

ANALISIS SENTIMEN ULASAN APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL DI GOOGLE PLAY STORE DENGAN BERT

Ahmad Farhan, Istiadi, Aviv Yuniar Rahman
Teknik Informatika, Universitas Widyagama Malang
Jl. Taman Borobudur Indah 3 Malang
ahmd.frhn2627@gmail.com

ABSTRAK

Dalam era digital, layanan administrasi kependudukan mengalami transformasi dengan hadirnya aplikasi Identitas Kependudukan Digital (IKD), yang dikembangkan oleh Ditjen Dukcapil untuk menggantikan KTP fisik dalam berbagai transaksi. Namun, sejak peluncurannya di Google Play Store, aplikasi ini menerima berbagai ulasan dari pengguna, yang mencerminkan pengalaman mereka terhadap layanan yang diberikan. Ulasan ini dapat menjadi indikator penting dalam mengevaluasi kualitas aplikasi, tetapi analisis manual terhadap ribuan ulasan sangat tidak efisien. Oleh karena itu, diperlukan metode otomatis yang mampu mengkategorikan opini pengguna secara akurat. Penelitian ini menerapkan analisis sentimen menggunakan model Bidirectional Encoder Representations from Transformers (BERT) untuk mengklasifikasikan 15.000 ulasan pengguna ke dalam kategori positif, netral, atau negatif. Teknik Synthetic Minority Over-sampling Technique (SMOTE) diterapkan untuk menyeimbangkan distribusi kelas sentimen. Model dilatih dengan fine-tuning pada IndoBERT menggunakan batch size 16, learning rate $2e-5$, dan epoch 10. Hasil evaluasi menunjukkan akurasi 84,69%, yang lebih unggul dibandingkan metode tradisional seperti Naïve Bayes dan SVM. Studi ini memberikan wawasan bagi pengembang dalam meningkatkan layanan aplikasi berdasarkan opini pengguna, serta menegaskan keunggulan BERT dalam analisis sentimen berbahasa Indonesia.

Kata kunci : Analisis Sentimen, BERT, IndoBERT, Google Play Store, SMOTE, Identitas Kependudukan Digital.

1. PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan besar dalam berbagai aspek kehidupan, termasuk dalam sistem administrasi kependudukan. Inovasi di bidang ini semakin berkembang seiring dengan meningkatnya kebutuhan masyarakat akan layanan yang lebih cepat, efisien, dan mudah diakses secara digital. Salah satu inovasi yang diterapkan oleh pemerintah Indonesia adalah pengembangan Aplikasi Identitas Kependudukan Digital (IKD) oleh Direktorat Jenderal Kependudukan dan Pencatatan Sipil (DITJEN DUKCAPIL) di bawah Kementerian Dalam Negeri [1]. Aplikasi ini bertujuan untuk memberikan kemudahan bagi masyarakat dalam mengakses informasi kependudukan secara digital, menggantikan kebutuhan akan KTP fisik dalam berbagai transaksi administratif [2].

Seiring dengan peluncurannya di Google Play Store, aplikasi ini telah mendapatkan berbagai ulasan dari pengguna yang mencerminkan pengalaman dan tingkat kepuasan mereka terhadap layanan yang diberikan. Ulasan pengguna menjadi sumber data yang berharga dalam mengevaluasi kualitas suatu aplikasi, karena dapat memberikan gambaran mengenai kelebihan dan kekurangan aplikasi berdasarkan pengalaman langsung pengguna [3]. Dengan jumlah ulasan yang terus meningkat, diperlukan metode analisis yang sistematis untuk mengolah dan mengkategorikan opini pengguna ke dalam kelompok sentimen tertentu.

Salah satu metode yang digunakan untuk mengolah data ulasan pengguna adalah analisis sentimen, yaitu teknik yang digunakan untuk

mengklasifikasikan opini pengguna ke dalam kategori positif, netral, atau negatif. Dengan menerapkan analisis sentimen, pengembang dapat memahami pola opini pengguna secara lebih sistematis dan mendapatkan wawasan yang berguna dalam pengembangan dan peningkatan aplikasi [4]. Namun, dengan jumlah ulasan yang besar, analisis secara manual menjadi tidak efisien. Oleh karena itu, dibutuhkan pendekatan berbasis kecerdasan buatan yang dapat mengolah data teks dalam jumlah besar secara otomatis dan akurat [5].

Berbagai metode machine learning telah digunakan dalam penelitian analisis sentimen, seperti Naïve Bayes, Support Vector Machine (SVM), dan Long Short-Term Memory (LSTM). Metode-metode tersebut telah menunjukkan hasil yang cukup baik dalam klasifikasi teks, tetapi memiliki keterbatasan dalam memahami konteks secara menyeluruh [6]. Oleh karena itu, penelitian terbaru lebih banyak menggunakan model deep learning yang lebih canggih, salah satunya adalah Bidirectional Encoder Representations from Transformers (BERT). Model ini memiliki keunggulan dalam memahami makna kata dalam sebuah kalimat dengan lebih baik karena menggunakan arsitektur berbasis self-attention [7].

Penelitian sebelumnya menunjukkan bahwa BERT mampu menghasilkan akurasi yang lebih tinggi dibandingkan metode tradisional dalam tugas analisis sentimen berbahasa Indonesia [8]. Model ini dapat menangkap hubungan antar kata secara dua arah, memungkinkan pemahaman konteks yang lebih mendalam dalam teks. Oleh karena itu, model ini menjadi pilihan utama dalam berbagai penelitian

terbaru di bidang pemrosesan bahasa alami (Natural Language Processing).

