



**ANALISIS SENTIMEN PENGGUNA ANDROID TERHADAP APLIKASI
INDOMARET POINKU MENGGUNAKAN METODE *MAXIMUM
ENTROPY***

SKRIPSI

ANDHIKA RIZQ PULUBUHU

NIM. 2010512061

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

APRIL 2024



**ANALISIS SENTIMEN PENGGUNA ANDROID TERHADAP APLIKASI
INDOMARET POINKU MENGGUNAKAN METODE *MAXIMUM
ENTROPY***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

ANDHIKA RIZQ PULUBUHU

NIM. 2010512061

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

APRIL 2024

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun dirujuk telah saya nyatakan benar. Saya menyatakan dibawah ini :

Nama : Andhika Rizq Pulubuhu
NIM : 2010512061
Program Studi : S1 Sistem Informasi
Judul Skripsi/TA : Analisis Sentimen Pengguna Android Terhadap Aplikasi Indomaret
Poinku Menggunakan Metode *Maximum Entropy*
Tanggal : 14 Juni 2024

Bila mana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 14 Juni 2024

Yang menyatakan,



(Andhika Rizq Pulubuhu)

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang menandatangani pernyataan ini :

Nama : Andhika Rizq Pulubuhu

NIM : 2010512061

Fakultas : Ilmu Komputer

Studi : S1- Sistem Informasi


Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Eksklusif (Non-exclusive Royalty Free Right) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN PENGGUNA ANDROID TERHADAP APLIKASI
INDOMARET POINKU MENGGUNAKAN METODE *MAXIMUM
ENTROPY***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 14 Juni 2024

Yang Menyatakan,


(Andhika Rizq Pulubuhu)

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Dengan ini menyatakan bahwa Skripsi/Tugas Akhir mahasiswa berikut dengan identitas

Nama : Andhika Rizq Pulubuhu
NIM : 2010512061
Program Studi : S1 Sistem Informasi
Judul Skripsi/TA : Analisis Sentimen Pengguna Android Terhadap Aplikasi Indomaret
Poinku Menggunakan Metode *Maximum Entropy*

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi S1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.

Dosen Pembimbing 1

(Dr. Bambang Saras Yulistiawan, S.T., M.Kom.)

Dosen Pembimbing 2

(Bambang Triwahyono, S.Kom., M.Si.)

Dosen Penguji 1

(Dr. Ermatita, M.Kom.)

Dosen Penguji 2

(Iin Ernawati S.Kom., M.Si.)



Dekan

(Prof. Dr. Ir. Supriyanto, M.Sc., IPM.)

Ketua Program Studi S1 Sistem Informasi

(Anita Muliawati, S.Kom., MTL)

Ditetapkan di : Jakarta

Tanggal Ujian : 07 Juni 2024

**ANALISIS SENTIMEN PENGGUNA ANDROID TERHADAP APLIKASI
INDOMARET POINKU MENGGUNAKAN METODE *MAXIMUM
ENTROPY***

Andhika Rizq Pulubuhu

ABSTRAK

Indomaret Poinku merupakan program membership dengan menggunakan aplikasi *mobile* yang dimana *customer* dapat menukarkan poin dan *stamp*. Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi Indomaret Poinku di Android guna memberikan wawasan tentang kepuasan, masalah, dan area peningkatan yang perlu dilakukan oleh pengembang. Data dikumpulkan dari 6,435 ulasan pengguna di Google Play Store. Hasil penelitian Berdasarkan data ulasan pengguna aplikasi Indomaret Poinku dari Januari hingga Juni 2023, terdapat variasi jumlah ulasan dan *rating* bintang dengan puncak tertinggi pada bulan Februari dan Maret yang dipengaruhi oleh campaign Indomaret. Proses klasifikasi ulasan melalui preprocessing data, pembobotan kata menggunakan TF-Idf, dan *K-Fold Cross Validation* menunjukkan bahwa klasifikasi tanpa pembobotan TF-Idf memberikan hasil akurasi yang lebih baik dengan akurasi rata rata dari 5 iterasi adalah 91,6%, dengan rata-rata *precision* sebesar 0.923, rata-rata *recall* sebesar 0.949, dan rata-rata *F1-Score* sebesar 0.936. Analisis asosiasi teks mengungkap bahwa ulasan positif mencerminkan kelebihan aplikasi, sementara ulasan negatif menyoroti masalah penggunaan. Mayoritas ulasan memuji layanan dan promosi, namun juga menyoroti masalah teknis dan kesulitan penggunaan yang perlu diperbaiki. Rekomendasi untuk meningkatkan kualitas aplikasi meliputi optimalisasi program *membership*, peningkatan layanan pelanggan, perbaikan teknis aplikasi, dan implementasi fitur saran pengguna berdasarkan ulasan positif, serta perbaikan kinerja aplikasi, proses verifikasi, dan manajemen pembaruan aplikasi berdasarkan ulasan negatif.

