

# Prediksi Pengunduran Diri Karyawan Menggunakan Metode Algoritma *Random Forest*

Bima Restu Prasetyo<sup>1</sup>, Lusy Pebi Apiliani<sup>2</sup>, Citra Nur Intan<sup>3</sup>, Kenny Jonathan<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Bisnis Digital, Institut Informatika dan Bisnis Darmajaya,  
email: <sup>1</sup>bima.23121400252@mail.darmajaya.ac.id, <sup>2</sup>lusy.2312140069@mail.darmajaya.ac.id,  
<sup>3</sup>citra.2312140055@mail.darmajaya.ac.id, <sup>4</sup>Kenny.2312140089@mail.darmajaya.ac.id



## Article History:

Received : 23-07-2025  
Revised : 12-08-2025  
Accepted : 29-11-2025  
Online : 29-11-2025



This is an open access article under the  
[CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

## ABSTRAK

Pengunduran diri karyawan merupakan permasalahan yang penting dalam manajemen sumber daya manusia karena berdampak langsung pada produktivitas dan efisiensi operasional perusahaan. Oleh karena itu, diperlukan sistem prediksi berbasis data yang mampu mengidentifikasi risiko pengunduran diri secara dini. Penelitian ini bertujuan untuk membangun model klasifikasi pengunduran diri karyawan menggunakan algoritma Random Forest dengan bantuan perangkat lunak RapidMiner. Dataset yang digunakan diperoleh dari IBM HR Analytics Employee Attrition Dataset. Proses penelitian meliputi pembersihan data, transformasi atribut, pemodelan, serta evaluasi kinerja menggunakan confusion matrix dan metrik evaluasi seperti akurasi, precision, dan recall. Hasil penelitian menunjukkan bahwa model Random Forest mampu menghasilkan *accuracy* sebesar 91,04%, *precision* kelas “Yes” sebesar 100%, dan *recall* sebesar 44,37%. Selain itu, ditemukan bahwa variabel JobLevel dan TotalWorkingYears memiliki pengaruh signifikan terhadap status attrition. Dengan demikian, model ini dapat digunakan sebagai alat bantu dalam mendeteksi risiko pengunduran diri karyawan dan merancang strategi retensi yang lebih efektif dan berbasis data.

**Kata Kunci:** Pengunduran Diri Karyawan, Random Forest, Prediksi Karyawan, Data Mining.

## ABSTRACT

Employee attrition is a critical issue in human resource management as it directly affects a company's productivity and operational efficiency. Therefore, a data-driven prediction system is needed to identify potential employee resignation risks at an early stage. This study aims to build an employee attrition classification model using the Random Forest algorithm, implemented in the RapidMiner software. The dataset used in this study is derived from the IBM HR Analytics Employee Attrition Dataset. The research process includes data cleaning, attribute transformation, model building, and performance evaluation using a confusion matrix and metrics such as accuracy, precision, and recall. The results show that the Random Forest model achieved an accuracy of 91.04%, a precision of 100% for the “Yes” class, and a recall of 44.37%. Furthermore, it was found that the variables JobLevel and TotalWorkingYears significantly influence attrition status. Therefore, this model can serve as a decision support tool in identifying employee attrition risks and designing more effective, data-driven retention strategies.

**Keywords:** Attrition, Random Forest, Employee Prediction, Data Mining.

## 1. PENDAHULUAN

Pada era persaingan global saat ini, organisasi di Indonesia dihadapkan pada tantangan yang kompleks dalam menjaga stabilitas sumber daya manusia (SDM) mereka. Tingginya tingkat pengunduran diri atau disebut juga *employee pengunduran diri* dapat mengganggu kontinuitas operasional, menimbulkan biaya tambahan dari proses rekrutmen dan pelatihan ulang, serta memperburuk budaya kerja yang ada [1]. Selain itu, tingginya tingkat pengunduran diri dapat berdampak negatif terhadap semangat kerja karyawan yang masih bertahan, menciptakan suasana kerja yang kurang kondusif dan tidak stabil. Fenomena ini menjadi semakin signifikan seiring dengan mobilitas tinggi generasi milenial dan Gen Z, di mana survei menunjukkan bahwa 35 % dari karyawan generasi milenial di Indonesia berencana mengundurkan diri dalam dua tahun ke depan di mana pergerakan karyawan semakin dinamis dan persaingan untuk merekrut talenta terbaik menjadi semakin sengit. Data ini menunjukkan bahwa generasi milenial dan Gen Z memiliki ciri khas tersendiri yang berdampak pada tingginya tingkat pengunduran diri karyawan, seperti kecenderungan memilih lingkungan kerja yang fleksibel, pentingnya budaya perusahaan, serta kebutuhan

