

**PERBANDINGAN METODE RANDOM FOREST DAN  
METODE K-NEAREST NEIGHBOR (KNN) PADA  
KLASIFIKASI PENDERITA PENYAKIT PARKINSON**

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



Oleh:

Zainudin

NIM : 09021381621079

**Jurusan Teknik Informatika**

**FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA**

2021

## LEMBAR PENGESAHAN TUGAS AKHIR

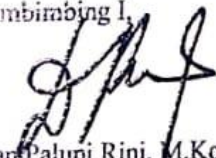
### PERBANDINGAN METODE RANDOM FOREST DAN METODE K-NEAREST NEIGHBOR (KNN) PADA KLASIFIKASI PENDERITA PENYAKIT PARKINSON

Oleh:

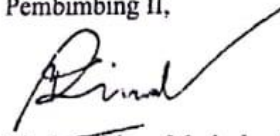
ZAINUDIN  
NIM : 09021381621079

Palembang, 19 April 2021

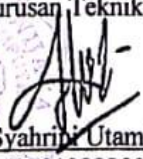
Pembimbing I,

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

Pembimbing II,

  
Mastura Diana Marieska, S.T, M.T.  
NIP. 198603212018032001

Mengetahui,  
Ketua Jurusan Teknik Informatika

  
Alvi Syahri Utami, M.Kom  
NIP. 197812222006042003



## TANDA LULUS SIDANG SKRIFSI

Pada hari Senin, tanggal 19 April 2021 telah dilaksanakan ujian sidang skripsi oleh jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Zainudin

NIM : 09021381621079

Judul : Perbandingan Metode *Random Forest* dan *K-Nearest Neighbor* (KNN) Pada Klasifikasi Penderita Penyakit Parkinson

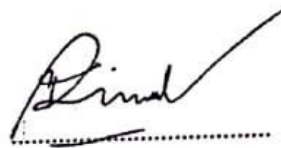
1. Pembimbing I

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



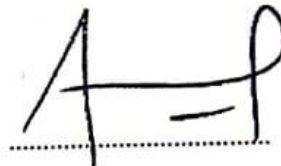
2. Pembimbing II

Mastura Diana Marieska, M.T.  
NIP. 198603212018032001



3. Penguji I

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



4. Penguji II

Rizki Kurniati, M.T.  
NIP. 199107122019032016



Mengetahui,

Ketua Jurusan Teknik Informatika



Alvi Syahrini Ujumi, M.Kom  
NIP. 197812222006042003

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Zainudin

NIM : 09021381621079

Program Studi : Teknik Informatika

Judul : Perbandingan Metode *Random Forest* dan *K Nearest Neighbor*  
(KNN) Pada Klasifikasi Penderita Penyakit Parkinson

Hasil Pengecekan Software *iThenticate/Turnitin* : 15%

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 sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

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



Palembang, 19 April 2021



Zainudin

NIM. 09021381621079

## HALAMAN MOTTO DAN PERSEMBAHAN

Motto :

“God destroyed your plan so that your plan does not destroy you”

Kupersembahkan karya tulis ini kepada :

- Orang tuaku tercinta
- Keluarga besarku
- Sahabat dan teman seperjuanganku
- Jurusan Teknik Informatika
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

## ABSTRACT

Parkinson's disease is the second most common neurodegenerative disease affecting 2-3% of the human population aged 65 and over. Parkinson's disease can be diagnosed through medical technology, but early diagnosis can also be made using machine learning and data mining. The Random Forest method and the KNN are two methods part of the classification, the way the two methods work have differences in data classification. The Random Forest method will convert the data into a pattern in the form of decision trees, then the number of votes is carried out. Meanwhile, KNN classifies data according to the proximity between the training data and the test data. Because the way the two methods work is different, this study will compare using data on people with Parkinson's disease. Based on the research results of the Random Forest method with several tested parameters, the optimal accuracy value is 89.42%, the precision value is 90.16%, and the recall value is 88.7%. KNN method with the highest classification results at k 1 and 3 yields accuracy 93.85% precision 97.29% and recall 90.97%. The results obtained were not much different between the two methods, this shows that the two methods were able to classify people with Parkinson's disease.

