



Analisis Sentiment Kepuasan Masyarakat Terhadap Identitas Kependudukan Digital Dengan Multinomial *Naïve Bayes*

Miko Kastomo Putro¹, Ali Mustopa², Joko Dwi santoso³, Eka Wahyu Sholeha⁴

^{1,2,3} Universitas Amikom Yogyakarta, Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta, 55281, Indonesia.

⁴ Politeknik Negeri Tanah Laut, Jl. Ahmad Yani No.Km.06, Pemuda, Kec. Pelaihari, Kabupaten Tanah Laut, Kalimantan Selatan 70815

Info Artikel

Riwayat Artikel:

Diajukan 9 Juli 2024

Direvisi 20 Juli 2024

Diterima 24 Juli 2024

Publikasi 31 Juli 2024

Kata Kunci:

Analisis Sentiment

Apliasi IKD

Google Playstore

Multinomial *Naïve Bayes*

ABSTRAK

Baru baru ini pemerintah menerbitkan sebuah kebijakan baru yaitu KTP Digital. Penggunaan KTP digital ini bertujuan untuk meningkatkan efisiensi dan keamanan data penduduk. Maka dari itu peneliti akan menganalisa kepuasan masyarakat yang sudah menggunakan Aplikasi KTP Digital atau IKD dengan menggunakan Metode *Naïve Bayes*. Penelitian ini juga bertujuan untuk bisa mengetahui kelemahan dari aplikasi IKD menurut pengguna dan diharapkan penelitian ini bisa menjadi sebuah pedoman untuk pemerintah jika ingin melakukan upgrade sistem IKD.

ABSTRACT

Recently the government issued a new policy, namely the Digital KTP. The use of this digital KTP aims to increase the efficiency and security of population data. Therefore, researchers will analyze the satisfaction of people who have used the Digital KTP Application or IKD using the *Naïve Bayes* Method. This study also aims to be able to find out the weaknesses of the IKD application according to users and it is hoped that this study can be a guideline for the government if they want to upgrade the IKD system.

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.



Corresponding Author:

Miko Kastomo Putro

Email: miko@amikom.ac.id

1. PENDAHULUAN

Pada jaman sekarang, semua fitur akan menjadi digitalisasi, dari wallet digital contohnya Ovo, Dana dan shopeepay yang banyak digunakan untuk pembayaran dalam berbelanja atau kita ingin membeli makanan. Bahkan bukan hanya sebuah dompet menjadi digital, tetapi Kartu Tanda Pengenal (KTP) juga dimasukkan kedalam sebuah bentuk digitalisasi atau disebut dengan IKD atau Identitas Kependudukan Digital. Dengan menggunakan aplikasi IKD kita bisa menggunakan pelayanan pemerintah tanpa harus datang ke kantor atau cukup menggunakan IKD kita sudah bisa mendapatkan layanan yang kita inginkan. Tetapi ada menjadi masalah ketika banyak orang yang tidak mengetahui kegunaan aplikasi tersebut secara umum. Dalam era digital seperti sekarang, identitas digital menjadi sangat penting untuk berbagai keperluan, termasuk transaksi online, pengaksesan layanan publik digital, dan keamanan identitas secara umum. Identitas digital yang terkelola dengan baik dapat membantu mengurangi risiko kejahatan identitas dan meningkatkan efisiensi dalam berbagai proses administratif. Maka dari permasalahan tersebut. IKD memungkinkan integrasi yang lebih baik antara berbagai sistem dan layanan digital, baik di sektor pemerintah maupun swasta. Hal ini dapat meningkatkan efisiensi administrasi publik dan menyediakan akses yang lebih mudah bagi individu terhadap layanan yang mereka butuhkan. Peneliti berniat untuk membuat sebuah penelitian yang bertujuan untuk

mengetahui apakah Masyarakat sudah puas dengan penggunaan IKD menggunakan *Multinomial Naïve Bayes*. Meskipun IKD memberikan banyak manfaat dalam hal efisiensi dan keamanan, juga penting untuk mempertimbangkan implikasi sosial dan etika dari penggunaan teknologi ini. Perlindungan terhadap privasi individu dan keadilan dalam penggunaan IKD harus selalu diperhatikan[1].

Analisis sentimen adalah proses komputasional untuk mengidentifikasi, mengekstrak, atau menafsirkan sentimen atau opini yang terkandung dalam teks. Ini adalah bidang yang berkembang pesat dalam ilmu data dan pemrosesan bahasa alami (*Natural Language Processing, NLP*), dengan aplikasi yang luas dalam bisnis, media sosial, dan riset pasar. Dengan pertumbuhan pesat penggunaan media sosial dan platform daring lainnya, volume besar data teks yang mencakup opini, ulasan, dan komentar terus meningkat. Analisis sentimen membantu mengatasi tantangan dalam mengelola dan memahami data teks dalam skala besar ini. Opini pelanggan dapat memiliki dampak besar pada reputasi dan keberlanjutan bisnis. Analisis sentimen membantu perusahaan untuk mengidentifikasi umpan balik pelanggan dengan lebih cepat dan secara efisien, sehingga memungkinkan mereka untuk merespons secara proaktif terhadap masalah yang muncul. Dengan demikian, analisis sentimen adalah alat yang penting dalam memahami opini publik dan berbagai aspek sosial dan bisnis yang relevan, dengan potensi untuk meningkatkan pengambilan keputusan dan respons terhadap dinamika pasar dan sosial[2].