akan peluang pengembangan karier. Secara konseptual, pengunduran diri merujuk pada proses keluarnya tenaga kerja dalam suatu organisasi, yang mencakup pengunduran diri sukarela maupun penghapusan posisi kerja seperti pensiun [2]. Dalam literatur manajemen SDM, pengunduran diri karyawan sering dikaji berdasarkan berbagai faktor, seperti faktor personal, organisasi, dan lingkungan kerja [3].

Secara umum, pengunduran diri karyawan dapat didefinisikan sebagai keputusan sukarela seorang individu untuk meninggalkan organisasi, yang dipengaruhi oleh variabel seperti kepuasan kerja, beban kerja, keseimbangan kehidupan dan pekerjaan, serta peluang karier. Sementara itu, algoritma *Random Forest* merupakan metode ensemble berbasis pohon keputusan yang menggabungkan banyak pohon lalu memilih hasil mayoritas. Metode ini unggul dalam menangani data berdimensi tinggi, mampu memitigasi risiko overfitting, serta menyediakan indikator pentingnya variabel dalam proses klasifikasi [4]. Oleh sebab itu, memahami secara mendalam faktor-faktor yang menyebabkan karyawan keluar menjadi hal yang krusial bagi pihak manajemen sumber daya manusia.

Walaupun semakin banyak perusahaan yang menyadari pentingnya memprediksi tingkat pengunduran diri karyawan, masih banyak yang mengalami kesulitan dalam mengenali faktor-faktor yang memengaruhi keputusan karyawan untuk tetap bekerja atau meninggalkan organisasi. Beragam teknik analisis data telah diterapkan, namun tidak semuanya mampu menangani kompleksitas data yang tinggi serta hubungan antar variabel yang bersifat non-linear [5]. Sejumlah penelitian terdahulu telah menelusuri penerapan model dalam memprediksi tingkat pengunduran diri karyawan, di antaranya adalah memanfaatkan teknologi dan data analitik. Dengan memanfaatkan algoritma seperti *Random Forest*, perusahaan dapat mengenali pola-pola spesifik yang memengaruhi keputusan karyawan untuk tetap bekerja atau keluar dari organisasi [6].

Algoritma *Random Forest* merupakan metode ensemble berbasis pohon keputusan yang menggabungkan banyak pohon lalu memilih hasil mayoritas. Metode ini unggul dalam menangani data berdimensi tinggi, mampu memitigasi risiko overfitting, serta menyediakan indikator pentingnya variabel dalam proses klasifikasi. Algoritma *Random Forest* telah menunjukkan akurasi dan presisi tinggi dibandingkan dengan algoritma lain seperti *Gradient Boosting* dan *Decision Tree* [3]. Pendekatan yang berlandaskan data ini membantu perusahaan membangun model prediksi yang lebih tepat dan merancang strategi retensi karyawan yang lebih efisien. Kelebihan tersebut menjadikannya pilihan yang ideal untuk menganalisis fenomena ketidakhadiran karyawan. Satu penelitian terkini yang menggunakan *Random Forest* melaporkan akurasi mencapai 93 % dan nilai AUC 0,98, melebihi algoritma lain seperti KNN dalam memprediksi atrisi karyawan berdasarkan faktor-faktor signifikan [7].

Melalui penerapan teknik analisis yang lebih maju, organisasi dapat memprediksi potensi kehilangan karyawan dan menyusun strategi retensi yang lebih terarah. Hal ini berdampak pada meningkatnya kepuasan kerja serta memperkuat posisi kompetitif organisasi [8]. Pendekatan ini diharapkan mampu memberikan kontribusi nyata dalam membangun sistem prediksi atrisi yang tidak hanya akurat, tetapi juga aplikatif dan berkelanjutan. Selain itu, pendekatan ini dapat menjadi dasar yang kuat untuk pengambilan keputusan yang lebih efektif dan berbasis data dalam pengelolaan sumber daya manusia yang lebih optimal.