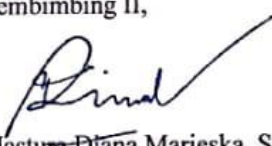
**Keywords:** Classification, Random Forest, K-Nearest Neighbor (KNN), Parkinson's Diseases

Palembang, 19 April 2021

Pembimbing I,

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

Pembimbing II,

  
Mastura Diana Marieska, S.T., M.T.  
NIP. 198603212018032001

Mengetahui,  
Ketua Jurusan Teknik Informatika

  
Alvi Syahrini Utami, M.Kom  
NIP. 197812222006042003



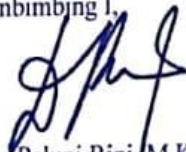
## ABSTRAK

Penyakit Parkinson adalah penyakit neurodegeneratif paling umum kedua mempengaruhi 2-3% populasi manusia usia 65 keatas. Penyakit Parkinson dapat diagnosis melalui teknologi kedokteran, akan tetapi diagnosis awal dapat juga dilakukan menggunakan mesin pembelajaran dan data mining. Metode *Random Forest* dan KNN dua metode bagian dari klasifikasi, cara kerja kedua metode memiliki perbedaan dalam pengklasifikasian data. Metode *Random Forest* akan mengubah data menjadi pola yang berbentuk pohon-pohon keputusan selanjutnya dilakukan *vote* terbanyak. Sedangkan KNN mengklasifikasikan data menurut jarak kedekatan antara data latih dengan data uji. Karena cara kerja kedua metode berbeda penelitian ini akan membandingkan menggunakan data penderita penyakit Parkinson. Berdasarkan hasil penelitian metode *Random Forest* dengan beberapa parameter yang diujikan didapat nilai *accuracy* optimal sebesar 89,42 % nilai *precision* 90,16%, dan nilai *recall* 88,7%. Metode KNN hasil klasifikasi tertinggi pada k 1 dan 3 menghasilkan *accuracy* 93,85% *precision* 97,29% dan *recall* 90,97%. Hasil yang didapatkan tidak jauh berbeda antar kedua metode, hal ini menunjukkan kedua metode mampu mengklasifikasikan penderita penyakit Parkinson.

**Kata Kunci:** Klasifikasi, *Random Forest*, *K-Nearest Neighbor* (KNN), Penyakit Parkinson

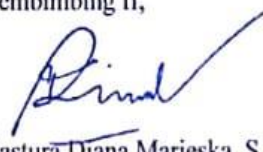
Palembang, 19 April 2021

Pembimbing I,



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

Pembimbing II,



Mastura Diana Marieska, S.T., M.T.  
NIP. 198603212018032001

Mengetahui,  
Ketua Jurusan Teknik Informatika



Alvi Svahri Utami, M.Kom  
NIP. 197812222006042003

## **KATA PENGANTAR**

Puji syukur kepada Allah SWT 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. Bapak Jaidan Jauhari, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika dan Ibu Mastura Diana Marieska, M. T. Selaku Sekretaris Jurusan Teknik Informatika.
3. Ibu, ayah, dan kakak-kakak saya yang selalu memberikan dukungan baik moril maupun materil, serta keluarga besar yang selalu mendoakan.
4. Ibu Dian Palupi Rini, M.Kom., Ph.D. selaku dosen pembimbing I yang telah membimbing, mengarahkan, dan memberikan motivasi dalam proses perkuliahan dan pengerjaan Tugas Akhir.



