



Klasifikasi Pegawai Terbaik Triwulan pada BPS Provinsi Gorontalo Menggunakan Algoritma *Naïve Bayes*

Hanna Fidri Mardiny^{1,*}, Frangky Tupamahu², Hilmansyah Gani³, Khairul Fathan Habie⁴

^{1,2,3,4}Sistem Informasi, Fakultas Sains dan Ilmu Komputer, Universitas Muhammadiyah Gorontalo, Indonesia

Informasi Artikel

Sejarah Artikel:

Submit: 23 Juni 2026

Revisi: 03 Juli 2026

Diterima: 10 Juli 2026

Diterbitkan: 13 Juli 2026

Kata Kunci

Naïve Bayes, Klasifikasi, Data Mining, Evaluasi Kinerja, Pegawai Terbaik

Korespondensi

E-mail: hannaafidri@gmail.com

A B S T R A C T

The selection of the best employee is an important performance evaluation process aimed at improving employee motivation and productivity. At the Statistics Indonesia (BPS) of Gorontalo Province, the selection process still involves subjective considerations, which may affect the consistency and objectivity of decision-making. Therefore, a data-driven approach is needed to support the evaluation process. This study aimed to implement the Naïve Bayes algorithm to classify the best quarterly employee based on employee performance assessment data. The dataset consisted of performance records from 54 employees of BPS Gorontalo Province collected from the first quarter of 2023 to the fourth quarter of 2024. The classification process utilized BerAKHLAK behavioral indicators, discipline indicators, and Employee Performance Achievement (CKP) as predictor variables, while the target variable was employee status, namely best employee and non-best employee. The research stages included data preparation, data transformation, training and testing data partitioning, model development using the Naïve Bayes algorithm, and model evaluation using a confusion matrix with accuracy, precision, recall, and F1-score metrics. The evaluation results showed that the proposed model achieved an accuracy, precision, recall, and F1-score of 100%, indicating high classification performance on the dataset used in this study. These findings demonstrate that the Naïve Bayes algorithm is effective in classifying employee performance and can be utilized as a decision-support tool for determining the best quarterly employee. The implementation of this method is expected to enhance the objectivity, consistency, and transparency of employee performance evaluation at Statistics Indonesia (BPS) of Gorontalo Province.

Abstrak

Pemilihan karyawan terbaik merupakan proses evaluasi kinerja penting yang bertujuan untuk meningkatkan motivasi dan produktivitas karyawan. Di Badan Pusat Statistik (BPS) Provinsi Gorontalo, proses seleksi masih melibatkan pertimbangan subjektif, yang dapat mempengaruhi konsistensi dan objektivitas pengambilan keputusan. Oleh karena itu, diperlukan pendekatan berbasis data untuk mendukung proses evaluasi. Penelitian ini bertujuan untuk menerapkan algoritma *Naïve Bayes* untuk mengklasifikasikan karyawan triwulanan terbaik berdasarkan data penilaian kinerja karyawan. Kumpulan data terdiri dari catatan kinerja dari 54 karyawan BPS Provinsi Gorontalo yang dikumpulkan sejak kuartal I-2023 hingga kuartal IV-2024. Proses klasifikasi menggunakan indikator perilaku BerAKHLAK, indikator disiplin, dan Employee Performance Achievement (CKP) sebagai variabel prediktor, sedangkan variabel sasaran adalah status karyawan yaitu karyawan terbaik dan karyawan non terbaik. Tahapan penelitian meliputi persiapan data, transformasi data, pelatihan dan pengujian partisi data, pengembangan model menggunakan algoritma *Naive Bayes*, dan evaluasi model menggunakan matriks kebingungan dengan akurasi, presisi, ingatan, dan metrik skor F1. Hasil evaluasi menunjukkan bahwa model yang diusulkan mencapai akurasi, presisi, ingatan, dan skor F1 sebesar 100%, menunjukkan kinerja klasifikasi yang tinggi pada kumpulan

data yang digunakan dalam penelitian ini. Temuan ini menunjukkan bahwa algoritma Naive Bayes efektif dalam mengklasifikasikan kinerja karyawan dan dapat digunakan sebagai alat pendukung keputusan untuk menentukan karyawan triwulanan terbaik. Penerapan metode ini diharapkan dapat meningkatkan objektivitas, konsistensi, dan transparansi evaluasi kinerja pegawai di Badan Pusat Statistik (BPS) Provinsi Gorontalo.

This is an open access article under the CC-BY-SA



1. Pendahuluan

Perkembangan teknologi informasi telah memberikan dampak yang signifikan terhadap berbagai aspek organisasi, mencakup pada pengelolaan sumber daya manusia. Pemanfaatan teknologi informasi memungkinkan aktivitas pengolahan data dilakukan secara lebih cepat, akurat, dan terstruktur sehingga mampu menghasilkan informasi sebagai pendukung pengambilan keputusan secara objektif dan dapat dipertanggungjawabkan. Salah satu implementasi teknologi informasi yang banyak digunakan dalam pengolahan data adalah data mining. Data mining merupakan tahapan mengidentifikasi pola serta pengetahuan dari kumpulan data sehingga menghasilkan informasi yang dapat dimanfaatkan sebagai landasan dalam pengambilan keputusan. Salah satu teknik dalam data mining yang banyak diterapkan adalah teknik klasifikasi, yaitu proses pengelompokan data ke dalam kelas tertentu berdasarkan karakteristik masing-masing data.

