

Model Prediksi Perhitungan Gaji Karyawan Menggunakan Algoritma Decision Tree dengan Google Colab

Elsye Vegadinata ^{*1}, Fitriasih ²

^{1,2} Program Studi D3 Teknik Informatika, Politeknik Purbaya, Tegal, Indonesia

E-mail: ¹ elsy25vega@gmail.com

Abstrak

Penelitian ini bertujuan untuk membangun model prediksi perhitungan gaji karyawan menggunakan algoritma Decision Tree yang diimplementasikan dengan Google Colab. Data simulasi sebanyak 100 sampel dibangkitkan menggunakan Python dengan variabel kinerja, lembur, absensi, dan tunjangan sebagai prediktor, serta total gaji sebagai target. Data dibagi menjadi 80% data latih dan 20% data uji. Hasil analisis menunjukkan bahwa kinerja merupakan faktor dominan dengan korelasi 0.875 terhadap gaji, sedangkan absensi memiliki korelasi negatif sebesar -0.053. Model Decision Tree mencapai akurasi 90%, precision 88.89%, dan recall 88.89% dengan confusion matrix $\begin{bmatrix} 10 & 1 \\ 1 & 8 \end{bmatrix}$. Penelitian ini membuktikan bahwa algoritma Decision Tree efektif untuk memprediksi gaji karyawan berbasis kinerja dan dapat diimplementasikan dengan mudah menggunakan Google Colab. Penelitian ini masih menggunakan data simulasi sehingga perlu validasi lebih lanjut dengan data riil perusahaan.

Kata kunci— Decision Tree, Data Mining, Prediksi Gaji, Google Colab, Kinerja Karyawan

1. PENDAHULUAN

Dalam proses pengolahan dan analisis data pada penelitian ini, digunakan platform *Google Colab* sebagai tools utama dalam menjalankan algoritma *data mining*. *Google Colab* dipilih karena merupakan lingkungan pemrograman berbasis *cloud* yang memungkinkan pengguna menjalankan kode *Python* tanpa perlu melakukan instalasi perangkat lunak tambahan pada komputer lokal. Selain itu, *Google Colab* menyediakan berbagai pustaka (library) yang mendukung analisis data dan *machine learning*, sehingga memudahkan proses pengolahan data, pengujian model, serta visualisasi hasil analisis secara efisien dan fleksibel.

Perkembangan teknologi data mining dalam manajemen sumber daya manusia (SDM) mendorong perusahaan untuk memanfaatkan data internal secara sistematis dalam mendukung pengambilan keputusan strategis, termasuk dalam perhitungan gaji karyawan. Sistem penggajian yang akurat tidak hanya berperan dalam menjaga kepatuhan terhadap regulasi ketenagakerjaan, tetapi juga meningkatkan kepuasan dan memotivasi karyawan untuk tetap produktif [1]

Berbagai penelitian sebelumnya menunjukkan bahwa kinerja karyawan menjadi salah satu faktor utama dalam menentukan kompensasi atau penggajian. Penelitian yang dilakukan oleh Nitesh et al. menunjukkan bahwa penerapan algoritma *Decision Tree* [2] dapat digunakan untuk menganalisis dan memprediksi kenaikan gaji berdasarkan kinerja karyawan, dengan hasil bahwa variabel kinerja memiliki pengaruh yang signifikan terhadap keputusan pemberian kompensasi. Penelitian lain yang dilakukan oleh Hidayah dan Rozi juga membuktikan bahwa algoritma C4.5 mampu mengklasifikasikan dan memetakan tingkat kinerja karyawan secara sistematis serta memberikan tingkat akurasi yang baik dalam proses pengambilan keputusan [3]. Selain itu, beberapa penelitian juga mengungkapkan bahwa faktor lain seperti tingkat kehadiran (absensi) dan kedisiplinan kerja turut memengaruhi produktivitas serta status kerja karyawan dalam suatu organisasi [4][5].

Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada analisis kinerja karyawan atau faktor produktivitas secara terpisah, sehingga belum banyak yang mengintegrasikan berbagai variabel seperti kinerja, absensi, dan disiplin kerja secara bersamaan dalam sistem perhitungan gaji yang lebih komprehensif. Oleh karena itu, diperlukan penelitian lebih lanjut yang dapat menggabungkan berbagai faktor tersebut dengan memanfaatkan metode data mining sehingga dapat menghasilkan sistem penggajian yang lebih objektif, sistematis, dan akurat dalam mendukung pengambilan keputusan manajemen [6].

