

**Klasifikasi Genre Novel Berdasarkan Sinopsis dengan  
Metode *Random Forest***

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



Oleh:

Ahmad Munirul Ihwan  
NIM : 09021381722153

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2022**

## LEMBAR PENGESAHAN SKRIPSI

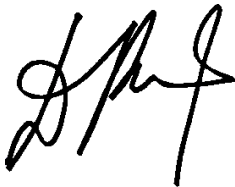
### KLASIFIKASI GENRE NOVEL BERDASARKAN SINOPSIS DENGAN METODE *RANDOM FOREST*

Oleh :

Ahmad Munirul Ihwan  
NIM: 09021381722153

Palembang, 2 Februari 2022

Pembimbing I,



Dian Palupi Rini, M.Kom., Ph.D.  
NIP. 197802232006042002

Pembimbing II,



Novi Yusliani, M.T.  
NIP. 198211082012122001

Mengetahui,  
Ketua Jurusan Teknik Informatika,



Alvi Syahrini Utami, M.Kom  
NIP. 197812222006042003

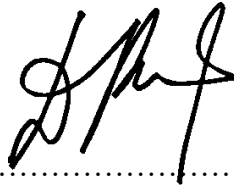
## TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari Jumat tanggal 7 Januari 2022 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Ahmad Munirul Ihwan  
NIM : 09021381722153  
Judul : Klasifikasi Genre Novel Berdasarkan Sinopsis Dengan Metode *Random Forest*

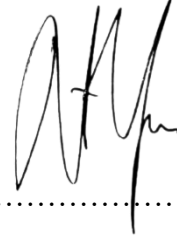
1. Pembimbing I

Dian Palupi Rini, M.Kom., Ph.D.  
NIP. 197802232006042002



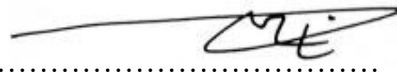
2. Pembimbing II

Novi Yusliani, M.T.  
NIP. 198211082012122001



3. Penguji I

Osvari Arsalan, S.Kom, M.T.  
NIP. 198806282018031001



4. Penguji II

Kanda Januar Miraswan, M.T.  
NIP. 199001092019031012



Mengetahui,  
Ketua Jurusan Informatika,  
  
Alvi Syahrini Utami, M.Kom.  
NIP. 197812222006042003



## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Ahmad Munirul Ihwan  
NIM : 09021381722153  
Program Studi : Teknik Informatika  
Judul Skripsi : Klasifikasi Genre Novel Berdasarkan Sinopsis Dengan  
Metode *Random Forest*

Hasil Pengecekan Software *iThenticate/Turnitin* : **9%**

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

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



Palembang, 2 Februari 2022



Ahmad Munirul Ihwan  
NIM. 09021381722153

## MOTTO DAN PERSEMBAHAN

### MOTTO:

*There is a calmness to a life lived in gratitude, a quiet  
joy.*

*[-Ralph H. Blum]*

*Within you, there is a stillness and sanctuary to which  
you can retreat at any time and be yourself.*

*[-Hermann Hesse]*

*Kupersembahkan karya tulis ini kepada:*

- ✓ *Allah SWT & Nabi Muhammad SAW*
- ✓ *Ayah & Ibuku tercinta*
- ✓ *Keluarga besarku*
- ✓ *Dosen pembimbing & penguji*
- ✓ *Teman-temanku*
- ✓ *Universitas Sriwijaya*

## **KATA PENGANTAR**

Puji syukur kepada Allah atas berkat dan rahmat-Nya yang telah diberikan kepada Penulis sehingga dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun secara tidak langsung. Penulis ingin menyampaikan rasa terima kasih kepada:

1. Kedua orang tuaku mama dan papa yang selalu mendoakan serta memberikan dukungan baik moril maupun materil.
2. Bapak Jaidan Jauhari, S.Pd., M.T, selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Ibu Alvi Syahrini, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya
4. Ibu Dian Palupi Rini, M.Kom., Ph.D. selaku dosen pembimbing I dan Ibu Novi Yusliani, M.T. selaku pembimbing II yang telah membimbing dalam proses perkuliahan dan pengerjaan Tugas Akhir.
5. Bapak Osvari Arsalan, S.Kom, M.T. selaku dosen penguji I dan Bapak Kanda Januar Miraswan, M.T. selaku dosen penguji II yang telah memberikan masukan dan ilmu pengetahuan.