Di lingkungan Badan Pusat Statistik (BPS) Provinsi Gorontalo, evaluasi kinerja pegawai dilaksanakan secara berkala sebagai salah satu upaya untuk meningkatkan kualitas sumber daya manusia dan pelayanan publik. Salah satu bentuk evaluasi tersebut adalah pemilihan pegawai terbaik triwulan yang bertujuan memberikan apresiasi kepada pegawai dengan capaian kinerja terbaik berdasarkan indikator penilaian yang telah ditetapkan. Untuk mendukung proses tersebut, BPS Provinsi Gorontalo telah menggunakan Sistem Penilaian Pegawai Berkinerja Terbaik (SIPEKERTI) sebagai sarana pencatatan, pengelolaan, dan rekapitulasi data penilaian pegawai. Namun demikian, proses penentuan pegawai terbaik masih menghadapi beberapa kendala. Meskipun data penilaian telah tersedia melalui SIPEKERTI, proses penetapan pegawai terbaik masih memerlukan pertimbangan pimpinan terhadap hasil penilaian yang ada. Kondisi tersebut berpotensi menimbulkan subjektivitas dalam pengambilan keputusan serta membutuhkan waktu yang relatif lebih lama ketika jumlah data pegawai yang dievaluasi semakin banyak. Selain itu, belum adanya mekanisme klasifikasi berbasis data menyebabkan pemanfaatan data historis penilaian belum dilakukan secara optimal untuk menghasilkan keputusan yang lebih objektif, konsisten, dan terukur.

Evaluasi kinerja Aparatur Sipil Negara (ASN) merupakan bagian penting dalam mewujudkan sistem manajemen kinerja yang profesional, objektif, terukur, akuntabel, partisipatif, dan transparan. Hal tersebut sejalan dengan Peraturan Pemerintah Republik Indonesia Nomor 30 Tahun 2019 tentang Penilaian Kinerja Pegawai Negeri Sipil yang menegaskan bahwa penilaian kinerja menjadi dasar pembinaan, pengembangan karier, pemberian penghargaan, serta peningkatan kinerja organisasi [1]. Selain itu, implementasi Core Values ASN BerAKHLAK sebagaimana diatur dalam Surat Edaran Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 20 Tahun 2021 menjadi landasan dalam membangun budaya kerja ASN yang berorientasi pelayanan, akuntabel, kompeten, harmonis, loyal, adaptif, dan kolaboratif [2]. Implementasi Core Values ASN BerAKHLAK tidak hanya menjadi landasan budaya kerja ASN, tetapi juga terbukti berkontribusi terhadap peningkatan kinerja pegawai pada instansi pemerintah [2], [3]. Sejalan dengan ketentuan tersebut, penelitian ini menggunakan indikator BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP) sebagai variabel dalam proses klasifikasi sehingga model yang dibangun tidak hanya didasarkan pada pendekatan data mining, tetapi juga selaras dengan sistem evaluasi kinerja ASN yang berlaku. Oleh

karena itu, penerapan metode klasifikasi berbasis data historis diharapkan mampu mendukung proses penentuan pegawai terbaik secara lebih objektif, konsisten, transparan, dan akuntabel.

Salah satu pendekatan yang dapat diterapkan untuk mengatasi permasalahan tersebut yaitu penerapan algoritma Naïve Bayes. Naïve Bayes merupakan metode klasifikasi berbasis probabilitas yang dikembangkan berdasarkan Teorema Bayes dengan asumsi independensi antaratribut. Metode ini banyak digunakan karena memiliki proses komputasi yang sederhana, mampu mengolah data dalam jumlah besar, serta menunjukkan kinerja yang baik pada berbagai kasus klasifikasi. Dengan memanfaatkan data historis penilaian pegawai, algoritma Naïve Bayes dapat digunakan guna melakukan klasifikasi terhadap pegawai ke dalam kategori pegawai terbaik dan non-terbaik sehingga mampu mendukung proses pengambilan keputusan secara lebih objektif dan sistematis.

Penelitian mengenai penerapan metode klasifikasi telah banyak dilakukan di berbagai bidang. Penelitian yang dilakukan oleh Tupamahu dan Suleman berhasil menerapkan metode klasifikasi menggunakan Jaringan Syaraf Tiruan pada citra digital motif Karawo. Hasil penelitian menunjukkan bahwa teknik klasifikasi mampu mengidentifikasi objek berdasarkan karakteristik data yang dimiliki secara efektif sehingga mampu mendukung proses pengambilan keputusan berbasis data [4]. Penelitian lain yang dilaksanakan oleh Liputo, Tupamahu, dan rekan menerapkan algoritma Decision Tree dengan K-Fold Cross Validation untuk memprediksi kondisi kesehatan mental siswa sekolah dasar. Hasil penelitian menunjukkan bahwa metode klasifikasi mampu menghasilkan model prediksi yang dapat dimanfaatkan sebagai pendukung dalam pengambilan keputusan berbasis data [5].

Dalam bidang klasifikasi berbasis data historis, Ibrahim, Gani, Lamusu, dan Humuhungo membandingkan algoritma Naïve Bayes dan C4.5 pada proses klasifikasi bantuan rumah sehat. Hasil penelitian mengindikasikan bahwa algoritma Naïve Bayes memiliki kinerja yang baik dalam mengolah data historis untuk menghasilkan keputusan klasifikasi yang akurat [6]. Selain itu, penelitian yang dilaksanakan oleh Yunus, Malik, dan Gani mengenai sistem pendukung keputusan penerimaan karyawan menunjukkan bahwa penerapan metode komputasional mampu meningkatkan objektivitas dan konsistensi proses pengambilan keputusan dibandingkan metode konvensional [7]. Beberapa penelitian lain juga menunjukkan efektivitas metode Naïve Bayes dalam proses evaluasi dan klasifikasi kinerja. Kenny dkk. berhasil menerapkan metode Naïve Bayes untuk evaluasi kinerja karyawan dan memperoleh tingkat akurasi yang baik dalam proses klasifikasi [8]. Penelitian Poerwondono dan Perwitosari menunjukkan bahwa algoritma Naïve Bayes mampu dimanfaatkan dalam proses penilaian kinerja karyawan secara objektif berdasarkan data historis yang tersedia [9]. Selain itu, Wardah dkk. membuktikan bahwa algoritma Naïve Bayes dapat dimanfaatkan dalam proses penentuan pegawai terbaik secara konsisten dan terukur [10].

