

**KOMPARASI KINERJA ALGORITMA *RANDOM FOREST*  
DAN *SUPPORT VECTOR MACHINE* DALAM ANALISIS  
POTENSI *TURNOVER* KARYAWAN**

Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Jurusan Teknik Informatika



Oleh :

Nauval Ahmad Zidane  
NIM : 09021382025163

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2024**

## LEMBAR PENGESAHAN SKRIPSI

### KOMPARASI KINERJA ALGORITMA *RANDOM FOREST* DAN *SUPPORT VECTOR MACHINE* DALAM ANALISIS POTENSI *TURNOVER* KARYAWAN

Oleh :

Nauval Ahmad Zidane  
NIM : 09021382025163

Indralaya, 14 Mei 2024

Pembimbing I,



Alvi Syahrini Utami, S.Si., M.Kom.  
NIP. 197812222006042003

Pembimbing II,



Junia Kurniati, S.Kom., M.Kom.  
NIP. 198906262024212001

Mengetahui,

Ketua Jurusan Teknik Informatika



Dr. M. Fachrudzi, S.Si., M.T.  
NIP. 198005222008121002

## TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jumat tanggal 3 Mei 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Nauval Ahmad Zidane  
NIM : 09021382025163  
Judul : KOMPARASI KINERJA ALGORITMA *RANDOM FOREST* DAN *SUPPORT VECTOR MACHINE* DALAM ANALISIS POTENSI *TURNOVER KARYAWAN*

Dan dinyatakan LULUS

1. Ketua Penguji

Al Farissi, S.Kom., M.Cs.  
NIP. 198512152014041001



2. Penguji

Samsuryadi, M.Kom., Ph.D.  
NIP. 197102041997021003



3. Pembimbing I

Alvi Syahrini Utami, S.Si., M.Kom.  
NIP. 197812222006042003



4. Pembimbing II

Junia Kurniati, S.Kom., M.Kom.  
NIP. 198906262024212001



Mengetahui,  
Ketua Jurusan Teknik Informatika

Dr. M. Fachrurrozi, S.Si., M.T.  
NIP. 198005222008121002

## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini :

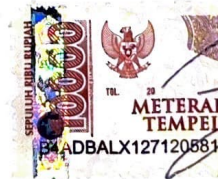
Nama : Nauval Ahmad Zidane  
NIM : 09021382025163  
Program Studi : Teknik Informatika  
Judul Skripsi : KOMPARASI KINERJA ALGORITMA *RANDOM FOREST*  
DAN *SUPPORT VECTOR MACHINE* DALAM ANALISIS  
POTENSI *TURNOVER* KARYAWAN

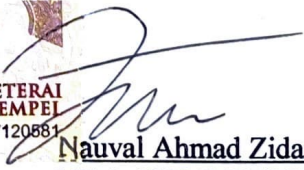
Hasil Pengecekan Software (iThenticate/Turnitin) : 17%

Menyatakan bahwa laporan projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan / plagiat. Apabila ditemukan unsur penjiplakan / plagiat dalam laporan projek ini, maka saya bersedia menerima saksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.

Palembang, 14 Mei 2024



  
Nauval Ahmad Zidane  
NIM. 09021382025163

## MOTTO DAN PERSEMBAHAN

*“Dan aku menyerahkan urusanku kepada Allah.”*

*(QS. Ghafir: 44)*

Kupersembahkan Skripsi ini kepada :

- Keluarga Besarku
- Teman Seperjuangan
- Dosen Pembimbing Akademik dan Skripsi
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

## ABSTRACT

*Employee turnover, whether voluntary or involuntary, has a negative impact on company costs and productivity which influences employee decisions to stay or move, which can be analyzed using data mining techniques. This research aims to compare the performance of the Random Forest and Support Vector Machine (SVM) algorithms in analyzing potential employee turnover to provide in-depth insights to organizations. In Random Forest, parameters in the form of the number of trees are used, and in Support Vector Machine, parameters in the form of C values are used. The research results show that the Random Forest classification method has higher performance than the Support Vector Machine (SVM) method in the dataset tested. Random Forest shows performance stability with accuracy ranging from 65.93% to 78%, as well as relatively consistent precision, recall and F1-Score values, even with variations in the number of trees. On the other hand, SVM shows an accuracy level ranging from 46.10% to 51.76%, and there are indications of overfitting.*

