

**KLASIFIKASI KARIES GIGI DI RUMAH SAKIT GIGI DAN  
MULUT BAITURRAHMAH MENGGUNAKAN METODE  
*RANDOM FOREST***

**SKRIPSI**



Oleh  
**MARTIA ROSADA**  
**19337010**

**PROGRAM STUDI S1 STATISTIKA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI PADANG  
2024**

## PERSETUJUAN SKRIPSI

### KLASIFIKASI KARIES GIGI DI RUMAH SAKIT GIGI DAN MULUT BAITURRAHIMAH MENGGUNAKAN METODE *RANDOM FOREST*

Nama : Martia Rosada  
NIM : 19337010  
Program Studi : S1 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 03 Juni 2024

Mengetahui:  
Kepala Departemen Statistika



Dr. Yenni Kurniawati, S.Si., M.Si  
NIP. 198402232010122005

Disetujui Oleh:  
Pembimbing



Zilrahmi, S.Pd., M.Si  
NIP. 198911062019032009

## PENGESAHAN LULUS UJIAN SKRIPSI

Nama : Martia Rosada  
NIM : 19337010  
Program Studi : S1 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

### KLASIFIKASI KARIES GIGI DI RUMAH SAKIT GIGI DAN MULUT BAITURRAHMAH MENGGUNAKAN METODE *RANDOM FOREST*

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Skripsi  
Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Negeri Padang

Padang, 03 Juni 2024

#### Tim Penguji

	Nama	Tanda Tangan
Ketua	: Zilrahmi, S.Pd., M.Si	
Anggota	: Dr. Syafriandi, M.Si	
Anggota	: Tessy Octavia Mukhti, M.Si.,	

## SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertandatangan di bawah ini:

Nama : Martia Rosada  
NIM : 19337010  
Program Studi : S1 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, skripsi saya dengan judul "**Klasifikasi Karies Gigi Di Rumah Sakit Gigi Dan Mulut Baiturrahmah Menggunakan Metode Random Forest**" adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan.

Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan negara.

Demikian pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,  
Kepala Departemen Statistika,



Dr. Yenni Kurniawati, S.Si., M.Si  
NIP. 198402232010122005

Saya yang menyatakan,



Martia Rosada  
NIM. 19337010

# Klasifikasi Karies Gigi Di Rumah Sakit Gigi Dan Mulut Baiturrahmah Menggunakan Metode *Random Forest*

Martia Rosada

## ABSTRAK

Karies gigi di Indonesia memiliki prevalensi yang tinggi, dengan 45,3% penduduk mengalaminya, dan di Sumatera Barat mencapai 43,9%. Jika tidak diatasi, karies gigi dapat menyebabkan komplikasi serius seperti infeksi, abses, kerusakan permanen, dan kehilangan gigi, serta dapat mempengaruhi kesehatan umum, termasuk penyakit jantung dan diabetes. Pencegahan karies gigi penting dilakukan dengan meningkatkan kesadaran masyarakat tentang kebersihan gigi dan mulut. Penelitian ini bertujuan untuk mengetahui ciri-ciri pasien mengalami karies gigi dan mengukur akurasi metode *random forest* dalam klasifikasi karies gigi.

*Random forest* merupakan metode *ensemble* yaitu pengembangan dari berapa metode *decision tree* dengan menggunakan *bootstrap* sampling. *Random forest* memiliki tiga parameter utama yaitu *mtry* (banyak variabel prediktor yang dipilih), *ntree* (banyak pohon yang ditumbuhkan), *node size* (jumlah amatan minimum dalam sebuah *terminal node*). Validasi dari penelitian ini menggunakan tingkat *error rate* OOB terkecil dan *Variabel Importance Measure* (VIM). Klasifikasi *random forest* pada penelitian ini menggunakan data rekam medis poli sakit gigi dan mulut di Rumah Sakit Baiturrahmah Padang.

Klasifikasi *random forest* pada penelitian ini menghasilkan tingkat kesalahan OOB 29,52% dan akurasinya 71%. Model optimal diperoleh menggunakan *mtry*=2 dan *ntree*=500. Dari penelitian ini dapat disimpulkan bahwa plak gigi, umur, dan kebiasaan menggosok gigi merupakan *variabel importance* atau faktor utama yang mempengaruhi hasil klasifikasi.

Keywords: OOB, Random Forest, VIM

# **Classification of Dental Caries in RSGM Baiturrahmah using the Random Forest Method**

**Martia Rosada**

## **ABSTRACT**

Dental caries in Indonesia has a high prevalence, with 45.3% of the population affected, and in West Sumatra, it reaches 43.9%. If left untreated, dental caries can lead to serious complications such as infections, abscesses, permanent damage, and tooth loss, and can also impact general health, including heart disease and diabetes. Preventing dental caries is crucial by raising public awareness about oral hygiene. This study aims to identify the characteristics of patients with dental caries and measure the accuracy of the Random Forest method in classifying dental caries.

