

# Optimalisasi Evaluasi Pelaksanaan Pelatihan Melalui Analisis Sentimen Otomatis Dengan Model Text Classification

Evi Septiana Pane<sup>a,1,\*</sup>, Cynthia Caroline<sup>b,2</sup>

<sup>a</sup>Balai Diklat Industri Surabaya, Jl. Gayung Kebonsari Dalam 12, Surabaya 60235

<sup>b</sup>Home Credit Indonesia, Plaza Oleos Lt. 8 Jalan TB Simatupang, Jakarta Selatan 12520

<sup>1</sup>evi-septiana@kemenperin.go.id\*; <sup>2</sup>pocaroline34@gmail.com

ARTICLE INFO

ABSTRACT / ABSTRAK

## Article history

Received:

Revised:

Accepted:

Lembaga pelatihan sering menghadapi tantangan dalam mengumpulkan dan menganalisis umpan balik peserta secara akurat, karena metode konvensional seperti kuesioner sering kali tidak mencerminkan pengalaman peserta yang sebenarnya. Peningkatan volume data dari platform evaluasi online memperumit analisis manual, dan komentar teks peserta sering diabaikan meskipun tersimpan secara digital. Oleh karena itu, inovasi teknologi diperlukan untuk memahami pengalaman peserta dengan lebih baik. Salah satu solusi inovatif adalah penerapan analisis sentimen menggunakan model klasifikasi teks berbasis machine learning. Proses analisis ini dimulai dengan pembersihan dan pemrosesan awal data teks untuk mendapatkan data yang bersih dan siap diekstraksi fiturnya. Mengingat distribusi data yang tidak seimbang antara kelas sentimen, metode BERT digunakan untuk menyeimbangkan data tersebut, sementara model klasifikasi Naïve Bayes digunakan untuk memprediksi sentimen dari data teks peserta. Dari penelitian ini diperoleh hasil prediksi sentimen komentar peserta dengan akurasi terbaik sebesar 95,7% dalam skenario pembagian data uji dan latih 70:30, yang diterapkan pada data hasil augmentasi. Ini menunjukkan bahwa model klasifikasi sentimen berbasis machine learning dapat diandalkan untuk membantu lembaga pelatihan memahami sentimen peserta secara otomatis. Selain itu, visualisasi word cloud dari hasil penelitian ini memberikan gambaran yang lebih jelas bagi penyelenggara pelatihan untuk menindaklanjuti evaluasi yang dilakukan.

*Training institutions often struggle with accurately collecting and analyzing participant feedback, as traditional methods like questionnaires frequently fail to capture the true experiences of participants. The growing volume of data from online evaluation platforms further complicates the manual analysis, and participants' text comments are often overlooked despite being stored digitally. This underscores the need for technological innovation to gain a deeper understanding of participant experiences. One promising solution is the application of sentiment analysis through a machine learning-based text classification model. This process starts with cleaning and preprocessing the text data to ensure it is ready for feature extraction. To address the issue of imbalanced data distribution between sentiment classes, the BERT method is employed to balance the data, while the Naïve Bayes classification model is used to predict sentiment from the participant's text. According to the experiment results, the sentiment prediction on trainee comment data text achieved the highest accuracy of 95.7% using a 70:30 split between training and test data, applied to the augmented dataset. This demonstrates that machine learning-based sentiment classification models are a reliable tool for training institutions to automatically gauge participant sentiment. Additionally, word cloud visualizations from the study offer training organizers a clearer understanding, enabling more informed follow-up actions based on the evaluations.*

This is an open-access article under the [CC-BY-SA](#) license.



**Kata Kunci:** analisis sentimen, evaluasi pelatihan, klasifikasi teks

**Keywords:** sentiment analysis, training evaluation, text classification

## 1. Pendahuluan

Hingga tahun 2024, sebanyak 5.494 lembaga pelatihan kerja telah terverifikasi oleh Pusat Pembinaan Kementerian Ketenagakerjaan. Sejumlah lembaga pelatihan kerja tersebut dapat menyelenggarakan pelatihan dengan kapasitas mencapai 4.620.252 peserta/tahun (*Beranda : Kementerian Ketenagakerjaan RI, n.d.*). Jumlah tersebut belum termasuk lembaga pelatihan untuk ASN yang telah di akreditasi oleh Lembaga Administrasi Negara (*Jumlah Lembaga Pelatihan Terakreditasi Menurut Program Diklat Dan Jenis Instansi - Portal Satu Data Indonesia, n.d.*) dan lembaga pendidikan dan pelatihan non-formal pada Kementerian Pendidikan Nasional serta lembaga pelatihan SDM yang tersebar di berbagai Kementerian/Lembaga/Instansi. Dengan semakin banyaknya jumlah lembaga pelatihan, maka pemerintah perlu rutin melakukan pengawasan mutu (*surveillance*) secara berkala baik

melalui mekanisme akreditasi ataupun verifikasi. Oleh sebab itu, lembaga pelatihan dituntut untuk terus melakukan perbaikan kualitas dari berbagai aspek penyelenggaraan pelatihan.

Terdapat berbagai upaya yang dapat dilakukan oleh sebuah lembaga pelatihan untuk melakukan perbaikan berkelanjutan dalam penyelenggaraan pelatihan. Salah satu diantaranya adalah dengan melakukan evaluasi penyelenggaraan pelatihan secara berkala setelah pelatihan dilaksanakan. Evaluasi penyelenggaraan pelatihan penting dikarenakan beberapa alasan. Evaluasi membantu mengetahui tingkat kepuasan peserta, menilai efektivitas program dalam mencapai tujuan organisasi, meningkatkan kualitas pelatihan melalui pengukuran pemahaman peserta, menilai perubahan perilaku kerja peserta, mengukur dampak pelatihan terhadap produktivitas organisasi, meningkatkan akuntabilitas dalam pengelolaan program, meningkatkan manfaat pelatihan, serta membantu pengembangan program pelatihan yang lebih efektif (Widoyoko, E. P, 2017). Semua ini memastikan pelatihan berkualitas, efektif, dan transparan.

Dalam pelaksanaan evaluasi, lembaga pelatihan dapat menggunakan metode evaluasi yang dianggap sesuai. Model evaluasi pelatihan yang paling sering digunakan adalah Kirkpatrick's training evaluation model (Widoyoko, E. P, 2017). Model Kirkpatrick's seringkali digunakan dalam evaluasi program pelatihan karena menawarkan pendekatan yang komprehensif dengan empat tingkatan evaluasi: reaksi, pembelajaran, perilaku, dan hasil. Untuk mengukur reaksi kepuasan peserta terhadap program pelatihan metode yang paling sering digunakan adalah menggunakan angket reaksi. Metode ini menggunakan angket (kuesioner) yang berisi pertanyaan-pertanyaan untuk mengetahui tingkat kepuasan peserta terhadap aspek edukatif, fasilitas, dan pelayanan terhadap peserta pelatihan. Metode angket seringkali digunakan karena termasuk yang paling mudah digunakan, dapat memperoleh hasil reaksi responden secara cepat dan juga dapat menjangkau seluruh kalangan peserta secara luas, terutama jika dilakukan secara online. Namun demikian, metode angket juga memiliki beberapa kekurangan, diantaranya adalah keterbatasan respon peserta dari jawaban akan pertanyaan tertutup yang seringkali muncul pada angket evaluasi peserta pelatihan. Untuk mengatasi hal tersebut, disediakan pula satu atau lebih pertanyaan terbuka dimana responden dapat memberikan komentar/saran masukan atas penyelenggaraan pelatihan yang telah diikuti. Hanya saja hasil jawaban dari pertanyaan terbuka tersebut sangat beragam antar responden yang membuat pengelola evaluasi kesulitan dalam membuat ringkasan atau menyajikan informasi dari data teks komentar/saran masukan responden.