Algoritma Decision Tree banyak digunakan dalam data mining karena kemampuannya menghasilkan model prediksi yang mudah dipahami oleh pengambil keputusan dari berbagai latar belakang disiplin ilmu. Putra menyatakan bahwa *Decision Tree* memiliki keunggulan dalam hal interpretabilitas dibandingkan algoritma lain seperti *Random Forest* dan *Support Vector Regression* [7]. Penelitian lain juga menunjukkan efektivitas *Decision Tree* dalam memprediksi potensi atrisi karyawan dengan tingkat akurasi yang baik [8].

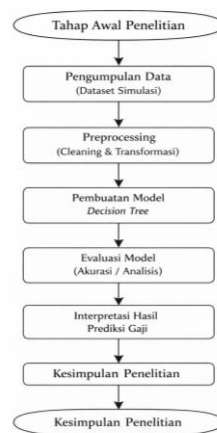
Meskipun berbagai penelitian telah mengkaji faktor-faktor yang memengaruhi pemberian kompensasi kepada karyawan, sebagian besar penelitian sebelumnya seperti yang dilakukan oleh Nitesh et al. serta Hidayah dan Rozi lebih berfokus pada analisis kinerja karyawan dan proses klasifikasi tingkat performa dalam organisasi. Penelitian-penelitian tersebut umumnya digunakan untuk mendukung evaluasi kinerja atau pengambilan keputusan terkait penempatan dan pengembangan karyawan. Namun demikian, kajian yang secara khusus mengarah pada pemanfaatan metode data mining untuk memprediksi atau membantu proses perhitungan gaji karyawan secara terstruktur masih relatif terbatas. Oleh karena itu, penelitian ini dilakukan untuk mengisi celah tersebut dengan menerapkan pendekatan data mining dalam mendukung sistem perhitungan gaji yang lebih sistematis dan berbasis data [9]. Selain itu, masih terbatas penelitian yang mendokumentasikan implementasi algoritma *Decision Tree* secara transparan menggunakan platform berbasis cloud seperti *Google Colab*. Dokumentasi yang sistematis sangat penting agar proses analisis dapat dilakukan secara konsisten serta memungkinkan hasil penelitian untuk diuji kembali atau diduplikasi oleh peneliti lain.

Berdasarkan celah tersebut, penelitian ini mengusulkan model prediksi perhitungan gaji karyawan menggunakan algoritma *Decision Tree* dengan memanfaatkan *Google Colab* sebagai platform implementasi. Penelitian ini menggunakan dataset simulasi yang merepresentasikan pola umum sistem penggajian pada perusahaan manufaktur dengan beberapa variabel utama, yaitu kinerja karyawan, lembur, absensi, dan tunjangan [10]. Kebaruan penelitian ini terletak pada dokumentasi implementasi yang transparan dan dapat diduplikasi melalui *Google Colab* serta analisis terhadap faktor-faktor yang paling berpengaruh dalam penentuan gaji karyawan.

Penelitian ini diharapkan dapat berkontribusi dalam pengembangan model prediksi gaji berbasis data mining pada manajemen sumber daya manusia.

2. METODE PENELITIAN

Penelitian ini dirancang menggunakan pendekatan data mining untuk membangun model prediksi perhitungan gaji karyawan dengan algoritma *Decision Tree*. Seluruh proses analisis dan implementasi model dilakukan menggunakan platform *Google Colab* dengan bahasa pemrograman *Python*. Alur penelitian ini meliputi tahap pengumpulan data, pra-proses data, pembangunan model, evaluasi model, hingga interpretasi hasil prediksi.