Penelitian ini bertujuan untuk menerapkan model BERT, khususnya IndoBERT, dalam analisis sentimen ulasan pengguna terhadap aplikasi IKD yang tersedia di Google Play Store. Data yang digunakan dalam penelitian ini terdiri dari 15.000 ulasan, yang telah melalui beberapa tahapan pemrosesan, termasuk pre-processing untuk membersihkan teks, pelabelan sentimen berdasarkan rating bintang, serta teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk menangani ketidakseimbangan kelas dalam dataset.

Dengan pendekatan ini, diharapkan penelitian ini dapat memberikan wawasan yang lebih dalam mengenai persepsi pengguna terhadap aplikasi IKD serta menjadi referensi dalam pemanfaatan model BERT untuk analisis sentimen dalam aplikasi layanan publik.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen adalah metode untuk menganalisis pendapat, sentimen, evaluasi, sikap, penilaian, dan emosi terkait layanan, produk, individu, atau kegiatan tertentu. Tujuannya adalah mengidentifikasi opini seseorang dengan memperhatikan topik tertentu. Tahapan dalam analisis sentimen melibatkan pengidentifikasian domain dataset, preprocessing, pemilihan fitur, pelabelan, klasifikasi, dan evaluasi [9]. Analisis Sentimen merupakan bentuk khusus dari data mining atau penambangan data yang memanfaatkan Natural Language Processing (NLP), analisis teks, dan Linguistik Komputasi untuk mengekstraksi dan menganalisis informasi subjektif yang berasal dari media sosial dan sumber-sumber lainnya. Fokusnya adalah mengidentifikasi arah atau polaritas kata, frasa, kalimat, dan dokumen, dan dapat diterapkan di berbagai sektor seperti bisnis, politik, dan psikologi. Karena Analisis Sentimen memiliki aplikasi yang luas, dapat digunakan sebagai landasan atau referensi dalam berbagai kegiatan manusia dan memainkan peran penting dalam membentuk perilaku manusia secara umum. Oleh karena itu, seringkali dalam proses pengambilan keputusan, kita cenderung mencari pendapat orang lain. Prinsip ini tidak hanya berlaku untuk individu, tetapi juga berlaku untuk organisasi [4].

Analisis sentimen adalah metode untuk mengevaluasi pemikiran, perasaan, dan penilaian pengguna berdasarkan informasi seperti opini tertulis. Tujuan utamanya adalah untuk mengidentifikasi pandangan yang diyakini oleh pengguna terkait suatu topik atau subjek tertentu. Sentiment analysis erat kaitannya dengan Natural Language Processing (NLP), yang merupakan kemampuan mesin untuk memahami bahasa manusia menggunakan algoritma komputer, matematika, dan linguistik komputasi. Contoh produk NLP termasuk Google Assistant dan

Siri. Python dianggap cocok untuk diterapkan dalam NLP karena ketersediaan library open source, kemudahan penggunaan, kompleksitas yang tidak terlalu tinggi, dan optimalisasi sistem berbasis NLP yang lebih sederhana dibandingkan paradigma pemrograman lainnya. NLP sendiri adalah teknik untuk menganalisis data teks dengan pendekatan berbasis aturan (rule-based) dan pembelajaran mesin. Dalam konteks analisis sentimen, polaritas dan subjektivitas menjadi metrik penting. Pelatihan model dapat dimonitor melalui grafik kerugian pelatihan (training loss), dan evaluasi akurasi memberikan gambaran tentang seberapa efektif model dalam menganalisis sentiment [10].

Bidang penelitian Text Mining dikenal sebagai Analisis Sentimen atau Opinion Mining, yang mencakup pembaruan terhadap suatu pendapat. Proses ini melibatkan penentuan perspektif pembicara atau penulis terhadap sejumlah topik atau keseluruhan polaritas dalam konteks tertentu. Penggalan opini, pendapat, penilaian, evaluasi, serta pemahaman terhadap keadaan afektif atau komunikasi emosional menjadi elemen kunci dalam mengambil sikap. Contoh penerapan dari analisis sentimen mencakup produk konsumen, layanan kesehatan, acara sosial dan politik, serta analisis tren dengan menggunakan pendekatan positif atau negatif [11]. Penelitian ini bertujuan untuk menyelidiki pandangan publik tentang Aplikasi Mobile Identitas Kependudukan Digital 'Ditjen Dukcapil Kemendagri' di Google Play Store.

2.2. Ulasan

Ulasan merupakan suatu pendapat atau komentar terkait suatu peristiwa, sebagaimana dijelaskan dalam Kamus Besar Bahasa Indonesia (KBBI). Ulasan dapat diartikan sebagai rangkaian kalimat atau teks yang memberikan penilaian atau komentar terhadap karya yang dihasilkan oleh seseorang [12].

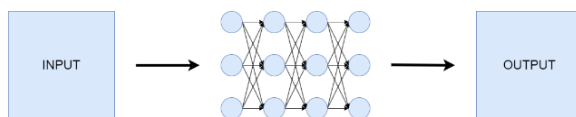
2.3. Identitas Kependudukan Digital

Identitas Kependudukan Digital (IKD) merupakan bentuk informasi elektronik yang dapat diakses melalui aplikasi pada smartphone. Inovasi ini diperkenalkan oleh Dirjen Kependudukan dan Catatan Sipil Kementerian Dalam Negeri untuk memanfaatkan proses digitalisasi serta mengatasi kendala pencetakan ulang KTP elektronik fisik yang memakan waktu lama. Sebagai solusi, masyarakat dapat menggunakan IKD sebagai alternatif pengganti KTP elektronik. Aplikasi ini dapat diunduh melalui platform Play Store atau App Store, dan aktivasinya melibatkan pemindaian barcode yang tersedia di Sistem Informasi Administrasi Kependudukan (SIAD) di kelurahan, kecamatan, atau layanan publik lain yang terkait dengan administrasi kependudukan. Keberadaan IKD mempermudah akses masyarakat dalam melakukan transaksi pelayanan publik dan privat hanya dengan menggunakan smartphone. Fungsi IKD meliputi sebagai bukti identitas, otentifikasi identitas, dan otorisasi identitas [13].

