



Tren Sentimen Media Sosial Terhadap Merek Mobil Terlaris di Indonesia Menggunakan BERT, SVM, dan Random Forest

Rizqi Maulidi¹, Taufik Edy Sutanto²

^{1,2}Matematika, Universitas Islam Negeri Syarif Hidayatullah Jakarta

¹rizqimaulidi44@gmail.com

Corresponding author: ²taufik.sutanto@uinjkt.ac.id

Abstract: The use of cars as a primary means of transportation remains a top choice for many Indonesians. Social media, as a platform where people express their opinions, can serve as a valuable source for automotive manufacturers to conduct sentiment analysis. This study analyzes sentiments from social media platforms such as Facebook, X (Twitter), YouTube, LinkedIn, TikTok, and Instagram. It focuses on the top five best-selling car brands in Indonesia in 2023, with data spanning from January 2023 to May 2023, utilizing the BERT model. The sentiment prediction results show an average accuracy of 85%, recall of 84.86%, precision of 84.79%, and an F1-Score of 85.07%. Toyota, the top seller, displayed an increase in both positive and negative sentiments from January 2023 to April 2023. Daihatsu showed consistent positive sentiment and a decrease in negative sentiment since March 2023. Honda experienced an increase in positive sentiment from January 2023 to May 2023 and a decrease in negative sentiment since March 2023. Mitsubishi saw only an increase in positive sentiment from January 2023 to April 2023, with non-constant decreases in negative sentiment since January 2023. Suzuki showed an increase in positive sentiment from February to March 2023, while negative sentiment was generally much lower compared to the other brands.

Keywords: Automotive Industry, BERT, Best-Selling Car Brands, Sentiment Analysis, Social Media,

Abstrak: Penggunaan mobil sebagai alat transportasi masih menjadi pilihan utama bagi sebagian besar warga Indonesia. Media sosial sebagai tempat masyarakat menyampaikan opini dapat menjadi sumber evaluasi berharga bagi produsen mobil untuk melakukan analisis sentimen. Penelitian ini menganalisis sentimen dari media sosial seperti Facebook, X (Twitter), YouTube, LinkedIn, Tiktok, dan Instagram. Studi ini memfokuskan pada lima merek mobil terlaris di Indonesia pada tahun 2023 dengan data dari Januari 2023 hingga Mei 2023 menggunakan model BERT. Hasil prediksi sentimen menunjukkan rata-rata akurasi sebesar 85%, recall 84,86%, presisi 84,79%, dan F1-Score 85,07%. Toyota dengan penjualan terlaris, menunjukkan kenaikan sentimen positif dan negatif dari Januari 2023 hingga April 2023. Daihatsu menunjukkan konsistensi sentimen positif dan penurunan sentimen negatif sejak Maret 2023. Honda mengalami peningkatan sentimen positif dari Januari 2023 hingga Mei 2023 dan penurunan sentimen negatif sejak Maret 2023. Mitsubishi hanya mengalami peningkatan sentimen positif dari Januari 2023 hingga April 2023 dan penurunan tidak konstan pada sentimen negatif sejak Januari 2023. Suzuki mengalami peningkatan sentimen positif dari Februari hingga Maret 2023, sementara sentimen negatif cenderung jauh lebih rendah dibandingkan merek lainnya.

Kata kunci: Analisis Sentimen, BERT, Industri Otomotif, Media Sosial, Merek Mobil Terlaris

I. PENDAHULUAN

Data yang diunggah di media sosial mengandung banyak kegunaan yang bisa diteliti dan dianalisis untuk menemukan informasi baru yang bermanfaat. Data yang tersedia di media sosial bisa dimanfaatkan untuk mendapatkan informasi khusus, sehingga menarik untuk analisis lebih mendalam [1]. Penggunaan kendaraan mobil masih menjadi pilihan utama bagi sebagian besar masyarakat Indonesia. Dalam era digital yang semakin berkembang, media sosial menjadi tempat masyarakat untuk memberikan pendapat mengenai suatu produk sehingga banyak diantara masyarakat yang menjadikan ulasan dari media sosial sebagai keputusan pembelian [2]. Hal ini menjadi tantangan kepada para produsen produk otomotif untuk membuat produk yang memenuhi ekspektasi masyarakat, selain itu media sosial juga dapat menjadi tempat marketing dan evaluasi bagi para produsen untuk meningkatkan kualitas produknya [3]. Media sosial yang digunakan dalam penelitian ini yaitu Facebook, X (Twitter), YouTube, LinkedIn, Tiktok, dan Instagram. Data yang digunakan mulai bulan Januari 2023 sampai bulan April 2024. Produk mobil sering menjadi fokus perbincangan karena merupakan aset yang penting



