Smart Comp: Jurnalnya Orang Pintar Komputer
Volume 15, Nomor 3, Juli 2026, hlm. 626-637
Terakreditasi Sinta, Peringkat 4, SK No. 295/C/C3/KPT/2026
DOI : 10.30591/smartcomp.v15i3.9669

P-ISSN : 2089-676X
E-ISSN : 2549-0796

Deteksi Ulasan Palsu Berbahasa Indonesia Menggunakan Model DistilBERT: Studi Kasus pada *Platform E-Commerce*

Nurani Istiaen[*1], Ivana Lucia Kharisma[2], Somantri[3], Kamdan[4], Elsy Rahajeng[5]
[1,2,3,4)]Teknik Informatika, Universitas Nusa Putra
[5)]Sistem Informasi, Universitas Islam Negeri Syarif Hidayatullah
Email: [*1]nurani.istiaeni_ti21@nusaputra.ac.id, [2]ivana.lucia@nusaputra.ac.id,
[3]somantri@nusaputra.ac.id, [4]kamdan@nusaputra.ac.id, [5]elsy.rahajeng@uinjkt.ac.id

(Naskah masuk: *26 Agustus 2025*, diterima untuk diterbitkan: *12 Juni 2026*)

Abstrak: *Kemajuan teknologi telah mempermudah berbagai aspek kehidupan manusia, termasuk di sektor bisnis dan ekonomi. Kehadiran e-commerce memungkinkan interaksi antara penjual dan pembeli berlangsung lebih efektif serta efisien. Dalam proses pengambilan keputusan pembelian, ulasan pelanggan memiliki peran penting sebagai sumber informasi mengenai kualitas produk maupun layanan. Oleh karena itu, keaslian ulasan sangat krusial dalam menjaga kepercayaan pengguna sekaligus integritas platform. Penelitian ini menggunakan dataset berisi 887 ulasan berbahasa Indonesia yang dikumpulkan dari salah satu toko di satu platform e-commerce. Sistem deteksi ulasan palsu dibangun dengan memanfaatkan model DistilBERT berbasis pemrosesan bahasa alami (NLP) yang diawali dengan tahapan preprocessing. Model dilatih dalam beberapa skenario, baik tanpa maupun dengan teknik oversampling, lalu dievaluasi menggunakan accuracy, precision, recall, F1-Score, confusion matrix, serta ROC-AUC. Hasil evaluasi menunjukkan bahwa oversampling memberikan performa terbaik dengan akurasi 97% dan ROC-AUC 0.97–0.99. Model mampu mengenali ulasan asli (OR) dengan presisi 0.99 serta mendeteksi ulasan palsu (CG) dengan recall 0.96. Pada skenario 50 epoch, ROC-AUC tertinggi dicapai dengan nilai 0.9921. Selanjutnya, model diimplementasikan ke dalam aplikasi web berbasis Streamlit untuk membantu pengguna mengambil keputusan pembelian yang lebih bijak di e-commerce.*

Kata Kunci – *Ulasan Palsu; DistilBERT; Pemrosesan Bahasa Alami; Perdagangan Elektronik; Streamlit*

Fake Review Detection in the Indonesian Language Using the DistilBERT Model: A Case Study on an E-Commerce Platform

Abstract: *Technological advancements have simplified various aspects of human life, including the business and economic sectors. The emergence of e-commerce enables interactions between sellers and buyers to take place more effectively and efficiently. In the purchasing decision-making process, customer reviews play an important role as a source of information regarding product and service quality. Therefore, the authenticity of reviews is crucial for maintaining user trust as well as the integrity of the platform. This research utilizes a dataset containing 887 Indonesian-language reviews collected from one store on an e-commerce platform. A fake review detection system was developed using the DistilBERT model based on natural language processing (NLP), beginning with preprocessing stages. The model was trained under several scenarios, both with and without oversampling techniques, and then evaluated using accuracy, precision, recall, F1-Score, confusion matrix, and ROC-AUC. The evaluation results indicate that oversampling produced the best performance, achieving 97% accuracy and an ROC-AUC between 0.97 and 0.99. The model was able to identify genuine reviews (OR) with a precision of 0.99 and detect fake reviews (CG) with a recall of 0.96. In the 50-epoch scenario, the highest ROC-AUC was achieved at 0.9921. Furthermore, the model was implemented into a Streamlit-based web application to assist users in making wiser purchasing decisions on e-commerce platforms.*

Keywords – *Fake Review; DistilBERT; Natural Language Processing (NLP); E-Commerce; Streamlit*

1. PENDAHULUAN

Kemajuan teknologi dan kemudahan informasi saat ini memengaruhi semua aspek kehidupan. Misalnya, toko online adalah contoh bisnis yang menggunakan sosial media selain untuk mendapatkan informasi dan unjuk diri. Seiring perkembangan teknologi, banyak situs jual beli online menggabungkan berbagai toko online menjadi satu situs, membuat pembeli lebih mudah

mendapatkan barang yang mereka inginkan [1]. Selain itu, istilah "toko online" telah berkembang menjadi "e-commerce", yang mencakup seluruh ekosistem perdagangan elektronik. Ketika banyak platform e-commerce muncul, konsumen dapat dengan mudah mendapatkan berbagai produk dan layanan [2]. Selain itu, pembayaran digital menjadi lebih mudah saat berbelanja online [3].

Menurut laporan yang dirilis oleh Center of Economic and Law Studies pada 22 Desember 2024 menunjukkan bahwa transaksi pembayaran digital di Indonesia dapat mencapai nilai Rp2.491,68 triliun hanya pada tahun 2024 [4]. Capaian tersebut mencerminkan pesatnya pertumbuhan aktivitas ekonomi digital, termasuk dalam sektor e-commerce. Namun, ketika seseorang ingin mempertimbangkan untuk melakukan transaksi melalui e-commerce, salah satu cara untuk mendapatkan informasi tentang produk dan layanan adalah dengan melihat ulasan yang ada. Dari ulasan itu, seseorang dapat mengetahui kualitas dan kinerja produk tersebut sebelumnya [5]. Ulasan tentang produk adalah faktor penting yang memengaruhi keputusan pembelian, ulasan pelanggan lain memberikan informasi berharga tentang kualitas juga keunggulan. Ulasan membantu calon pembeli memutuskan apakah suatu barang layak dibeli atau tidak. karena kredibilitas dan kebenaran ulasan sangat penting untuk menjaga kredibilitas platform e-commerce [6].