Random Forest is an ensemble method that develops multiple decision trees using bootstrap sampling. It has three main parameters: *mtry* (the number of predictor variables selected), *ntree* (the number of trees grown), and *node size* (the minimum number of observations in a terminal node). The validation of this study uses the smallest OOB error rate and Variable Importance Measure (VIM). The Random Forest classification in this study uses medical records data from the dental and oral clinic at Baiturrahmah Hospital Padang.

The Random Forest classification in this study resulted in an OOB error rate of 29.52% and an accuracy of 71%. The optimal model was obtained using *mtry*=2 and *ntree*=500. The study concludes that dental plaque, age, and tooth brushing habits are important variables or key factors influencing the classification results.

Keywords: OOB, Random Forest, VIM

## KATA PENGANTAR

Segala puji dan rasa syukur penulis ucapkan kepada Allah SWT yang telah memberikan rahmat iman, islam dan kesehatan. Serta telah memberkahi nikmat pikiran sehingga penulis dapat menyelesaikan Skripsi dengan judul “Klasifikasi Karies Gigi Di Rumah Sakit Gigi Dan Mulut Baiturrahmah Menggunakan Metode *Random Forest*”. Skripsi ini merupakan salah satu persyaratan untuk memperoleh gelar Sarjana Statistika. Sholawat dan salam penulis ucapkan untuk Nabi Muhammad Salallahu Alaihi Wassalam.

Dalam penyusunan dan penulisan skripsi ini penulis banyak mendapatkan bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh sebab itu, dalam kesempatan ini penulis dengan senang hati menyampaikan banyak terima kasih kepada:

1. Ibu Zilrahmi, S.Pd, M.Si, selaku dosen pembimbing akademik dan pembimbing skripsi yang telah meluangkan banyak waktu dan tenaga untuk memberikan ilmu, saran, dan arahan dalam menyelesaikan penelitian ini.
2. Bapak Dr. Syafriandi, M.Si., dan Ibu Tessy Octavia Mukhti, M.Si., Dosen Penguji Skripsi yang telah memberikan masukan dan saran demi kesempurnaan skripsi ini.
3. Ibu Dr. Yenni Kurniawati, M.Si., Kepala Departemen Statistika sekaligus Koordinator Program Studi S1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.

4. Bapak dan Ibu Dosen, Staf Pengajar, dan Karyawan Departemen Statistika FMIPA UNP yang telah membimbing dan berbagi ilmu pengetahuan kepada penulis selama berada di bangku perkuliahan.
5. Teristimewa teruntuk kedua orang tua dan keluarga yang telah memberikan semangat, nasehat, dukungan, dan bantuan moril maupun materi kepada penulis.
6. Semua sahabat, teman, serta rekan-rekan seperjuangan yang memberikan semangat kepada penulis.

Semoga dorongan, bimbingan dan bantuan yang diberikan kepada penulis menjadi amal ibadah yang diridhai Allah SWT. Penulis menyadari bahwa skripsi ini masih terdapat kekurangan. Oleh karena itu, masukan dan saran yang bersifat membangun dari berbagai pihak sangat diharapkan demi kesempurnaan penulisan dimasa mendatang. Akhir kata, penuli berharap semoga skripsi ini dapat memberikan sumbangan pemikiran bagi setiap pembaca dan dapat bermanfaat bagi semua pihak.

Padang, 11 Mei 2024

Penulis

Martia Rosada

## DAFTAR ISI

<b>ABSTRAK</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>KATA PENGANTAR</b> .....	iii
<b>DAFTAR ISI</b> .....	v
<b>DAFTAR TABEL</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	viii
<b>DAFTAR LAMPIRAN</b> .....	ix
<b>BAB I PENDAHULUAN</b> .....	1
A. Latar Belakang Masalah.....	1
B. Batasan Masalah Penelitian.....	8
C. Rumusan Masalah Penelitian .....	8
D. Tujuan Penelitian .....	9
E. Manfaat Penelitian .....	9
<b>BAB II KERANGKA TEORITIS</b> .....	10
A. Machine Learning .....	10
B. Recursive Partitioning atau Tree Based Method.....	11
C. Classification and Regression Trees (CART).....	15
D. Random Forest .....	18
<b>BAB III METODOLOGI PENELITIAN</b> .....	37
A. Jenis Penelitian.....	37
B. Jenis dan Sumber Data .....	37
C. Variabel Penelitian .....	37
D. Struktur Data Penelitian .....	38
E. Teknik Analisis Data.....	39
F. Diagram Alur Penelitian .....	41
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	42
A. Deskripsi Data.....	42

B. Analisis Random Forest .....	49
C. Pembahasan.....	53
<b>BAB V PENUTUP</b> .....	<b>56</b>
A. Kesimpulan .....	56
B. Saran.....	56
<b>DAFTAR PUSTAKA</b> .....	<b>57</b>

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
1. Komponen dalam <i>decision tree</i> .....	12
2. Kriteria Tingkat Kesalahan Prediksi .....	24
3. Kriteria Tingkat Akurasi Prediksi .....	25
4. Confusion Matrix .....	28
5. Variabel Penelitian .....	37
6. Struktur Data Penelitian .....	37
7. Nilai Laju Galat OOB(%) .....	49
8. Nilai VIM .....	50
9. Confusion Matrix .....	51