dan sering kali merepresentasikan gaya hidup dan status seseorang. Dalam konteks ini, analisis sentimen menjadi kunci untuk memahami tingkat kepuasan konsumen terhadap produk mobil.

Dalam penelitian ini akan menggunakan model BERT (*Bidirectional Encoder Representations from Transformers*) yaitu sebuah model bahasa yang dikembangkan oleh Google yang memanfaatkan arsitektur *Transformer* untuk menghasilkan representasi kata-kata yang lebih baik dalam pemrosesan bahasa alami (*Natural Language Processing*, NLP) [4]. Arsitektur BERT didasarkan pada arsitektur *Transformer*, yang merupakan sebuah model jaringan saraf yang dirancang khusus untuk memahami dan menghasilkan teks [5]. *Transformer* menggunakan *self-attention mechanism* untuk mengidentifikasi hubungan antara kata-kata dalam teks [6], mengatasi masalah kehilangan informasi jarak jauh yang sering terjadi dalam arsitektur rekuren pada model sebelumnya. Pendekatan *Masked Language Model* (MLM): Salah satu fitur kunci dari BERT adalah bahwa modelnya dilatih menggunakan pendekatan *Masked Language Model*. Dalam pendekatan ini, sebagian token dalam teks input secara acak (dilakukan *masking*), dan model bertugas untuk memprediksi token yang di-*mask* berdasarkan konteks dari token-token lainnya dalam kalimat tersebut. Pendekatan ini memungkinkan BERT untuk memahami konteks dari kata-kata dalam teks, baik secara individual maupun dalam konteks kalimat [4]. Selain menggunakan pendekatan MLM, BERT juga dilatih menggunakan pendekatan *Next Sentence Prediction*. Pendekatan ini melibatkan pelatihan model untuk memprediksi apakah sebuah kalimat kedua adalah kelanjutan dari kalimat pertama dalam sebuah pasangan kalimat. Hal ini memungkinkan BERT untuk memahami hubungan antara kalimat-kalimat dalam teks [7].

Pada penelitian sebelumnya telah dilakukan analisis sentimen menggunakan *Naive Bayes* dan memiliki hasil yang cukup baik [8], namun kelemahan utama *Naive Bayes* adalah asumsi independensi yang kuat antar fitur. Dalam konteks analisis teks, ini berarti bahwa *Naive Bayes* menganggap bahwa kemunculan satu kata dalam teks tidak terkait dengan kemunculan kata lainnya. Karena asumsi independensi, *Naive Bayes* tidak dapat menangani korelasi antara fitur dengan baik. Jika ada fitur yang sangat berkorelasi, *Naive Bayes* akan menganggap mereka sebagai independen, yang bisa mengarah pada hasil yang kurang akurat [9]. Oleh karena itu penelitian ini menggunakan model BERT yang merupakan salah satu teknologi pemrosesan bahasa alami (*Natural Language Processing*) yang dikenal dengan kemampuannya dalam memahami konteks dan hubungan antar kata dalam teks [10]. BERT dirancang untuk memahami konteks dari kata dalam kalimat dengan lebih baik karena menggunakan arsitektur *Transformer* yang dapat mempertimbangkan kedua arah (*bidirectional*). Ini berarti BERT dapat menangkap hubungan antara kata-kata dalam konteks yang lebih luas [11], sedangkan *Naive Bayes* menggunakan pendekatan probabilistik sederhana yang menganggap kata-kata independen satu sama lain [12]. Oleh karena itu meskipun akurasi BERT dalam penelitian ini lebih rendah dibandingkan *Naive Bayes*, namun model BERT memiliki banyak keunggulan yang dapat membuatnya lebih unggul dalam situasi dan jenis data tertentu. Selain itu dalam penelitian ini juga menggunakan model Support Vector Machine dan Random Forest sebagai perbandingan kinerja model BERT.

