

**KLASIFIKASI KATEGORI KEHADIRAN KARYAWAN
MENGUNAKAN ALGORITMA *GRADIENT BOOSTED
TREES* (Studi Kasus: PT Guna Kemas Indah)**

SKRIPSI

Mutia Safitri

20210050070



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK KOMPUTER DAN DESAIN
UNIVERSITAS NUSA PUTRA
SUKABUMI
JULI 2025**

**KLASIFIKASI KATEGORI KEHADIRAN KARYAWAN
MENGUNAKAN ALGORITMA *GRADIENT BOOSTED
TREES* (Studi kasus: PT Guna Kemas Indah)**

SKRIPSI

*Diajukan Untuk Memenuhi Salah Satu Syarat Dalam Menempuh
Gelar Sarjana Komputer*

Mutia Safitri

20210050070



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK KOMPUTER DAN DESAIN
UNIVERSITAS NUSA PUTRA
SUKABUMI
JULI 2025**

PERNYATAAN PENULIS

JUDUL : KLASIFIKASI KATEGORI KEHADIRAN KARYAWAN
MENGUNAKAN ALGORITMA *GRADIENT BOOSTED
TREES* (Studi Kasus: PT Guna Kemas Indah)

Nama : Mutia Safitri

NIM : 20210050070

“Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti - bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”

Sukabumi, 22 Juli 2025



Mutia Safitri
Penulis

PENGESAHAN SKRIPSI

JUDUL : KLASIFIKASI KATEGORI KEHADIRAN KARYAWAN
MENGUNAKAN ALGORITMA *GRADIENT BOOSTED
TREES* (Studi Kasus: PT Guna Kemas Indah)

Nama : Mutia Safitri

NIM : 20210050070

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 22 Juli 2025. Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom).

Sukabumi, 22 Juli 2025

Pembimbing I

Pembimbing II

Sudin Saepudin, M.Kom

NIDN. 0414088608

Carti Irawan, S.T., M.Kom

NIDN. 0401108606

Ketua Penguji

Ketua Program Studi Sistem Informasi

Falentino Sembiring, M.Kom

NIDN. 0408029102

Falentino Sembiring, M.Kom

NIDN. 0408029102

Plh. Dekan Fakultas Teknik, Komputer dan Desain

Ir. Paikun, S.T., M.T., IPM, ASEAN Eng

NIDN. 0402037401

HALAMAN PERUNTUKAN

*Skripsi ini kupersembahkan untuk **Bapak** dan **Mamah** tercinta yang telah memberikan seluruh cinta kasihya serta doa yang tiada henti. Teruntuk **Bapak** sosok yang selalu kuat didepan mata, meski mungkin lelah tak pernah disuarakan, terimakasih atas kerja keras, semangat, dan kepercayaanmu padaku. Dan teruntuk **Mamah**, sumber doaku yang tak pernah putus, hangat pelukmu jadi tempat pulang paling nyaman.*

*Skripsi ini kupersembahkan juga kepada **Kedua Tete** yang selalu memberikan semangat, dukungan dan motivasi. Terimakasih atas setiap nasihat yang kalian berikan, atas teladan yang tanpa sadar telah menjadi arah langkahku.*

*Dan skripsi ini kupersembahkan juga untuk **Diriku Sendiri**, yang telah melewati begitu banyak rasa lelah, kecewa, bahkan ingin menyerah. Terimakasih kerana sudah tetap bertahan, terus bangkit, dan tidak berhenti berjuang meski jalan terasa berat. Terimakasih telah percaya bahwa setiap usaha, sekecil apapun itu, tetap berarti. Terimakasih karena telah menyelesaikan sesuatu yang dulu terasa mustahil.*

ABSTRACT

Employee attendance is an important factor in human resource management, because it affects the productivity and efficiency of the company. However, recording and analyzing employee attendance often experiences obstacles, especially in terms of the accuracy and effectiveness of the system used. This study aims to develop an employee attendance classification model using the Gradient Boosted Trees algorithm to increase the accuracy of grouping attendance categories such as Present, Permission, Sick, Leave, and Absent into the High, Medium, Low attendance categories.

The research method includes collecting employee attendance data at PT Guna Kemas Indah in 2024. Model evaluation uses the Accuracy, Precision, Recall, and Confusion Matrix metrics. The results of the study show that the developed model has an accuracy of 100% with a mean precision of 100% and a mean recall of 100%.

Keywords : *Employee Attendance, Gradient Boosted Trees, Classification, Data Mining, Machine Learning*

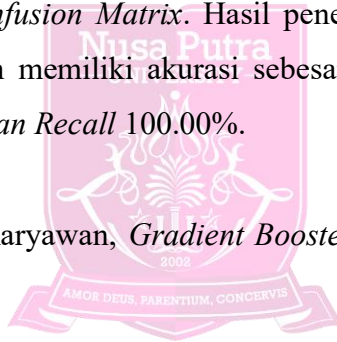


ABSTRAK

Kehadiran karyawan merupakan faktor penting dalam manajemen sumber daya manusia karena berpengaruh pada produktivitas dan efisiensi perusahaan. Namun, pencatatan dan analisis kehadiran karyawan seringkali mengalami kendala, terutama dalam hal akurasi dan efektivitas sistem yang digunakan. Penelitian ini bertujuan untuk mengembangkan model klasifikasi kehadiran karyawan menggunakan algoritma *Gradient Boosted Trees* guna meningkatkan ketepatan dalam pengelompokan kategori kehadiran, seperti Hadir, Izin, Sakit, Cuti dan Absen kedalam pengelompokan Kategori kehadiran Tinggi, Sedang dan Rendah.