**Keyword :** Comparison, Performance, Random Forest, Support Vector Machine, Turnover

Indralaya, 14 Mei 2024

Pembimbing I,



Alvi Syahrini Utami, S.Si., M.Kom.  
NIP. 197812222006042003

Pembimbing II,



Junia Kurniati, S.Kom., M.Kom.  
NIP. 198906262024212001

Mengetahui,

Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.  
NIP. 198005222008121002



## ABSTRAK

*Turnover* karyawan, baik sukarela maupun tidak, memiliki dampak negatif pada biaya dan produktivitas perusahaan yang memengaruhi keputusan karyawan untuk bertahan atau berpindah, yang dapat di analisis menggunakan teknik *data mining*. Penelitian ini bertujuan membandingkan kinerja algoritma *Random Forest* dan *Support Vector Machine* (SVM) dalam analisis potensi *turnover* karyawan untuk memberikan wawasan yang mendalam kepada organisasi. Pada *Random Forest* digunakan parameter berupa jumlah pohon, dan pada *Support Vector Machine* digunakan parameter berupa nilai C. Hasil penelitian menunjukkan bahwa metode klasifikasi *Random Forest* memiliki performa yang lebih tinggi daripada metode *Support Vector Machine* (SVM) dalam dataset yang diuji. *Random Forest* menunjukkan stabilitas kinerja dengan akurasi berkisar antara 65,93% hingga 78%, serta nilai *precision*, *recall*, dan *F1-Score* yang relatif konsisten, bahkan dengan variasi jumlah pohon. Di sisi lain, SVM menunjukkan tingkat akurasi berkisar antara 46,10% hingga 51,76%, dan terdapat indikasi *overfitting*.

**Kata Kunci** : Komparasi, Kinerja, *Random Forest*, *Support Vector Machine*, *Turnover*

Indralaya, 14 Mei 2024

Pembimbing I,



Alvi Syahrini Utami, S.Si., M.Kom.  
NIP. 197812222006042003

Pembimbing II,



Junia Kurniati, S.Kom., M.Kom.  
NIP. 198906262024212001

Mengetahui,  
Ketua Jurusan Teknik Informatika



Dr. M. Fachrudrozi, S.Si., M.T.  
NIP. 198005222008121002

## KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberikan berkat, rahmat, dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan judul “Komparasi Kinerja Algoritma *Random Forest* Dan *Support Vector Machine* Dalam Analisis Potensi *Turnover* Karyawan”. Tugas akhir ini disusun sebagai syarat dalam menyelesaikan studi Strata-1 program studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Penulis ingin mengucapkan terima kasih kepada semua pihak yang telah membantu dan memberikan dukungan, bimbingan, kerjasama dan motivasi kepada penulis dalam menyelesaikan tugas akhir ini. Oleh karena itu, ungkapan rasa syukur penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Taufiq Marwa, S.E., M.Si., selaku Rektor Universitas Sriwijaya.
2. Bapak Prof. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Dr. M. Fachrurrozi, S.Si., M.T., selaku Ketua Jurusan Teknik Informatika.
4. Ibu Alvi Syahrini Utami, S.Si., M.Kom. dan Ibu Junia Kurniati, S.Kom., M.Kom., selaku dosen pembimbing 1 dan dosen pembimbing 2 yang telah memberikan masukan, arahan, motivasi, bimbingan serta waktunya dalam proses penyelesaian laporan tugas akhir ini.



5. Bapak Kanda Januar Miraswan, S.Kom., M.T., selaku dosen pembimbing akademik atas dukungannya selama perkuliahan.
6. Ayah dan Mama yaitu Yusparman dan Virni Mutiara, serta adik penulis yaitu Kemal dan Haikal yang selalu memberikan doa, dukungan, motivasi dan semangat untuk penulis.
7. Tritanti Salsabila, terima kasih karena sudah bersedia menemani dan mendukung saya hingga saat ini.
8. Rekan-rekan Infixion Teknik Informatika Angkatan 2020, terutama rekan-rekan Mbah Yono.

Penulis menyadari masih banyak kekurangan dalam menyelesaikan laporan tugas akhir ini. Penulis berharap semoga laporan tugas akhir ini dapat bermanfaat khususnya bagi penulis dan Jurusan Teknik Informatika Universitas Sriwijaya.

