

Sentiment Analysis on Social Media using Bidirectional Encoder from Transformers (Case Study : Covid – 19 Omicron)

Dody Pradipta¹, Eko Widodo²

¹Program Studi Teknik Komputer, ²Program Studi Manajemen
Universitas Islam Kadiri, Jl. Sersan Suharmaji No.38, Kota Kediri
Email : pradiptadody@uniska-kediri.ac.id

ABSTRAK

Indonesia tercatat sebagai negara dengan jumlah pengguna internet terbesar di Dunia, sehingga adanya perubahan komunikasi dimasyarakat. Kemajuan jaman membuat interaksi manusia banyak dihabiskan pada media sosial, diantaranya Instagram, facebook, twitter (X) dan Tiktok mendominasi ruang digital. Tingginya interaksi pada sosial media, tidak bisa dipungkiri kerap kali bias, menimbulkan banyak persepsi baik karena subjektivitas, pandangan pribadi, maupun emosi sesaat. Pandemi Covid-19 memaksa pemerintah mencari solusi untuk berbagai masalah yang muncul, namun kebijakan yang diambil seringkali memicu gejolak di dunia maya. Penelitian ini bertujuan untuk melihat seberapa baik performa BERT dalam mengklasifikasikan komentar, BERT (*Bidirectional Encoder Representations from Transformers*) memanfaatkan pendekatan *word embedding* untuk memahami konteks kata dalam dua arah. Arsitektur BERT menggunakan transformers, mekanisme yang memberikan makna kontekstual pada kata, membantu BERT dalam mengklasifikasikan komentar. Keunggulan BERT terletak pada pemahaman konteks kata yang mendalam melalui *embedding* kontekstual, dengan *encoder transformers* yang membaca kalimat dua arah.. Data yang digunakan adalah komentar pada instagram dan facebook didapatkan data sebanyak 3.522 data. Fokus penelitian ini adalah Covid – 19 varian *omicron*, topik ini diambil karena polaritas reaksi pada masyarakat. Model pada penelitian ini menggunakan 4 proporsional dataset, untuk melihat kontribusi data latih untuk model yang dikembangkan, sedangkan untuk evaluasi menggunakan dataset yang dilabeli secara manual. Penelitian ini mengungkapkan bahwa BERT dapat digunakan untuk mengklasifikasikan sentimen dengan hasil yang baik, model yang dibangun mendapatkan rentang efisiensi pada *epoch* 3 sampai 5, meskipun nilai terbaik berada *epoch* 8, dengan akurasi sebesar 90%. Mayoritas sentimen menunjukkan sentimen negatif ditengarai timbul panik dimasyarakat terkait penanganan Pandemi Covid – 19.

Kata Kunci : Analisis Sentimen, BERT, Covid – 19

ABSTRACT

Indonesia is recorded as one of the countries with the largest number of internet users in the world, leading to changes in communication within society. The advancement of the times has shifted much of human interaction to social media, with platforms such as Instagram, Facebook, Twitter (X), and TikTok dominating the digital space. The high level of interaction on social media often leads to biases, creating various perceptions influenced by subjectivity, personal views, or momentary emotions. The Covid-19 pandemic forced the government to find solutions to various emerging issues, but the policies implemented frequently sparked unrest in the online

world. This research aims to evaluate the performance of BERT in classifying comments. BERT (Bidirectional Encoder Representations from Transformers) utilizes a word embedding approach to understand the context of words in both directions. BERT's architecture employs transformers, a mechanism that provides contextual meaning to words, aiding in comment classification. The strength of BERT lies in its deep understanding of word context through contextual embeddings, with transformers' encoders reading sentences bidirectionally. The data used in this study consists of 3,522 comments from Instagram and Facebook. The focus of this research is on the Covid-19 Omicron variant, chosen due to the polarity of public reactions. The model in this study uses 4 proportional datasets to assess the contribution of training data to the developed model, while manual labeling is applied for evaluation. The findings reveal that BERT is effective in classifying sentiment with good results, achieving efficiency between epochs 3 and 5, although the best performance occurs at epoch 8 with an accuracy of 90%. The majority of sentiments are negative, likely due to public panic over the handling of the Covid-19 pandemic.

Keywords : *Sentiment Analysis, BERT, Covid – 19*

1. PENDAHULUAN

Indonesia tercatat sebagai salah satu negara dengan jumlah pengguna internet terbesar di Dunia. APJII sebagai asosiasi pengguna internet Indonesia mencatat bahwa pengguna internet di Indonesia di tahun 2024 mencapai 221.563.479 jiwa dari total populasi 278.696.200 jiwa penduduk Indonesia tahun 2023. Pertumbuhan akses internet yang pesat ini didorong oleh penetrasi perangkat seluler dan kemudahan akses jaringan di berbagai wilayah, termasuk daerah terpencil. Media Sosial menjadi salah satu platform yang paling banyak digunakan oleh masyarakat Indonesia, platform seperti Instagram, facebook, twitter (X) dan Tiktok mendominasi ruang digital. Hal ini menjadikan Indonesia sebagai salah satu negara dengan aktivitas media sosial yang tinggi di Dunia.

