



SKRIPSI

PREDIKSI SENTIMEN DAN PEMODELAN TOPIK DARI ULASAN APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL (IKD)

PRIMANDIKA HAKIKI

NPM 18082010036

DOSEN PEMBIMBING

Amalia Anjani Arifiyanti, S.Kom., M.Kom.

Dhian Satria Yudha Kartika, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA
2025**



SKRIPSI

PREDIKSI SENTIMEN DAN PEMODELAN TOPIK DARI ULASAN APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL (IKD)

PRIMANDIKA HAKIKI

NPM 18082010036

DOSEN PEMBIMBING

Amalia Anjani Arifiyanti, S.Kom., M.Kom.

Dhian Satria Yudha Kartika, S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA
2025**

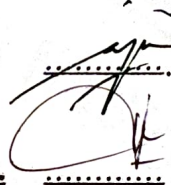
LEMBAR PENGESAHAN

PREDIKSI SENTIMEN DAN PEMODELAN TOPIK DARI ULASAN APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL (IKD)

Oleh:
PRIMANDIKA HAKIKI
NPM. 18082010036

Telah dipertahankan dihadapan dan diterima oleh Tim Penguju Skripsi Prodi
Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional
Veteran Jawa Timur Pada tanggal 04 Juni 2025

Amalia Anjani Arifiyanti, S.Kom., M.Kom.
NIP. 19920812 201803 2 001

 (Pembimbing I)

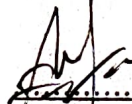
Dhian Satria Yudha Kartika, S.Kom., M.Kom.
NPT. 201 19860522 249

 (Pembimbing II)

Rizka Hadiwiyantri, S.Kom., M.Kom., MBA.
NIP. 19860727 201803 2 001

 (Ketua Penguji)

Anindo Saka Fitri, S.Kom., M.Kom.
NIP. 19930325 202406 2 001

 (Penguji II)

Iqbal Ramadhani Mukhlis, S.Kom., M.Kom.
NIP. 19930305 202406 1 002

 (Penguji III)

Mengetahui,
Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

“Halaman ini sengaja dikosongkan.”

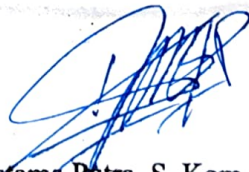
LEMBAR PERSETUJUAN

PREDIKSI SENTIMEN DAN PEMODELAN TOPIK DARI ULASAN APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL (IKD)

Oleh:
PRIMANDIKA HAKIKI
NPM. 18082010036

Telah disetujui untuk mengikuti Ujian Skripsi

Mengetahui,
Koordinator Program Studi Sistem Informasi
Fakultas Ilmu Komputer



Agung Brastama Putra, S. Kom., M. Kom.
NIP. 19851124 2021211 003

“Halaman ini sengaja dikosongkan.”

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan dibawah ini:

Nama Mahasiswa : Primandika Hakiki
NPM : 18082010036
Program : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 04 Juni 2025
Yang Membuat Pernyataan,



PRIMANDIKA HAKIKI
NPM.18082010036

“Halaman ini sengaja dikosongkan.”

ABSTRAK

Nama Mahasiswa / NPM : Primandika Hakiki / 18082010036
Judul Skripsi : Prediksi Sentimen dan Pemodelan Topik Dari
Ulasan Aplikasi Identitas Kependudukan
Digital (IKD)
Dosen Pembimbing : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

Di era yang sudah serba digital seperti saat ini, membuat pemerintah mau tidak mau untuk melakukan digitalisasi dalam penyediaan layanan terhadap masyarakat. Contoh digitalisasi yang dilakukan oleh pemerintah Indonesia adalah pada Kartu Tanda Penduduk (KTP). Pada tahun 2023, pemerintah meluncurkan Identitas Kependudukan Digital (IKD) sebagai kelanjutan langkah digitalisasi pada KTP. Berdasarkan pantauan pada Google PlayStore dan Apple AppStore, aplikasi IKD memiliki ulasan yang beragam. Ada yang memuji karena proses aktivasi cepat, namun ada juga yang memberikan ulasan buruk karena aktivasi aplikasi harus datang ke kecamatan ataupun Dispendukcapil. Skripsi ini bertujuan untuk melakukan prediksi sentimen dengan menggunakan metode Naive Bayes dan PNN, serta pemodelan topik dengan menggunakan metode LDA. Hasil prediksi sentimen menunjukkan skenario keempat memiliki akurasi paling tinggi dengan 97,12%, serta f1-score nilai positif 97% dan nilai negatif 97%. Hasil pemodelan topik menunjukkan nilai koherensi terbaik untuk kelas positif diperoleh topik 2 dengan nilai 0,4076. dan untuk kelas negatif diperoleh topik 10 dengan nilai 0,5564. Interpretasi hasil pemodelan topik ulasan positif meliputi verifikasi identitas, saran pengembangan aplikasi, kemudahan penggunaan, digitalisasi KTP, dan pengalaman penggunaan (UX). Sedangkan untuk ulasan negatif meliputi kendala teknis aplikasi, akses & instalasi aplikasi, kendala jaringan, fitur dan keamanan data, serta pendaftaran aplikasi.