Munculnya ulasan palsu menjadi masalah besar bagi industri e-commerce. Ulasan palsu sering dibuat untuk menipu pelanggan dan meningkatkan penjualan produk tertentu, baik oleh penjual yang tidak jujur maupun pihak ketiga yang ingin merusak reputasi competitor [7]. E-commerce, penginapan, restoran, dan layanan kesehatan adalah beberapa industri di mana ulasan palsu berbasis AI muncul. Menurut Transparency Company, perusahaan pengawas teknologi Amerika, ulasan AI telah meningkat sejak pertengahan 2023. Dalam laporan terbarunya, Transparency Company menyelidiki 73 juta ulasan tentang hukum, kesehatan, dan rumah tangga. Hasilnya menunjukkan bahwa hampir 14 persen dari ulasan tersebut terindikasi palsu, dan sekitar 2,3 juta di antaranya dianggap dibuat oleh kecerdasan buatan sepenuhnya atau sebagian. Ini berdampak negatif pada penjual etis dan membuat konsumen membuat pilihan yang salah [8]. Maka dari itu, sistem yang efisien diperlukan untuk mengidentifikasi ulasan palsu dan melindungi pelanggan dan penjual yang jujur.

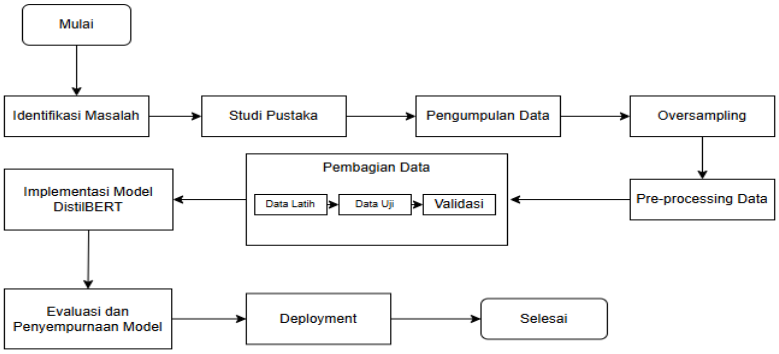
Berbagai metode digunakan dalam penelitian deteksi informasi palsu. Support Vector Machine (SVM), Naïve Bayes [2], dan Random Forest [9] adalah beberapa metode yang digunakan, serta model berbasis transformer seperti BERT dan RoBERTa [7]. Studi penelitian yang membandingkan model transformer yang diajarkan menemukan bahwa RoBERTa memiliki akurasi tertinggi (90,8%), sedangkan DistilBERT unggul dalam efisiensi waktu. Pengujian dalam penelitian SVM dan Naïve Bayes menunjukkan bahwa meningkatkan jumlah dataset dan menyeimbangkan data dapat meningkatkan akurasi hasil pengujian. Hasil penelitian menggunakan model Random Forest menunjukkan bahwa model Transformer menangani masalah deteksi ulasan palsu pada dataset Ott lebih baik daripada model Deep Learning. Namun, hasil Deep Learning masih cukup baik, sehingga disarankan untuk menggunakan dataset yang lebih besar. Kemampuan BERT untuk memahami konteks dan bahasa yang kompleks, klasifikasi berita palsu sangat akurat pada penelitian [10]. Penelitian ini menggunakan model DistilBERT untuk mendeteksi ulasan palsu pada salah satu toko e-commerce. Model ini termasuk dalam kategori deep learning dan merupakan bagian dari machine learning berbasis Natural Language Processing (NLP) dan merupakan versi lebih ringan dari BERT (Bidirectional Encoder Representations from Transformers) [11]. Tujuan dari penelitian ini adalah untuk membuat proses lebih cepat. Bahkan dengan jumlah data yang terbatas, model ini dapat digunakan untuk menentukan apakah ulasan adalah asli atau palsu berdasarkan pola linguistik yang ada karena memahami konteks dan sifat bahasa yang terkandung dalam teks [12].

Penelitian ini akan menggunakan data ulasan pelanggan dari toko Herbilogy di Tokopedia, platform e-commerce lokal. Dataset ini sebelumnya digunakan untuk mendeteksi ulasan palsu menggunakan SVM dan Naive Bayes, menemukan bahwa SVM memiliki akurasi lebih tinggi sebesar 94,38% dalam mengklasifikasikan ulasan untuk salah satu toko di Tokopedia [2]. Diharapkan bahwa dataset ini akan memungkinkan model DistilBERT untuk dilatih untuk mengidentifikasi fitur yang membedakan ulasan asli dari palsu. Untuk mengetahui efektivitas

model, setelah pelatihan selesai, kinerjanya akan dinilai menggunakan metrik accuracy, precision, recall, dan f1-score. Selain itu, untuk membedakan tiap kelas dan mengevaluasi jumlah prediksi yang benar dan salah, confusion matrix dan ROC-AUC akan digunakan. Hasil yang diharapkan dari penelitian ini adalah pengembangan sistem deteksi ulasan palsu yang akurat dan efisien menggunakan model DistilBERT.

2. METODE PENELITIAN

Tahapan – tahapan penelitian yang dilakukan oleh penulis digambarkan pada sebuah flowchart pada gambar dibawah ini:



Gambar 1. Tahapan Penelitian

2.1. Identifikasi Masalah

Selama proses ini, penulis melakukan penyelidikan dan mempelajari informasi dari masalah yang muncul. Tujuan dari penelitian ini adalah untuk memperoleh informasi yang mendalam dan mendalam yang dapat digunakan sebagai dasar untuk penelitian dan pengembangan sistem.