Multinomial Naïve Bayes untuk membantu klasifikasi dalam bentuk sentiment analisis. *Multinomial Naïve Bayes* adalah suatu algoritma khusus dari *Naïve Bayes* yang digunakan dalam text mining untuk mengklasifikasikan teks berdasarkan probabilitas munculnya kelas dalam dokumen. *Multinomial Naïve Bayes* adalah sebuah metode klasifikasi dalam machine learning yang digunakan khususnya untuk data yang direpresentasikan sebagai vektor fitur dengan nilai non-negatif, seperti dalam kasus analisis teks. Metode ini merupakan variasi dari *Naïve Bayes* yang mengasumsikan distribusi multinomial untuk setiap fitur (atau kata dalam konteks analisis teks)[3].

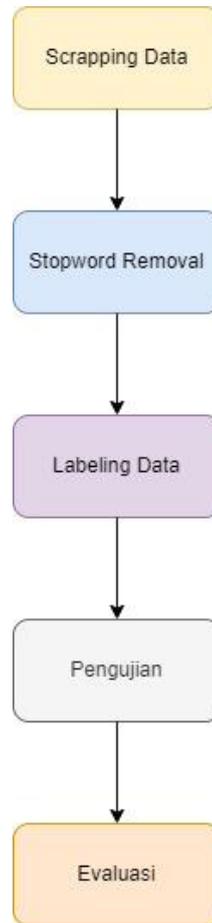
Pada penelitian sebelumnya berjudul analisis sentimen dengan *Naïve Bayes* terhadap komentar aplikasi Tokopedia oleh Rita pada September 2019, Penelitian ini menggunakan metode *Naïve Bayes* untuk menghasilkan sentimen positif dan negatif terhadap komentar pengguna aplikasi Tokopedia di Playstore. Pengujian berdasarkan nilai *class negative*, *class positive*, *Recall*, dan *Accuracy* pada analisis sentimen. dengan nilai *Accuracy performance* yang baik sebesar 97,13%, dengan nilai *precision* 1 Sementara pada *Class Recall* dihasilkan nilai 95,49% (positive class: negative). Dan nilai AUC 0,980[4].

Penelitian yang lain berjudul Penerapan Algoritma *Naïve Bayes* untuk Analisis Sentimen Review Data Twitter BMKG Nasional oleh Darwis tahun 2021, Pertumbuhan Twitter terus meningkat setiap waktu, sehingga hal tersebut dimanfaatkan para pengguna Twitter untuk menyampaikan informasi berupa kritik maupun saran kepada pelayanan yang diberikan BMKG Nasional dengan lebih mudah. Metode yang digunakan dalam penelitian ini adalah klasifikasi data adalah *Naïve Bayes Classifier* (NBC). Sistem yang dikembangkan dengan menggunakan data internal yang diambil dari internet/Twitter untuk proses penentuan kalimat termasuk opini positif, netral atau negatif. Penentuan tersebut digolongkan sebagai proses pengklasifikasian. Serta menggunakan Application Python 3.74. Hasil Penelitian ini masuk kedalam *finned grained sentiment analysis* yaitu analisis pada suatu kalimat komentar. Data tersebut akan diproses menggunakan text mining, kemudian dilanjutkan dengan mengklasifikasikan Tweet ke dalam tiga kelas, yaitu positif, negatif, dan netral. Klasifikasi ini menggunakan algoritma *naïve bayes*. Klasifikasi dapat memberikan kemudahan bagi pengguna untuk melihat opini positif, negatif, dan netral. Hasil uji akurasi pada metode *Naïve Bayes* untuk klasifikasi yaitu 69.97% [5].

Penelitian selanjutnya yang berjudul Analisis Sentimen menggunakan Metode *Naïve Bayes Classifier* terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding oleh Sinta tahun 2021. Salah satu algoritme untuk melakukan klasifikasi analisis sentimen, yaitu menggunakan metode *Naïve Bayes Classifier* di mana merupakan metode sederhana yang memiliki performa yang cepat dalam melatih data, mudah dalam implementasinya, serta memiliki efektifitas yang tinggi. Dalam proses klasifikasi akan digunakan seleksi fitur menggunakan algoritme N-gram dan DF-Thresholding untuk mengurangi dimensi fitur pada data. Tujuan dari penelitian ini untuk mengetahui pengaruh penggunaan algoritme DFThresholding pada tingkat akurasi algoritme *Naïve Bayes Classifier* menggunakan metode N-gram. Hasil penelitian menunjukkan adanya pengurangan fitur sebanyak 16.312 fitur menjadi 43 fitur dengan nilai akurasi tertinggi pada kombinasi unigram dan bigram, yaitu sebesar 49%, *precision* sebesar 0,23, *Recall* sebesar 0,26 serta f-measure sebesar 0,24[6].