**FORM EVALUASI KINERJA  
PENYELENGGARA DIKLAT ELEKTRO  
TAHUN 2019**

9. Pelayanan panitia terhadap Peserta dan instruktur \*

1 2 3 4 5

Sangat Kurang Memuaskan      Sangat Memuaskan

10. Komentar/Saran \*

Your answer

11. Manfaat Mengikuti Diklat \*

Your answer

Back Submit Clear form

**Gambar 1** Contoh tangkapan layar angket GForm evaluasi penyelenggaraan pelatihan

Dari metode angket (gForm) evaluasi penyelenggara pada gambar 1 diatas, dapat terlihat bahwa penilaian dengan pertanyaan tertutup masih menjadi rujukan untuk melakukan evaluasi dan penyajian hasil evaluasi. Sementara hasil pertanyaan terbuka, yakni komentar/saran peserta serta manfaat pelatihan bagi peserta yang berupa inputan teks lebih sering diabaikan. Dikarenakan untuk pengolahan data teks tersebut masih mengandalkan operator evaluasi yang mengambil secara subjektif dari data komentar peserta tersebut. Dengan demikian, hasil evaluasi menjadi kurang valid. Oleh sebab itu, penelitian ini bertujuan mengoptimalkan pengolahan dan penyajian data evaluasi khususnya untuk data yang berbentuk teks. Metode yang diusulkan yakni analisis sentimen dari komentar peserta secara otomatis dengan klasifikasi teks dengan model *learning*, serta penyajian ringkasan data yang lebih mudah dipahami dalam bentuk *wordcloud*.

### 1.1. Rumusan Masalah

Lembaga penyelenggara pelatihan saat ini perlu secara aktif mengumpulkan umpan balik relevan dari peserta. Akan tetapi, metode konvensional melalui kuesioner penilaian evaluasi oleh peserta dirasa kurang cukup untuk memberikan gambaran akurat tentang pengalaman peserta selama proses pelatihan. Oleh karena itu, diperlukan terobosan berbasis teknologi untuk menggali perasaan dan pengalaman peserta, sehingga lembaga pelatihan dapat memastikan program mereka tetap relevan dan efektif dengan mengintegrasikan teknologi yang tepat. Hal ini akan membantu dalam perbaikan berkelanjutan program pelatihan dan membuat peserta merasa didengar serta dipahami dalam pengembangan keterampilan mereka.

Dalam beberapa tahun terakhir, terjadi ledakan volume data yang belum pernah terjadi sebelumnya, terutama didorong oleh menjamurnya platform online, media sosial, dan saluran komunikasi digital. Peserta pelatihan terlibat dalam berbagai cara, berbagi pengalaman, pendapat, dan masukan melalui berbagai media, termasuk teks, gambar, dan video (Wankhade et al., 2022). Lonjakan data ini menimbulkan tantangan yang signifikan bagi penyelenggara, karena banyaknya masukan dari peserta menjadi terlalu besar untuk dianalisis secara manual dan efektif.

Data evaluasi penyelenggaraan pelatihan selama ini telah tersimpan dalam bentuk digital. Namun demikian, data kualitatif teks komentar peserta belum pernah diolah atau dianalisis lebih lanjut, karena fokus pengolahan dan analisis lebih pada aspek kuantitatif. Dari umpan balik peserta, hanya sejumlah indikator kuantitatif tertentu yang menjadi fokus evaluasi, sedangkan data kualitatif, seperti teks komentar peserta, sering diabaikan. Padahal pertanyaan terbuka seperti teks komentar peserta sering kali efektif untuk mendapatkan umpan balik yang autentik dan mendalam. Respons pertanyaan terbuka, di mana responden memberikan jawaban dalam format teks bebas tanpa kategori yang ditentukan sebelumnya dapat mengungkap keragaman pendapat serta nuansa dalam tanggapan, dan membantu menangkap alasan di balik data survei kualitatif (Rouder et al., 2021). Namun demikian, tantangan dalam analisis dan pelaporan data kualitatif dapat diatasi dengan teknik visualisasi yang sederhana. Pengkodean data dan visualisasi yang efektif dapat memperkaya narasi data kuantitatif dan memberikan konteks yang lebih dalam menggali informasi untuk analisis yang bermakna. Salah satu metode pengolahan dan penggalian data teks kualitatif yang terkini adalah analisis sentimen. Penerapan analisis sentimen membantu pengolahan data kualitatif dengan cara mengidentifikasi dan menganalisis opini atau pendapat yang terkandung dalam data tekstual. Selain itu, dalam analisis sentimen terdapat metode *word cloud*. *Word cloud* digunakan untuk mengidentifikasi kata-kata yang paling sering muncul dalam teks (Sinclair & Cardew-Hall, 2008), yang dapat membantu dalam mengklasifikasi sentimen positif, negatif, atau netral. Sehingga, *word cloud* dapat membantu visualisasi dalam memahami tema dan sentimen yang dominan dalam data tekstual.

Dengan menggunakan pemrosesan bahasa alami (Natural Language Processing / NLP) dan pembelajaran mesin (*machine learning*) untuk melakukan filter dan mengekstraksi informasi berharga tentang sentimen, preferensi, dan emosi para peserta (Wankhade et al., 2022). Analisis sentimen memungkinkan penyelenggara untuk merespons tren dan masalah secara real-time, dan dapat dijadikan solusi untuk mengekstrak makna dari data komentar peserta secara otomatis, memberikan kontribusi signifikan dalam evaluasi pelatihan. Pendekatan ini tidak hanya mengatasi subjektivitas dalam penilaian manual tetapi juga meningkatkan efisiensi pengolahan data besar-besaran dan memastikan perhatian setara untuk aspek kualitatif dan kuantitatif dalam evaluasi.

Akibat dari jumlah data teks yang cukup besar serta belum tergalinya aspek kualitatif dari komentar pada evaluasi dari peserta, maka rumusan permasalahan yang diangkat pada penelitian ini yaitu 'bagaimana melakukan analisis sentimen dari umpan balik evaluasi peserta pelatihan dengan model machine learning untuk klasifikasi teks?'

## 1.2. Tujuan

Penelitian ini merupakan salah satu terobosan inovatif dari lembaga pelatihan yang bertujuan meningkatkan pelayanan dan kepuasan peserta melalui analisis sentimen dengan metode pembelajaran mesin. Tujuan dari penelitian ini adalah mengembangkan dan membangun model sentiment analysis atas feedback peserta pelatihan terhadap kinerja penyelenggara. Berdasarkan tujuan tersebut, terdapat dua kontribusi utama dari penelitian ini, yakni:

1. Sebuah model prediksi sentiment analysis dengan metode machine learning dari data umpan balik peserta pelatihan yang diungkapkan dalam kuesioner (online) kinerja penyelenggara pelatihan.
2. Membuat word cloud dari kumpulan komentar feedback peserta pelatihan. Word Cloud merupakan representasi visual dari data kata, di mana kata-kata digambarkan dalam ukuran berbeda berdasarkan frekuensi kemunculannya dalam teks tertentu. Semakin sering sebuah kata muncul dalam teks, semakin besar dan tebal kata tersebut muncul