Tingginya interaksi pada sosial media, tidak bisa dipungkiri kerap kali mengundang bias, baik karena subjektivitas, pandangan pribadi, maupun emosi sesaat. Konteks ranah sosial – politik, komentar – komentar di media sosial sering mencerminkan sentimen yang kuat terhadap suatu peristiwa atau kebijakan, baik dalam bentuk dukungan maupun kritik. Bias yang muncul dalam komentar ini dapat menjadi cerminan penting untuk memahami bagaimana publik memandang kebijakan yang dikeluarkan oleh pemerintah atau lembaga terkait.

Kebijakan publik yang memicu banyak kontroversi adalah penanganan terhadap Covid – 19, terutama dalam pembatasan aktivitas publik yang berdampak

pada keterbatasan mobilitas dan kegiatan di luar rumah. Kebijakan seperti Pembatasan Sosial Berskala Besar (PSBB) dan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) memicu beragam reaksi di masyarakat. Banyak yang mendukung langkah tersebut sebagai bentuk perlindungan kesehatan, namun tidak sedikit pula yang merasa keberatan, terutama karena dampak ekonomi yang ditimbulkan.

Penelitian (Syukur et al., 2021) mengungkapkan bahwa krisis kesehatan ini tidak hanya memengaruhi sektor kesehatan, tetapi juga menimbulkan tantangan besar di bidang ekonomi. Oleh karena itu, diperlukan langkah-langkah kebijakan yang komprehensif untuk mengatasi dampak ekonomi yang timbul di tengah upaya penanganan krisis kesehatan.

Pandemi membawa dampak yang luas dan signifikan mengungkapkan bahwa perlambatan pertumbuhan ekonomi terjadi di seluruh Indonesia, dengan banyak sektor usaha mengalami penurunan omzet, produksi yang merosot, serta tantangan dalam pemasaran. Selain itu, sektor ketenagakerjaan juga terdampak secara serius, di mana banyak pekerja dirumahkan untuk jangka waktu yang tidak pasti. Hal ini menyebabkan meningkatnya angka pengangguran dan ketidakpastian ekonomi, terutama bagi mereka yang bekerja di sektor informal dan UMKM, yang sangat bergantung pada interaksi langsung dengan konsumen. Kompleksitas situasi selama pandemi menuntut pemerintah untuk mengambil langkah – langkah kebijakan yang cepat dan tepat guna mengatasi dampak kesehatan serta memitigasi dampak ekonomi yang terjadi.

Sentiment Analysis merupakan teknik yang digunakan untuk menganalisis pandangan, perasaan, dan penilaian pengguna berdasarkan informasi seperti opini yang tertulis. Teknik ini bertujuan untuk memahami keyakinan pengguna terhadap suatu topik atau subjek tertentu (Chinnasamy et al., 2022). *Sentiment Analysis* memiliki hubungan erat dengan *Natural Language Processing* (NLP), yang merupakan kemampuan komputer untuk memahami bahasa manusia melalui algoritma, matematika, dan linguistik komputasional. Beberapa contoh aplikasi NLP termasuk Google Assistant dan Siri. Python sering digunakan dalam NLP karena tersedianya banyak *library open source*, kemudahan penggunaan, serta sistem optimalisasi yang tidak serumit paradigma pemrograman lainnya (Jalaj Thanaki,

2017).

NLP digunakan untuk menganalisis data tekstual dengan pendekatan berbasis aturan dan machine learning. Dalam *Sentiment Analysis*, metrik penting yang digunakan adalah polaritas dan subjektivitas. Pelatihan model biasanya dipantau menggunakan diagram garis *training loss*, sedangkan evaluasi akurasi mengukur efektivitas model dalam menganalisis sentimen. Model yang digunakan dalam penelitian ini adalah BERT (*Bidirectional Encoder Representations from Transformers*) yang tetap mempertahankan kemampuan pemrosesan Bahasa.

Segala bentuk kebijakan pemerintah terhadap penanganan Covid – 19 tidak luput dari opini publik dari masyarakat. Baik yang mendukung terhadap kebijakan maupun kontra terhadap kebijakan yang diambil oleh pemerintah. Penelitian (Kurniasari et al., 2021) yang meneliti tentang komentar pada instagram menggunakan SVM (*Support Vector Machine*) mendapatkan temuan bahwa metode SVM dapat digunakan untuk mengklasifikasikan komentar dengan akurasi 59%.

Penelitian (Samsir et al., 2021) meneliti tentang “Analisis Sentimen Pembelajaran Daring pada Twitter di Masa Pandemi Covid-19 Menggunakan Metode Naive Bayes”, ditemukan bahwa 30% komentar positif, 69% komentar negatif dan 1% komentar netral. Tingginya sentimen negatif menunjukkan bahwa terdapat kekecewaan terhadap pembelajaran secara daring sehingga perlu diadakan evaluasi. Hasil penelitian ini menyoroti adanya ketidakpuasan yang signifikan di kalangan siswa dan orang tua terhadap implementasi pembelajaran daring selama pandemi. Evaluasi yang komprehensif terhadap metode dan pelaksanaan pembelajaran daring sangat penting untuk meningkatkan kualitas dan efektivitasnya di masa mendatang.