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
1. Persentase penyakit gigi dan mulut di Sumbar .....	4
2. Struktur pohon decision tree .....	13
3. Struktur pohon klasifikasi algoritma CART .....	16
4. Ilustrasi Struktur Random Forest .....	21
5. Pembagian sampel Bootstrap dan OOB pada data sampel asli.....	23
6. Skor untuk kriteria plak gigi .....	31
7. Tingkat Keparahan Karies Gigi .....	36
8. Diagram alur penelitian.....	41
9. Jenis Karies Gigi .....	42
10. Jenis Karies Berdasarkan Jenis Kelamin.....	43
11. Jenis Karies Berdasarkan Umur .....	44
12. Jenis Karies Berdasarkan Kebiasaan Menyikat Gigi .....	45
13. Jenis Karies Berdasarkan Plak Gigi .....	46
14. Karakteristik Pasien berdasarkan Tempat Tinggal .....	47
15. Karakteristik Pasien berdasarkan Status Pasien.....	48
16. Hasil Out of Bag Error .....	50

## DAFTAR LAMPIRAN

<b>Lampiran</b>	<b>Halaman</b>
1. Surat keterangan dari RSGM Baiturrahmah Padang .....	62
2. Data Karies Gigi di RSGM Baiturrahmah Padang .....	63
3. Syntax Paket Random Forest .....	64
4. Output R Studio .....	65

## **BAB I PENDAHULUAN**

### **A. Latar Belakang Masalah**

Kesejahteraan suatu negara dapat dilihat dari tingkat kesehatan masyarakatnya, kesehatan merupakan aspek yang sangat penting yang harus terpenuhi dalam kehidupan masyarakat. Hal ini tercermin dari kemampuan negara untuk menyediakan akses yang merata terhadap pelayanan kesehatan, memperhatikan kesejahteraan fisik dan mental seluruh masyarakatnya, serta mendorong gaya hidup yang sehat melalui edukasi dan kebijakan yang mendukung. Organisasi Kesehatan Dunia atau WHO mendefinisikan kesehatan sebagai kesejahteraan fisik, mental, dan juga sosial yang lengkap, dan tidak memiliki penyakit atau kelemahan (WHO, 2020). Undang-Undang nomor 36 tahun 2009, mendefinisikan kesehatan sebagai keadaan sehat, baik secara fisik, mental, spiritual maupun sosial untuk memungkinkan setiap orang dapat hidup produktif secara sosial dan juga ekonomi (Undang-undang tentang kesehatan tahun 2009). Oleh karena itu, pentingnya untuk memenuhi aspek kesehatan sehingga terbentuknya masyarakat yang sejahtera.

Kesehatan yang perlu diperhatikan selain kesehatan tubuh secara umum, juga kesehatan gigi dan mulut. Karena kesehatan gigi dan mulut dapat mempengaruhi kesehatan tubuh lainnya. Salah satu indikator kesehatan gigi dan mulut adalah tingkat kebersihan gigi dan mulut. Hal tersebut dapat dilihat secara klinis dari ada tidaknya sisa-sisa makanan yang menempel pada gigi, seperti pelikel, materi alba, debris, kalkulus, dan plak gigi (Sintawati, 2008). Ketika kebersihan gigi dan

mulut tidak terjaga maka akan menyebabkan masalah atau penyakit gigi dan mulut seperti, periodontitis, karies gigi, abses gigi, gingivitis dan masalah kesehatan gigi dan mulut lainnya.

Menurut Departemen Kesehatan Republik Indonesia 2013, penyakit gigi dan mulut termasuk dalam 10 jenis penyakit yang sering dialami oleh masyarakat (Depkes, 2013). Permasalahan gigi dan mulut yang banyak dialami masyarakat adalah karies atau gigi berlubang. Karies merupakan penyakit rusaknya jaringan keras gigi oleh aktivitas metabolisme bakteri dalam plak yang menyebabkan terjadinya demineralisasi (Listriana, 2017). Bakteri yang menyebabkan karies yaitu *streptococcus mutans*, memetabolisme sisa-sisa makanan yang mengandung gula dan pati menjadi asam. Asam ini kemudian melarutkan mineral dari enamel gigi, yang merupakan lapisan terluar dan paling keras dari gigi. Proses ini dikenal sebagai demineralisasi.

Karies terjadi ketika bakteri dalam plak memfermentasi karbohidrat yang terdapat pada sisa makanan, menghasilkan asam yang merusak lapisan keras gigi. Proses ini mengakibatkan penurunan mineral pada enamel gigi, yang merupakan lapisan pelindung terluar dari gigi. Seiring waktu, penurunan mineral ini dapat melemahkan enamel dan menyebabkan pembentukan lubang pada gigi (Selwitz, 2007). Jika tidak segera ditangani, lubang ini dapat membesar dan menyebabkan rasa sakit, infeksi, dan bahkan kehilangan gigi.