5. Ibu Mastura Diana Marieska, M.T. selaku dosen pembimbing II yang telah membimbing, mengarahkan, dan memberikan motivasi dalam proses perkuliahan dan pengerjaan Tugas Akhir.
6. Bapak M. Fachrurrozi, S.Si., M.T. selaku dosen penguji I yang telah memberikan masukan dan dorongan dalam proses pengerjaan Tugas Akhir.
7. Ibu Rizki Kurniati, M.T. selaku dosen penguji II yang telah memberikan masukan dan dorongan dalam proses pengerjaan Tugas Akhir.
8. Bapak Rifkie Primartha, M.T. selaku pembimbing akademik saya yang telah memberikan arahan dan motivasi dalam menjalani perkuliahan dari awal hingga akhir.
9. Seluruh dosen Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
10. Kak Ricy dan Mbak Wiwin yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan, serta seluruh staf administrasi yang telah membantu dalam pengurusan surat dan berkas lainnya.
11. Putra, Ade, Rifal, Bagus, Bayu, Satria, Dendi, Wahyu, Isa, Cesil, Desi serta teman-teman IF angkatan 2016 yang sudah bersama-sama berjuang dalam menempuh ilmu sejak awal perkuliahan.
12. Teman-teman anggota BPH HMIF yang memberikan saya pengalaman berorganisasi selama perkuliahan.

13. Beserta seluruh pihak yang tidak dapat disebutkan satu persatu.

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, April 2021

Zainudin

## DAFTAR ISI

	Halaman
HALAMAN JUDUL .....	i
LEMBAR PENGESAHAN TUGAS AKHIR.....	viii
TANDA LULUS UJIAN SIDANG TUGAS AKHIR.....	iii
HALAMAN PERNYATAAN.....	iv
HALAMAN MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT .....	vi
ABSTRAK .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	xi
DAFTAR TABEL .....	xvi
DAFTAR GAMBAR .....	xviii
BAB I PENDAHULUAN .....	I-1
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang.....	I-1
1.3 Rumusan Masalah.....	I-4
1.4 Tujuan Penelitian .....	I-4
1.5 Manfaat Penelitian .....	I-5
1.6 Batasan Masalah .....	I-5
1.7 Sistematika Penulisan.....	I-5
1.8 Kesimpulan.....	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan.....	II-1
2.2 Landasan Teori .....	II-1
2.2.1 Penyakit Parkinson.....	II-1
2.2.2 Data Mining .....	II-3
2.2.3 Klasifikasi .....	II-4

2.2.4 Metode Random Forest .....	II-4
2.2.4.1 Decision Tree .....	II-5
2.2.5 Metode K-Nearest Neighbor (KNN).....	II-7
2.2.6 Min-Max Normalization.....	II-8
2.2.7 K-Fold Cross Validation.....	II-9
2.2.8 Compusion Matrix .....	II-9
2.2.9 Rational Unified Process (RUP) .....	II-10
2.3 Penelitian Lain Yang Relevan .....	II-12
2.3.1 Traffic Accident Detection Using Random Forest Classifier (2018).....	II-12
2.3.2 Perbandingan Algoritma Klasifikasi dalam Pengklasifikasian Data Penyakit Jantung Koroner (2019).....	II-12
2.3.3 Sistem Klasifikasi Pada Penyakit Parkinson Dengan Menggunakan Metode K-Nearest Neighbor (2020).....	II-13
2.3.4 A comparative study on thyroid disease detection using K- Nearest Neighbor and Naive Bayes classification techniques (2019).....	II-13
2.4 Kesimpulan.....	II-14
<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-2
3.3 Tahapan Penelitian.....	III-2
3.3.1 Kerangka Kerja .....	III-2
3.3.2 Kriteria Pengujian .....	III-5
3.3.3 Format Data Pengujian .....	III-6
3.3.4 Alat Yang Digunakan Dalam Pelaksanaan Penelitian .....	III-6
3.3.5 Pengujian Penelitian .....	III-7
3.3.6 Analisis Hasil Pengujian dan Kesimpulan.....	III-7