Metode penelitian meliputi pengumpulan data absensi karyawan PT Guna Kemas Indah pada tahun 2024. Evaluasi model menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *Confusion Matrix*. Hasil penelitian menunjukkan bahwa model yang dikembangkan memiliki akurasi sebesar 100.00%, dengan *Mean Precision* 100.00% dan *Mean Recall* 100.00%.

Kata kunci : Kehadiran Karyawan, *Gradient Boosted Trees*, Klasifikasi, *Data Mining*, *Machine Learning*



KATA PENGANTAR

Puji Syukur kami panjatkan ke hadirat Allah SWT karena atas rahmat, karunia, dan hidayah-Nya, penulis dapat menyelesaikan penyusunan skripsi yang berjudul **“KLASIFIKASI KATEGORI KEHADIRAN KARYAWAN MENGGUNAKAN ALGORITMA *GRADIENT BOOSTED TREES* (Studi Kasus: PT Guna Kemas Indah)”** sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Fakultas Teknik Komputer dan Desain Universitas Nusa Putra.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa keberhasilan ini tidak lepas dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terimakasih kepada:

1. Bapak DR. Kurniawan, S.T., M.Si., MM., Selaku Rektor Universitas Nusa Putra Sukabumi.
2. Bapak Ir. Paikun, S.T., M.T., IPM., Sebagai Dekan Fakultas Teknik Komputer dan Desain Universitas Nusa Putra Sukabumi.
3. Bapak Falentino Sembiring, M.Kom., Selaku Ketua Program Studi Sistem Informasi Universitas Nusa Putra Sukabumi.
4. Bapak Sudin Saepudin, M.Kom., Selaku Dosen Pembimbing I Universitas Nusa Putra Sukabumi yang telah sabar membimbing dan memberikan arahan selama penyusunan skripsi ini.
5. Ibu Carti Irawan, S.T., M.Kom., Selaku Dosen Pembimbing II Universitas Nusa Putra Sukabumi yang telah sabar membimbing dan memberikan arahan selama penyusunan skripsi ini.
6. Dosen Penguji 1 Bapak Falentino Sembiring, M.Kom. dan Dosen Penguji 2 Ibu Arny Lattu, S.Pd.Kom., M.Kom.
7. Seluruh jajaran Dosen Program Studi Sistem Informasi Universitas Nusa Putra Sukabumi.
8. Pimpinan dan Staff HRD PT Guna Kemas Indah, yang telah memberikan izin, data, serta dukungan dalam proses pengumpulan informasi yang sangat membantu dalam penyusunan skripsi ini.

9. Kedua Orang tua tercinta terhebat, yang selalu berjuang untuk kehidupan penulis, untuk kasih sayang dan cintanya yang tulus, doa yang tak pernah putus, materi, semangat, motivasi dan dukungan yang tiada henti.
10. Kedua Kaka penulis yaitu Liana dan Fitriani yang selalu memberikan dukungan disituasi sulit penulis dan menjadi motivasi untuk menyelesaikan skripsi ini.
11. Support system terbaik Rd. M. Rafli Nursiddiq S.Ds. yang selalu menemani, memberikan dukungan dan motivasi untuk menyelesaikan skripsi ini.
12. Teman-teman mahasiswa Program Studi Sistem Informasi angkatan 2021, khususnya teman-teman tercinta diantaranya Restu, Nabila, Kaysa dan Sindy yang telah kebersamai selama perkuliahan ini.
13. Teman terdekat penulis yaitu Fathin, Karin, Widia dan Meyra yang telah memberikan semangat pada penulis dalam proses penelitian ini.
14. Serta masih banyak lagi pihak-pihak yang sangat berpengaruh dalam proses penyusunan skripsi ini.

Dengan ini semua semoga atas izin Allah SWT semua pihak yang telah membantu proses penyelesaian skripsi ini kebaikannya Allah balas dengan pahala yang berlipat ganda.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi penyempurnaan dimasa yang akan datang. Semoga skripsi ini dapat memberikan manfaat dan kontribusi positif khususnya dalam bidang *data mining*. Amin Yaa Rabbal 'Alamiin.