Palembang, 14 Mei 2024

Nauval Ahmad Zidane

## DAFTAR ISI

Halaman

LEMBAR PENGESAHAN SKRIPSI .....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI.....	iii
HALAMAN PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	x
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xv
DAFTAR ISTILAH .....	xvi
DAFTAR LAMPIRAN.....	xviii
BAB I PENDAHULUAN.....	I-1
1.1    Pendahuluan .....	I-1
1.2    Latar Belakang .....	I-1
1.3    Rumusan Masalah.....	I-3
1.4    Tujuan Penelitian .....	I-4
1.5    Manfaat Penelitian .....	I-4
1.6    Batasan Masalah.....	I-4
1.7    Sistematika Penulisan .....	I-5
1.8    Kesimpulan.....	I-6
BAB II KAJIAN LITERATUR.....	II-1
2.1    Pendahuluan .....	II-1
2.2    Landasan Teori.....	II-1
2.2.1    Kinerja.....	II-1
2.2.2 <i>Turnover</i> Karyawan .....	II-2
2.2.3    Komparasi.....	II-3
2.2.4 <i>Random Forest</i> .....	II-3
2.2.5 <i>Support Vector Machine</i> .....	II-5
2.2.6    CART (Classification and Regression Trees) .....	II-9
2.2.7    Metode Pengujian.....	II-11
2.3    Metode Pengembangan Perangkat Lunak .....	II-12
2.3.1    Fase Insepsi.....	II-13
2.3.2    Fase Elaborasi.....	II-13

2.3.3	Fase Konstruksi .....	II-13
2.3.4	Fase Transisi .....	II-14
2.4	Penelitian Lain yang Relevan.....	II-14
2.5	Kesimpulan.....	II-16
<b>BAB III METODOLOGI PENELITIAN .....</b>		<b>III-1</b>
3.1	Pendahuluan .....	III-1
3.2	Pengumpulan Data .....	III-1
3.2.1	Jenis dan Sumber Data.....	III-1
3.2.2	Metode Pengumpulan Data .....	III-3
3.3	Tahapan Penelitian.....	III-4
3.3.1	Pengumpulan Data .....	III-4
3.3.2	Kerangka Kerja Penelitian.....	III-5
3.3.3	Alat Bantu Penelitian .....	III-23
3.3.4	Mengembangkan Perangkat Lunak .....	III-24
3.3.5	Melakukan Pengujian.....	III-24
3.3.6	Menganalisis Hasil Pengujian dan Membuat Kesimpulan.....	III-25
3.3.7	Membuat Laporan Penelitian .....	III-25
3.4	Manajemen Proyek Penelitian.....	III-26
3.5	Kesimpulan.....	III-26
<b>BAB IV PENGEMBANGAN PERANGKAT LUNAK .....</b>		<b>IV-1</b>
4.1	Pendahuluan .....	IV-1
4.2	Fase Insepsi .....	IV-1
4.2.1	Pemodelan Bisnis .....	IV-1
4.2.2	Kebutuhan Sistem .....	IV-2
4.2.3	Analisis dan Desain .....	IV-3
4.2.3.1	Analisis dan Kebutuhan Perangkat Lunak.....	IV-4
4.2.3.2	Analisis Data .....	IV-4
4.2.3.3	Analisis <i>Random Forest</i> .....	IV-5
4.2.3.4	Analisis <i>Support Vector Machine (SVM)</i> .....	IV-6
4.2.3.5	Analisis Confusion Matrix.....	IV-7
4.2.3.6	Desain Perangkat Lunak .....	IV-9
4.3	Fase Elaborasi .....	IV-16
4.3.1	Pemodelan Bisnis .....	IV-17
4.3.1.1	Perancangan Data .....	IV-17
4.3.1.2	Perancangan Antar Muka .....	IV-17
4.3.2	Kebutuhan Sistem .....	IV-19
4.3.3	Diagram .....	IV-19