## 2. METODE

Metodologi pengolahan data dilakukan dengan pendekatan data mining yang mencakup tahapan data preparation dan data analysis. Proses data preparation atau seleksi atribut dilakukan guna menentukan variabel-variabel yang paling berpengaruh terhadap prediksi pengunduran diri, berdasarkan informasi gain dan korelasi antar variabel. Proses pengolahan data, pelatihan model, dan evaluasi dilakukan menggunakan perangkat lunak RapidMiner Studio. RapidMiner dipilih karena memiliki antarmuka visual yang interaktif dan mendukung berbagai algoritma machine learning, termasuk *Random Forest*, sehingga mempermudah pengguna dalam membangun workflow analisis data tanpa harus menulis kode secara manual.

### 2.1 RapidMiner

RapidMiner merupakan perangkat lunak mandiri yang berfungsi untuk melakukan analisis data dan aktivitas data mining. Aplikasi ini dapat dengan mudah diintegrasikan dengan berbagai bahasa pemrograman. Karena dibangun menggunakan bahasa Java, RapidMiner dapat dijalankan di berbagai sistem operasi. Perangkat ini juga menyediakan antarmuka pengguna (UI) yang memudahkan perancangan alur kerja analisis data. Alur tersebut akan disimpan dalam bentuk file XML yang menggambarkan langkah-langkah analisis yang ingin dilakukan pengguna terhadap data. RapidMiner kemudian membaca file tersebut untuk mengeksekusi proses analisis secara otomatis [9].

### 2.2 Random Forest

Algoritma *Random Forest* adalah model yang membentuk sejumlah pohon keputusan independen dengan cara memilih sampel data latih secara acak menggunakan metode *bootstrap* pada setiap nodenya.

Setelah itu, proses klasifikasi dilakukan dengan menggabungkan hasil dari seluruh pohon keputusan (ensemble), dan keputusan akhir ditentukan berdasarkan hasil yang paling sering muncul atau mayoritas [7]. Berikut adalah langkah-langkah dalam menggunakan *Random Forest* memakai RapidMiner, sebagai berikut:

1. Persiapan Data dan Import dataset yang akan dianalisis.
2. Lakukan eksplorasi data awal untuk memahami struktur dan konten data, serta menangani missing values
3. Tambah Operator Nominal to Numerical di dalam proses desain untuk mengubah kategori nominal menjadi bentuk numerik, tambahkan operator dari panel operator.
4. Bagi data menjadi data training dan data testing dengan menambahkan operator split data.
5. Seret operator *Random Forest* dari panel operator.
6. Tambahkan operator Apply Model untuk menerapkan model ke data uji, lalu tambahkan Performance untuk mengukur performa model.
7. Hubungkan semua operator sesuai alur proses, pastikan semua input dan output terhubung dengan
8. benar.
9. Jalankan Proses.

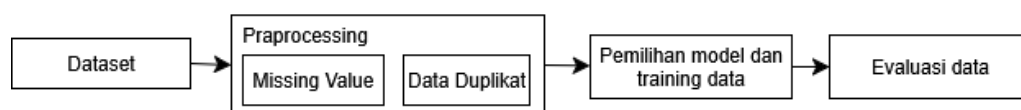
### 2.3 Data Mining

*Data mining* merupakan suatu proses yang melibatkan penggunaan satu atau lebih teknik *machine learning* untuk secara otomatis menganalisis dan mengekstraksi pengetahuan dari data. Selain itu, data mining juga dapat dipahami sebagai serangkaian aktivitas yang bertujuan mengumpulkan data dalam jumlah besar untuk menemukan informasi bermakna yang dapat dimanfaatkan [10]. Secara umum, data mining berfokus pada pencarian pola atau tren tertentu dalam database berskala besar guna mendukung proses pengambilan keputusan di masa depan. Pola-pola ini dikenali melalui perangkat atau sistem khusus yang mampu memberikan analisis data yang informatif dan bernilai wawasan, yang kemudian bisa ditelusuri lebih lanjut menggunakan alat bantu pengambilan keputusan lainnya [11].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Dataset

Proses dimulai dengan pengumpulan dataset sebagai bahan utama dalam analisis. Data yang digunakan pada penelitian ini merupakan dataset terbuka yang diperoleh dari platform Kaggle dengan judul *IBM HR Analytics Employee Pengunduran diri & Performance* yang dipublikasikan oleh Pavan Subhasht pada tahun 2020. Dataset ini terdiri dari 1.470 entri dan 35 atribut yang mencakup berbagai informasi karyawan, seperti usia, jenis kelamin, jabatan, lama bekerja, tingkat pendapatan, status promosi, kepuasan kerja, serta label *Pengunduran diri* yang menunjukkan apakah karyawan mengundurkan diri atau tidak. Dataset IBM HR telah dimanfaatkan dalam sejumlah penelitian sebelumnya, salah satunya oleh Gunawan dan timnya pada tahun 2024 [12].