BERT (*Bidirectional Encoder Representations from Transformers*) memanfaatkan pendekatan *Word Embedding* untuk memahami konteks kata dalam dua arah, baik dari kiri ke kanan maupun sebaliknya. Hal ini memungkinkan BERT untuk menangkap makna yang lebih mendalam dari sebuah kata berdasarkan konteks kalimat secara keseluruhan. Penelitian (Kurniasari et al., 2021) juga menyebutkan bahwa penggunaan *Word Embedding* dapat meningkatkan akurasi dalam berbagai tugas pemrosesan bahasa alami, seperti analisis sentimen, klasifikasi teks, dan penjawaban otomatis. Kemampuannya yang lebih unggul dalam memahami konteks, BERT telah menjadi salah satu model andalan dalam bidang NLP (*Natural Language*

Processing) yang digunakan secara luas dalam berbagai aplikasi modern.

BERT menghasilkan representasi untuk setiap kata dalam kalimat sebagai *output* dan dapat meningkatkan kinerja model pada tugas – tugas berurutan kompleks dalam NLP. Arsitektur BERT melibatkan *Transformers* yang diungkapkan pada *paper* “*Attention Is All You Need*” (Vaswani et al., 2017), BERT adalah salah satu model penyematan (*Embedding*) teks canggih yang diterbitkan oleh Google (Sudharsan Ravichandiran, 2021). BERT adalah model representasi berbasis *fine – tuning* pertama yang mencapai kinerja canggih pada rangkaian besar tugas *sentence – level* dan *token – level* dan mengungguli banyak arsitektur *task – specific* (Devlin et al., 2019).

Pemilihan BERT didasarkan pada penelitian berikut ini “*An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews*” (Ray et al., 2021), “*Sentiment Analysis on The Impact of Coronavirus in Social Life using The BERT Model*” (Singh et al., 2021), “*An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US*” (Zhang et al., 2022), “*Transformer-based deep learning models for the sentiment*” (Tabinda Kokab et al., 2022), “*Predictive Intelligence in Harmful News Identification by BERT-Based Ensemble Learning Model with Text Sentiment Analysis*” (Lin et al., 2022), “*BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews*” (Nugroho et al., 2021), “*Sentimental analysis on user’s reviews using BERT*” (Selvakumar & Lakshmanan, 2022). Seluruh penelitian tersebut menyebutkan bahwa BERT menempati hasil penelitian yang tertinggi dibandingkan model dengan algoritma lainnya.

BERT telah terbukti sangat efektif dalam berbagai tugas NLP, seperti menjawab pertanyaan, pembuatan teks, dan klasifikasi kalimat (Vidya Chandradev et al., 2023). Keunggulan BERT terletak pada kemampuannya memahami konteks kata secara mendalam melalui model *embedding* berbasis konteks yang memanfaatkan *encoder* dari *Transformers* untuk membaca kalimat dari dua arah. Pendekatan *bidirectional* ini memungkinkan BERT memahami hubungan antar kata lebih baik dibandingkan model sebelumnya, sehingga meningkatkan akurasi dalam berbagai aplikasi NLP. Dengan demikian, BERT menjadi model pilihan dalam banyak penelitian dan aplikasi praktis di bidang pemrosesan bahasa alami.

Berdasarkan fenomena di atas, peneliti bertujuan untuk menganalisis sentimen komentar di media sosial menggunakan model BERT (*Bidirectional Encoder Representations from Transformers*). BERT sangat cocok untuk digunakan dalam analisis sentimen karena kemampuannya dalam memahami konteks kata secara dua arah (*bidirectional*), yaitu dengan mempertimbangkan kata-kata sebelum dan sesudah dalam sebuah kalimat. Hal ini memungkinkan BERT untuk menangkap makna yang lebih kompleks dibandingkan model terdahulu, sehingga dapat menghasilkan analisis sentimen yang lebih akurat. Penelitian ini, BERT diharapkan mampu mengidentifikasi nuansa dan pola sentimen yang mungkin terlewatkan oleh metode lain, terutama dalam menangani komentar yang sering kali ambigu atau memiliki konteks yang rumit. Dengan demikian, penggunaan BERT tidak hanya relevan, tetapi juga memberikan hasil yang lebih terpercaya dalam memahami opini masyarakat terhadap kebijakan publik.

