



**ANALISIS SENTIMEN ULASAN MASYARAKAT DI  
INTERNET TERHADAP WISATA ALAM PUNTI KAYU  
KOTA PALEMBANG MENGGUNAKAN METODE  
*RANDOM FOREST***

**SKRIPSI**

**RIZKI KURNIAWAN SAPUTRA**

**2020.11.0006**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER DAN SAINS  
UNIVERSITAS INDO GLOBAL MANDIRI**

**2024**

**ANALISIS SENTIMEN ULASAN MASYARAKAT DI  
INTERNET TERHADAP WISATA ALAM PUNTI KAYU  
KOTA PALEMBANG MENGGUNAKAN METODE  
*RANDOM FOREST***



**SKRIPSI**

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

**Oleh:**

**RIZKI KURNIAWAN SAPUTRA**

**2020.11.0006**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER DAN SAINS  
UNIVERSITAS INDO GLOBAL MANDIRI  
2024**

**LEMBAR PENGESAHAN SKRIPSI**

**ANALISIS SENTIMEN ULASAN MASYARAKAT DI  
INTERNET TERHADAP WISATA ALAM PUNTI KAYU  
KOTA PALEMBANG**

Oleh

**RIZKI KURNIAWAN SAPUTRA**

**NPM : 2020.11.0006**

**Palembang , 28 Juni 2024**

**Pembimbing I**



**Dr. Shinta Puspasari, S.Si.,  
M.Kom NIK : 2015.01.0132**

**Pembimbing II**



**Lastris Widya Astuti, M.Kom,  
NIK:2003.01.0063**

**Mengetahui,**

**Dekan Fakultas Ilmu Komputer dan Sains**

FAKULTAS ILMU KOMPUTER DAN SAINS


**Rudi Heriansyah, S.T., M.Eng., Ph.D,  
NIK:2022.01.0315**

## LEMBAR PERSETUJUAN DEWAN PENGUJI

Pada hari Jum'at tanggal 5 Juli 2024 telah dilaksanakan ujian sidang skripsi :

Nama : Rizki Kurniawan Saputra

NPM : 2020.11.0006

Judul : Analisis Sentimen Ulasan Masyarakat Di Internet Terhadap  
Wisata Alam Pundi Kayu Kota Palembang Menggunakan  
Metode *Random Forest*

Oleh Prodi Teknik Informatika Fakultas Ilmu Komputer dan Sains Universitas  
Indo Global Mandiri Palembang

Palembang, 18 Juli 2024

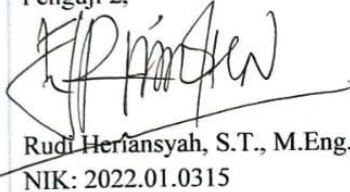
Penguji 1,



Dr. Herri Setiawan, M.Kom

NIK: 2003.01.0060

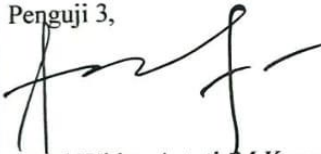
Penguji 2,



Rudi Herfiansyah, S.T., M.Eng., Ph.D

NIK: 2022.01.0315

Penguji 3,



Lastri Widya Astuti, M.Kom

NIK: 2003.01.0063

Menyetujui,  
Ka. Prodi Teknik Informatika



Zaid Romegar Mair, S.T., M.Cs

NIK: 2021.01.0307



SURAT KETERANGAN REVISI SKRIPSI  
PROGRAM STUDI TEKNIK INFORMATIKA (S1)  
FASILKOM DAN SAINS UNIVERSITAS INDO GLOBAL MANDIRI

---

Kami yang bertanda tangan dibawah ini, menerangkan bahwa :

Nama : Rizki Kurniawan Saputra  
NPM : 2020.11.0006  
Judul : Analisis Sentimen Ulasan Masyarakat Di Internet Terhadap  
Wisata Alam Punti Kayu Kota Palembang Menggunakan  
Metode *Random Forest*