**Gambar 1.** Tahap Penelitian

Dalam studi tersebut, dataset ini dimanfaatkan untuk mengkaji faktor-faktor yang berperan dalam memengaruhi tingkat *turnover* karyawan di sebuah perusahaan [13].

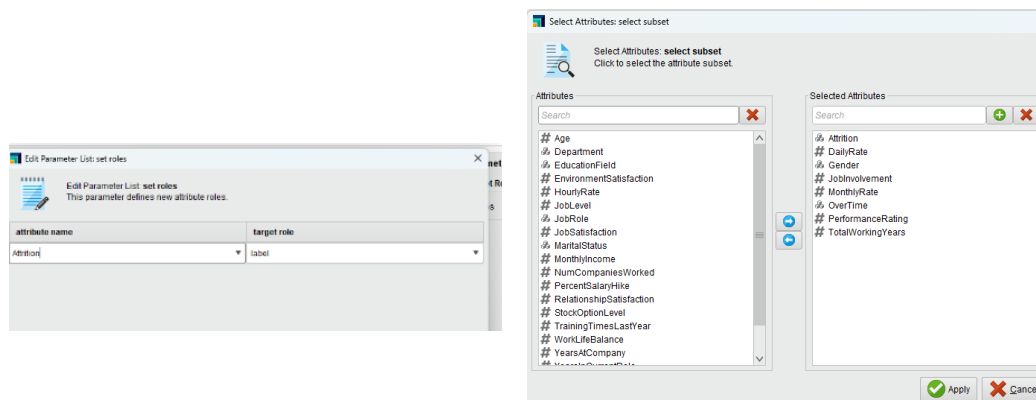
### 3.2 Praprocessing

Tahapan praproses dilakukan untuk memastikan kualitas data yang akan digunakan dalam proses pelatihan model. Tahap praproses bertujuan untuk menyiapkan data mentah menjadi data yang siap digunakan oleh algoritma *Random Forest*. Data yang memiliki nilai kosong atau tidak tersedia akan mempengaruhi hasil analisis dan akurasi model. Oleh karena itu, dilakukan penanganan missing value, melalui operator *Replace Missing Values* untuk menangani nilai yang kosong atau tidak terisi, dan juga melalui operator *Detect Outlier* untuk mendeteksi dan mengeliminasi outlier yang dapat mempengaruhi hasil model.



Gambar 2. Praprocessing Data

Pada Gambar 2. Dilakukan pemilihan atribut yang akan digunakan dalam proses *modeling* dengan menggunakan *Operator Select Attributes*. Proses ini bertujuan untuk memprediksi pengunduran diri. Atribut target Pengunduran diri diatur sebagai label menggunakan operator *Set Role* agar dapat dikenali sebagai target klasifikasi oleh algoritma pembelajaran mesin seperti Gambar 3. Setelah itu, seluruh atribut kategorikal seperti *JobRole*, *Gender*, *Department*, dan sebagainya dikonversi menjadi nilai numerik menggunakan operator *Nominal to Numerical* agar dapat diproses oleh algoritma *Random Forest* yang hanya dapat menangani input numerik.

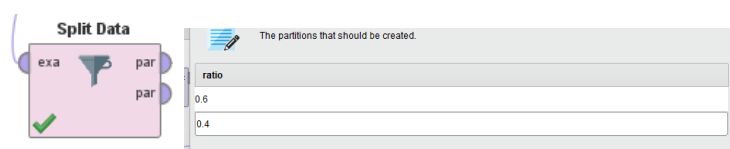


Gambar 3. Atribut yang dijadikan label dan atribut yang akan di memprediksi Pengunduran diri

Pada Gambar 2, Metode yang dipilih untuk prediksi pengunduran diri karyawan adalah dengan Pengunduran diri sebagai label serta menggunakan 7 variabel utama untuk mengetahui pengaruh variabel terhadap keputusan pengunduran diri karyawan yakni *DailyRate*, *Gender*, *JobInvolvement*, *MouthlyRate*, *OverTime*, *PerformanceRating*, dan *TotalWorkingYears*.