4.3.3.1	Diagram Aktivitas .....	IV-19
4.3.3.2	Sequence Diagram .....	IV-21
4.4	Fase Konstruksi .....	IV-25
4.4.1	Kebutuhan Sistem .....	IV-25
4.4.2	Diagram Kelas .....	IV-25
4.4.3	Implementasi .....	IV-26
4.4.3.1	Implementasi Kelas .....	IV-26
4.4.3.2	Implementasi Antarmuka .....	IV-27
4.5	Fase Transisi .....	IV-29
4.5.1	Pemodelan Bisnis .....	IV-30
4.5.2	Rencana Pengujian .....	IV-30
4.5.2.1	Rencana Pengujian <i>Use Case</i> Melakukan <i>Load Data</i> dan <i>Classification</i> .....	IV-30
4.5.2.2	Rencana Pengujian <i>Use Case View Predict</i> .....	IV-31
4.5.2	Implementasi .....	IV-31
4.5	Kesimpulan .....	IV-34
BAB V HASIL DAN PEMBAHASAN .....		V-1
5.1	Pendahuluan .....	V-1
5.2	Data Hasil Percobaan .....	V-1
5.2.1	Konfigurasi Percobaan .....	V-1
5.2.2	Data Hasil Konfigurasi .....	V-2
5.2.2.1	Hasil Pengujian Parameter Metode <i>Random Forest</i> .....	V-2
5.2.2.2	Hasil Pengujian Parameter Metode <i>Support Vector Machine</i> .....	V-3
5.3	Analisis Hasil Pengujian .....	V-4
5.3.1	Analisis Hasil Pengujian Metode <i>Random Forest</i> .....	V-4
5.3.2	Analisis Hasil Pengujian Metode <i>Support Vector Machine (SVM)</i> .....	V-6
5.4	Kesimpulan .....	V-9
BAB VI KESIMPULAN DAN SARAN .....		VI-1
6.1	Pendahuluan .....	VI-1
6.2	Kesimpulan .....	VI-1
6.3	Saran .....	2
DAFTAR PUSTAKA .....		xix
LAMPIRAN .....		xxi

## DAFTAR TABEL

Halaman

<b>Tabel II-1.</b> Confusion Matrix .....	II-11
<b>Tabel III-1.</b> Sampel Data Training.....	III-2
<b>Tabel III-2.</b> Sampel Data Testing.....	III-2
<b>Tabel III-3.</b> Sampel Data Random Forest.....	III-7
<b>Tabel III-4.</b> Mean Random Forest .....	III-7
<b>Tabel III-5.</b> Normalisasi Data Random Forest .....	III-8
<b>Tabel III-6.</b> Node 1 Gender & Industry .....	III-9
<b>Tabel III-7.</b> Node 1.1 (industry=Pharma) Proffesion & Coach .....	III-10
<b>Tabel III-8.</b> Node 1.2 (industry=Banks) Traffic & Coach .....	III-10
<b>Tabel III-9.</b> Node 1.2.1 (industry=Banks, traffic=rabrecNErab) way & head_gender .....	III-10
<b>Tabel III-10.</b> Node 1.3 (industry=Retail) Greywage & Way .....	III-11
<b>Tabel III-11.</b> Node 1.3.1 (industry=Retail, greywage=white) way & head_gender .....	III-11
<b>Tabel III-12.</b> Node 1.3.1.2 Profession & Coach .....	III-11
<b>Tabel III-13.</b> Node 1.4 (industry=Manufacture) Greywage & Profession .....	III-11
<b>Tabel III-14.</b> Node 1 (root) Gender & Stag .....	III-13
<b>Tabel III-15.</b> Node 1.1 (stag=0) Age & Way .....	III-13
<b>Tabel III-16.</b> Node 1.1.1 (stag=0, age=0) Independent & Selfcontrol .....	III-13
<b>Tabel III-17.</b> Node 1.1.1.2 (stag=0, age=0, independent=1) Way & Coach .....	III-14
<b>Tabel III-18.</b> Node 1.1.1.2.2 (stag=0, age=0, independent=1, way=bus) Industry & Head_Gender .....	III-14
<b>Tabel III-19.</b> Node 1.1.2 (stag=0, age=1) Extraversion & Selfcontrol .....	III-14
<b>Tabel III-20.</b> Node 1.2 (stag=1) Anxiety & Novator .....	III-14
<b>Tabel III-21.</b> Node 1.2.1 (stag=1, anxiety=0) Coach & Novator .....	III-15
<b>Tabel III-22.</b> Node 1.2.2 (stag=1, anxiety=1) Coach & Novator .....	III-15
<b>Tabel III-23.</b> Node 1 (root) Anxiety & Novator .....	III-17
<b>Tabel III-24.</b> Node 1.1 (anxiety=0) Greywage & Way .....	III-17
<b>Tabel III-25.</b> Node 1.1.1 (anxiety=0, way=bus) Coach & Age .....	III-17
<b>Tabel III-26.</b> Node 1.2 (anxiety=1) Extraversion & Stag .....	III-18
<b>Tabel III-27.</b> Node 1.1.2 (anxiety=0, way=car) Traffic & Age .....	III-18
<b>Tabel III-28.</b> Sampel Data <i>Support Vector Machine</i> (SVM) .....	III-20
<b>Tabel III-29.</b> Normalisasi Data .....	III-20
<b>Tabel III-30.</b> Hasil Perhitungan K .....	III-21