2.2. Studi Pustaka

Peneliti mencari data secara manual pada sumber informasi dan data yang terkait dengan konsep, analisis, dan simulasi untuk melakukan studi pustaka. Sumber-sumber ini termasuk buku, jurnal, penelitian serupa, dan website yang dianggap dapat diandalkan.

2.3. Pengumpulan Data

2.3.1. Observasi

Dalam pengumpulan data, observasi melibatkan pengamatan langsung fenomena atau peristiwa yang terjadi [13]. Tujuan observasi ini adalah untuk memverifikasi atau memahami situasi yang sedang terjadi, seperti banyaknya ulasan palsu yang dapat memengaruhi persepsi konsumen dan kepercayaan mereka terhadap produk dan toko [14]. Proses observasi ini juga menghasilkan data yang nyata, jelas, dan dapat diandalkan. Pengumpulan data yang dilakukan oleh penulis dalam penelitian deteksi ulasan palsu diantaranya memilih dataset ulasan penjualan toko Herbilogy dari platform Tokopedia yang berisi 887 data ulasan yang di scraping pada tanggal 4 November tahun 2022 dan sebelumnya juga pernah digunakan untuk penelitian, sehingga penulis ingin mengembangkan dan memperkuat penelitian ini dengan menggunakan model yang berbeda.

2.3.2. Interview

Penulis melakukan verifikasi manual dataset dan kemudian mewawancarai pemilik data yang juga seorang anotator dengan pengalaman tiga tahun di bidang data dan juga pernah melakukan penelitian tentang subjek ulasan palsu sebelumnya.

2.3.3. Tinjauan Literatur

Penulis juga melakukan tinjauan literatur dan mendapatkan berbagai informasi dengan membaca jurnal di berbagai platform yang terkait dengan pembahasan yang diperlukan untuk referensi dalam penulisan ini. Tujuannya adalah untuk menghasilkan penelitian terbaru, memvalidasi teori, dan membantu akademisi membuat keputusan berdasarkan bukti [15].

2.4. Oversampling

Ketidakseimbangan jumlah, atau ketidakseimbangan data, dapat menyebabkan model hanya belajar dari kelas yang jumlahnya lebih dominan, yang merupakan masalah besar bagi penelitian ini. Kondisi tersebut dapat menyebabkan model menjadi bias terhadap kelas mayoritas dan mengabaikan ciri-ciri kelas minoritas. Akibatnya, kemampuan model untuk membedakan kelas menjadi sangat buruk. Untuk mengatasi masalah ini, sebuah teknik oversampling yang menambahkan jumlah data ke kelas minoritas agar jumlahnya sebanding dengan kelas mayoritas digunakan. Oversampling dapat memperbesar data minoritas [16]. Tujuannya adalah agar model dapat melakukan klasifikasi yang lebih adil dan seimbang dengan mendapatkan data yang cukup dari kedua kelas.

Ketika data sudah berimbang, diharapkan bahwa performa model terhadap kedua kelas akan meningkat, terutama dalam nilai recall dan f1-score untuk kelas minoritas [17].Tujuan analisis adalah untuk mengejar akurasi keseluruhan dan memastikan bahwa perhatian yang seimbang diberikan kepada setiap kelas [18].

2.5. Pre-processing Data

Preprocessing adalah tahap yang paling penting karena menentukan kualitas data yang digunakan. Preprocessing membersihkan dan mengatur data acak semula untuk menjadi data yang relevan dan rapi [19]. Cleaning, case folding dan tokenizing adalah tahapan pre-processing data yang dilakukan pada penelitian ini.

2.6. Pembagian Data

Data yang berjumlah 887 ulasan terdiri dari data berlabel palsu (Computer Generated) sebanyak 655 data dan data berlabel asli (Original) sebanyak 155 data. Nantinya dataset dibagi menjadi tiga bagian, yaitu 70% data latih (training), 15% data uji (testing) dan 15% data validasi (validation) untuk mendapatkan hasil yang optimal [20].

2.7. Implementasi dan Training Model

Pada tahap pelatihan model, parameter seperti bobot dan bias diperbarui berdasarkan data pelatihan selama fase pelatihan model. Pembaruan ini dilakukan berulang kali untuk mengurangi kesalahan prediksi model terhadap data. Algoritma optimisasi Adam (Adaptive Moment Estimation), salah satu algoritma optimasi yang paling populer dan banyak digunakan dalam pelatihan model deep learning, digunakan untuk mengontrol dan mengarahkan pembaruan parameter tersebut. Algoritma ini berfungsi untuk mengurangi nilai fungsi kerugian, yang mengukur seberapa jauh prediksi model sesuai dengan nilai sebenarnya pada data latih.