### 3.3 Pemilihan Model dan Training Data

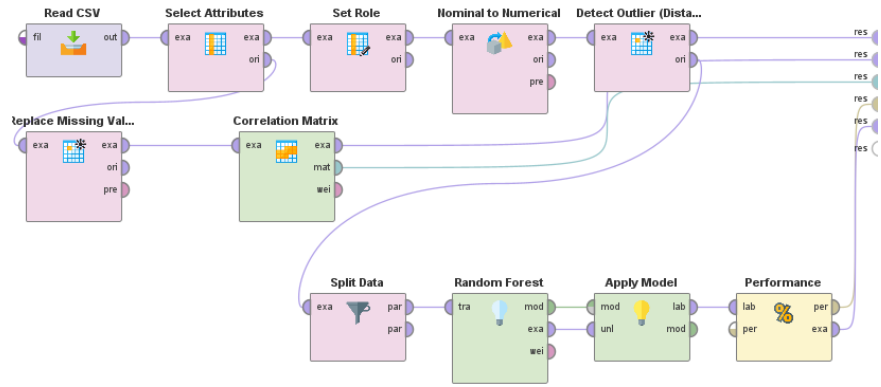
Setelah proses pembersihan data dan transformasi atribut dilakukan, tahap selanjutnya adalah mempersiapkan data untuk training data dan pengujian model. Untuk itu, data yang telah siap digunakan akan dibagi menjadi dua bagian, yaitu *data training* dan *data testing*, dengan tujuan agar model yang dibangun dapat diuji performanya terhadap data yang belum pernah dilihat sebelumnya. Hal ini dilakukan untuk memastikan bahwa model tidak hanya belajar dari data yang tersedia, tetapi juga mampu melakukan generalisasi terhadap data baru.



Gambar 4. Tahap Pemisahan data Training dan Testing

Pembagian data dilakukan dengan menggunakan operator *Split Data* pada RapidMiner. Operator ini memungkinkan pembagian data ke dalam beberapa *subset* berdasarkan rasio yang ditentukan oleh pengguna. Dalam penelitian ini, rasio yang digunakan adalah 60:40, artinya 60% dari total data digunakan untuk *training* (pelatihan model), sementara 40% sisanya digunakan untuk *testing* (pengujian performa model). Rasio ini dipilih karena dianggap cukup seimbang untuk memberikan data pelatihan yang memadai, sekaligus menyediakan data uji yang cukup untuk mengevaluasi akurasi dan kinerja model. Gambar 4 memperlihatkan konfigurasi dari operator *Split Data* yang digunakan dalam proses ini. Kolom

ratio diisi dengan dua nilai: 0.6 dan 0.4, yang secara berurutan merepresentasikan proporsi data untuk training dan testing. Hasil dari proses ini menjadi masukan utama untuk tahap berikutnya, yaitu pelatihan model *Random Forest* menggunakan *data training*, dan evaluasi model menggunakan *data testing*. Algoritma *Random Forest* dipilih karena keunggulannya dalam menangani data berdimensi tinggi, ketahanannya terhadap overfitting, dan kemampuannya menangani hubungan non-linear antar variabel.



Gambar 5. Permodelan *Random Forest*

Setelah pemilihan algoritma, selanjutnya merupakan inti dari proses data mining, yaitu proses pemodelan. Pemodelan dilakukan dengan menyusun alur kerja pada RapidMiner seperti yang ditunjukkan pada Gambar 5. Setelah model terbentuk, operator *Apply Model* digunakan untuk menerapkan model terhadap data testing, dan hasil prediksi tersebut dianalisis menggunakan operator *Performance*. Operator *Performance* akan menghitung berbagai metrik evaluasi seperti akurasi, precision, dan recall yang digunakan untuk menilai seberapa baik model mengenali pola pengunduran diri karyawan berdasarkan data yang ada.