6. Bapak Osvari Arsalan, S.Kom, M.T. selaku dosen pembimbing akademik yang telah memberikan bimbingan, arahan dan motivasi dalam proses perkuliahan dan pengerjaan Tugas Akhir.
7. Seluruh dosen Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
8. Mba Wiwin Juliani selaku admin Teknik Informatika beserta para staf Fakultas Ilmu Komputer Universitas Sriwijaya yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan.
9. Muhammad Surya Netrane, Gomgom Oktorio Siregar, Muhammad Fadel Zahran Zindabad, dan Edo Pratama sahabat seperjuangan yang telah menjadi saksi kehidupan perkuliahan penulis yang telah berbagi canda tawa dan yang selalu sabar dalam mengajarkan, mendengarkan, memberi saran dan selalu bersedia membantu penulis sejak awal perkuliahan sampai selesai.
10. Haidar Ali, Egi Kurniadi, Muhammad Hafizh Sytar, Yasmin Azzahrah Lubis teman baik penulis yang telah menemani masa perkuliahan penulis dan memberi motivasi untuk penulis.
11. Teman-teman kelas dan jurusan Teknik Informatika yang telah berbagi keluh kesah, motivasi, semangat, dan canda tawa selama masa perkuliahan.
12. Muhammad Jodi Ramadhan, Bentar Satria Ganesyah, Rizky Ramadhan, Pugano Rizky Lintang, Mega Pertiwi, Siti Amirah Saskia, selaku teman

seperjuangan semasa perkuliahan yang telah banyak memberikan bantuan, dukungan serta semangat.

13. Fitri Amelia yang selalu menjadi penyemangat dan menjadi motivasi penulis untuk menyelesaikan kuliah, yang selalu mendengarkan kisah keluh kesah perkuliahan maupun kehidupan pribadi penulis.
14. Semua pihak lainnya yang tidak dapat disebutkan satu-persatu yang telah memberikan doa dan dukungan, serta banyak berperan bagi penulis terutama dalam penyelesaian tugas akhir ini.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya. Akhir kata semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Palembang, 2 Februari 2022



Ahmad Munirul Ihwan

NIM. 09021381722153



# NOVEL GENRE CLASSIFICATION BASED ON SYNOPSIS WITH RANDOM FOREST METHOD

By:  
Ahmad Munirul Ihwan  
09021381722153

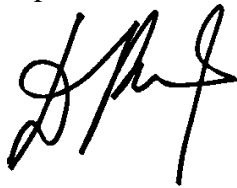
## ABSTRACT

This research is a system that can classify a text in the form of a synopsis of a novel into its genre group. Genre is the story style in a novel, there are many genres in the novel including the genres of history, horror, mystery, psychology, romance and others. Before the synopsis classification is carried out, it will go through a preprocessed text process and then take the weight value using the TF-IDF method, after that it will be classified using the Random Forest Classifier method. The results of the classification using Random Forest 50 trees have the greatest accuracy, namely 60.8% with precision, recall, f-1, history 66.1%, 78%, 71.56%, horror 46.15%, 48%, 47.06%, mystery 57.5%, 46%, 51.11%, psychology 66.04%, 70%, 67.96%, romance 67.39%, 62%, 64.58%.

**Keywords:** Classification, Genre, Random Forest Classifier, TF-IDF

Palembang, February 2<sup>nd</sup> 2022

Supervisor I,



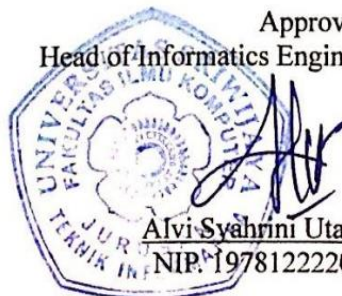
Dian Palupi Rini, M.Kom., Ph.D.  
NIP. 197802232006042002