Karies disebabkan oleh faktor primer dan faktor sekunder. Faktor primer atau faktor utama merupakan faktor yang memiliki peran langsung dalam proses terjadinya karies seperti, *host* (penjamu), substrat mikroorganisme, dan waktu.

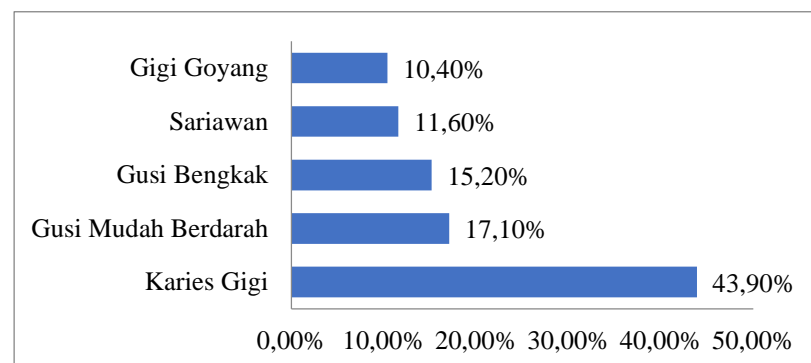
Sedangkan faktor sekunder atau faktor resiko merupakan faktor yang tidak memiliki peran langsung dalam proses terjadinya karies seperti usia, jenis kelamin, pekerjaan, pendidikan, lingkungan, tingkah laku, kultur (budaya), dan penyakit sistemik (Listrianah, 2017). Penyakit karies ditandai dengan rusaknya lapisan email gigi yang disebabkan oleh menumpuknya karang gigi pada lapisan gigi. Ketika karang gigi yang tidak dibersihkan akan menyebabkan gigi menjadi keropos dan berlubang.

Karies gigi yang tidak segera ditangani akan menyebabkan peradangan dan berpengaruh terhadap kesehatan tubuh. Hal ini dapat diketahui dari Penelitian yang dilakukan oleh ketua umum pengurus besar persatuan dokter gigi Indonesia, Dr. drg. Zaura Rini Anggraeni, MSD, menyatakan bahwa gigi berlubang yang tidak segera ditangani dapat memicu penyakit kronis pada jantung seperti infeksi katup jantung. Selain itu, gigi berlubang juga dapat berdampak negatif pada organ-organ lain seperti ginjal, paru-paru, lambung, bahkan dapat memperparah kondisi gula darah pada penderita *diabetes mellitus* (Wahyuni, 2016).

Menurut *Federation Dentaire Internationale* (FDI) tahun 2019, permasalahan yang banyak terjadi pada gigi dan mulut adalah karies gigi. Masalah karies gigi merupakan masalah yang banyak dialami oleh setiap penduduk di dunia (FDI, 2019). Begitu pula permasalahan karies gigi di Indonesia mempunyai tingkat prevalensi yang cukup tinggi. Berdasarkan hasil riset kesehatan dasar (Riskesdas) tahun 2018, sebanyak 57,6% dari 265 juta jiwa penduduk Indonesia mengalami masalah sakit gigi dan mulut. Masalah terbesar penyakit gigi dan mulut di

Indonesia adalah gigi rusak, berlubang, atau sakit gigi dengan persentase mencapai 45,3% (Riskesdas, 2018).

Masalah karies gigi untuk provinsi Sumatera Barat juga mengalami masalah yang serius, berdasarkan hasil riset kesehatan dasar (Riskesdas) tahun 2018, sebanyak 43,9% dari 5.519.245 jiwa penduduk Sumatera Barat mengalami masalah sakit gigi dan mulut. Masalah sakit gigi dan mulut terbesar adalah mengalami karies gigi dengan persentase 43,9%, masalah pembengkakan gusi (abses) 15,2%, dari penelitian tersebut yang mendapatkan penanganan medis dan konsultasi dengan dokter gigi hanya 5,2%. (Riskesdas, 2018). Gambar 1 menyajikan persentase penyakit gigi dan mulut di Sumatera Barat.



(Sumber: Riskesdas, 2018)

**Gambar 1.** Persentase penyakit gigi dan mulut di Sumbar

Berdasarkan Gambar 1 dapat diketahui bahwa masalah karies gigi merupakan penyakit yang paling banyak dialami oleh masyarakat Sumatera Barat. Sebanyak 43,90% dari 5.519.245 jiwa masyarakat Sumatera Barat mengalami karies gigi. Hal ini berarti dari 5.519.245 jiwa penduduk 2.422.948 jiwa mengalami karies gigi. Namun, sebagian besar masyarakat belum mengetahui bahaya yang akan timbul akibat gigi berlubang sehingga banyak yang tidak peduli dengan kesehatan giginya. Cara untuk meningkatkan kesadaran masyarakat yaitu dengan melakukan

edukasi kepada masyarakat agar mengetahui faktor-faktor yang mempengaruhi terjadinya karies gigi. Dengan meningkatnya kesadaran masyarakat tentang pentingnya perawatan gigi dan mulut di Sumatera Barat maka masalah karies gigi dapat berkurang.

