

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Tinjauan Pustaka**

##### **2.1.1. Analisis Sentimen**

Menurut (Purnamasari dkk., 2023), “Analisis sentimen merupakan metode dalam data tekstual yang memanfaatkan Pemrosesan Bahasa Alami (NLP) dan Pembelajaran Mesin (ML) untuk mengevaluasi teks secara otomatis dan mengidentifikasi emosi penulis, seperti positif, negatif, atau netral. Sebagai bagian dari NLP, analisis sentimen berkembang seiring interaksi manusia dengan komputer, ekstraksi informasi, dan interpretasi emosi dari data media sosial daring yang terus bertambah. Tujuannya adalah untuk mengenali kata atau frasa yang mencerminkan sikap dalam teks. Analisis sentimen umumnya digunakan untuk mengekstrak berbagai karakteristik data teks, baik dari data terstruktur maupun tidak terstruktur, dan menganalisisnya untuk memahami pikiran, opini, dan perasaan yang terkandung di dalamnya.”

##### **2.1.2. Sentimen**

Menurut (Pristika & Rozi, 2024), “Sentimen adalah pendapat atau pandangan yang didasarkan pada perasaan yang berlebih-lebihan terhadap sesuatu. Sentimen komentar netizen mengacu pada pendapat, perasaan, atau pandangan yang dinyatakan oleh pengguna media sosial (netizen) terhadap suatu topik, peristiwa, atau konten tertentu. Sentimen ini bisa berupa positif, negatif, atau netral, tergantung pada bagaimana netizen merespons suatu hal. Sentimen dalam media sosial berupa komentar-komentar yang dilontarkan netizen terhadap suatu postingan di salah satu media sosial. Sentimen netral berisi opini biasa. Sentimen

positif berisi opini dan komentar yang mendukung, sebaliknya untuk sentimen negatif berisi opini dan komentar yang tidak mendukung.”




### 2.1.3. UML (*Unified Modeling Language*)




Menurut (Limantoro & Kristiadi, 2021) “UML merupakan kepanjangan dari *Unified Modeling Language* yang memiliki arti bahasa pemodelan standar. UML dapat digunakan untuk mendeskripsikan sesuatu. UML terdapat simbol penulisan yang membantu penulis untuk menjelaskan mengenai hasil yang diekspektasikan. Simbol-simbol yang dimaksud disebut dengan notasi dan proses.”

### 2.1.4. *Use Case Diagram*

Menurut (Noviantoro dkk., 2022), “*Use case diagram* merupakan model untuk perilaku (*behaviour*) dari sistem informasi yang akan dibangun. *Use case* digunakan untuk mengetahui fungsi apa yang ada dalam sistem informasi dan siapa yang memiliki kewenangan untuk menggunakan fungsi tersebut. Simbol-simbol yang digunakan dalam *Use Case Diagram* terlihat pada tabel berikut:”

Tabel 2. 1 Tabel *Use Case Diagram*





No	Notasi	Keterangan	Simbol
1	<i>Actor</i>	Menggambarkan orang, sistem atau external entitas yang menyediakan atau menerima informasi dari sistem	
2	<i>Use Case</i>	Menerangkan apa yang dikerjakan sistem, bukan bagaimana sistem mengerjakannya.	
3	<i>Association</i>	Menggambarkan bagaimana actor terlibat dalam <i>use case</i> .	