Gambar 1. Alur Penelitian

2.1 Data Penelitian

Penelitian ini menggunakan data simulasi yang dibangkitkan secara sintesis menggunakan library *numpy* pada *Python*. Data dirancang dengan pola yang mencerminkan karakteristik umum penggajian di perusahaan manufaktur skala menengah. Jumlah data yang dibangkitkan sebanyak 100 sampel karyawan dengan variabel sebagai berikut:

Tabel 1. Variabel Dataset Simulasi

No	Variable	Deskripsi	Rentang Nilai
1	Kinerja	Skor penilaian kinerja karyawan	60 – 100
2	Lembur	Jumlah jam lembur per bulan	0 – 20 jam
3	Absensi	Jumlah hari tidak hadir per bulan	0 – 10 hari
4	Tunjangan	Besaran tunjangan yang diterima karyawan	Rp200.000-Rp800.000
5	Gaji	Total gaji karyawan (variabel target)	Berdasarkan hasil perhitungan

2.2 Implementasi Google Colab

Penelitian ini diimplementasikan menggunakan *Google Colab (Google Collaboratory)*, sebuah platform berbasis *cloud* yang memungkinkan eksekusi kode *Python* secara interaktif tanpa memerlukan instalasi perangkat lunak tambahan. *Google Colab* dipilih karena mendukung reproduksibilitas penelitian dan memudahkan dokumentasi seluruh proses analisis.

1. Library *Python* yang digunakan dalam implementasi ini meliputi:
2. *pandas* dan *numpy*: Untuk manipulasi data dan pembangkitan data simulasi
3. *scikit-learn*: Untuk membangun model *Decision Tree* serta menghitung metrik evaluasi
4. *matplotlib* dan *graphviz*: Untuk memvisualisasikan struktur pohon keputusan
5. *seaborn*: Untuk visualisasi heatmap korelasi

Kode program dalam penelitian ini dikembangkan menggunakan bahasa pemrograman *Python* pada platform *Google Colab*. Model prediksi dibangun menggunakan algoritma *Decision Tree* dengan variabel kinerja, lembur, absensi, dan tunjangan sebagai fitur utama. Secara matematis, algoritma ini bekerja dengan membagi data berdasarkan nilai entropi dan information gain untuk menghasilkan struktur pohon keputusan yang digunakan dalam memprediksi besaran gaji karyawan.

2.3 Algoritma Decision Tree

Algoritma *Decision Tree* yang digunakan mengadopsi pendekatan C4.5 yang diimplementasikan menggunakan *DecisionTreeRegressor* dari library *scikit-learn*. Pohon keputusan dibangun dengan membagi data secara rekursif berdasarkan nilai information gain tertinggi hingga mencapai kondisi stopping criteria (kedalaman maksimum pohon = 4 dan jumlah sampel minimal pada daun = 2). Untuk menentukan atribut terbaik sebagai node pembagi, algoritma C4.5 menggunakan konsep *Entropy* dan *Information Gain*. *Entropy* digunakan untuk

mengukur tingkat ketidakpastian dalam suatu dataset, sedangkan *Information Gain* digunakan untuk mengukur seberapa besar pengurangan ketidakpastian setelah dilakukan pembagian berdasarkan atribut tertentu.

Perhitungan entropy ditunjukkan pada Persamaan

$$H(S) = - \sum_{i=1}^n p_i \log_2 p_i \quad (1)$$

Keterangan:

S = himpunan data

p_i = proporsi kelas ke-i dalam dataset

Sedangkan perhitungan information gain ditunjukkan pada Persamaan

$$\text{Gain}(S,A) = H(S) - \sum_{v=1}^n \left(\frac{|S_v|}{|S|} \right) H(S_v) \quad (2)$$

Keterangan:

A = atribut yang diuji

S_v = subset data hasil pembagian berdasarkan atribut A

$|S|$ = jumlah total data

Langkah-langkah pembangunan model dengan algoritma *Decision Tree* adalah sebagai berikut:

1. Pembangkitan data simulasi sebanyak 100 sampel
2. Pembagian data menjadi 80% data latih (training) dan 20% data uji (testing)
3. Pelatihan model *Decision Tree* menggunakan data latih
4. Prediksi gaji menggunakan data uji prediksi gaji menggunakan data uji

2.4 Evaluasi Model

Evaluasi performa model dilakukan menggunakan metrik standar dalam data mining. Karena prediksi gaji merupakan masalah regresi, dilakukan konversi ke klasifikasi dua kelas (gaji tinggi dan gaji rendah) berdasarkan nilai rata-rata gaji sebagai *threshold*. Metrik evaluasi yang digunakan meliputi:

1. *Akurasi*: Mengukur persentase prediksi yang benar terhadap total prediksi
2. *Precision*: Mengukur proporsi prediksi positif yang benar-benar positif
3. *Recall*: Mengukur proporsi data positif yang berhasil diprediksi dengan benar
4. *Confusion Matrix*: Menampilkan distribusi prediksi benar dan salah dalam bentuk matriks

Selain evaluasi model, dilakukan pula analisis korelasi untuk mengidentifikasi faktor-faktor dominan yang memengaruhi besaran gaji karyawan.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi model prediksi perhitungan gaji karyawan menggunakan algoritma *Decision Tree* berdasarkan dataset simulasi yang telah disusun. Analisis dilakukan terhadap beberapa variabel utama, yaitu kinerja, lembur, absensi, dan tunjangan, untuk mengidentifikasi pola hubungan yang memengaruhi penentuan gaji karyawan. Selain memaparkan hasil pembangunan dan evaluasi model, bagian ini juga membahas interpretasi hasil analisis guna mengetahui faktor-faktor yang paling dominan dalam menentukan besaran gaji karyawan.

3.1 Analisis Korelasi Variabel

Analisis korelasi dilakukan untuk mengidentifikasi hubungan antara variabel prediktor (kinerja, lembur, absensi, tunjangan) dengan variabel target (gaji). Hasil analisis korelasi ditunjukkan pada Tabel 2.

Tabel 2 Korelasi Variabel terhadap Total Gaji.

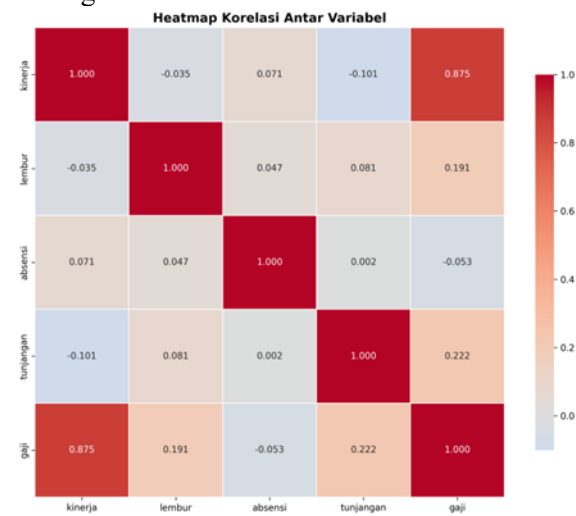
Variabel	Korelasi dengan Gaji	Interpretasi
Kinerja	0.875	Positif sangat kuat, faktor dominan

Tunjangan	0.222	Positif sedang
Lembur	0.191	Positif lemah
Absensi	-0.053	Negatif sangat lemah

Berdasarkan Tabel 2, variabel kinerja memiliki nilai korelasi tertinggi sebesar 0.875, yang menunjukkan hubungan positif dan sangat kuat terhadap total gaji. Hal ini menegaskan bahwa kinerja merupakan faktor dominan dalam menentukan besaran gaji karyawan. Semakin tinggi kinerja karyawan, semakin besar pula gaji yang diterima.

Variabel tunjangan dan lembur menunjukkan korelasi positif dengan nilai masing-masing 0.222 dan 0.191, yang mengindikasikan bahwa kedua variabel tersebut berkontribusi terhadap peningkatan gaji meskipun tidak sekuat kinerja. Sementara itu, variabel absensi menunjukkan korelasi negatif sebesar -0.053, yang berarti semakin tinggi tingkat ketidakhadiran karyawan, gaji cenderung menurun [11].

Visualisasi heatmap korelasi antar variabel ditampilkan pada Gambar 2 untuk memudahkan pemahaman hubungan antar variabel secara keseluruhan.