Mahasiswa yang namanya tercantum diatas, telah selesai merevisi penulisan SKRIPSI

Palembang, 18 Juli 2024

Penguji 1,

Dr Herri Setiawan, M.Kom  
NIK: 2003.01.0060

Penguji 2,

Rudi Heriansyah, S.T., M.Eng., Ph.D  
NIK: 2022.01.0315

Penguji 3,

Lastri Widya Astuti, M.kom  
NIK: 2003.01.0063

Menyetujui,  
Ka. Prodi Teknik Informatika

Zaid Romegat Mair, S.T., M.Cs  
NIK: 2021.01.0307

## ABSTRAK

Pariwisata merupakan sektor penting dalam perekonomian Kota Palembang, dengan Punti Kayu sebagai salah satu destinasi wisata alam populer. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Wisata Alam Punti Kayu berdasarkan ulasan di *Google Maps* menggunakan metode *Random Forest*. Data ulasan dikumpulkan menggunakan *Instant Data Scraper* berjumlah 1000 data ulasan dari *Google Maps* dan digunakan untuk membangun model klasifikasi yang mengkategorikan sentimen menjadi positif atau negatif. Berdasarkan hasil pengujian awal dapat diketahui dengan menggunakan 1000 dataset ulasan, Pengujian pertama, menggunakan metode *Random Forest* tanpa *K-Fold Cross-Validation*, dengan pembagian data 80:20 (800 data *training* dan 200 data *testing*), menghasilkan akurasi sebesar 86%. Selanjutnya, pada pengujian yang sama dengan metode *Random Forest* yang dioptimalkan menggunakan *K-Fold Cross-Validation*, tercatat peningkatan akurasi menjadi 93% dengan *fold* terbaik yaitu *fold* ke-5. Kesimpulannya, analisis sentimen menggunakan metode *Random Forest* terbukti efektif dalam melakukan analisis sentimen.

**Kata Kunci:** Punti Kayu, Analisis sentimen, *Random Forest*, *K-Fold Cross-Validation*

## **ABSTRACT**

*Tourism is an important sector in the economy of Palembang City, with Pundi Kayu being one of the popular nature tourism destinations. This research aims to analyze public sentiment towards Pundi Kayu Nature Tourism based on reviews on Google Maps using the Random Forest method. Review data were collected using Instant Data Scraper, amounting to 1000 review data from Google Maps, and used to build a classification model that categorizes sentiment into positive or negative. Based on the initial test results, it can be seen that using a dataset of 1000 reviews, the first test, using the Random Forest method without K-Fold Cross-Validation, with an 80:20 data split (800 training data and 200 testing data), resulted in an accuracy of 86%. Furthermore, in the same test with the Random Forest method optimized using K-Fold Cross-Validation, an increase in accuracy to 93% was recorded with the best fold being the 5rd fold. In conclusion, sentiment analysis using the Random Forest method proved effective in conducting sentiment analysis.*

*Keywords: Pundi Kayu, Sentiment analysis, Random Forest, K-Fold Cross-Validation*

## KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan penelitian pada skripsi ini tepat pada waktunya, tidak lupa juga shalawat beserta salam kepada Nabi Muhammad SAW beserta pengikutnya hingga dan insyaallah kita semua dapat bertemu di surga nanti pada akhir zaman.

Penelitian pada skripsi yang penulis buat dengan judul “Analisis Sentimen Ulasan Masyarakat di Internet Terhadap Wisata Alam Pundi Kayu Kota Palembang Menggunakan Metode *Random Forest*” disusun guna memenuhi persyaratan untuk menyelesaikan Pendidikan Strata I Program Studi Teknik Informatika, Fakultas Ilmu Komputer dan Sains, Universitas Indo Global Mandiri Palembang.