Kata Kunci: IKD, prediksi sentimen, pemodelan topik, NB, PNN, LDA

“Halaman ini sengaja dikosongkan.”

ABSTRACT

Student Name / NPM : Primandika Hakiki / 18082010036

Thesis Title : Sentiment Prediction and Topic Modeling from
Identitas Kependudukan Digital (IKD) Application
Reviews

Advisor : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

In the current digital era, the government has no choice but to digitize the provision of services to the public. An example of digitalization carried out by the Indonesian government is the Population Identity Card (KTP). In 2023, the government launched the Digital Population Identity (IKD) as a continuation of the digitalization steps on the KTP. Based on monitoring on the Google PlayStore and Apple AppStore, the IKD application has various reviews. Some praise it because the activation process is fast, but some give bad reviews because the application activation must come to the sub-district or Dispendukcapil. This thesis aims to predict sentiment using the Naive Bayes and PNN methods, as well as topic modeling using the LDA method. The results of the sentiment prediction show that the fourth scenario has the highest accuracy with 97.12%, and a positive f1-score of 97% and a negative value of 97%. The results of topic modeling show that the best coherence value for the positive class is topic 2 with a value of 0.4076. and for the negative class, topic 10 is obtained with a value of 0.5564. Interpretation of positive review topic modeling results includes identity verification, application development suggestions, ease of use, ID card digitization, and user experience (UX). While negative reviews include application technical constraints, application access & installation, network constraints, features and data security, and application registration..

Keywords: IKD, sentiment prediction, topic modeling, NB, PNN, LDA

“Halaman ini sengaja dikosongkan.”

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas limpahan rahmat, karunia, serta hidayah-Nya, sehingga penulis dapat menyelesaikan seluruh rangkaian laporan skripsi dengan judul "Prediksi Sentimen dan Pemodelan Topik dari Ulasan Aplikasi Identitas Kependudukan Digital (IKD)". Skripsi ini diajukan untuk memenuhi persyaratan dalam memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Penulis menyadari bahwa kelancaran dan penyelesaian skripsi ini tidak terlepas dari bimbingan, kontribusi, dan doa dari banyak pihak. Maka dari itu, dengan segala kerendahan hati, penulis mengucapkan ucapan terima kasih kepada:

1. Kedua orang tua penulis yang tak henti-hentinya memberikan dukungan doa, moril, materiel, serta motivasi dalam setiap tahap penyelesaian skripsi ini.
2. Bapak Agung Brastama Putra, S.Kom., M.Kom. selaku koordinator program studi Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom. selaku dosen pembimbing I yang telah meluangkan waktu untuk membimbing dan mengarahkan penulis sehingga seluruh laporan skripsi ini dapat diselesaikan.
4. Bapak Dhian Satria Yudha Kartika, S.Kom., M.Kom. selaku dosen pembimbing II yang telah membimbing, mengarahkan, dan memberikan saran kepada penulis dalam proses pengerjaan skripsi ini.
5. Segenap jajaran dosen Program Studi Sistem Informasi yang telah memberikan bimbingan serta mencurahkan ilmu dan pengetahuannya selama masa studi penulis.
6. Kepada orang yang paling dekat dengan saya yang selalu memberikan support, doa, motivasi serta sebagai pendorong saya untuk menyelesaikan skripsi penulis.

7. Teman-teman kelas perkuliahan yang selalu mendorong dan memberikan semangat kepada penulis.
8. Berbagai pihak yang tidak dapat penulis sebutkan satu per satu, atas segala bentuk bantuan dan dukungan yang telah diberikan dalam penyelesaian skripsi ini.

Dalam penyusunan laporan ini, penulis mengakui bahwa masih terdapat berbagai keterbatasan dan ketidaksempurnaan, baik dari aspek isi pembahasan maupun kaidah penulisan. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf atas segala kekurangan tersebut. Semoga laporan ini dapat memberi manfaat bagi penulis secara pribadi, serta bagi para pembaca dan kemajuan ilmu pengetahuan pada umumnya.