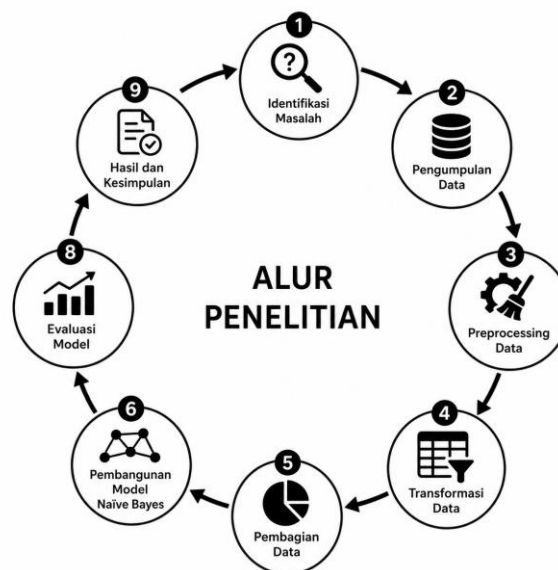
Meskipun penelitian sebelumnya telah menerapkan metode Naïve Bayes pada beragam kasus klasifikasi dan evaluasi kinerja, sebagian besar penelitian tersebut dilakukan pada sektor swasta, pendidikan, maupun klasifikasi data umum. Hingga saat ini belum terdapat penelitian yang secara khusus menerapkan metode Naïve Bayes untuk pemilihan pegawai terbaik triwulan pada BPS Provinsi Gorontalo menggunakan indikator penilaian Aparatur Sipil Negara (ASN) yang terdiri atas nilai BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP). Selain itu, penelitian ini memanfaatkan data historis hasil penilaian pegawai yang tersimpan dalam SIPEKERTI sebagai dasar pembentukan model klasifikasi. Integrasi indikator penilaian ASN dengan metode Naïve Bayes pada proses pemilihan pegawai terbaik triwulan menjadi kebaruan penelitian yang diharapkan mampu menghasilkan proses evaluasi yang lebih objektif, sistematis, dan berbasis data historis.

Berdasarkan permasalahan dan gap penelitian yang telah diuraikan, penelitian ini bertujuan untuk menerapkan metode Naïve Bayes dalam melakukan klasifikasi terhadap pegawai terbaik triwulan pada BPS Provinsi Gorontalo menggunakan variabel BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP). Kinerja model dievaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score untuk mengetahui tingkat performa kemampuan model dalam

melakukan klasifikasi terhadap pegawai terbaik secara objektif. Hasil penelitian diharapkan dapat menjadi alternatif pendukung keputusan yang membantu pimpinan dalam menentukan pegawai terbaik secara lebih konsisten, transparan, dan akuntabel.

2. Metode Penelitian

Penelitian ini dilaksanakan di Badan Pusat Statistik (BPS) Provinsi Gorontalo dengan memanfaatkan data historis penilaian pegawai yang digunakan dalam proses pemilihan pegawai terbaik triwulan. Penelitian ini menerapkan pendekatan kuantitatif melalui teknik data mining yang berfokus pada penerapan algoritma Categorical Naïve Bayes dalam mengklasifikasikan pegawai terbaik berdasarkan data historis penilaian kinerja pegawai. Penggunaan Categorical Naïve Bayes didasarkan pada karakteristik data penelitian yang telah ditransformasikan ke dalam bentuk data kategorikal sebelum proses pembentukan model klasifikasi, sehingga probabilitas setiap atribut dihitung berdasarkan frekuensi kemunculannya pada masing-masing kelas target. Pendekatan kuantitatif digunakan karena penelitian bertujuan mengolah data numerik untuk menghasilkan model klasifikasi yang dapat digunakan sebagai pendukung pengambilan keputusan secara objektif [11]. Data penelitian diperoleh dari Sistem Penilaian Pegawai Berkinerja Terbaik (SIPEKERTI) yang mencakup periode Triwulan I Tahun 2023 sampai dengan Triwulan IV Tahun 2024. Tahapan penelitian mencakup identifikasi masalah, pengumpulan data, preprocessing data, transformasi data, pembagian data, pembangunan model Naïve Bayes, pengujian model, evaluasi model, dan interpretasi hasil penelitian. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

2.1 Metode Pengumpulan Data

Pengumpulan data dilaksanakan melalui observasi, wawancara, dan studi dokumentasi. Observasi dilaksanakan untuk melakukan pengamatan secara langsung terhadap proses pemilihan pegawai terbaik triwulan yang berlangsung di BPS Provinsi Gorontalo. Wawancara dilakukan kepada pihak yang terlibat dalam proses penilaian guna memperoleh informasi mengenai mekanisme penilaian, kriteria yang digunakan, serta proses pengambilan keputusan. Studi dokumentasi dilakukan dengan mengumpulkan data historis penilaian pegawai yang tersimpan dalam SIPEKERTI serta berbagai referensi yang berkaitan dengan data mining, klasifikasi, dan metode Naïve Bayes [12].

2.2 Tahapan Penelitian

Dataset penelitian terdiri atas data penilaian 54 pegawai aktif BPS Provinsi Gorontalo. Variabel yang digunakan meliputi nilai BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP) sebagai atribut prediktor, sedangkan variabel target berupa kategori pegawai terbaik dan non-terbaik. Tahap preprocessing dilakukan melalui proses data cleaning, pemeriksaan data yang tidak lengkap, transformasi data, dan pengkodean atribut agar dapat diproses menggunakan metode klasifikasi [13], [14].

Tahap transformasi data dilakukan untuk mengubah nilai numerik pada setiap atribut menjadi data kategorikal sehingga sesuai dengan karakteristik algoritma Naïve Bayes yang digunakan dalam penelitian ini. Variabel BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP) dikelompokkan ke dalam beberapa kategori berdasarkan interval penilaian yang digunakan pada sistem penilaian pegawai. Hasil transformasi tersebut selanjutnya digunakan sebagai atribut masukan dalam proses pembentukan model klasifikasi.