Pertama dan paling utama, ucapan terima kasih kepada Allah SWT dan Nabi Muhammad SAW atas berkah dan karunia-Nyalah penulis bisa menyelesaikan penelitian pada skripsi ini. Selanjutnya, ucapan terima kasih penulis tujukan kepada yang terhormat:

1. Orangtua dan saudara yang telah memberikan doa, restu serta dukungan yang sangat besar selama menjalani kehidupan hingga perkuliahan di Universitas Indo Global Mandiri.
2. Bapak Dr. H. Marzuki Alie, S.E., M.M, selaku Rektor Universitas Indo Global Mandiri.
3. Bapak Rudi Heriansyah, S.T., M.Eng., Ph.D, selaku Dekan Fakultas Ilmu Komputer Universitas Indo Global Mandiri.
4. Bapak Zaid Romegar Mair, S.T., M.Cs, Selaku Ketua Program Studi Teknik Informatika Universitas Indo Global Mandiri.
5. Ibu Dr. Shinta Puspasari, S.Si., M.Kom, selaku Dosen Pembimbing I yang telah membimbing saya selama pembuatan Skripsi ini.
6. Ibu Lastri Widya Astuti, M.Kom, selaku Dosen Pembimbing II yang telah membimbing saya selama pembuatan Skripsi ini.



7. Ibu Dr. Shinta Puspasari, S.Si., M.Kom, sebagai Dosen Pembimbing Akademik.
8. Bapak/Ibu Dosen Program Studi Informatika yang telah memberikan banyak pembelajaran selama masa perkuliahan.
9. Teman-teman yang telah memberikan dukungan materi maupun moral sehingga penulis bisa menyelesaikan penelitian pada skripsi ini.

Penulis menyadari bahwa penelitian pada skripsi ini masih jauh dari kata sempurna, oleh karena itu kritik beserta saran untuk perbaikan dan pengembangan sangat dibutuhkan. Akhir kata, semoga penelitian pada skripsi ini bermanfaat bagi semua pihak. Penulis mengucapkan terima kasih.

Palembang,  
Penulis,

Rizki Kurniawan Saputra  
2020.11.0006

## DAFTAR ISI

<b>HALAMAN JUDUL LUAR</b> .....	<b>i</b>
<b>HALAMAN JUDUL DALAM</b> .....	<b>ii</b>
<b>HALAMAN PENGESAHAN SKRIPSI</b> .....	<b>iii</b>
<b>LEMBAR PERSETUJUAN DEWAN PENGUJI</b> .....	<b>iv</b>
<b>SURAT KETERANGAN REVISI SKRIPSI</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>KATA PENGANTAR</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiii</b>
<b>DAFTAR TABEL</b> .....	<b>xiv</b>
<b>DAFTAR RUMUS</b> .....	<b>xvi</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xvii</b>
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan dan Manfaat Penelitian.....	3
1.4.1 Tujuan Penelitian .....	3
1.4.2 Manfaat Penelitian .....	3
1.5 Sistematika Penulisan.....	4
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Wisata Alam Punti Kayu .....	5
2.2 Analisis Sentimen.....	5
2.3 Ulasan .....	6
2.4 <i>Data Mining</i> .....	7
2.5 Klasifikasi.....	8
2.6 <i>Random Forest</i> .....	8
2.7 <i>TF-IDF</i> .....	11