2. METODE

Pada penelitian kali ini, peneliti menggunakan metode penelitian sebagai berikut yang dapat dilihat pada Gambar 1.



Gambar 1. Metode yang digunakan

Berikut penjelasan dari metode penelitian yang digunakan :

A. *Scrapping Data*

Pada proses ini mengumpulkan data yang ada di ulasan playstore dengan cara teknik scrapping data dimana *library* atau Pustaka yang disebut *google play scraper* berbasis python. *Scrapping* Google Play berarti mengumpulkan data secara otomatis dari Google Play Store, seperti informasi aplikasi, ulasan pengguna, peringkat, dan lain-lain. *Web scraping* Google Play dapat dilakukan untuk tujuan riset, analisis pasar, atau pengembangan aplikasi. Namun, penting untuk dicatat bahwa scrapping Google Play harus dilakukan dengan mematuhi kebijakan layanan Google, dan jika memungkinkan, lebih baik menggunakan API resmi yang disediakan oleh GoogleAPI *Google-Play-Scraper* adalah metode untuk mengambil data dari Google Play Store tanpa dependensi eksternal menggunakan bahasa pemrograman python. Data yang diambil dapat berupa informasi aplikasi seperti judul aplikasi, developer url, kategori aplikasi,keseluruhan rating dan review, deskripsi thumbnail, rating konten dan screenshot aplikasi. Selain informasi dari aplikasi API *Google-Play-Scraper* juga dapat mengambil data ulasan pengguna seperti nama, foto, rating, tanggal, comment likes, dan comment[11].

Data yang discrapping kita batasi 600 data, dimana nanti kita bakal mengambil variabel nomer, ulasan dan ratingnya. Jika ulasan tersebut memiliki rating 1-2 maka itu termasuk koment negative atau tidak

kepuasan masyarakat terhadap aplikasi tersebut, untuk rating 3 akan dihapus saja menjadikan data final berisi sekitar 560 Data.

B. *Stopword Removal*

Stopword Removal adalah langkah pra-pemrosesan umum dalam pemrosesan bahasa alami (*Natural Language Processing* atau NLP) dan penambangan teks. *Stopword* adalah kata-kata yang sering digunakan dalam suatu bahasa (seperti "dan," "yang," "adalah," dll.) yang biasanya tidak membawa makna signifikan dan sering dihapus dari data teks untuk fokus pada kata-kata yang lebih bermakna[11].

Pada tahap ini juga ditambahkan *case folding*, *case folding* adalah proses mengonversi semua huruf dalam teks menjadi huruf kecil (*lowercase*) untuk memastikan bahwa teks tersebut konsisten dalam hal huruf besar dan huruf kecil. Proses ini penting dalam pemrosesan teks dan analisis bahasa alami (NLP) karena membantu dalam penghapusan variasi yang disebabkan oleh perbedaan kapitalisasi.

C. *Labeling Data*

Setelah data diproses maka saatnya untuk memberikan label pada komentar atau ulasan yang ada di google play store, proses ini dinamakan *labeling data*. *Labeling data* adalah proses memberikan tanda atau label pada data untuk mengidentifikasi berbagai fitur atau kategori di dalamnya. *Labeling* ini sangat penting dalam machine learning dan data analysis, terutama dalam tugas-tugas yang melibatkan supervised learning, di mana model dilatih pada data yang telah dilabeli untuk memprediksi atau mengklasifikasikan data baru[3].

D. Pengujian

Setelah semua data sudah siap maka masuk proses pengujian dataset. Pada proses ini akan memulai proses implementasikan teknik *Multinomial Naïve Bayes* untuk membantu klasifikasi dalam bentuk sentiment analisis. *Multinomial Naïve Bayes* adalah suatu algoritma khusus dari *Naïve Bayes* yang digunakan dalam text mining untuk mengklasifikasikan teks berdasarkan probabilitas munculnya kelas dalam dokumen. Prosesnya dimulai dengan menyertakan data latih untuk proses pembelajaran, diikuti dengan menghitung probabilitas kemunculan suatu kelas pada data latih menggunakan persamaan khusus[3].

$$P(V_j) = \text{doc } j / \text{training} \quad (1)$$

Penjelasan :

$P(V_j)$: Probabilitas dokumen terhadap kategori atau kelas kata.

$\text{doc } j$: Jumlah dokumen memiliki kategori atau kelas "j" yaitu positif atau negative dalam data latih.

training : jumlah dokumn dalam data latih

Selanjutnya menghitung probabilitas kata i termasuk dalam kelas tertentu dapat dicapai dengan menerapkan rumus no 2.

$$P(x_i|V_j) = \frac{n_{k+1}}{n + |\text{kosakata}|} \quad (2)$$

Penjelasan:

$P(x_i|V_j)$: probilitas kata "xi" dalam suatu dokumen berdasarkan V_j

n_k : frekuensi kehadiran kata "xi" dalam dokumen pada kategori atau kelas dokumen V_j .

n : total jumlah kata dalam dokumen

$|\text{kosakata}|$: jumlah kata pada data latih

E. Analisa

Setelah melakukan pengujia, masuk pada proses Analisa dari hasil pengujian tadi. Menganalisa untuk mendapatkan kesimpulan apakah masyarakat puas dengan penggunaan platform IKD atau Identitas Kependudukan Digital

3. HASIL DAN DISKUSI

Setelah menentukan metode penelitian yang bakal digunakan, selanjutnya peneliti akan melakukan testing sesuai dengan metode penelitian yang sudah dibuat.