3.4 Metode Pengembangan Perangkat Lunak .....	III-8
3.4.1 Fase Insepsi.....	III-8
3.4.2 Fase Elaborasi .....	III-9
3.4.3 Fase konstruksi.....	III-9
3.4.4 Fase Transisi .....	III-10
3.5 Manajemen Proyek Penelitian .....	III-10
3.6 Kesimpulan.....	III-15
<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 Metode Random Forest .....	IV-5
4.2.3.4 Analisis Metode K-Nearest Neighbor (KNN).....	IV-11
4.2.3.5 Desain Perangkat Lunak .....	IV-15
4.3 Fase Elaborasi.....	IV-24
4.3.1 Pemodelan Bisnis .....	IV-24
4.3.1.1 Perancangan Data .....	IV-24
4.3.1.1 Perancangan Antarmuka.....	IV-24
4.3.2 Kebutuhan Sistem .....	IV-25
4.3.3 Diagram .....	IV-26
4.3.3.1 Diagram Aktivitas.....	IV-26
4.3.3.2 Diagram Alur .....	IV-28
4.4 Fase Konstruksi .....	IV-31
4.4.1 Kebutuhan Sistem .....	IV-32
4.4.2 Diagram Kelas .....	IV-32
4.4.3 Implementasi.....	IV-34

4.3.3.1 Implementasi Kelas.....	IV-34
4.3.3.2 Implementasi Antarmuka .....	IV-36
4.5 Fase Transisi .....	IV-38
4.5.1 Pemodelan Bisnis .....	IV-38
4.5.2 Kebutuhan Sistem .....	IV-38
4.5.3 Rencana Pengujian .....	IV-39
4.5.3.1 Rencana Pengujian Use Case Memuat Data RF.....	IV-39
4.5.3.2 Rencana Pengujian Use Case Memuat Data KNN ....	IV-39
4.5.3.3 Rencana Pengujian Use Case Proses Klasifikasi Random Forest .....	IV-40
4.5.3.4 Rencana Pengujian Use Case Menguji klasifikasi menggunakan <i>K-Nearest Neighbor</i> (KNN).....	IV-40
4.5.4 Implementasi.....	IV-41
4.6 Kesimpulan.....	IV-44
BAB V HASIL DAN ANALISIS PENELITIAN .....	V-1
5.1 Pendahuluan.....	V-1
5.2 Data Hasil Percobaan/Penelitian.....	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 Random Forest.....	V-2
5.2.2.2 Hasil Pengujian Parameter Metode <i>K-Nearest Neighbor</i> (KNN) .....	V-2
5.3 Analisis Hasil Pengujian .....	V-3
5.3.1 Analisis Hasil Pengujian Klasifikasi Menggunakan Metode Random Forest.....	V-3
5.3.2 Analisis Hasil Pengujian Metode <i>K-Nearest Neighbor</i> (KNN) .....	V-4
5.4 Kesimpulan .....	V-5
BAB VI KESIMPULAN DAN SARAN .....	VI-1
6.1 Pendahuluan.....	VI-1

6.2 Kesimpulan.....	VI-1
6.3 Saran.....	VI-2
DAFTAR PUSTAKA .....	VII-1

## DAFTAR TABEL

	Halaman
Tabel II-1 Confusion Matrix .....	II-16
Tabel III-1 Rancangan tabel hasil pengujian metode Random Forest .....	III-6
Tabel III-2 Rancangan tabel hasil pengujian metode K-Nearest Neighbor (KNN) .....	III-6
Tabel III-3. Tabel Work Breakdown Structure (WBS) .....	III-11
Tabel IV-1 Kebutuhan Fungsional .....	IV-3
Tabel IV-2 Kebutuhan Non-Fungsional .....	IV-3
Tabel IV-3. Data Latih Random Forest .....	IV-5
Tabel IV-4 Data Uji Random Forest.....	IV-11
Tabel IV-5 Data Latih K-Nearest Neighbor (KNN).....	IV-12
Tabel IV-6 Data Uji K-Nearest Neighbor (KNN) .....	IV-12
Tabel IV-7 Definisi Aktor .....	IV-16
Tabel IV-8 Definisi Use Case.....	IV-16
Tabel IV-9 Skenario Memuat Data RF.....	IV-17
Tabel IV-10 Skenario Memuat Data KNN .....	IV-19
Tabel IV-11 Skenario Use Case Menguji Klasifikasi Menggunakan Random Forest .....	20
Tabel IV-12 Skenario Use Case Menguji klasifikasi menggunakan K-Nearest Neighbor (KNN).....	IV-22
Tabel IV-13 Tabel Implementasi Kelas .....	IV-34
Tabel IV-14 Rencana pengujian Use Case Memuat Data RF .....	IV-39