## 1.3. Manfaat

Analisis sentimen terhadap umpan balik peserta evaluasi pelatihan berdampak signifikan, memberikan wawasan kepuasan peserta terhadap penyelenggara, materi, dan program (Lundqvist et al., 2020). Penggunaan teknologi ini diharapkan memberikan wawasan lebih mendalam terkait preferensi peserta dan aspek-aspek yang mempengaruhi kepuasan mereka. Penelitian ini bukan hanya alat evaluasi, tetapi juga sarana peningkatan kualitas pelatihan melalui pendekatan canggih, membantu pimpinan membaca dan menginterpretasi umpan balik peserta. Bagi penyelenggara, penelitian ini mengevaluasi kepuasan peserta, sarana prasarana, kualitas layanan, dan program pelatihan secara keseluruhan, membantu memahami dampak komunikasi dan pengorganisasian terhadap motivasi belajar peserta dengan dampak yang diharapkan yang lebih baik. Secara detail, berikut beberapa dampak yang diharapkan dari penelitian ini:

1. Mengidentifikasi kekuatan dan kelemahan  
Umpan balik peserta pelatihan, sumber informasi kunci bagi penyelenggara, dapat meningkatkan kualitas pembelajaran. Umpan balik positif menyoroti keberhasilan pelaksanaan program, termasuk efektivitas panitia, ketersediaan bahan praktek, dan pelayanan. Sebaliknya, umpan balik negatif mencerminkan area perbaikan, seperti masalah ketersediaan air, kebersihan ruangan, ketidakpuasan terhadap pengajar. Evaluasi sentimen peserta, sebagaimana diungkapkan (Dalipi et al., 2021), menjadi informasi penting untuk memastikan penyelenggara dapat merespons perbaikan yang diperlukan untuk memajukan kualitas keseluruhan pelatihan.
2. Meningkatkan kepuasan peserta pelatihan  
Analisis sentimen dari umpan balik peserta pelatihan membantu identifikasi area perbaikan dan tindakan perbaikan untuk meningkatkan pengalaman peserta (Chahar, 2023). Sentimen negatif menjadi panduan konstruktif untuk perbaikan masa mendatang, meningkatkan kualitas dan kepuasan peserta.
3. Pengambilan keputusan berdasarkan data (data-driven policy)  
Analisis sentimen penting bagi pengambil kebijakan dalam mengambil keputusan penting, melakukan alokasi sumber daya secara efektif, dan perbaikan berbasis prioritas. Penelitian ini berdampak pada peningkatan kualitas pengambilan kebijakan di lembaga pelatihan dengan memanfaatkan ulasan peserta untuk perbaikan berkelanjutan.

Pada akhirnya, analisis sentimen dapat membantu dalam mengevaluasi efektivitas penyelenggara dan pada akhirnya dapat meningkatkan kredibilitas dari lembaga pelatihan.

## 1.4. Studi Literatur

Analisis sentimen adalah salah satu bidang penelitian yang berkembang paling cepat dalam ilmu komputer. Pertumbuhan ini didorong oleh meningkatnya jumlah data, terutama teks, yang tersedia di internet. Secara signifikan, penelitian di bidang ini berkembang pesat setelah tahun 2004, dengan 99% makalah diterbitkan setelah tahun tersebut. Awalnya, topik penelitian analisis sentimen berfokus pada ulasan produk di e-commerce. Penelitian ini menjadi dasar bagi perkembangan analisis sentimen yang kemudian meluas hingga mencakup subjektivitas teks dan bidang terkait lainnya. Baru-baru ini, topik analisis sentimen telah bergeser ke berbagai teks dan opini di media sosial, mencakup bidang seperti pasar saham, pendidikan, dan kesehatan (Mäntylä et al., 2018).

Analisis sentimen, yang juga disebut sebagai penggalian opini (*opinion mining*), adalah proses menganalisis teks untuk menentukan nada emosional atau sentimen yang diungkapkan di dalamnya. Hal ini melibatkan identifikasi apakah teks tersebut menyampaikan sentimen positif, negatif, atau netral. Analisis sentimen

menggunakan teknik pemrosesan bahasa alami NLP dan pembelajaran mesin ML untuk mengekstrak informasi subjektif dari teks, seperti ulasan pelanggan, posting pada media sosial, dan bentuk bahasa tertulis atau lisan lainnya. Tujuannya adalah untuk memahami bagaimana perasaan orang tentang suatu produk, layanan, atau ide, yang dapat membantu organisasi meningkatkan produk, layanan, dan pengalaman pelanggan mereka.

Ruang lingkup penelitian analisis sentimen di Indonesia dapat dikategorikan kedalam tiga area, yaitu: ulasan terhadap penggunaan aplikasi (Herjanto & Carudin, 2024; Umair & Susanto, 2024; Wahyudi & Kusumawardhana, 2021), evaluasi dosen (Amrustian et al., 2022; Rusli, 2020; Santoso et al., 2017), ulasan terhadap layanan jasa dan produk (Aulia Rahman et al., 2023; Gian, 2024; Negara et al., 2020). Dengan kata kunci pencarian ‘analisis sentimen ulasan pelatihan’ pada Google Scholar, ditemukan satu artikel analisis sentimen data kritik dan saran pelatihan aplikasi teknologi informasi yang dilakukan oleh (Kabir et al., n.d.). Rangkuman dari beberapa penelitian analisis sentimen yang telah terbit dalam beberapa tahun terakhir dapat dilihat pada Tabel 1.

**Tabel 1.** Rangkuman penelitian analisis sentimen di Indonesia dalam beberapa tahun terakhir

Pengarang	Objek Analisis Sentimen	Kelas Sentimen	Data berimbang (Ya/Tidak); Melakukan Penyeimbangan Kelas (Ya/Tidak)	Metode Ekstraksi Fitur	Metode Klasifikasi	Word Cloud (Ada/Tidak Ada)
(Santoso et al., 2017)	Evaluasi Penilaian Dosen	Positif, Negatif, Netral	Tidak; Tidak	TF-IDF	SVM	Tidak ada
(Rusli, 2020)	Komentar Evaluasi Dosen	Positif, Negatif	Ya; Tidak	Word2vec	SVM	Tidak ada
(Negara et al., 2020)	Ulasan pengguna maskapai	Positif, Negatif	Ya; Tidak	TF-IDF, Information Gain	Naïve Bayes	Tidak ada
(Wahyudi & Kusumawardhana, 2021)	Ulasan pengguna aplikasi Grab	Positif, Negatif	Tidak; Tidak	N-Gram	SVM	Tidak ada
(Amrustian et al., 2022)	Evaluasi Pengajaran Dosen	Positif, Negatif	Tidak; Ya	Word Vector	LSTM	Tidak ada
(Aulia Rahman et al., 2023)	Ulasan pengguna Jasa Logistik J&T	Positif, Negatif	Tidak; Tidak	TF-IDF	Multinomial Naïve Bayes, SVM	Ada
(Herjanto & Carudin, 2024)	Ulasan pengguna aplikasi SIREKAP	Positif, Negatif, Netral	Tidak; Tidak	TF-IDF	Random Forest	Ada
(Umair & Susanto, 2024)	Ulasan pengguna aplikasi BRImo	Positif, Negatif, Netral	Tidak; Tidak	Tokenizing, Stemming	Naïve Bayes	Tidak ada
(Gian, 2024)	Ulasan pengguna produk Airism Uniqlo	Positif, Negatif	Tidak; Tidak	Word Count	-	Tidak ada
(Kabir et al., n.d.)	Ulasan Kritik dan Saran Pelatihan	Positif, Negatif, Netral	Ya; Tidak	TF-IDF	SVM	Tidak ada