A. Dataset

Pada proses ini peneliti dengan bantuan python akan mengambil ulasan yang ada pada playstore IKD dengan menggunakan Teknik *Scrapping Data*. Pada teknik pengambilan ini peneliti hanya membatasi 600 data saja yang diambil. Selanjutnya data scrapping akan disimpan dengan format csv yang ditunjukkan pada gambar 2.

| | reviewId | userName | userImage | content | score | thumbsUpCount | reviewCreatedVersion | at | replyContent | repliedAt | appVersion |
|---|--------------------------------------|------------------|--|---|-------|---------------|----------------------|---------------------|--------------|-----------|------------|
| 0 | 32c0006e-9032-40aa-82ba-e3871d148b3b | Didin Misbahudin | https://play-ih.googleusercontent.com/va/ALV-U... | Kayanya aplikasi ini termasuk datanya yang dij... | 3 | 0 | 1.2.2 | 2024-07-09 00:25:08 | None | None | 1.2.2 |
| 1 | ada32f97-d90f-4b6e-a75a-5f2417362981 | The Slebew | https://play-ih.googleusercontent.com/va/ACg80c... | Suram kesalahan koneksi terus | 1 | 0 | 1.2.2 | 2024-07-09 00:09:40 | None | None | 1.2.2 |
| 2 | fab8e23a-83e5-43cd-a694-2c7a92e89a64 | dhee must | https://play-ih.googleusercontent.com/va/ALV-U... | Sudah senang padahal, kirain Indonesia makin m... | 1 | 0 | 1.2.2 | 2024-07-08 22:44:59 | None | None | 1.2.2 |
| 3 | c30c0662-40d1-4437-812e-0eb2592c0219 | Jangan Di hps | https://play-ih.googleusercontent.com/va/ALV-U... | gw udah install tapi pas gw mo buka apps nya l... | 1 | 0 | 1.2.2 | 2024-07-08 21:28:28 | None | None | 1.2.2 |
| 4 | 5beb622-afa5-471a-a8ac-e29e1782b60a | Eka Yulianti | https://play-ih.googleusercontent.com/va/ACg80c... | Aplikasi ga guna, harus scan barkode ke dukcap... | 1 | 0 | None | 2024-07-08 21:14:36 | None | None | None |

Gambar 2. Contoh data awal dengan menggunakan teknik *Scrapping*

Data yang didapatkan tadi masih dalam bentuk kasar atau masih ada beberapa variable yang tidak kita butuhkan. Maka dari itu kita hanya membutuhkan username, comment dan score / rating Bintang. Hasilnya ditunjukkan pada gambar 3.

| | userName | content | score |
|----|---------------------|---|-------|
| 0 | didik s | udah ribet ditambah aplikasi nya cacat | 1 |
| 1 | Abdul Manap | Di bank masih belum bisa di pake | 1 |
| 2 | Hirzuda Prastiawan | Terjadi kesalahan koneksi. | 1 |
| 3 | Karim Hidayat S. | Aplikasi ga guna.. koneksi bagus kok dibilang ... | 1 |
| 4 | Budi Triyono 378 | tolong perbaiki sistemnya, sangat susah sekali... | 1 |
| 5 | Yustina Tio Pratama | Sekiranya ga bisa mempermudah rakyat, jangan d... | 1 |
| 6 | ATEP SOFWAN | Urusan ktp el jadi mudah | 5 |
| 7 | Agus Setiono | bagus. menunggu perkembangan selanjutnya.... | 5 |
| 8 | Didin Misbahudin | Kayanya aplikasi ini termasuk datanya yang dij... | 3 |
| 9 | The Slebew | Suram kesalahan koneksi terus | 1 |
| 10 | dhee must | Sudah senang padahal, kirain Indonesia makin m... | 1 |
| 11 | Jangan Di hps | gw udah install tapi pas gw mo buka apps nya l... | 1 |
| 12 | Eka Yulianti | Aplikasi ga guna, harus scan barkode ke dukcap... | 1 |
| 13 | janoko 53 | Coba2 dlu kalo bagus sya kasi bintang 5 | 4 |
| 14 | Sulaeman Eman | susah masuk | 5 |
| 15 | Tedi Julianda | nggak bisa di buka memuat mulu | 1 |
| 16 | aruel wonk | Ngapain bikin aplikasi begini , kalau harus ke... | 1 |
| 17 | AS Sabar | Kena apa sih lelet amat bukannya. | 3 |
| 18 | david andika | udah coba hapus data, rupanya 1 hp cuma bisa 1... | 1 |
| 19 | Udang Galau | Update mei 2023 yang lalu. Hellow.... Ini udah... | 1 |
| 20 | Paiman 12 | Aplikasi nya gak bisa di buka,mungkin dia lelah | 1 |