Supervisor II,



Novi Yusliani, M.T.  
NIP. 198211082012122001

Approve,  
Head of Informatics Engineering Departement,



Alvi Syahrini Utami, M.Kom  
NIP. 197812222006042003

# KLASIFIKASI GENRE NOVEL BERDASARKAN SINOPSIS DENGAN METODE *RANDOM FOREST*

Oleh:  
Ahmad Munirul Ihwan  
09021381722153

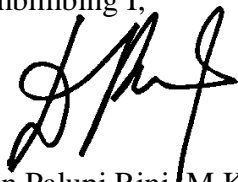
## ABSTRAK

Penelitian ini yaitu sebuah sistem yang dapat mengklasifikasikan teks berupa sinopsis novel kedalam kelompok genrenya. Genre adalah gaya cerita dalam sebuah novel, terdapat banyak genre pada novel diantaranya genre *history*, *horror*, *mystery*, *psychology*, *romance* dan lain-lain. Sebelum dilakukan klasifikasi sinopsis akan melalui proses teks-praproses kemudian mengambil nilai bobotnya dengan metode *TF-IDF* setelah itu baru di klasifikasi dengan metode *Random Forest Classifier*. Hasil dari klasifikasi dengan menggunakan *Random Forest* 50 pohon memiliki akurasi terbesar yaitu 60,8% dengan *precision*, *recall*, *f-1*, *history* 66,1%, 78%, 71,56%, *horror* 46,15%, 48%, 47,06%, *mystery* 57,5%, 46%, 51,11%, *psychology* 66,04%, 70%, 67,96%, *romance* 67,39%, 62%, 64,58%.

**Kata Kunci:** Genre, Klasifikasi, *Random Forest*, *TF-IDF*

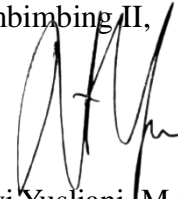
Palembang, 2 Februari 2022

Pembimbing I,



Dian Palupi Rini, M.Kom., Ph.D.  
NIP. 197802232006042002

Pembimbing II,



Novi Yusliani, M.T.  
NIP. 198211082012122001

Mengetahui,  
Ketua Jurusan Teknik Informatika,



Alvi Syahrini Utami, M.Kom  
NIP. 197812222006042003

## DAFTAR ISI

HALAMAN JUDUL .....	i
LEMBAR PENGESAHAN SKRIPSI .....	ii
TANDA LULUS UJIAN SIDANG TUGAS AKHIR.....	iii
HALAMAN PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
ABSTRACT .....	ix
ABSTRAK .....	x
DAFTAR ISI .....	x
DAFTAR GAMBAR .....	xiv
DAFTAR TABEL .....	xv
BAB 1 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-3
1.5    Manfaat Penelitian .....	I-3
1.6    Batasan Masalah .....	I-4
1.7    Sistematika Penulisan .....	I-4
1.8    Kesimpulan.....	I-5
BAB II KAJIAN LITERATUR .....	<b>II-Error! Bookmark not defined.</b>
2.1    Pendahuluan .....	<b>II-Error! Bookmark not defined.</b>
2.2    Landasan Teori .....	<b>II-Error! Bookmark not defined.</b>
2.2.1 <i>Machine Learning</i> .....	<b>II-Error! Bookmark not defined.</b>
2.2.2 <i>Text Pre-processing</i> .....	<b>II-Error! Bookmark not defined.</b>
2.2.3    Pembobotan TF-IDF.....	<b>II-Error! Bookmark not defined.</b>
2.2.4 <i>Decision Tree</i> .....	<b>II-Error! Bookmark not defined.</b>
2.2.5 <i>Random Forest</i> .....	<b>II-Error! Bookmark not defined.</b>
2.2.6 <i>Confusion Matrix</i> .....	<b>II-Error! Bookmark not defined.</b>
2.3    Penelitian Lain yang Relevan.....	<b>II-Error! Bookmark not defined.</b>

