

**ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI PEDULILINDUNGI
MENGUNAKAN LATENT DIRICHLET ALLOCATION DAN SUPPORT
VECTOR MACHINE**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh :

Ana Rohmah Zaidah

NIM : H76217050

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA
2021**

PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini,

Nama : Ana Rohmah Zaidah

NIM : H76217050

Program Studi : Sistem Informasi

Angkatan : 2017

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul : “ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI PEDULILINDUNGI MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN SUPPORT VECTOR MACHINE “. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian Pernyataan keaslian ini saya buat dengan sebenar - benarnya.

Surabaya, 10 Januari 2022

Yang menyatakan,



(ANA ROHMAH ZAIDAH)

H76217050

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

NAMA : Ana Rohmah Zaidah

NIM : H76217050

JUDUL : ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI
PEDULILINDUNGI MENGGUNAKAN LATENT DIRICHLET
ALLOCATION DAN SUPPORT VECTOR MACHINE

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 29 November 2021

Menyetujui,

Dosen pembimbing I



(Mujib Ridwan, S.Kom.,MT)
NIP.198604272014031004

Dosen pembimbing II



(Ahmad Yusuf, M.Kom)
NIP.199001202014031003

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Ana Rohmah Zaidah ini telah dipertahankan
di depan Tim Penguji Skripsi
Surabaya, 21 Desember 2021.

Mengesahkan,
Dewan Penguji

Dosen Penguji I



Dwi Rolliawati.MT
NIP.197909272014032001

Dosen Penguji II



Khalid.M.Kom
NIP.197906092014031002

Dosen Penguji III



Mujib Ridwan.S.Kom..M.T
NIP.198604272014031004

Dosen Penguji IV



Ahmad Yusuf.M.Kom
NIP.199001202014031003

Mengetahui

Dekan Fakultas Sains dan Teknologi
UN Sunan Ampel Surabaya




Prof. Dr. H. Evi Fatimatur Rusydiyah, M.Ag
NIP.197312272005012003



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpustakaan@uinsby.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : ANA ROHMAH ZAIDAH
NIM : H76217050
Fakultas/Jurusan : SAINS DAN TEKNOLOGI/SISTEM INFORMASI
E-mail address : anarohmahzaidah77@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Sekripsi Tesis Desertasi Lain-lain

(.....)

yang berjudul :

ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI
PEDULILINDUNGI MENGGUNAKAN DIRICHLET ALLOCATION DAN SUPPORT
VECTOR MACHINE.

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 10 Januari 2022

Penulis

(ANA ROHMAH ZAIDAH)

				berdasarkan probabilitas tertinggi.
2	(Astuti, 2020)	Analisis sentiment berbasis aspek pada aplikasi Tokopedia menggunakan LDA dan <i>naïve bayes</i> .	Penentuan jumlah aspek didapat dari <i>clustering</i> topik LDA berdasarkan <i>aspect-based sentiment analisi</i> , didapat 4 aspek yaitu kebermanfaatan, pelayanan, pengalaman belanja, dan tampilan. Output aspek LDA digunakan untuk anotasi aspek kemudian diklasifikasikan dengan <i>naïve bayes</i> untuk menentukan masing – masing sentimen aspek.	Penelitian yang sudah dilakukan melakukan pengembangan yang sebagaimana ditulis di saran penelitian ini, salah satunya menggunakan suatu teknik lain agar data diurutkan berdasarkan komentar terbaru, dan melakukan <i>treatment</i> pada proses <i>scraping</i> web ulasan agar mendapatkan data berdasarkan komentar yang terbaru.
3	(Putra, 2017b)	Analisis topik informasi publik media sosial di Surabaya menggunakan pemodelan <i>latent dirichlet</i>	Berdasarkan hasil penelitian menunjukkan topik 4 merupakan jumlah topik terbaik dalam membentuk topik model didasarkan pada nilai <i>perplexity</i>	Hasil penelitian menunjukkan pengaruh positif terhadap pemodelan topik dan penelitian sebelumnya menyarankan

		<i>allocation</i> (LDA)	dan proses <i>stemming</i> sangat berpengaruh positif terhadap pembentukan model.	untuk menormalisasi sebelum melakukan pemodelan topik, sehingga pada penelitian selanjutnya menggunakan tambahan slang word lexicon berisi kata – kata gaul yang sudah dinormalisasikan dan dibuat sendiri disesuaikan dataset.
4	(Kabiru & Sari, 2019)	Analisa konten media sosial <i>E-commerce</i> pada Instagram menggunakan metode sentiment analysis dan LDA-based topik modeling (studi kasus : Shopee Indonesia)	Hasil dari penelitian mendapatkan 6 topik utama dan melakukan analisis sentimen terhadap topik modeling yang berdasarkan sentimen positif dan negatif.	Hasil dari penelitian sudah menghasilkan banyak saran dan <i>insight</i> untuk kelemahan dan kelebihan aplikasi, akan tetapi belum ada pengujian tentang penentuan topik, proses penamaan topik hasil <i>clustering</i> .
5	(Firdaus et al., 2020)	Analisis sentiment dan	Penelitian ini bertujuan untuk	Pada penelitian sudah menjelaskan