II. METODE PENELITIAN

Penelitian ini dilakukan dengan menggunakan data dari media sosial seperti Facebook, X (Twitter), YouTube, LinkedIn, Tiktok, dan Instagram mulai bulan Januari 2023 sampai bulan April 2024. Berikut tahapan dan metode yang digunakan dalam penelitian ini.



II.1. Pengambilan Sampel Relevan

Dalam paper ini memperkenalkan Analitik Berbasis Relevansi untuk memahami persepsi merek dan dinamika pasar dalam batasan regulasi data media sosial yang ketat. Metodologi inovatif ini terdiri dari dua komponen utama: Pengambilan Sampel Relevan dan Pembelajaran Mesin Berbasis Pengambilan Sampel Relevan. Setiap elemen dirancang untuk mengatasi kompleksitas etis dalam pengumpulan data sekaligus memastikan analisis mendalam terhadap konten yang dihasilkan oleh pengguna.

II.2. Pra Pemrosesan Data

Dalam proses ini dilakukan agar model dapat lebih mudah untuk memberikan label sentimen kepada data yang akan digunakan tahapan dari proses ini adalah *cleaning* yaitu proses pembersihan data dari berbagai karakter di dalam data seperti menghilangkan emoji, simbol, huruf, dan menghapus data duplikat, *case folding* yaitu proses mengubah huruf dalam data menjadi huruf kecil semua, *Tokenization* yaitu pemecahan menjadi token-token. BERT menggunakan *tokenizer* khusus yang menghasilkan token berdasarkan model BERT yang telah dilatih sebelumnya, *Lemmatization* Mengubah token ke bentuk dasarnya, selanjutnya adalah proses *stemming* yaitu proses yang mengubah kata imbuhan menjadi kata dasar setelah data sudah dilakukan melalui tahap pra pemrosesan data maka data tersebut siap diberikan label sentimen dengan model BERT.

II.3. Pemberian Label Sentimen Menggunakan Model BERT

Dalam penelitian ini pelabelan sentiment menggunakan model *BERT* khusus yang sebelumnya sudah dilatih untuk klasifikasi sentimen teks berbahasa Indonesia [13]. *Hyperparameter* tidak diatur secara eksplisit karena menggunakan model *pre-trained* ini secara langsung untuk prediksi tanpa melakukan *fine-tuning*. *Tokenizer* yang digunakan adalah *AutoTokenizer* dari model *pre-trained* yang sama.

II.4. Pemberian Label Sentimen Menggunakan Model SVM

Dalam penelitian ini, data dibagi menjadi data latih dan data uji dengan perbandingan sebesar 80:20. Kemudian teks mentah diubah menjadi vektor hitungan kata menggunakan *CountVectorizer*, kemudian menjadi representasi *TF-IDF* untuk mempertimbangkan frekuensi dan kepentingan kata. Model *SVM* yang digunakan adalah *LinearSVC* untuk klasifikasi teks, bekerja dengan menemukan *hyperplane* optimal yang memisahkan kelas-kelas sentimen (positif, negatif, dan netral) berdasarkan fitur yang diekstraksi dari teks.

II.5. Pemberian Label Sentimen Menggunakan Model Random Forest

Pertama, data dibagi menjadi set pelatihan dan pengujian dengan rasio 80:20. *Pipeline* klasifikasi kemudian dibentuk dengan menggunakan *CountVectorizer* untuk mengubah teks mentah menjadi vektor hitungan kata, kemudian *TF-IDF* yang mempertimbangkan frekuensi dan kepentingan kata, serta *RandomForestClassifier* dengan 100 pohon keputusan sebagai model klasifikasinya. Model dilatih menggunakan data pelatihan serta prediksi dilakukan pada data pengujian.