2.4. Deep Learning

Deep learning merupakan metode dari machine learning yang juga mampu untuk melakukan analisis sentiment. Metode deep learning membutuhkan waktu yang lebih lama pada saat training, tetapi kelebihan metode deep learning ialah dapat membuat model mempelajari konteks dari kata. Deep Learning digunakan dalam komputer untuk mengklasifikasikan gambar, suara, teks, atau video. Deep Learning melibatkan pengujian dan pelatihan komputer dengan menggunakan dataset yang memiliki label dan jumlah yang besar, kemudian mengubah data tersebut menjadi nilai piksel pada gambar sebagai representasi internal atau vektor fitur, kemudian menggunakan pengklasifikasi tersebut untuk mendeteksi pola pada input [5].

Deep Learning merupakan pembelajaran representasi untuk membangun arsitektur jaringan syaraf tiruan dengan lapisan banyak, termasuk lapisan input, lapisan tersembunyi, dan lapisan output. Lapisan tersembunyi dengan banyak lapisan dapat membuat komposisi algoritma dalam proses pembelajaran [14]. Ilustrasi dari deep learning dapat dilihat pada Gambar 1 di bawah ini.



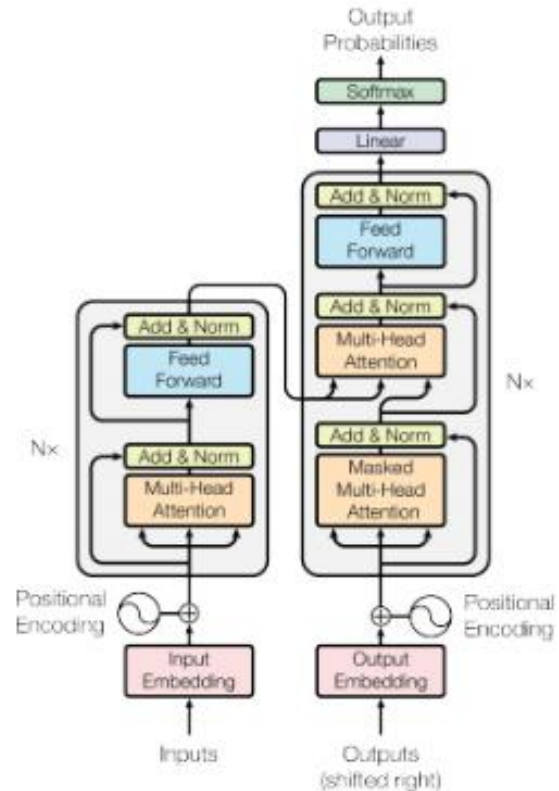
Gambar 1. Deep Learning

2.5. Bidirectional Encoder from Transformers (BERT)

Dalam ranah Deep Learning, model yang diperkenalkan oleh peneliti Google AI Language adalah BERT, yaitu model representasi bahasa terlatih. BERT menggunakan pendekatan Transformers, suatu sistem yang mengamati interaksi kata-kata dalam suatu teks dengan konteksnya [11]. BERT dapat diterapkan dalam berbagai masalah pemrosesan bahasa alami, termasuk klasifikasi teks. Pemanfaatan BERT dapat dilakukan melalui proses fine-tuning dengan menambahkan lapisan yang sesuai dengan output yang diinginkan. Pendekatan ini lebih efisien dalam penggunaan sumber daya komputer dibandingkan dengan pelatihan model BERT dari awal melalui tahap pre-training. Analisis sentimen, sebagai sub-masalah dari klasifikasi teks, dapat dianggap memiliki metode pelatihan yang serupa dengan klasifikasi teks secara umum menggunakan model BERT [15].

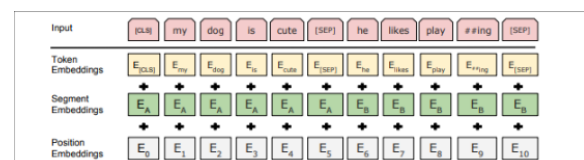
Arsitektur BERT terdiri dari multi-layer bidirectional Transformer encoder. Transformer adalah sebuah mekanisme attention yang mempelajari hubungan kontekstual diantara kata atau sub kata dalam teks. Transformer menggunakan mekanisme self-attention untuk memahami representasi input dan output. Transformer memiliki 2 mekanisme terpisah, yaitu encoder untuk membaca teks input dan decoder yang menentukan prediksi. Tiap layer encoder memiliki dua sub layer, yang pertama adalah multi-

head self-attention mechanism, dan yang kedua fully connected feed forward network. Pada encoder, setiap input akan melewati self-attention layer. Layer ini membantu encoder untuk melihat kata lain yang ada pada kalimat. Output dari attention layer kemudian digunakan sebagai input ke feed-forward neural network [5].