		topik modeling dalam aplikasi ruang guru	mengetahui kepuasan pelanggan yang datanya diambil dari <i>website</i> youtube dengan output datanya tentang sentimen positif dan negatif jenjang pendidikan.	masing – masing topik, tapi belum dijelaskan secara spesifik hasil <i>clustering</i> topik dan pendekatan algoritma sentimen yang digunakan.
6	(Onan et al., 2016)	<i>LDA-based topic modeling in text sentiment classification : An empirical analysis</i>	Pada penelitian ini memeriksa kinerja prediktif dari algoritma klasifikasi (<i>naïve bayes</i> , SVM, <i>regresi logistic</i> , <i>radial basis function network</i> dan <i>k-nearest neighbor</i> algoritma untuk klasifikasi sentiment teks ketika representasi berbasis LDA	Pada penelitian yang akan dilakukan menggunakan metode LDA dan SVM.
7	(Ye et al., 2017)	<i>Sentiment analysis using modified LDA</i>	Penelitian yang dilakukan top modeling dengan LDA dan sentimen analisis dengan SVM dengan data <i>review</i> film,dengan menunjukkan hasil <i>clustering</i> topik 10 -	Berdasarkan hasil penelitian menunjukkan pengaruh yang baik antara LDA dan sentimen analisis, kemudian untuk penelitian yang akan

			50 dan precision sampai 86%.	dilakukan menggunakan objek serta <i>treatment data processing</i> yang berbeda dan belum dijelaskan pada penelitian sebelumnya.
8	(Ilmawan, 2020)	Perbandingan metode klasifikasi <i>support vector machine</i> dan <i>naïve bayes</i> untuk analisis sentimen pada ulasan tekstual di <i>google play store</i> .	Penelitian dilakukan dengan membandingkan antara algoritma <i>svm</i> dan <i>naïve bayes</i> dengan tujuan melihat tingkat <i>accuracy</i> yang lebih tinggi, yang dimana kesimpulan penelitian menunjukan <i>svm</i> memiliki tingkat <i>accuracy</i> yang lebih tinggi	Berdasarkan hasil <i>accuracy</i> yang tinggi <i>SVM</i> bisa melakukan proses sentimen ulasan pengguna dengan baik, sehingga memperkuat dasar menggunakan algoritma <i>SVM</i> .
9	(Fransiska et al., 2020)	<i>Sentiment analysis provider by.U on google play store reviews TF.IDF and support vector machine</i>	Hasil penelitian dari review aplikasi by.U di <i>google play store</i> menggunakan <i>SVM</i> dengan proses penentuan N-Fold sampai 5 kemudian menghasilkan rata –	Penelitian menghasilkan bahwa <i>SVM</i> mempunyai <i>accuracy</i> yang baik sehingga pada penelitian yang akan dilakukan

		<i>(SVM) method.</i>	rata <i>accuracy</i> yang tinggi yaitu 83,3%	menggunakan skenario pengujian yang sama tapi berbeda tujuan penelitian.
10.	(Mustopa et al., 2020)	<i>Analysis of user reviews for the Pedulilindungi application on google play using the support vector machine and naïve bayes algorithm on particle swarm optimization</i>	Pada penelitian yang sudah dilakukan bertujuan untuk membandingkan proses algoritma SVM dan <i>naïve bayes</i> dengan bantuan algoritma <i>particle swarm optimization</i> untuk meningkatkan <i>accuracy</i> , dan menunjukkan <i>accuracy</i> SVM 93.00% lebih tinggi dari pada <i>naïve bayes</i> yaitu 69,00%.	Berdasarkan hasil penelitian yang menunjukkan <i>accuracy</i> yang lebih baik dengan algoritma SVM maka penelitian yang dilakukan menggunakan algoritma SVM.
11	(Akhtar et al., 2017)	<i>Aspect based Sentiment Oriented Summarization of Hotel Reviews</i>	Penelitian berbasis aspek menggunakan ulasan restoran, penentuan nama aspek dilakukan dengan menggunakan topword/ kata yang sering muncul dari setiap clustering,	Penelitian yang sudah dilakukan menggunakan acuan probabilitas kata yang mewakili setiap topik, kemudian hitung dominasi menggunakan library LDA