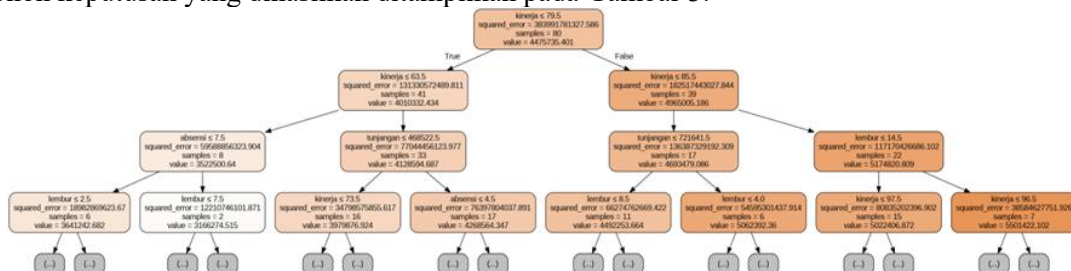


Gambar 2 Heatmap Korelasi Antar Variabel

Hasil analisis korelasi ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa kinerja merupakan faktor utama dalam penentuan kompensasi karyawan [12]. Pola ini juga mencerminkan penerapan prinsip performance-based pay dalam sistem penggajian, di mana karyawan dengan kinerja tinggi mendapatkan penghargaan finansial yang lebih besar.

3.2 Model Decision Tree

Model *Decision Tree* dibangun menggunakan data latih sebanyak 80 sampel. Struktur pohon keputusan yang dihasilkan ditampilkan pada Gambar 3.



Gambar 3 Pohon Keputusan Prediksi Gaji Karyawan.

Berdasarkan Gambar 2, variabel kinerja muncul sebagai root node (node akar) yang pertama kali membagi data. Hal ini mengkonfirmasi hasil analisis korelasi bahwa kinerja merupakan faktor paling dominan dalam menentukan gaji. Node akar membagi data berdasarkan nilai threshold kinerja, misalnya kinerja ≤ 79.5 dan kinerja > 79.5 .

Pada node cabang berikutnya, model mempertimbangkan variabel lain seperti lembur, absensi, dan tunjangan untuk memperhalus prediksi. Struktur pohon yang dihasilkan

menunjukkan bahwa karyawan dengan kinerja tinggi cenderung memiliki gaji tinggi, terutama jika didukung dengan jam lembur yang memadai dan tingkat absensi yang rendah.

Keunggulan *Decision Tree* terletak pada interpretabilitasnya, di mana manajemen dapat dengan mudah memahami faktor-faktor yang memengaruhi gaji karyawan tanpa memerlukan latar belakang statistik yang mendalam [13].

3.3 Simulasi Prediksi

Simulasi prediksi dilakukan pada 5 sampel data uji untuk membandingkan gaji aktual dengan gaji hasil prediksi model. Hasil simulasi ditampilkan pada Tabel 3.

Tabel 3 Simulasi Prediksi Gaji Karyawan (5 Sampel).

Kinerja	Lembur	Absensi	Tunjangan (Rp)	Gaji Aktual (Rp)	Gaji Prediksi (Rp)	Selisih (Rp)
67	3	0	377.247	3.835.984	3.867.726	31.742
61	8	1	589.957	3.989.712	3.601.465	388.247
75	4	0	798.135	4.752.959	4.440.270	312.689
66	0	3	318.834	3.650.908	3.867.726	216.818
87	0	7	430.229	4.719.688	4.907.036	187.348

Berdasarkan Tabel 3, terlihat bahwa gaji prediksi yang dihasilkan model memiliki selisih yang relatif kecil dibandingkan dengan gaji aktual. Rata-rata selisih prediksi berkisar antara Rp30.000 hingga Rp400.000, yang menunjukkan bahwa model mampu menangkap pola hubungan antara variabel input dan besaran gaji dengan cukup akurat.

Karyawan dengan kinerja tinggi (87) cenderung memiliki gaji tinggi meskipun jam lembur rendah, sedangkan karyawan dengan kinerja sedang (61-67) sangat dipengaruhi oleh kontribusi lembur dan tunjangan. Hal ini mengkonfirmasi bahwa model tidak hanya menghafal data, tetapi benar-benar belajar dari pola yang ada.