Gambar 2. Arsitektur Transformers

BERT menggunakan WordPiece embeddings dengan vocabulary sebanyak 30.000 token. Konsep WordPiece embedding awalnya dikembangkan untuk menangani masalah segmentasi bahasa Jepang dan Korea pada sistem pengenalan suara Google. Segmentasi ini dapat menghasilkan banyak kata yang tidak termasuk dalam vocabulary. WordPiece dilatih untuk secara otomatis memahami unit kata dari data besar tanpa menimbulkan masalah Out of Vocabulary (OOV). Kata-kata yang sering digunakan tetap diperlakukan sebagai satu unit kata utuh, sementara kata-kata yang jarang digunakan dipisahkan menjadi subkata bahkan karakter [5].

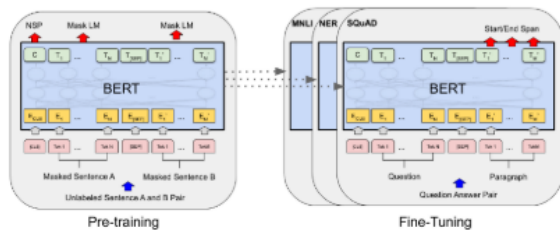


Gambar 3. Representasi Input Bert

Pada gambar 3 bisa dilihat representasi input BERT. Token pertama dari setiap urutan adalah token klasifikasi kelas khusus ([CLS]). Untuk membedakan setiap kalimat, token pemisah ([SEP]) ditambahkan

pada akhir setiap kalimat. Selanjutnya, segment embedding ditambahkan pada setiap token untuk membedakan kata-kata antara kalimat A dan kalimat B. Langkah berikutnya adalah menambahkan positional embedding untuk menandai posisi masing-masing token dalam kalimat. Input untuk encoder BERT diperoleh dengan menggabungkan Token Embedding, Segment embedding, dan Positional Embedding [5].

Framework BERT melibatkan dua tahap, yaitu tahap pre-training dan fine-tuning. Pada tahap pre-training, model mempelajari bahasa dan konteksnya dengan menjalankan dua tugas sekaligus, yaitu Mask Language Modelling (MLM) dan Next Sentence Prediction (NSP). MLM memungkinkan representasi untuk menyatukan konteks dari sisi kiri dan sisi kanan, memungkinkan pelatihan transformator dua arah yang mendalam. Sementara itu, NSP menggabungkan representasi pasangan teks yang telah menjalani pre-training [5].



Gambar 4. Ilustrasi Tahapan Pre-Training dan Fine-Tuning

2.6. Web Scraping

Salah satu metode untuk mengambil data yang tidak terstruktur dari suatu web secara otomatis dengan perintah program adalah web scraping. Biasanya, hasilnya disimpan dalam bentuk spreadsheet atau nilai komplemen (CSV) [12].

2.7. Confusion Matrix

Matriks kebingungan (confusion matrix) adalah sebuah tabel yang memuat informasi tentang data pengujian yang telah diklasifikasikan dengan benar serta data pengujian yang salah klasifikasinya. Matriks kebingungan juga digunakan sebagai metode untuk menilai kinerja model yang telah dibuat. Dalam matriks kebingungan, terdapat empat istilah utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) [16].

2.8. Accuracy

Akurasi (Accuracy) digunakan sebagai metrik untuk mengukur kinerja suatu algoritma dengan cara yang mudah dipahami. Biasanya, akurasi model dievaluasi setelah parameter-parameter model telah ditentukan, dan hasilnya dinyatakan dalam bentuk persentase. Di sisi lain, Validation Accuracy mengacu pada tingkat akurasi pada data uji (test set) dan mencerminkan sejauh mana model yang telah dibangun mampu melakukan prediksi dengan baik [18].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

2.9. Precision

Precision merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Precision merupakan metrik evaluasi yang digunakan dalam machine learning dan statistik untuk mengukur sejauh mana model dapat mengidentifikasi dengan tepat kelas positif [18].

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

2.10. Recall

Recall merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Recall merupakan metrik yang mengukur kemampuan model untuk mengidentifikasi semua objek dari kelas positif [18].

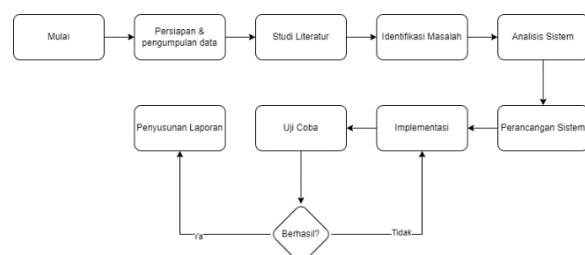
$$Recall = \frac{TP}{TP + FN} \quad (3)$$

2.11. F1 Score

F1 Score merupakan perbandingan rata-rata presisi dan recall yang dibobotkan. F1 Score merupakan metrik evaluasi yang menggabungkan precision dan recall dalam satu angka tunggal untuk memberikan gambaran menyeluruh tentang kinerja model klasifikasi. Metrik ini sangat berguna ketika terdapat ketidakseimbangan antara kelas positif dan negatif dalam data. F1 Score mencapai nilai maksimum 1 jika model memiliki presisi dan recall yang sempurna, dan nilai minimum 0 jika model tidak dapat mengklasifikasikan kelas positif dengan benar [18].

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

3. METODE PENELITIAN



Gambar 5. Tahapan Penelitian

Pada Gambar 5 di atas, terlihat tahapan-tahapan yang akan dilalui pada penelitian, yang mana tahapan tersebut adalah tahapan umum. Pertama, diawali dengan tahapan persiapan dan pengumpulan data ulasan aplikasi IKD untuk dijadikan dataset pada penelitian. Kemudian dilakukan studi literatur untuk mencari literatur-literatur yang berhubungan dan mendukung bagi penelitian. Kemudian dilanjutkan pada tahapan identifikasi masalah dalam penelitian yang akan dilakukan. Selanjutnya dilanjutkan dengan tahapan analisis, lebih tepatnya menganalisis sistem

yang akan digunakan dalam penelitian. Setelah dilakukan analisis sistem, selanjutnya melakukan perancangan sistem yang akan digunakan untuk analisis sentiment ulasan aplikasi IKD menggunakan BERT. Setelah berhasil merancang sistem, dilanjutkan dengan implementasi. Kemudian dilakukan uji coba sistem, jika berhasil maka dilanjutkan sampai penyusunan laporan.