2	0.1	1.0	linear	7.093.393.0 57.110.860	7.129.433.475 .116.260
3	0.1	10.0	linear	7.093.393.0 57.110.860	7.129.433.475 .116.260
4	0.1	100.0	linear	7.093.393.0 57.110.860	7.129.433.475 .116.260
5	1.0	0.01	linear	8.784.210.5 26.315.780	9.915.433.651 .655.330
6	1.0	0.1	linear	8.784.210.5 26.315.780	9.915.433.651 .655.330
7	1.0	1.0	linear	8.784.210.5 26.315.780	9.915.433.651 .655.330
8	1.0	10.0	linear	8.784.210.5 26.315.780	9.915.433.651 .655.330
9	1.0	100.0	linear	8.784.210.5 26.315.780	9.915.433.651 .655.330
10	10.0	0.01	linear	8.795.072.7 88.353.860	9.992.953.608 .844.600
11	10.0	0.1	linear	8.795.072.7 88.353.860	9.992.953.608 .844.600
12	10.0	1.0	linear	8.795.072.7 88.353.860	9.992.953.608 .844.600
13	10.0	10.0	linear	8.795.072.7 88.353.860	9.992.953.608 .844.600
14	10.0	100.0	linear	8.795.072.7 88.353.860	9.992.953.608 .844.600
15	100.0	0.01	linear	8.710.414.3 33.706.600	9.990.603.432 .581.380
16	100.0	0.1	linear	8.710.414.3 33.706.600	9.990.603.432 .581.380
17	100.0	1.0	linear	8.710.414.3 33.706.600	9.990.603.432 .581.380
18	100.0	10.0	linear	8.710.414.3 33.706.600	9.990.603.432 .581.380
19	100.0	100.0	linear	8.710.414.3 33.706.600	9.990.603.432 .581.380
20	1000.0	0.01	linear	8.562.486.0 02.239.640	1.0
21	1000.0	0.1	linear	8.562.486.0 02.239.640	1.0
22	1000.0	1.0	linear	8.562.486.0 02.239.640	1.0
23	1000.0	10.0	linear	8.562.486.0 02.239.640	1.0
24	1000.0	100.0	linear	8.562.486.0 02.239.640	1.0

- Awad, M., & Khanna, R. (2015). *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*. Springer nature.
- COVID-19, satgas penanganan covid. (2020). [SALAH] WHO Sebut Penularan Corona Tak Lagi Hanya Lewat Droplet Tapi Juga Udara—Hoax Buster / Satgas Penanganan COVID-19 [Pemerintah]. covid19.go.id. <https://covid19.go.id/p/hoax-buster/salah-who-sebut-penularan-corona-tak-lagi-hanya-lewat-droplet-tapi-juga-udara>
- Create Dataset for Sentiment Analysis by Scraping Google Play App Reviews using Python / Curiously—Hacker's Guide to Machine Learning*. (2021). <https://curiously.com/posts/create-dataset-for-sentiment-analysis-by-scraping-google-play-app-reviews-using-python/>
- Firdaus, M. R., Rizki, F. M., Gaus, F. M., & Susanto, I. K. (2020). Analisis Sentimen Dan Topic Modelling Dalam Aplikasi Ruangguru. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 4(1), 66–76.
- Fransiska, S., Rianto, R., & Gufroni, A. I. (2020). Sentiment Analysis Provider By. U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method. *Scientific Journal of Informatics*, 7(2), 203–212.
- Ginanti, M. S., & Evayani, E. (2016). Perancangan Basis Data Pendapatan Cargo dan Pencatatan Data Pelanggan Menggunakan Visual Basic 6.0 pada PT Intan Lubuk Pertama. *Jurnal Ilmiah Mahasiswa Ekonomi Akuntansi*, 1(2), 30–44.
- Ilmawan, L. B. (2020). *Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store*. 12(2), 8.

- Onan, A., Korukoglu, S., & Bulut, H. (2016). LDA-based Topic Modelling in Text Sentiment Classification: An Empirical Analysis. *Int. J. Comput. Linguistics Appl.*, 7(1), 101–119.
- Paramadita, S., Umar, A., & Kurniawan, Y. J. (2020). Analisa PES^{TEL} terhadap Penetrasi Gojek di Indonesia. *Jurnal Pengabdian Dan Kewirausahaan*, 4(1).
- PeduliLindungi*. (2021). <https://pedulilindungi.id/>
- Putra, I. (2017a). *Analisis Topik Informasi Publik Media Sosial di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (LDA)* [PhD Thesis]. Institut Teknologi Sepuluh Nopember.
- Putra, I. (2017b). *Analisis Topik Informasi Publik Media Sosial di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (LDA)* [PhD Thesis]. Institut Teknologi Sepuluh Nopember.
- Suyanto, D. & others. (2017). Data Mining untuk klasifikasi dan klusterisasi data. *Bandung: Informatika Bandung*.
- Thet, T. T., Na, J.-C., & Khoo, C. S. (2010). Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science*, 36(6), 823–848.
- Vaksin Dashboard*. (2021). <https://vaksin.kemkes.go.id/#/vaccines>
- Wahyurianto, F., Arwani, I., & Soebroto, A. A. (2019). Pembangunan Aplikasi Informasi Kesehatan Masyarakat Kota Malang Berbasis Mobile Native Android. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer* E-ISSN, 2548, 964X.