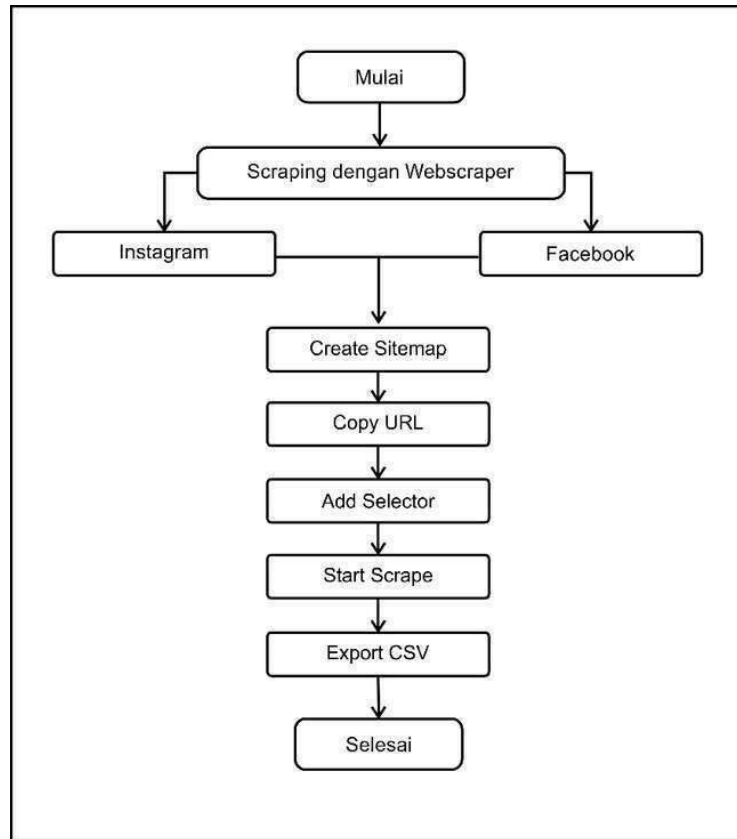
2. METODE

Jenis, Sifat Dan Pendekatan Penelitian

Merupakan penelitian eksperimen dan studi kasus, dimana penelitian eksperimen adalah penelitian yang dilakukan dengan menerapkan serangkaian tindakan untuk membuktikan suatu konsep (Prof. Dr. Sugiyono, 2018). Perihal ini penulis melaksanakan eksperimen dengan *scraping* informasi pada instagram dan facebook kemudian diklasifikasikan menggunakan BERT.

Metode Pengumpulan Data

Pengumpulan data dengan cara *scraping* data komentar instagram dan facebook dengan menggunakan webscraper, berdasarkan postingan pada instagram dan facebook yang mengangkat berita terhadap covid – 19 varian *omicron*.



Gambar 1: Alur Pengumpulan Data

Hasil *scraping* menghasilkan 3.522 data, komentar yang didapatkan adalah serangkaian opini publik pada instagram dan facebook yang mengangkat berita Covid – 19 varian omicron. Topik ini diambil sebagai fokus analisis, karena polaritas reaksi pada masyarakat, dengan komentar yang bersifat pro maupun kontra. Data ini memberikan peluang untuk menganalisis pola sentimen yang ditemukan, sekaligus menjadi data latih yang efektif dalam mengukur kinerja model yang telah dibuat.

Meskipun pandemi covid – 19 varian omicron berhasil kita lewati, analisis terhadap komentar tetap bermanfaat dalam menilai efektivitas model yang dirancang untuk menangani dan memperbaiki model BERT. Menggunakan data ini, model BERT dapat diuji kemampuannya dalam mengklasifikasikan sentimen yang sering kali beragam dan berpotensi bias (ambigu), sehingga tetap relevan untuk analisis performa model NLP (*Natural Language Processing*). Berikut ini adalah data penelitian yang digunakan.

Tabel 1. Data Penelitian Covid – 19

No.	Data Penelitian
1.	Semua adalah Bisnis
2.	Mmg Cuma flu biasa, itu mha dr dulu
3.	Betul, virus omicron hanya bahaya bagi orang yang belum vaksin, jika yg sdh vaksin walaupun ada terhinggapi virus varian omicron mudah disembuhkan.
...	...
3520	Mudah2an sudah berakhir deh, walaupun sudah vaksin tetap saja beberapa kegiatan harus colok2 hidung yg kadang bikin agak bengek setelahnya
3521	Kita tunggu semesta yang bertindak utk org2 tidak
adil 3522	keuntungan sudah diraup, semua sudah terencana

Sebagaimana yang ditunjukkan pada Tabel 1, keseluruhan data tersebut tidak dapat langsung digunakan untuk klasifikasi. Data tersebut memerlukan proses *preprocessing* terlebih dahulu agar siap untuk dianalisis lebih lanjut. Tahapan *preprocessing*, seperti pembersihan data, penghapusan nilai yang tidak relevan, dan transformasi data, diperlukan untuk memastikan bahwa data yang diolah sesuai dengan format yang dibutuhkan oleh model. Hal ini penting untuk meningkatkan akurasi model dan memastikan bahwa data yang diinput dapat menghasilkan klasifikasi yang optimal.