Dari tabel 1 diatas, dapat dilihat bahwa sebagian besar penelitian analisis sentimen dilakukan terhadap ulasan dan evaluasi terhadap produk/jasa layanan/aplikasi, hanya satu penelitian yang ditemukan membahas ulasan kritik dan saran dari pelaksanaan pelatihan (Kabir et al., n.d.). Dari segi kelas sentimen, sebagian besar penelitian bertujuan untuk melakukan klasifikasi pada kelas sentimen positif dan negatif. Hal ini bisa berbeda di setiap penelitian, bergantung pada ketersediaan data yang dapat dikumpulkan. Jika dilihat lebih lanjut pada distribusi jumlah data antar kelas sentimen, sebagian besar penelitian memiliki distribusi kelas data yang tidak berimbang. Namun demikian, hanya satu penelitian yang dilakukan oleh (Amrustian et al., 2022), yang menerapkan metode

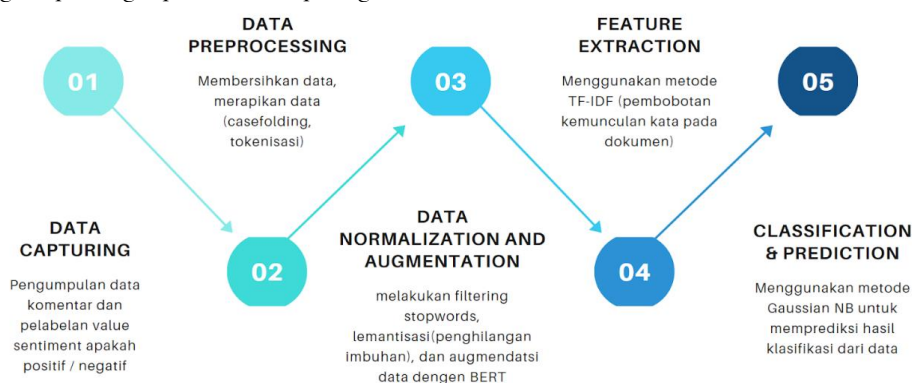
penyeimbangan data agar distribusi data antar kelas sentimen lebih merata. Metode penyeimbangan data yang digunakan yakni *resample*. Proses *resampling* menghasilkan distribusi sampel yang unik berdasarkan data aktual.

Lebih lanjut dari tabel 1, dilihat dari segi metode ekstraksi fitur dari data teks, yang banyak digunakan di sebagian besar penelitian adalah metode TF-IDF (*Term Frequency Inverse Document Frequency of records*). Sementara untuk metode klasifikasi/prediksi sentimen, dua metode klasifikasi yang paling sering digunakan adalah SVM (*Support Vector Machine*) dan Naïve Bayes. Untuk mendukung hasil analisis sentimen, *word cloud* kerap kali digunakan sebagai visualisasi pelengkap dari data teks yang dilakukan penggalian, pada sebagian besar penelitian di tabel 1, sayangnya tidak menampilkan hasil *word cloud* dari masing-masing kelas sentimen untuk memperkaya hasil analisis sentimen.

Berdasarkan beberapa penelitian analisis sentimen yang telah dirangkum pada tabel 1 diatas, penelitian analisis sentimen yang dilakukan disini berbeda dalam hal metode penyeimbangan distribusi data antar kelas sentimen dan ketersediaan *word cloud* sebagai visualisasi lanjutan untuk melengkapi pemahaman pengguna akhir. Berbeda dengan teknik *resampling* dalam menyeimbangkan distribusi data, pada penelitian ini dilakukan augmentasi data atau penambahan data menggunakan metode BERT (*Bidirectional Encoder Representations from Transformers*). BERT merupakan sebuah model pengolahan bahasa alami (NLP) berbasis teknik *deep learning* yang dikembangkan oleh Google di tahun 2018. Metode BERT digunakan dalam data augmentation untuk meningkatkan ketahanan dan performa model pengolahan bahasa alami (NLP).

## 2. Metodologi

Pendekatan yang diusulkan untuk mencapai tujuan dari penelitian ini menggunakan tahapan dasar analisis sentimen pada data teks. Metodologi dari penelitian ini mengikuti alur sentimen analisis yang meliputi beberapa langkah penting seperti terlihat pada gambar 2 berikut.



**Gambar 2** Metodologi tahapan melakukan sentiment analysis pada penelitian ini

### 2.1. Sumber Data

Sebagai sumber data utama dari penelitian ini adalah data umpan balik dari peserta pelatihan yang diisi melalui form evaluasi pelatihan. Terdapat tiga (3) jenis evaluasi yang dilakukan pada setiap penyelenggaraan pelatihan, yaitu: evaluasi peserta terhadap instruktur (tenaga pengajar), evaluasi peserta terhadap kinerja penyelenggara, dan evaluasi instruktur terhadap capaian pembelajaran peserta.

Secara khusus, penelitian ini dimulai dari sentimen analisis terhadap hasil evaluasi peserta terhadap kinerja penyelenggara pelatihan yang dilaksanakan di Balai Diklat Industri Surabaya sepanjang tahun 2018 dan 2019. Balai Diklat Industri Surabaya memiliki dua skema pelaksanaan pelatihan, yakni pelatihan yang berlokasi di workshop dan asrama Balai Diklat (in-site) dan pelatihan yang berlokasi di perusahaan mitra industri (on-site). Untuk menghindari bias ketidakseragaman lokasi pelatihan pada skema diklat on-site, maka pada penelitian ini dipergunakan terlebih dahulu evaluasi kinerja penyelenggara pelatihan yang diadakan dengan skema in-site. Di tahun 2019, pelaksanaan skema diklat in-site yang digunakan sebagai sample data pada penelitian analisis sentimen kali ini adalah data komentar dari evaluasi peserta yang berasal dari 12 angkatan pelatihan (masing-masing angkatan terdiri atas maksimal 50 peserta).

### 2.2. Pengumpulan dan pelabelan data

Tahapan pengumpulan data dilakukan melalui penyebaran kuesioner online dengan tools Google Form. Panitia penyelenggara membagikan tautan survei kuesioner kepada peserta di akhir masa pelatihan. Peserta mengisi evaluasi terhadap kinerja penyelenggaraan pelatihan yang sesuai seperti pada tautan form berikut:

<https://bit.ly/EvaluasiPenyelenggara2019> . Pada form tersebut terdapat sembilan (9) pertanyaan pilihan ganda dengan opsi jawaban skala likert (mulai dari sangat kurang memuaskan hingga sangat memuaskan) dan dua (2) pertanyaan terbuka yang menanyakan tentang ‘Komentar/Saran’ dan ‘Manfaat mengikuti diklat’ untuk perbaikan penyelenggaraan pelatihan di masa mendatang.

Isian data evaluasi peserta yang terekam pada Google Sheet kemudian akan diolah secara manual oleh panitia pelatihan. Saat ini, yang dilakukan oleh panitia terhadap data evaluasi adalah sebatas menyajikan hasil pengolahan data kuantitatif dari jawaban pertanyaan pilihan ganda. Sementara hasil jawaban pertanyaan terbuka, disampaikan begitu saja tanpa ada pengolahan atau analisis lebih lanjut. Padahal, seringkali jawaban dari pertanyaan terbuka memberikan insight yang lebih tepat untuk perbaikan dari penyelenggaraan pelatihan dari perspektif peserta.