2.4	Kesimpulan.....	II-Error! Bookmark not defined.
<b>BAB III METODOLOGI PENELITIAN..... III-Error! Bookmark not defined.</b>		
3.1	Pendahuluan .....	III-Error! Bookmark not defined.
3.2	Pengumpulan Data.....	III-Error! Bookmark not defined.
3.3	Tahapan Penelitian.....	III-Error! Bookmark not defined.
3.3.1	Kerangka Kerja .....	III-Error! Bookmark not defined.
3.3.2	Kriteria Pengujian .....	III-Error! Bookmark not defined.
3.3.3	Format Data Pengujian.....	III-Error! Bookmark not defined.
3.3.4	Alat yang Digunakan dalam Pelaksanaan Penelitian .....	III-Error! Bookmark not defined.
3.3.5	Pengujian Penelitian.....	III-Error! Bookmark not defined.
3.3.6	Analisis Hasil Pengujian dan Membuat Kesimpulan.....	III-Error! Bookmark not defined.
3.4	Metode Pengembangan Perangkat Lunak .....	III-Error! Bookmark not defined.
<b>BAB IV PENGEMBANGAN PERANGKAT LUNAK . IV-Error! Bookmark not defined.</b>		
4.1	Pendahuluan .....	IV-Error! Bookmark not defined.
4.2	Fase Insepsi .....	IV-Error! Bookmark not defined.
4.2.1	Pemodelan Bisnis .....	IV-Error! Bookmark not defined.
4.2.2	Kebutuhan Sistem .....	IV-Error! Bookmark not defined.
4.2.3	Analisis Kebutuhan dan Desain .....	IV-Error! Bookmark not defined.
4.3	Fase Elaborasi.....	IV-Error! Bookmark not defined.
4.3.1	Pemodelan Bisnis .....	IV-Error! Bookmark not defined.
4.4	Fase Konstruksi .....	IV-Error! Bookmark not defined.
4.4.1	Kebutuhan Sistem .....	IV-Error! Bookmark not defined.
4.4.2	<i>Diagram Class</i> .....	IV-Error! Bookmark not defined.
4.4.3	Implementasi.....	IV-Error! Bookmark not defined.
4.4.4	Implementasi Kelas .....	IV-Error! Bookmark not defined.
4.4.5	Implementasi Antarmuka.....	IV-Error! Bookmark not defined.
4.5	Fase Transisi.....	IV-Error! Bookmark not defined.
4.5.1	Pemodelan Bisnis .....	IV-Error! Bookmark not defined.
4.5.2	Rencana Pengujian .....	IV-Error! Bookmark not defined.
4.5.3	Rencana Pengujian Klasifikasi Genre Novel Berdasarkan Sinopsis	

dengan metode <i>Random Forest</i> .....	IV-Error! Bookmark not defined.
4.5.4 Implementasi.....	IV-Error! Bookmark not defined.
4.6 Kesimpulan.....	IV-Error! Bookmark not defined.
<b>BAB V HASIL DAN ANALISIS PENELITIAN.....</b>	<b>V-Error! Bookmark not defined.</b>
5.1 Pendahuluan .....	V-Error! Bookmark not defined.
5.2 Hasil Percobaan Penelitian.....	V-Error! Bookmark not defined.
5.2.1 Hasil <i>Confusion Matrix</i> dari Klasifikasi <i>Random Forest</i> .....	V-Error! Bookmark not defined.
5.3 Kesimpulan.....	V-Error! Bookmark not defined.
<b>BAB VI KESIMPULAN DAN SARAN .....</b>	<b>VI-Error! Bookmark not defined.</b>
6.1 Kesimpulan.....	VI-Error! Bookmark not defined.
Kesimpulan yang didapat dari penelitian ini ialah :....	VI-Error! Bookmark not defined.
6.2 Saran .....	VI-Error! Bookmark not defined.
<b>DAFTAR PUSTAKA.....</b>	<b>vi</b>

## DAFTAR GAMBAR

- Gambar II-1. *Decision Tree* ..... **II-Error! Bookmark not defined.**
- Gambar II-2. Proses *Random Forest* (<https://www.researchgate.net>).... **II-Error! Bookmark not defined.**
- Gambar III-1. Kerangka Kerja..... **III-Error! Bookmark not defined.**
- Gambar IV-5. Contoh *DecisionTree*..... **IV-Error! Bookmark not defined.**
- Gambar IV-6. Contoh *Random Forest* ..... **IV-Error! Bookmark not defined.**
- Gambar IV-8. Diagram Aktivitas Membuat File .arff dan pembobotan Sinopsis ..... **IV-Error! Bookmark not defined.**
- Gambar IV-9. Diagram Aktivitas Klasifikasi *Random Forest*..... **IV-Error! Bookmark not defined.**
- Gambar IV-10. Gambar Rancangan Antar Muka ..... **IV-Error! Bookmark not defined.**
- Gambar IV-11. Gambar *Diagram Sequence* Memuat Data Sinopsis.... **IV-Error! Bookmark not defined.**
- Gambar IV-12. Gambar *Diagram Sequence* Proses dan Membuat File **IV-Error! Bookmark not defined.**
- Gambar IV-13. Gambar *Diagram Sequence* Memuat Data Bobot ..... **IV-Error! Bookmark not defined.**
- Gambar IV-14. Gambar *Diagram Sequence* Klasifikasi Random Forest ..... **IV-Error! Bookmark not defined.**
- Gambar IV-15. Gambar *Diagram Class*..... **IV-Error! Bookmark not defined.**
- Gambar IV-16. Gambar Implementasi Antarmuka ..... **IV-Error! Bookmark not defined.**