4	<i>Generalization</i>	Dibuat Ketika ada sebuah keadaan yang lain atau perlakuan khusus	
5	<i>Include</i>	Relasi use case dimana proses bersangkutan akan dijalankan ke proses yang dituju	
6	<i>Extend</i>	Relasi use case tambahan ke sebuah use case	

### 2.1.5. Activity Diagram

Menurut (Noviantoro dkk., 2022), “Diagram yang menggambarkan atau memodelkan berbagai aktivitas aliran dalam sistem yang sedang dikembangkan, serta bagaimana setiap aliran berawal dari keputusan yang mungkin terjadi, dan bagaimana setiap aktivitas berakhir. Simbol-simbol yang digunakan dalam *Activity Diagram* terlihat pada tabel berikut:”

Tabel 2. 2 Tabel *Activity Diagram*

Gambar	Keterangan
	<i>Start Point</i> , diletakkan pada pojok kiri atas dan merupakan awal aktivitas
	<i>End Point</i> , akhir aktivitas
	<i>Activities</i> , menggambarkan suatu proses atau kegiatan bisnis
	<i>Fork</i> atau percabangan, digunakan untuk menunjukkan kegiatan yang secara paralel atau untuk menggabungkan dua kegiatan paralel menjadi satu

	<i>Join</i> penggabungan atau <i>rake</i> , digunakan untuk menunjukkan adanya dekomposisi
	<i>Decision Points</i> , menggambarkan pilihan untuk pengambilan keputusan, <i>True</i> dan <i>False</i>
	<i>Swimlane</i> , pembagian <i>activity diagram</i> untuk menunjukkan siapa melakukan apa.

### 2.1.6. Python

Menurut (Retno Tri Vулandari, 2017) “*Python* berkembang cukup pesat karena memiliki berbagai kelebihan seperti aspek *readability*, multifungsi, interoperabilitas dan juga dukungan komunitas yang memadai.

#### 1. Keterbacaan atau *Readability*

Pengembang Bahasa pemrograman *Python* menekankan aspek *readability* pada sintaks serta kode-kode dasar programnya, itu berarti bahwa *Python* mudah dibaca dan dipahami. *Python* memiliki sintaks yang sederhana, sehingga mudah ditulis, mudah diingat dan juga digunakan ulang.

#### 2. Efisien

Memiliki aspek keterbacaan yang baik dan programmer *friendly*, memberikan efisiensi bagi programmer dalam membuat program. *Python* juga memiliki *library* yang lengkap sehingga untuk mendapatkan program yang sama, *Python* Code akan lebih simpel dibandingkan dengan kode yang ditulis dengan bahasa pemrograman lainnya seperti misalnya Java, C, C# maupun C++.

#### 3. Multifungsi

*Python* adalah bahasa pemrograman yang sangat fleksibel. Berbagai jenis aplikasi dapat dikembangkan dengan *Python*, seperti *website*, aplikasi jaringan,

aplikasi bidang robotika, dan bahkan aplikasi kecerdasan buatan. Kekuatan *Python* terletak pada dukungan *library* yang luas, menyediakan modul siap pakai untuk berbagai keperluan pengembangan.

#### 4. Interoperabilitas

Keunggulan *Python* terletak pada interoperabilitasnya yang tinggi. *Python* dapat berinteraksi dengan bahasa pemrograman lain, seperti Java, C, atau C++, begitu pula sebaliknya. Kemampuan ini memungkinkan integrasi kode antar bahasa pemrograman yang berbeda.

#### 5. Dukungan Komunitas

*Python*, sebagai program *open source*, didukung oleh komunitas yang sangat solid. Kekuatan komunitas ini menjadi pendorong utama perkembangan *Python* hingga saat ini. Interaksi yang baik dalam komunitas memudahkan pengguna untuk saling bertanya dan berbagi pengetahuan tentang pemrograman *Python*.”

### 2.1.7. Text Mining

Menurut (Djamaludin dkk., 2022), “*Text mining* sebagai bagian dari data mining, yang berfokus pada penggalian informasi dari teks, umumnya bersumber dari dokumen. Tujuannya adalah mengidentifikasi kata kunci yang merepresentasikan isi dokumen, sehingga dapat dianalisis keterkaitan antar dokumen. Selain itu, *text mining* memanfaatkan komputer untuk memproses informasi yang sudah ada secara eksplisit, dengan tujuan menghasilkan informasi baru.”