Gambar 3. Final Dataset

B. Stopword Removal

Karena dataset tersebut masih memiliki kelemahan salah satunya dimana data tersebut masih memiliki tata penulisan yang campur alias tidak rapi, maka disini menggunakan proses *case folding* dimana proses ini bertujuan untuk mengubah seluruh huruf besar atau kapital ke dalam bentuk huruf kecil. Hasil dari *case folding* ini ditunjukkan pada gambar 4.

| | userName | content | score | Case_Folding |
|---|---------------------|---|-------|---|
| 0 | didik s | udah ribet ditambah aplikasi nya cacat | 1 | udah ribet ditambah aplikasi nya cacat |
| 1 | Abdul Manap | Di bank masih belum bisa di pake | 1 | di bank masih belum bisa di pake |
| 2 | Hirzuda Prastiawan | Terjadi kesalahan koneksi. | 1 | terjadi kesalahan koneksi |
| 3 | Karim Hidayat S. | Aplikasi ga guna.. koneksi baguskok dibilang ... | 1 | aplikasi ga guna koneksi baguskok dibilang ko... |
| 4 | Budi Triyono 378 | tolong perbaiki sistemnya, sangat susah sekali... | 1 | tolong perbaiki sistemnya sangat susah sekali ... |
| 5 | Yustina Tio Pratama | Sekiranya ga bisa mempermudah rakyat, jangan d... | 1 | sekiranya ga bisa mempermudah rakyat jangan di... |
| 6 | ATEP SOFWAN | Urusan ktp el jadi mudah | 5 | urusan ktp el jadi mudah |
| 7 | Agus Setiono | bagus. menunggu perkembangan selanjutnya.... | 5 | bagus menunggu perkembangan selanjutnya |
| 8 | Didin Misbahudin | Kayanya aplikasi ini termasuk datanya yang dij... | 3 | kayanya aplikasi ini termasuk datanya yang dij... |
| 9 | The Slebew | Suram kesalahan koneksi terus | 1 | suram kesalahan koneksi terus |

Gambar 4. Data melalui proses *Case folding*

Setelah melalui proses *case folding*, karena masih ada beberapa data yang memiliki kata umum yang biasanya muncul dalam jumlah yang besar dan kadang dianggap tidak memiliki makna, maka kita perlu menghapus kata kata yang memiliki informasi rendah dari sebuah teks, agar focus pada kata-kata yang penting saja proses ini dinamakan *Stopword Removal*. Hasil *Stopword Removal* ditunjukkan gambar 5.

| | userName | content | score | Case_Folding | Stopword_Removal |
|----|---------------------|---|-------|---|---|
| 0 | didik s | udah ribet ditambah aplikasi nya cacat | 1 | udah ribet ditambah aplikasi nya cacat | udah ribet ditambah aplikasi nya cacat |
| 1 | Abdul Manap | Di bank masih belum bisa di pake | 1 | di bank masih belum bisa di pake | bank pake |
| 2 | Hirzuda Prastiawan | Terjadi kesalahan koneksi. | 1 | terjadi kesalahan koneksi | kesalahan koneksi |
| 3 | Karim Hidayat S. | Aplikasi ga guna.. koneksi baguskok dibilang ... | 1 | aplikasi ga guna koneksi baguskok dibilang ko... | aplikasi ga koneksi baguskok dibilang koneksi... |
| 4 | Budi Triyono 378 | tolong perbaiki sistemnya, sangat susah sekali... | 1 | tolong perbaiki sistemnya sangat susah sekali ... | tolong perbaiki sistemnya susah dibuka |
| 5 | Yustina Tio Pratama | Sekiranya ga bisa mempermudah rakyat, jangan d... | 1 | sekiranya ga bisa mempermudah rakyat jangan di... | ga mempermudah rakyat bikin ribetlah yah tulis... |
| 6 | ATEP SOFWAN | Urusan ktp el jadi mudah | 5 | urusan ktp el jadi mudah | urusan ktp el mudah |
| 7 | Agus Setiono | bagus. menunggu perkembangan selanjutnya.... | 5 | bagus menunggu perkembangan selanjutnya | bagus menunggu perkembangan |
| 8 | Didin Misbahudin | Kayanya aplikasi ini termasuk datanya yang dij... | 3 | kayanya aplikasi ini termasuk datanya yang dij... | kayanya aplikasi datanya dijebol hacker |
| 9 | The Slebew | Suram kesalahan koneksi terus | 1 | suram kesalahan koneksi terus | suram kesalahan koneksi |
| 10 | dhee must | Sudah senang padahal, kirain Indonesia makin m... | 1 | sudah senang padahal kirain indonesia makin ma... | senang kirain indonesia maju eeeh gak taunya ... |
| 11 | Jangan Di hps | gw udah install tapi pas gw mo buka apps nya l... | 1 | gw udah install tapi pas gw mo buka apps nya l... | gw udah install pas gw mo buka apps nya loadin... |
| 12 | Eka Yullanti | Aplikasi ga guna, harus scan barkode ke dukcap... | 1 | aplikasi ga guna harus scan barkode ke dukcapi... | aplikasi ga scan barkode dukcapil ribet dibuka... |
| 13 | janoko 53 | Coba2 dlu kalo bagus sya kasi bintang 5 | 4 | coba dlu kalo bagus sya kasi bintang | coba dlu kalo bagus sya kasi bintang |
| 14 | Sulaeman Eman | susah masuk | 5 | susah masuk | susah masuk |
| 15 | Tedi Julianda | nggak bisa di buka memuat mulu | 1 | nggak bisa di buka memuat mulu | nggak buka memuat mulu |
| 16 | aruel wonk | Ngapain bikin aplikasi begini , kalau harus ke... | 1 | ngapain bikin aplikasi begini kalau harus ke ... | ngapain bikin aplikasi dukcapil |
| 17 | AS Sabar | Kena apa sih lelet amat bukannya | 3 | kena apa sih lelet amat bukannya | kena sih lelet bukannya |
| 18 | david andika | udah coba hapus data, rupanya 1 hp cuma bisa 1... | 1 | udah coba hapus data rupanya hp cuma bisa ak... | udah coba hapus data hp akun bikin ktp fisik o... |
| 19 | Udang Galau | Update mei 2023 yang lalu. Hellow.... Ini udah... | 1 | update mei yang lalu hellow ini udah kemana ... | update mei hellow udah kemana aja lo |