Untuk penelitian sentimen analisis ini, data yang digunakan berupa data text hasil dari pertanyaan terbuka, khususnya pada kolom ‘Komentar/Saran’. Sebanyak 587 baris data text ‘Komentar/Saran’ berhasil dikumpulkan untuk diproses lebih lanjut. Dari kumpulan data tersebut diberi label kategori sentimen secara manual berdasarkan polaritas sentimen {positif dan negatif}. Polaritas sentimen ditentukan berdasar keberadaan kata-kata yang mewakili dari setiap label. Tabel I menunjukkan beberapa contoh hasil data ‘Komentar’ beserta label kategori sentimen yang diberikan secara manual.

	komen	kategori
	mohon segera memperbaiki tempat-	yang rusak negatif
Bimbingan yang bapak berikan	sangatlah bagus d...	positif
	Air minum sering telat diisi	negatif
	sangat puas	positif
	Luar biasa	positif
	...	...
	Trimah kasih atas bimbingannya di Surabaya	positif
Penguji harus bisa membakar semangat peserta		positif
Trainer.nya menerangkan sangat teliti dan saya...		positif
	kamar sebagian tidak ada kuncinya	negatif
	pelayanan dan fasilitas cukup baik	positif

**Gambar 3** Contoh tangkapan layar data mentah hasil pengumpulan dari GForm Kuesioner Evaluasi

### 2.3. Pra-pemrosesan data

Data teks komentar peserta pelatihan merupakan jenis data yang tidak terstruktur. Untuk mengekstrak informasi berguna dari data teks yang tidak terstruktur, perlu dilakukan beberapa tahap langkah pra-pemrosesan dengan tujuan untuk menghilangkan kesalahan ejaan, kesalahan tata bahasa, singkatan, dll terhadap teks asli. Selama tahap pra-pemrosesan, aktivitas tersebut diatas dilakukan dengan bantuan package Python NLTK (Bird et al., 2009). Aktivitas detail yang dilakukan dalam tahap pra-pemrosesan adalah sebagai berikut:

- Menghilangkan tanda baca dan angka atau karakter khusus lainnya. Hal ini perlu dihapus karena karakter tersebut bukan merupakan informasi berguna terkait dengan analisis sentimen.
- Melakukan tokenisasi. Tokenisasi adalah proses pemisahan (pemotongan) teks kalimat ke daftar kata.
- Konversi teks. Setelah tokenisasi, kata-kata dikonversi ke dalam huruf kecil.
- Menghilangkan stopwords. Stopwords pada dasarnya merupakan sekumpulan kata-kata yang dihiraukan saat melakukan pemrosesan teks dan umumnya dimasukkan ke dalam daftar stop lists. Daftar stop lists ini berisikan kata-kata umum yang memiliki fungsi tetapi tidak memiliki makna, seperti kata-kata penghubung "dan," "atau," "tapi," "akan," dan sebagainya.

### 2.4. Augmentasi data dengan metode BERT

Tujuan dilakukan augmentasi data dalam analisis sentimen adalah untuk meningkatkan ketahanan dan kemampuan generalisasi model prediksi analisis sentimen. Pada penelitian ini augmentasi juga dilakukan akibat dari jumlah data yang tidak seimbang antar kategori label positif dan negatif. Dataset dengan data label tidak seimbang akan berpengaruh pada saat pembangunan model klasifikasi teks dengan machine learning. Dampak dari ketidakseimbangan jumlah kelas (kategori) pada dataset adalah terjadinya kesalahan klasifikasi kelas minoritas (jumlah data sedikit), sehingga dapat mempengaruhi kinerja klasifikasi secara keseluruhan (Jiawei & Micheline, 2006). Jumlah baris data yang telah di augmentasi terlihat seperti pada Gambar 4.

Teknik augmentasi data yang digunakan untuk tujuan ini adalah BERT (Bidirectional Encoder Representations from Transformers). BERT adalah rangkaian model bahasa yang diperkenalkan oleh para peneliti di Google sejak tahun 2018. Ini adalah model pembelajaran mesin untuk pemrosesan NLP berbasis

teknik neural network. BERT dirancang untuk membantu komputer memahami makna bahasa ambigu dalam teks dengan menggunakan konteks sekitar (Devlin et al., 2019)

Penggunaan BERT untuk augmentasi data secara otomatis dilakukan dengan menghasilkan titik data baru dengan memanfaatkan kemampuannya untuk memahami dan menulis ulang teks. Hal ini dapat dicapai dengan menggunakan model bahasa bertopeng (masked language model). BERT mengganti atau menyisipkan kata-kata dalam teks input, menciptakan variasi baru dari kata sambil mempertahankan makna aslinya. Gambar berikut menunjukkan penerapan BERT augmentasi data pada salah satu baris data.

'pintu kamar mohon di perbaiki' → 'pintu or kamar ka mohon di ka perbaiki'

Gambar 4 Contoh Hasil Augmentasi dari metode BERT

## 2.5. Ekstraksi fitur dengan metode TF-IDF

Setelah pra-pemrosesan data, ekstraksi fitur diterapkan pada seluruh dataset sebelum dimasukkan kedalam model klasifikasi teks untuk proses training dan testing. Pada tahap ekstraksi fitur, teks yang telah diproses sebelumnya diubah menjadi vektor fitur numerik menggunakan teknik TF-IDF (*Term Frequency-Inverse Document Frequency*). TF-IDF merupakan metode pembobotan gabungan antara term frequency dan inverse document frequency (*An Introduction to TF-IDF*, n.d.). Metode TF-IDF digunakan untuk memilih fitur sebagai hasil ringkasan dari data komentar peserta pelatihan, dengan penerapannya pada seleksi fitur bobot setiap kata.

Dengan matriks TF-IDF ditentukan bobot pentingnya sebuah kata dari keseluruhan kata pada dokumen dalam sebuah korpus (dataset). TF-IDF memberikan bobot yang lebih tinggi pada kata-kata yang sering muncul dalam sekumpulan dokumen berlabel dengan polaritas sentimen tertentu tetapi paling sedikit terjadi dalam korpus. Secara umum, terdapat berbagai package yang tersedia untuk menghitung metrik TF-IDF. Dalam penelitian ini digunakan vektorizer TF-IDF dari scikit-learn (Pedregosa et al., 2011). Berikut rumus untuk mendapatkan bobot setiap kata dalam matriks TF-IDF:

$$w_{x,y} = tf_{x,y} \times \log \left( \frac{N}{df_x} \right)$$

**TF-IDF**  
Term x within document y

$tf_{x,y}$  = frequency of x in y  
 $df_x$  = number of documents containing x  
N = total number of documents

Gambar 5 Formula untuk mendapatkan skor TF-IDF sebagai salah satu fitur untuk prediksi klasifikasi

## 2.6. Klasifikasi teks dengan metode Naïve Bayes

Naive bayes termasuk salah satu algoritma klasifikasi supervised learning yang populer digunakan untuk analisis sentimen karena kesederhanaan dan efisiensinya (Jurafsky & Martin, n.d.). Naive Bayes adalah algoritma klasifikasi probabilistik yang menggunakan frekuensi kata dan fitur untuk menghitung probabilitas suatu dokumen termasuk dalam kelas sentimen tertentu (Webb et al., 2010). Naive Bayes bekerja baik untuk kasus klasifikasi teks dan pemfilteran spam. Algoritma ini mengasumsikan bahwa fitur (kata-kata) tidak bergantung satu sama lain, oleh karena itu disebut "naif". Untuk menggunakan Naive Bayes dalam analisis sentimen, langkah pertama adalah membagi kumpulan data menjadi kumpulan pelatihan dan pengujian. Kemudian, model Naive Bayes dilatih pada set pelatihan dan dievaluasi menggunakan set pengujian. Algoritma ini memanfaatkan asumsi independensi fitur berdasarkan label kelas dan menyederhanakan penghitungan. Model yang dilatih kemudian digunakan untuk memprediksi sentimen data baru. Meski terkenal dengan kesederhanaannya, Naïve Bayes bisa mengungguli metode klasifikasi canggih lainnya untuk aplikasi tertentu.