Surabaya, 04 Juni 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xv
DAFTAR TABEL	xix
DAFTAR GAMBAR.....	xxi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	5
1.4 Tujuan	5
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA.....	9
2.1 Dasar Teori.....	9
2.1.1 App Store	9
2.1.2 Text Mining.....	10
2.1.3 Analisis Sentimen	10
2.1.4 Identitas Kependudukan Digital (IKD).....	10
2.1.5 Text Preprocessing.....	11
2.1.6 Naïve Bayes	12
2.1.7 Probabilistic Neural Network (PNN).....	12
2.1.8 Latent Dirichlet Allocation (LDA).....	14
2.1.9 TF-IDF	15
2.1.10 TF-RF.....	16
2.1.11 Confusion Matrix	16

2.2	Penelitian Terdahulu	18
BAB III METODOLOGI PENELITIAN		23
3.1	Studi Literatur	24
3.2	Analisis Kebutuhan	24
3.2.1	Kebutuhan Data.....	24
3.2.2	Kebutuhan Software dan Hardware	24
3.3	Pengumpulan Data	24
3.4	Penyaringan Data	28
3.5	Pelabelan Data.....	28
3.6	Data Preprocessing.....	29
3.7	Pembagian Data untuk Pembuatan Model	31
3.7.1	Pembobotan TF-IDF	31
3.7.2	Pembobotan TF-RF	32
3.8	Pembagian Data berdasarkan Kategori Sentimen	32
3.9	Prediksi Sentimen	32
3.10	Pemodelan Topik	32
3.11	Evaluasi Model.....	33
3.12	Hasil Ulasan	34
3.13	Visualisasi	34
3.14	Pembuatan Laporan.....	34
BAB IV HASIL DAN PEMBAHASAN		35
4.1	Implementasi Kebutuhan	35
4.1.1	Kebutuhan Data.....	35
4.1.2	Kebutuhan Hardware dan Software	35
4.2	Pembangunan Model.....	36
4.2.1	Pengumpulan Data	36
4.2.2	Penyaringan Data	38
4.2.3	Pelabelan Data.....	39
4.2.4	Preprocessing Data.....	42
4.2.5	Pembagian Data Untuk Pembuatan Model	52

4.2.6	Pembagian Data Berdasarkan Kategori Sentimen	53
4.2.7	Pembobotan TF-IDF	54
4.2.8	Pembobotan TF-RF	55
4.2.9	Pemodelan Sentimen Naïve Bayes	56
4.2.10	Pemodelan Sentimen Probabilistic Neural Network (PNN)	59
4.2.11	Pemodelan Topik Latent Dirichlet Allocation (LDA)	63
4.2.12	Evaluasi Model Naïve Bayes dan PNN	68
4.2.13	Evaluasi Model Latent Dirichlet Allocation (LDA)	73
4.2.14	Pembahasan Hasil	74
4.3	Visualisasi	78
BAB V PENUTUP		87
5.1	Kesimpulan	87
5.2	Saran	88
DAFTAR PUSTAKA		89

“Halaman ini sengaja dikosongkan.”

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	17
Tabel 2. 2 Penelitian Terdahulu.....	18
Tabel 3. 1 Contoh Data dari Apple AppStore	25
Tabel 3. 2 Contoh Data dari Google PlayStore	25
Tabel 4. 1 Hasil Penyaringan Data	38
Tabel 4. 2 Hasil Pelabelan Data	39
Tabel 4. 3 Contoh Hasil <i>Case Folding</i>	42
Tabel 4. 4 Contoh Hasil Normalisasi Kata	44
Tabel 4. 5 Contoh Hasil <i>Cleansing</i>	45
Tabel 4. 6 Contoh Hasil <i>Tokenizing</i>	47
Tabel 4. 7 Contoh Hasil <i>Stopword Removal</i>	49
Tabel 4. 8 Contoh Hasil Stemming	51
Tabel 4. 9 Hasil Perhitungan Nilai Koherensi.....	73
Tabel 4. 10 Hasil Seluruh Pemodelan Sentimen	74
Tabel 4. 11 Nilai dari Topik Koherensi.....	76
Tabel 4. 12 Hasil Klasifikasi Web Aplikasi	82
Tabel 4. 13 Hasil Topik Web Aplikasi.....	84

“Halaman ini sengaja dikosongkan.”