Gambar 5. Data melalui proses *Stopword Removal*

C. Labeling Data

Selanjutnya setelah kita mendapat dataset yang dibutuhkan, maka kita memasuki pada proses pelabelan untuk menentukan mana yang termasuk koment positif dan negative berupa rating score yang ada pada ulasan di playstore. Untuk skor rating yang tercantum dibawah 3 bintang maka kita asumsikan itu sebuah ulasan yang negative atau ketidak puasan Masyarakat, selanjutnya jika ulasan rating tersebut terdapat skor diatas 3 maka ulasan atau koment tersebut masuk pada label Positif. Hasil pada proses label ini tercantum pada gambar 6.

| | userName | content | score | Case_Folding | Stopword_Removal | Label |
|----|---------------------|---|-------|---|---|---------|
| 0 | didik s | udah ribet ditambah aplikasi nya cacat | 1 | udah ribet ditambah aplikasi nya cacat | udah ribet ditambah aplikasi nya cacat | Negatif |
| 1 | Abdul Manap | Di bank masih belum bisa di pake | 1 | di bank masih belum bisa di pake | bank pake | Negatif |
| 2 | Hirzuda Prastilawan | Terjadi kesalahan koneksi. | 1 | terjadi kesalahan koneksi | kesalahan koneksi | Negatif |
| 3 | Karim Hidayat S. | Aplikasi ga guna.. koneksi baguskok dibilang ... | 1 | aplikasi ga guna koneksi baguskok dibilang ko... | aplikasi ga koneksi baguskok dibilang koneksi... | Negatif |
| 4 | Budi Triyono 378 | tolong perbaiki sistemnya, sangat susah sekali... | 1 | tolong perbaiki sistemnya sangat susah sekali ... | tolong perbaiki sistemnya susah dibuka | Negatif |
| 5 | Yustina Tio Pratama | Sekiranya ga bisa mempermudah rakyat, jangan d... | 1 | sekiranya ga bisa mempermudah rakyat jangan di... | ga mempermudah rakyat bikin ribetlah yah tulis... | Negatif |
| 6 | ATEP SOFWAN | Urusan ktp el jadi mudah | 5 | urusan ktp el jadi mudah | urusan ktp el mudah | Positif |
| 7 | Agus Setiono | bagus. menunggu perkembangan selanjutnya.... | 5 | bagus menunggu perkembangan selanjutnya | bagus menunggu perkembangan | Positif |
| 9 | The Slebew | Suram kesalahan koneksi terus | 1 | suram kesalahan koneksi terus | suram kesalahan koneksi | Negatif |
| 10 | dhee must | Sudah senang padahal, kirain Indonesia makin m... | 1 | sudah senang padahal kirain indonesia makin ma... | senang kirain indonesia maju eeeh gak taunya ... | Negatif |
| 11 | Jangan Di hps | gw udah install tapi pas gw mo buka apps nya l... | 1 | gw udah install tapi pas gw mo buka apps nya l... | gw udah install pas gw mo buka apps nya loadin... | Negatif |
| 12 | Eka Yulianti | Aplikasi ga guna, harus scan barkode ke dukcap... | 1 | aplikasi ga guna harus scan barkode ke dukcapl... | aplikasi ga scan barkode dukcapil ribet dibuka... | Negatif |
| 13 | Janoko 53 | Coba2 dlu kalo bagus sya kasi bintang 5 | 4 | coba dlu kalo bagus sya kasi bintang | coba dlu kalo bagus sya kasi bintang | Positif |
| 14 | Sulaeman Eman | susah masuk | 5 | susah masuk | susah masuk | Positif |
| 15 | Tedi Juliana | nggak bisa di buka memuat mulu | 1 | nggak bisa di buka memuat mulu | nggak buka memuat mulu | Negatif |
| 16 | aruel wonk | Ngapain bikin aplikasi begini , kalau harus ke... | 1 | ngapain bikin aplikasi begini kalau harus ke ... | ngapain bikin aplikasi dukcapil | Negatif |
| 18 | david andika | udah coba hapus data, rupanya 1 hp cuma bisa 1... | 1 | udah coba hapus data rupanya hp cuma bisa ak... | udah coba hapus data hp akun bikin ktp fisik o... | Negatif |
| 19 | Udang Galau | Update mei 2023 yang lalu. Hellow.... Ini udah... | 1 | update mei yang lalu hellow ini udah kemana ... | update mei hellow udah kemana aja lo | Negatif |