*Text Mining* adalah salah satu bidang khusus dalam *data mining* yang memiliki makna menambang data berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat

mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen. Disisi lain *teks mining* dapat bekerja dalam komputer dengan tujuan mengolah informasi lama secara eksplisit sehingga menghasilkan temuan informasi baru.

### **2.1.8. Data Mining**

Menurut (Agustin dan Budiman., 2024), “Secara sederhana, *data mining* merupakan proses ekstraksi untuk mendapatkan informasi penting yang tersirat dan belum diketahui. *Data mining* dimulai dengan proses pemilihan, pembersihan, *pre-processing*, dan transformasi data. Semua ini merupakan bagian dari proses yang ada dalam KDD.”

Menurut (Teguh Wahyono, 2018), “Proses tahapan data mining memiliki tahap sebagai berikut:

1. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*)
2. Integrasi data (penggabungan data dari beberapa sumber)
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)
4. Aplikasi teknik Data Mining, proses ekstraksi pola dari data yang ada
5. Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan)
6. Presentasi pengetahuan (dengan teknik visualisasi)

Dalam proses pencarian pengetahuan, tahap ini melibatkan pengecekan apakah pola atau informasi baru yang ditemukan sesuai dengan fakta atau hipotesis yang telah ada sebelumnya.”

### 2.1.9. CRISP-DM (*Cross-Industry Standard Process for Data Mining*)

Menurut (Ginantra dkk., 2021), “Dalam CRISP-DM, suatu proyek data mining mempunyai siklus hidup yang dibagi ke dalam enam tahapan yang berurutan dan bersifat adaptif. Keluaran dari tahapan sebelumnya akan sangat berpengaruh bagi tahapan berikutnya.

Adapun penjelasan mengenai tahapan dalam CRISP-DM, yaitu:

1. *Business Understanding*

Tahapan pertama dalam CRISP-DM ini merupakan tahapan yang cukup penting, karena di dalam tahapan ini memerlukan pengetahuan dari objek bisnis, proses membangun dan mendapatkan data serta mencocokkan tujuan permodelan dengan tujuan bisnis.

2. *Data Understanding*

Tahap ini dilaksanakan untuk memberikan fondasi analitik pada suatu penelitian dengan menyusun ringkasan atau *summary* dan mengidentifikasi potensi berbagai masalah yang ada di dalam data. Tahap data understanding ini juga harus dilakukan dengan teliti dan hati-hati, karena dikhawatirkan jika terdapat kendala pada tahapan ini yang belum terjawab, maka dapat mengganggu tahap modeling. *Summary* dari data berguna untuk mengkonfirmasi apakah data sudah terdistribusi seperti yang diharapkan.

3. *Data Preparation*

Tahap *data preparation* ini memerlukan pemikiran matang dan upaya tinggi untuk memperbaiki masalah dalam data dan dibuat *variable derived* serta memastikan apakah data sudah tepat untuk algoritma yang digunakan. Tahap ini sering mengalami peninjauan ulang ketika menemukan kendala pada pembangunan

model, sehingga dilaksanakan iterasi hingga menemukan hal yang sesuai dengan data yang dimaksud.

#### 4. *Modelling*

Pada tahap ini, dilakukan metode statistika dan *machine learning* untuk menentukan teknik, alat bantu serta algoritma data mining yang akan diterapkan. Kemudian langkah selanjutnya adalah menerapkan teknik dan algoritma tersebut pada data dengan alat bantu. Yang perlu digaris bawahi di sini, beberapa teknik memungkinkan untuk digunakan pada data mining yang memiliki permasalahan yang sama. Jika diperlukan penyesuaian data terhadap metode data mining, kita dapat kembali ke tahapan *data preparation*.

#### 5. *Evaluation*

Tahap *evaluation* ini merupakan tahap evaluasi dengan melaksanakan interpretasi terhadap *output* dari *data mining* yang dihasilkan dalam tahapan sebelumnya. Evaluasi di sini bertujuan agar model yang sudah ditentukan dapat sesuai dengan tujuan yang ingin dipenuhi pada fase pertama.