## 2.7. Pengukuran kinerja model klasifikasi dan prediksi sentimen

Terdapat beberapa ukuran umum untuk menilai kinerja model klasifikasi teks yang sudah diterapkan pada dataset. Pada penelitian ini, ukuran yang digunakan adalah akurasi ketepatan prediksi model terhadap data ujicoba dan confusion matrix. Akurasi merupakan ukuran proporsi data yang diklasifikasikan dengan benar dari total data. Akurasi adalah metrik paling sederhana dan seringkali digunakan untuk mengevaluasi model klasifikasi. Confusion matrix merupakan tabel yang digunakan untuk menggambarkan kinerja suatu model klasifikasi. Ini memberikan rincian jumlah data yang diprediksi model klasifikasi dengan benar dan salah (*Confusion Matrix*, n.d.). Hal ini sangat berguna untuk kumpulan data yang tidak seimbang.

### 3. Hasil dan Pembahasan

#### 3.1. Hasil Pengumpulan Data

Data yang berhasil dikumpulkan dari lembar evaluasi penyelenggaraan pelatihan sebanyak 600 baris (50 peserta X 12 angkatan pelatihan). Namun data mentah yang kemudian digunakan sebagai dataset untuk analisis sentimen adalah sebanyak 587 baris, 13 baris dinyatakan tidak valid dikarenakan isian data yang kurang lengkap atau 'null'. Gambar dibawah ini adalah hasil tangkapan layar untuk dua kelompok data komentar positif dan negative yang telah di bersihkan dilakukan pemrosesan awal.

Tabel 1 dibawah menunjukkan hasil rekapitulasi jumlah baris data yang digunakan dalam penelitian ini mencakup dataset awal yang dikumpulkan (RAW), dataset yang telah dilakukan pemrosesan awal untuk klasifikasi dan prediksi sentimen (KL) dan dataset hasil augmentasi sebagai bahan perbandingan hasil prediksi (AUG). Jumlah baris dataset KL dari pengumpulan dataset RAW berkurang dikarenakan, dataset sentiment netral tidak disertakan dalam tujuan prediksi dan klasifikasi sentiment. Sehingga jumlah data yang diklasifikasi dengan model machine learning sebanyak 433 baris data. Sedangkan dataset AUG adalah jumlah baris data yang dihasilkan oleh proses augmentasi BERT agar jumlah baris data antara kelompok komentar positif dan kelompok komentar negative seimbang, dengan total hasil dataset sebanyak 573 baris data.

**Tabel 2.** Rekapitulasi Jumlah Baris Data Hasil Pengumpulan

Kode Data	Jenis Data	Kategori Sentimen (Jumlah baris data)			Total Baris Data
		Positif	Negatif	Netral	
RAW	Dataset Mentah	295	147	145	587
KL	Dataset Klasifikasi yang telah diproses	289	144	-	433
AUG	Dataset Hasil Augmentasi	289	284	-	573

A	B
Comment	VALUE
alat kebersihan yang ada di asrama putri kurang lengkap ada kain pel dan sapunya ada yang rusak	negatif
untuk fasilitas olahraga kurang sepakbola	negatif
kurang lengkap untuk alat praktek	negatif
Saran saya agar bdi bisa memperbaiki fasilitas kamar yang kurang baik	negatif
untuk fasilitas olahraga kurang sepakbola	negatif
Sarana buat olahraga kurang	negatif
rasa makananya kurang memuaskan	negatif
tempat tidur untuk peserta diklat kurang memuaskan karena terlalu kecil	negatif
air sering mati	negatif
Comment	VALUE
Puas	positif
menjadi inspirasi	positif
selalu jadi yang terbaik untuk kedepan	positif
sangat baik untuk diadakan diklat seperti ini bagi yang ingin mengetahui lebih dalam tentang elek tronika	positif
sangat memuaskan	positif
Sangat membantu dalam pendidikan	positif

**Gambar 6** Hasil taangkapan layar dua kelompok data yang telah dikumpulkan

Pada proses klasifikasi dan prediksi analisis sentiment komentar, dataset yang dipergunakan adalah dataset KL dan dataset AUG sebagai bahan perbandingan untuk menunjukkan apakah terdapat pengaruh dari proses augmentasi data terhadap hasil kinerja model klasifikasi sentiment komentar peserta.

#### 3.2. Hasil Klasifikasi dan Prediksi Analisis Sentimen

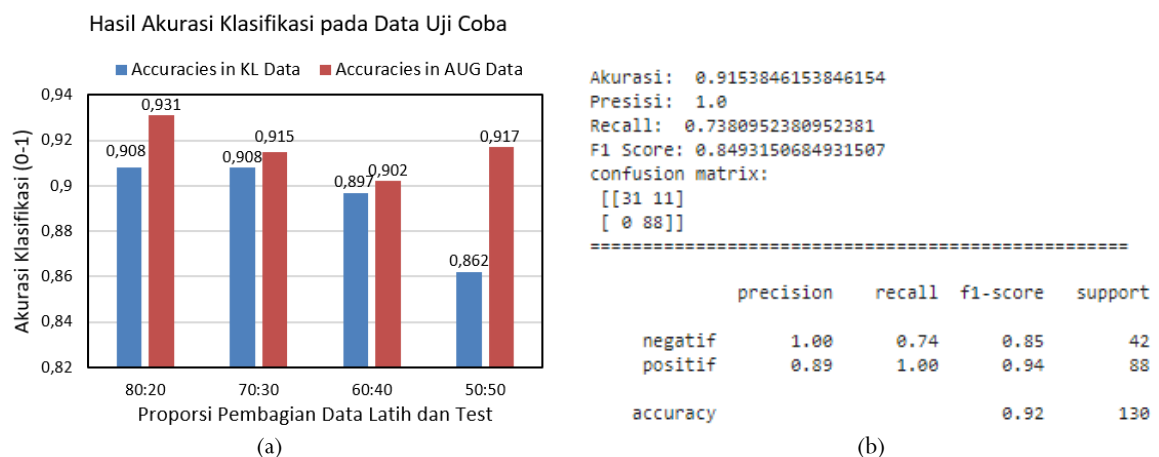
Pada tahap klasifikasi dan prediksi hasil yang diperoleh yakni pengukuran hasil kinerja dari model klasifikasi dan prediksi sentiment komentar evaluasi peserta dan system berbasis web (*localhost*) sebagai tampilan antar muka untuk melakukan input prediksi sentiment peserta secara langsung. Tabel 2 berikut menyajikan hasil kinerja klasifikasi dari model Naïve Bayes pada dataset KL dan AUG yang dilakukan terhadap data Latihan. Terdapat empat scenario pembagian proporsi data latih dan data ujicoba (test), yakni 80:20, 70:30, 60:40 dan 50:50.

Skenario pembagian proporsi ini dibuat dengan tujuan untuk menguji keandalan model klasifikasi terhadap jumlah data agar tidak terjadi *overfitting* (model klasifikasi dengan hasil yang terlalu akurat namun hanya untuk proporsi data latih besar). Seperti terlihat pada table, proporsi pembagian data latih dan test 50:50 mencapai akurasi terbaik yakni 0.954 (semakin mendekati 1 akurasi semakin baik). Dari Tabel 2 ini juga dapat dilihat bahwa hasil klasifikasi data AUG selalu meningkat dibandingkan dengan klasifikasi pada dataset KL.