Sukabumi, 22 Juli 2025

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik UNIVERSITAS NUSA PUTRA, saya yang bertanda tangan dibawah ini:

Nama : Mutia Safitri
NIM : 20210050070
Program Studi : Sistem Informasi
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, meyetujui untuk memberikan kepada Universitas Nusa Putra **Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalty- Free Right*)** atas karya ilmiah saya yang berjudul : *KLASIFIKASI KATEGORI KEHADIRAN KARYAWAN MENGGUNAKAN ALGORITMA GRADIENT BOOSTED TREES* (Studi Kasus: P.T Guna Kemas Indah) beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Nusa Putra berhak menyimpan, mengalih media/format kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : SUKABUMI

Pada Tanggal : Juli 2025

Yang Menyatakan

(Mutia Safitri)

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN PENULIS	i
PENGESAHAN SKRIPSI.....	iii
HALAMAN PERUNTUKAN	iv
ABSTRACT	iv
ABSTRAK	v
KATA PENGANTAR.....	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	15
1.1 Latar Belakang	15
1.2 Rumusan Masalah	17
1.3 Batasan Masalah.....	17
1.4 Tujuan Penelitian.....	18
1.5 Manfaat Penelitian.....	18
1.6 Sistematika Penulisan.....	19
BAB II TINJAUAN PUSTAKA	Error! Bookmark not defined.
2.1 Penelitian Terkait.....	Error! Bookmark not defined.
2.2 Landasan Teori	Error! Bookmark not defined.
2.2.1 Konsep Dasar Kehadiran Karyawan.....	Error! Bookmark not defined.
2.2.2 <i>Data Mining</i>	Error! Bookmark not defined.
2.2.3 Klasifikasi	Error! Bookmark not defined.
2.2.4 Algoritma <i>Gradient Boosted Trees</i> ...	Error! Bookmark not defined.
2.2.5 PT Guna Kemas Indah	Error! Bookmark not defined.
2.2.5 RapidMiner	Error! Bookmark not defined.
2.3 Kerangka Berpikir	Error! Bookmark not defined.
BAB III METODOLOGI PENELITIAN	Error! Bookmark not defined.

3.1 Tahapan Penelitian	Error! Bookmark not defined.
3.2 Alat dan Bahan	Error! Bookmark not defined.
3.3 Metode Pengumpulan Data	Error! Bookmark not defined.
3.3.1 Sumber Data	Error! Bookmark not defined.
3.3.2 Teknik Pengumpulan Data	Error! Bookmark not defined.
3.3.3 Penerapan Algoritma <i>Gradient Boosted Trees</i>	Error! Bookmark not defined.
3.4 Jadwal Penelitian.....	Error! Bookmark not defined.
BAB IV HASIL DAN PEMBAHASAN	Error! Bookmark not defined.
4.1 <i>Data Selection</i>	Error! Bookmark not defined.
4.2 <i>Preprocessing Data</i>	Error! Bookmark not defined.
4.2.1 Transformasi Struktur	Error! Bookmark not defined.
4.2.2 Penanganan <i>Missing Values</i>	Error! Bookmark not defined.
4.2.3 <i>Feature Engineering</i>	Error! Bookmark not defined.
4.3 Pembagian Data.....	Error! Bookmark not defined.
4.4 Penerapan Algoritma <i>Gradient Boosted Trees</i>	Error! Bookmark not defined.
4.5 Hasil Analisis.....	Error! Bookmark not defined.
4.5.1 <i>Confusion Matrix</i>	Error! Bookmark not defined.
4.5.2 <i>Precision dan Recall</i>	Error! Bookmark not defined.
4.5.3 <i>Accuracy Model</i>	Error! Bookmark not defined.
BAB V PENUTUP.....	21
5.1 Kesimpulan.....	21
5.2 Saran.....	22
DAFTAR PUSTAKA.....	22
LAMPIRAN.....	Error! Bookmark not defined.

DAFTAR TABEL

Tabel 3.1 Alat dan Bahan	Error! Bookmark not defined.
Tabel 3.2 Potongan <i>Dataset</i> Absensi 2024.....	18
Tabel 3.3 Penjelasan Komponen Rumus.....	23
Tabel 3.4 Kategori Kehadiran Berdasarkan Kebijakan Internal Perusahaan	23
Tabel 3.5 Jadwal Penelitian.....	26
Tabel 4.1 Hasil Transformasi Struktur di RapidMiner.....	Error! Bookmark not defined.
Tabel 4.2 Potongan <i>Dataset</i> Hasil Transformasi Struktur.....	32
Tabel 4.3 Potongan <i>Dataset</i> Hasil Penanganan <i>Missing Values</i>	36
Tabel 4.4 Potongan <i>Dataset</i> Hasil <i>Feature Engineering</i>	40
Tabel 4.5 Potongan <i>Dataset</i> Data Latih	44
Tabel 4.6 Hasil <i>Precision</i> dan <i>Recall</i> (<i>Threshold Rule-based</i>).....	Error! Bookmark not defined.
Tabel 4.7 Hasil <i>Precision</i> dan <i>Recall</i> (Kebijakan Internal Perusahaan).....	Error! Bookmark not defined.
Tabel 4.8 Hasil <i>Accuracy</i> Model (<i>Threshold Rule-based</i>).....	Error! Bookmark not defined.
Tabel 4.9 Hasil <i>Accuracy</i> Model (Kebijakan Internal Perusahaan)	Error! Bookmark not defined.