Tabel 1. Hasil Korelasi pada Operator *Correlation Matrix*

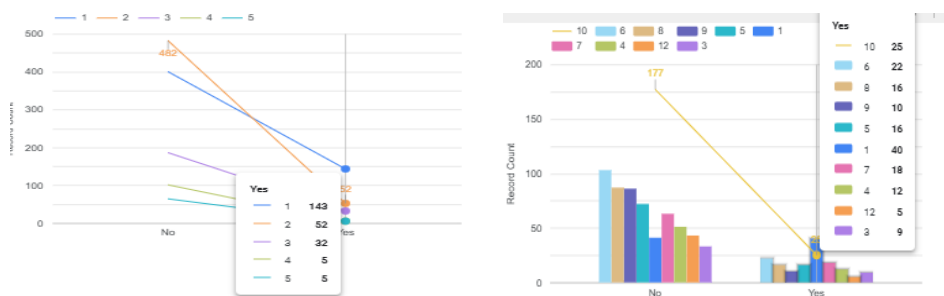
| First Attribute | Secound Attribute | Correlation |
|-----------------|-------------------|-------------|
| JobLevel        | MonthlyIncome     | 0.950       |
| JobLevel        | TotalWorkingYears | 0.782       |
| MothlyIncome    | TotalWorkingYears | 0.773       |
| Age             | TotalWorkingYears | 0.680       |

Berdasarkan analisis korelasi, ditemukan bahwa atribut *JobLevel* memiliki korelasi sangat tinggi terhadap *MonthlyIncome* (0.950), serta berkorelasi kuat terhadap *TotalWorkingYears* (0.782). Hal ini menunjukkan bahwa semakin tinggi jabatan dan pengalaman kerja seorang karyawan, maka semakin tinggi pula penghasilan yang diperoleh. Korelasi ini memperkuat temuan bahwa karyawan dengan level dan gaji yang tinggi memiliki kecenderungan untuk tidak mengundurkan diri, karena telah berada pada posisi yang stabil secara professional seperti hasil visualisasi pada Gambar 7.

|                  |          |         |                 |
|------------------|----------|---------|-----------------|
| accuracy: 91.04% |          |         |                 |
|                  | true Yes | true No | class precision |
| pred. Yes        | 63       | 0       | 100.00%         |
| pred. No         | 79       | 740     | 90.35%          |
| class recall     | 44.37%   | 100.00% |                 |

Gambar 6. Nilai Akurasi *Random Forest*.

Gambar 6. Dengan parameter konfigurasi number of trees = 100, criterion = gini\_index, dan maximal depth = 10, model *Random Forest* berhasil mencapai akurasi sebesar 91.04%. Nilai precision untuk kelas pengunduran diri (Pengunduran diri = Yes) mencapai 100%, artinya seluruh prediksi pengunduran diri benar-benar terjadi. Namun, recall-nya hanya sebesar 44.37, hasil ini merupakan peningkatan signifikan dari model sebelumnya, dan dapat dijadikan dasar untuk pengembangan model selanjutnya.



**Gambar 7.** Korelasi Pengunduran diri dengan Variabel JobLevel dan TotalWorkingYears

Berdasarkan analisis distribusi data antara variabel *JobLevel* dan Pengunduran diri, ditemukan bahwa mayoritas karyawan yang mengundurkan diri berasal dari level jabatan terendah, yaitu *JobLevel* 1 (143 orang). Jumlah tersebut menurun secara signifikan seiring dengan meningkatnya level jabatan, di mana pada *JobLevel* 5 hanya terdapat 5 kasus pengunduran diri. Hal ini menunjukkan bahwa posisi jabatan memiliki korelasi negatif terhadap kecenderungan karyawan untuk resign, di mana karyawan dengan posisi lebih rendah cenderung memiliki tingkat pengunduran diri yang lebih tinggi. Menunjukkan bahwa karyawan dengan posisi rendah dan pengalaman kerja pendek merupakan kelompok dengan risiko pengunduran diri tertinggi. Oleh karena itu, perusahaan dapat memanfaatkan informasi ini untuk menyusun strategi retensi karyawan secara lebih tepat sasaran, seperti program onboarding yang lebih intensif, pelatihan untuk pengembangan karier awal, serta peningkatan keterlibatan kerja bagi karyawan level rendah dan berpengalaman pendek.

