

**PENERAPAN ALGORITMA *RANDOM FOREST* DALAM
MENGKLASIFIKASIKAN *TURNOVER* KARYAWAN**

TUGAS AKHIR



Nhisha Aulia

1202002006

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE**

2024

**PENERAPAN ALGORITMA *RANDOM FOREST* DALAM
MENGKLASIFIKASIKAN *TURNOVER* KARYAWAN**

TUGAS AKHIR

**Diajukan sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer**



Nhisha Aulia

1202002006

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE**

2024

HALAMAN PERNYATAAN ORISINALITAS

Tugas Akhir ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : **Nhisha Aulia**

NIM : **1202002006**

Tanda Tangan :



Tanggal : **28 Agustus 2024**

HALAMAN PENGESAHAN


Tugas akhir ini diajukan oleh


Nama : Nhisha Aulia
NIM : 1202002006
Program Studi : Sistem Informasi
Fakultas : Teknik dan Ilmu Komputer
Judul Skripsi : Penerapan *Algoritma Random Forest* dalam
Mengklasifikasikan *Turnover* Karyawan

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie.

DEWAN PENGUJI

Pembimbing 1 : Zakiul Fahmi Jailani, S.Kom., M.Sc. ()

Pembimbing 2 : Brian Arnesto Sitorus, S.Kom., M.Sc. ()

Penguji 1 : Dita Nurmadewi, S.Kom, M.Kom. ()

Penguji 2 : Dr. Shidiq Al Hakim, S.T, M.Eng. ()

Ditetapkan di : Jakarta

Tanggal : 28 Agustus 2024

UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir dengan judul **“Penerapan Algoritma *Random Forest* dalam Mengklasifikasikan *Turnover* Karyawan”**. Penyusunan tugas akhir ini dilakukan untuk memenuhi syarat kelulusan guna memperoleh gelar Sarjana Komputer dari Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie.

Dalam menyusun tugas akhir ini, penulis menerima banyak bantuan, bimbingan, dan dukungan dari berbagai pihak. Dengan tulus, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Zakiul Fahmi Jailani, S.Kom., M.Sc. selaku Dosen Pembimbing 1 yang telah menyarankan dan membuat penulis mengambil topik ini. Terimakasih karena telah memberi arahan, motivasi, mendukung penulis, memberikan ilmu baru, dan meluangkan waktunya untuk membantu penulis melewati segala rintangan dalam proses menyelesaikan penyusunan tugas akhir ini.
2. Bapak Brian Arnesto Sitorus, S.Kom., M.Sc. selaku Dosen Pembimbing 2 yang juga telah sangat memberikan bantuan, saran, segala arahan serta meluangkan waktunya untuk membimbing dan mendukung penulis agar dapat menyelesaikan tugas akhir ini tepat waktu.
3. Ibu Dr. Siti Rohajawati, S.Kom, M.Kom, Ibu Dita Nurmawati, S.Kom, M.Kom serta Bapak Dr. Shidiq Al Hakim, S.T, M.Eng. selaku Dosen Penguji yang telah meluangkan waktu dan masukan kepada penulis.
4. Prof. Dr. Hoga Saragih S.T., M.T., IPM., CIRR., MIEEE., M.Th., Ph.D selaku Ketua Program Studi Sistem Informasi Universitas Bakrie.
5. Kedua orang tua yang sangat berjasa dalam hidup penulis, Mama Tini dan Papa Dedet Suparman. Terima kasih atas kepercayaan dan segala pengorbanan, cinta, do'a, dukungan baik moral maupun material yang selalu diberikan kepada penulis.

6. Seluruh Dosen Sistem Informasi yang telah membantu penulis dan telah meluangkan waktu, tenaga untuk memberikan ilmunya selama perkuliahan
7. Nadyatul Hikmah, Khairunissa Maudina, Mery Andani Hutagalung, Christophora Putri, Endah Sripawening Hasanah, Nafila Desca Fitria, Anggun Setyaningrum, Putri Khoirunisa, Dwi Itsna Fadya, Delliana Marsya, Yuni Sophia serta kakak tingkat penulis yang selalu baik membantu yaitu Ka Vyrilia Amanda, Ka Yuliza Maulina dan teman perkuliahan di prodi Sistem Informasi 2020 yang telah berjuang bersama-sama dengan penulis.
8. Pihak-pihak lain yang tidak dapat disebutkan satu persatu yang telah membantu penulis dalam memberikan dukungan dan bantuan dalam proses penyusunan tugas akhir ini.
9. Terakhir terimakasih untuk diri sendiri, karena telah bekerja dan berusaha keras untuk dapat lulus tepat waktu. Terimakasih telah berjuang dan mampu melewati proses yang tidak mudah sampai akhirnya mampu menyelesaikan tugas akhir ini.

