

# Analisis Sentimen Terhadap Pembelajaran Secara Daring Pasca Pandemi Covid-19 Menggunakan Metode IndoBERT

Muhammad Nur Hidayat<sup>1</sup>, Rully Pramudita<sup>2\*</sup>

<sup>1,2</sup> Teknik Informatika; Universitas Bina Insani; Jalan Siliwangi No.6 Rawa Panjang Kota Bekasi; 021-82436886/021-82436996; e-mail: [hidayatmuhammad344@gmail.com](mailto:hidayatmuhammad344@gmail.com), [rullypramudita@binainsani.ac.id](mailto:rullypramudita@binainsani.ac.id)

\* Korespondensi: e-mail: [rullypramudita@binainsani.ac.id](mailto:rullypramudita@binainsani.ac.id)

Diterima: 01 Januari 2024 ; Review: 13 Januari 2024; Disetujui: 15 Januari 2024

Cara sitasi: Hidayat M.N., Pramudita R. (2023). Analisis Sentimen Terhadap Pembelajaran Secara Daring Pasca Pandemi Covid-19 Menggunakan Metode IndoBERT. Information Management for Educators and Professionals. Vol 8 (2): 161-170

**Abstrak:** Pandemi Covid-19 yang sebelumnya terjadi pada tahun 2020-2022, membuat pemerintah Indonesia mengubah metode pembelajaran yang sebelumnya dilakukan secara tatap muka menjadi pembelajaran secara daring atau *online*. Perubahan tersebut memaksa para pelajar untuk mampu beradaptasi dengan sistem pembelajaran secara daring tersebut, hal itu memunculkan berbagai opini dari masyarakat. Pembelajaran secara daring juga dilakukan di Universitas Bina Insani dan pihak Universitas ingin mengetahui tanggapan mahasiswanya mengenai pembelajaran daring yang diberlakukan. Hal itu dilakukan dengan cara membagikan sebuah survei berupa google form kepada mahasiswa lalu hasil survei tersebut akan dikaji oleh staff Universtas sebagai masukan untuk perbaikan sistem pembelajaran kedepannya. Analisis sentimen dilakukan untuk mengetahui pendapat atau opini bernilai positif, negatif atau netral. Untuk mengklasifikasi sebuah kalimat diperlukan adanya metode yang bisa mengelompokan suatu sentimen. Metode IndoBERT merupakan metode yang digunakan pada penelitian ini untuk mendapatkan hasil sentimen dengan metode pengujian menggunakan *Confusion Matrix* yang nantinya akan menghitung nilai akurasi dari metode IndoBERT. Hasil pengujian yang dilakukan dalam penelitian ini menghasilkan nilai *accuracy* yang cukup tinggi yaitu sebesar 87% dengan nilai *precision* sebesar 87%, *recall* sebesar 91% dan *F1-score* sebesar 89%. Pengujian dilakukan dengan menguji sebanyak 100 kalimat dengan berbagai sentimen.

**Kata kunci:** Analisis Sentimen, *Confusion Matrix*, *Google Form*, IndoBERT, Kuliah Online

**Abstract:** The Covid-19 pandemic, which previously occurred in 2020-2022, made the Indonesian government change the learning method from face-to-face to online learning. This change forces students to be able to adapt to the online learning system, it raises various opinions from the community. Online learning is also carried out at Bina Insani University and the University wants to know the responses of its students regarding online learning. This is done by distributing a survey in the form of a google form to students and then the survey results will be reviewed by the University staff as input for future learning system improvements. Sentiment analysis is done to find out opinions or opinions are positive, negative or neutral. To classify a sentence, a method is needed that can classify a sentiment. The IndoBERT method is the method used in this research to get sentiment results with a testing method using *Confusion Matrix* which will calculate the accuracy value of the IndoBERT method. The test results conducted in this study resulted in a fairly high accuracy value of 87% with a precision value of 87%, recall of 91% and F1-score of 89%. Testing was done by testing 100 sentences with various sentiments.