Tabel IV-15 Rencana pengujian Use Case Memuat Data KNN .....	IV-39
Tabel IV-16 Rencana Pengujian Use Case Mneguji Klasifikasi Menggunakan Random Forest .....	IV-40
Tabel IV-17 Rencana Pengujian Use Case Menguji klasifikasi menggunakan K- Nearest Neighbor (KNN) .....	IV-40
Tabel IV-18 Hasil Pengujian Use Case Memuat Data RF.....	IV-41
Tabel IV-19 Hasil Pengujian Use Case Memuat Data KNN .....	IV-42
Tabel IV-20. Hasil Pengujian Use Case Proses Klasifikasi Random Forest ....	IV-42
Tabel IV-21 Hasil Pengujian Use Case Proses Klasifikasi Metode K-Nearest Neighbor (KNN).....	IV-43
Tabel V-1 Hasil Perhitungan Klasifikasi Metode Random Forest.....	V-2
Tabel V-2 Hasil Perhitungan Metode K-Nearest Neighbor (KNN) .....	V-3

## DAFTAR GAMBAR

	Halaman
Gambar II-1 Fitur diekstraksi dari sampel data suara.....	II-2
Gambar II-2 Proses metode Random Forest.....	II-5
Gambar II-3 Model Rational Unified Process.....	II-11
Gambar III-1 Tahapan kerangka kerja pada metode Random Forest.....	III-3
Gambar III-2 Tahapan kerangka kerja metode KNN.....	III-4
Gambar III-3 Tahapan pengujian penelitian .....	III-7
Gambar IV-1 Diagram Use Case.....	IV-15
Gambar IV-2 Rancangan Antarmuka tab RF .....	IV-25
Gambar IV-3 Rancangan Antarmuka tab KNN .....	IV-25
Gambar IV-4 Diagram aktivitas memuat data.....	IV-27
Gambar IV-5 Diagram aktivitas RF.....	IV-27
Gambar IV-6 Diagram aktivitas KNN.....	IV-28
Gambar IV-7 Diagram Sequence Memuat Data.....	IV-29
Gambar IV-8 Diagram Sequence RF.....	IV-30
Gambar IV-9 Diagram Sequence KNN.....	IV-31
Gambar IV-10 Diagram Kelas.....	IV-33
Gambar IV-11 Implementasi Antarmuka Metode RF .....	IV-37
Gambar IV-12 Implementasi Antarmuka Metode KNN.....	IV-37
Gambar V-1 Grafik hasil akurasi klasifikasi metode Random Forest.....	V-3
Gambar V-2 Grafik hasil akurasi klasifikasi metode KNN.....	V-4

Gambar V-3 Perbandingan hasil akurasi klasifikasi metode Random Forest dan metode K-Nearest Neighbor (KNN)..... V-5

# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Pembahasan pada bab ini adalah menjelaskan dasar dari penelitian yang akan dilakukan mengenai perbandingan metode *Random Forest* dan KNN pada klasifikasi penderita penyakit Parkinson. Dasar dari penelitian ini akan dijelaskan diantaranya latar belakang masalah penelitian yang akan dilakukan, rumusan masalah yang menjadi alasan melakukannya penelitian, tujuan penelitian yang menggambarkan dari rumusan masalah yang ada dan manfaat penelitian yang merupakan hasil dari penelitian, serta batasan masalah penelitian yang bertujuan untuk membatasi ruang lingkup masalah yang ada pada penelitian supaya penelitian tetap pada area agar tidak luas cakupannya.