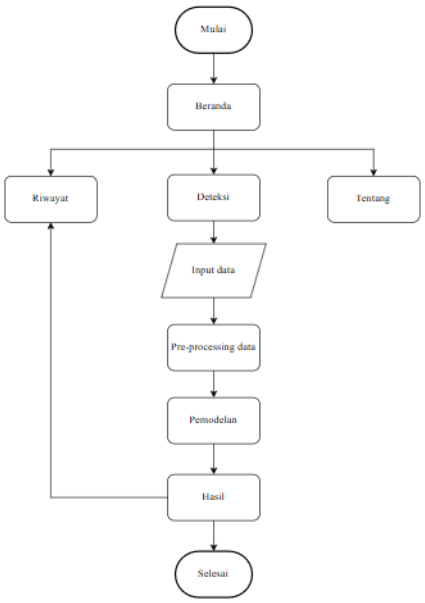
Setiap epoch menunjukkan satu kali putaran model penuh melalui seluruh dataset pelatihan [21]. Model melakukan beberapa iterasi terhadap batch data di setiap epoch. Secara bertahap, parameter model diperbarui berdasarkan gradien yang dihitung dari fungsi kerugian. Dengan cara ini, model dapat belajar pola dan hubungan dalam data dengan lebih baik. Memilih jumlah epoch yang tepat sangat penting karena jika model dilatih dengan terlalu sedikit epoch, model mungkin tidak belajar dengan baik dan hasil prediksinya mungkin tidak akurat (underfitting). Sebaliknya, jika terlalu banyak epoch digunakan, model dapat menghafal data latih secara berlebihan, menyebabkan kinerja yang buruk pada data baru atau data uji [22]. Untuk mencegah overfitting, kinerja model dipantau dengan data validasi yang tidak digunakan dalam pelatihan langsung. Teknik awal penundaan dapat digunakan, misalnya menghentikan pelatihan ketika kinerja data validasi mulai menurun meskipun kinerja data latih terus meningkat. Dengan cara ini, model dapat menjaga keseimbangan antara kemampuan generalisasi ke data baru dan jumlah belajar yang cukup. Untuk membuat prediksi yang akurat dan tahan terhadap perubahan data yang terjadi di dunia nyata, penggunaan algoritma optimisasi yang tepat dan kontrol jumlah epoch sangat penting dalam proses pelatihan model.

2.8. Evaluasi Model

Setelah proses training selesai, langkah berikutnya adalah melakukan evaluasi terhadap performa model untuk mengetahui sejauh mana model mampu melakukan prediksi terhadap data yang belum pernah dilihat sebelumnya, yaitu data uji. Evaluasi ini sangat penting karena bertujuan untuk mengukur kemampuan generalisasi model. Beberapa metrik evaluasi yang akan digunakan untuk mengukur performa model adalah accuracy, precision, recall, f1-score, confusion matrix dan ROC-AUC.

2.9. Deployment

Untuk implementasi, Streamlit akan digunakan sebagai antarmuka web interaktif yang memungkinkan pengguna mengakses dan menguji model deteksi ulasan palsu secara langsung melalui tampilan sederhana.



Gambar 2. Alur Sistem Web

3. HASIL DAN PEMBAHASAN

3.1. Pengolahan Data

Meskipun dataset yang digunakan dalam penelitian ini terdiri dari 887 baris data, terdapat 77 data duplikat yang harus imputasi terlebih dahulu. Selanjutnya, diketahui bahwa distribusi kelasnya tidak seimbang, kelas CG memiliki jumlah data yang jauh lebih besar, yaitu 655 dibandingkan dengan kelas OR, yang hanya memiliki 155. Kondisi ini dapat menyebabkan model bias terhadap kelas mayoritas (CG) dan mengabaikan karakteristik kelas minoritas (OR). Akibatnya, kinerja model menjadi sangat buruk ketika mencari kelas OR. Dengan melakukan oversampling, data dari kelas OR diperbanyak. Ini dilakukan dengan menyalin data secara acak menggunakan pustaka Python, `pandas` dan `sklearn.utils.resample`. Berikut perbandingan jumlah data sebelum dan sesudah dilakukan oversampling:

Tabel 1. Pembagian data sebelum dan sesudah oversampling

Pembagian Data	Sebelum	Sesudah
Data latih (70%)	567	917
Data uji (15%)	122	197
Data validasi (15%)	121	196
Jumlah Data	**887**	**1.310**

3.2. Text Processing

Text processing yang dilakukan adalah cleaning dan case folding untuk memastikan teks bersih dan seragam sebelum dimasukkan ke dalam model. Meskipun DistilBERT dapat mengolah

teks mentah, terlalu banyak simbol yang tidak relevan dalam data dapat memengaruhi hasil prediksi dan mengurangi akurasi model [23]. Teks promosi dengan tautan atau simbol seperti `!!!` dapat menjadi indikasi penting; namun, jika terlalu banyak atau tidak konsisten, itu membuat lebih sulit bagi model untuk mengenali pola [7]. Karena model dan tokenizer sudah mampu mengenali struktur bahasa, proses pengolahan teks kontemporer berbasis Natural Language Processing (NLP) seperti DistilBERT tidak memerlukan pengolahan teks yang terlalu kompleks [24].

3.3. Pemodelan DistilBERT

3.3.1. DistilBertTokenizer

Setelah text processing, langkah selanjutnya yaitu pemodelan yang diawali dengan proses tokenisasi menggunakan `DistilBertTokenizer`. Pertama, teks harus dibagi menjadi bagian kecil, seperti kata atau frasa, yang kemudian akan digunakan sebagai token selama tahap analisis [25]. Selanjutnya, token ditambahkan ke ID numerik yang dibuat berdasarkan kosakata bawaan (vocabulary) DistilBERT. Tokenizer secara otomatis menambahkan token tertentu, seperti [CLS] di awal kalimat dan [SEP] sebagai pemisah antarsegmen, untuk menyesuaikan format input yang dibutuhkan oleh model. Selain itu, tokenizer membuat attention mask, yaitu array yang menandai token dengan informasi penting (diberikan nilai 1) dan token tambahan, seperti padding (diberikan nilai 0). Dengan fitur ini, model dapat berkonsentrasi pada elemen yang relevan selama proses pelatihan dan inferensi. Dalam kebanyakan kasus, hasil tokenisasi disusun dalam bentuk struktur lexicon yang mengandung input_ids dan attention_mask, yang dapat dimasukkan langsung ke dalam model DistilBERT sebagai input.