Setelah tahapan preprocessing selesai dilakukan, dataset kemudian dipisahkan menjadi data latih (training data) dan data uji (testing data) menggunakan pendekatan Stratified Train-Test Split dengan proporsi 80%:20%. Pendekatan ini digunakan untuk mempertahankan distribusi kelas agar tetap seimbang pada data latih dan data uji sehingga model yang dihasilkan mampu merepresentasikan karakteristik data dengan lebih baik [13], [14]. Data latih dimanfaatkan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengevaluasi kemampuan model dalam memprediksi data yang belum pernah diproses sebelumnya.

2.3 Metode Naïve Bayes

Metode Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang dikembangkan berdasarkan Teorema Bayes dengan asumsi independensi antaratribut. Algoritma ini banyak digunakan pada proses klasifikasi karena memiliki tingkat kompleksitas yang rendah, mampu bekerja secara efektif pada data berukuran kecil maupun besar, serta memberikan kinerja yang baik dalam berbagai kasus prediksi dan klasifikasi [6], [12]. Pada penelitian ini, algoritma Naïve Bayes diterapkan pada data yang telah ditransformasikan ke dalam bentuk atribut kategorikal, yaitu kategori Tinggi, Sedang, dan Rendah untuk setiap variabel prediktor. Transformasi tersebut dilakukan agar setiap atribut dapat diproses berdasarkan distribusi probabilitas kemunculannya pada masing-masing kelas target, yaitu Pegawai Terbaik dan Non-Terbaik. Pendekatan ini sesuai dengan konsep supervised machine learning pada data kategorikal yang dijelaskan oleh Géron [14], di mana atribut kategorikal digunakan sebagai masukan untuk membangun model klasifikasi probabilistik. Persamaan Naïve Bayes yang digunakan dalam penelitian ini ditunjukkan pada Persamaan (1).

$$P(Y|X) = \frac{P(Y|X) \times P(Y)}{P(X)} \quad (1)$$

Keterangan:

$P(Y|X)$ = probabilitas suatu kelas berdasarkan atribut yang diamati;

$P(X|Y)$ = probabilitas atribut terhadap kelas tertentu;

$P(Y)$ = probabilitas awal setiap kelas;

$P(X)$ = probabilitas data yang diamati.

Proses klasifikasi dilakukan dengan menghitung probabilitas setiap kelas berdasarkan nilai atribut yang dimiliki oleh pegawai. Kelas dengan nilai probabilitas terbesar ditetapkan sebagai hasil klasifikasi, yaitu kategori Pegawai Terbaik atau Non-Terbaik [12].

2.4 Evaluasi Model

Evaluasi model dilaksanakan dengan menggunakan Confusion Matrix untuk mengukur kemampuan algoritma Naïve Bayes dalam melakukan klasifikasi terhadap pegawai terbaik. Pengukuran performa model dilaksanakan menggunakan metrik Accuracy, Precision, Recall, dan F1-Score yang dihitung berdasarkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) [15].

Nilai Accuracy dihitung menggunakan Persamaan (2).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Nilai Precision dihitung menggunakan Persamaan (3).

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Nilai Recall dihitung menggunakan Persamaan (4).

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

Nilai F1-Score dihitung menggunakan Persamaan (5).

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

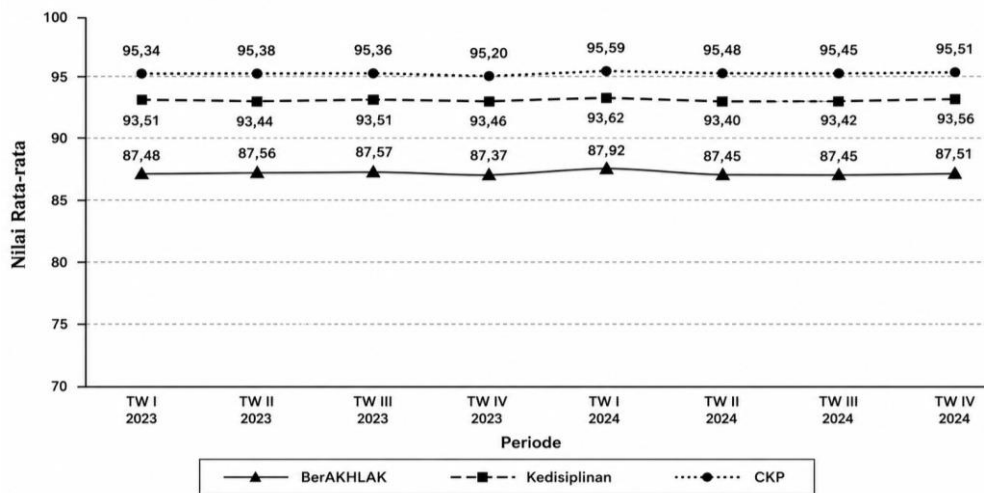
Pada tahap awal, dataset dipisahkan menjadi data latih dan data uji menggunakan pendekatan Stratified Train-Test Split dengan proporsi 80:20 untuk membangun dan menguji model [13]. Selanjutnya, untuk memperoleh gambaran performa model yang lebih stabil serta mengurangi potensi bias akibat pembagian data, dilakukan evaluasi tambahan menggunakan metode Stratified 5-Fold Cross Validation [13], [16]. Pendekatan ini memungkinkan setiap subset data digunakan secara bergantian sebagai data latih dan data uji sehingga hasil evaluasi lebih representatif terhadap kemampuan model dalam melakukan klasifikasi [13], [16]. Nilai accuracy, precision, recall, dan F1-score yang dilaporkan merupakan rata-rata hasil pengujian dari seluruh proses validasi.

3. Hasil dan Pembahasan

3.1 Hasil Penelitian

Penelitian ini menggunakan data historis penilaian pegawai BPS Provinsi Gorontalo periode Triwulan I Tahun 2023 sampai dengan Triwulan IV Tahun 2024 yang diperoleh dari Sistem Penilaian Pegawai Berkinerja Terbaik (SIPEKERTI). Dataset terdiri dari 54 pegawai aktif yang dievaluasi berdasarkan tiga atribut utama yaitu BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP). Ketiga atribut tersebut digunakan sebagai variabel prediktor untuk menentukan kategori pegawai terbaik dan non-terbaik.