### **1.2 Latar Belakang**

Penyakit Parkinson adalah penyakit neurodegeneratif paling umum kedua yang mempengaruhi 2-3% dari populasi manusia pada usia 65 dan lebih tua (Poewe, et al., 2017). Menurut Parkinson's Disease Foundation, 7-10 juta orang di seluruh dunia terinfeksi penyakit ini. Penyakit Parkinson akibat dari penipisan neuron nigrostriatal dopaminergik, kerusakan terjadi pada gerakan (tremor, kekakuan, lambat gerakan, postur yang tidak stabil). Perubahan non-dopaminergik juga dapat mempengaruhi bahasa, dan suasana hati yang dapat mempengaruhi

komunikasi (Miller, 2009) dan dikembangkan oleh (Naranjo, Perez, Roca, & Martin, 2016) .

Penyakit Parkinson dapat diagnosis melalui teknologi kedokteran, akan tetapi diagnosis awal dapat juga dilakukan menggunakan mesin pembelajaran (*machine learning*) dan data mining. Penelitian terdahulu (Astuti & Ferinanto, 2016) yang membahas diagnosis penyakit parkinson berdasarkan kombinasi algoritme data mining yang menghasilkan nilai akurasi terbaik 96,923%. Pada penelitian ini menggunakan klasifikasi yang merupakan bagian dari data mining. Klasifikasi bertujuan untuk memprediksi kelas dari objek yang label kelasnya belum diketahui (Han & Kamber, 2012). Sehingga dapat memprediksi orang terkena penyakit Parkinson atau tidak dengan menggunakan teknik klasifikasi pada data mining.

Penelitian terdahulu yang membahas penyakit Parkinson oleh R.M. Aburizal Arif (2018) yang menggunakan metode Naive Bayes menghasilkan nilai akurasi 68,08% dan peningkatan akurasi setelah dioptimasi dengan *Particle Swarm Optimization* menghasilkan nilai akurasi sebesar 73,08%. Penelitian tersebut memiliki cukup banyak data dan tingkat akurasi yang dihasilkan pada penelitian tersebut dapat ditingkatkan. Sehingga perlu metode yang dapat menghasilkan akurasi yang lebih tinggi, Penelitian terdahulu (Januarsyah, Zuhairi, & Malik, 2019) yang membahas perbandingan metode Naive Bayes dan *Random Forest*. Metode *Random Forest* menghasilkan akurasi 74,28% yang mana lebih baik daripada metode Naive Bayes memiliki akurasi 49,06%, C4.5 yang memiliki

akurasi 57,53%, metode *Bayesian Network* memiliki akurasi 48,07% dan metode terakhir *Decision Stump* yang menghasilkan akurasi 49,95%.

Penelitian lain yang memiliki hasil analisis bahwa *Random Forest* mampu menghasilkan akurasi tinggi pada klasifikasi ketidaktepatan waktu kelulusan mahasiswa yang mencapai prediksi klasifikasi sebesar 93.23% (Suwardika & Suniantara, 2019). Metode klasifikasi lain yang diharapkan dapat menghasilkan akurasi yang tinggi yaitu *K-Nearest Neighbor* (KNN). Penelitian sebelumnya yang membahas perbandingan metode *K-Nearest Neighbor* (KNN) dan Naive Bayes yang berjudul “*A comparative study on thyroid disease detection using K-nearest neighbor and Naive Bayes classification techniques*” yang menghasilkan akurasi *K-Nearest Neighbor* (KNN) sebesar 93,44% dan Naive Bayes sebesar 22,56% (Chandel, Kunwar, Sabitha, Choudhury, & Mukherjee, 2017).

Maka pada penelitian ini metode yang dipilih ialah metode *Random Forest* dan *K-Nearest Neighbor* (KNN). Metode *Random Forest* merupakan salah satu teknik klasifikasi data dalam jumlah besar dan dikenal memiliki kinerja sangat baik (Balagatabi, Ibrahim, & Balagatabi, 2015). Sedangkan *K-Nearest Neighbor* (KNN) merupakan metode yang efektif digunakan untuk data *training* yang besar, bahkan dengan menggunakan data *training* yang sedikit, metode tersebut dapat berfungsi dengan baik (Chandel, Kunwar, Sabitha, Choudhury, & Mukherjee, 2017).

