Fungsionalitas Sistem

Pada pengujian Fungsionalitas sistem akan dilakukan percobaan untuk menguji semua fungsi yang ada pada sistem apakah bekerja dengan normal atau tidak. Pengujian dilakukan pada 3 browser pada perangkat dengan spesifikasi CPU yang digunakan adalah Core i5-8265U 8th, RAM 8 GB. Hasil dari pengujian fungsionalitas sistem dapat ditunjukkan pada Tabel 4.13 sebagai berikut :

Tabel 4.13 Pengujian Fungsionalitas Sistem

Fungsi	Browser		
	Microsoft Edge	Google Chrome	Firefox Developer Edition
Autentikasi	√	√	√
Dashboard	√	√	√
Data Training	√	√	√

Tambah Data Training	√	√	√
Ubah Data Training	√	√	√
Hapus Data Training	√	√	√
Algoritma C4.5	√	√	√
Algoritma Naïve Bayes	√	√	√
Confution Matrix	√	√	√

4.13. Pengujian Pengguna

Pada pengujian pengguna, merupakan proses pengujian sistem dengan cara melakukan langsung kepada pengguna yang nantinya akan menggunakan sistem tersebut, dalam kasus ini pengguna yang menggunakan hanya 1 orang saja yaitu dari pihak direktur untuk mengetahui data sebelum serta hasil prediksi dari data sebelumnya. Hasil dari pengujian pengguna dapat dilihat pada tabel 4.14 berikut :

Tabel 4.14 Pengujian Pengguna

No.	Pertanyaan	Hasil Uji			
		SS	S	KS	TS
1.	Apakah aplikasi perbandingan algoritma C4.5 dan metode naïve bayes untuk menentukan karyawan terbaik mudah digunakan?	1	0	0	0
2.	Apakah fitur dari aplikasi ini mudah di pahami?	1	0	0	0
3.	Apakah informasi yang di sajikan pada aplikasi ini mudah dipahami?	0	1	0	0
4.	Apakah warna tombol, jenis font, dan warna font terlihat jelas dan sesuai?	1	0	0	0
Jumlah		3	1	0	0
Total Respon		4			

Keterangan :

SS : Sangat Setuju

S : Setuju

KS : Kurang Setuju

TS : Tidak Setuju

Berdasarkan tabel 4.14 setelah dilakukan pengujian langsung pada pengguna dapat di lihat 75% pengguna mengatakan sangat setuju terhadap fitur dan kenyamanan dari website yang telah digunakan, serta 25% untuk setuju bahkan tidak ada pengguna memiliki nilai kurang setuju maupun tidak setuju.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil implementasi dari perancangan aplikasi perbandingan algoritma C4.5 dan Naïve Bayes untuk menentukan karyawan terbaik dengan 4 atribut, dapat ditarik kesimpulan antara lain :

1. Dengan menggunakan data karyawan pada bulan Juli hingga November 2020 sebanyak 162 data, Algoritma C4.5 dan Naïve Bayes memiliki kinerja sama baik nya dengan nilai Akurasi 99,38%, Presisi 99,12% dan *recall* 100%
2. Dengan menggunakan data karyawan pada bulan Juli hingga Desember 2020 sebanyak 187 data, Algoritma C4.5 memiliki keakurasian yang lebih baik dibandingkan dengan Algoritma Naïve Bayes dengan masing masing nilai 98,91% dan 96,21%. Dengan hasil ini algoritma C4.5 lebih digunakan untuk mengklasifikasikan data dibandingkan dengan Algoritma Naïve Bayes.
3. Dengan menggunakan data karyawan pada bulan Juli hingga Desember 2020 sebanyak 187 data, Algoritma C4.5 memiliki presisi yang lebih baik dibandingkan dengan Algoritma Naïve Bayes dengan masing masing nilai 99,17% dan 95,23%. Dengan hasil ini algoritma C4.5 lebih digunakan untuk untuk memprediksi data dibandingkan Naïve Bayes
4. Dengan menggunakan data karyawan pada bulan Juli hingga Desember 2020 sebanyak 187 data, Algoritma C4.5 dan Naïve Bayes memiliki kinerja sama baik nya untuk menemukan kembali sebuah informasi dengan nilai *recall* 99,17%.

5.2 Saran

Untuk pengembangan aplikasi yang lebih baik, ada beberapa saran pada aplikasi komparasi Algoritma C4.5 dan Naïve Bayes sebagai berikut :

1. Untuk mengetahui perbandingan dapat menggunakan model *ROC Curve*.
2. Untuk melakukan Klasifikasi dapat menggunakan algoritma lain seperti Algoritma Jaringan Saraf Tiruan atau K-NN.

DAFTAR PUSTAKA

- [1] Retno, Yustina Wahyu Utami, 2014, Implementasi Algoritma C4.5 Dalam Pembuatan Aplikasi Penunjang Keputusan Penerimaan Pegawai CV. Dinamika Ilmu. Jurnal Teknologi dan Komunikasi.
- [2] Marudut, Victor Mulia Siregar, 2018, Sistem Pendukung Keputusan Penentuan Insentif Bulanan Pegawai dengan Menggunakan Metode Naïve Bayes. Jurnal Sistem Informasi
- [3] Imam, Dimas Baihaqi, 2019, Perbandingan Metode Naïve Bayes dan C4.5 Untuk

- Memprediksi Mortalitas Pada peternakan Ayam Broiler. *Jurnal Teknik Industri, Mesin, Elektro dan Ilmu Komputer*.
- [4] Pauziah, Ulfa, 2017, Analisis Penentuan Karyawan terbaik Menggunakan Metode Algoritma Naïve Bayes (Studi Kasus PT. XYZ).
- [5] Saefudin, Sri Lestari, 2015, Sistem Pendukung Keputusan Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5. *Jurnal Program Teknik Informatika Vol. 2*.