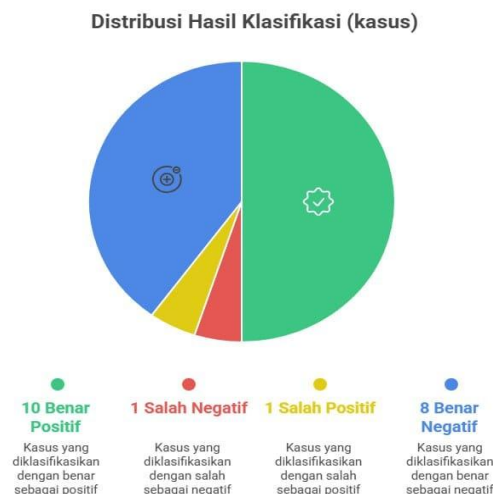
3.4 Evaluasi Model

Evaluasi performa model *Decision Tree* dilakukan menggunakan data uji sebanyak 20 sampel. Hasil evaluasi ditampilkan pada Tabel 4.

Tabel 4 Evaluasi Performa Model Decision Tree

Metrik Evaluasi	Hasil
Akurasi	90.00%
Precision	88.89%
Recall	88.89%

Confusion matrix yang dihasilkan adalah sebagai berikut:



Interpretasi confusion matrix:

1. True Negatif (TN) = 10: Model benar memprediksi 10 karyawan dengan gaji rendah
2. False Positif (FP) = 1: Model salah memprediksi 1 karyawan (seharusnya gaji rendah, diprediksi tinggi)

-
3. False Negatif (FN) = 1: Model salah memprediksi 1 karyawan (seharusnya gaji tinggi, diprediksi rendah)
 4. True Positif (TP) = 8: Model benar memprediksi 8 karyawan dengan gaji tinggi

Nilai akurasi sebesar 90% menunjukkan bahwa model mampu memprediksi gaji dengan tepat pada 18 dari 20 sampel data uji. *Precision* 88.89% mengindikasikan bahwa dari seluruh prediksi gaji tinggi yang dihasilkan model, 88.89% benar-benar merupakan karyawan dengan gaji tinggi. *Recall* 88.89% menunjukkan bahwa model berhasil menangkap 88.89% dari seluruh karyawan yang sebenarnya memiliki gaji tinggi [14].

Hasil evaluasi ini menegaskan bahwa model *Decision Tree* memiliki performa yang sangat baik dan reliabel untuk digunakan sebagai alat bantu prediksi gaji karyawan. Tingkat kesalahan prediksi yang rendah (hanya 2 dari 20 sampel) menunjukkan bahwa model mampu menggeneralisasi pola dengan baik.

3.5 Pembahasan

Hasil penelitian ini menunjukkan bahwa algoritma *Decision Tree* efektif digunakan untuk memprediksi gaji karyawan berdasarkan faktor kinerja, lembur, absensi, dan tunjangan. Temuan utama penelitian ini adalah:

1. Pertama, kinerja merupakan faktor dominan dengan korelasi 0.875 terhadap gaji. Hal ini sejalan dengan penelitian Nitesh et al [15] yang membuktikan bahwa kinerja memiliki pengaruh signifikan terhadap kenaikan gaji. Temuan ini juga mendukung penelitian Hidayah dan Rozi [3] yang menunjukkan bahwa algoritma C4.5 mampu memetakan kinerja karyawan secara sistematis.
2. Kedua, absensi memiliki korelasi negatif terhadap gaji, meski pun dengan nilai yang relatif kecil (-0.053). Hal ini sesuai dengan penelitian Alibasyah et al [16] dan Sahputra et al [17] yang menekankan bahwa tingkat kehadiran memengaruhi penilaian kelayakan karyawan.
3. Ketiga, model *Decision Tree* mencapai akurasi 90% dengan precision dan recall di atas 88%. Hasil ini lebih tinggi dibandingkan penelitian serupa yang dilakukan oleh Darma dan Wahyudin [18] yang melaporkan akurasi di kisaran 85-87% untuk prediksi gaji menggunakan algoritma C4.5.

Keunggulan utama *Decision Tree* terletak pada interpretabilitasnya, sebagaimana dinyatakan oleh Putra [6] bahwa *Decision Tree* lebih mudah dipahami dibandingkan algoritma seperti *Random Forest* dan *Support Vector Regression*. Struktur pohon yang dihasilkan (Gambar 2) memungkinkan manajemen untuk memahami secara visual faktor-faktor yang memengaruhi gaji karyawan.