Tabel 2. Contoh hasil DistilBertTokenizer

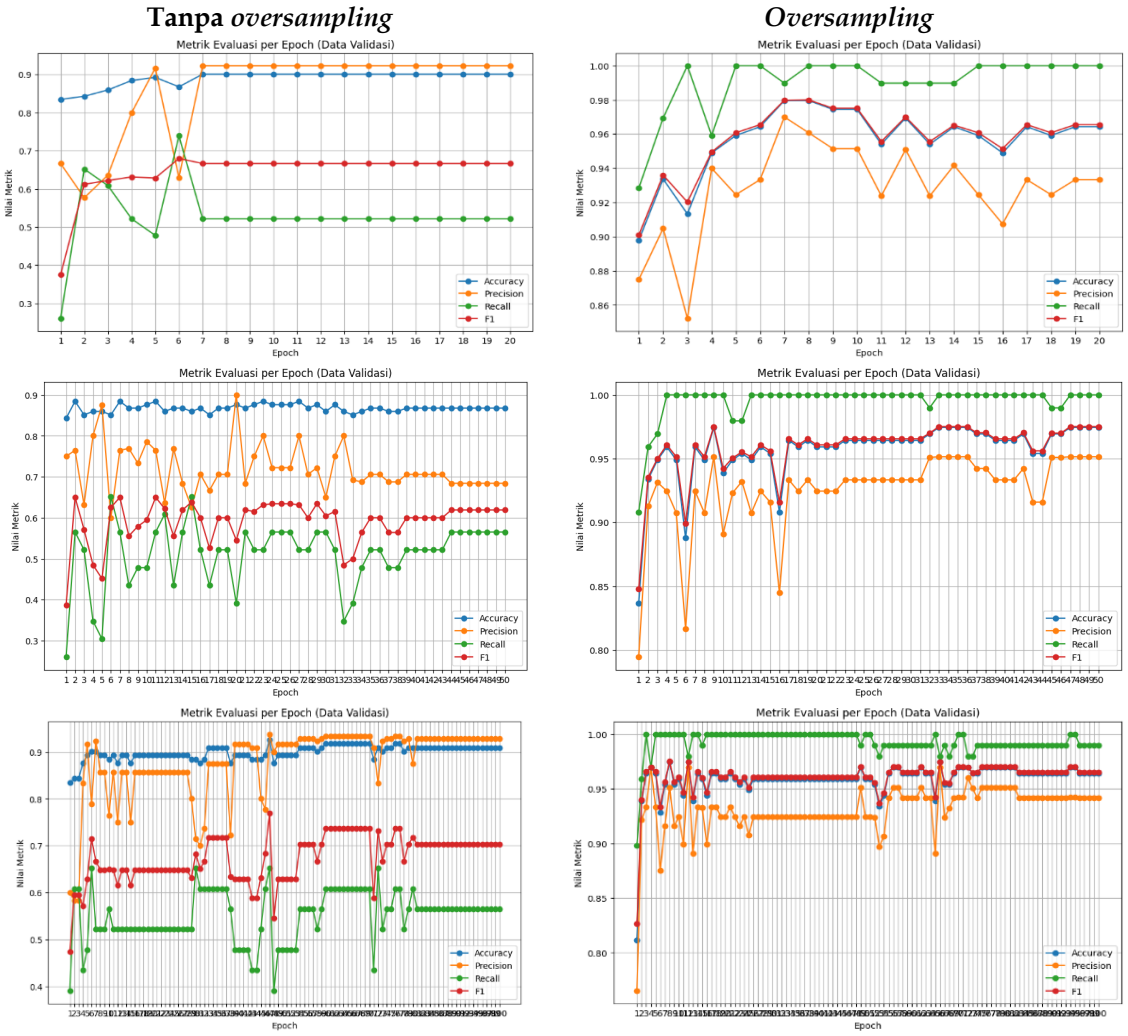
review	***input_ids***	***attention_mask***
Kalau minum ini bagusnya malam karena pasti diare, hari pertama diare parah banget sampe mau nangis. Hari kedua cuma buang air besar normal tapi sekitar 2-3x sehari, saya...	[101, 25444, 35166, 13484, 10465, 10592, 56332, 10251, 10676, 51697, 15786, 91688, 10671, 10246, 117, 18370, 14253, 10671, 10246, 10220, 10237, 17937, 10308, 21083, 11355, 43024, 55607, 10291, 119, 39769, 19059, 16008, 10113, 11499, 11889, 12566, 15096, 16626, 64747, 19825, 123, 118, 124, 10686, 10126, 44830, 117, 64981, 119, 119, 119, 102, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
pengirimannya cepat, cuma yang saya rasa setelah 2 minggu pemakaian ga ada perubahan malah jadi susah bab	[101, 66558, 28908, 45357, 10676, 65160, 117, 16008, 10113, 10265, 64981, 45772, 18044, 123, 74646, 11161, 23892, 57364, 11887, 15290, 55442, 73682, 17760, 10846, 12257, 15688, 10457, 102, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
beli buat suami, kata dia bikin lancar bab semoga bisa ngebantu kontrol berat badan dia ya	[101, 65817, 10116, 11499, 10526, 10603, 10500, 117, 21907, 10671, 11342, 12130, 12228, 15983, 15688, 10457, 11531, 25201, 17103, 10743, 101929, 40468, 55605, 60954, 51463, 10671, 10549, 102, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

3.3.2. Setup Model

Model `DistilBertForSequenceClassification` yang digunakan adalah varian khusus dari arsitektur DistilBERT yang tersedia dalam pustaka Hugging Face Transformers yang dirancang untuk memenuhi tugas klasifikasi teks. Dengan menggunakan objek `TrainingArguments`, proses pelatihan dikonfigurasi untuk mengatur berbagai parameter penting dalam proses pembelajaran melalui modul Trainer Hugging Face. Dalam penelitian ini, model dilatih selama 20, 50, dan 100 epoch, masing-masing dengan batch size 16 untuk proses pelatihan dan evaluasi. Untuk menjaga stabilitas selama pembelajaran, pembaruan bobot model dilakukan secara bertahap dengan rate pembelajaran 2e-5 (0.00002).