Kondisi ini menekankan pentingnya fasilitas kesehatan gigi seperti Rumah Sakit Gigi dan Mulut (RSGM) Baiturrahmah dalam menyediakan layanan diagnostik dan perawatan yang diperlukan. RSGM Baiturrahmah memiliki lokasi yang strategis dan mudah diakses yaitu berlokasi di Jalan Bypass Km. 15 Air Pacah, Padang. RSGM Baiturrahmah dipilih sebagai tempat penelitian untuk mewakili kasus karies gigi di Sumatera Barat karena jangkauan pasiennya yang luas dan beragam, mencakup berbagai kelompok usia dan latar belakang sosial ekonomi. Banyak pasien rujukan yang berobat ke RSGM Baiturrahmah karena memiliki fasilitas yang lengkap dan tenaga medis yang berpengalaman. RSGM Baiturrahmah dapat memberikan data yang representatif dan akurat tentang prevalensi dan karakteristik karies gigi di Sumatera Barat. Data dari penelitian ini penting untuk menyusun strategi yang efektif dalam pencegahan dan pengobatan karies gigi, guna meningkatkan kesehatan gigi dan mulut masyarakat secara keseluruhan.

Penanganan karies gigi dilakukan secara cepat dan tepat, diperlukan metode yang mampu mengklasifikasi karies gigi berdasarkan gejalanya. Metode klasifikasi ini sangat berguna untuk mengetahui faktor utama yang menjadi penyebab karies gigi, sehingga dokter dapat merancang strategi pengobatan yang

tepat dan efektif. Dengan demikian, langkah pencegahan dan perawatan dapat dilakukan dengan lebih optimal, mengurangi risiko komplikasi lebih lanjut.

Klasifikasi adalah proses penempatan objek ke dalam satu set kategori, berdasarkan sifat objek yang bersangkutan (Larose, 2013). Metode klasifikasi digunakan untuk pembelajaran fungsi-fungsi berbeda yang memetakan masing-masing data terpilih ke dalam salah satu dari kelompok kelas yang telah ditetapkan sebelumnya. Klasifikasi dapat didefinisikan sebagai pekerjaan yang melakukan pelatihan/pembelajaran terhadap fungsi target yang memetakan setiap set atribut ke satu jumlah label kelas yang tersedia.

Metode klasifikasi yang umum digunakan dalam analisis data dan pembelajaran mesin, di antaranya adalah *decision tree*, *K-Nearest Neighbors* (KNN), *Support Vector Machine* (SVM), *naive bayes*, dan *random forest*. Masing-masing metode memiliki kelebihan dan kekurangannya sendiri dalam hal akurasi, kecepatan, dan interpretabilitas. Pada penelitian ini menggunakan metode *random forest* karena kemampuannya untuk mengatasi *overfitting* yang sering terjadi pada *decision tree* tunggal, mampu menghasilkan prediksi yang stabil dan akurat dalam kasus klasifikasi kompleks dengan dataset besar (Breiman, 2001). Selain itu, *random forest* juga dapat mengakomodasi berbagai jenis data, baik yang bersifat kategorikal maupun numerik, serta memberikan perkiraan kepentingan variabel untuk meningkatkan interpretasi hasil.

*Random forest* melakukan klasifikasi dengan cara pendekatan metode *ensemble* dari berbagai pohon melalui banyaknya kemunculan untuk mencapai keputusan akhir (Breiman, 2001). Teknik ini berawal dari penggunaan *decision*

*tree* sebagai dasar. Dalam *random forest*, banyak pohon keputusan (*decision tree*) dibangun dalam sebuah hutan (*forest*) sehingga pohon-pohon tersebut memberikan hasil klasifikasi yang lebih baik dan akurat. Setiap pohon yang dibentuk menggunakan subset acak dari data latih, yang membantu mengurangi *overfitting* dan meningkatkan generalisasi model (Liaw & Wiener, 2002).

Selama proses klasifikasi, setiap pohon dalam hutan memberikan suara (*voting*) untuk kelas yang dianggap paling sesuai berdasarkan data yang dimilikinya. Keputusan akhir diambil berdasarkan suara mayoritas dari seluruh pohon yang ada, sehingga menghasilkan prediksi yang lebih stabil dan akurat dibandingkan dengan menggunakan satu pohon keputusan saja (Breiman, 2001). Pendekatan ini menjadikan *random forest* salah satu metode yang sangat efektif dalam berbagai aplikasi machine learning, terutama ketika menghadapi data yang kompleks dan beragam.