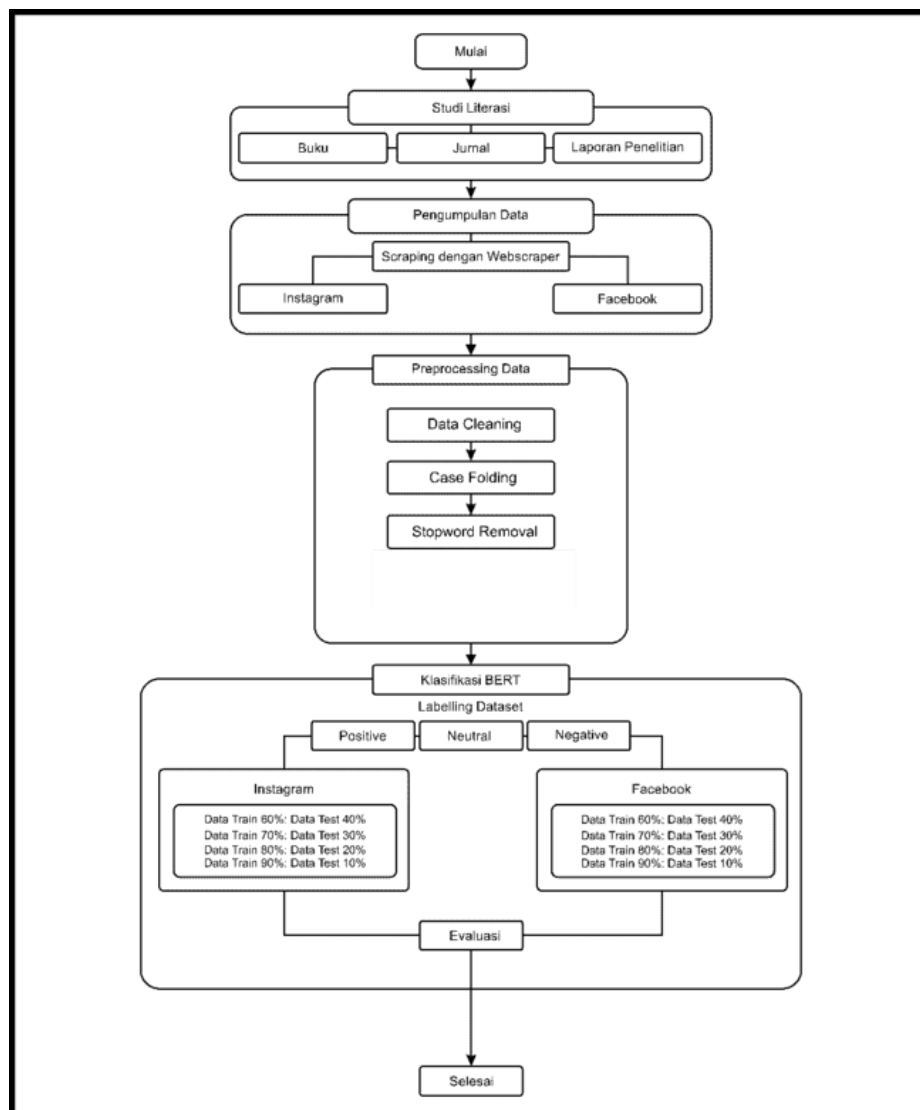
Tabel 2. Data Latih

No.	Data Penelitian	Label
1.	stop berita meresahkan jangan buat masyarakat ikut cemas	Positif
2.	dia bisa liat virus gak keliatan	Netral
3.	skill melariskan vaksin dosis adakah lepas vaksin teori konspirasi sudah macam doktrin	Negatif
4.	siapa yang panik ajg	Negatif

Proses evaluasi model menggunakan *Confusion Matrix* untuk mengidenfitikasi jumlah klasifikasi yang benar (*true positive* dan *true negative*) serta kesalahan klasifikasi (*false positive* dan *false negative*). Melalui pendekatan ini, performa model dalam mengklasifikasikan data dapat diukur secara objektif dan hasil yang diperoleh dapat memberikan gambaran nyata tentang keefektifan model BERT.

Dataset yang digunakan untuk evaluasi adalah dataset yang dilabeli secara manual, proses pelabelan manual ini sangat penting untuk memastikan bahwa data memiliki label yang akurat baik positif, negatif maupun netral. Untuk mengurangi bias penelitian yang bersifat subjektif, pelabelan secara manual dilakukan dengan metode yang sistematis dan berstandar. Proses pelabelan melibatkan beberapa evaluator independen yang diberi pedoman jelas tentang kriteria pengelompokan sentimen, sehingga memastikan konsistensi penilaian. Pendekatan ini bertujuan untuk menghasilkan label yang objektif dan valid, sehingga dapat meminimalkan potensi bias dan meningkatkan keakuratan dalam mengukur performa model yang diuji.