II.6. Evaluasi Model

Evaluasi model menggunakan perhitungan *confusion matrix* dengan data sampel yang diambil menggunakan metode *stratified sampling* dengan tujuan setiap kategori sentimen dalam populasi diwakili secara proporsional dalam sampel, sehingga hasil penelitian lebih akurat dan dapat digeneralisasi ke seluruh populasi. Diambil sepuluh data yang mewakili setiap kategori sentimen yaitu positif, negatif, dan netral, selanjutnya dilakukan perhitungan rata-rata *accuracy*, *recall*, *presicion*, serta



F1-Score, serta *area under the curve* dan *cross validation* dengan lima *fold*. untuk melihat kinerja model dengan lebih jelas.

Tabel 1. Parameter Matrik Evaluasi

| Matrik | Metode |
|------------------|-------------------------|
| <i>accuracy</i> | $(TP+TN)/(TP+FP+FN+TN)$ |
| <i>precision</i> | $TP/(TP+FP)$ |
| <i>Recall</i> | $TP/(TP+FN)$ |
| <i>F1-Score</i> | $2*(R*P)/(R+P)$ |

II.7. Analisis Hasil

Hasil akhir dari penelitian ini yang merupakan tren sentimen dari model yang memiliki performa terbaik karena paling mempresentasikan sentimen sebenarnya, per bulannya dari setiap merek mobil terlaris kemudian dianalisis untuk mencari insight yang bermanfaat untuk para produsen mobil maupun masyarakat.

III. HASIL DAN PEMBAHASAN

III.1. Hasil Pra Pemrosesan Data

Dari hasil *cleaning*, *case folding*, *tokenization*, *lemmatization*, kemudian *stemming* maka diperoleh hasil seperti dibawah ini.

Tabel 2. Tabel Hasil Pra Pemrosesan Data

| Data Awal | Data Bersih |
|---|---|
| Free Aksesoris (Kaca Film, Karpet Dasar, Karpet Ori, Dompot Honda, Gantungan Kunci, Sensor Parkir, APAR, dll). *syarat & ketentuan berlaku (MELAYANI Mobil Honda Payakumbuh added a... - Mobil Honda Payakumbuh Mobil Honda Payakumbuh added a new photo. | free aksesoris kaca film karpet dasar karpet ori dompot honda gantungan kunci sensor parkir apar dll syarat ketentuan berlaku melayani mobil honda payakumbuh added a mobil honda payakumbuh mobil honda payakumbuh added a new photo |

III.2. Hasil Labeling Model

Dari hasil ini didapatkan total sentimen untuk setiap model yang digunakan sebagai perbandingan setelah melalui tahap pra pemrosesan Data.

Tabel 3. Tabel Total Sentimen Dari Setiap Model

| Model | Positif | Negatif | Netral |
|----------------------|---------|---------|--------|
| <i>BERT</i> | 1131 | 217 | 4054 |
| <i>SVM</i> | 1110 | 175 | 4117 |
| <i>Random Forest</i> | 964 | 171 | 4267 |

III.3. Evaluasi Model

Pada proses ini evaluasi model dilakukan dengan cara mengambil beberapa sampel yang mewakili setiap sentimen dengan teknik *stratified* sampling, kemudian dengan menggunakan *confusion matrix* akan rata-rata *accuracy*, *recall*, *presicion*, dan *F1-Score*. Selain itu dilakukan juga perhitungan *area under the curve* dan *cross validation* dengan lima *fold*. untuk melihat kinerja model dengan lebih jelas.

Tabel 4. Tabel Matrik Evaluasi

| Model | <i>Accuracy</i> | <i>recall</i> | <i>presicion</i> | <i>F1-Score</i> |
|----------------------|-----------------|---------------|------------------|-----------------|
| <i>BERT</i> | 85% | 84,86% | 84,79% | 85,07% |
| <i>SVM</i> | 75% | 75,74% | 77,33% | 75,67% |
| <i>Random Forest</i> | 70% | 71,11% | 76,19 | 70,79% |