Berdasarkan penjelasan dan penelitian yang telah disebutkan, maka pada penelitian ini akan dilakukan penelitian menggunakan metode *Random Forest* dan

*K-Nearest Neighbor* (KNN). Penelitian ini bertujuan untuk mengetahui metode klasifikasi yang terbaik untuk penderita penyakit Parkinson.

### 1.3 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah mengetahui perbandingan hasil klasifikasi yang dihasilkan metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN) pada penderita penyakit Parkinson.

Pertanyaan penelitian yang diperlukan untuk menjawab permasalahan di atas adalah sebagai berikut :

1. Bagaimana menerapkan metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN) pada klasifikasi penderita penyakit Parkinson?
2. Bagaimana perbandingan hasil klasifikasi metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN) pada klasifikasi penderita penyakit Parkinson?

### 1.4 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Menerapkan metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN) sebagai metode pada klasifikasi penderita penyakit Parkinson.
2. Mengetahui perbandingan yang dihasilkan metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN) pada klasifikasi penderita penyakit Parkinson.

### 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Sistem dapat digunakan untuk media pembelajaran data mining dengan menggunakan metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN).
2. Hasil penelitian dapat dijadikan sumber referensi untuk memahami metode *Random Forest* dan metode *K-Nearest Neighbor* (KNN).

### 1.6 Batasan Masalah

Batasan masalah dari penelitian ini adalah:

1. Data yang digunakan pada penelitian ini didapat melalui situs penyedia *datasets* yaitu *UCI Machine Learning*.
2. Metode pencarian jarak antara data uji dan latih pada metode *K-Nearest Neighbor* (KNN) menggunakan *Eucliden Distance*.

### 1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yaitu sebagai berikut:

## BAB I. PENDAHULUAN

Pada bab ini akan diuraikan mengenai latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah/ruang lingkup, sistematika penulisan, dan kesimpulan.



## **BAB II. KAJIAN LITERATUR**

Pada bab ini akan membahas seluruh dasar-dasar teori yang digunakan mulai dari penyakit Parkinson, definisi klasifikasi dan metode *Random Forest* dan *K-Nearest Neighbor* (KNN).

## **BAB III. METODELOGI PENELITIAN**

Pada bab ini akan membahas mengenai tahap-tahap yang akan diterapkan pada penelitian. Setiap rencana dari tahapan penelitian dideskripsikan secara rinci berdasarkan kerangka kerja. Dilanjutkan dengan perancangan manajemen proyek dalam pelaksanaan penelitian.

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

Pada bab ini membahas mengenai analisis serta perancangan perangkat lunak yang akan dibuat pada penelitian ini.

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

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

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

Pada bab ini berisi kesimpulan dari semua uraian pada bab sebelumnya dan juga berisi saran-saran yang diharapkan berguna dalam penerapan metode klasifikasi dan pengembangan penelitian kedepannya.

### **1.8 Kesimpulan**

Pada bab ini telah dijelaskan secara umum mengenai penelitian yang akan dilakukan, meliputi latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penulisan.