Beberapa penelitian terdahulu terkait penggunaan algoritma *Random Forest* telah dilakukan. Penelitian oleh Depari (2022) yang membandingkan model *Decision Tree*, *Naive Bayes*, dan *Random Forest* untuk prediksi klasifikasi penyakit jantung. Hasil penelitian menunjukkan bahwa model *naive bayes* memiliki nilai akurasi sebesar 71%, model *decision tree* sebesar 72%, dan model *random forest* menghasilkan nilai akurasi tertinggi sebesar 75%. Penelitian oleh Ismanto (2021) tentang Komparasi Kinerja Algoritma C4.5, *Random Forest*, dan *Gradient Boosting* untuk Klasifikasi Komoditas. Hasil penelitian menunjukkan bahwa model C4.5 memiliki nilai akurasi sebesar 90%, sementara model *random forest* mencapai nilai akurasi yang lebih tinggi sebesar 98,96%. Penelitian Hanum

dan Zailani (2020) tentang penerapan algoritma klasifikasi *random forest* untuk penentuan kelayakan pemberian kredit di koperasi mitra sejahtera dengan hasil akurasi sebesar 87,88%.

Berdasarkan penelitian terdahulu tentang penerapan algoritma *random forest* yang memiliki tingkat akurasi yang cukup baik dalam pengujian model pada pohon keputusan. Maka, dilakukan penelitian ilmiah dengan judul “Klasifikasi Karies Gigi Di Rumah Sakit Gigi Dan Mulut Baiturrahmah Menggunakan Metode *Random Forest*”.

## **B. Batasan Masalah Penelitian**

Batasan masalah pada penelitian ini adalah sebagai berikut.

1. Data yang digunakan bersumber dari RSGM Baiturrahmah Padang, yaitu data yang diteliti dari bulan Januari-Oktober 2023.
2. Objek penelitian menggunakan data pasien sakit gigi dan mulut yang terdiri dari 300 tindakan dan 7 variabel yakni variabel Karies gigi, jenis kelamin, umur, Plak Gigi, status pasien (pernah berobat sebelumnya), frekuensi menggosok gigi dalam sehari, dan tempat tinggal.
3. Analisis yang digunakan adalah algoritma *random forest*.
4. Menggunakan  $mtry_1=2$ ,  $mtry_2=3$ ,  $mtry_3=4$  dan  $ntree$  100, 200, 300, 500

## **C. Rumusan Masalah Penelitian**

Rumusan masalah pada penelitian ini adalah sebagai berikut.

1. Bagaimana ciri-ciri pasien yang mengalami karies gigi di RSGM Baiturrahmah?

2. Bagaimana tingkat akurasi dari algoritma *random forest* dalam mengidentifikasi penyakit karies gigi?

#### **D. Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut.

1. Mengetahui ciri-ciri pasien yang mengalami karies gigi di RSGM Baiturrahmah.
2. Mengetahui tingkat akurasi yang dihasilkan dari algoritma *random forest* dalam mengidentifikasi penyakit karies gigi.

#### **E. Manfaat Penelitian**

Manfaat penelitian ini adalah sebagai berikut.

1. Bagi penulis, dapat menambah pengetahuan tentang pengolahan data dan dapat menerapkan ilmu yang dipelajari selama melaksanakan perkuliahan di Universitas Negeri Padang (UNP).
2. Bagi pembaca atau mahasiswa lainnya, penelitian ini diharapkan mampu memberikan manfaat dalam pembelajaran atau rujukan dalam penelitian mengenai metode klasifikasi menggunakan algoritma *random forest*.
3. Bagi masyarakat, hasil penelitian dapat digunakan untuk menyusun program penyuluhan kesehatan gigi yang lebih baik untuk masyarakat. Selain itu, juga dapat membantu meningkatkan kesadaran masyarakat tentang pentingnya menjaga kesehatan gigi dan mulut.

## **BAB V PENUTUP**

### **A. Kesimpulan**

Berdasarkan hasil penelitian diperoleh beberapa kesimpulan sebagai berikut:

1. Ciri-ciri pasien yang mengalami karies gigi di Rumah Sakit Baiturrahmah dapat diketahui dari urutan variabel terpenting dalam pengklasifikasian penyakit karies gigi dengan menggunakan perhitungan Mean Decrease Gini (MDG) adalah variabel plak gigi yang merupakan penciri dari pasien yang mengalami karies gigi.
2. Tingkat keakuratan prediksi pada random forest tersebut sebesar 71% dan tingkat kesalahan prediksi yang diperoleh sebesar 29,52%. Sehingga, kriteria keakuratan prediksi random forest tersebut dapat dikatakan baik.

### **B. Saran**

Berdasarkan penelitian yang telah dilakukan, adapun saran yang diajukan sebagai berikut.

1. Untuk pengembangan penelitian selanjutnya dapat ditambahkan beberapa atribut dan menggunakan lebih banyak data dalam mengklasifikasikan pasien karies gigi. Hal ini berguna untuk menghasilkan klasifikasi yang lebih baik.
2. Untuk penelitian selanjutnya, diharapkan menggunakan metode klasifikasi yang berbeda agar bisa melihat perbedaan akurasinya.