#### 6. *Deployment*

Tahap *deployment* atau rencana penggunaan model merupakan tahapan yang penting dalam proses CRISP-DM. Perencanaan untuk tahap *deployment* sudah dimulai sejak proses *Business Understanding* (tahapan pertama) dilakukan. Tahap *deployment* ini tidak hanya menghasilkan suatu model, tapi juga mengonversi skor putusan serta menggabungkan keputusan dalam sistem operasional. Namun terbentuknya model tidak menandakan proyek selesai begitu saja. Model dibangun dari data yang mewakili pada jangka waktu tertentu, sehingga ada kemungkinan

perubahan karakteristik data seiring perubahan waktu. Model pun harus dipantau serta dapat diganti dengan model lain yang sudah diperbaiki atau dimodifikasi.”

#### **2.1.10. *Natural Language Processing (NLP)***

Menurut (Purnamasari dkk., 2023), “*Natural Language Processing (NLP)*, atau pemrosesan bahasa alami, bertujuan untuk memungkinkan komputer menganalisis, memahami, dan menghasilkan bahasa manusia. Tantangan utama dalam NLP adalah mengembangkan kemampuan komputer untuk belajar dan menggunakan bahasa sebagaimana manusia.

NLP bertujuan untuk memperoleh pemahaman mendalam dari data atau konten yang tidak terstruktur yang diekspresikan dalam bahasa sehari-hari, contohnya bahasa Inggris, bahasa Indonesia, dan bahasa-bahasa lainnya. Perbedaan struktur pada tiap bahasa, termasuk tata bahasa dan sintaksisnya, memerlukan teknik NLP khusus. Dengan teknik ini, komputer dapat membaca teks, mendengarkan ucapan, menafsirkan makna, mengukur sentimen atau melakukan *opinion mining*, dan akhirnya menentukan bagian-bagian penting dalam sistem.

NLP diterapkan pada data tekstual, mulai dari input pengguna di platform internet hingga analisis berbagai dokumen. Proses ini melibatkan ekstraksi informasi penting untuk mendukung pengambilan keputusan yang lebih baik. NLP diterapkan dalam berbagai situasi untuk menyelesaikan beragam jenis masalah, antara lain pencarian serangkaian karakter, pengenalan entitas, analisis sentimen, pengembangan mesin pencari, dan pengelolaan pesan elektronik.”

#### **2.1.11. Model**

Menurut (Devlin dkk., 2019) “Model dalam *machine learning* adalah representasi matematis dari pola atau hubungan yang dipelajari dari data. Model ini

digunakan untuk membuat prediksi atau keputusan pada data baru yang belum pernah dilihat sebelumnya. Dalam konteks BERT, model adalah representasi kompleks dari pemahaman bahasa yang dipelajari dari jutaan teks.”

#### **2.1.12. Deep Learning**

Menurut (Purnamasari dkk., 2023), “*Deep Learning* telah mencapai kemajuan signifikan dalam pengenalan ucapan dan gambar. Pemanfaatan jaringan saraf dalam *deep learning* juga merevolusi bidang *Natural Language Processing* (NLP). Dengan banyaknya data dan opini yang dihasilkan dan disebarluaskan setiap hari melalui internet, analisis sentimen menjadi salah satu area penelitian yang sangat berkembang dalam NLP.

*Deep Learning* mengandalkan jaringan saraf yang besar dengan banyak lapisan unit pemrosesan. Dengan memanfaatkan peningkatan daya komputasi dan penyempurnaan teknik pelatihan, *Deep Learning* mampu mengidentifikasi pola-pola yang rumit dalam data yang sangat besar. Contoh penerapannya adalah pengenalan gambar dan suara.”

#### **2.1.13. Pre-processing**

Menurut (Purnamasari dkk., 2023), “Tahap ini merupakan tahap yang juga terpenting untuk memastikan bahwa data yang digunakan pada model *machine learning* merupakan data yang tepat.