## DAFTAR PUSTAKA

- Astuti, T., & Ferinanto, T. (2016). Diagnosis Penyakit Parkinson Berdasarkan Kombinasi Algoritma Data Mining dan Seleksi Fitur. Lombok Raya Mataram: SEMNASTIKOM.
- Balagatabi, Z. N., Ibrahim, R., & Balagatabi, H. N. (2015). Comparison of Decision Tree Methods in Classification of Researcher's Cognitive styles in Academic Environment. 223-231.
- Chandel, K., Kunwar, V., Sabitha, S., Choudhury, T., & Mukherjee, S. (2017). A comparative study on thyroid disease detection using K-nearest neighbor and Naive Bayes classification techniques. *Special Issue Redset 2016 Of CSIT* .
- Dogru, N., & Subasi, A. (2018). Traffic Accident Detection Using Random Forest Classifier.
- Han, J., & Kamber, M. (2012). *Data Mining: Concepts and Techniques*. University of Illinois at Urbana-Champaign: Elsevier.
- Januarsyah, M. F., Zuhairi, E., & Malik, R. F. (2019). Perbandingan Algoritma Random Forest, Decision Stump, Naïve Bayes, Bayesian Network dan Algoritma C4.5 Untuk . *Universitas Sriwijaya*, (pp. Vol.5 No.1 ISBN : 978-979-587-846-9). Palembang, Indonesia.
- Kacha, A., Mertens, C., Grenez, F., Skodda, S., & Schoentgen, J. (2017). On the harmonic-to-noise ratio as an acoustic cue of vocal timbre of. *Biomedical Signal Processing and Control* , 7.
- Miller. (2009). Communication changes in Parkinson's disease changes in communication Parkinson's disease. *Speech Therapy, Phoniatics and Audiology Magazine* , 37-46.
- Mubarok, F., Harlina, & Hadijah, I. (2015). Perbandingan Antara Metode RUP dan Prototype Dalam Aplikasi Penerimaan Siswa Baru Berbasis Web. *Citec Journal* , Vol. 2, No. 2.
- Naranjo, L., Perez, C. J., Roca, Y. C., & Martin, J. (2016). Addressing voice recording replications for Parkinson's disease detection. 286-292.
- poewe, w., seppi, k., m, c., halliday, g. h., brundin, p., volkmann, j., et al. (2017). Parkinson disease. 1.
- Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkman, J., et al. (2017). Parkinson disease. *Nature Reviews Disease Primesrs* , 3 17013.

- Prakoso, R. D., Wiriaatmadja, B. S., & Wibowo, W. F. (2020). Sistem Klasifikasi Pada Penyakit Parkinson Dengan Menggunakan Metode K-Nearest Neighbor. Hal 63-68.
- Ramauna, M. D., & Mahmudy, W. F. (2015). Optimasi Persediaan Barang Dalam Produksi Jilbab Menggunakan Algoritma Genetika. *DORO: Repository Jurnal Mahasiswa PTIIK Universitas Brawijaya* , vol. 5, no. 14.
- Sakar, B. E., Isenkul, M. E., Sakar, C. O., Sertbas, A., Gurgun, F., Delil, S., et al. (2013). Collection and Analysis of a Parkinson Speech Dataset With Multiple Types of Sound Recordings. *Biomedical And Health Informatics* .
- Saqib, P., Qamar, U., Aslam, A., & Ahmad, A. (2019). Hybrid of Filters and Genetic Algorithm- Random Forests Based Wrapper Approach for Feature Selection and Prediction. 190–199.
- Suwardika, G., & Suniantara, K. P. (2019). Analisis Random Forest Pada Klasifikasi Cart Ketidaktepatan Waktu Kelulusan Mahasiswa . *Barekeng* , 179-186.
- Upadhya, S. S., Cheeran, A., & Nirmal, J. H. (2017). Statistical comparison of Jitter and Shimmer voice features for healthy and Parkinson affected persons. *Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)* , 22-24.
- Utami, M. P., Nhita, F., & Aditsania, A. (2019). Prediksi Penyebaran Penyakit Demam Berdarah Dengue (DBD) di Kabupaten Bandung menggunakan Hybrid Random Forest(RF) dan Genetic Algorithm(GA) . 9963.
- Wibisono, A. B., & Fahrurozi, A. (2019). Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner. *Jurnal Ilmiah Teknologi dan Rekayasa* , vol 24, No 3.
- Wibowo, A. T., Saikhu, A., & Soelaiman, R. (2016). Implementation Of Image Contained Spam Detection Algorithm Using Svm And Random Forest.
- Yuli, M. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. 213-219.