## DAFTAR PUSTAKA

- Ahmad, A. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Teknologi Indonesia*, 1-5.
- Adnyani, N., & Artawa, B. (2016). Pengaruh Penyakit Gigi Dan Mulut Terhadap Halitosis. *Jurnal Kesehatan Gigi*, 24-28.
- Ayyadevara, V. K. (2018). *Pro Machine Learning Algorithms*. India: Hyderabad, Andhra Pradesh.
- Alhamda, S. (2011). Status Kebersihan Gigi dan Mulut dengan Status Karies Gigi (Kajian pada Murid Kelompok Umur 12 Tahun di Sekolah Dasar Negeri Kota Bukittinggi). *Berita Kedokteran Masyarakat*, 108 - 115.
- Breiman, L. (2001). *Random Forest*. Berkeley: Statistics Department University of California, 5-32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification And Regression Tree*. New York: Chapman And Hall.
- Breiman, L., & Friedman, J. (1984). *Classification and Regression Tree*. New York: Chapman and Hall.
- Buulolo, E. (2020). *Data Mining untuk Perguruan Tinggi*. Yogyakarta: Deepbulish (Group Penerbitan CV Budi Utama).
- Departemen Kesehatan (Depkes). 2009. Kategori Usia. Diakses September 2023.
- D. Ribeiro, A. P., P. Maciel, I., & de Souza Hilgert, A. L. (2018). Caries assessment spectrum treatment: the severity score. *International Dental Journal*, 84-90.
- Depari, D. H., Widiastiwi, Y., & Santoni, M. M. (2022). Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *Jurnal Informatik*, 239-248.
- Genuer, R., & Poggi, J. M. (2020). *Random Forests with R*. Switzerland: Springer Nature .
- Genuer, R., & Poggi, J. M. (2008). *Random Forests: some methodological insights*. France: Inria.

- Hamoud, A. K. (2018). Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis. *International Journal of Interactive Multimedia and Artificial Intelligence*, 26-31.
- Han, J., Kamber, M., & Pei, J. (2006). *Data Mining Concepts and Techniques Second Edition*. San Francisco: Morgan Kaufmann.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques Third Edition*. USA: Morgan Kaufmann.
- Hanun, N., & Zailani, A. (2020). Penerapan Algoritma Klasifikasi Random Forest Untuk Penentuan Kelayakan Pemberian Kredit Di Koperasi Mitra Sejahtera. *Journal Of Technology Information*, 6(1), 7-14.
- Hidayah, U., & Sifaunajah, A. (2019). *Cara Mudah Memahami Algoritma K-Nearest Neighbor Studi Kasus Visual Basic 6.0*. Jombang: LPPM Universitas K.H A.Waahab.larose
- Husna, N., & Prasko. (2019). Efektivitas Penyuluhan Kesehatan Gigi Dengan Menggunakan Media Busy Book Terhadap Tingkat Pengetahuan Kesehatan Gigi Dan Mulut. *Jurnal Kesehatan Gigi*, 51-55.
- Ismanto, E., & Novalia, M. (2021). Komparasi Kinerja Algoritma C4.5, Random Forest, dan Gradient Boosting untuk Klasifikasi Komoditas. *Techno.Com*, 400-410.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Prentice International Inc.
- Janitza, S., & Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PLoS ONE*, 1-31.
- Kusrini, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Yogyakarta: CV Andi Offset.
- Kemenkes, R. (2016). Permenkes 89 tahun 2015 tentang upaya kesehatan gigi dan mulut.
- Kumala, A. D., & Hasiva, W. Y. (2019). Hubungan Tingkat Pengetahuan Tentang Cara Memelihara Kesehatan Gigi Dan Mulut Dengan Ketersediaan Alat Menyikat Gigi. *Jurnal Kesehatan Gigi*, 23-26.
- Kurniawan, U. S., Akmalia, N. S., & Syah Novildan, N. H. (2023). Peningkatan Pengetahuan Kesehatan Gigi dan Mulut Anak di Dusun Botokan,

- Sendangrejo, Minggu, Sleman . Jurnal Pengabdian kepada Masyarakat, 1-7.
- Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data : an introduction to data mining*. New Jersey: John Willey.
- Lewis, R. J. (2000). Pengantar Analisis Klasifikasi dan Pohon Regresi (CART). *SanFrancisco, California*.
- Listrianah, & Putri, N. (2017). Indeks karies gigi ditinjau dari penyakit umum dan sekresi saliva pada anak di Sekolah Dasar Negeri 30 Palembang , Jurnal Kesehatan Palembang, 12(2), 136-148)
- Lukacs J .R dan Largaespada L L. (2006). *Explaining Sex Differences in Dental Caries Prevalence: Saliva, Hormones, and Life History Etiologies*. American Journal of Human Biology. Wiley-Liss, inc;18:540-55)
- Lukacs, J.R. (2011). Sex differences in dental caries experience: clinical evidence, complex etiology. *Clin Oral Invest* , 649–656.
- Liaw, A., & Wiener, M. (2002). *Classification and regression by randomForest*. *R news*, 2(3), 18-22.
- Linasari, Meilenra K. Hubungan perilaku menyikat gigi pada malam hari dengan tingkat keparahan gingivitis pada remaja di Bandar Lampung. Jurnal Ilmiah Keperawatan Sei Batik. 2018;14(2):200-4.
- Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, 213-219.
- Marsland, S. (2015). *Machine Learning An Algorithmic Perspective Second Edition*. UK: CRC Press.
- Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horiz. B4*, 51-62.
- Muller, A., & Guido, S. (2016). *Introduction to machine learning with Python* . United States of Amerika: O'Reilly Media, Inc.
- Primajaya, A., & Sari, B. N. (2018). Random Forest Algorithm for Prediction of Precipitation. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*, 1(1), 27-31.