#### 4. KESIMPULAN

Penelitian ini berhasil membuktikan bahwa penerapan algoritma *Random Forest* mampu membangun model prediksi pengunduran diri karyawan dengan performa yang baik. Melalui proses pengolahan data, pembersihan atribut, transformasi data, dan evaluasi menggunakan perangkat lunak RapidMiner, diperoleh akurasi model sebesar 91,04% dengan nilai precision pada kelas pengunduran diri mencapai 100% dan recall sebesar 44,37%. Hasil ini menunjukkan bahwa model dapat memprediksi seluruh karyawan yang diprediksi resign dengan tepat. Berdasarkan hasil korelasi antar variabel serta distribusi kelas, ditemukan bahwa atribut *JobLevel* dan *TotalWorkingYears* memiliki kontribusi signifikan dalam mempengaruhi keputusan pengunduran diri. Karyawan dengan jabatan rendah (*JobLevel* 1) dan pengalaman kerja yang pendek (1–3 tahun) menunjukkan tingkat pengunduran diri yang jauh lebih tinggi dibandingkan kelompok lainnya.

Temuan ini memperkuat bahwa faktor jabatan dan pengalaman kerja dapat menjadi indikator risiko resign, dan perlu mendapat perhatian khusus dari manajemen SDM. Penelitian ini membuka peluang lanjutan untuk mengembangkan sistem prediksi pengunduran diri yang lebih akurat dengan memperluas variabel input dan menerapkan teknik balancing data guna meningkatkan recall. Di samping itu, penerapan model ini dapat dimanfaatkan perusahaan sebagai dasar pengambilan keputusan dalam merancang strategi retensi karyawan secara lebih terukur, efektif, dan berbasis data.

#### DAFTAR PUSTAKA

- [1] M. B. Setiawan and A. Rahmatulloh, "Analisis Perbandingan Model Random Forest dan XGBoost dalam Memprediksi Turnover Karyawan," *Jurnal Sistem Informasi*, vol. 15, no. 2.
- [2] G. Dessler, *Human Resource Management*, 17th ed. Pearson Education, 2023.
- [3] A. Maehendrayuga, A. Setyanto, and Kusnawi, "Analisa Prediksi Turnover Karyawan menggunakan Machine Learning," *bit-Tech*, vol. 7, no. 2, pp. 648–659, Dec. 2024, doi: 10.32877/bt.v7i2.1999.
- [4] D. Fernando and R. G. Guntara, "Model Klasifikasi Penyebab Turnover Karyawan Menggunakan Kerangka Kerja CRISP-DM," *J-INTECH*, vol. 12, no. 02, pp. 383–392, Dec. 2024, doi: 10.32664/j-intech.v12i02.1502.
- [5] K. Arai, S. Kapoor, and R. Bhatia, Eds., *Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys) Volume 2*, vol. 869. in *Advances in Intelligent Systems and Computing*, vol. 869. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-01057-7.

- [6] B. Raharjo, "Deep Learning dengan Python," *Penerbit Yayasan Prima Agus Teknik*, pp. 1–131, Feb. 2022.
- [7] A. R. B. J. Jamroni, Wahyu Hadikristanto, and Muhamad Fatchan, "Analisis Faktor dan Prediksi Atrisi untuk Optimalisasi Retensi Karyawan Menggunakan Machine Learning," *bit-Tech*, vol. 7, no. 3, pp. 1057–1067, Apr. 2025, doi: 10.32877/bt.v7i3.2301.
- [8] N. B. Yahia, J. Hlel, and R. Colomo-Palacios, "From Big Data to Deep Data to Support People Analytics for Employee Attrition Prediction," *IEEE Access*, vol. 9, pp. 60447–60458, 2021, doi: 10.1109/ACCESS.2021.3074559.
- [9] T. P. Panjaitan *et al.*, "Klasterisasi Data Penjualan Menggunakan Algoritma K- Mean Dengan RapidMiner," 2024.
- [10] D. F. Pasaribu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *bios*, vol. 2, no. 1, pp. 11–20, Mar. 2021, doi: 10.37148/bios.v2i1.17.
- [11] G. F. Mandias, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Di Universitas Klabat Dengan Metode Klasifikasi".
- [12] C. I. Gunawan, N. D. Maharani, and D. W. Irawanto, "Modeling Affective Commitment and Job Satisfaction as A Mediating Role in Predicting Turnover Intention among Gen Z Employees," *jbti*, vol. 15, no. 2, pp. 143–160, Aug. 2024, doi: 10.18196/jbti.v15i2.20881.
- [13] Irawan *et al.*, *Pengembangan Sumber Daya Manusia Yang Berkelanjutan (Strategi Untuk Meningkatkan Keterampilan dan Kompetensi dalam Organisasi)*. Purbalingga: Eureka Media Aksara, 2021.