**Tabel 2.** Hasil Akurasi klasifikasi pada Dataset KL dan AUG yang dilakukan pada data latih. (a) metode Naïve Bayes.

Proporsi Pembagian Data Latih dan Test	Hasil Akurasi klasifikasi pada data latih		
	Metode Naïve Bayes		Metode SVM
	KL Data	AUG Data	KL Data
80:20	0.948	0.948	0.920
70:30	0.947	<b>0.957</b>	0.908
60:40	0.942	0.954	0.856
50:50	<b>0.954</b>	0.949	0.823

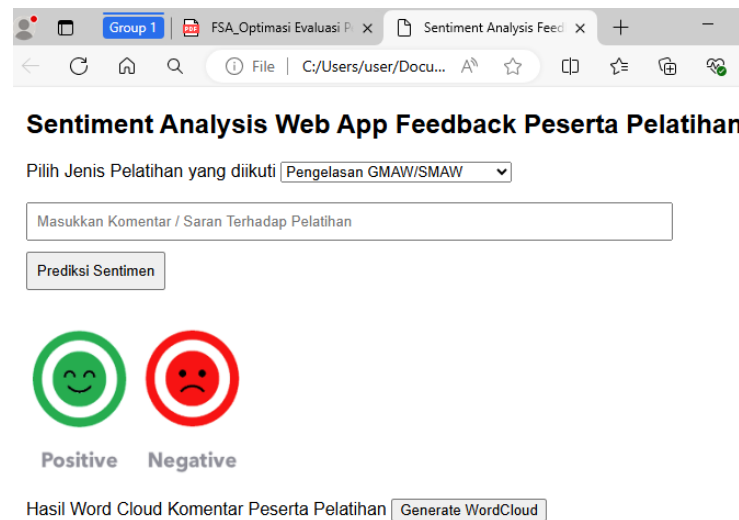
Sedangkan hasil dari klasifikasi pada data uji coba baik untuk dataset KL dan dataset AUG dapat dilihat perbandingannya pada gambar grafik berikut. Hasil yang dapat ditarik menjadi kesimpulan dari grafik gambar 6(a) juga serupa dengan hasil pada data latih (Tabel 2), yakni, klasifikasi pada jenis dataset AUG meningkat jika dibanding dataset KL. Hal ini berarti bahwa proses augmentasi data memberikan dampak pada peningkatan akurasi klasifikasi. Skor hasil akurasi pada data uji coba meskipun lebih rendah daripada data latih, namun tetap masih termasuk dalam batas ambang baik (yakni akurasi > 0.7) untuk sebuah model klasifikasi. Hal ini dapat diartikan bahwa model klasifikasi sentiment komentar beserta berkinerja baik dan handal untuk digunakan memprediksi sentiment dari komentar peserta pada form evaluasi penyelenggaraan pelatihan. Untuk memperkuat kesimpulan tersebut, hasil pada gambar 6(b) menunjukkan kinerja lebih lengkap dari hasil klasifikasi pada dataset AUG terhadap data uji coba dengan scenario proporsi pembagian data 70:30. Dapat dilihat pada confusion matrix di Gambar 6(b) jumlah hasil prediksi baris data pada setiap kelompok sentiment. Selain kinerja akurasi, kinerja presisi dari model klasifikasi sentiment komentar ini juga baik, yakni mencapai 1.0 (maksimal). Presisi menunjukkan seberapa sering model klasifikasi dapat memprediksi secara tepat untuk kelas target sentiment. Dari Tabel 2 dan Gambar 6 dapat diambil kesimpulan bahwa model klasifikasi sentiment komentar peserta ini dapat bekerja dengan baik dan digunakan untuk prediksi. Sehingga panitia penyelenggara tidak perlu melakukan anotasi manual apakah komentar peserta terhadap evaluasi tersebut positif/negatif.



**Gambar 7** Hasil Akurasi Klasifikasi pada Data Uji coba untuk dataset KL dan dataset AUG. (a) Grafik perbandingan akurasi;(b) Confusion matrix dari klasifikasi pada dataset AUG dengan proporsi pembagian data 70:30

Untuk memudahkan panitia penyelenggara pelatihan melakukan analisis sentiment dari komentar peserta pelatihan, maka dibuat sebuah web antar muka yang berisi menu untuk melakukan prediksi dari komentar peserta pelatihan per-baris data komentar (Gambar 7). Sementara, input komentar peserta masih dimasukkan per-baris komentar peserta. Kedepan, komentar peserta akan dimasukkan secara *bulk* (bersamaan) untuk satu periode

pelaksanaan pelatihan atau satu angkatan yang di ekspor datanya dari Gform evaluasi. Tampilan ini masih termasuk versi awal untuk menguji hasil prediksi jika ada inputan komentar peserta yang akan di analisis sentimennya.



**Gambar 8** Hasil tangkapan layar untuk web antar muka prediksi sentiment komentar peserta pelatihan.

Berdasarkan hasil klasifikasi dan prediksi analisis sentimen pada Tabel 2, hasil akurasi klasifikasi pada data latih, dapat ditarik kesimpulan bahwa metode analisis sentimen dengan klasifikasi model Naïve Bayes, masih terbukti lebih handal dalam model prediksi sentimen dengan akurasi yang lebih baik dibandingkan dengan metode SVM. Hasil ini juga dibuktikan oleh penelitian analisis sentimen sejenis yang juga menggunakan metode Naïve Bayes sebagai model klasifikasi (Amelia & Yustiana, 2024; Negara et al., 2020; Umair & Susanto, 2024). Dari segi pencapaian akurasi, hasil analisis sentimen pada penelitian ini mencapai hasil yang paling baik. Meskipun hal ini sangat bergantung dari berbagai hal, namun demikian pada kategori klasifikasi jenis kelas sentimen biner (positif dan negatif) capaian hasil akurasi usulan metode dari studi ini mencapai nilai terbaik diantara penelitian terdahulu. Hal tersebut dapat disebabkan adanya perbedaan metode pembersihan data, pemberian label sentimen yang lebih cermat, pemrosesan data awal, adanya penyeimbangan distribusi data antar kelas sentimen, hingga pemilihan model klasifikasi. Pada Tabel 3 berikut memuat hasil komparasi capaian akurasi klasifikasi sentimen dari beberapa studi penelitian sejenis.

**Tabel 3.** Komparasi dengan penelitian analisis sentimen (positif & negatif) sejenis di Indonesia dalam beberapa tahun terakhir

Pengarang	Objek Analisis Sentimen	Data berimbang (Ya/Tidak); Melakukan Penyeimbangan Kelas (Ya/Tidak)	Metode Ekstraksi Fitur	Metode Klasifikasi	Hasil Capaian Akurasi (maks. 1)
(Rusli, 2020)	Komentar Evaluasi Dosen	Ya; Tidak	Word2vec	SVM	0.848
(Negara et al., 2020)	Ulasan pengguna maskapai	Ya; Tidak	TF-IDF, Information Gain	Naïve Bayes	0.865
(Wahyudi & Kusumawardhana, 2021)	Ulasan pengguna aplikasi Grab	Tidak; Tidak	N-Gram	SVM	0.855
(Amrustian et al., 2022)	Evaluasi Pengajaran Dosen	Tidak; Ya	Word Vector	LSTM	0.918
(Aulia Rahman et al., 2023)	Ulasan pengguna Jasa Logistik J&T	Tidak; Tidak	TF-IDF	Multinomial Naïve Bayes	0.728
Penelitian ini	Ulasan komentar peserta pelatihan	Tidak; Ya;	TF-IDF	Naïve Bayes	0.957



Selain itu, visualisasi data teks dalam bentuk *word cloud* dari keseluruhan komentar umpan balik peserta memberikan nilai tambah dengan mempermudah proses evaluasi dan pemahaman oleh para pengambil keputusan. Grafik visual ini menampilkan representasi visual dari kumpulan komentar peserta, di mana kata-kata digambarkan dalam ukuran berbeda berdasarkan frekuensi kemunculannya. Dengan demikian, hasil penelitian analisis sentimen ini membantu penyelenggara pelatihan dalam membaca dan menginterpretasikan hasil komentar evaluasi peserta, meskipun data berjumlah besar dan dalam format teks. Pada akhirnya, tindak lanjut perbaikan hasil evaluasi oleh lembaga pelatihan menjadi lebih tepat sasaran dan sesuai dengan harapan peserta pelatihan.

