



***TOPIC MODELING ULASAN NEGATIF PENGGUNA APLIKASI
MYMRTJ DENGAN METODE LATENT DIRICHLET ALLOCATION***

SKRIPSI

**MAULANA YUSUF
NIM. 2010512034**

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
2024**



***TOPIC MODELING ULASAN NEGATIF PENGGUNA APLIKASI
MYMRTJ DENGAN METODE LATENT DIRICHLET ALLOCATION***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana
komputer**

**MAULANA YUSUF
NIM. 2010512034**

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
2024**

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Maulana Yusuf
NIM : 2010512034
Tanggal : 16 April 2024

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 16 April 2024

Yang Menyatakan,



Maulana Yusuf

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini:

Nama : Maulana Yusuf
NIM : 2010512034
Fakultas : Ilmu Komputer
Program 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:

***TOPIC MODELING* ULASAN NEGATIF PENGGUNA APLIKASI
MYMRTJ DENGAN METODE *LATENT DIRICHLET ALLOCATION***

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

Dibuat di : Jakarta
Pada tanggal : 16 April 2024
Yang Menyatakan,



(Maulana Yusuf)

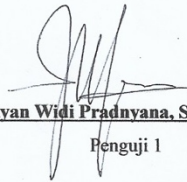
..

LEMBAR PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Maulana Yusuf
NIM : 2010512034
Program Studi : S1 – Sistem Informasi
Judul Tugas Akhir : *Topic Modeling* Ulasan Negatif Pengguna Aplikasi MyMRTJ Dengan Metode
Latent Dirichlet Allocation

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



I Wawan Widi Pradnyana, S.Kom., M.TI.

Penguji 1



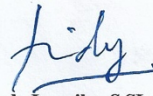
Zatin Niqotaini, S.Tr.Kom., M.Kom.

Penguji 2



Rio Wirawan, S.Kom, M.M.S.I.

Pembimbing 1



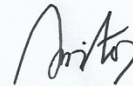
Nindy Irzavika, S.SI., M.T.

Pembimbing 2



Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM

Dekan FIK UPN Veteran Jakarta



Anita Muliawati, S.Kom., M.TI.

Ketua Program Studi S1 Sistem Informasi

Ditetapkan di : Jakarta

Tanggal Ujian : 3 April 2024

***TOPIC MODELING* ULASAN NEGATIF PENGGUNA
APLIKASI MYMRTJ DENGAN METODE *LATENT*
*DIRICHLET ALLOCATION***

Maulana Yusuf

ABSTRAK

Transportasi umum sudah menjadi alat mobilitas bagi sebagian orang untuk bepergian di kota besar seperti di DKI Jakarta. Penggunaan ini dibarengi dengan kebutuhan penggunaan aplikasi transportasi umum sebagai media mutakhir perkembangan moda transportasi umum terkait. Aplikasi MyMRTJ telah menjadi salah satu aplikasi terkemuka yang digunakan oleh pengguna MRT di Daerah Khusus Ibukota Jakarta, Indonesia. Namun, dalam penggunaan sehari-hari, tidak jarang adanya ulasan negatif dari pengguna pada platform Google Play Store dan App Store terkait dengan pengalaman mereka dalam menggunakan aplikasi MyMRTJ. Ulasan negatif ini mencakup beragam masalah hingga ketidakpuasan terhadap layanan pelanggan. Penting untuk melakukan analisis mendalam terhadap ulasan negatif pengguna aplikasi pada MyMRTJ menggunakan metode yang tepat. Dalam penelitian ini, penulis menggunakan metode *Latent Dirichlet Allocation* untuk menganalisis topik yang ada pada ulasan. Hasil penelitian yang dilakukan dengan melakukan iterasi sebanyak 2 hingga 15 topik dengan 10 percobaan acak, didapatkan nilai topik tertinggi sebanyak 3 topik dengan skor koherensi rata-rata 0,734 dan skor tertinggi sebesar 0,75. Dari interpretasi topik, dihasilkan 3 topik dengan topik Registrasi & Login, Metode Pembayaran, dan Ticketing & Sistem.

Kata Kunci: *Latent Dirichlet Allocation, MyMRTJ, Topic Modeling*

TOPIC MODELING OF NEGATIVE REVIEWS FROM MYMRTJ APPLICATION USERS USING LATENT DIRICHLET ALLOCATION METHOD

Maulana Yusuf

ABSTRACT

Public transportation has become a mobility tool for many people to travel in major cities such as DKI Jakarta. This usage is accompanied by the need for using public transportation application as a modern medium for related public transportation mode developments. The MyMRTJ application has become one of the leading applications used by MRT users in the Special Capital Region of Jakarta, Indonesia. However, in daily usage, there are often negative reviews from users on Google Play Store and the App Store regarding their experience with the MyMRTJ application. These negative reviews cover various issues up to dissatisfaction with customer service. It is important to conduct in-depth analysis of user's negative reviews of the MyMRTJ application using appropriate methods. In this research, the author used the Latent Dirichlet Allocation method to analyze the topics present in the reviews. The research results, conducted by iterating from 2 to 15 topics with 10 random experiments, yielded the highest topic number of 3 topics with an average coherence score of 0,734 and the highest score of 0,75. From the topic interpretation, 3 topics were generated: Registration & Login, Payment Methods, and Ticketing & System.