Hasil analisis menunjukkan bahwa nilai CKP memiliki tingkat konsistensi yang tinggi pada seluruh periode pengamatan dengan rentang nilai 91,60–99,90. Variabel Kedisiplinan juga menunjukkan pola yang relatif stabil dengan rentang nilai 87,50–98,70. Sementara itu, atribut BerAKHLAK memiliki variasi nilai yang lebih besar dibandingkan atribut lainnya dengan rentang nilai 73,80–94,50. Variasi tersebut menunjukkan bahwa atribut BerAKHLAK memiliki kemampuan yang lebih baik dalam membedakan karakteristik pegawai dibandingkan atribut lainnya. Sebaliknya, rentang nilai CKP dan Kedisiplinan yang relatif sempit menunjukkan bahwa sebagian besar pegawai memiliki capaian kinerja dan tingkat kedisiplinan yang tinggi, sehingga distribusi data pada kedua atribut tersebut cenderung lebih homogen.



Gambar 2. Distribusi Nilai BerAKHLAK, Kedisiplinan, dan CKP Periode 2023–2024

Berdasarkan Gambar 2 terlihat bahwa nilai CKP dan Kedisiplinan cenderung stabil pada kategori tinggi selama periode penelitian. Sebaliknya, nilai BerAKHLAK menunjukkan variasi yang lebih besar sehingga berpotensi menjadi atribut pembeda utama dalam proses klasifikasi pegawai terbaik. Perbedaan tingkat variasi antaratribut tersebut menunjukkan bahwa BerAKHLAK memberikan informasi yang lebih beragam dalam membedakan kelas pegawai, sedangkan CKP dan Kedisiplinan lebih menggambarkan karakteristik umum pegawai pada dataset penelitian.

3.2 Transformasi dan Pembagian Data

Sebelum dilakukan proses klasifikasi, data numerik ditransformasikan ke dalam kategori Tinggi, Sedang, dan Rendah berdasarkan interval penilaian yang digunakan dalam sistem evaluasi kinerja pegawai. Transformasi dilakukan untuk mempermudah proses perhitungan probabilitas pada algoritma Naïve Bayes serta mengubah atribut numerik menjadi data kategorikal yang sesuai dengan proses klasifikasi. Hasil transformasi menunjukkan bahwa atribut BerAKHLAK memiliki distribusi kategori yang bervariasi pada setiap triwulan, sedangkan atribut Kedisiplinan didominasi kategori Tinggi dan Sedang. Pada atribut CKP, seluruh data berada pada kategori Tinggi pada seluruh periode pengamatan.

Tabel 1. Hasil Transformasi Data Triwulan III Tahun 2024

Atribut	Tinggi	Sedang	Rendah	Total
BerAKHLAK	18	21	15	54
Kedisiplinan	46	8	0	54
CKP	54	0	0	54

Sumber: Hasil Transformasi Data Penelitian, 2026

Tabel 1 menunjukkan bahwa atribut CKP memiliki distribusi yang homogen pada kategori Tinggi, sedangkan atribut BerAKHLAK memiliki distribusi yang lebih beragam. Kondisi ini menunjukkan bahwa atribut BerAKHLAK memiliki kontribusi yang lebih besar dalam membentuk probabilitas klasifikasi dibandingkan atribut CKP. Sebaliknya, dominasi kategori Tinggi pada atribut CKP mengindikasikan bahwa variasi data pada atribut tersebut relatif rendah sehingga informasi yang diberikan dalam membedakan kelas pegawai menjadi lebih terbatas. Setelah proses transformasi selesai dilakukan, dataset dibagi menjadi data latih (training data) dan data uji (testing data) menggunakan metode Stratified Train-Test Split dengan proporsi 80:20 untuk menjaga distribusi kelas tetap seimbang [11]. Pendekatan ini dipilih agar proporsi kelas Pegawai Terbaik dan Non-Terbaik tetap terjaga pada data latih maupun data uji sehingga model dapat mempelajari karakteristik masing-masing kelas secara lebih representatif.

3.3 Hasil Klasifikasi Naïve Bayes

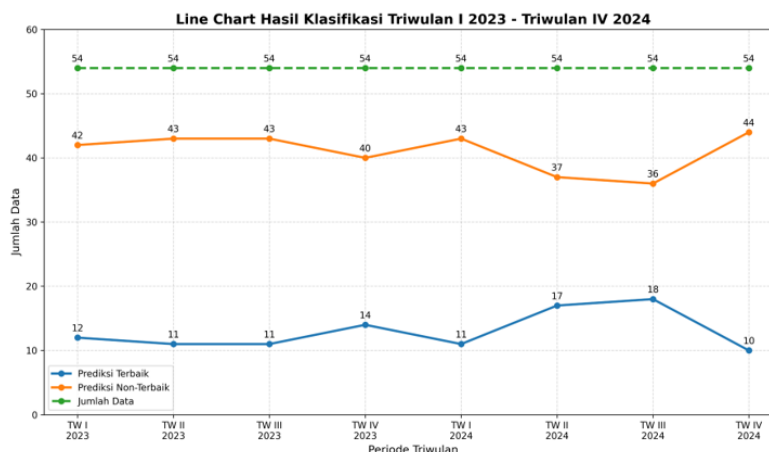
Proses klasifikasi dilakukan menggunakan algoritma Naïve Bayes berdasarkan probabilitas prior dan probabilitas likelihood dari setiap atribut. Hasil perhitungan menunjukkan bahwa kombinasi atribut BerAKHLAK Tinggi, Kedisiplinan Tinggi, dan CKP Tinggi selalu menghasilkan probabilitas posterior terbesar pada kelas Terbaik. Sebaliknya, kombinasi atribut BerAKHLAK Sedang maupun Rendah menghasilkan probabilitas posterior yang lebih tinggi pada kelas Non-Terbaik. Hasil tersebut menunjukkan bahwa atribut BerAKHLAK memiliki peran yang lebih dominan dalam membedakan kelas Pegawai Terbaik dan Non-Terbaik dibandingkan atribut lainnya.