DAFTAR GAMBAR

Gambar 1.1 Grafik Jumlah Karyawan	15
Gambar 2.1 Proses <i>Knowledge Discovery In Database</i>	Error! Bookmark not defined.
Gambar 2.2 Kerangka Berpikir	Error! Bookmark not defined.
Gambar 2.3 Kerangka Berpikir	Error! Bookmark not defined.
Gambar 3.1 Tahapan Penelitian	Error! Bookmark not defined.
Gambar 3.2 <i>Confusion Matrix</i>	Error! Bookmark not defined.
Gambar 4.1 Operator <i>Read CSV</i>	Error! Bookmark not defined.
Gambar 4.2 Proses Data <i>Selection</i> (1).....	Error! Bookmark not defined.
Gambar 4.3 Proses Data <i>Selection</i> (2).....	Error! Bookmark not defined.
Gambar 4.4 Operator <i>Pivot</i>	Error! Bookmark not defined.
Gambar 4.5 Parameter <i>Grup By Attributes</i>	Error! Bookmark not defined.
Gambar 4.6 Parameter <i>Aggregation Attributes</i>	Error! Bookmark not defined.
Gambar 4.7 Operator <i>Replace Missing Values</i>	Error! Bookmark not defined.
Gambar 4.8 Parameter <i>Missing Values</i>	Error! Bookmark not defined.
Gambar 4.9 Operator <i>Generate Attributes</i>	Error! Bookmark not defined.
Gambar 4.10 Parameter <i>Generate Attributes</i> (1) ...	Error! Bookmark not defined.
Gambar 4.11 Parameter <i>Generate Attributes</i> (2)....	Error! Bookmark not defined.
Gambar 4.12 Parameter <i>Generate Attributes</i> (3) ...	Error! Bookmark not defined.
Gambar 4.13 Operator <i>Set Role</i> dan Split Data.....	Error! Bookmark not defined.
Gambar 4.14 Target Label.....	Error! Bookmark not defined.
Gambar 4.15 <i>Ratio</i> Pembagian <i>Dataset</i>	Error! Bookmark not defined.
Gambar 4.16 Model Klasifikasi <i>Gradient Boosted Trees</i>	Error! Bookmark not defined.
Gambar 4.17 Proses Algoritma <i>Gradient Boosted Trees</i>	Error! Bookmark not defined.
Gambar 4.18 Hasil <i>Confusion Matrix</i> (<i>Threshold Rule-based</i>).....	Error! Bookmark not defined.
Gambar 4.19 Hasil <i>Confusion Matrix</i> (Kebijakan Internal Perusahaan)	Error! Bookmark not defined.



DAFTAR LAMPIRAN

Lampiran 1. Foto Gedung PT Guna Kemas Indah.	Error! Bookmark not defined.
Lampiran 2. Alamat PT Guna Kemas Indah	Error! Bookmark not defined.
Lampiran 3. Pengambilan Data Absensi Karyawan 2024	Error! Bookmark not defined.0
Lampiran 4. Surat Pengantar Untuk Perusahaan....	Error! Bookmark not defined.
Lampiran 5. Surat Penerimaan Dari Perusahaan.....	62
Lampiran 6. <i>Curriculum Vitae</i>	63
Lampiran 7. <i>Letter of Acceptance</i> (LOA)	64
Lampiran 8. Berita Acara Revisi Skripsi.....	65
Lampiran 9. Cek Hasil Plagiarisme	66



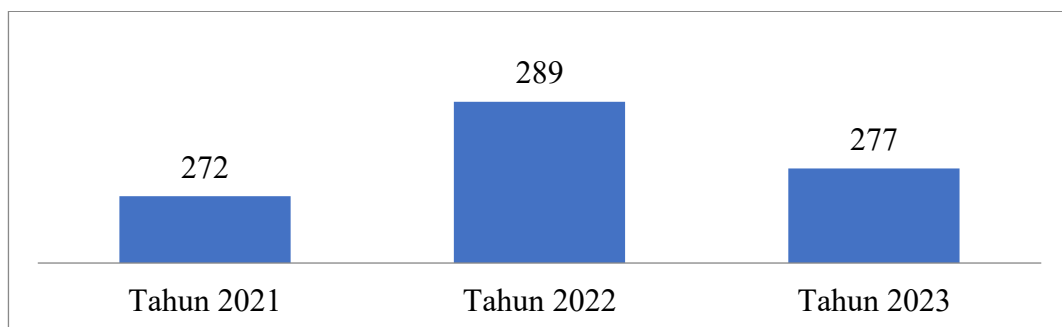
BAB I PENDAHULUAN

1.1 Latar Belakang

Sistem absensi dan informasi karyawan merupakan sebuah komponen penting dalam manajemen sumber daya manusia di suatu perusahaan. Kehadiran karyawan yang tepat waktu dan informasi karyawan yang akurat sangat diperlukan untuk mendukung produktivitas dan efisiensi operasional perusahaan. [1]

Produktivitas kerja dan efisiensi karyawan menjadi salah satu faktor utama bagi perusahaan untuk terus berkembang menjadi lebih baik [2]. Oleh karena itu, pemantauan dan analisis pola kehadiran karyawan menjadi aspek yang sangat penting dalam manajemen sumber daya manusia.