2.8 <i>Confusion Matrix</i> .....	11
2.9 <i>K-Fold</i> .....	12
2.10 <i>Colabs</i> .....	13
2.11 <i>Python</i> .....	14
2.12 Penelitian Terdahulu.....	18
<b>BAB III METODOLOGI PENELITIAN</b>	
3.1 Tahapan Penelitian .....	24
3.2 Kebutuhan Perangkat Keras dan Perangkat Lunak .....	25
3.3 Pengumpulan Data .....	25
3.4 <i>Preprocessing</i> .....	26
3.4.1 <i>Cleaning</i> .....	27
3.4.2 <i>Case Folding</i> .....	27
3.4.3 <i>Stemming</i> .....	28
3.4.4 <i>Stopword Removal</i> .....	29
3.4.5 <i>Tokenize</i> .....	30
3.5 Implementasi .....	30
3.6 Pengujian .....	40
<b>BAB IV HASIL DAN PEMBAHASAN</b>	
4.1 Pengumpulan data .....	42
4.2 Hasil proses preprosesing Data .....	43
4.2.1 <i>Cleaning Data</i> .....	44
4.2.2 <i>Case Folding</i> .....	46
4.2.3 <i>Tokenize</i> .....	49
4.2.4 <i>Stemming</i> .....	51
4.2.5 <i>Stopword Removal</i> .....	54
4.2.6 Proses Pelabelan <i>TF-IDF</i> .....	57
4.3 Implementasi <i>Random Forest</i> .....	60
4.3.1 Pengujian <i>Random Forest</i> .....	61
4.3.2 Pengujian <i>K-fold cross validation</i> .....	65
4.4 Kesimpulan Pengujian.....	67
<b>BAB V PENUTUP</b>	

5.1 Kesimpulan.....	68
5.2 Saran.....	68
<b>DAFTAR PUSTAKA .....</b>	<b>69</b>
<b>LAMPIRAN</b>	

## DAFTAR GAMBAR

Gambar 2.1. Ilustrasi Proses dari <i>Random Forest</i> .....	9
Gambar 2.2. Ilustrasi Proses dari <i>K-Fold Cross Validation</i> .....	13
Gambar 2.3. <i>Google Colabs</i> .....	14
Gambar 2.4. <i>Numpy</i> .....	15
Gambar 2.5. <i>Scipy</i> .....	16
Gambar 2.6. <i>scikit learn</i> .....	17
Gambar 2.7. <i>Pandas</i> .....	18
Gambar 2.8. <i>Matplotlib</i> .....	18
Gambar 3.1. Tahapan Penelitian.....	24
Gambar 3.2. Implementasi <i>Random Forest</i> .....	31
Gambar 3.3. Implementasi <i>Random Forest</i> .....	34
Gambar 3.4. <i>Root</i> Pohon Keputusan 1.....	36
Gambar 3.5. <i>Node 1.1</i> Pohon Keputusan 1.....	38
Gambar 3.6. <i>Node</i> Pohon 1.1.1 Pohon Keputusan 1.....	40
Gambar 4.1. pemanggilan Data.....	43
Gambar 4.2. <i>Cleaning data</i> .....	44
Gambar 4.3. <i>Case Folding</i> .....	46
Gambar 4.4. <i>Tokenize</i> .....	49
Gambar 4.5. Install Sastrawi untuk <i>NLP Stemming</i> .....	51
Gambar 4.6. <i>Stemming</i> .....	52
Gambar 4.7. <i>Stopword Removal</i> .....	55
Gambar 4.8. Pembobotan <i>TF-IDF</i> .....	57
Gambar 4.9. Hasil Pembobotan <i>TF-IDF</i> .....	58
Gambar 4.10. Pelebelan <i>TF-IDF</i> .....	59
Gambar 4.11. Pemanggilan Dataset.....	60
Gambar 4.12. Pemanggilan Fungsi <i>Random Forest</i> .....	61
Gambar 4.13. Proses Menghitung dan Menampilkan Akurasi <i>Random Forest</i> ....	62
Gambar 4.14. <i>Confusion Matrix</i> .....	63
Gambar 4.15. kode program Melakukan <i>K-fold cross validation</i> .....	66