Berdasarkan metrik evaluasi, model *BERT* menunjukkan performa terbaik dengan akurasi 85%, *recall* 84,86%, presisi 84,79%, dan *F1-Score* 85,07%, menunjukkan keseimbangan yang baik antara *recall* dan presisi. Model *SVM* memiliki akurasi 75%, *recall* 75,74%, presisi 77,33%, dan *F1-Score* 75,67%, yang menunjukkan performa yang cukup baik namun lebih rendah dibandingkan *BERT*. Model *Random Forest* memiliki akurasi 70%, *recall* 71,11%, presisi 76,19%, dan *F1-Score* 70,79%, menunjukkan performa terendah di antara ketiga model. Oleh karena itu, *BERT* menjadi model terbaik untuk klasifikasi sentimen dengan kinerja yang paling konsisten dan tinggi, sementara *SVM* dan *Random Forest* dapat dipertimbangkan dengan kesadaran akan penurunan performa.

Tabel 5. Tabel *Area Under The Curve*

| Model | Positif | Negatif | Netral |
|----------------------|---------|---------|---------|
| <i>BERT</i> | 86,96 % | 96,25 % | 82,94 % |
| <i>SVM</i> | 75,24% | 92,50% | 76,98% |
| <i>Random Forest</i> | 71,05% | 90% | 73,81% |

Kemudian dari tabel *Area Under Curve (AUC)*, model *BERT* menunjukkan performa terbaik dalam klasifikasi sentimen dengan *AUC* tertinggi untuk sentimen negatif (0.9625), netral (0.8294), dan positif (0.8696), terutama unggul dalam klasifikasi sentimen negatif. Model *SVM* juga memberikan hasil yang baik, terutama untuk sentimen negatif (0.9250), namun performanya menurun untuk sentimen netral (0.7698) dan positif (0.7524). Model *Random Forest* memiliki performa terendah di antara ketiga model, dengan *AUC* terendah pada sentimen positif (0.7105) dan netral (0.7381), meskipun masih cukup baik untuk sentimen negatif (0.9000). Oleh karena itu, *BERT* direkomendasikan untuk klasifikasi sentimen dengan akurasi tinggi, sementara *SVM* dan *Random Forest* mungkin dipertimbangkan dengan pertimbangan performa dan interpretabilitas.

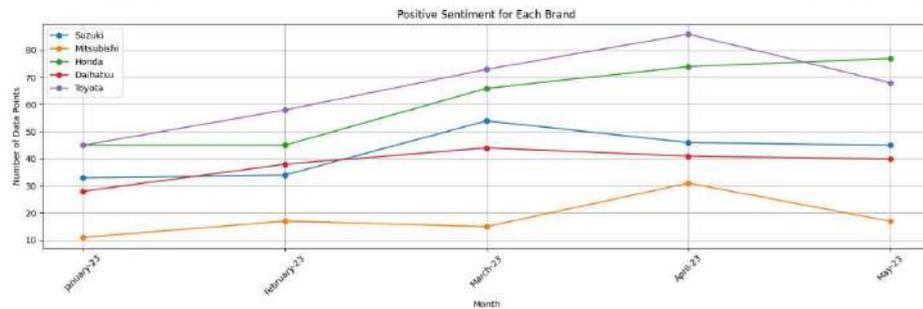
Tabel 6. Tabel *Cross-Validation*

| <i>Fold</i> | <i>BERT</i> | <i>SVM</i> | <i>Random Forest</i> |
|----------------|-------------|------------|----------------------|
| 1 | 91,67% | 41,67% | 58,33% |
| 2 | 100% | 66,67% | 33,33% |
| 3 | 100% | 58,33% | 58,33% |
| 4 | 66,67% | 75% | 58,33% |
| 5 | 66,67% | 66,67% | 41,67% |
| <i>Average</i> | 85% | 61,67% | 50% |