Implementasi menggunakan *Google Colab* memberikan nilai tambah dalam hal transparansi dan reproduksibilitas penelitian. Seluruh kode program dan data simulasi dapat diakses publik, memungkinkan peneliti lain untuk memverifikasi dan mengembangkan hasil penelitian ini lebih lanjut. Pendekatan ini sejalan dengan prinsip *reproducible research* yang semakin penting dalam komunitas ilmiah [9].

Namun demikian, penelitian ini memiliki keterbatasan utama karena menggunakan data simulasi. Meskipun data dirancang dengan pola yang mencerminkan kondisi riil perusahaan, hasil yang diperoleh belum tentu sama persis dengan implementasi pada data riil. Oleh karena itu, diperlukan validasi lebih lanjut dengan data aktual perusahaan sebelum model dapat diimplementasikan secara operasional [10].

Implikasi manajerial dari penelitian ini adalah bahwa perusahaan dapat memanfaatkan model prediksi gaji berbasis *Decision Tree* untuk:

1. Mengevaluasi keadilan sistem penggajian yang ada
 2. Mengidentifikasi faktor-faktor yang perlu ditingkatkan untuk mendorong produktivitas karyawan
 3. Merencanakan anggaran gaji secara lebih akurat
 4. Memberikan transparansi kepada karyawan mengenai komponen-komponen yang memengaruhi gaji mereka
-

4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa:

1. Algoritma *Decision Tree* efektif digunakan untuk membangun model prediksi perhitungan gaji karyawan dengan tingkat akurasi 90%, precision 88.89%, dan recall 88.89%.
2. Faktor kinerja merupakan variabel paling dominan dalam menentukan besaran gaji karyawan dengan nilai korelasi 0.875, diikuti oleh tunjangan (0.222) dan lembur (0.191), sedangkan absensi memiliki korelasi negatif (-0.053).
3. Implementasi menggunakan *Google Colab* memungkinkan proses pembangunan model yang transparan, terdokumentasi dengan baik, dan dapat direproduksi oleh peneliti lain.
4. Model *Decision Tree* yang dihasilkan mampu menangkap pola hubungan antara variabel input dan besaran gaji, yang tercermin dari kecilnya selisih antara gaji aktual dan gaji prediksi pada simulasi 5 sampel karyawan.

Penelitian ini membuktikan bahwa pendekatan data mining dengan algoritma *Decision Tree* dapat mendukung pengelolaan gaji yang berbasis kinerja dan memberikan dasar bagi pengembangan sistem penggajian serta evaluasi sumber daya manusia di masa mendatang.

5. SARAN

Untuk pengembangan penelitian selanjutnya, disarankan beberapa hal berikut:

1. Melakukan validasi model menggunakan data riil perusahaan untuk memastikan keandalan model dalam kondisi operasional yang sebenarnya.
2. Mengintegrasikan model prediksi yang telah dibangun ke dalam sistem penggajian berbasis web atau mobile untuk memudahkan implementasi praktis di perusahaan.
3. Menambahkan variabel prediktor lain seperti masa kerja, tingkat pendidikan, dan jabatan struktural untuk meningkatkan akurasi prediksi.
4. Membandingkan performa algoritma *Decision Tree* dengan algoritma lain seperti *Random Forest*, *Gradient Boosting*, atau *Neural Network* untuk mendapatkan model dengan akurasi terbaik.