## DAFTAR TABEL

Tabel 2.1. <i>Confusion Matrix</i> .....	12
Tabel 2.2. Penelitian Terdahulu .....	19
Tabel 3.1. Contoh Dataset Ulasan.....	26
Tabel 3.2. Cleaning Data.....	27
Tabel 3.3. Case Folding .....	28
Tabel 3.4. Stemming .....	28
Tabel 3.5. Stopword Removal.....	29
Tabel 3.6. Tokenize.....	30
Tabel 3.7. Pelabelan Dataset Awal .....	31
Tabel 3.8. Frekuensi Kemunculan (TF) .....	32
Tabel 3.9. <i>Perhitungan</i> IDF .....	32
Tabel 3.10. Perhitungan (TF-IDF) .....	34
Tabel 3.11. Perhitungan Pohon Keputusan 1 .....	35
Tabel 3.12. Hasil Perhitungan Root Pohon Keputusan 1 .....	36
Tabel 3.13. Perhitungan Node 1.1 Pohon Keputusan 1 .....	37
Tabel 3.14. Hasil Perhitungan Node 1.1 Pohon Keputusan 1 .....	38
Tabel 3.15. Perhitungan <i>Node</i> 1.1.1 Pohon Keputusan 1.....	39
Tabel 3.16. Hasil Perhitungan Node 1.1.1 Pohon Keputusan 1 .....	39
Tabel 3.17. Tabel <i>Confusion Matrix</i> .....	41
Tabel 4.1. Tabel Data Sebelum <i>preprocessing</i> .....	42
Tabel 4.2. Sebelum di <i>Cleaning</i> .....	44
Tabel 4.3. Setelah di <i>Cleaning</i> .....	45
Tabel 4.4. sebelum di <i>Case Folding</i> .....	46
Tabel 4.5. sesudah di <i>Case Folding</i> .....	48
Tabel 4.6. sebelum di <i>Tokenize</i> .....	49
Tabel 4.7. Sesudah di <i>Tokenize</i> .....	50
Tabel 4.8. Sebelum di <i>Stemming</i> .....	52
Tabel 4.9. Sesudah di <i>Stemming</i> .....	53
Tabel 4.10. Sebelum di Stopword Removal.....	55

Tabel 4.11. Sesudah di Stopword Removal .....	56
Tabel 4.12. Hasil Pelebelan <i>TF-IDF</i> .....	59
Tabel 4.13. Hasil Pengujian <i>Random Forest</i> .....	64
Tabel 4.14. Menampilkan hasil <i>K-fold cross validation</i> .....	66
Tabel 4.15. Menampilkan hasil Terbaik <i>K-fold cross validation</i> .....	67

## DAFTAR RUMUS

(2.1) Perhitungan nilai entorpy .....	10
(2.2) Perhitungan gain .....	10
(2.3) Perhitungan IDF .....	11
(2.4) Perhitungan TFIDF.....	11
(2.5) Perhitungan Acuraccy .....	12
(2.6) Perhitungan Recall.....	12
(2.7) Perhitungan Precision.....	12
(2.8) Perhitungan F1 - Score .....	12
(3.1) Perhitungan IDF pada data sampel.....	33
(3.2) Perhitungan TFIDF pada data sampel .....	34
(3.3) Perhitungan root pohon keputusan .....	35
(3.4) Perhitungan node 1.1 .....	37
(3.5) Perhitungan node 1.1.1 .....	39
(3.6) Perhitungan Acuraccy .....	41
(3.7) Perhitungan Recall.....	41
(3.8) Perhitungan Precision.....	41
(3.9) Perhitungan F1 - Score .....	41
(4.1) Rumus Acuraccy .....	63
(4.2) Rumus Recall.....	63
(4.3) Rumus Precision.....	63
(4.4) Rumus F1 - Score .....	63
(4.5) Perhitungan Acuraccy .....	63
(4.6) Perhitungan Recall.....	63
(4.7) Perhitungan Precision.....	64
(4.8) Perhitungan F1- Score .....	64



## **DAFTAR LAMPIRAN**

Lampiran 1 Daftar Riwayat Hidup

Lampiran 2 Kartu Bimbingan

Lampiran 3 Surat Pernyataan Tidak Plagiat