Kata Kunci: Analisis Sentimen, Klasifikasi, *K-Fold Cross Validation*, TF-IDF, *Maximum Entropy*

**ANALYSIS OF ANDROID USER SENTIMENT TOWARDS THE
INDOMARET POINKU APPLICATION USING THE MAXIMUM ENTROPY
METHOD**

Andhika Rizq Pulubuhu

ABSTRACT

Indomaret Poinku is a membership program using a mobile application where customers can exchange points and stamps. This research aims to analyze the sentiment of users of the Indomaret Poinku application on Android to provide insight into satisfaction, problems and areas of improvement that need to be made by the developer. Data collected from 6,435 user reviews on the Google Play Store. Research results Based on data from user reviews of the Indomaret Poinku application from January to June 2023, there are variations in the number of reviews and star ratings with the highest peaks in February and March which are influenced by the Indomaret campaign. Although the majority of reviews provide positive ratings, there are a significant number of negative reviews related to technical issues and app usage. The review classification process through data preprocessing, word weighting using TF-Idf, and K-Fold Cross Validation shows that classification without TF-Idf weighting provides better accuracy results with the average accuracy of 5 iterations being 91.6%, with an average of The average precision is 0.923, the average recall is 0.949, and the average F1-Score is 0.936. Text association analysis revealed that positive reviews reflected app advantages, while negative reviews highlighted usage issues. The majority of reviews praise the service and promotions, but also highlight technical issues and usage difficulties that need to be fixed. Recommendations for improving application quality include optimizing the membership program, improving customer service, improving application technical aspects, and implementing a user suggestion feature based on positive reviews, as well as improving application performance, verification processes, and application update management based on negative reviews.

Keywords: Sentiment Analysis, Classification, K-Cross Validation, TF-IDF, Maximum Entropy

KATA PENGANTAR

Puji syukur kehadirat Allah SWT. beserta junjungan kita Nabi Muhammad SAW karena atas berkat dan rahmat-Nya telah memberikan kesehatan dan kekuatan sehingga penulis dapat menyelesaikan penelitian dengan judul “Analisis Sentimen Pengguna Android Terhadap Aplikasi Indomaret Poinku Menggunakan Metode Maximum Entropy”. Laporan skripsi ini disusun sebagai salah satu syarat untuk mengerjakan skripsi pada Jurusan S1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jakarta.

Dalam pelaksanaan proses pembuatan proposal penelitian ini tidak dapat lepas dari dukungan berbagai pihak, oleh karena itu, dalam kesempatan kali ini juga penulis mengucapkan terima kasihnya kepada:

1. Pak Dr. Bambang Saras Yulistiawan, S.T., M.Kom. selaku Dosen Pembimbing 1 atas bimbingan, saran, dan motivasi yang diberikan.
2. Pak Bambang Triwahyono, S.Kom., M.Si. selaku Dosen Pembimbing 2 atas bimbingan, saran, dan motivasi yang diberikan.
3. Ibu Anita Muliawati, S.Kom., M.TI. selaku Ketua Program Studi S1 Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jakarta.
4. Segenap Dosen Jurusan S1 Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jakarta yang telah memberikan ilmunya kepada penulis.
5. Pihak Indomaret Poinku Karena sudah mengizinkan dilakukan penelitian ini.
6. Orang tua, saudara-saudara, serta keluarga atas doa, bimbingan, serta motivasi yang telah diberikan selama ini.
7. Teman-teman seperjuangan di jurusan S1 Sistem Informasi Kelas B, atas semua dukungan semangat dan kerjasamanya.
8. Teman-teman yang membantu sebagai annotator dalam penelitian ini

Penulis mengakui adanya kelemahan dan keterbatasan dalam penyusunan proposal penelitian ini. Oleh karena itu, penulis berharap mendapatkan masukan serta kritik yang membangun guna meningkatkan kualitasnya. Diharapkan bahwa hasil akhir dari proposal skripsi ini akan memberikan kontribusi yang berarti dalam bidang pendidikan dan implementasinya di lapangan, serta memiliki potensi untuk dikembangkan lebih lanjut. Penulis mengakhiri pernyataannya dengan harapan bahwa proposal penelitian ini akan bermanfaat bagi pembaca dan pihak-pihak terkait lainnya.