## DAFTAR TABEL

- Tabel II-1. Tabel *Confusion Matrix*..... II-Error! **Bookmark not defined.**
- Tabel II-2. Tabel *Confusion Matrix Multilabel* Kelas Merah..... II-Error! **Bookmark not defined.**
- Tabel III-1. Tabel Format Data Pengujian .... III-Error! **Bookmark not defined.**
- Tabel III-2. Tabel Analisis Hasil Pengujian.. III-Error! **Bookmark not defined.**
- Tabel IV-1. Tabel Kebutuhan Fungsional..... IV-Error! **Bookmark not defined.**
- Tabel IV-2. Tabel Kebutuhan non-Fungsional..... IV-Error! **Bookmark not defined.**
- Tabel IV-3. Tabel Bobot tfidf..... IV-Error! **Bookmark not defined.**
- Tabel IV-4. Tabel Contoh Perhitungan *Entropy* dan *Gain*..IV-Error! **Bookmark not defined.**
- Tabel IV-5. Tabel Definisi Aktor..... IV-Error! **Bookmark not defined.**
- Tabel IV-6. Definisi Use Case ..... IV-Error! **Bookmark not defined.**
- Tabel IV-7. Skenario Memuat Data File Sinopsis ..... IV-Error! **Bookmark not defined.**
- Tabel IV-8. Skenario Membuat File Berformat .arff... IV-Error! **Bookmark not defined.**
- Tabel IV-9. Skenario Memuat File Berformat .arff..... IV-Error! **Bookmark not defined.**
- Tabel IV-10. Skenario Membuat File Berformat .arff. IV-Error! **Bookmark not defined.**
- Tabel IV-11. Implementasi Kelas..... IV-Error! **Bookmark not defined.**
- Tabel IV-12. Rencana Pengujian Klasifikasi Genre Novel Berdasarkan Sinopsis dengan Metode *Random Forest* ..... IV-Error! **Bookmark not defined.**
- Tabel IV-13. Pengujian Klasifikasi Genre Novel Berdasarkan Sinopsis dengan Metode *Random Forest* ..... IV-Error! **Bookmark not defined.**

Tabel V-1. Hasil Klasifikasi *Random Forest* dengan 25 Pohon ..... V-Error!  
**Bookmark not defined.**

Tabel V-2. *Confusion Matrix* dengan 25 Pohon.....V-Error! **Bookmark not defined.**

Tabel V-3. Hasil Klasifikasi *Random Forest* dengan 50 Pohon ..... V-Error!  
**Bookmark not defined.**

Tabel V-4. *Confusion Matrix* dengan 50 Pohon.....V-Error! **Bookmark not defined.**

Tabel V-5. Hasil Klasifikasi *Random Forest* dengan 75 Pohon ..... V-Error!  
**Bookmark not defined.**

Tabel V-6. *Confusion Matrix* dengan 75 Pohon.....V-Error! **Bookmark not defined.**

Tabel V-7. Hasil Klasifikasi *Random Forest* dengan 100 Pohon ..... V-Error!  
**Bookmark not defined.**

Tabel V-8. *Confusion Matrix* dengan 100 Pohon.....V-Error! **Bookmark not defined.**



# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Dalam bab pendahuluan ini, diuraikan pokok-pokok pikiran berupa latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan yang diuraikan secara singkat bab perbab untuk membuat tugas akhir klasifikasi genre berdasarkan sinopsis novel dengan metode *random forest*.

### **1.2 Latar Belakang**

Novel merupakan hasil pikiran pengarang yang sengaja dirancang untuk menyatakan buah pikiran atau ide. Kemudian diolah penulis dengan menghubungkan kejadian atau peristiwa disekelilingnya, bisa juga merupakan pengalaman orang lain maupun pengalaman penulis. Pola penulisan mengalir secara bebas yang tidak terikat oleh kaidah seperti yang terdapat pada puisi (Yanti, 2015). Novel memiliki beberapa genre antara lain genre romantis, horor, misteri, fiksi-non fiksi serta masih banyak lagi. Novel yang sudah terbit di pasaran pun sudah lebih dari ribuan, namun untuk mengklasifikasi novel-novel tersebut sesuai genre masih dilakukan secara manual. Oleh sebab itu dibutuhkan suatu sistem yang dapat mengklasifikasi novel-novel tersebut secara otomatis. Hal ini bertujuan agar sistem ini nantinya dapat membantu pembaca, penulis, ataupun

penerbit untuk mengetahui arah genre yang akan ditulis maupun dibaca (Sari, 2019).