- Permana, A., Wahyuddin S, Santoso, L., Wibowo, G., Wardhani Khrisna, A., Rahmadden, et al. (2023). *Machine Learning*. Padang: Global Ekstusif Teknologi.
- Riset Kesehatan Dasar (Risksdas). 2018. Kesehatan Gigi Dan Mulut. Diakses Agustus 2023.
- Rahardjo, A. K., Widjiastuti, I., & Arif, E. (2016). Prevalensi Karies Gigi Posterior Berdasarkan Kedalaman, Usia Dan Jenis Kelamin Di Rsgm Fkg Unair Tahun 2014. *Conservative Dentistry Journal*, 66-70.
- Rezki, S., & Pawarti. (2014). Pengaruh Ph Plak Terhadap Angka Kebersihan Gigi Dan Angka Karies Gigi Anak Di Klinik Pelayanan Asuhan Poltekkes Pontianak Tahun 2013. *Odonto Dental Journal*, 13-18.
- Syamsu, S., Muhajirin, & Wijaya, N. S. (2019). Rules Generation Untuk Klasifikasi Data Bakat dan Minat Berdasarkan Rumpun Ilmu Dengan Decision Tree. *Jurnal Teknologi Informasi dan Komunikasi*, 9, 40-51.
- Saifullah. (2017). Analisa Terhadap Perbandingan Algoritma Decision Tree dengan Algoritma Random Tree untuk Pre-Processing Data. *Jurnal Sains Komputer & Informatika*, 180-185.
- Santoso, B. (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Saptiwi, B., Hanafi, M., & Purwitasari, D. (2019). Perilaku Pemeliharaan Kesehatan Gigi Dan Mulut Terhadap Status Kebersihan Gigi Dan Mulut (Ohi-S) Warga Samin Surosentiko Kabupaten Blora. *Jurnal Kesehatan Gigi*, 68-71.
- Syukron, A., & Subekti, A. (2018). Penerapan Metode Random Over-Under Sampling dan Random Forest untuk Klasifikasi Penilaian Kredit. *Jurnal Informatika*, 5, 175-185.
- Suwelo, L. S (1997). Peranan Pelayanan Kesehatan Gigi Anak Dalam Menunjang Kualitas Sumber Daya Manusia Indonesia Di Masa Mendatang. Jakarta: Universitas Indonesia.
- Schouten, K., Frasinca, F., & Dekker, R. (2016). An Information Gain-Driven Feature Study for Aspect-Based Sentiment Analysis. *in Natural Language Processing and Information Systems*, 48-59.

- Undang-Undang Nomor 36 Tahun 2009 Tentang Kesehatan. Jakarta. Kementrian Kesehatan Republik Indonesia.
- Wahyuni, E., & Rizky, M. (2016). Sistem Pakar Untuk Mendiagnosis Penyakit Gigi dan Mulut dengan Metode Pohon Keputusan. *Prosiding*, 151-160.
- Widiastuti, E., & Prasetyo, A. (2020). Faktor-faktor penyebab tingginya prevalensi karies gigi pada lansia di Puskesmas X. *Jurnal Kesehatan Gigi*, 8(3), 102-110.
- Yahya, Y., & Zuliana, R. (2018). Prediksi Jumlah Penggunaan BBM Perbulan Menggunakan Algoritma Decition Tree (C4. 5) Pada Kantor Dinas Lingkungan Hidup dan Kebersihan Kecamatan Selong Kabupaten Lombok Timur. *Jurnal Informatika dan Teknologi*, 1, 56-63.
- Yandi, S., & Sari, W. (2019). Distribusi Penyakit Gigi Dan Mulut Dalam Pelaksanaan Bulan Kesehatan Gigi Nasional (Bkgn) Fakultas Kedokteran Gigi Universitas Baiturrahmah Padang 2018. *Menara Ilmu*, 8(10), 75-80.
- Yuliana, L., & Setyowati, A. (2020). *Peran Pemeriksaan Gigi Rutin dalam Menjaga Kesehatan Gigi dan Mulut*. *Jurnal Kesehatan Gigi*, 8(2), 101-110.
- Ying Chan, A. K., Tamrakar, M., & Jiang, C. M. (2021). A Systematic Review on Caries Status of Older Adults. *Int. J. Environ. Res. Public Health*, 2-15.
- Zhang, H. (2004). Recursive Partitioning and Tree-based Methods. *Econstor*, 1-29.
- Zhang, H., & Singer, B. H. (2010). *Recursive Partitioning and Applications Second Edition*. USA: Springer Science.