**Keywords:** Sentiment Analysis, *Confusion Matrix*, *Google Form*, IndoBERT, Online Lectures

## 1. Pendahuluan

Pada tahun 2020 – 2022, virus corona atau dikenal dengan Covid-19 menginfeksi jutaan orang di dunia sampai Organisasi Kesehatan Dunia (WHO) menyatakan penyebaran virus Covid-19 sebagai pandemi global. Mengingat penyebaran virus Covid-19 yang cepat dan berskala besar, mengakibatkan Indonesia mengalami keadaan darurat nasional yang berdampak pada seluruh sektor kehidupan, tidak terkecuali pendidikan. Dalam dunia pendidikan, pemerintah Indonesia mencetuskan kebijakan sistem pembelajaran daring yang mewajibkan kegiatan pembelajaran dilaksanakan secara daring (*online*) dari rumah dengan tujuan meminimalkan penyebaran virus dan memastikan bahwa Kegiatan Belajar Mengajar tetap terlaksana meski di kondisi pandemi [1].

Metode pembelajaran daring tidak terlepas dari pemanfaatan teknologi seperti jaringan internet untuk menunjang proses pembelajaran secara daring. Perubahan tersebut memaksa para pelajar untuk mampu beradaptasi dengan sistem pembelajaran daring, hal itu memunculkan berbagai opini dari masyarakat terutama kalangan pelajar, baik yang positif, netral bahkan negatif. Saat ini, Universitas Bina Insani menerapkan metode pembelajaran secara *hybrid*, yaitu metode pembelajaran yang menggabungkan antara pembelajaran secara daring dengan Pembelajaran Tatap Muka (PTM). Untuk mengetahui opini terkait kebijakan pembelajaran secara daring, umumnya staff BAA (Biro Administrasi Akademik) akan membagikan sebuah survei berupa *google form* yang akan diisi oleh mahasiswa. Hasil survei tersebut akan dibaca dan dikaji oleh staff sebagai masukan untuk perbaikan sistem pembelajaran kedepannya. Cara tersebut dirasa kurang tepat karena proses mendapatkan sebuah sentimen membutuhkan waktu yang cukup lama dan kurang efisien. Penelitian ini bertujuan untuk menciptakan sebuah platform website yang dapat mengklasifikasikan data berupa opini menjadi sebuah sentimen positif, netral dan negatif secara otomatis. Diharapkan website ini dapat mempermudah dalam proses mendapatkan sentimen dan bisa menjadi sistem penunjang keputusan yang tepat.

Analisis sentimen adalah sebuah proses mendeteksi, mengekstraksi, dan mengklasifikasikan pendapat tentang sesuatu hal. Analisis sentimen melibatkan opini yang dibuat oleh pengguna internet untuk dikumpulkan dan diperiksa. Analisis sentimen berbentuk sebuah teks dan bisa dikategorikan sebagai sentimen positif, negatif atau netral [2]. Analisis sentimen bertujuan menganalisis pendapat seseorang yang berkenaan dengan suatu topik [3].

Machine learning (pembelajaran mesin) adalah bagian dari kecerdasan buatan yang membuat sistem memiliki kemampuan belajar secara otomatis dan meningkatkan kemampuannya berdasarkan pengalaman tanpa di program. Data yang diolah juga tidak hanya teks, tetapi dapat juga berupa gambar, video ataupun aktivitas-aktivitas pengguna selama berselancar atau mengakses internet [4]. Berdasarkan teknik pembelajarannya, machine learning terbagi menjadi *Supervised Learning* dan *Unsupervised Learning*. *Supervised Learning* adalah sebuah teknik dimana aplikasi akan diberikan sebuah label sebagai pembelajaran terhadap masukan dan keluaran yang akan diberikan. *Unsupervised Learning* adalah sebuah teknik dimana aplikasi tidak diberikan label sebagai pembelajaran dan membiarkan algoritma menentukan struktur dan pola tersembunyi pada data yang tidak memiliki label [5].

*Python* adalah bahasa pemrograman yang banyak digunakan dalam aplikasi *web*, pengembangan perangkat lunak, ilmu data, dan *machine learning*. Hal ini disebabkan oleh kesederhanaan sintaksisnya, fleksibilitasnya, serta berbagai pustaka dan modul yang kuat yang dapat digunakan oleh para pengembang [6].

*Google Colab* adalah platform pemrograman berbasis *cloud* yang menggunakan bahasa pemrograman *python*. *Google Colab* memberikan kemudahan kepada pengguna dalam visualisasi data karena tidak perlu melakukan instalasi perangkat lunak pada komputer anda [7]. *Google Colab* menyediakan *processor* dengan spesifikasi tinggi (GPU dan TPU) gratis kepada penggunanya dengan tujuan memudahkan para *researcher* menjalankan program yang membutuhkan spesifikasi tinggi secara *online* [8].