Tabel 2. Ringkasan Hasil Klasifikasi Pegawai

Periode	Terbaik	Non-Terbaik	Total
TW I 2023	12	42	54
TW II 2023	11	43	54
TW III 2023	11	43	54
TW IV 2023	14	40	54
TW I 2024	11	43	54
TW II 2024	17	37	54
TW III 2024	18	36	54
TW IV 2024	10	44	54

Sumber: Hasil Klasifikasi Naïve Bayes, 2026

Hasil klasifikasi menunjukkan bahwa jumlah pegawai yang termasuk kategori Terbaik mengalami fluktuasi pada setiap triwulan. Jumlah tertinggi terjadi pada Triwulan III Tahun 2024 sebanyak 18 pegawai, sedangkan jumlah terendah terjadi pada Triwulan IV Tahun 2024 sebanyak 10 pegawai. Meskipun demikian, kelas Non-Terbaik tetap mendominasi pada seluruh periode pengamatan. Distribusi kelas tersebut menunjukkan bahwa dataset penelitian memiliki ketidakseimbangan (class imbalance) ringan, namun setiap kelas tetap terwakili pada seluruh periode pengamatan sehingga proses pembentukan model klasifikasi dapat dilakukan secara konsisten.



Gambar 3. Grafik Hasil Klasifikasi Pegawai Terbaik dan Non-Terbaik

Berdasarkan Gambar 3 terlihat bahwa jumlah pegawai pada kelas Non-Terbaik selalu lebih besar dibandingkan kelas Terbaik pada seluruh triwulan. Pola tersebut sejalan dengan distribusi data aktual yang menunjukkan dominasi kelas Non-Terbaik dalam dataset penelitian. Meskipun demikian, keberadaan kedua kelas pada setiap periode pengamatan memungkinkan algoritma Naïve Bayes mempelajari karakteristik masing-masing kelas berdasarkan kombinasi atribut BerAKHLAK, Kedisiplinan, dan CKP sebelum dilakukan proses evaluasi model.

3.4 Evaluasi Model

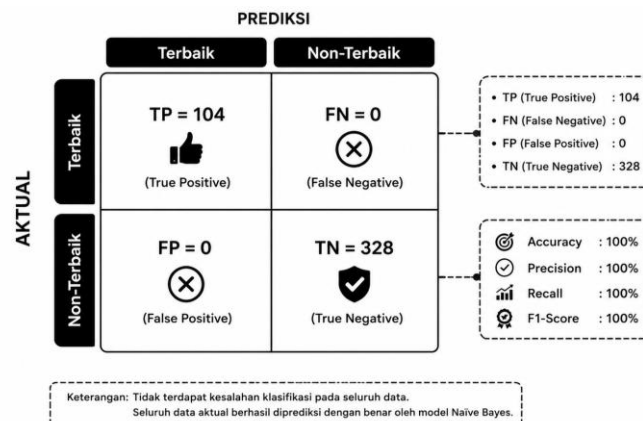
Evaluasi model dilakukan menggunakan Confusion Matrix dengan metrik Accuracy, Precision, Recall, dan F1-Score. Pengukuran dilakukan untuk mengetahui kemampuan algoritma Naïve Bayes dalam mengklasifikasikan pegawai terbaik berdasarkan atribut BerAKHLAK, Kedisiplinan, dan CKP. Selain menggunakan pembagian data latih dan data uji melalui pendekatan Stratified Train-Test Split, evaluasi juga dilakukan menggunakan Stratified 5-Fold Cross Validation untuk memperoleh gambaran performa model yang lebih stabil pada beberapa subset data.

Tabel 3. Ringkasan Evaluasi Model

	Nilai
Accuracy	100%
Precision	100%
Recall	100%
F1-Score	100%

Sumber: Hasil Evaluasi Model, 2026

Hasil evaluasi menunjukkan bahwa model berhasil mengklasifikasikan seluruh data dengan benar pada seluruh periode pengamatan. Tidak ditemukan nilai False Positive maupun False Negative, sehingga nilai Accuracy, Precision, Recall, dan F1-Score masing-masing mencapai 100%. Hasil tersebut diperoleh secara konsisten pada proses evaluasi menggunakan Confusion Matrix serta didukung oleh penerapan Stratified 5-Fold Cross Validation, yang menunjukkan bahwa model memiliki performa yang stabil pada setiap subset data yang diuji. Kondisi tersebut menunjukkan bahwa model mampu mengenali pola data pegawai terbaik dan non-terbaik berdasarkan data historis yang digunakan.



Gambar 4. Confusion Matrix Hasil Klasifikasi

Berdasarkan Confusion Matrix pada Gambar 4, seluruh data uji berhasil diklasifikasikan ke dalam kelas yang sesuai dengan data aktual. Tidak ditemukannya kesalahan klasifikasi (misclassification) menunjukkan bahwa model mampu membedakan karakteristik Pegawai Terbaik dan Non-Terbaik pada dataset penelitian berdasarkan kombinasi atribut BerAKHLAK, Kedisiplinan, dan CKP.

3.6 Pembahasan

Temuan penelitian mengindikasikan bahwa algoritma Naïve Bayes mampu melakukan klasifikasi terhadap pegawai terbaik triwulan pada BPS Provinsi Gorontalo dengan tingkat performa yang sangat tinggi. Kemampuan model dalam menghasilkan klasifikasi secara akurat dipengaruhi oleh karakteristik data yang tersedia, terutama atribut BerAKHLAK yang memiliki variasi nilai lebih tinggi dibandingkan atribut Kedisiplinan dan CKP. Variasi tersebut menyebabkan atribut BerAKHLAK menunjukkan kemampuan diskriminatif relatif lebih baik untuk membedakan pegawai

yang termasuk dalam kategori Terbaik dan Non-Terbaik. Sebaliknya, atribut CKP memiliki kontribusi yang relatif kecil karena seluruh data berada pada kategori Tinggi sehingga tidak memberikan perbedaan yang signifikan antar kelas.