Gambar 6. Data melalui proses *labeling*

D. Pengujian

Setelah melakukan pelabelan dan preprocessing data, akhirnya data yang terkumpul akhirnya melakukan perhitungan multinomial *Naïve Bayes*. Hasil dari pengujian berada pada gambar 7. Dimana hasil yang didapatkan *Accuracy* sebesar 91%, *Precision* 89.4%, *Recall* 98.7%

```
MultinomialNB Accuracy: 0.911504424778761
MultinomialNB Precision: 0.8941176470588236
MultinomialNB Recall: 0.987012987012987
MultinomialNB f1_score: 0.9382716049382716
```

Gambar 7. Hasil Pengujian Multinomial *Naïve Bayes*

Setelah melakukan perhitungan Multinomial *Naïve Bayes*, selanjutnya mengevaluasi dari pengujian Multinomial *Naïve Bayes* dengan menggunakan *Confusion Matrix*. Hasil dari *Confusion Matrix* ditunjukkan gambar 8.

```
confusion_matrix:
[[76  1]
 [ 9 27]]
=====
```

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| Negatif | 0.89 | 0.99 | 0.94 | 77 |
| Positif | 0.96 | 0.75 | 0.84 | 36 |
| accuracy | | | 0.91 | 113 |
| macro avg | 0.93 | 0.87 | 0.89 | 113 |
| weighted avg | 0.92 | 0.91 | 0.91 | 113 |

Gambar 8. Hasil *Confusion Matrix*

Pada gambar 8 menunjukkan hasil *Confusion Matrix* dimana class negatif memiliki tingkatan *precision* sebanyak 89%, *Recall* 99% dan *F1-score* 94%. Sedangkan untuk class positif memiliki tingkatan *precision* 96%, 75%, 84% dengan *Accuracy Confusion Matrix*nya 91% atau cukup baik.

E. Analisa Hasil Pengujian

Pada Gambar 8, hasil analisis sentimen menunjukkan rincian prediksi model terhadap data ulasan aplikasi IKD di Playstore. Dari hasil prediksi, dapat diuraikan sebagai berikut:

Prediksi Sentimen Negatif:

Terdapat 76 ulasan yang diprediksi sebagai sentimen negatif dan ternyata benar-benar negatif (*True Negative*). Ada 1 ulasan yang diprediksi sebagai sentimen negatif namun sebenarnya adalah sentimen positif (*False Negative*).

Prediksi Sentimen Positif:

Sebanyak 9 ulasan diprediksi sebagai sentimen positif, namun seharusnya masuk ke dalam kategori sentimen negatif (*False Positive*).

Prediksi Benar Sentimen Negatif:

Ada 27 ulasan yang diprediksi sebagai sentimen negatif dan prediksi tersebut terbukti benar (*True Negative*). Hasil analisis ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengidentifikasi ulasan dengan sentimen negatif, meskipun masih ada kesalahan prediksi yang perlu diperhatikan. Kesalahan prediksi, baik *False Negative* maupun *False Positive*, perlu diminimalisir untuk meningkatkan akurasi model dalam klasifikasi sentimen.

Secara keseluruhan, meskipun terdapat beberapa kesalahan dalam prediksi, model ini masih menunjukkan performa yang cukup baik dengan kemampuan mengidentifikasi ulasan negatif secara akurat. Hasil analisis ini juga menekankan pentingnya peningkatan model dan pemahaman lebih lanjut terhadap data untuk mengurangi tingkat kesalahan dan meningkatkan keakuratan prediksi.

4. KESIMPULAN

Kesimpulan yang didapatkan dari penelitian ini menunjukkan bahwa aplikasi IKD di Playstore memiliki lebih banyak ulasan dengan sentimen negatif dibandingkan dengan sentimen positif, dengan perbandingan 362 ulasan negatif dan 201 ulasan positif. Hal ini mengindikasikan bahwa aplikasi IKD masih memiliki beberapa bug atau kelemahan. Sebagian besar ulasan negatif mengarah pada masalah gangguan jaringan yang sering terjadi saat pengguna melakukan registrasi data.

Hasil analisis performa model menunjukkan tingkat akurasi yang cukup tinggi dengan nilai *Accuracy* sebesar 91%, *Precision* sebesar 89.4%, dan *Recall* sebesar 98.7%. Ini menunjukkan bahwa model yang digunakan dalam penelitian ini mampu mengklasifikasikan sentimen dengan tingkat akurasi yang tinggi. Meski demikian, tingginya jumlah ulasan negatif menunjukkan bahwa aplikasi IKD perlu melakukan perbaikan signifikan, terutama dalam hal stabilitas jaringan saat registrasi, untuk meningkatkan kepuasan pengguna.