3.4. Implementasi dan Evaluasi

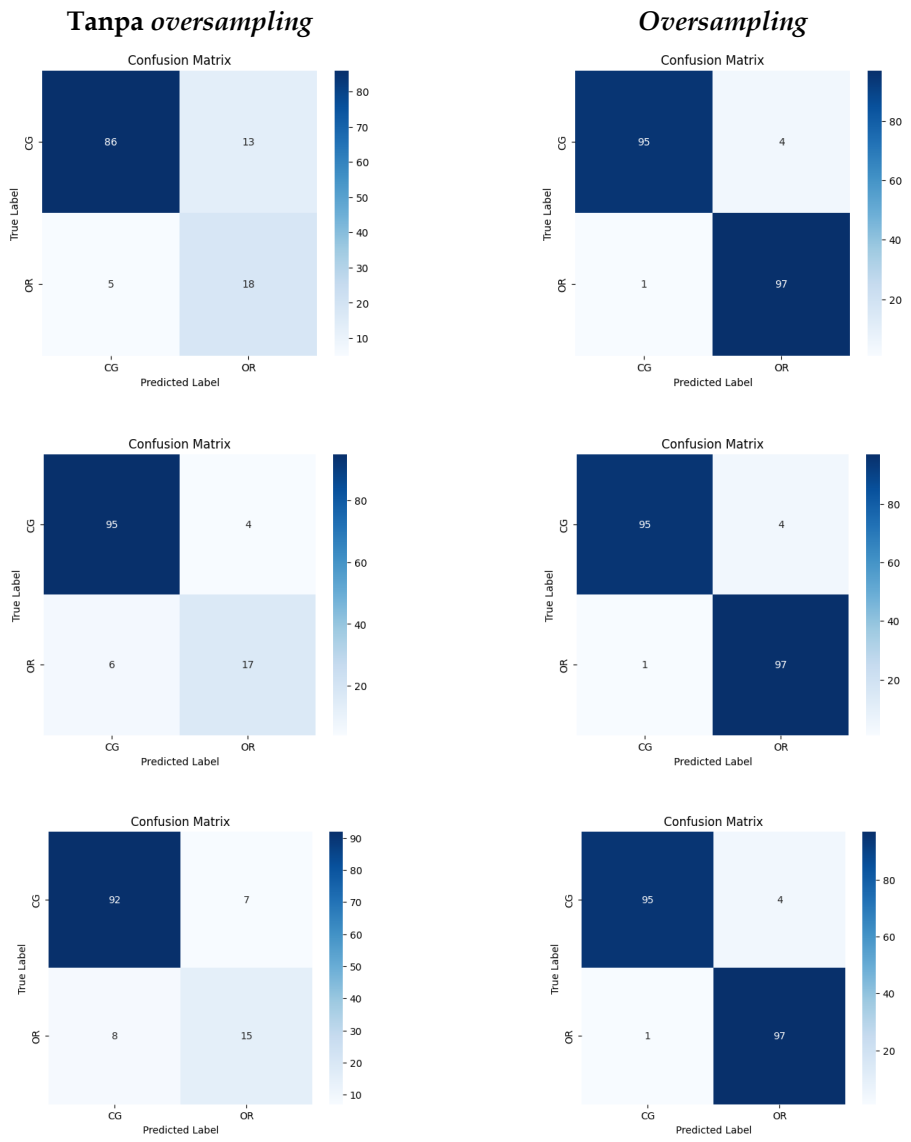
Implementasi model dilakukan pada berbagai scenario pelatihan dan pengujian tanpa dan dengan menggunakan oversampling selama 20, 50 dan 100 epoch. Berikut hasil evaluasi pada setiap skenario.



Gambar 3. Perbandingan hasil skenario

Hasil evaluasi terhadap enam skenario pelatihan (20, 50, dan 100 epoch) menunjukkan perbedaan yang signifikan antara model yang dilatih tanpa oversampling dan yang dilatih dengan oversampling. Tanpa oversampling, model memang mampu mencapai ketepatan dan ketepatan yang tinggi, yaitu di atas 0.85 hingga 0.90, terutama pada pelatihan 100 epoch. Namun, seluruh skenario menunjukkan recall yang rendah dan tidak stabil di sekitar 0.40 hingga 0.63, yang mengakibatkan f1-score rendah dan mengindikasikan model masih cenderung bias dan belum optimal.

Sedangkan, hasil pengujian dengan oversampling menunjukkan peningkatan performa yang signifikan dan konsisten sepanjang skenario. Model memiliki kemampuan yang sangat andal untuk mengidentifikasi ulasan OR (true positive), dengan f1-score yang stabil di atas 0.95 dan recall 1.00 atau mendekati. Meskipun pada epoch awal, precision sempat berubah-ubah, metrik secara keseluruhan menjadi lebih stabil seiring bertambahnya epoch.



Gambar 4. Perbandingan confusion matrix

Pada tiga skenario dengan oversampling, confusion matrix mencatat hanya 5 kesalahan dari total 197 data, dengan 95 prediksi benar dan 4 salah untuk kelas CG, serta 97 benar dan hanya 1 salah untuk kelas OR.

Tabel 3. Perbandingan metrik evaluasi

Skenario	***epoch***	***Optimizer***	***Batch size***	**Metrik evaluasi**				
				accuracy	***precision***	***recall***	***f1-score***	***roc-auc***
oversampling	20	*Adam*	16	0.97	0.99 (CG) 0.96 (OR)	0.96 (CG) 0.99 (OR)	0.97	0.9797
	50	*Adam*	16	0.97	0.99 (CG) 0.96 (OR)	0.96 (CG) 0.99 (OR)	0.97	0.9921
	100	*Adam*	16	0.97	0.99 (CG) 0.96 (OR)	0.96 (CG) 0.99 (OR)	0.97	0.9876
tanpa *oversampling*	20	*Adam*	16	0.85	0.95 (CG) 0.58 (OR)	0.87 (CG) 0.78 (OR)	0.91 (CG) 0.67 (OR)	0.8759
	50	*Adam*	16	0.92	0.94 (CG) 0.81 (OR)	0.96 (CG) 0.74 (OR)	0.95 (CG) 0.77 (OR)	0.9563
	100	*Adam*	16	0.88	0.92 (CG) 0.68 (OR)	0.93 (CG) 0.65 (OR)	0.92 (CG) 0.67 (OR)	0.8997

Hasil evaluasi skenario tanpa oversampling lebih rendah daripada hasil skenario dengan menggunakan oversampling, yang menunjukkan precision mencapai 0,99 (CG) dan 0,96 (OR),

sedangkan recall seimbang dengan 0,96 (CG) dan 0,99 (OR). Dengan skor f1-score sebesar 0,97 di kedua kelas, ini menunjukkan bahwa sensitivitas dan presisi seimbang. Dengan nilai akurasi total yang konsisten mencapai 97%.