DAFTAR PUSTAKA

- [1] S. Syam *et al.*, *Data Mining: Teori dan Penerapannya dalam Berbagai Bidang*. PT. Sonpedia Publishing Indonesia, 2024.
- [2] I. Himawati and F. Fitriasih, "Sistem Informasi Pembayaran Pada Lembaga Pendidikan Anak Hebat (Ahe) Unit Brekat Berbasis Web," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 3610–3618, 2024, doi: 10.36040/jati.v8i3.9763.
- [3] F. Fitriasih, M. Ainurrohman, F. Fahrudin, and Y. Nurasri, "Pengujian Keamanan Website JDIIH Kab.Tegal Menggunakan Acunetix dengan Standar ISO/IEC/27001:2013," *J. Pengabd. Sos.*, vol. 2, no. 3, pp. 3320–3323, 2025, doi: 10.59837/agthp967.
- [4] A. N. Z. Hidayah and A. F. Rozi, "Penerapan Data Mining Dalam Menentukan Kinerja Karyawan Terbaik Dengan Menggunakan Metode Algoritma C4. 5 (Studi Kasus: Universitas Mercu Buana Yogyakarta)," *J. Inf. Syst. Artif. Intell.*, vol. 1, no. 2, pp. 117–127, 2021.
- [5] A. Aditama *et al.*, "Klasifikasi Potensi Turnover Karyawan Berdasarkan Data Kinerja Menggunakan Algoritma C4. 5 Studi Kasus Pt. Republic Technology Nusantara," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 2, pp. 2723–2730, 2025.
- [6] N. R. Silaen *et al.*, "Kinerja karyawan," 2021.
- [7] A. Nelson, A. Oktalia, L. Willyanto, and M. Ella, "Analisis jam kerja dan waktu lembur pada perusahaan manufaktur Kota Batam," *E-Bisnis J. Ilm. Ekon. dan Bisnis*, vol. 16, no. 1, pp. 181–189, 2023.
- [8] A. G. Mulia, "Sistem informasi absensi berbasis web di Politeknik Negeri Padang," *J.*

-
- Teknol. Inf. Indones.*, vol. 5, no. 1, pp. 11–17, 2020.
- [9] M. Fahrurrozi and S. Sriyanto, “Prediksi Loyalitas Karyawan Dengan Menggunakan Fitur Optimasi Pembobotan PSO dan Algoritma Decision Tree C4. 5,” *TEKNIKA*, vol. 17, no. 1, pp. 195–203, 2023.
 - [10] R. N. Pratama, R. Risdianti, and N. F. Diniaty, “Sistem Informasi Penggajian Berbasis Web Pada Perusahaan Inkindo Provinsi Kalimantan Selatan,” *J. INTEKNA Inf. Tek. dan Niaga*, vol. 22, no. 02, pp. 76–94, 2022.
 - [11] I. M. Ienda, A. S. Wati, and L. R. Rahmi, “Pemanfaatan Algoritma Decision Tree Pada Machine Learning Dalam Penentuan Klasifikasi Kinerja Karyawan Pada CV Duta Media,” *JUPITER J. Penelit. Ilmu dan Teknol. Komput.*, vol. 17, no. 1, pp. 169–177, 2025.
 - [12] F. F. Harryanto and S. Hansun, “Penerapan Algoritma C4. 5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE,” *Jatisi*, vol. 3, no. 2, pp. 95–103, 2017.
 - [13] A. Perdana, *Data analytics: keterampilan teknis akuntan dan auditor di era digital*. Madza Media, 2020.
 - [14] D. A. Darma and I. Wahyudin, “Analisis dan implementasi data mining untuk menentukan gaji karyawan tetap serta honorer memakai prosedur algoritma k-means clustering dan c4. 5,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 7, no. 2, pp. 280–293, 2022.
 - [15] I. Jayanto, “Analisis Perbandingan Algoritma Decision Tree untuk Prediksi Karyawan dengan Potensi Atrisi di PT. XYZ,” *J. Inform. Komputer, Bisnis dan Manaj.*, vol. 22, no. 1, pp. 49–59, 2024.
 - [16] D. A. Sahputra, M. R. R. Saelan, L. A. Utami, and W. Gata, “Penentuan Faktor Kelayakan Penerimaan Karyawan Menggunakan Algoritma Decision Tree pada Perusahaan PT. Personel Alih Daya,” *J. Sains Dan Inform.*, vol. 6, no. 2, pp. 148–156, 2020.
 - [17] F. Fariyani, U. Nurbaiti, and I. Yulianti, “Comparison of Random Forest, K-Nearest Neighbors, Decision Tree, and Neural Network for Predicting Rainfall,” *J. INFORMATICS Telecommun. Eng.*, vol. 9, no. 1, pp. 24–35, 2025.
 - [18] A. Alibasyah, A. Ajiz, G. Dwilestari, and E. Wahyudin, “Penerapan Algoritma Decision Tree dalam Penentuan Karyawan Kontrak,” *MEANS (Media Inf. Anal. dan Sist.)*, pp. 124–129, 2022.
-