Jakarta, 15 April 2024


Andhika Rizq Pulubuhu

DAFTAR ISI

SKRIPSI	i
PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian.....	4
1.4. Manfaat Penelitian.....	4
1.5. Ruang Lingkup Penelitian	4
1.6. Luaran.....	5
1.7. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Kajian Literatur	7
2.2. Landasan Teori	14
2.2.1 Indomaret	14
2.2.1.1 Indomaret Poinku.....	14
2.2.2 Web Scraping	15
2.2.3 Machine Learning	16
2.2.4 Data Mining	17

2.2.5	Python	19
2.2.6	Text mining	20
2.2.6.1	<i>Text Preprocessing</i>	20
2.2.7	Pembobotan Kata (<i>Term Weighting</i>).....	23
2.2.8	Analisis Sentimen	25
2.2.9	Klasifikasi	26
2.2.9.1	Ukuran Evaluasi Model.....	26
2.2.9.2	<i>K-Fold Cross Validation</i>	29
2.2.10	Maximum Entropy	30
2.2.10.1	Definisi <i>Entropy</i>	30
2.2.10.2	Prinsip <i>Maximum Entropy</i>	32
2.2.10.3	Algoritma Klasifikasi dengan <i>Maximum Entropy</i>	32
2.2.11	Wordcloud.....	33
2.2.12	Asosiasi Text.....	34
BAB III METODOLOGI PENELITIAN		37
3.1.	Alat Penelitian	37
3.1.1	Spesifikasi Kebutuhan Perangkat Keras (<i>Hardware</i>)	37
3.1.2	Spesifikasi Kebutuhan Perangkat Lunak (<i>Software</i>).....	37
3.2.	Populasi Penelitian	37
3.3.	Jenis dan Sumber Data	37
3.4.	Variabel Penelitian	37
3.5.	Diagram Alir Penelitian.....	38
3.5.1.	Identifikasi Masalah	38
3.5.2.	Studi Pustaka.....	39
3.5.3.	Pengumpulan Data	39
3.5.4.	Analisis Deskriptif	39
3.5.5.	Pelabelan Data.....	39
3.5.6.	<i>Data Preprocessing</i>	41
3.5.6.1	Spelling Normalization	41
3.5.6.2	<i>Case Folding</i>	42
3.5.6.3	<i>Cleansing</i>	42
3.5.6.4	<i>Stemming</i>	42
3.5.6.5	<i>Tokenizing</i>	43

3.5.6.6	<i>Filtering</i>	43
3.5.7.	Pembobotan Kata	44
3.5.8.	<i>Split Data</i>	45
3.5.9.	Klasifikasi	45
3.5.10.	Evaluasi	45
3.5.11.	Pembuatan <i>Wordcloud</i>	46
3.5.12.	Asosiasi Kata	46
3.5.13.	Pembuatan Dashboard	46
3.6.	Jadwal Pelaksanaan	46
BAB IV HASIL DAN PEMBAHASAN		48
4.1.	Pengumpulan Data	48
4.2.	Analisis Deskriptif.....	49
4.3.	Pelabelan Data	53
4.4.	<i>Preprocessing Data</i>	56
4.4.1.	<i>Preprocessing Data</i>	56
4.4.2.	<i>Text Preprocessing</i>	56
4.4.2.1.	<i>Case Folding</i>	56
4.4.2.2.	<i>Cleansing</i>	57
4.4.2.3.	<i>Tokenizing</i>	58
4.4.2.4.	<i>Spelling Normalization</i>	58
4.4.2.5.	<i>Stemming</i>	59
4.4.2.6.	<i>Filltering</i>	60
4.5.	Pembobotan Kata	60
4.6.	Proses Klasifikasi	63
4.6.1.	<i>Split Data</i>	63
4.6.2.	Klasifikasi	64
4.7.	Evaluasi	68
4.7.1.	Performa Analisis.....	68
4.7.2.	Prediksi.....	72
4.8.	Asosiasi Teks	74
4.8.1.	Sentimen Positif	74
4.8.2.	Sentimen Negatif.....	76
4.9.	Pembuatan Dashboard & <i>Insight</i>	79

4.9.1. Dashboard	79
4.9.2. <i>Insight</i>	80
4.9.3. Rekomendasi.....	86
BAB V PENUTUP.....	90
5.1. Kesimpulan.....	90
5.2. Saran.....	92
DAFTAR PUSTAKA.....	94
RIWAYAT HIDUP.....	99
LAMPIRAN.....	100