Nilai ROC-AUC juga baik namun sangat berbeda dan semua menunjukkan kemampuan klasifikasi yang sangat baik. Paling tinggi, nilai ROC-AUC ditunjukkan pada skenario dengan menggunakan oversampling, mencapai 0,9797 pada 20 epoch, naik menjadi 0,9921 pada 50 epoch, dan sedikit turun menjadi 0,9876 pada 100 epoch.

3.5. *Implementasi Sistem*

Antarmuka sistem yang telah dirancang dan diprogram menggunakan Streamlit ditunjukkan pada beberapa Gambar dibawah ini.

Gambar 5. Implementasi Menu Beranda

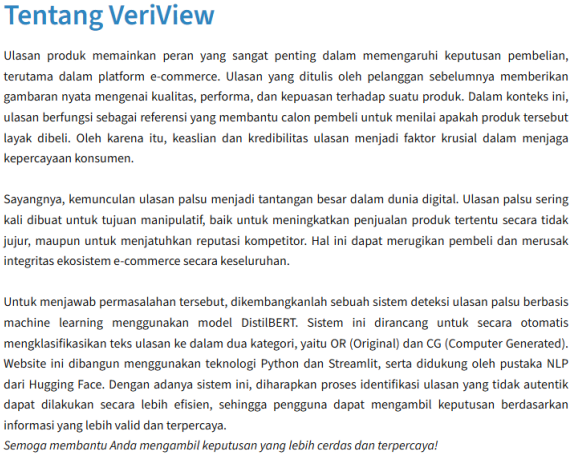
Halaman pengantar yang memberikan penjelasan singkat tentang tujuan pembuatan sistem. Pengguna akan menemukan penjelasan tentang cara website menggunakan model berbasis DistilBERT untuk mendeteksi ulasan palsu (CG) atau asli (OR) secara otomatis di bagian ini. Selain itu, menu Beranda menawarkan arahan untuk menuju menu Deteksi untuk memulai penggunaan fitur utama web.

Gambar 6. Implementasi Menu Deteksi

Pengguna dapat memasukkan teks ulasan yang ingin diperiksa pada gambar 4.8, dan web akan memeriksa apakah ulasan tersebut asli (OR) atau palsu (CG). Selain itu, Anda dapat memilih untuk memasukkan ulasan secara manual atau mengunggah file.csv.

Gambar 7. Implementasi Menu Riwayat

Menu yang merupakan bagian dari website yang berfungsi untuk menyimpan, menampilkan, dan mengelola catatan hasil prediksi atau aktivitas yang pernah dilakukan pengguna. Pada menu ini, setiap ulasan yang telah dianalisis oleh sistem akan tersimpan secara otomatis, lengkap dengan detail seperti teks ulasan, tanggal, waktu, serta hasil deteksi.



Gambar 8. Implementasi Menu Tentang

Menu yang berisi informasi latar belakang tentang sistem yang dibangun. Mencakup tujuan pengembangan sistem, manfaat penggunaannya untuk mendukung validitas ulasan online, dan teknologi utama yang digunakan seperti Streamlit, Python, dan Hugging Face Transformers libraries.

4. KESIMPULAN

Untuk model deteksi ulasan palsu menggunakan model DistilBERT, *oversampling* meningkatkan kinerja secara signifikan dibandingkan model tanpa *oversampling*. Hasil pengujian menunjukkan bahwa kelas ulasan palsu (CG) memiliki nilai *precision* sebesar 0,99 dan kelas ulasan asli (OR) memiliki nilai recall sebesar 0,96 dan nilai precision sebesar 0,99. Nilai *f1-score* kedua kelas sama-sama tinggi, 0,97, dan akurasi keseluruhan 97%. Meskipun demikian, pelatihan 50 *epoch* memiliki ROC-AUC tertinggi dengan nilai 0,9921, yang menunjukkan kemampuan klasifikasi model yang sangat baik. Akibatnya, oversampling telah terbukti efektif dalam meningkatkan keseimbangan data dan kinerja model deteksi.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada semua pihak yang secara langsung maupun tidak langsung telah memberikan kontribusi dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

[1] A. Lailatul Rachmawati, "Analisis Pengaruh E-commerce Terhadap Perilaku Konsumtif Mahasiswa (Studi Kasus Pada Mahasiswa Di Prodi Manajemen Universitas Tidar)," *J. Online Mhs. Manaj.*, vol. 1, no. 1, 2019, [Online]. Available: http://jom.untidar.ac.id/index.php/market/article/download/642/pdf

[2] H. Alamsyah, Y. Cahyana, and A. R. Pratama, "Deteksi Fake Review Menggunakan Metode Support Vector Machine dan Naïve Bayes Di Tokopedia," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 12, no. 2, pp. 585–598, 2023.

[3] R. Febriyanti, P. Ariwibowo, and D. Nurmalasari F, "Pengaruh E-Commerce Dan Digital Payment Terhadap Perilaku Konsumtif Siswa Di Sman 11 Depok," *Econ. Learn. Exp. Soc. Think. Educ. J.*, vol. 3, no. 2, pp. 123–130, 2024, doi: 10.58890/eleste.v3i2.173.