PT Guna Kemas Indah ialah perusahaan manufaktur yang fokus pada produksi kemasan plastik. Perusahaan ini pencatatan kehadiran karyawan sering kali dilakukan menggunakan sistem kehadiran digital, seperti penggunaan *fingerprint* dan kartu RFID, memungkinkan pencatatan yang lebih akurat dan efisien. Namun, meskipun teknologi kehadiran semakin berkembang, perusahaan masih kesulitan dalam menganalisis kehadiran karyawan secara akurat. Berikut jumlah karyawan ditahun 2021, 2022 dan 2023:



Gambar 1.1 Grafik Jumlah Karyawan

Seiring berkembangnya zaman, berkembang pula teknologi-teknologi yang mempermudah manusia melakukan suatu hal [3], salah satunya adalah kehadiran. Analisis data kehadiran karyawan dapat dilakukan secara lebih sistematis dan cerdas. *Machine learning* memungkinkan perusahaan untuk

mengidentifikasi serta mengklasifikasikan kategori kehadiran karyawan. Salah satu algoritma yang dipilih karena efektif dalam tugas klasifikasi adalah *Gradient Boosted Trees* (GBT). Algoritma ini mampu menghasilkan model prediktif yang akurat dengan mengombinasikan banyak pohon keputusan (*decision trees*) dalam proses pembelajaran yang berulang.

Gradient Boosted Trees telah digunakan dalam berbagai klasifikasi, termasuk klasifikasi kualitas air, klasifikasi status desa, klasifikasi serangan DDos, dan yang lainnya. Namun, penerapannya dalam klasifikasi kehadiran karyawan masih belum dieksplorasi. Dengan menggunakan algoritma ini, perusahaan dapat mengoptimalkan manajemen kehadiran karyawan, serta memberikan wawasan yang lebih baik bagi pengambil keputusan dalam meningkatkan efisiensi operasional.

Penelitian ini bertujuan untuk membangun model klasifikasi yang mampu mengelompokkan kehadiran karyawan seperti Hadir, Izin, Sakit, Cuti, dan Absen kedalam pengelompokan kategori kehadiran Tinggi, Sedang dan Rendah. Dengan menggunakan *dataset* kehadiran karyawan yang telah dikumpulkan, model ini akan dilatih untuk mengenali pola-pola tertentu yang mempengaruhi klasifikasi kehadiran. Beberapa faktor yang akan dipertimbangkan dalam analisis meliputi kategori kehadiran, total jumlah karyawan dan jumlah kehadiran efektif.

Dalam penelitian ini, *Gradient Boosted Trees* dipilih karena kemampuannya dalam menangani data yang memiliki kompleksitas tinggi serta keunggulannya dalam menangani fitur yang bersifat heterogen. Algoritma ini mampu mempelajari hubungan yang kompleks antara berbagai variabel dan secara iteratif meningkatkan akurasi model. Dengan menerapkan teknik ini, diharapkan model yang dihasilkan dapat memberikan prediksi yang lebih akurat dibandingkan metode konvensional lainnya.

Selain membangun model klasifikasi, penelitian ini juga akan mengevaluasi kinerja model yang dihasilkan dengan menggunakan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *confusion matrix*. Evaluasi ini penting untuk memastikan bahwa model yang dikembangkan tidak hanya memiliki performa tinggi pada data pelatihan, tetapi juga mampu melakukan generalisasi dengan baik terhadap data baru.

Hasil dari penelitian ini diharapkan dapat memberikan manfaat bagi perusahaan dalam meningkatkan efisiensi sistem kehadiran karyawan. Dengan adanya sistem klasifikasi otomatis berbasis *machine learning*, perusahaan dapat lebih cepat dan akurat dalam mengelola data kehadiran karyawan, mengurangi risiko manipulasi data, serta mengoptimalkan strategi manajemen sumber daya manusia. Selain itu, penelitian ini juga memberikan kontribusi dalam pengembangan ilmu pengetahuan di bidang kecerdasan buatan, khususnya dalam penerapan *machine learning* untuk analisis data kehadiran karyawan.