PERNYATAAN KONTRIBUSI PENULIS

Penulis pertama bertanggung jawab atas pengembangan konsep dan metodologi penelitian. Mereka merancang kerangka kerja dan strategi yang digunakan dalam penelitian ini. **Penulis kedua** mengurus aspek teknis, termasuk penggunaan software yang relevan dan administrasi proyek, memastikan semua prosedur dan perangkat lunak berjalan dengan lancar. **Penulis ketiga** fokus pada penulisan naskah, dengan tugas tambahan untuk meninjau dan menyunting konten yang telah ditulis untuk memastikan kualitas dan konsistensi. Demikian pula, **penulis keempat** juga bertanggung jawab atas penulisan naskah serta meninjau dan menyuntingnya, bekerja sama dengan penulis ketiga untuk menyempurnakan hasil akhir penelitian..

PERNYATAAN KONFLIK KEPENTINGAN

Kami, para penulis, dengan ini menyatakan bahwa kami tidak memiliki kepentingan finansial atau hubungan pribadi yang dapat mempengaruhi hasil penelitian yang dilaporkan dalam makalah ini. Semua data dan hasil yang disajikan dalam penelitian ini merupakan hasil kerja kami yang objektif dan berdasarkan fakta yang kami peroleh selama proses penelitian.

Dengan ini kami menyatakan bahwa:

1. Kami tidak memiliki kepentingan finansial yang berkaitan dengan subjek penelitian ini.
2. Kami tidak memiliki hubungan pribadi dengan individu atau organisasi yang dapat mempengaruhi penilaian dan hasil dari penelitian ini.
3. Penelitian ini dilakukan dengan integritas dan sesuai dengan prinsip-prinsip etika penelitian yang berlaku.

Kami berharap bahwa makalah ini dapat memberikan kontribusi yang berharga bagi pengembangan ilmu pengetahuan dan dapat digunakan sebagai referensi yang terpercaya oleh pembaca dan peneliti lainnya.

REFERENSI

- [1] Ditjen Dukcapil Kemendagri, “Identitas Kependudukan Digital Wujudkan Layanan Publik yang Inklusif”. 12 Januari 2024. <https://dukcapil.kemendagri.go.id/blog/read/identitas-kependudukan-digital-wujudkan-layanan-publik-yang-inklusif>.
- [2] Liu, Bing. 2015. “Sentiment Analysis: Mining Opinions, Sentiments, and Emotions”, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, no. May: 1– 367.
- [3] Manning, christopher d., Prabhakar Raghavan and Hinrich Schutze. 2009. “An Introduction To Informational Retrieval”. *Information Retrieval*.
- [4] R. Apriani, D Gustianm. “analisis sentimen dengan *Naïve Bayes* terhadap komentar aplikasi tokopedia”, *Jurnal Rekayasa Teknologi Nusa Putra*. Vol. 6, No. 1, September 2019: Hal 54 – 62.
- [5] D. Da, dkk. “Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional”, *Jurnal TEKNO KOMPAK*, Vol. 15, No. 1, P-ISSN: 1412-9663, E-ISSN : 2656-3525, Hal. 131-145.
- [6] S. Kusuma, dkk. “Analisis Sentimen menggunakan Metode *Naïve Bayes Classifier* terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding”, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN: 2548-964X* Vol. 5, No. 12, Desember 2021, hlm. 5582-5590.
- [7] Buntoro, Ghulam Asrofi, Teguh Bharata Adji, dan Adhistya Erna Purnamasari. 2016. *Sentiment Analysis Candidates of Indonesian Presiden 2014 with Five Class Attribute*. *International Journal of Computer Applications*. Vol. 136 No. 2. Ronowijayan :Universitas Muhammadiyah Ponorogo
- [8] F. Ratnawati, “ Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter”, *JURNAL INOVTEK POLBENG – SERI INFORMATIKA*, VOL. 3, NO. 1, JUNI 2018.
- [9] Yudho Baskoro, Setyoko, Achmad Ridok and Muhammad Tanzil Furqon. 2015. “Pencarian Pasal Pada Kitab UndangUndang Hukum Pidana (Kuhp) Berdasarkan Kasus Menggunakan Metode Cosine Similarity Dan LatentSemantic Indexing (Lsi)”. *Journal of Enviromental Engineering and Sustainable Technology 2*: 83–88.<<https://doi.org/10.21776/ub.jeest.2015.002.02.4>>.
- [10] Nurjanah, Winda Estu, Rizal Setya Perdana and Mochammad Ali Fauzi. 2017. “Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat Pada Media Sosial Twitter Menggunakan Metode K-Nearest Neighbor Dan Pembobotan Jumlah Retweet”. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya 1* No. 12: 1750–57.”
- [11] S.I Nurhafida, F Sembiring. “Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)”. *Jurnal Sains Komputer & Informatika (J-SAKTI)*. Volume 6 Nomor 1, Maret 2022, pp. 317-327