[4] R. S. Nailul Huda, Dyah Ayu, "Digital Economy Outlook 2025," *https://celios.co.id/*, 2024. https://celios.co.id/digital-economy-outlook-2025/

[5] A. Awalina, F. A. Bachtiar, and I. Indriati, "Klasifikasi Ulasan Palsu Menggunakan Borderline Over Sampling (BOS) dan Support Vector Machine (SVM) (Studi Kasus : Ulasan Tempat Makan)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 2, pp. 419–426, 2022, doi: 10.25126/jtiik.2022925692.

[6] A. Atthahahirah, "Pengaruh Ulasan Produk Terhadap Kepuasan Pengguna Aplikasi Female Daily," *Innov. J. Soc. Sci. Res.*, vol. 3, pp. 7135–7147, 2023.

[7] A. Awalina, F. A. Bachtiar, and F. Utaminingrum, "Perbandingan Pretrained Model Transformer pada Deteksi Ulasan Palsu," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 597–604, 2022, doi: 10.25126/jtiik.2022935696.

[8] Mentari Puspadini, "Jangan Tertipu, Begini Cara Mendeteksi Review Palsu di Internet," 2025. https://www.cnbcindonesia.com/lifestyle/20250412133257-33-625470/jangan-tertipu-begini-cara-mendeteksi-review-palsu-di-internet

[9] S. Nurohanisah, R. Astuti, and F. Muhammad Basysyar, "Deteksi Berita Palsu Menggunakan Algoritma Random Forest," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 422–428, 2024, doi: 10.36040/jati.v8i1.8418.

[10] E. C. Sukmawati, L. Suryaningrum, and D. Angelica, "Klasifikasi Berita Palsu Menggunakan Model Bidirectional Encoder Representations From Transformers (BERT)," vol. 6, no. 2, pp. 76–85, 2024.

[11] Hugging Face, "DistilBERT," 2019. https://huggingface.co/docs/transformers/main/en/model_doc/distilbert

[12] Mahira Putri, T. E. Sutanto, and S. Inna, "Studi Empiris Model BERT dan DistilBERT Analisis Sentimen pada Pemilihan Presiden Indonesia," *Indones. J. Comput. Sci.*, vol. 12, no. 5, 2023, doi: 10.33022/ijcs.v12i5.3445.

[13] Syafnidawaty, "Observasi," 2020. https://raharja.ac.id/2020/11/10/observasi/ (accessed Jun. 24, 2025).

[14] A. H. Fauziya, M. Isa, S. Manajemen, and F. Ekonomi, "PENGARUH FAKE BUYER DAN FAKE REVIEW TERHADAP PURCHASE INTENTION PRODUK FASHION PADA MARKETPLACE SHOPEE DENGAN TRUST SEBAGAI VARIABEL INTERVENING," vol. 17, no. 3, pp. 8–11, 2024.

[15] J. H. Yam, "Kajian Penelitian: Tinjauan Literatur Sebagai Metode Penelitian," *J. Emp.*, vol. 4, no. 1, pp. 61–70, 2024.

[16] N. A. Verdikha, T. B. Adji, and A. E. Permanasari, "Komparasi Metode Oversampling Untuk Klasifikasi Teks Ujaran Kebencian," *Semin. Nas. Teknol. Inf. dan Multimed. 2018*, pp. 85–90, 2018, [Online]. Available: https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/download/2062/1871

[17] L. Qadrini, H. Hikmah, and M. Megasari, "Oversampling, Undersampling, Smote SVM dan Random Forest pada Klasifikasi Penerima Bidikmisi Sejawa Timur Tahun 2017," *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 386–391, 2022, doi: 10.47065/josyc.v3i4.2154.

[18] I. G. Agung, R. Wira, and I. W. Santiyasa, "Pengaruh Penanganan Ketidakseimbangan Kelas pada Prediksi Cacat Perangkat Lunak dengan Teknik Oversampling," vol. 3, no. November,

pp. 143–152, 2024.
[19] M. B. Prayogi and G. Masitoh, "Analisis Sentimen Ulasan Pengguna Aplikasi Alfagift Menggunakan Random Forest," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 10, no. 2, pp. 158–170, 2025, doi: 10.14421/jiska.2025.10.2.158-170.
[20] cloudfactory, "Dataset Split in Machine Learning." https://wiki.cloudfactory.com/docs/mp-wiki/splits/data-splitting-in-machine-learning? (accessed Jun. 24, 2025).
[21] rage.pythai.net, "Fine-tuning Hyperparameters: exploring Epochs, Batch Size, and Learning Rate for Optimal Performance," 2025. https://rage.pythai.net/fine-tuning-hyperparameters-a-deep-dive-into-epochs-batch-size-and-learning-rate-for-optimal-performance/
[22] Gian Paolo Santopaolo, "Understanding Key Hyperparameters When Fine-Tuning an LLM," *genmind.ch*, 2025. https://genmind.ch/posts/understanding-key-hyperparameters-when-fine-tuning-an-llm/
[23] Drishti, "Introduction to DistilBERT in Student Model," 20222. https://www.analyticsvidhya.com/blog/2022/11/introduction-to-distilbert-in-student-model/ (accessed Jan. 30, 2025).
[24] B. Pranay Kumar, S. Ahmed, and M. Sadanandam, "DistilBERT: A Novel Approach to Detect Text Generated by Large Language Models (LLM)," no. Llm, 2024, [Online]. Available: https://doi.org/10.21203/rs.3.rs-3909387/v1
[25] A. Fergina, I. L. Kharisma, P. A. S, and A. I. Taek, "Analisis Sentimen Ulasan Pengguna Aplikasi Webtoon Menggunakan Text Mining Dan Algoritma," *J. Inf. dan Teknol.*, vol. 5, no. 2, pp. 164–170, 2023, doi: 10.60083/jidt.v5i2.355.