Penulis menyadari bahwa tugas akhir ini belum sempurna dan memiliki banyak kekurangan. Oleh karena itu, adanya kritik dan saran diharapkan oleh penulis sebagai pembangun agar selanjutnya dapat memperbaiki kedepannya. Semoga tugas akhir ini dapat bermanfaat dan menjadi bahan masukan kepada para pembaca.

Jakarta, 28 Agustus 2023



Nhisha Aulia

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai sivitas akademik Universitas Bakrie, saya yang bertanda tangan di bawah ini :

Nama : Nhisha Aulia
NIM : 1202002006
Program Studi : Sistem Informasi
Fakultas : Teknik dan Ilmu Komputer
Jenis Tugas Akhir : Skripsi

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie **Hak Bebas Royalti Non eksklusif (Non-exclusive RoyaltyFree Right)** atas karya ilmiah saya yang berjudul :

“Penerapan Algoritma *Random Forest* dalam Mengklasifikasikan *Turnover Karyawan*”

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Bakrie berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 28 Agustus 2024

Yang menyatakan



(Nhisha Aulia)

**PENERAPAN ALGORITMA *RANDOM FOREST* DALAM
MENGKLASIFIKASIKAN *TURNOVER* KARYAWAN**

Nhisha Aulia

ABSTRAK

Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan *turnover* karyawan dengan menggunakan penerapan algoritma *Random Forest*. *Dataset* yang digunakan mencakup informasi mengenai tingkat kepuasan karyawan, evaluasi terakhir, jumlah proyek yang dikerjakan, rata-rata jam kerja bulanan, lama waktu bekerja di perusahaan, kecelakaan kerja, promosi dalam lima tahun terakhir, dan tingkat gaji. Dari hasil statistik deskriptif ditemukan bahwa rata-rata tingkat kepuasan karyawan adalah 61.28%, rata-rata evaluasi terakhir adalah 71.61%, dan tingkat *turnover* karyawan tercatat sebesar 23.78%. Analisis korelasi menunjukkan bahwa tingkat kepuasan karyawan, lama waktu bekerja di perusahaan, dan gaji yang rendah merupakan faktor utama yang mempengaruhi *turnover*. Evaluasi model *Random Forest* yang dibentuk menunjukkan performa yang sangat baik dalam melakukan prediksi dengan *accuracy* 99%, *recall* 99%, *precision* 96%, dan *f1-score* 98%. Penelitian ini memberikan manfaat teoritis dengan memperkaya literatur tentang faktor-faktor *turnover* karyawan dan mengembangkan metode prediksi yang baru. Selain itu, penelitian ini juga memberikan manfaat praktis dengan membantu perusahaan merancang kebijakan yang efektif untuk mengurangi *turnover* dan meningkatkan retensi karyawan melalui penggunaan teknologi canggih dalam manajemen sumber daya manusia.

Kata kunci: *Turnover* Karyawan, *Random Forest*, Klasifikasi, Evaluasi Model.

***APPLICATION OF THE RANDOM FOREST ALGORITHM IN
CLASSIFYING EMPLOYEE TURNOVER***

Nhisha Aulia

ABSTRACT

The study aims to identify and classify employee turnover using the application of the Random Forest algorithm. The dataset used includes information on employee satisfaction levels, final evaluations, the number of projects completed, average monthly working hours, length of time working at the company, workplace accidents, promotions over the past five years, and salary levels. From the descriptive statistics, it was found that the average employee satisfaction level is 61.28%, the average final evaluation is 71.61%, and the employee turnover rate is 23.78%. Correlation analysis revealed that employee satisfaction levels, length of time working at the company, and low salaries are the primary factors influencing turnover. The evaluation of the formed Random Forest model showed excellent performance in prediction with an accuracy of 99%, recall of 99%, precision of 96%, and an F1-score of 98%. This study provides theoretical benefits by enriching the literature on factors influencing employee turnover and developing new predictive methods. Additionally, it offers practical benefits by helping companies design effective policies to reduce turnover and improve employee retention through the use of advanced technology in human resource management.