Menurut Arshad dalam (Fitrianto, 2019), Sinopsis adalah ringkasan cerita yang mengutamakan alur atau plot yang tepat dan menarik dari suatu episode tertentu, cerpen, novel, atau drama. Sinopsis bisa juga diartikan sebagai ikhtisar karangan yang biasanya diterbitkan bersama-sama dengan karangan aslinya sehingga membuat pembaca tertarik untuk membaca isi utuh karangan tersebut.

Menurut Larose dalam (Fitrianto, 2019), Klasifikasi merupakan proses menempatkan objek atau konsep tertentu kedalam satu set kategori. Berdasarkan sifat objek atau konsep yang bersangkutan. Metode klasifikasi ditujukan untuk pembelajaran fungsi-fungsi berbeda yang memetakan masing-masing data terpilih kedalam salah satu dari kelompok kelas yang telah ditentukan sebelumnya. Dinamakan *Random Forest* karena merupakan keturunan dari pendekatan ID3 untuk membangun pohon keputusan. *Random Forest* merupakan algoritma yang sesuai digunakan untuk masalah klasifikasi pada *machine learning* dan data mining.

Adapun penelitian tentang perbandingan metode klasifikasi yang dilakukan (Sodik et al., 2020), berdasarkan pembahasan, maka diperoleh metode terbaik untuk klasifikasi data bank customers adalah metode klasifikasi *Random Forest*. Hal ini dikarenakan karena ketepatan klasifikasinya yang paling tinggi daripada yang lain dengan nilai akurasi 86.2%, klasifikasi dengan metode Regresi Logistik 82%, klasifikasi dengan metode *K-Nearest Neighbor* 83,9%, klasifikasi dengan metode *Super Vector Machine* 80,1%, klasifikasi dengan metode *Naive bayes*

83,6%, klasifikasi dengan metode *Decision Tree* 79,1%. Berdasarkan pada penjelasan yang telah diuraikan maka pada penelitian ini. Peneliti akan meneliti tentang klasifikasi genre novel berdasarkan sinopsis dengan metode *Random Forest*.

### **1.3 Rumusan Masalah**

Rumusan masalah penelitian ini melakukan klasifikasi genre berdasarkan sinopsis dengan metode perhitungan *TF-IDF* dan metode *Random Forest*, berdasarkan masalah tersebut maka terdapat beberapa pertanyaan:

1. Bagaimana implementasi metode *TF-IDF* dan metode *Random Forest* untuk klasifikasi genre berdasarkan sinopsis novel?
2. Bagaimana hasil akurasi dari metode *TF-IDF* dan *Random Forest* pada klasifikasi genre berdasarkan sinopsis novel?

### **1.4 Tujuan Penelitian**

Tujuan penelitian ini adalah sebagai berikut:

1. Melakukan klasifikasi genre berdasarkan sinopsis novel dengan metode *TF-IDF* dan *Random Forest*.
2. Menganalisis hasil akurasi yang didapatkan dalam melakukan klasifikasi genre berdasarkan sinopsis novel dengan metode *TF-IDF* dan *Random Forest*.

### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah sebagai berikut:

1. Hasil klasifikasi penelitian ini dapat dijadikan sebagai bahan kajian dan referensi dalam menggunakan metode *TF-IDF* dan *Random Forest*.

2. Hasil akurasi dan prediksi dari klasifikasi penelitian ini dapat dijadikan pembandingan dengan penelitian menggunakan metode lain.

### **1.6 Batasan Masalah**

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. Penelitian ini menggunakan *dataset* yang ada dalam *website* <https://www.goodreads.com>. Dengan total data sebanyak 750 data.
2. Data latih yang digunakan sebanyak 500 data yang terbagi 100 per genre dari 5 genre yaitu *History, Horror, Mystery, Psychology, Romance*.
3. *Data uji yang digunakan sebanyak 250 data yang terbagi 50 per kategori genre.*

### **1.7 Sistematika Penulisan**

Sistematika penulisan tugas akhir ini adalah sebagai berikut:

#### **BAB I. PENDAHULUAN**

Pada bab ini diuraikan mengenai latar belakang, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan dalam penelitian klasifikasi genre berdasarkan sinopsis novel dengan metode *random forest*.