*Natural Language Processing* (NLP) adalah salah satu bidang ilmu komputer, kecerdasan buatan dan bahasa (linguistik) yang mempelajari mengenai bagaimana komputer dapat berinteraksi dengan bahasa alami manusia [9].

BERT adalah sebuah algoritma *deep learning* keluaran Google yang merupakan invasi dari model *Transformer* dimana model tersebut memproses sebuah kata pada kalimat dengan mempelajari konteks berdasarkan kata-kata yang ada [10]. Untuk analisis sentimen menggunakan BERT, itu bisa dilakukan secara maksimal jika datasetnya bahasa Inggris. Untuk

dataset berbahasa Indonesia digunakan modifikasi dari model BERT Base yaitu Model IndoBERT yang sudah dilatih dengan dataset bahasa Indonesia yang lebih besar dan beragam [11].

*Text pre-processing* adalah teknik *data mining* yang melibatkan transformasi data mentah menjadi format yang mudah dimengerti. Hal itu dilakukan untuk membuat data menjadi lebih bersih sehingga proses klasifikasi dapat memberikan hasil yang lebih optimal [12]. Dalam penelitian ini, *Text pre-processing* yang dilakukan adalah *data filtering & cleaning, case folding* dan *tokenization*.

*ReactJS* adalah pustaka *javascript* yang digunakan untuk mengembangkan komponen antarmuka pengguna (UI). *React* juga termasuk pustaka *javaScript* yang paling mudah untuk dipelajari karena menawarkan kemudahan dan efektivitas dalam pengalaman pengembangan aplikasi [13].

*Confusion Matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep *data mining* [14]. *Confusion matrix* terdiri dari informasi yang membandingkan hasil klasifikasi dengan mengukur nilai kebenaran. Terdapat empat istilah sebagai representasi hasil klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP) merupakan nilai positif yang dideteksi dengan benar, *True Negative* (TN) merupakan nilai negatif yang dideteksi dengan benar, *False Positive* (FP) merupakan data negatif namun terdeteksi secara positif dan *False Negative* (FN) merupakan data positif namun terdeteksi secara negatif [15].

*Website* adalah kumpulan semua halaman *web* yang fungsinya untuk menampilkan berbagai informasi dalam bentuk tulisan, gambar dan suara dari sebuah *domain* yang terbentuk dalam suatu rangkaian yang saling terkait [16].

*Wordcloud* adalah gambar yang menunjukkan daftar kata-kata yang digunakan dalam sebuah teks dalam bentuk abstraksi visual yang intuitif. Umumnya semakin banyak kata yang digunakan dalam sebuah teks, maka akan semakin besar bentuk representasi kata tersebut dalam gambar [17].

## 2. Metode Penelitian

Pada penelitian ini, konsep garis besar dari tahapan untuk membuat sebuah *website* klasifikasi opini ditampilkan dengan sebuah langkah-langkah penelitian. Berikut adalah langkah-langkahnya:

### **Pengumpulan data**

Pada tahap ini dilakukan pengumpulan data dengan cara membuat survei berupa *google form* dan survei tersebut akan diisi oleh mahasiswa. Data hasil survei tersebut akan menjadi input untuk proses klasifikasi opini.

### **Planning**

Selanjutnya adalah melakukan *planning* atau perencanaan untuk menganalisis masalah yaitu menentukan cara terbaik untuk klasifikasi opini serta untuk mempresentasikan hasil klasifikasi dalam suatu *website*.

### **Design**

Selanjutnya adalah merancang model IndoBERT untuk sistem klasifikasi opini. Selain itu, dilakukan perancangan GUI untuk *website* sebagai representasi hasil dari klarifikasi opini yang telah dilakukan.

### **Coding**

Selanjutnya, dilakukan tahap *coding* atau pemrograman yang terbagi atas 2 bagian yaitu pembuatan model dengan bahasa pemrograman *python* serta pembuatan tampilan GUI dengan bahasa pemrograman *ReactJS*.

### **Testing**

Langkah terakhir yaitu melakukan testing atau uji coba model dan *website* untuk mengecek apakah model dan aplikasi *website* berjalan dengan baik.

### **Release**

Jika hasil dari uji coba sudah sesuai keinginan, maka dapat dilakukan *release* berdasarkan kinerja yang dihasilkan dan dinyatakan berhasil setelah diuji.