Alur Penelitian



Gambar 2: Alur Penelitian

Penelitian ini bertujuan untuk mengukur akurasi dari model BERT dalam mengklasifikasikan sentimen pada komentar sosial media instagram dan facebook. Terdapat serangkaian tahapan sebelum dilakukan proses klasifikasi yaitu *Data Cleaning*, *Case Folding* dan *Stopword Removal*. Pengujian metode BERT menggunakan *confusion matrix*, dengan menggunakan dataset yang telah dilabeli secara manual, untuk mendapatkan hasil yang optimal dataset dirancang dengan 4 proporsional penelitian, dengan tujuan melihat bagaimana performa model dalam mengklasifikasikan sentimen. Melalui pendekatan ini, peneliti dapat mengidentifikasi tingkat akurasi, *Precision*, recall, dan F1-score model, serta mengevaluasi seberapa baik BERT dalam menangani variasi data yang digunakan pada analisis sentimen.

Tahap Analisis Sentimen

Analisis dilakukan setelah tahapan *preprocessing*, BERT memiliki keunggulan adanya *layer*, dimana setiap lapisannya memiliki fungsinya tersendiri. *Layer* yang dibangun pada model ini meliputi *BertForSequenceClassification*, *BertModel*, *BertEmbeddings*, *BertEncoder*, *BertLayer*, *BertSelfAttention*, *BertIntermediate* dan *BertOutput*. Kompleksitas dari arsitektur yang dibangun diharapkan dapat meningkatkan performa BERT dalam mengklasifikasikan sentimen sehingga memiliki akurasi yang tinggi dalam mengklasifikasikan setiap kalimat.

Keunggulan mengklasifikasikan menggunakan model BERT adalah adanya token *unique* yang ditandai dengan [cls] (*decoder*) pada awal lapisan dan [sep] (*encoder*), token tersebut berisikan bobot representasi teks untuk melatih model BERT dalam melakukan tugas klasifikasi (Naufaldi Hafidhigbal et al., 2024). Sentimen ini berfokus pada sentimen bahasa Indonesia, sehingga seluruh data akan ditransformasikan ke bentuk dasar bahasa Indonesia, *library* yang digunakan adalah “*indobert-base-p1*” untuk proses klasifikasinya.

Sebelum diterapkan pada dataset utama, dilakukan uji coba untuk memastikan proses klasifikasi berjalan dengan baik. Uji coba ini mencakup beberapa skenario *epoch*, yang bertujuan untuk menentukan jumlah *epoch* optimal yang dapat menghasilkan model dengan akurasi terbaik. Dalam skenario ini, model dilatih pada jumlah *epoch* yang optimal berada pada *epoch* ke – 8. Pada *epoch* adalah skenario terbaik, tidak menunjukkan tanda – tanda *overfitting* atau *underfitting*, sehingga

diputuskan bahwa *epoch* 8 adalah pilihan terbaik untuk melanjutkan ke dataset utama

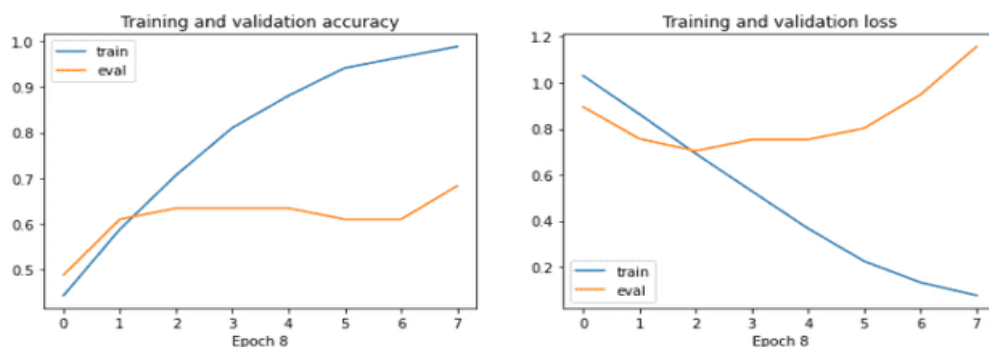
Melalui uji coba ini, model juga dapat dievaluasi apakah mengalami *overfitting* atau *underfitting* pada dataset, sehingga peneliti dapat menyesuaikan parameter lain yang diperlukan sebelum menerapkan model pada dataset yang lebih besar dan kompleks (Jason Brownlee, 2019). Proses ini sangat penting untuk memastikan bahwa model dapat bekerja secara optimal dan efisien saat diimplementasikan pada data yang sesungguhnya.