<b>Tabel III-31.</b> Hasil Perhitungan Matriks Hessian .....	III-21
<b>Tabel III-32.</b> Data Latih .....	III-22
<b>Tabel III-33.</b> Data Uji .....	III-22
<b>Tabel III-34.</b> Hasil Perhitungan Dot Product .....	III-22
<b>Tabel III-35.</b> Tabel Hasil Evaluasi Random Forest .....	III-24
<b>Tabel III-36.</b> Tabel Hasil Evaluasi Support Vector Machine.....	III-25
<b>Tabel IV-1.</b> Kebutuhan Fungsional .....	IV-3
<b>Tabel IV-2.</b> Kebutuhan Non-Fungsional .....	IV-3
<b>Tabel IV-3.</b> Tabel Confusion Matrix Data Training .....	IV-7
<b>Tabel IV-4.</b> Hasil Evaluasi Data Training .....	IV-7
<b>Tabel IV-5.</b> Tabel Confusion Matrix Data Testing .....	IV-8
<b>Tabel IV- 6.</b> Hasil Evaluasi Data Testing .....	IV-8
<b>Tabel IV-7.</b> Definisi Aktor.....	IV-10
<b>Tabel IV-8.</b> Definisi Use Case.....	IV-10
<b>Tabel IV-9.</b> Skenario Load Data .....	IV-11
<b>Tabel IV-10.</b> Skenario View Classification .....	IV-13
<b>Tabel IV-11.</b> Skenario View Predict .....	IV-15
<b>Tabel IV-12.</b> Implementasi Kelas .....	IV-26
<b>Tabel IV-13.</b> Rencana Pengujian Use Case Load Data dan Classification .....	IV-30
<b>Tabel IV-14.</b> Rencana Pengujian Use Case View Predict .....	IV-31
<b>Tabel IV-15.</b> Pengujian Use Case Load Data .....	IV-32
<b>Tabel IV-16.</b> Pengujian Use Case View Predict .....	IV-33
<b>Tabel V-1.</b> Hasil Perhitungan Klasifikasi Metode Random Forest .....	V-2
<b>Tabel V-2.</b> Hasil Perhitungan Klasifikasi Metode Support Vector Machine .....	V-3



## DAFTAR GAMBAR

Halaman

<b>Gambar II-1.</b> Struktur Random Forest (Kamila et al., 2023) .....	II-4
<b>Gambar II-2.</b> Garis Hyperplane (G. G. Ghiffary, 2020) .....	II-6
<b>Gambar III-1.</b> Rincian Kegiatan Penelitian .....	III-4
<b>Gambar III-2.</b> Kerangka Kerja Penelitian .....	III-5
<b>Gambar III-3.</b> Gambar Pohon 1 .....	III-12
<b>Gambar III-4.</b> Gambar Pohon 2 .....	III-16
<b>Gambar III-5.</b> Gambar Pohon 3 .....	III-19
<b>Gambar IV-1.</b> Diagram Use Case.....	IV-9
<b>Gambar IV-2.</b> Rancangan Antarmuka Halaman Load Data .....	IV-17
<b>Gambar IV-3.</b> Rancangan Antarmuka Halaman Classification.....	IV-18
<b>Gambar IV-4.</b> Rancangan Antarmuka Halaman Predict.....	IV-18
<b>Gambar IV-5.</b> Diagram Aktivitas View Predict .....	IV-20
<b>Gambar IV-6.</b> Diagram Aktivitas Load Data dan Classification.....	IV-20
<b>Gambar IV-7.</b> Diagram Sequence Load Data .....	IV-22
<b>Gambar IV-8.</b> Diagram Sequence View Predict .....	IV-23
<b>Gambar IV-9.</b> Diagram Sequence Hasil Classification.....	IV-24
<b>Gambar IV-10.</b> Diagram Kelas.....	IV-26
<b>Gambar IV-11.</b> Implementasi Antarmuka Halaman Load Data .....	IV-27
<b>Gambar IV-12.</b> Implementasi Antarmuka Halaman Classification .....	IV-28
<b>Gambar IV-13.</b> Implementasi Antarmuka Halaman Predict .....	IV-29
<b>Gambar V-1.</b> Grafik Hasil Klasifikasi Random Forest .....	V-5
<b>Gambar V-2.</b> Grafik Hasil Klasifikasi Support Vector Machine.....	V-7
<b>Gambar V-3.</b> Komparasi Hasil Klasifikasi Random Forest dan SVM.....	V-8