### 3. Hasil dan Pembahasan

Hasil dan pembahasan terdiri dari analisis kebutuhan sistem, tahapan *machine learning*, perancangan dan implementasi sistem serta pengujian sistem yang dirancang. Adapun pembahasan mengenai beberapa sub-bab, sebagai berikut:

#### Analisis kebutuhan sistem

Analisis kebutuhan sistem terdiri dari dua kebutuhan yaitu kebutuhan fungsional dan kebutuhan non-fungsional. Kebutuhan fungsional adalah informasi mengenai pengembangan yang dilakukan untuk dapat memberikan solusi dari permasalahan yang terjadi, sedangkan kebutuhan non-fungsional adalah informasi mengenai spesifikasi perangkat yang digunakan dalam penelitian hingga implementasi. Adapun uraian mengenai analisis kebutuhan sistem, sebagai berikut:

- a. Analisis kebutuhan fungsional  
Berikut adalah kebutuhan fungsional sistem dari pada penelitian ini adalah sebagai berikut:
  - 1) Sistem dapat mengunggah data berupa opini berbentuk file CSV atau *excel*.
  - 2) Sistem dapat menampilkan hasil sentimen berdasarkan data input yang dikumpulkan sebelumnya melalui survei mahasiswa.
- b. Analisis kebutuhan non-fungsional  
Berikut adalah kebutuhan non-fungsional sistem dari pada penelitian ini adalah sebagai berikut:
  - 1) Tampilan aplikasi *web* yang mudah digunakan agar pengguna tidak kesulitan saat mengakses sistem tersebut.
  - 2) Sistem dapat dijalankan di beberapa *web browser* diantaranya *Internet Explorer*, *Google Chrome* dan *Mozilla Firefox*.

#### Tahapan *Machine Learning*

Tahap ini merupakan pembahasan rangkaian proses untuk mendapatkan hasil sentimen yang terbagi atas 3 kelas yaitu positif, netral dan negatif. Klasifikasi opini dilakukan dengan menggunakan metode IndoBERT. Adapun tahapan-tahapan *machine learning* yang dilakukan sebagai berikut:

- a. Pengumpulan dataset  
Data yang digunakan dalam penelitian ini merupakan dataset hasil survei yang dilakukan kepada mahasiswa. Pada penelitian ini, proses pengumpulan opini dilakukan menggunakan membagikan *google form* ke mahasiswa dengan topik mengenai kuliah online. Data yang didapatkan sebanyak 35 opini yang diambil dari tanggal 03 September 2023 sampai 10 September 2023.
- b. *Text pre-processing*  
Selanjutnya adalah tahapan dimana data yang didapat dari survei akan ditransformasi menjadi data dengan format yang dapat dipahami oleh sistem. *Text pre-processing* dilakukan menjadi beberapa tahapan yaitu:

#### *Data filtering & cleaning*

Pada tahap ini, data yang sudah dikumpulkan akan di-identifikasi dan dibersihkan dari atribut kata yang tidak diperlukan serta kata duplikat. Tahap ini berguna untuk mengurangi redundansi data serta kesalahan pada data yang nanti akan diklasifikasikan. Berikut adalah tabel untuk memperlihatkan perbedaan sebelum dan sesudah diubah.

Tabel 1. Hasil *Data Filtering & Cleaning*

Opini	<i>Filtering &amp; Cleaning</i>
Selama pandemi covid-19 kegiatan kuliah terasa berbeda dari yang biasanya.	Selama pandemi covid kegiatan kuliah terasa berbeda dari yang biasanya

Sumber : Hasil Penelitian (2023)

#### *Case folding*

Selanjutnya, pada tahap ini semua data akan diubah hurufnya menjadi huruf kecil (*lowercase*). Berikut adalah tabel untuk memperlihatkan perbedaan sebelum dan sesudah diubah.

Tabel 2. Hasil *Case Folding*

<i>Filtering &amp; Cleaning</i>	<i>Case Folding</i>
Selama pandemi covid kegiatan kuliah terasa berbeda dari yang biasanya	selama pandemi covid kegiatan kuliah terasa berbeda dari yang biasanya

Sumber : Hasil Penelitian (2023)

**Tokenization**

Selanjutnya, pada tahap ini semua data dilakukan proses pemecahan pada kalimat menjadi satuan kata. Berikut adalah tabel untuk memperlihatkan perbedaan sebelum dan sesudah diubah.