DAFTAR GAMBAR

Nomor	Judul Gambar	Halaman
Gambar 2.1	Tahapan <i>knowledge discovery in database (KDD)</i>	19
Gambar 2.2	Contoh <i>Wordcloud</i>	34
Gambar 3.1	Diagram Alir Penelitian.....	38
Gambar 3.2	<i>Review</i> Aplikasi.....	40
Gambar 4.1	Ulasan pengguna aplikasi Indomaret poinku	48
Gambar 4.2	Kode <i>Scrapping data</i>	48
Gambar 4.3	Hasil <i>Scrapping data</i>	49
Gambar 4.4	Diagram Jumlah Ulasan Aplikasi.....	49
Gambar 4.5	<i>Campaign</i> Indomaret Febuari dan Maret	50
Gambar 4.6	Diagram <i>Rating</i> Pengguna Aplikasi	51
Gambar 4.7	Diagram Proporsi Jumlah <i>Rating</i> Pengguna Per Bulan	52
Gambar 4.8	Hasil Pelabelan tahap 1	54
Gambar 4.9	Hasil Pelabelan tahap 2	55
Gambar 4.10	Data dengan nilai <i>null</i> dan data yang kurang relevan	56
Gambar 4.11	Hasil Perhitungan TF-IDF Dokumen 1, Dokumen 2, dan Dokumen 3	62
Gambar 4.12	Kode dan hasil Python untuk TF-IDF	63
Gambar 4.13	Konsep <i>K-Fold Cross Validation</i>	64
Gambar 4.14	Klasifikasi Model <i>Maxent</i> dengan TF-IDF Dengan Python	65
Gambar 4.15	Klasifikasi Model <i>Maxent</i> dengan TF-IDF Dengan Python	66
Gambar 4.16	Perbandingan Klasifikasi model dengan dataset split	67
Gambar 4.17	<i>Confusion Matrix</i> iterasi 1	70
Gambar 4.18	Prediksi Manual	73
Gambar 4.19	Prediksi <i>File CSV</i>	74
Gambar 4.20	<i>Dashboard</i> 1 Sentimen Analisis – Aplikasi Indomaret Poinku.....	79
Gambar 4.21	<i>Dashboard</i> 2 Sentimen Analisis – Aplikasi Indomaret Poinku.....	80
Gambar 4.22	<i>Line Chart</i>	80
Gambar 4.23	<i>Donut Chart</i>	81
Gambar 4.24	<i>Bar Chart</i>	82
Gambar 4.25	<i>Word Cloud</i> Positif.....	82
Gambar 4.26	<i>Word Cloud</i> Negatif	83
Gambar 4.27	Tabel <i>Supported Sentiment</i>	84

DAFTAR TABEL

Nomor	Judul Tabel	Halaman
Table 2.1	Perbandingan Penelitian Terdahulu	9
Table 2.2	<i>Confusion Matrix</i>	27
Table 2.3	<i>K-Fold Cross Validation</i>	29
Table 2.4	<i>Document Term Matrix</i>	35
Table 2.5	Perhitungan Korelasi Kata2 dan Kata5	36
Table 3.1	Tahap 1 Pelabelan Data	40
Table 3.2	Tahap 2 Pelabelan Data	41
Table 3.3	Proses <i>Spelling normalization</i>	41
Table 3.4	Proses <i>Case Folding</i>	42
Table 3.5	Proses <i>Cleansing</i>	42
Table 3.6	Proses <i>Stemming</i>	43
Table 3.7	Proses <i>Tokenizing</i>	43
Table 3.8	Proses <i>Filtering</i>	44
Table 3.9	Jadwal Pelaksanaan	46
Table 4.1	Hasil Pelabelan tahap 1	54
Table 4.2	Contoh <i>rating</i> dan respon pengguna	54
Table 4.3	Pelabelan data tahap 2	55
Table 4.4	Hasil Pelabelan tahap 2	56
Table 4.5	Proses <i>Case Folding</i>	57
Table 4.6	Proses <i>Cleansing</i>	57
Table 4.7	Proses <i>Tokenizing</i>	58
Table 4.8	Proses <i>Spelling Normalization</i>	58
Table 4.9	Proses <i>Stemming</i>	59
Table 4.10	Proses <i>Filtering</i>	60
Table 4.11	Dokumen 1,2, dan 3.....	61
Table 4.12	TF untuk tiap Token/kata pada Dokumen 1	61
Table 4.13	Akurasi Model <i>Maxent</i> dengan TF-IDF Tiap Iterasi	65
Table 4.14	Akurasi Model <i>Maxent</i> Tiap Iterasi	66
Table 4.15	Akurasi dengan dataset split	67
Table 4.16	Evaluasi Model Tanpa Pembobotan Kata	69
Table 4.17	Evaluasi Model Tanpa Pembobotan Kata	69
Table 4.18	<i>Confusion Matrix</i> iterasi 1	71
Table 4.19	Prediksi Manual.....	73
Table 4.20	Asosiasi Sentimen Positif.....	74
Table 4.21	Asosiasi Sentimen Negatif	77

DAFTAR LAMPIRAN

Nomor	Judul Tabel
1.	Lembar Hasil Turnitin
2.	<i>Dashboard Hasil</i>
3.	<i>Source Code Scrapping Data</i>
4.	<i>Source Code Preprocessing Data</i>
5.	<i>Source Code Asosiasi Kata</i>
6.	<i>Source Code Klasifikasi Sentimen</i>