## DAFTAR ISTILAH

- Overfitting* : Terjadi ketika model pembelajaran mesin terlalu kompleks dan terlalu "mengingat" data pelatihan dengan sangat baik, sehingga kinerjanya menurun saat diuji pada data baru yang belum pernah dilihat sebelumnya
- Role Perception* : Cara individu atau anggota kelompok memahami dan mengartikan peran atau tanggung jawab mereka dalam suatu situasi atau konteks tertentu.
- Supervised Learning* : Jenis pembelajaran mesin di mana model belajar dari data yang sudah dilabeli, yaitu data yang telah dikaitkan dengan output yang diinginkan
- Hyperplane* : Batas atau pemisah di ruang fitur yang memisahkan kelas atau kelompok data
- Root Node* : Node pertama dalam struktur pohon keputusan
- Internal Node* : Node dalam pohon keputusan yang memiliki anak atau cabang, berisi pengujian kondisi yang menentukan cabang mana yang harus diambil
- Leaf Node* : Tempat di mana prediksi atau keputusan akhir dibuat
- Regresi : Metode statistik untuk memodelkan dan memahami hubungan antara variabel
- Bootstrap Sampling* : Teknik pengambilan sampel dengan penggantian
- Split Point* : Nilai atau titik di mana data dibagi menjadi dua atau lebih kelompok dalam konteks penggunaan seperti pohon keputusan atau algoritma pembelajaran mesin lainnya
- GINI Splitting Index* : Matrik yang digunakan dalam pohon keputusan untuk menentukan bagaimana cara membagi data menjadi subset yang lebih kecil berdasarkan fitur tertentu.
- Ensemble Learning* : Teknik pembelajaran mesin yang menggabungkan beberapa model untuk meningkatkan kinerja prediksi.
- Fluktuasi : Perubahan yang tidak teratur atau tidak terduga dalam data atau perilaku suatu sistem.
- Leaf Node* : Node terminal dalam pohon keputusan yang tidak memiliki anak (child node) dan merupakan hasil akhir dari pengambilan keputusan.
- Child Node* : Node dalam pohon keputusan yang berasal dari node induk (parent node) dan memiliki cabang-cabang untuk kondisi-kondisi spesifik.

- Pure* : Dalam konteks klasifikasi, sebuah node dianggap "murni" jika semua sampel di dalamnya termasuk dalam satu kelas atau kategori.
- Wrapper Method* : Metode dalam seleksi fitur yang menggunakan model pembelajaran untuk mengevaluasi setiap subset fitur.
- Tabular : Berkaitan dengan data yang disusun dalam bentuk tabel, dengan baris-baris mewakili contoh data dan kolom-kolom mewakili fitur-fitur atau atribut-atribut.

## DAFTAR LAMPIRAN

Halaman

<b>Lampiran 1.</b> Kode Program.....	xxi
<b>Lampiran 2.</b> User Guide Program .....	xxii
<b>Lampiran 3.</b> Rencana Kegiatan Penelitian .....	xxiv
<b>Lampiran 4.</b> Dataset Training dan Testing .....	xxvii
<b>Lampiran 5.</b> Sampel Data <i>Confusion Matrix</i> .....	xxviii

# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Pada bab ini akan dibahas latar belakang penelitian terkait dengan komparasi kinerja algoritma Random Forest dan Support Vector Machine dalam menganalisis potensi turnover karyawan. Selain itu, akan dijelaskan rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan. Bab ini juga akan memberikan gambaran umum mengenai isi bab-bab selanjutnya. Pendahuluan akan dimulai dengan menjelaskan masalah yang sedang dihadapi dan cara penyelesaiannya.

### **1.2 Latar Belakang**

Peran sumber daya manusia (SDM) menjadi semakin strategis dalam konteks bisnis yang terus berubah dan penuh tantangan. Karyawan tidak hanya pelaku operasional, melainkan juga aset berharga yang memberikan dampak signifikan pada produktivitas, inovasi, dan reputasi perusahaan. Kehadiran dan kualitas karyawan berkontribusi langsung pada keberhasilan suatu organisasi.

*Turnover* karyawan mengacu pada perputaran anggota tim dalam organisasi. Baik itu kepergian karyawan secara sukarela untuk mencari peluang baru, atau kepergian yang tidak diinginkan yang mungkin disebabkan oleh berbagai alasan seperti konflik atau ketidakcocokan dengan budaya perusahaan (Apriantini et al., 2021). *Turnover* karyawan memiliki implikasi yang luas pada berbagai aspek

bisnis. Dampak negatif termasuk biaya pelatihan dan perekrutan baru, penurunan produktivitas, hilangnya pengetahuan internal, dan bahkan kerusakan pada budaya perusahaan. Selain itu, keberlanjutan operasional dan kemampuan bersaing juga dapat terpengaruh oleh fluktuasi anggota tim yang signifikan.