Temuan penelitian ini mengindikasikan bahwa metode Naïve Bayes dapat memanfaatkan pola data historis untuk menghasilkan keputusan klasifikasi yang konsisten. Kondisi ini terjadi karena algoritma Naïve Bayes menghitung probabilitas setiap atribut terhadap kelas tertentu sehingga keputusan yang dihasilkan tidak hanya didasarkan pada satu atribut, tetapi merupakan kombinasi dari seluruh atribut yang digunakan. Dengan demikian, proses penentuan pegawai terbaik dapat dilakukan secara lebih objektif dibandingkan pendekatan manual yang sangat bergantung pada pertimbangan subjektif penilai.

Temuan penelitian ini sejalan dengan penelitian Ibrahim, Gani, Lamusu, dan Humolungo [6] yang mengindikasikan bahwa algoritma Naïve Bayes memiliki kinerja yang baik dalam mengolah data historis untuk menghasilkan klasifikasi secara akurat. Temuan penelitian ini juga sejalan dengan penelitian Kenny dkk. [8] yang berhasil menerapkan metode Naïve Bayes untuk evaluasi kinerja karyawan serta penelitian Wardah dkk. [10] yang mengindikasikan bahwa algoritma Naïve Bayes efektif diterapkan pada proses penentuan pegawai terbaik. Hasil penelitian Sangaji [17] juga menunjukkan bahwa algoritma Naïve Bayes mampu melakukan klasifikasi penilaian kinerja pegawai secara efektif sehingga dapat dimanfaatkan sebagai pendukung pengambilan keputusan berbasis data. Perbedaan penelitian ini dibandingkan penelitian terdahulu terletak pada objek penelitian serta variabel yang digunakan, yaitu penilaian Aparatur Sipil Negara (ASN) yang terdiri atas indikator BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP).

Nilai Accuracy, Precision, Recall, dan F1-Score sebesar 100% menunjukkan bahwa seluruh data uji berhasil diklasifikasikan dengan benar oleh model Naïve Bayes. Hasil tersebut mengindikasikan bahwa pola pada dataset penelitian dapat dikenali dengan baik oleh model berdasarkan kombinasi atribut BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP). Untuk meningkatkan keandalan hasil evaluasi, penelitian ini tidak hanya menggunakan pendekatan Stratified Train-Test Split, tetapi juga menerapkan Stratified 5-Fold Cross Validation sehingga performa model diuji pada beberapa subset data yang berbeda. Penggunaan confusion matrix sebagai dasar evaluasi performa model juga sejalan dengan penelitian Rudianto dkk. [18] yang memanfaatkan metrik accuracy, precision, recall, dan F1-score untuk membandingkan performa algoritma klasifikasi pada evaluasi kinerja karyawan. Kesamaan pendekatan evaluasi tersebut menunjukkan bahwa metrik-metrik tersebut merupakan ukuran yang umum digunakan untuk menilai keandalan model klasifikasi dalam mendukung pengambilan keputusan berbasis data.

Meskipun demikian, capaian performa sebesar 100% perlu diinterpretasikan secara hati-hati karena dapat dipengaruhi oleh karakteristik dataset yang relatif homogen, khususnya pada atribut CKP yang seluruhnya berada pada kategori Tinggi, serta jumlah dataset yang masih terbatas, yaitu 54 data pegawai. Oleh karena itu, hasil penelitian ini menunjukkan bahwa model memiliki performa yang sangat baik pada dataset yang digunakan, namun kemampuan generalisasinya masih perlu diuji menggunakan dataset yang lebih besar, lebih beragam, serta berasal dari periode maupun instansi yang berbeda agar diperoleh gambaran performa model yang lebih komprehensif.

Penelitian ini memiliki beberapa keterbatasan. Dataset yang digunakan hanya berasal dari satu instansi, yaitu BPS Provinsi Gorontalo, dengan jumlah data sebanyak 54 pegawai sehingga belum sepenuhnya merepresentasikan karakteristik penilaian kinerja ASN pada instansi lain. Selain itu, distribusi nilai pada beberapa atribut, khususnya CKP, cenderung homogen sehingga variasi data yang digunakan dalam proses klasifikasi masih terbatas. Penelitian selanjutnya disarankan menggunakan dataset yang lebih besar, melibatkan lebih banyak periode penilaian, serta membandingkan algoritma Naïve Bayes dengan metode klasifikasi lainnya agar diperoleh evaluasi performa model yang lebih komprehensif.

Secara praktis, penerapan metode Naïve Bayes pada penelitian ini memberikan alternatif pendukung keputusan yang lebih objektif dalam proses pemilihan pegawai terbaik di lingkungan BPS Provinsi Gorontalo. Model klasifikasi yang dihasilkan tidak dimaksudkan untuk menggantikan peran pengambil keputusan, tetapi sebagai alat bantu yang menyediakan rekomendasi berdasarkan analisis data historis sehingga proses evaluasi kinerja dapat dilakukan secara lebih konsisten, transparan, dan akuntabel.

4. Kesimpulan

Merujuk pada hasil penelitian yang telah dilakukan, algoritma Naïve Bayes berhasil diterapkan untuk mengklasifikasikan pegawai terbaik triwulan pada BPS Provinsi Gorontalo menggunakan atribut BerAKHLAK, Kedisiplinan, dan Capaian Kinerja Pegawai (CKP). Proses klasifikasi dilakukan menggunakan data historis penilaian pegawai yang diperoleh dari Sistem Penilaian Pegawai Berkinerja Terbaik (SIPEKERTI). Hasil penelitian menunjukkan bahwa model mampu mengidentifikasi kategori Pegawai Terbaik dan Non-Terbaik secara akurat berdasarkan pola data yang terdapat pada dataset penelitian.