3.1. Persiapan

Mempersiapkan keperluan dalam melakukan penelitian, yang dimana dalam hal ini mempersiapkan segala kebutuhan dalam melakukan penelitian.

3.2. Studi Pustaka

Pada tahap ini, proses mencakup pencarian referensi dan konsep dasar yang mendukung penelitian, yang akan membantu dalam pemahaman dan pengembangan sistem yang menerapkan metode Bidirectional Encoder Representation From Transformers (BERT).

3.3. Perencanaan

Tahap perencanaan merupakan langkah awal dalam penelitian yang bertujuan untuk menyusun rencana kerja secara sistematis serta membangun kerangka kerja yang terstruktur dan terorganisir, seperti yang ditunjukkan pada Gambar 5. Proses ini mencakup perumusan rencana penelitian, penentuan lokasi penelitian, jenis data yang akan dikumpulkan, serta pemilihan peralatan dan bahan yang diperlukan agar penelitian dapat menghasilkan temuan yang optimal dan bermanfaat.

3.4. Pengumpulan Data

Tahap awal dimulai dengan mengumpulkan data yang dibutuhkan. Data yang dibutuhkan dalam penelitian tugas akhir ini merupakan data opini pengguna aplikasi Identitas Kependudukan Digital di Google Play Store. Proses pengambilan data dilakukan menggunakan metode Web Scraping. Dataset ini kemudian disimpan dalam bentuk file csv untuk mempermudah proses pengolahan data. Dari hasil pengumpulan data didapatkan sebanyak 15.000 data. Supaya data dapat digunakan secara optimal, maka dilakukan pengurangan atribut dengan menghilangkan atribut user karena tidak digunakan dalam proses analisis sentimen.

3.5. Labelisasi Dataset

Pada tahap labelling, setiap ulasan aplikasi IKD diklasifikasikan ke dalam tiga kategori sentimen berdasarkan score (rating bintang) yang diberikan oleh pengguna di Google Play Store. Dalam penelitian ini, ulasan dengan rating 4 dan 5 dikategorikan sebagai sentimen positif, rating 3 sebagai sentimen netral, dan rating 1 serta 2 sebagai sentimen negatif. Proses labelling ini bertujuan untuk memberikan kelas yang jelas pada setiap ulasan sehingga dapat digunakan sebagai data latih dalam model BERT. Dengan adanya

label yang terdefinisi dengan baik, model dapat mempelajari pola sentimen dalam teks secara optimal dan menghasilkan prediksi yang lebih akurat.

3.6. Pre-processing Dataset

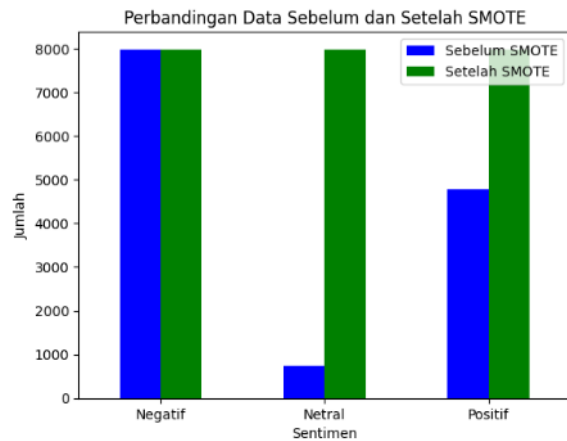
Pada tahap pre-processing, data ulasan aplikasi IKD yang diperoleh dari Google Play Store melalui teknik scraping akan dibersihkan dan dipersiapkan sebelum digunakan dalam analisis sentimen menggunakan BERT. Proses ini mencakup beberapa langkah utama, yaitu penghapusan karakter khusus, tanda baca, dan angka yang tidak relevan, normalisasi teks untuk menyamakan variasi kata, serta penghapusan kata-kata umum (stopwords) yang tidak berkontribusi pada analisis sentimen. Selain itu, dilakukan tokenization untuk memecah teks menjadi unit kata atau sub-kata yang sesuai dengan model BERT, serta padding agar panjang setiap ulasan seragam sebelum dimasukkan ke dalam model. Tahap pre-processing ini bertujuan untuk meningkatkan kualitas data sehingga model dapat memahami pola sentimen dengan lebih baik dan menghasilkan prediksi yang lebih akurat.

3.7. Pembagian Dataset

Pada tahap split data, dataset dibagi menjadi data latih (training set) dan data uji (testing set) dengan berbagai rasio untuk mengevaluasi kinerja model secara lebih komprehensif. Dalam penelitian ini, data dibagi dengan beberapa skenario, mulai dari 90% untuk data latih dan 10% untuk data uji, hingga 10% untuk data latih dan 90% untuk data uji. Variasi rasio ini digunakan untuk menganalisis sejauh mana jumlah data latih memengaruhi performa model BERT dalam melakukan prediksi sentimen. Data latih digunakan untuk membangun model, sementara data uji berfungsi untuk mengukur akurasi dan keandalan model dalam mengklasifikasikan ulasan aplikasi IKD yang belum pernah dilihat sebelumnya.