Sebagai kelanjutan dari penelitian ini, diperlukan pengumpulan data yang lebih banyak lagi serta penerapan model analisis sentimen lainnya, termasuk model berbasis *unigram* dan *tree kernel*, untuk menunjukkan efektivitas dalam mengklasifikasikan opini. Hal ini sekaligus menjawab tantangan dalam mendeteksi komentar peserta yang mengandung sarkasme.

## Daftar Referensi

- Amelia, E. E., & Yustiana, I. (2024). *Analisis Sentimen Pada Ulasan Produk UNIQLO dengan Algoritma Naive Bayes*. 8.
- Amrustian, M. A., Widayat, W., & Wirawan, A. M. (2022). Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(1), 535. <https://doi.org/10.30865/mib.v6i1.3527>
- An Introduction to TF-IDF: What It Is & How to Use It*. (n.d.). Semrush Blog. Retrieved 4 July 2024, from <https://www.semrush.com/blog/tf-idf/>
- Aulia Rahman, H., Santoso, R., & Widiarini, T. (2023). Analisis Sentimen Pada Perusahaan Penyedia Jasa Logistik J&T Menggunakan Algoritma Multinomial Naive Bayes dan Support Vector Machine. *Jurnal Gaussian*, 12(2), 242–253. <https://doi.org/10.14710/j.gauss.12.2.242-253>
- Beranda: Kementerian Ketenagakerjaan RI. (n.d.). Retrieved 3 July 2024, from [https://binalavotas.kemnaker.go.id/lembaga\\_pelatihan\\_kerja](https://binalavotas.kemnaker.go.id/lembaga_pelatihan_kerja)
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. O'Reilly Media, Inc.
- Chahar, R. (2023). Evaluating the Impact of Sentiments in Decision Making: A Review. *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, 893–900.
- Confusion Matrix*. (n.d.). Retrieved 4 July 2024, from <https://socs.binus.ac.id/2020/11/01/confusion-matrix/>
- Conner, C., Samuel, J., Kretinin, A., Samuel, Y., & Nadeau, L. (n.d.). *Textual Visualization in Big Data Analytics*.
- Dalipi, F., Zdravkova, K., & Ahlgren, F. (2021). Sentiment Analysis of Students' Feedback in MOOCs: A Systematic Literature Review. *Frontiers in Artificial Intelligence*, 4, 728708. <https://doi.org/10.3389/frai.2021.728708>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv.
- Gian. (2024, May 17). Distribusi Sentimen dalam Ulasan Pelanggan untuk AIRism Cotton Oversize V Neck Half Sleeve T-Shirt.... *Medium*. <https://medium.com/@21611176/distribusi-sentimen-dalam-ulasan-pelanggan-untuk-airism-cotton-oversize-v-neck-half-sleeve-t-shirt-80a7e387c8d2>
- Herjanto, M. F. Y., & Carudin, C. (2024). ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI SIREKAP PADA PLAY STORE MENGGUNAKAN ALGORITMA RANDOM FOREST CLASSIFER. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(2). <https://doi.org/10.23960/jitet.v12i2.4192>
- Jiawei, H., & Micheline, K. (2006). *Data mining: Concepts and techniques*. Morgan kaufmann. [http://elib.vku.udn.vn/bitstream/123456789/2541/1/2006.%20The%20Morgan%20Kaufmann%20Series%20in%20Data%20Management%20Systems%29%20Jiawei%20Han%2C%20Micheline%20Kamber%2C%20Jian%20Pei%20-%20Data%20Mining\\_%20Concepts%20and%20Techniques-Morgan%20Kaufmann%20%282006%29.pdf](http://elib.vku.udn.vn/bitstream/123456789/2541/1/2006.%20The%20Morgan%20Kaufmann%20Series%20in%20Data%20Management%20Systems%29%20Jiawei%20Han%2C%20Micheline%20Kamber%2C%20Jian%20Pei%20-%20Data%20Mining_%20Concepts%20and%20Techniques-Morgan%20Kaufmann%20%282006%29.pdf)

- Jumlah Lembaga Pelatihan Terakreditasi Menurut Program Diklat dan Jenis Instansi—Portal Satu Data Indonesia.* (n.d.). Retrieved 3 July 2024, from <https://katalog.data.go.id/dataset/jumlah-lembaga-pelatihan-terakreditasi-menurut-program-diklat-dan-jenis-instansi>
- Jurafsky, D., & Martin, J. H. (n.d.). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.*
- Kabir, A. H. A., Basuki, S., & Wicaksono, G. W. (n.d.). *Analisis Sentimen Data Kritik Dan Saran Pelatihan Aplikasi Teknologi Informasi Menggunakan Algoritma Support Vector Machine.* 1(1).
- Lundqvist, K., Liyanagunawardena, T., & Starkey, L. (2020). Evaluation of student feedback within a MOOC using sentiment analysis and target groups. *The International Review of Research in Open and Distributed Learning*, 21(3), 140–156.
- Negara, A. B. P., Muhandi, H., & Putri, I. M. (2020). Analisis Sentimen Maskapai Penerbangan Menggunakan Metode Naive Bayes dan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(3), 599. <https://doi.org/10.25126/jtiik.2020711947>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830.
- Rouder, J., Saucier, O., Kinder, R., & Jans, M. (2021). What to Do With All Those Open-Ended Responses? Data Visualization Techniques for Survey Researchers. *Survey Practice*, 14(1), 1–9. <https://doi.org/10.29115/SP-2021-0008>
- Rusli, M. (2020). EKSTRAKSI FITUR MENGGUNAKAN MODEL WORD2VEC PADA SENTIMENT ANALYSIS KOLOM KOMENTAR KUISIONER EVALUASI DOSEN OLEH MAHASISWA. *KLIK - KUMPULAN JURNAL ILMU KOMPUTER*, 7(1), 35. <https://doi.org/10.20527/klik.v7i1.296>
- Santoso, V. I., Virginia, G., & Lukito, Y. (2017). PENERAPAN SENTIMENT ANALYSIS PADA HASIL EVALUASI DOSEN DENGAN METODE SUPPORT VECTOR MACHINE. *Jurnal Transformatika*, 14(2), 72. <https://doi.org/10.26623/transformatika.v14i2.439>
- Sinclair, J., & Cardew-Hall, M. (2008). The folksonomy tag cloud: When is it useful? *Journal of Information Science*, 34(1), 15–29. <https://doi.org/10.1177/0165551506078083>
- Umair, M., & Susanto, E. R. (2024). Analisis Sentimen Ulasan Pengguna Pada Aplikasi BRImo BRI Menggunakan Metode Klasifikasi Algoritma Naive Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(2), 1149. <https://doi.org/10.30865/mib.v8i2.7381>
- Wahyudi, R., & Kusumawardhana, G. (2021). Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *JURNAL INFORMATIKA*, 8(2).
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>
- Webb, G. I., Keogh, E., & Miikkulainen, R. (2010). Naïve Bayes. *Encyclopedia of Machine Learning*, 15(1), 713–714.
- Widoyoko, E. P. (2017). *Evaluasi Program Pelatihan.* Pustaka Pelajar.