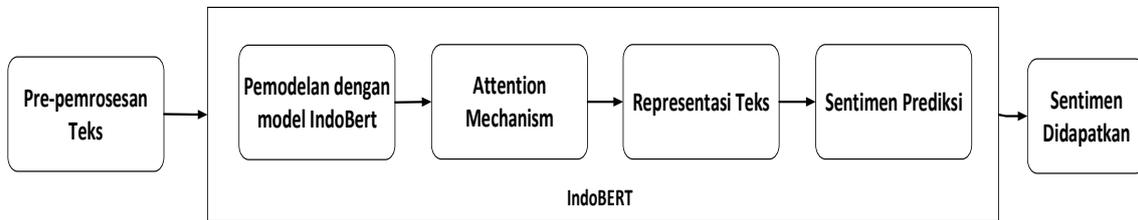
Tabel 3. Hasil *Tokenization*

<i>Case Folding</i>	<i>Tokenization</i>
selama pandemi covid kegiatan kuliah terasa berbeda dari yang biasanya	"selama", "pandemi", "covid", "kegiatan", "kuliah", "terasa", "berbeda", "dari", "yang", "biasanya"

Sumber : Hasil Penelitian (2023)

c. Implementasi model IndoBERT

Arsitektur IndoBERT yang digunakan untuk proses klasifikasi opini untuk menghasilkan sentimen, dapat digambarkan pada gambar 1 berikut :



Sumber : Hasil Penelitian (2023)

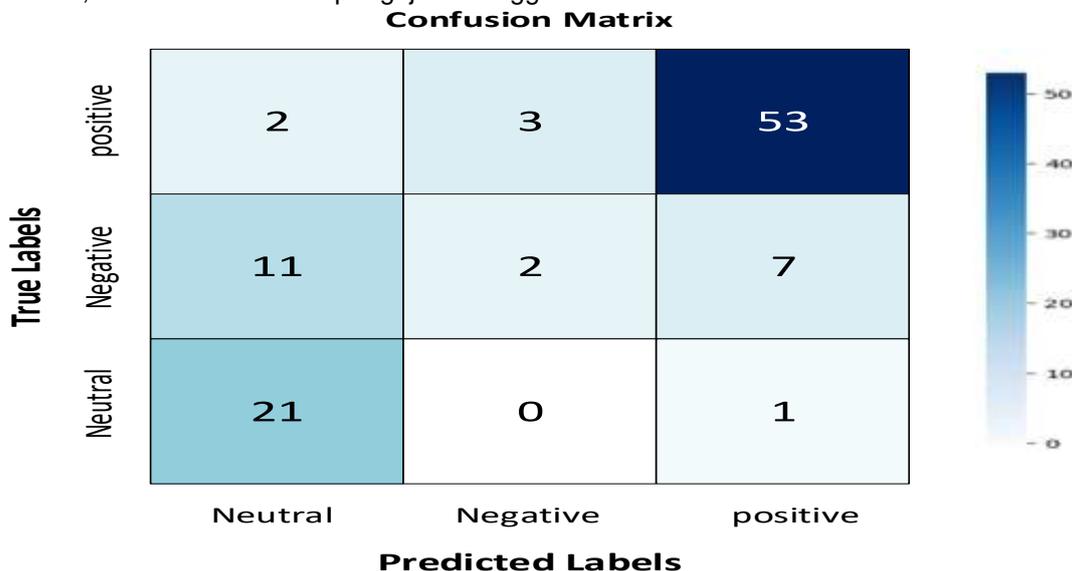
Gambar 1. Model IndoBERT

Berikut adalah penjelasan gambar 1 :