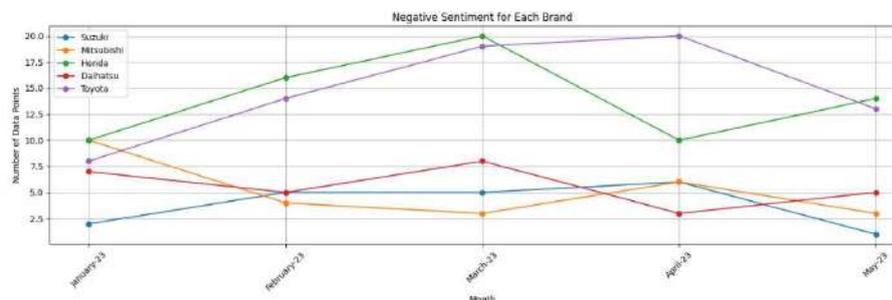
Dari hasil *cross-validation*, model *BERT* memiliki performa terbaik dengan rata-rata skor 0.85, meskipun terdapat variasi yang signifikan antara *fold* (skor 0.6667 hingga 1.0). Model *SVM* menunjukkan performa yang lebih rendah dengan rata-rata skor 0.6167, yang lebih konsisten namun tidak setinggi *BERT*. Model *Random Forest* memiliki performa paling rendah dengan rata-rata skor 0.5, menunjukkan variasi yang lebih kecil namun performa keseluruhan yang kurang memadai. Sehingga model *BERT* menjadi model terbaik untuk tugas klasifikasi ini,

III.4. Analisis Hasil

Pada proses evaluasi model model *BERT* diketahui menjadi model dengan performa paling knsisten. Oleh karena itu akan akan dilakukan tren sentimen serta *wordcloud* unigram positif dan negative dari setiap merek mobil di Indonesia berdasarkan model *BERT*, akan di analisis untuk mendapatkan insight.



Gambar 2. Gambar Tren Sentimen Positif Dari Setiap Merek



Gambar 3. Gambar Tren Sentimen Negatif Dari Setiap Merek

Dari hasil tren sentimen diatas diketahui Toyota yang merupakan merek dengan penjualan terlaris mengalami peningkatan sentimen positif menunjukkan keberhasilan promosi pemasaran atau peluncuran produk baru, sementara kenaikan sentimen negatif bisa disebabkan oleh isu atau keluhan pelanggan yang muncul seiring meningkatnya jumlah pengguna. Daihatsu menunjukkan kestabilan dalam sentimen positif yang mengindikasikan kepuasan pelanggan yang baik dan penurunan sentimen negatif yang mengindikasikan adanya perbaikan dalam layanan pelanggan atau peningkatan kualitas produk sejak Maret 2023. Honda menunjukkan peningkatan sentimen positif yang mengindikasikan keberhasilan strategi pemasaran atau peluncuran produk yang diterima baik sementara untuk sentimen negatifnya menunjukkan penurunan sejak Maret 2023, namun sentimen negatifnya berada di puncak pada bulan April 2023 sama dengan toyota yang merupakan merek terlaris. Mitsubishi mengalami peningkatan sentimen positif yang menunjukkan keberhasilan meningkatnya kepuasan masyarakat sejak bulan Januari sampai April 2023. Penurunan yang tidak konsisten pada sentimen negatif mungkin menunjukkan adanya fluktuasi dalam kualitas produk atau layanan yang perlu diperbaiki. Suzuki mengalami peningkatan sentimen positif sejak bulan Februari sampai bulan Maret 2023 yang menunjukkan adanya penerimaan baik dari masyarakat. Sentimen negatifnya memuncak pada bulan April 2023 Mungkin terkait dengan masalah tertentu yang muncul di bulan tersebut.

IV. KESIMPULAN

Berdasarkan analisis dan pengujian yang dilakukan pada penelitian ini, maka didapatkan kesimpulan model *BERT* menunjukkan kinerja terbaik dengan rata-rata akurasi sebesar 85%, recall 84,86%, presisi 84,79%, dan F1-Score 85,07%. Berdasarkan wordcloud, sentimen positif terhadap Mitsubishi menunjukkan keberhasilan promosi, sementara sentimen negatif mengindikasikan masalah kehilangan kunci. Toyota mendapat banyak perhatian di Instagram, dengan isu pelaku dan korban mendominasi sentimen negatif. Honda mendapat respon positif untuk mobil Brio, tetapi perlu evaluasi di Instagram untuk sentimen negatif. Suzuki dipuji untuk slogan dan Ertiga, namun perlu evaluasi di



Instagram. Daihatsu dominan di Instagram dengan isu aki yang mendominasi sentimen negatif. Tren sentimen menunjukkan keberhasilan promosi Toyota dan Daihatsu dengan peningkatan sentimen positif, sementara Honda dan Suzuki menunjukkan fluktuasi dengan puncak sentimen negatif pada April 2023. Mitsubishi juga menunjukkan peningkatan kepuasan sejak Januari, namun fluktuasi sentimen negatif menunjukkan perlu adanya perbaikan kualitas. Namun perlu dilakukan penelitian lebih lanjut untuk mengetahui secara pasti apa yang menjadi kepuasan ataupun kekecewaan pengguna media sosial terhadap suatu produk otomotif.