Evaluasi model menggunakan Confusion Matrix menunjukkan nilai Accuracy, Precision, Recall, dan F1-Score masing-masing sebesar 100%. Hasil tersebut menunjukkan bahwa metode Naïve Bayes mampu melakukan klasifikasi dengan akurat pada dataset penelitian yang digunakan. Meskipun demikian, hasil tersebut perlu diinterpretasikan sesuai dengan karakteristik dataset yang digunakan, yaitu data historis penilaian sebanyak 54 pegawai dari BPS Provinsi Gorontalo dengan distribusi nilai yang relatif homogen pada beberapa atribut, sehingga kemampuan generalisasi model pada dataset yang lebih beragam masih memerlukan pengujian lebih lanjut. Penerapan metode ini dapat membantu meningkatkan objektivitas, konsistensi, dan transparansi dalam proses pemilihan pegawai terbaik triwulan di lingkungan BPS Provinsi Gorontalo sebagai pendukung pengambilan keputusan berbasis data. Pada penelitian selanjutnya, disarankan menambahkan jumlah data, melibatkan lebih banyak periode penilaian maupun instansi yang berbeda, serta menggunakan atribut penilaian yang lebih bervariasi agar kemampuan generalisasi model dapat dievaluasi pada kondisi data yang lebih kompleks.

Daftar Pustaka

- [1] "Peraturan Pemerintah Republik Indonesia Nomor 30 Tahun 2019 tentang Penilaian Kinerja Pegawai Negeri Sipil," Pemerintah Republik Indonesia, Jakarta, 2019.
- [2] "Surat Edaran Menteri PANRB Nomor 20 Tahun 2021 tentang Implementasi Core Values ASN BerAKHLAK dan Employer Branding ASN 'Bangga Melayani Bangsa,'" Kementerian PANRB, Jakarta, 2021.
- [3] A. C. Rahmat Kurniawan, D. Lituhyu, and D. Rostyaningsih, "Efektivitas Core Value BerAKHLAK dalam Meningkatkan Kinerja Pegawai Pada Badan Kepegawaian Daerah Kabupaten Rembang," *J. Public Policy Manag. Rev.*, vol. 13, no. 4, pp. 59–79, 2024.
- [4] F. Tupamahu and S. Suleman, "Klasifikasi Citra Digital Motif Karawo Berbasis Tekstur Menggunakan Jaringan Syaraf Tiruan," *J. Teknol. Inf. Indones.*, vol. 6, no. 1, 2021.
- [5] S. Liputo, F. Tupamahu, W. Hasyim, S. A. Sabiku, R. Parman, and A. Hanapi, "Prediction of Elementary School Students' Mental Health using Decision Tree Algorithm with K-Fold Cross-Validation in Bone Bolango Regency, Gorontalo Province," *J. La Multiapp*, vol. 4, no. 6, pp. 253–259, 2023, doi: 10.37899/journallamultiapp.v4i6.1005.
- [6] I. Ibrahim, H. Gani, R. Lamusu, and Y. Humolungo, "Perbandingan Algoritma Naïve Bayes dan C4.5 untuk Klasifikasi Bantuan Rumah Sehat," *J. Ilmu Komput.*, vol. 2, no. 1, pp. 72–80, 2022, doi: 10.31314/juik.v2i1.1477.

- [7] Y. Yunus, O. Maliki, and H. Gani, "Implementasi Metode AHP dalam Sistem Pendukung Keputusan untuk Penerimaan Karyawan di Sashop Kota Gorontalo," *J. Ilmu Komput.*, vol. 5, no. 1, pp. 70–79, 2025, doi: 10.31314/juik.v5i1.4036.
- [8] K. Kenny, D. Arisandi, and T. Sutrisno, "Evaluasi Penilaian Kinerja Karyawan dengan Metode Naïve Bayes," *Comput. J. Comput. Sci. Inf. Syst.*, vol. 8, no. 1, pp. 110–118, 2024, doi: 10.24912/computatio.v8i1.20287.
- [9] E. Poerwandono and F. J. Perwitosari, "Application of Data Mining for Employee Performance Assessment Using the Naïve Bayes Classification Method," *Saintek*, vol. 5, no. 1, pp. 120–122, 2023, doi: 10.55338/saintek.v5i1.1416.
- [10] N. N. Wardah, A. M. Yunita, A. Sugiarto, and R. Rizky, "Implementasi Metode Naïve Bayes Untuk Menentukan Pegawai Terbaik Di Universitas Mathla'ul Anwar Banten," *J. Tek. Inform. Unis*, vol. 9, no. 1, pp. 31–40, 2023, doi: 10.33592/jutis.v9i1.1291.
- [11] A. Algiffary and T. Sutabri, "Implementasi Machine Learning dengan Algoritma Naive Bayes Terhadap Sistem Informasi Pelayanan Pemberkasan Kepegawaian di BKPSDM Kota Palembang," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 284–301, 2023.
- [12] M. R. Khoirudin, M. Hasbi, B. Widada, K. Akhyar, and K. Sandradewi, "Klasifikasi Kelayakan Pegawai Kontrak Menjadi Pegawai Tetap Menggunakan Metode Naive Bayes," *J. Teknol. Inf. dan Komun.*, vol. 12, no. 2, pp. 7–15, 2024, doi: 10.30646/tikomsin.v12i2.863.
- [13] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Morgan Kaufmann, 2017.
- [14] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. O'Reilly Media, 2022.
- [15] E. Wahyudi, R. F. Wijaya, and Khairul, "K-Means and Naive Bayes Algorithms for Evaluation of Education Personnel Performance Based on SPMI Standards," *Sink. J. dan Penelit. Tek. Inform.*, vol. 8, no. 3, pp. 1872–1883, 2024, doi: 10.33395/sinkron.v8i3.13890.
- [16] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 4th ed. Morgan Kaufmann, 2023.
- [17] D. N. S. S. Sangaji, "Penerapan Algoritma Naïve Bayes untuk Klasifikasi Penilaian Kinerja Pegawai di Kedai XYZ," *Kohesi J. Sains dan Teknol.*, vol. 7, no. 1, pp. 61–70, 2025.
- [18] R. Rudianto, T. Wahyudi, and P. Handayani, "Perbandingan Akurasi C4.5 dan Naive Bayes Untuk Evaluasi Kinerja Karyawan PT Catur Sentosa Adiprana," *JUSIKOM PRIMA*, vol. 5, no. 2, pp. 61–66, 2022.