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

Pada bab ini akan membahas dasar-dasar teori yang digunakan dalam penelitian serta beberapa kajian literatur mengenai penelitian lain yang relevan pada penelitian ini.

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

Pada bab ini akan dibahas mengenai tahapan yang akan dilaksanakan pada penelitian ini. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci dengan mengacu pada suatu kerangka kerja. Di akhir bab ini berisi perancangan manajemen proyek pada pelaksanaan penelitian.

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

Pada bab ini akan dibahas mengenai rancangan dan lingkungan implementasi, implementasi program hasil text praprocess, pembobotan kata dan klasifikasi, hasil eksekusi serta hasil pengujian.

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

Pada bab ini, hasil pengujian berdasarkan langkah-langkah yang telah direncanakan disajikan. Analisis diberikan sebagai basis dari kesimpulan yang diambil dalam penelitian.

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

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab-bab sebelumnya dan juga berisi saran yang diharapkan dapat berguna dalam penerapan penelitian selanjutnya.

#### **1.8 Kesimpulan**

Di bab ini telah dijelaskan Pendahuluan, Latar Belakang, Rumusan Masalah, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penulisan untuk membuat penelitian mengklasifikasikan genre novel berdasarkan sinopsis dengan metode

*random*

*forest.*

## DAFTAR PUSTAKA

- Ahmad, A. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Jurnal Teknologi Indonesia*, 1(1), 1–5.
- Fitri, M. (2013). Perancangan Sistem Temu Balik Informasi Dengan Metode Pembobotan Kombinasi Tf-Idf Untuk Pencarian Dokumen Berbahasa Indonesia. *Jurnal Sistem Dan Teknologi Informasi*, Vol. 1(1), 1–6.
- Fitrianto, A. (2019). Journal of Social Science Teaching, Vol. 3 No. 1 Tahun 2019 Program Studi Tadris Ilmu Pengetahuan Sosial Institut Agama Islam Negeri Kudus. *Journal of Social Science Teaching*, 3(1), 86–95.
- Indriani, A. (2014). Klasifikasi Data Forum dengan menggunakan Metode Naïve Bayes Classifier. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI) Yogyakarta*, 1(1), 21–2014. [www.bluefame.com](http://www.bluefame.com),
- Jaka, A. T. (2015). Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining. *Informatika UPGRIS*, 1, 1–9.
- Rahman, M. F., Alamsah, D., Darmawidjadja, M. I., & Nurma, I. (2017). Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN). *Jurnal Informatika*, 11(1), 36. <https://doi.org/10.26555/jifo.v11i1.a5452>
- Rahmayanti, V., Basuki, S., & Hilman, H. (2019). Klasifikasi sinopsis novel menggunakan metode naïve bayes classifier. *Jurnal Repositor*, 1(2), 125. <https://doi.org/10.22219/repositor.v1i2.799>
- Sari, Y. A. (2019). Prediksi Rating Novel Baru Berdasarkan Sinopsis Menggunakan Genre Based Collaborative Filtering dan Text Similarity. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer Vol.*, 3(March), 2768–2773.
- Setiawati, D., Taufik, I., Jumadi, J., & Zulfikar, W. B. (2016). Klasifikasi Terjemahan Ayat Al-Quran Tentang Ilmu Sains Menggunakan Algoritma Decision Tree Berbasis Mobile. *Jurnal Online Informatika*, 1(1), 24. <https://doi.org/10.15575/join.v1i1.7>
- Sodik, F., Dwi, B., & Kharisudin, I. (2020). Perbandingan Metode Klasifikasi

Supervised Learning pada Data Bank Customers Menggunakan Python.  
*Jurnal Matematika*, 3, 689–694.

Wuisan, D. S. . (2020). Journal of technology information. *Journal Of Technology Information*, 6(1), 29–34.  
<https://ojs.uajy.ac.id/index.php/IJIS/article/view/1704/1195>

Yanti, C. S. (2015). Religiositas Islam Dalam Ratu Yang Bersujud Karya Amrizal Mochamad Mahdavi. *Religiositas Islam Dalam Novel*, 3(15), 1–15.