Dalam menghadapi tantangan ini, organisasi membutuhkan pemahaman mendalam tentang faktor-faktor yang mempengaruhi keputusan karyawan untuk tetap bertahan atau berpindah. Namun, di tengah kompleksitas lingkungan yang beragam, identifikasi faktor-faktor ini secara manual semakin sulit. Di sinilah peran analisis data dan teknik data *mining* masuk.

Teknik analisis data *mining* memberikan kesempatan untuk menggali pola tersembunyi dalam data yang dihasilkan oleh karyawan dan lingkungan kerja. Dengan memanfaatkan algoritma klasifikasi, organisasi dapat meramalkan potensi *turnover* karyawan dengan menggabungkan faktor-faktor seperti kinerja, karakteristik demografis, interaksi sosial, dan umpan balik.

Dua algoritma yang menonjol dalam analisis klasifikasi penelitian ini adalah *Random Forest* dan *Support Vector Machine* (SVM). Kedua algoritma tersebut memiliki pendekatan yang berbeda dalam menyelesaikan masalah klasifikasi dan memiliki karakteristik yang dapat mempengaruhi kinerja mereka dalam konteks analisis potensi *turnover* karyawan.

*Random Forest* adalah algoritma yang berdasarkan konsep *ensemble learning*, yang menggabungkan hasil dari beberapa pohon keputusan yang berbeda untuk menghasilkan prediksi akhir. *Random Forest* dapat mengatasi *overfitting* dan memiliki kemampuan mengenali hubungan kompleks dalam data, membuatnya



menjadi pilihan yang kuat dalam analisis klasifikasi (Manurung et al., 2021). Sedangkan *Support Vector Machine* (SVM) adalah algoritma yang bekerja dengan mencari pemisah linier atau non-linier yang optimal antara dua kelas data. SVM memiliki kemampuan untuk mengatasi data yang kompleks dan tidak linear dengan menggunakan fungsi kernel (Abiyu, 2023).

Oleh karena itu, tujuan utama dari penelitian ini adalah untuk melakukan perbandingan kinerja antara algoritma *Random Forest* dan *Support Vector Machine* (SVM) dalam analisis potensi *turnover* karyawan. Penelitian ini akan menitikberatkan pada evaluasi akurasi prediksi, kemampuan mengatasi data yang tidak seimbang, kompleksitas model, serta interpretasi hasil dari kedua algoritma. Hasil dari penelitian ini diharapkan akan memberikan wawasan yang mendalam kepada organisasi dalam memilih algoritma yang paling optimal untuk mengidentifikasi dan mengelola potensi *turnover* karyawan, sehingga dapat memberikan manfaat yang signifikan dalam menjaga keberlanjutan dan kesuksesan bisnis.

### **1.3 Rumusan Masalah**

Rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana menentukan metode yang lebih efektif dalam analisis potensi *turnover* karyawan antara metode *Random Forest* dan *Support Vector Machine*?
2. Bagaimana penerapan dan hasil perbandingan dari akurasi hasil analisis potensi *turnover* karyawan menggunakan metode *Random Forest* dan *Support Vector Machine*?

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut :

1. Mengimplementasikan penerapan metode *Random Forest* dan *Support Vector Machine* untuk menganalisis potensi *turnover* karyawan.
2. Mengetahui hasil analisis secara komparatif oleh metode *Random Forest* dan *Support Vector Machine* untuk menganalisis potensi *turnover* karyawan.

#### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah sebagai berikut :

1. Sistem yang dihasilkan dapat berguna bagi bidang *human resources* untuk menganalisis potensi *turnover* karyawan.
2. Penelitian dapat menjadi referensi dan kontribusi bagi penelitian analisis potensi *turnover* karyawan selanjutnya.

#### **1.6 Batasan Masalah**

Berdasarkan latar belakang dan identifikasi masalah di atas dan menghindari penyimpangan dari ruang permasalahan, maka diperlukan batasan masalah yang berisi ruang lingkup penelitian. Maka dari itu, proses yang akan dibahas antara lain adalah :

1. Penelitian ini menggunakan data *Employee Turnover* yang diambil dari *Kaggle*.
2. Menggunakan kernel RBF (*Radial Basic Function*).