Keyword: *Latent Dirichlet Allocation, MyMRTJ, Topic Modeling*

KATA PENGANTAR

Dengan mengucapkan puji dan syukur ke hadirat Allah SWT atas segala rahmat, karunia dan hidayah-Nya, serta Shalawat dan salam tak lupa saya sebagai penulis ucapkan kepada Nabi Muhammad SAW, keluarga serta sahabatnya sehingga penulis dapat menyelesaikan skripsi ini dengan judul penelitian “*Topic Modeling* Ulasan Negatif Pengguna Aplikasi MyMRTJ Dengan Metode *Latent Dirichlet Allocation*”.

Skripsi ini dapat selesai berkat adanya bimbingan dan bantuan dari berbagai pihak, maka penulis dapat menyelesaikannya dengan baik. Untuk itu, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Tuhan Yang Maha Esa atas segala berkat dan karunianya
2. Kedua orang tua saya dan adik saya yang telah memberikan dukungan moril dan doa.
3. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM. selaku Dekan Fakultas Ilmu Komputer UPN Veteran Jakarta.
4. Ibu Anita Muliawati, S.Kom., MTI. selaku Ketua Jurusan S1 Sistem Informasi UPN Veteran Jakarta.
5. Bapak Rio Wirawan, S.Kom., M.M.S.I selaku Dosen Pembimbing I.
6. Ibu Nindy Irzavika, S.SI., M.T selaku Dosen Pembimbing II.
7. Teman-teman yang telah memberikan semangat serta dukungan hingga penulis menyelesaikan laporan ini.

Penulis menyadari bahwa penyusunan skripsi ini masih memiliki kekurangan, baik dari materi ataupun teknik penyajiannya. Oleh karena itu, diharapkan saran dan kritik yang membangun agar penulis dapat menjadi lebih baik lagi di masa mendatang.

Jakarta, 20 Maret 2024

Penulis

Maulana Yusuf

DAFTAR ISI

PERNYATAAN ORISINALITAS.....	i
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR.....	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian	3
1.5 Ruang Lingkup Penelitian.....	4
1.6 Luaran Yang Diharapkan	4
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 Aplikasi	6
2.2 <i>Mass Rapid Transit</i>	6
2.3 <i>Data Scraping</i>	7
2.4 <i>Text Mining</i>	7
2.5 <i>Topic Modeling</i>	8
2.6 <i>Data Labelling</i>	9
2.7 <i>Data Preprocessing</i>	9
2.7.1 <i>Case Folding</i>	9
2.7.2 <i>Remove Punctuation</i>	10
2.7.3 <i>Tokenization</i>	10
2.7.4 <i>Spelling Normalization</i>	11
2.7.5 <i>Lemmatization</i>	11
2.7.6 <i>Stopwords</i>	11
2.8 Ekstraksi Fitur	12
2.9 <i>Latent Dirichlet Allocation</i>	12

2.10	<i>Coherence Score</i>	14
2.11	Google Play Store	14
2.12	App Store	15
2.13	Python	15
2.14	<i>Dashboard</i>	16
2.15	<i>Wordcloud</i>	16
2.16	Python Streamlit.....	17
2.17	Penelitian Sebelumnya	17
BAB III METODE PENELITIAN.....		21
3.1	Alur Penelitian	21
3.1.1	Identifikasi Masalah	21
3.1.2	Studi Literatur	22
3.1.3	Pengumpulan Data	22
3.1.4	<i>Data Labeling</i>	22
3.1.5	<i>Data Preprocessing</i>	23
3.1.6	Ekstraksi Fitur	23
3.1.7	Pembuatan Model <i>Latent Diriclet Allocation</i>	23
3.1.8	Evaluasi Model.....	24
3.1.9	Implementasi <i>Dashboard</i> Web.....	25
3.1.10	Analisis dan Interpretasi Topik	25
3.2	Alat dan Bahan Penelitian.....	25
3.3	Jadwal Penelitian.....	26
BAB IV HASIL DAN PEMBAHASAN		28
4.1	Pengumpulan Data	28
4.2	<i>Data Labeling</i>	30
4.3	<i>Data Preprocessing</i>	33
4.3.1	<i>Case Folding</i>	33
4.3.2	<i>Remove Punctuation</i>	34
4.3.3	<i>Tokenization</i>	36
4.3.4	<i>Spelling Normalization</i>	37
4.3.5	<i>Lemmatization</i>	38
4.3.6	<i>Stopword Removal</i>	39
4.4	Ekstraksi Fitur	41
4.5	Pembuatan Model dengan <i>Latent Dirichlet Allocation</i>	41
4.6	Evaluasi Model.....	42
4.7	<i>Dashboard</i>	44