DAFTAR PUSTAKA

1. M. Liebenlito, N. Inayah, E. Choerunnisa, T. E. Sutanto, and S. Inna, “Active Learning on Indonesian Twitter Sentiment Analysis Using Uncertainty Sampling,” *Journal of Applied Data Sciences*, vol. 5, no. 1, pp. 114–121, Jan. 2024, doi: 10.47738/jads.v5i1.144.
2. A. Fahmi Pangestu Ramadhan and F. Munawar, “PENGARUH MEDIA SOSIAL INSTAGRAM DAN KUALITAS PELAYANAN TERHADAP KEPUTUSAN PEMBELIAN (STUDI KASUS PADA ARTOMORO MOTOR CIMAHU),” vol. 6, no. 3, p. 2022.
3. S. I. Isnawati, M. I. Pratiwi, J. Ramadany Bangsa, K. Pemasaran, and V. Pemasaran, “PELATIHAN KONTEN MARKETING PADA INDUSTRI OTOMOTIF DENGAN MEDIA VIDEO MARKETING DI PT WAHANA INVESTASINDO SALATIGA Informasi Artikel Abstrak Kata kunci,” 2023. [Online]. Available: <http://jurnal.unipasby.ac.id/index.php/abadimas~240~>
4. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Oct. 2018, [Online]. Available: <http://arxiv.org/abs/1810.04805>
5. P. Su and K. Vijay-Shanker, “Investigation of improving the pre-training and fine-tuning of BERT model for biomedical relation extraction,” *BMC Bioinformatics*, vol. 23, no. 1, Dec. 2022, doi: 10.1186/s12859-022-04642-w.
6. A. D. Rendragraha, M. Arif Bijaksana, and A. Romadhony, “Pendekatan Metode Transformers untuk Deteksi Bahasa Kasar dalam Komentar Berita Online Indonesia.”
7. X. Wu, T. Zhang, L. Zang, J. Han, and S. Hu, “‘Mask and Infill’: Applying Masked Language Model to Sentiment Transfer,” Aug. 2019, [Online]. Available: <http://arxiv.org/abs/1908.08039>
8. D. Rustiana Program Studi Sistem Komputer Perguruan Tinggi Raharja and N. Rahayu Magister Teknologi Informatika Perguruan Tinggi Raharja, “ANALISIS SENTIMEN PASAR OTOMOTIF MOBIL: TWEET TWITTER MENGGUNAKAN NAÏVE BAYES,” *Jurnal SIMETRIS*, vol. 8, 2017.
9. F. S. Pamungkas and I. Kharisudin, “Analisis Sentimen dengan SVM,” vol. 4, pp. 628–634, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
10. Koroteev MV, “BERT: A Review of Applications in Natural Language Processing and Understanding.”
11. R. Qasim, W. H. Bangyal, M. A. Alqarni, and A. Ali Almazroi, “A Fine-Tuned BERT-Based Transfer Learning Approach for Text Classification,” *J Healthc Eng*, vol. 2022, 2022, doi: 10.1155/2022/3498123.
12. S. Rana, R. Kanji, and S. Jain, “Comparison of SVM and Naïve Bayes for Sentiment Classification using BERT data,” in *2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, 2022, pp. 1–5. doi: 10.1109/IMPACT55510.2022.10029067.
13. J. Khatib Sulaiman, M. Misbah Effendy, T. Edy Sutanto, M. Liebenlito, and U. Syarif Hidayatullah Jakarta, “2023 | page 4054 Efektivitas Variabel Demografi Pengguna Twitter dalam Prediksi Pilpres Indonesia,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 6, 2014.