## **1.7 Sistematika Penulisan**

Pada Sistematika penulisan penelitian ini adalah sebagai berikut:

### **BAB I. PENDAHULUAN**

Bab satu ini membahas mengenai latar belakang, tujuan penelitian, manfaat penelitian, batasan masalah, metodologi penelitian, dan sistematika penulisan.

### **BAB II. KAJIAN LITERATUR**

Bab dua membahas dasar-dasar mengenai teori yang digunakan dalam penelitian, seperti definisi kinerja, *turnover* karyawan, komparasi, algoritma *Random Forest*, algoritma *Support Vector Machine*, dan penelitian lainnya.

### **BAB III. METODOLOGI PENELITIAN**

Bab tiga menjelaskan tahapan penelitian dengan rinci yang mengacu pada kerangka kerja. Pada akhir bab ini berisi tentang perancangan manajemen proyek pada pelaksanaan penelitian.

### **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Bab empat menganalisis dan membahas pengembangan *software* mengenai analisis kebutuhan dan desain pada *software*, hingga pembuatan dan pengujian *software*.

### **BAB V. HASIL DAN ANALISIS PENELITIAN**

Bab lima menguraikan hasil dan analisis penelitian yang menjadi dasar bagi penarikan kesimpulan.

## **BAB VI. KESIMPULAN DAN SARAN**

Bab enam berisi kesimpulan dan saran untuk penelitian selanjutnya.

### **1.8 Kesimpulan**

Pada riset ini mencakup latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan. Berdasarkan uraian di atas dapat disimpulkan bahwa dalam penelitian ini terdapat perbedaan yang signifikan antara penerapan metode *Random Forest* dan *Support Vector Machine* dalam menganalisis potensi *turnover* karyawan. Temuan ini mewakili kontribusi asli penelitian ini terhadap pemahaman dan penerapan praktis di lapangan.

## DAFTAR PUSTAKA

- Abiyyu, A. S., & Lhaksana, K. M. 2023. Perbandingan Metode Seleksi Fitur untuk Mengoptimasi Model Support Vector Machine dalam Memprediksi Turnover Pegawai. *eProceedings of Engineering*, 10(2).
- Apriantini, S. A. M., Imbayani, I. G. A., & Ribek, P. K. 2021. Pengaruh Job Stress, Kepuasan Kerja dan Komitmen Organisasi terhadap Turnover Intention Karyawan Hotel Kamandalu Ubud. *EMAS*, 2(2).
- Akbardipura, F., Manurung, D. D. E., Sandi, F., Ashfahan, H., & Prasvita, D. S. 2021. Prediksi Pengunduran Diri Karyawan Perusahaan “Y” Menggunakan Random Forest. In *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*. Vol. 2(2):202-213.
- Kustianing, U. 2021. Pengembangan Model Pembelajaran Online Menggunakan Learning Management System Bintaraloka Dalam Meningkatkan Kemampuan Berpikir Kritis Siswa SMP Negeri 3 Malang. *Jurnal Pendidik Indonesia*, 4(2):41–54.
- Nurkhaliza, A. A., & Wijayanto, A. W. 2022. Perbandingan Algoritma Klasifikasi Support Vector Machine dan Random Forest pada Prediksi Status Indeks Mitigasi dan Kesiapsiagaan Bencana (IMKB) Satuan Kerja BPS di Indonesia Tahun 2020. *Maret*, 7(1):54-59.
- Palupi, E. S. 2021. Employee Turnover Classification Using Pso-Based Naïve Bayes and Naïve Bayes Algorithm in Pt. Mastersystem Infotama. *Jurnal Riset Informatika*, 3(3):233–240.

- Trimarsiah, Y., & Arafat, M. 2017. Analisis dan Perancangan Website sebagai Sarana Informasi pada Lembaga Bahasa Kewirausahaan dan Komputer Akmi Baturaja. *Jurnal Ilmiah MATRIK*, 19(1):1–10.
- Wahyuningtias, P., Utami, H. W., Raihan, U. A., Hanifah, H. N., & Adanson, Y. N. (2022). Comparison Of Random Forest And Support Vector Machine Methods On Twitter Sentiment Analysis (Case Study: Internet Selebgram Rachel Venny Escape From Quarantine). *Jurnal Teknik Informatika (JUTIF)*, 3(1):141-145.
- Widarsono, A., & Adhi Saputra, R. 2017. Analisis Dan Perancangan Sistem Informasi Akuntansi Penerimaan Kas Ke Sekolah Dengan Menggunakan Metode System Development Life Cycle (Sdlc). *Jurnal ASET (Akuntansi Riset)*, 4(2):843.