3. HASIL DAN PEMBAHASAN



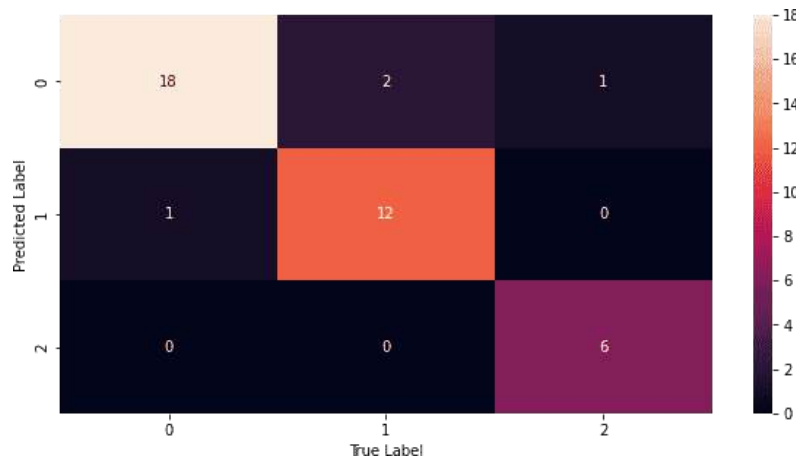
Gambar 3: Wordcloud

Gambar 3 merupakan *wordcloud* dari keseluruhan dataset cuitan terkait berita covid – 19 varian omicron melalui media sosial instagram dan facebook. Berdasarkan *wordcloud* diketahui bahwa terjadi kepanikan terhadap fenomena pandemi covid – 19 varian omicron dimasyarakat. Ditunjukkan terjadi penurunan nilai kepercayaan masyarakat terhadap penanganan yang dilakukan oleh pemerintah. Terkesan bahwa virus ini adalah rekayasa yang didalangi hanya untuk mendapatkan keuntungan bagi beberapa pihak yang tidak bertanggung jawab.



Gambar 4: Train dan Validation Loss

Gambar 4 menunjukkan hasil pelatihan model *machine learning* yang memanfaatkan metode *epoch* untuk mengukur akurasi dan *loss* pada data pelatihan dan validasi. Grafik pertama di sebelah kiri menggambarkan *Training and Validation Accuracy*, yang menunjukkan akurasi model selama proses pelatihan (*training*) dan evaluasi (*validation*). Garis biru mewakili akurasi pada data pelatihan, sedangkan garis *orange* menunjukkan akurasi pada data validasi. Dari grafik ini terlihat bahwa akurasi pelatihan terus meningkat secara signifikan seiring bertambahnya *epoch*. Namun, akurasi validasi mulai stagnan setelah *epoch* ke-3, menunjukkan bahwa setelah beberapa *epoch* tambahan, model tidak memberikan peningkatan yang berarti pada data validasi, yang mengindikasikan adanya *overfitting*. Setelah dilakukan proses *train* secara bertahap didapatkan nilai *epoch* 8 mendapatkan nilai yang paling optimal, ditunjukkan dengan nilai akurasi mencapai 90%.



Gambar 5: *Confusion Matrix*

Gambar 5 merupakan evaluasi *Confusion Matrix* data evaluasi yang telah diberi label secara manual. Hasil direpresentasikan dalam bentuk *confusion matrix* pada model yang dilatih dengan model dataset yang telah didapatkan pada sosial media instagram dan facebook. Dilihat pada Gambar 5 terdapat komentar jumlah prediksi yang benar positif, benar negatif dan prediksi yang salah dimana model mengklasifikasikan *instance* negatif sebagai positif dan sebaliknya. Sehingga didapatkan performa sebagai berikut:

→ Accuracy score is 90%.

Classification Report:					
	precision	recall	f1-score	support	
0	0.95	0.86	0.90	21	
1	0.86	0.92	0.89	13	
2	0.86	1.00	0.92	6	
accuracy			0.90	40	
macro avg	0.89	0.93	0.90	40	
weighted avg	0.90	0.90	0.90	40	

Gambar 6: Evaluasi Model

Gambar 6 menunjukkan model yang dihasilkan mencapai akurasi sebesar 90%, yang menunjukkan kemampuan model memprediksi benar 90% dari keseluruhan data uji. *Precision* rata-rata model sebesar 0.90 menunjukkan bahwa sebagian besar prediksi yang dihasilkan sesuai dengan kelas sebenarnya, artinya model memiliki kemampuan yang baik dalam mengurangi *false positive*. Sebaliknya, *recall* rata-rata juga mencapai 0.90, yang menandakan kemampuan model dalam menangkap hampir semua instance yang relevan dari setiap kelas, atau dengan kata lain, model mampu meminimalkan *false negative*.

F1 – score rata – rata sebesar 0.90 menunjukkan keseimbangan yang kuat antara *Precision* dan *recall*, yang berarti model tidak hanya cermat dalam memprediksi kelas dengan benar tetapi juga efektif dalam menangkap seluruh *instance* relevan. Kinerja model pada setiap kelas cukup konsisten, dengan F1-score tertinggi pada kelas 2 sebesar 0.92, yang berarti bahwa model sangat baik dalam mengidentifikasi *instance* pada kelas tersebut. Sementara itu, untuk kelas 0 dan 1, F1-score masing – masing berada di angka 0.90 dan 0.89, menandakan bahwa performa model di kedua kelas ini masih sangat baik dan seimbang antara *Precision* dan *recall*.

Secara keseluruhan, meskipun terdapat sedikit perbedaan dalam *Precision* dan *recall* di setiap kelas, performa model tetap konsisten dan kuat di seluruh metrik evaluasi, terutama dengan F1-score rata – rata yang menunjukkan keseimbangan antara kemampuan model dalam membuat prediksi yang benar dan menangkap semua *instance* yang relevan.