3.8. Penyeimbangan Data dengan SMOTE

Pada tahap ini, dilakukan penyeimbangan data menggunakan Synthetic Minority Over-sampling Technique (SMOTE) untuk mengatasi ketidakseimbangan kelas dalam dataset. SMOTE bekerja dengan membuat sampel sintetis pada kelas minoritas agar distribusi data lebih seimbang, sehingga model BERT dapat belajar lebih baik dalam mengenali pola sentimen. Hasil penerapan SMOTE dapat dilihat pada Gambar 6, yang menunjukkan peningkatan jumlah data pada kelas netral dan positif agar setara dengan kelas negatif. Sebelum SMOTE, jumlah data netral jauh lebih sedikit dibandingkan kelas lainnya, yang berpotensi menyebabkan bias dalam model. Setelah SMOTE diterapkan, distribusi data menjadi lebih seimbang di semua kelas.



Gambar 6. Perbandingan Data Sebelum dan Sesudah SMOTE

Selain itu, Tabel 1 menyajikan jumlah data sebelum dan sesudah proses oversampling. Setelah penerapan SMOTE, setiap kelas memiliki 7.980 data, sehingga model dapat belajar dengan lebih adil tanpa dipengaruhi oleh dominasi salah satu kelas.

Tabel 1. Tabel Perbandingan Sebelum dan Setelah SMOTE

Sentimen	Sebelum SMOTE	Setelah SMOTE
Negatif	7980	7980
Netral	741	7980
Positif	4779	7980

3.9. Implementasi BERT

Penelitian ini menggunakan BERT untuk klasifikasi sentimen ulasan aplikasi IKD. Proses diawali dengan tokenisasi menggunakan BERT tokenizer, yang mengubah teks ulasan menjadi token numerik dengan panjang maksimum 128 token. Setiap token diberi input ID dan attention mask untuk memastikan pemrosesan yang optimal oleh model. Data ulasan kemudian dipisahkan dengan berbagai rasio, mulai dari 90% training : 10% testing hingga 10% training : 90% testing, dengan hasil terbaik pada 70% training : 30% testing. Ketimpangan jumlah kelas dalam data diatasi dengan SMOTE, yang menyeimbangkan distribusi sentimen negatif, netral, dan positif agar model dapat belajar lebih efektif.

Model yang digunakan adalah BERT Base Uncased dengan BertForSequenceClassification, yang memiliki 12 lapisan Transformer dan 110 juta parameter. Model dilatih menggunakan optimizer AdamW, learning rate $2e-5$, batch size 16, dan Cross-Entropy Loss selama 10 epoch. Pelatihan dilakukan dengan GPU untuk mempercepat proses komputasi. Setelah pelatihan, model BERT mampu mengklasifikasikan sentimen ulasan secara otomatis ke dalam tiga kategori, yaitu negatif, netral, dan positif. Implementasi ini membuktikan bahwa pendekatan berbasis transformer efektif dalam analisis sentimen, memberikan hasil klasifikasi yang lebih baik dibandingkan metode konvensional.

3.10. Pengujian dan Evaluasi

Setelah tahapan implementasi diselesaikan, langkah selanjutnya adalah pengujian dan evaluasi. Pengujian dilakukan dengan menguji model BERT yang telah dilatih menggunakan data uji untuk menilai kemampuannya dalam mengklasifikasikan sentimen ulasan aplikasi IKD. Pada tahap ini, hasil evaluasi dianalisis untuk menginterpretasikan kinerja model dalam memahami dan memprediksi sentimen berdasarkan teks ulasan yang tersedia.

4. HASIL DAN PEMBAHASAN

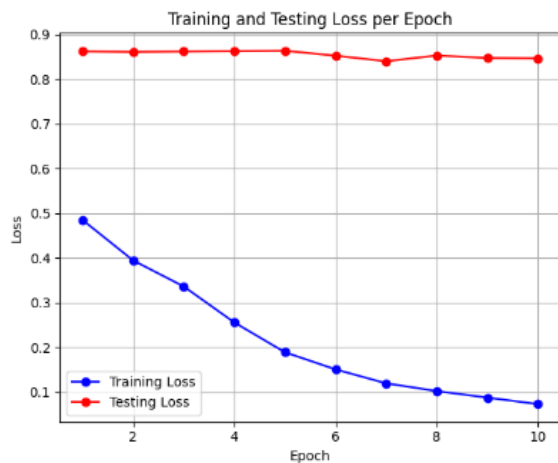
Penelitian ini menggunakan model Bidirectional Encoder Representations from Transformers (BERT) untuk melakukan analisis sentimen terhadap ulasan aplikasi Identitas Kependudukan Digital (IKD) yang diperoleh dari Google Play Store. Data yang digunakan berjumlah 15.000 ulasan, yang telah melalui tahap preprocessing untuk menghilangkan elemen yang tidak relevan, seperti tanda baca, angka, dan kata tidak baku. Selain itu, metode Synthetic Minority Over-sampling Technique (SMOTE) diterapkan untuk menyeimbangkan distribusi kelas sentimen dalam dataset agar model dapat belajar dengan lebih optimal. Dalam tahap pelatihan, data dibagi menjadi beberapa skenario rasio training dan testing untuk mengevaluasi kinerja model.