1. **Pre-pemrosesan teks**  
 Data berisi opini yang telah dilakukan *text pre-processing* akan digunakan sebagai input untuk pemrosesan pembuatan sentimen. Kalimat tersebut akan memasuki proses tokenisasi terlebih dahulu, lalu akan diberikan token-token khusus sebagai penanda untuk *hidden layer* saat melakukan klasifikasi teks, seperti token [CLS] untuk menandai awal kalimat dan token [SEP] untuk menandai akhir kalimat. Selanjutnya kalimat itu akan melalui proses *embedding* yaitu proses konversi kata menjadi bentuk vektor numerik agar sistem memahami kata tersebut dalam bentuk vektor numerik.
2. **Pemodelan dengan IndoBERT**  
 Setelah representasi vektor kata-kata diperoleh, kalimat akan dimasukkan ke dalam model IndoBERT. Model akan melewati serangkaian lapisan neuron yang berguna bagi sistem untuk mempelajari data dan mengekstraksi fitur-fitur penting dari teks.
3. **Mekanisme Attention**  
 Saat melewati setiap lapisan neuron, ada mekanisme *attention* yang memungkinkan model untuk menekankan kata-kata yang paling bermakna dan relevan dengan konteks kalimat. Ini membantu model memahami hubungan antara kata-kata dan makna keseluruhan kalimat.
4. **Representasi teks**  
 Setelah melewati lapisan-lapisan tersebut, kalimat akan direpresentasikan dalam bentuk vektor yang menggambarkan pemahaman model tentang teks tersebut. Vektor ini mencakup informasi tentang makna dan hubungan antar kata-kata dalam kalimat.
5. **Sentimen prediksi**  
 Setelah itu, vektor representasi kalimat tersebut akan dikirim ke lapisan keluaran (*output layer*). Hasil dari lapisan keluaran ini adalah label sentimen yang menunjukkan apakah kalimat tersebut memiliki sentimen positif, negatif, atau netral. Label sentimen yang dihasilkan (misalnya: positif) adalah hasil dari keputusan model berdasarkan analisis fitur-fitur dalam kalimat.

## d. Pengujian model

Pengujian ini dilakukan untuk mengetahui tingkat akurasi dalam melakukan klasifikasi opini. Pengujian model dilakukan menggunakan *confusion matrix* untuk mengukur performa model yang digunakan. Pengujian model pada penelitian ini digunakan 100 kalimat dengan berbagai sentimen, berikut adalah hasil pengujian menggunakan confusion matrix:



Sumber : Hasil Penelitian (2023)

Gambar 2. Confusion Matrix

Berdasarkan gambar 2, dapat disimpulkan bahwa terdapat nilai TP (*True Positive*) sebanyak 53 data, TN (*True Negative*) sebanyak 34 data, FP (*False Positive*) sebanyak 8 data dan FN (*False Negative*) sebanyak 5 data. Dari nilai-nilai tersebut, dapat digunakan untuk perhitungan nilai *accuracy*, *precision*, *recall* dan *F1-score*. Berikut adalah perhitungan nilai-nilai yang dibutuhkan untuk pengujian model.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{53+34}{53+34+8+5} = 0,87$$

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{53}{53+8} = 0,87$$

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{53}{53+5} = 0,91$$

$$\text{F1-Score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} = 2 \times \frac{(0,87 \times 0,91)}{(0,87+0,91)} = 0,89$$

Sumber : Hasil Penelitian (2023)

Gambar 3. Hasil Perhitungan Accuracy, Precision, Recall dan F1-Score

Berdasarkan perhitungan di atas, didapatkan nilai akurasi yang cukup tinggi yaitu 0,87 atau 87%, nilai *precision* sebesar 0,87 atau 87%, nilai *recall* sebesar 0,91 atau 91% dan nilai *F1-Score* sebesar 0,89 atau 89%. Dari hasil perhitungan tersebut menunjukkan bahwa model ini dapat mengklasifikasikan sentimen dengan baik dengan hasil akurasi yang cukup tinggi.

e. Perancangan dan implementasi sistem

Perancangan sistem adalah gambaran rancangan *website* yang akan dibuat untuk diimplementasikan. Sementara, Implementasi sistem adalah pemaparan mengenai tampilan dan fungsi dari setiap halaman untuk memperjelas bentuk dari implementasi antarmuka. Adapun beberapa gambaran perancangan dan implementasi sistem dari *website* ini:

**Tampilan Halaman Home**

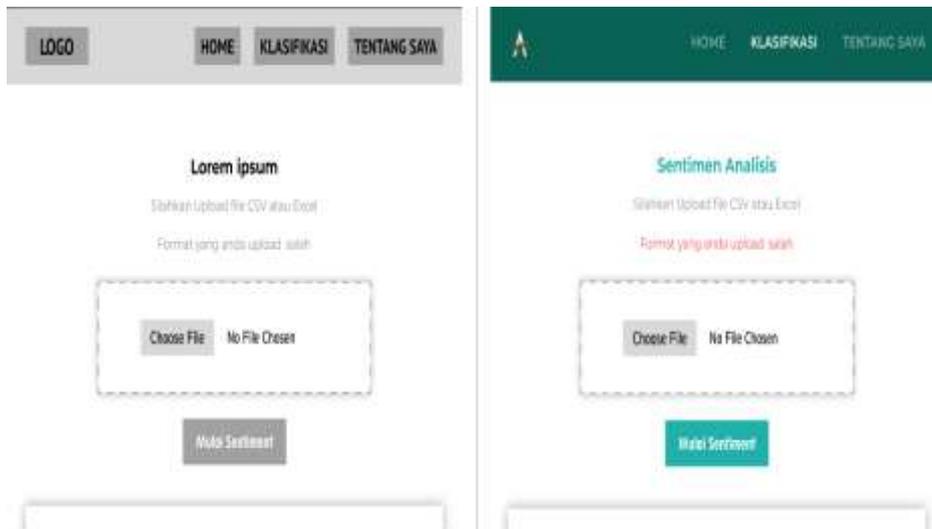
Halaman *Home* muncul di awal saat pengguna mengakses *website* ini. Halaman ini berisi deskripsi mengenai *website* ini. Berikut adalah rancangan dan implementasi dari halaman *home*.



Gambar 4. Halaman Home

**Tampilan Halaman Klasifikasi**

Halaman ini merupakan halaman yang berfungsi untuk melakukan klasifikasi dan mendapatkan sentimen berdasarkan input berupa file CSV atau *excel* berisi kumpulan opini.



Gambar 5. Halaman Klasifikasi

**Tampilan Halaman Hasil Klasifikasi**

Halaman ini merupakan halaman yang menampilkan hasil klasifikasi yang dihasilkan oleh sistem. Hasil klasifikasi tersebut berupa tabel berisi sentimen, persentase dari banyaknya data



No.	Pertanyaan yang diajukan	Jawaban yang diberikan
2.	Apakah sistem klasifikasi opini ini cukup memudahkan anda untuk menentukan sentimen?	Sistem ini cukup memudahkan untuk menentukan sentimen dengan cepat dan tepat.
3.	Apakah fitur yang terdapat pada sistem klasifikasi opini ini sudah sesuai dengan kebutuhan?	Sistem yang dibuat cukup mudah dimengerti dan sudah cukup memiliki fitur yang dibutuhkan.
4.	Menurut anda, apakah sistem ini memiliki tingkat akurasi yg tepat berdasarkan opini yang diberikan (jika dibandingkan dengan asumsi anda)?	Tingkat akurasi jika dibandingkan antara hasil prediksi dengan hasil yang sebenarnya sudah cukup tepat, walaupun terkadang masih ada yang salah dikarenakan adanya opini yang menggunakan singkatan kata yang tidak umum.
5.	Apakah sistem klasifikasi ini dapat memberikan manfaat?	Sistem ini bisa memberikan manfaat karena bisa menjadi sistem penunjang keputusan berdasarkan opini atau pendapat yang diberikan.

Sumber : Hasil Penelitian (2023)

#### 4. Kesimpulan

Berdasarkan rumusan masalah, dapat disimpulkan dari penelitian yang telah dilakukan adalah dari 35 data, didapatkan sentimen positif sebanyak 10 data, sentimen negatif sebanyak 23 data dan sentimen netral sebanyak 2 data. Dan sistem klasifikasi opini menggunakan metode IndoBERT dapat membantu untuk mengklasifikasikan opini secara akurat dan membaginya menjadi 3 kategori (positif, netral dan negatif) secara otomatis dengan nilai akurasi sebesar 0,87 atau 87%. Penelitian selanjutnya diharapkan dapat dikembangkan kembali lain waktu atau oleh peneliti selanjutnya. Seperti, dapat dilakukan penambahan fitur dan visualisasi data yang masih berkaitan dengan analisis sentimen sebagai sarana informasi yang lebih sederhana dan lebih mudah dipahami, dapat dilakukan upaya peningkatan kualitas dataset, seperti penambahan data dan melakukan *text pre-processing* dengan metode yang lebih baik untuk memastikan data yang digunakan berkualitas baik, dapat dilakukan penelitian menggunakan model IndoBERT yang lain atau metode klasifikasi lainnya untuk mengetahui metode terbaik dalam melakukan analisis sentimen.