4. SIMPULAN

Kesimpulan yang didapatkan dari penelitian ini adalah model BERT, dapat digunakan untuk mengklasifikasikan sentimen dengan hasil yang sangat baik. Model BERT yang dibangun dilatih mendapatkan nilai efisien pada rentang 3 – 5 *epoch* tidak ditemukan indikasi *overfitting* dan *underfitting*, meskipun temuan pada penelitian akurasi paling tinggi berada di *epoch* 8 dengan akurasi sebesar 90%. Mayoritas komentar terkait pemberitaan Covid – 19 menunjukkan sentimen negatif, ditengarai karena paniknya masyarakat dalam menghadapi pandemi, yang menyerang banyak aspek pendidikan, ekonomi dsb. Penelitian selanjutnya diharapkan dapat mengkombinasikan *fine – tuning* dan kombinasi data lebih beragam dengan *library* yang berbeda untuk membuktikan BERT dalam mengklasifikasikan sentimen.

5. DAFTAR PUSTAKA

- Chinnasamy, P., Suresh, V., Ramprathap, K., Jebamani, B. J. A., Srinivas Rao, K., & Shiva Kranthi, M. (2022). COVID-19 vaccine sentiment analysis using public opinions on Twitter. *Materials Today: Proceedings*, 64, 448–451. <https://doi.org/10.1016/j.matpr.2022.04.809>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- Jason Brownlee. (2019, August 12). *Overfitting and Underfitting With Machine Learning Algorithms*.
- Jalaj Thanaki. (2017). *Python Natural Language Processing: Advanced machine learning and deep learning techniques for natural language processing* (1st ed.).
- Kurniasari, I., Kusrini, K., & Fatta, H. Al. (2021). Analysis of Public Opinion Sentiment on Instagram regarding Covid-19 with SVM. *JTECS : Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem Dan Komputer*, 1(1), 67. <https://doi.org/10.32503/jtecs.v1i1.1416>
- Lin, S.-Y., Kung, Y.-C., & Leu, F.-Y. (2022). Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment analysis. *Information Processing & Management*, 59(2), 102872. <https://doi.org/10.1016/j.ipm.2022.102872>
- Naufaldi Hafidhigbal, Siti Ummi Masruroh, & Nenny Anggraini. (2024). *Implementasi model indobert untuk klasifikasi jenis tindak pidana korupsi berbasis web application*. UIN Syarif Hidayatullah Jakarta.
- Nugroho, K. S., Sukmadewa, A. Y., Wuswilahaken DW, H., Bachtiar, F. A., &

- Yudistira, N. (2021). BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews. *6th International Conference on Sustainable Information Engineering and Technology 2021*, 258–264. <https://doi.org/10.1145/3479645.3479679>
- Prof. Dr. Sugiyono. (2018). *Metode penelitian kuantitatif, kualitatif dan kombinasi (mixed methods)* (2nd ed., Vol. 10). Bandung : Alfabeta, 2018.
- Ray, B., Garain, A., & Sarkar, R. (2021). An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews. *Applied Soft Computing*, 98, 106935. <https://doi.org/10.1016/j.asoc.2020.106935>
- Samsir, S., Ambiyar, A., Verawardina, U., Edi, F., & Watrianthos, R. (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 157. <https://doi.org/10.30865/mib.v5i1.2580>
- Selvakumar, B., & Lakshmanan, B. (2022). Sentimental analysis on user's reviews using BERT. *Materials Today: Proceedings*, 62, 4931–4935. <https://doi.org/10.1016/j.matpr.2022.03.678>
- Singh, M., Jakhar, A. K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Social Network Analysis and Mining*, 11(1), 33. <https://doi.org/10.1007/s13278-021-00737-z>
- Sudharsan Ravichandiran. (2021). *Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT* (1st ed.).
- Syukur, M., Salam, M. N., & Junaidi, M. I. (2021). Dampak Pandemi Covid-19 terhadap Perekonomian Indonesia: Analisis terhadap Sektor Domestik dan Stabilitas Inflasi. *TRILOGI: Jurnal Ilmu Teknologi, Kesehatan, Dan Humaniora*, 2(3), 382–388. <https://doi.org/10.33650/trilogi.v2i3.3082>
- Tabinda Kokab, S., Asghar, S., & Naz, S. (2022). Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, 14, 100157. <https://doi.org/10.1016/j.array.2022.100157>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*.
- Vidya Chandradev, I Made Agus Dwi Suarjaya, & I Putu Agung Bayupati. (2023). Analisis Sentimen Review Hotel Menggunakan Metode Deep Learning BERT. *Jurnal Buana Informatika*, 14(02), 107–116. <https://doi.org/10.24002/jbi.v14i02.7244>
- Zhang, Y., Chen, K., Weng, Y., Chen, Z., Zhang, J., & Hubbard, R. (2022). An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US. *Expert Systems with Applications*, 198, 116882. <https://doi.org/10.1016/j.eswa.2022.116882>