4.7.1	Halaman Utama.....	44
4.7.2	Halaman Statistik Peringkat.....	45
4.7.3	Halaman Statistik Ulasan.....	45
4.7.4	Halaman Uraian Topik.....	48
4.8	Analisis dan Interpretasi Topik.....	48
4.8.1	Interpretasi Topik 1.....	48
4.8.2	Interpretasi Topik 2.....	51
4.8.3	Interpretasi Topik 3.....	54
4.8.4	Hasil Klasterisasi Ulasan Berdasarkan Bobot.....	57
4.8.5	Persebaran Topik pada Ulasan.....	59
4.8.6	Distribusi Topik.....	62
4.8.7	Persebaran Antar Kata pada Ulasan.....	63
4.8.8	Pergerakan Kata pada Ulasan.....	64
4.8.9	Sentimen Ulasan.....	65
BAB V KESIMPULAN DAN SARAN.....		66
DAFTAR PUSTAKA.....		68
RIWAYAT HIDUP.....		71
LAMPIRAN.....		72

DAFTAR TABEL

Tabel 2. 1 Penelitian Sebelumnya	17
Tabel 3. 1 Jadwal Penelitian	26
Tabel 4. 1 Hasil terjemahan Inggris-Indonesia pada data ulasan	28
Tabel 4. 2 Hasil sentimen oleh verifikator	31
Tabel 4. 3 Hasil case folding ulasan pengguna	34
Tabel 4. 4 Hasil remove punctuation ulasan pengguna	35
Tabel 4. 5 Hasil tokenization ulasan pengguna	36
Tabel 4. 6 Hasil spell normalization ulasan pengguna	37
Tabel 4. 7 Hasil lemmatization ulasan pengguna	38
Tabel 4. 8 Hasil stopwords removal ulasan pengguna	40
Tabel 4. 9 Tabel perhitungan coherence score pada model	43
Tabel 4. 10 Topik hasil model LDA	44
Tabel 4. 11 Kumpulan kata pada topik 1	48
Tabel 4. 12 Ulasan pada topik 1	50
Tabel 4. 13 Kumpulan kata pada topik 2	51
Tabel 4. 14 Ulasan pada topik 2	53
Tabel 4. 15 Kumpulan kata pada topik 3	54
Tabel 4. 16 Ulasan pada topik 3	55
Tabel 4. 17 Hasil klasterisasi ulasan	57

DAFTAR GAMBAR

Gambar 2. 1 Proses dari text mining.....	8
Gambar 2. 2 Visualisasi Proses LDA	13
Gambar 3. 1 Alur Penelitian.....	21
Gambar 4. 1 Hasil scraping data melalui Google Play Store dan App Store	30
Gambar 4. 2 Distribusi rating data ulasan	30
Gambar 4. 3 Persentase total sentimen	33
Gambar 4. 4 Hasil ekstraksi fitur menggunakan CountVectorizer.....	41
Gambar 4. 5 Grafik coherence score pada model	43
Gambar 4. 6 Halaman utama dashboard	44
Gambar 4. 7 Halaman Statistik Peringkat	45
Gambar 4. 8 Halaman Statistik Ulasan 1	46
Gambar 4. 9 Halaman Statistik Ulasan 2	46
Gambar 4. 10 Halaman Statistik Ulasan 3	46
Gambar 4. 11 Halaman Statistik Ulasan 4	47
Gambar 4. 12 Halaman Statistik Ulasan 5	47
Gambar 4. 13 Halaman Statistik Ulasan 6	47
Gambar 4. 14 Halaman Uraian Topik	48
Gambar 4. 15 Bobot pada topik 1	49
Gambar 4. 16 Wordcloud topik 1	49
Gambar 4. 17 Bobot pada topik 2	52
Gambar 4. 18 Wordcloud topik 2	52
Gambar 4. 19 Bobot pada topik 3	54
Gambar 4. 20 Wordcloud topik 3	55
Gambar 4. 21 Distribusi topik keseluruhan.....	59
Gambar 4. 22 Klasterisasi pada topik	59
Gambar 4. 23 Jangkauan pada ulasan Document 1102	60
Gambar 4. 24 Jangkauan pada ulasan Document 1305.....	60
Gambar 4. 25 Bobot topik Document 1305	61
Gambar 4. 26 Document 915 pada klasterisasi.....	61
Gambar 4. 27 Distribusi topik pada tahun 2023	62
Gambar 4. 28 Grafik topik pada tahun 2023.....	62
Gambar 4. 29 Visualisasi persebaran antar kata pada data ulasan	63
Gambar 4. 30 Visualisasi antar kata stasiun - tiket - beli	64
Gambar 4. 31 Pergerakan kata berdasarkan waktu dan topik	64
Gambar 4. 32 Sentimen Positif	65
Gambar 4. 33 Sentimen Negatif.....	65