#### Referensi

- [1] C. Ketut Mediana Ayu, S. I Made Agus Dwi, and W. Ketut Agung Cahyawan, "Analisis Sentimen Pembelajaran Daring Era Pandemi COVID-19 Menggunakan Naive Bayes Dan Support Vector Machine," vol. 5, pp. 1–6, 2023.
- [2] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 113, 2018, doi: 10.26418/jp.v4i2.27526.
- [3] F. F. Mailoa, "Analisis sentimen data twitter menggunakan metode text mining tentang masalah obesitas di indonesia," *J. Inf. Syst. Public Heal.*, vol. 6, no. 1, p. 44, 2021, doi: 10.22146/jisph.44455.
- [4] P. D. Kusuma, *Machine Learning Teori, Program, Dan Studi Kasus*. Deepublish, 2020.
- [5] et al. . A. Aditya Permana, Wahyuddin S, L. Willyanto Santoso, *Machine Learning*. Global Eksekutif Teknologi, 2023.
- [6] T. Suhesti, "Bahasa Pemrograman Python," *J. Ilmuti Ilmu Teknol. Inf.*, vol. 2, no. 11, pp. 1–6, 2022.
- [7] R. G. Guntara, "Visualisasi Data Laporan Penjualan Toko Online Melalui Pendekatan Data Science Menggunakan Google Colab," vol. 2, no. 6, pp. 2091–2100, 2023.
- [8] O. Nurdiawan, R. Herdiana, I. Ali, and M. Fijriani, "Kinerja Algoritma Convolutional Neural Network dalam Klasifikasi Covid-19 Varian Omicron Berdasarkan Citra Ct-Scan Thoax," vol. 9, no. 5, pp. 1472–1478, 2022, doi: 10.30865/jurikom.v9i5.4884.
- [9] V. R. Prasetyo, N. Benarkah, and V. J. Chrisintha, "Implementasi Natural Language Processing Dalam Pembuatan Chatbot Pada Program Information Technology Universitas Surabaya," *Teknika*, vol. 10, no. 2, pp. 114–121, 2021, doi: 10.34148/teknika.v10i2.370.
- [10] N. Adine, S. Casi, and D. Burhanuddin, "Deteksi Hate Speech Pada Twitter," *e-Proceeding Eng.*, vol. 10, no. 1, pp. 256–262, 2023.
- [11] H. K. Putra, M. Arif Bijaksana, and A. Romadhony, "Deteksi Penggunaan Kalimat

- Abusive Pada Teks Bahasa Indonesia Menggunakan Metode IndoBERT,” *e-Proceeding Eng.*, vol. 8, no. 2, pp. 3028–3038, 2021.
- [12] F. Sidik, I. Suhada, A. H. Anwar, and F. N. Hasan, “Analisis Sentimen Terhadap Pembelajaran Daring Dengan Algoritma Naive Bayes Classifier,” *J. Linguist. Komputasional*, vol. 5, no. 1, p. 34, 2022, doi: 10.26418/jlk.v5i1.79.
- [13] S. K. Murti and A. Sujarwo, “Membangun Antarmuka Pengguna Menggunakan ReactJs untuk Modul Manajemen Pengguna,” pp. 1–6, 2018.
- [14] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, “Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix,” *J. Inform. Upgris*, vol. 6, no. 2, pp. 66–75, 2021, doi: 10.26877/jiu.v6i2.6552.
- [15] L. Mutawalli, M. T. A. Zaen, and W. Bagye, “KLASIFIKASI TEKS SOSIAL MEDIA TWITTER MENGGUNAKAN SUPPORT VECTOR MACHINE (Studi Kasus Penusukan Wiranto),” *J. Inform. dan Rekayasa Elektron.*, vol. 2, no. 2, p. 43, 2019, doi: 10.36595/jire.v2i2.117.
- [16] T. A. Kinaswara, N. R. Hidayati, and F. Nugrahanti, “Rancang Bangun Aplikasi Inventaris Berbasis Website Pada Kelurahan Bantengan | Kinaswara | Prosiding Seminar Nasional Teknologi Informasi dan Komunikasi (SENATIK),” *Pros. Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2, no. 1, pp. 71–75, 2019, [Online]. Available: <http://prosiding.unipma.ac.id/index.php/SENATIK/article/view/1073>.
- [17] T. M. Fahrudin, A. R. F. Sari, A. Lisanthoni, and A. A. D. Lestari, “Analisis Speech-To-Text Pada Video Mengandung Kata Kasar Dan Ujaran Kebencian Dalam Ceramah Agama Islam Menggunakan Interpretasi Audiens Dan Visualisasi Word Cloud,” *Skatika*, vol. 5, no. 2, pp. 190–202, 2022, doi: 10.36080/skatika.v5i2.2942.