Keywords: Employee Turnover, Random Forest, Classification, Model Evaluation.

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
UCAPAN TERIMA KASIH	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	vii
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	3
1.4 Batasan Masalah.....	4
1.5 Tujuan Penelitian.....	4
1.6 Manfaat Penelitian.....	5
1.7 Sistematika Penulisan.....	6
BAB II	7
TINJAUAN PUSTAKA	7
2.1 <i>Turnover Intention</i> Karyawan	7
2.1.1 Indikator <i>Turnover Intention</i>	7
2.2 <i>Machine Learning</i>	8
2.3 Klasifikasi.....	8
2.4 Algoritma <i>Random Forest</i>	9

2.5	<i>Bootstrap Aggregation</i>	11
2.6	<i>Feature Engineering</i>	11
2.7	Transformasi.....	12
2.8	Evaluasi Model.....	12
2.9	Python.....	13
2.10	Penelitian Terdahulu	15
BAB III		19
METODE PENELITIAN		19
3.1	Kerangka Penelitian	19
3.1.1	Pengumpulan dan Input Data.....	19
3.1.2	Exploratory Data Analysis (EDA)	20
3.1.3	<i>Preprocessing</i>	20
3.1.4	Pembangunan Model <i>Random Forest</i>	21
3.1.5	Evaluasi dan Validasi	21
3.1.6	Kesimpulan Hasil Model.....	22
3.2	Metode Pengumpulan Data	22
3.3	Objek Penelitian	23
3.3.1	Data Karyawan.....	23
3.3.2	Data Perusahaan	24
3.3.3	Data <i>Turnover</i>	24
3.3.4	Faktor Eksternal	24
3.4	Alokasi Waktu Penelitian	24
BAB IV		25
HASIL DAN PEMBAHASAN		25
4.1	Hasil Penelitian.....	25
4.1.1	Hasil <i>Exploratory Data Analysis</i> (EDA)	25
4.1.1.1	Deskripsi Data.....	25
4.1.1.2	Pengecekan data <i>Null</i>	26
4.1.1.3	Perhitungan Jumlah Data	27

4.1.2	<i>Preprocessing Data</i>	28
4.1.3	Hasil Pembentukan Model <i>Machine Learning</i> dengan Algoritma <i>Random Forest</i>	36
4.1.4	Hasil Evaluasi Data Baru.....	39
4.2	Pembahasan.....	40
4.2.1	Hasil <i>Exploratory Data Analysis</i> (EDA).....	40
4.2.2	<i>Preprocessing Data</i>	41
4.2.3	Hasil Pembentukan Model <i>Machine Learning</i> dengan Algoritma <i>Random Forest</i>	42
4.2.4	Evaluasi Model.....	43
BAB V	44
KESIMPULAN DAN SARAN	44
5.1	Kesimpulan.....	44
5.2	Saran.....	45
DAFTAR PUSTAKA	46
LAMPIRAN	49

DAFTAR GAMBAR

Gambar 2.1 Machine Learning.....	8
Gambar 2.2 Arsitektur <i>Random Forest</i>	10
Gambar 3.1 Kerangka Penelitian	19
Gambar 4.1 Grafik Distribusi <i>dataset</i> pada dua kelas.....	27
Gambar 4.2 Heatmap Korelasi	30

DAFTAR TABEL

Tabel 3.1 Data Karyawan	23
Tabel 4.1 Statistik Deskriptif	26
Tabel 4.2 Output kolom null	26
Tabel 4.3 Kolom <i>salary</i>	29
Tabel 4.4 Percobaan Seleksi Fitur	33
Tabel 4.5 Normalisasi data	35
Tabel 4.6 Hasil Evaluasi Model	37
Tabel 4.7 Penggabungan hasil prediksi dengan data asli	38
Tabel 4.8 Data baru	40

DAFTAR SKRIP

Skrip 4.1 <i>Script Exploratory Data Analysis</i>	25
Skrip 4.2 <i>Script Preprocessing Data</i>	28
Skrip 4.3 <i>Script Pembentukan Model Random Forest dan Evaluasi</i>	36
Skrip 4.4 <i>Script Evaluasi Data Baru</i>	39

DAFTAR LAMPIRAN

Lampiran 1 Alokasi Waktu Penelitian	49
Lampiran 2 <i>Dataset</i> Karyawan.....	50
Lampiran 3 <i>Source Code</i>	51