Dengan demikian, penelitian ini tidak hanya bermanfaat bagi perusahaan dalam aspek praktis, tetapi juga memberikan wawasan baru dalam pemanfaatan teknologi *machine learning* untuk mendukung pengambilan keputusan berbasis data. Diharapkan penelitian ini dapat menjadi referensi bagi studi-studi selanjutnya yang ingin mengeksplorasi lebih dalam mengenai penerapan *machine learning* dalam manajemen sumber daya manusia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, terdapat beberapa permasalahan yang perlu diidentifikasi, antara lain yaitu:

1. Bagaimana membangun model klasifikasi untuk mengelompokkan kategori kehadiran karyawan menggunakan algoritma *Gradient Boosted Trees*?
2. Bagaimana implementasi proses klasifikasi kategori kehadiran karyawan?
3. Faktor-faktor apa saja yang berpengaruh terhadap klasifikasi kehadiran karyawan dalam *dataset* yang digunakan?
4. Seberapa akurat model *Gradient Boosted Trees* dalam mengklasifikasikan kategori kehadiran berdasarkan data yang ada?
5. Bagaimana performa model klasifikasi diukur menggunakan metric evaluasi seperti *confusion matrix*, *accuracy*, *precision*, *recall*?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah, beberapa batasan masalah yang diterapkan adalah sebagai berikut:

1. Penelitian ini hanya akan menggunakan algoritma *Gradient Boosted Trees*. Algoritma lain seperti *Decision Tree*, *Support Vector Machine (SVM)*, *Random Forest* atau algoritma lainnya tidak akan dibahas dalam penelitian ini.
2. Proses klasifikasi dan evaluasi model dilakukan menggunakan perangkat lunak RapidMiner sebagai *tools* utama dalam pemodelan data mining. Penelitian ini tidak membahas atau membandingkan aplikasi lainnya.
3. Penelitian ini dibatasi pada data yang diambil dari PT Guna Kemas Indah pada periode tertentu. Data dari perusahaan lain atau periode waktu yang berbeda tidak akan dipertimbangkan.
4. Penelitian ini hanya akan membahas katageri kehadiran yang ditemui di PT Guna Kemas Indah, yaitu: Hadir, Sakit, Izin dan Cuti.
5. Evaluasi model akan dilakukan menggunakan metrik seperti *confusion matrix*, *accuracy*, *precision*, *recall*. Metrik lain seperti *ROC-AUC* tidak akan digunakan dalam penelitian ini.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Membangun model klasifikasi kehadiran karyawan dengan menggunakan algoritma *Gradient Boosted Trees*.
2. Mengidentifikasi faktor-faktor penting yang mempengaruhi pola kehadiran karyawan.
3. Mengukur dan mengevaluasi performa model klasifikasi berdasarkan metrik seperti *confusion matrix*, *accuracy*, *precision*, *recall*.
4. Membangun dan menguji model klasifikasi menggunakan aplikasi RapidMiner.
5. Memberikan solusi sistematis untuk membantu perusahaan dalam pengelolaan kehadiran karyawan.

1.5 Manfaat Penelitian

Manfaat penelitian sebagai berikut:

1. Bagi Perusahaan: Meningkatkan efisiensi dalam manajemen kehadiran karyawan dengan menggunakan sistem klasifikasi berbasis *machine learning*.

Dan Meminimalkan risiko kesalahan dan manipulasi dalam pencatatan absensi karyawan.

2. Bagi Karyawan: Meningkatkan transparansi dalam pencatatan dan evaluasi kehadiran, mendorong kesadaran dan kedisiplinan karyawan terhadap pola kehadiran.
3. Bagi Peneliti dan Akademisi: Penelitian ini memberikan pengalaman langsung bagi peneliti dalam menerapkan algoritma *machine learning*, khususnya *Gradient Boosted Trees* pada data kehadiran. Dan hasil penelitian ini dapat menjadi referensi bagi peneliti lain yang tertarik melakukan penelitian lebih lanjut dalam SDM dan teknologi informasi.
4. Bagi Pengembangan Sistem Teknologi Informasi: Menjadi acuan dalam mengintegrasikan model *machine learning* kedalam sistem absensi berbasis digital.
5. Bagi Pengambil Kebijakan dan Manajemen SDM: Memberikan dasar pengambilan keputusan yang lebih objektif dan berbasis data dalam mengelola karyawan, seperti pemberian reward, peringatan, atau pembinaan berdasarkan pola kehadiran yang terklasifikasi secara otomatis dan akurat.

1.6 Sistematika Penulisan

Peneliti menyusun beberapa bab untuk memudahkan pembaca dalam mempelajari dan memahami hasil penelitian yang dilakukan:

BAB I : PENDAHULUAN

Pada bab ini, peneliti menjelaskan latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II : TINJAUAN PUSTAKA

Bab ini berisi tentang penelitian terkait menggunakan beberapa teori dan kerangka pemikiran.

BAB III : METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang digunakan dalam penelitian, termasuk tahapan-tahapan penelitian, teknik pengumpulan data, serta alat bantu yang digunakan dalam proses klasifikasi.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian yang diperoleh serta analisis terhadap hasil tersebut. Pembahasan dilakukan untuk menjawab rumusan masalah dan mengaitkan hasil dengan teori yang telah dikemukakan pada bab sebelumnya.

BAB V : PENUTUP

Bab ini memuat kesimpulan dari hasil penelitian serta saran-saran yang dapat diberikan berdasarkan temuan penelitian untuk pengembangan lebih lanjut.



BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil klasifikasi yang telah dilakukan dapat disimpulkan bahwa:

- a. Proses *Data Selection* dan *Preprocessing Data* telah dilakukan secara sistematis dan menyeluruh, mencakup transformasi struktur data, penanganan missing values, serta feature engineering. Hasil dari proses ini adalah dataset yang bersih, terstruktur, dan siap digunakan dalam pembangunan model klasifikasi.
- b. Dataset dibagi menjadi 80% data latih dan 20% data uji. Kolom tingkat kehadiran dijadikan sebagai label (target) klasifikasi, sedangkan kolom-kolom seperti jumlah hadir, sakit, izin, cuti, absen, dan total karyawan digunakan sebagai atribut pendukung (*features*) yang berperan penting dalam proses pembelajaran model. Pemilihan atribut ini didasarkan pada analisis faktor-faktor yang secara langsung memengaruhi pola kehadiran karyawan.
- c. Hasil pengujian model menggunakan *Confusion Matrix* menunjukkan bahwa algoritma Gradient Boosted Trees mampu mengklasifikasikan seluruh kategori tingkat kehadiran (Tinggi, Sedang, dan Rendah) dengan sangat akurat, tanpa terjadi kesalahan klasifikasi (false positive maupun false negative). Hal ini menunjukkan bahwa model berhasil memahami hubungan antara fitur dan label dengan sangat baik.
- d. Nilai *Precision* dan *Recall* untuk masing-masing kategori mencapai 100%, yang berarti model memiliki kemampuan klasifikasi yang sangat baik, stabil, dan konsisten dalam mengenali pola data. Selain itu, model menunjukkan kepekaan dan ketepatan yang tinggi terhadap data aktual yang diuji.

Secara keseluruhan, penelitian ini berhasil membangun model klasifikasi tingkat kehadiran karyawan menggunakan algoritma Gradient Boosted Trees secara efektif. Model tidak hanya mampu memberikan prediksi yang akurat, tetapi juga mampu mengidentifikasi fitur-fitur penting yang berkontribusi dalam menentukan pola kehadiran, seperti jumlah hadir, absen, serta persentase kehadiran.

5.2 Saran

Berdasarkan hasil penelitian, peneliti memberikan beberapa saran sebagai berikut:

- a. Penelitian ini menggunakan data dari satu instansi atau perusahaan. Untuk penelitian selanjutnya, disarankan agar model diuji pada dataset yang lebih beragam dan berskala besar dari berbagai sektor industri agar validitas dan generalisasi model dapat ditingkatkan.
- b. Selain *Gradient Boosted Trees*, penelitian mendatang dapat membandingkan performa dengan algoritma lain seperti *Random Forest*, *Neural Network*, *XGBoost*, atau *LightGBM* dan yang lainnya guna mengetahui pendekatan mana yang paling optimal dalam kasus klasifikasi kehadiran karyawan.
- c. Sistem klasifikasi ini dapat dikembangkan lebih lanjut menjadi bagian dari sistem monitoring kehadiran berbasis aplikasi atau *dashboard real-time*, agar hasil klasifikasi tidak hanya bersifat analisis tetapi juga dapat langsung dimanfaatkan dalam pengambilan keputusan dilingkungan kerja.



DAFTAR PUSTAKA

- [1] D. Darmawan, R. Hidayat, and A. Kurniawan, “Pengembangan Sistem Absensi dan Informasi Karyawan Berbasis Web,” vol. 1, no. 6, pp. 928–933, 2024.
- [2] H. A. Setyadi and S. Sundari, “Sistem Informasi Manajemen Kehadiran Dan Jam Kerja Karyawan Untuk Kelengkapan Perhitungan Gaji Karyawan,” *Indones. J. Comput. Sci.*, vol. 1, no. 1, pp. 28–33, 2022, doi: 10.31294/ijcs.v1i1.1114.
- [3] W. A. Wahyuni and S. Saepudin, “Penerapan *Data Mining Clustering* Untuk Mengelompokkan Berbagai Jenis Merk Mesin Cuci,” *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform. Univ. Nusa Putra*, vol. 1, no. 1, pp. 306–313, 2021.
- [4] A. R. Fauzi ERR and I. Nur, L. Fitriana, “Klasifikasi sekuensial untuk pengelompokan data absensi karyawan menggunakan algoritma *decision tree* dan *feature engineering*,” vol. 2, no. 1, pp. 457–470, 2025.
- [5] Y. I. Mahendra and R. E. Putra, “Penerapan Algoritma *Gradient Boosted Decision Tree* (GBDT) untuk Klasifikasi Serangan,” *JINACS (Journal Informatics Comput. Sci. ISSN*, vol. 06, pp. 158–166, 2024.
- [6] Theodorus Ikhtiar Hulu, “Klasifikasi Proses Menentukan Kelayakan Karyawan Baru dengan Metode C4.5 pada Perusahaan *Outsourcing* PT.SAHABAT DUA PEMUDA” vol. 2, no. 4, 2024.
- [7] Hasbi Firmanysah and Zainal Abidin, “PENERAPAN ALGORITMA *GRADIENT BOOSTED DECISION TREES* PADA *ADABOOST* UNTUK KLASIFIKASI STATUS DESA” *JNUS (Jurnal Informatika UPS)* vol. 1, no. 1, 2022 .
- [8] A. Handayani, “Model Algoritma *Boosted Gradient Trees* untuk Penentuan Kelayakan Pemberian Kredit Koperasi,” *Sci. Sacra J. Sains*, vol. 2, no. 2, pp. 300–305, 2022, [Online]. Available: <http://pijarpemikiran.com/index.php/Scientia>
- [9] S. Wahyuni and M. Sulaeman, “Penerapan Algoritma *Deep Learning* Untuk Sistem Absensi Kehadiran Deteksi Wajah Di PT Karya Komponen Presisi,”