Hasil dari setiap skenario dapat dilihat pada Tabel 2 berikut:

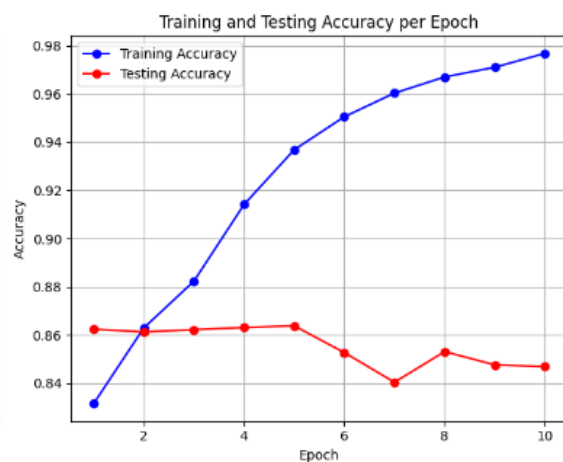
Tabel 2. Hasil Akurasi Model pada Berbagai Rasio Data Training dan Testing

Rasio Training : Testing	Akurasi (%)
90% : 10%	83.67%
80% : 20%	84.63%
70% : 30%	84.69%
60% : 40%	83.07%
50% : 50%	83.73%
40% : 60%	83.94%
30% : 70%	84.30%
20% : 80%	81.73%
10% : 90%	79.61%

Dari hasil di atas, skenario 70% training : 30% testing memberikan akurasi tertinggi 84.69%, menunjukkan bahwa rasio ini memberikan keseimbangan optimal antara jumlah data latih dan uji, di mana model memiliki cukup data untuk belajar sekaligus diuji secara komprehensif. Pada awal pelatihan, training loss cenderung tinggi, tetapi menurun seiring bertambahnya epoch, menunjukkan bahwa model semakin memahami pola dalam data. Visualisasi Training Loss dan Accuracy Loss menunjukkan tren pembelajaran model yang stabil dan menunjukkan peningkatan akurasi seiring waktu.

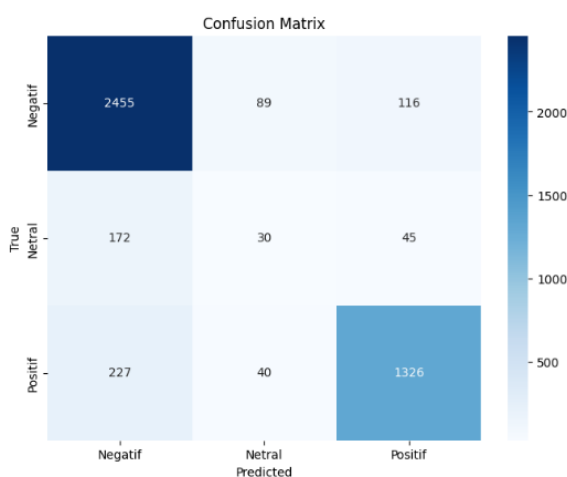


Gambar 7. Training dan Testing Loss per epoch



Gambar 8. Training dan Testing Accuracy per Epoch

Untuk mengevaluasi kinerja model lebih lanjut, digunakan Confusion Matrix. Hasil menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik, meskipun masih terdapat beberapa kesalahan dalam membedakan ulasan netral dengan ulasan positif dan negatif.



Gambar 9. Confusion Matrix

Selain itu, tabel berikut menampilkan metrik evaluasi berdasarkan precision, recall, dan F1-score untuk masing-masing kelas sentiment

Tabel 3. Metrik Evaluasi Klasifikasi Model BERT

Kelas	Precision	Recall	F1-Score	Support
0	0.86	0.92	0.89	2660
1	0.19	0.12	0.15	247
2	0.89	0.83	0.86	1593
Akurasi Keseluruhan			84.69%	
Macro Avg	0.65	0.63	0.63	4500
Weighted Avg	0.83	0.85	0.84	4500

Hasil penelitian ini kemudian dibandingkan dengan studi sebelumnya yang menggunakan algoritma K-Nearest Neighbors (KNN) untuk analisis sentimen ulasan IKD dan memperoleh akurasi 81%. Dibandingkan dengan metode tersebut, model BERT dalam penelitian ini menunjukkan peningkatan akurasi sebesar 3.69%. Hal ini menegaskan bahwa BERT, dengan kemampuannya memahami konteks dua arah dalam teks, memiliki keunggulan dalam analisis sentimen dibandingkan dengan metode klasik seperti KNN. Selain itu, penerapan SMOTE dalam penelitian ini berkontribusi dalam menyeimbangkan dataset, yang membantu model dalam mengenali pola sentimen dengan lebih baik. Dalam studi sebelumnya, KNN cenderung mengalami penurunan akurasi akibat ketidakseimbangan kelas dalam dataset, sedangkan dalam penelitian ini, metode oversampling telah mengurangi dampak ketidakseimbangan tersebut. Dengan hasil ini, dapat disimpulkan bahwa model BERT memiliki performa yang lebih unggul dibandingkan metode tradisional seperti KNN dalam analisis sentimen ulasan aplikasi IKD. Rasio 70% training : 30% testing menghasilkan akurasi tertinggi 84.69%, menunjukkan keseimbangan optimal dalam pembagian data. Penggunaan SMOTE juga terbukti membantu dalam meningkatkan kinerja model. Dengan hasil ini, model BERT dapat menjadi alternatif yang lebih efektif untuk menganalisis sentimen aplikasi layanan publik berbasis teks.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa model BERT mampu melakukan analisis sentimen ulasan aplikasi Identitas Kependudukan Digital (IKD) dengan akurasi tinggi, mencapai 84,69% pada rasio data training 70% dan testing 30%. Penerapan teknik Synthetic Minority Over-sampling Technique (SMOTE) berhasil menyeimbangkan distribusi kelas sentimen, sehingga model dapat mengenali pola opini pengguna dengan lebih baik. Keunggulan BERT dalam memahami konteks bahasa secara dua arah menjadikannya lebih efektif dibandingkan metode konvensional dalam analisis sentimen berbahasa Indonesia. Hasil penelitian ini memberikan wawasan bagi pengembang aplikasi untuk meningkatkan kualitas layanan

berdasarkan sentimen pengguna, serta menunjukkan bahwa pendekatan berbasis deep learning dapat menjadi solusi yang lebih akurat dalam menganalisis opini publik. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan beragam guna meningkatkan generalisasi model, serta mengeksplorasi teknik lain seperti transfer learning atau aspek-based sentiment analysis untuk analisis yang lebih mendalam.