DAFTAR GAMBAR

Gambar 1. 1 Aplikasi Identitas Kependudukan Digital (IKD)	2
Gambar 2. 1 Arsitektur <i>Probabilistic Neural Network</i> (PNN) [17]	13
Gambar 2. 2 Representasi Model LDA	15
Gambar 3. 1 Metodologi Penelitian.....	23
Gambar 3. 2 Diagram Alir Data <i>Preprocessing</i>	29
Gambar 3. 3 Diagram Skenario	33
Gambar 4. 1 <i>Source Code Scraping</i> Data PlayStore	36
Gambar 4. 2 <i>Source Code Scraping</i> Data AppStore	36
Gambar 4. 3 Contoh Hasil <i>Scraping</i> Data Google PlayStore.....	37
Gambar 4. 4 Contoh Hasil <i>Scraping</i> Data AppStore.....	37
Gambar 4. 5 <i>Source Code</i> Pada Tahap <i>Case Folding</i>	42
Gambar 4. 6 <i>Source Code</i> Pada Tahap Normalisasi Kata	43
Gambar 4. 7 <i>Source Code</i> Pada Tahap <i>Cleansing</i>	45
Gambar 4. 8 <i>Source Code</i> Pada Tahap <i>Tokenizing</i>	47
Gambar 4. 9 <i>Source Code</i> Pada Tahap <i>Stopword Removal</i>	49
Gambar 4. 10 <i>Source Code</i> Pada Tahap <i>Stemming</i>	51
Gambar 4. 11 <i>Source Code Random Over Sampling</i>	52
Gambar 4. 12 Hasil <i>Oversampling</i>	52
Gambar 4. 13 <i>Source Code</i> Pembagian Data <i>Holdout</i>	53
Gambar 4. 14 <i>Source Code</i> Pembagian Data <i>K-Fold Cross Validation</i> ..	53
Gambar 4. 15 <i>Source Code</i> Random Over Sampling.....	54
Gambar 4. 16 Hasil <i>Oversampling</i>	54
Gambar 4. 17 <i>Source Code</i> Pembobotan TF-IDF	55
Gambar 4. 18 <i>Source Code</i> Pembobotan Data TF-RF	56
Gambar 4. 19 <i>Source Code</i> dan Hasil Klasifikasi Skenario 1	57
Gambar 4. 20 <i>Source Code</i> dan Hasil Klasifikasi Skenario 2	57
Gambar 4. 21 <i>Source Code</i> dan Hasil Klasifikasi Skenario 3	58
Gambar 4. 22 <i>Source Code</i> dan Hasil Klasifikasi Skenario 4	59
Gambar 4. 23 <i>Source Code</i> dan Hasil Klasifikasi Skenario 5	60

Gambar 4. 24	<i>Source Code</i> dan Hasil Klasifikasi Skenario 6	61
Gambar 4. 25	<i>Source Code</i> dan Hasil Klasifikasi Skenario 7	62
Gambar 4. 26	<i>Source Code</i> dan Hasil Klasifikasi Skenario 8	63
Gambar 4. 27	<i>Source Code</i> Pemodelan Topik LDA	64
Gambar 4. 28	Hasil Pemodelan Topik LDA	64
Gambar 4. 29	<i>Source Code</i> Pembuatan <i>Wordcloud</i> Kelas Positif.....	65
Gambar 4. 30	<i>Source Code</i> Pembuatan <i>Wordcloud</i> Kelas Negatif.....	65
Gambar 4. 31	Visualisasi <i>Wordcloud</i> pada Kelas Positif.....	66
Gambar 4. 32	Visualisasi <i>Wordcloud</i> Kelas Negatif.....	66
Gambar 4. 33	<i>Source Code</i> pyLDAvis Kelas Positif	67
Gambar 4. 34	<i>Source Code</i> pyLDAvis Kelas Negatif.....	67
Gambar 4. 35	Visualisasi pyLDAvis Kelas Positif	67
Gambar 4. 36	Visualisasi pyLDAvis Kelas Negatif.....	68
Gambar 4. 37	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 1.....	68
Gambar 4. 38	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 2.....	69
Gambar 4. 39	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 3.....	70
Gambar 4. 40	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 4.....	70
Gambar 4. 41	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 5.....	71
Gambar 4. 42	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 6.....	71
Gambar 4. 43	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 7.....	72
Gambar 4. 44	<i>Source Code</i> Serta Hasil <i>Confusion Matrix</i> & <i>F1-Score</i> Skenario 8.....	72
Gambar 4. 45	<i>Source Code</i> Perhitungan Nilai Koherensi	73
Gambar 4. 46	<i>Source Code</i> Pelabelan Kata LDA	78
Gambar 4. 47	<i>Source Code</i> HTML	79

Gambar 4. 48 <i>Source Code</i> PHP	79
Gambar 4. 49 <i>Source Code</i> Python	79
Gambar 4. 50 Tampilan Visualisasi Grafik	80
Gambar 4. 51 Visualisasi <i>WordCloud</i>	80
Gambar 4. 52 Visualisasi LDA.....	80
Gambar 4. 53 Prediksi Sentimen	81
Gambar 4. 54 Prediksi Topik.....	81

“Halaman ini sengaja dikosongkan.”