- J. Inform. SIMANTIK*, vol. 7, no. 1, pp. 5–6, 2022, [Online]. Available: <https://simantik.panca-sakti.ac.id/index.php/simantik/article/view/127>
- [10] Annisa Ilham Fathimah, Sudin Saepudin, and Edwinanto, “Penerapan *Data Mining* Dengan Metode *Apriori* Pada Penjualan Sembako (Studi Kasus: Grosir Sembako Lina),” *J. Rekayasa Teknol. Nusa Putra*, vol. 8, no. 2, pp. 123–131, 2022, doi: 10.52005/rekayasa.v8i2.271.
- [11] S. Suliman, “Data Mining Terhadap Prestasi Belajar Mahasiswa Berdasarkan Pergaulan dan Sosial Ekonomi Dengan Algoritma *K-Means Clustering*,” *Simkom*, vol. 6, no. 1, pp. 1–11, 2021, doi: 10.51717/simkom.v6i1.48.
- [12] A. Pebdika, R. Herdiana, and D. Solihudin, “Klasifikasi Menggunakan Metode *Naive Bayes* Untuk Menentukan Calon Penerima Pip,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 452–458, 2023, doi: 10.36040/jati.v7i1.6303.
- [13] Lydia Wahid Rizkallah, “Meningkatkan performa *Gradient Boosting Trees* pada permasalahan regresi” *Journal of Big Data*, 2025, <https://doi.org/10.1186/s40537-025-01071-3>
- [14] Fauziah and A. S. Ramadhantya, “Penggunaan Rapidminer Untuk Memprediksi Kelulusan Mahasiswa Dengan *Algorithm Naive Bayes*,” vol. 10, no. 1, Juni 2024.
- [15] Atika Juhaedah Alifah, Sudin Saepudin, and Carti Irawan, “IMPLEMENTASI ALGORITMA *K-MEANS CLUSTERING* DALAM MENGANALISIS KEPUASAN MASYARAKAT MENGENAI PELAYANAN PUBLIK (STUDI KASUS: BALAI PENGUJIAN STANDAR INSTRUMEN TANAMAN INDUSTRI DAN PENYEGAR),” *Jurnal Teknik Informatika (JUTIF)*, Vol. 5, No. 4, August 2024, pp. 487–496, DOI: <https://doi.org/10.52436/1.jutif.2024.5.4.2125>.
- [16] A. Wahyudi, S. Ovelia Tampubolon, N. Afrilia Putri, A. Ghassa, E. Rasywir, and D. Kisbianty, “Penerapan *Data Mining* Algoritma *Naive Bayes Classifier* Untuk Mengetahui Minat Beli Pelanggan Terhadap INDIHOME,” *J. Inform. Dan Rekayasa Komputer (JAKAKOM)*, vol. 2, no. 2, pp. 240–247, 2022, doi: 10.33998/jakakom.2022.2.2.111.

- [17] Iwan Pii, Nana Suarna, and Nining Rahaningsih, "PENERAPAN *DATA MINING* PADA PENJUALAN PRODUK PAKAIAN DAMEYRA FASHION MENGGUNAKAN METODE *K-MEANS CLUSTERING*," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 7 No.1 2023.
- [18] Arya Wijaya, Ahmad Faqih, Dodi Solihudin, Cep Lukman Rohmat, and Sandy Eka Permana, "PENERAPAN *ASSOCIATION RULES* MENGGUNAKAN ALGORITMA APRIORI UNTUK IDENTIFIKASI POLA PEMBELIAN," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7 No. 6, 2023.
- [19] Fauzan, and Didi Juardi, "Penerapan *Data Mining* Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma *NaïveBayes*," *Jurnal Ilmiah Informatika (JIF)*, Vol. 09 No. 02 (2021).
- [20] R. Nurhidayat and K. E. Dewi, "Penerapan Algoritma *K-Nearest Neighbor* Dan Fitur Ekstraksi *N-Gram* Dalam Analisis Sentimen Berbasis Aspek," *Komputa J. Ilm. Komput. dan Inform.*, vol. 12, no. 1, pp. 91–100, 2023, doi: 10.34010/komputa.v12i1.9458.
- [21] M. Azhari, Z. Situmorang, and R. Rosnelly, "Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, *Random Forest*, SVM dan *Naive Bayes*," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 640, 2021, doi: 10.30865/mib.v5i2.2937.