DAFTAR PUSTAKA

- [1] R. Hidayat, R. N. Rahman, M. R. Perdana, and Arbansyah, "Analisis Sentimen Aplikasi Identitas Kependudukan Digital (IKD) Menggunakan Metode Naïve Bayes," *Jurnal Sistem Informasi dan Ilmu Komputer*, vol. 2, pp. 129–140, 2024. [Online]. Available: <https://doi.org/10.59581/jusiik-widyakarya.v2i1.2320>.
- [2] P. Aditiya, U. Enri, and I. Maulana, "Analisis Sentimen Ulasan Pengguna Aplikasi MyIM3 pada Situs Google Play Menggunakan Support Vector Machine," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, p. 1020, 2022.
- [3] Y. S. V. Putranta, B. Rahayudi, and W. Purnomo, "Analisis Sentimen Masyarakat terhadap Kebijakan Penghapusan Subsidi BBM pada Media Sosial Twitter menggunakan Algoritma Naive Bayes Classifier dengan Ekstraksi Fitur N-Gram TF-IDF," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, pp. 1501–1510, 2023.
- [4] B. Rahmatullah, *Sentiment Analysis Pelaksanaan Work From Home di Indonesia pada Masa Pandemi COVID-19 Menggunakan IndoBERT*. Surabaya: Institut Teknologi Sepuluh Nopember, 2021.
- [5] A. Yahyadi and F. Latifah, "Analisis Sentimen Twitter terhadap Kebijakan PPKM di Tengah Pandemi COVID-19 menggunakan Model LSTM," *Journal of Information System, Applied, Management, Accounting and Research*, vol. 6, pp. 464–470, 2022.
- [6] R. M. R. W. P. K. Atmaja and W. Yustanti, "Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers)," *JEISBI (Journal of Emerging Information Systems and Business Intelligence)*, vol. 2, pp. 55–62, 2021.
- [7] B. Kurniawan, A. A. Aldino, and A. R. Isnain, "Sentimen Analisis Terhadap Kebijakan Penyelenggara Sistem Elektronik (PSE) Menggunakan Algoritma Bidirectional Encoder Representations from Transformers (BERT)," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 3, pp. 98–106, 2022.
- [8] V. Chandradev, I. M. A. D. Suarjaya, and I. P. A. Vayupati, "Analisis Sentimen Review Hotel Menggunakan Metode Deep Learning BERT," *Jurnal Buana Informatika*, vol. 14, pp. 107–116, 2023.
- [9] D. Pradipta, Kusri, and H. Al Fatta, "Sentimen Analisis Komentar Covid-19 Varian Omicron pada Media Sosial Instagram dengan Bidirectional Encoder Representations from Transformers (BERT)," *JTECS: Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem & Komputer*, vol. 3, 2023. [Online]. Available: <https://doi.org/10.32503/jtecs.v3i1.3219>.
- [10] D. D. Audiansyah, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Aplikasi MyXL menggunakan Metode Support Vector Machine berdasarkan Ulasan Pengguna di Google Play Store," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, pp. 3987–3994, 2022.
- [11] V. S. Bella and D. Widodo, "Implementasi Aplikasi Identitas Kependudukan Digital (IKD) dalam Menunjang Pelayanan Publik Masyarakat di Kecamatan Tambaksari," *Jurnal SARAQ OPAT*, vol. 6, pp. 14–31, 2024. [Online]. Available: <https://doi.org/10.55542/saraqopat.v6i1.833>.
- [12] B. Yanto, L. Fimawahib, A. Supriyanto, B. H. Hayadi, and R. R. Pratama, "Klasifikasi Tekstur Kematangan Buah Jeruk Manis Berdasarkan Tingkat Kecenderungan Warna dengan Metode Deep Learning Convolutional Neural Network," *INOVTEK Polbeng - Seri Informatika*, vol. 6, p. 259, 2021. [Online]. Available: <https://doi.org/10.35314/isi.v6i2.2104>.
- [13] M. A. Abdurrazzaq and E. L. Tjong, "Analisis Sentimen KUHP Baru Pada Data Twitter Menggunakan Model BERT," *Jurnal Komunikasi, Sains dan Teknologi*, vol. 1, pp. 38–43, 2022.
- [14] A. Faizin, M. Lutfi, and Achmyatari, "Perbandingan Arsitektur Lenet dan Googlenet dalam Klasifikasi Diabetic Retinopathy pada Citra Retina Fundus," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, pp. 342–347, 2022. [Online]. Available: <https://doi.org/10.36040/jati.v6i1.4581>.
- [15] T. Setiawan and D. Avianto, "Implementasi Convolutional Neural Network untuk Pengenalan Warna Kendaraan," *Naskah Publikasi FTIE*, Universitas Teknologi Yogyakarta, 2020.
- [16] P. A. Nugroho, I. Fenriana, and R. Arijanto, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) pada Ekspresi Manusia," *Algor*, vol. 2, pp. 12–21, 2020.