Tahapan *pre-processing* terdiri dari:

1. *Data selection*, proses pemilihan fitur data yang akan digunakan.
2. *Cleaning*, proses dimana karakter dan tanda baca yang tidak diperlukan dihilangkan dari teks.
3. *Case folding*, proses mengubah seluruh huruf menjadi huruf kecil.

4. *Stopword removal*, proses menghilangkan kata yang sering muncul pada suatu teks dan dianggap tidak memiliki makna.
5. *Stemming*, proses mengubah kata menjadi bentuk kata dasarnya dengan menghilangkan kata imbuhan.
6. *Slang Word*, proses mengubah kata tidak baku pada Bahasa Indonesia menjadi kata dasar.”

#### **2.1.14. BERT (*Bidirectional Encoder Representations from Transformers*)**

Menurut (Kurniawan dkk., 2022), “BERT merupakan model *deep learning* yang telah meraih kesuksesan dalam berbagai tugas NLP. Dengan arsitektur enam lapisan *Transformer* bertumpuk pada setiap *encoder* dan *decoder*, proses pelatihan BERT menjadi kompleks, konfigurasinya rumit, membutuhkan waktu lama, dan berbiaya besar. Namun, model BERT yang sudah terlatih (*pre-trained*) oleh Google tersedia sebagai *open source* sehingga dapat dimanfaatkan tanpa harus membangun model dari awal. Proses BERT dimulai dengan representasi *embedding* kata dari lapisan *embedding*. Tiap lapisan selanjutnya melakukan serangkaian perhitungan *multi-headed attention* pada representasi kata dari lapisan sebelumnya untuk menghasilkan representasi perantara baru.”

#### **2.1.15. Firebase**

Menurut (Panjaitan & Pakpahan, 2021), “*Firebase*, sebuah *Backend as a Service* (BaaS) yang dikembangkan oleh Google, memfasilitasi pengembangan aplikasi bagi para pengembang perangkat lunak. Layanan-layanan *Firebase* yang dimanfaatkan dalam sistem yang dirancang antara lain otentikasi (*authentication*) yang memungkinkan pengguna untuk login dan terhubung langsung dengan server *Firebase*, *Firebase Storage* yang memfasilitasi pengembang untuk mengunggah

dan mengunduh berkas, serta *Firebase Cloud Function* sebagai kerangka kerja tanpa server yang memungkinkan kode dieksekusi di *back-end* secara otomatis sebagai respons terhadap pemicu fitur maupun permintaan HTTPS. *Firebase* menyediakan pustaka yang lengkap dan kompatibel dengan berbagai kerangka kerja seperti NodeJS, Java, Javascript, ReactJS, dan lain-lain.”

### 2.1.16. Confusion Matrix

Menurut (Suryati dkk., 2023), “*Confusion matrix* salah satu alat mengolah data untuk membandingkan hasil prediksi dengan label sesungguhnya. Evaluasi dengan *confusion matrix* menghasilkan nilai akurasi, presisi, *re-call* dan *F1-Score*. Pada evaluasi klasifikasi terdapat empat kemungkinan yang bisa terjadi dari hasil klasifikasi suatu data. Jika data positif dan diprediksi positif maka dihitung sebagai *true positive* dan jika data positif diprediksi negatif maka akan dihitung sebagai *false negative*. Pada data negatif jika diprediksi negatif dihitung sebagai *true negative* dan jika diprediksi positif maka akan dihitung sebagai *false positive*.

Tabel 2. 3 Tabel *Confusion Matrix*

Actual Label	Predicted Label	
	Positive (+)	Negative (-)
Positive (+)	True Positives (TP)	False Negatives (FN)
Negative (-)	False Positives (FP)	True Negatives (TN)

- a. Akurasi adalah ukuran kedekatan antara hasil prediksi dengan nilai yang sebenarnya. Sistem yang memiliki tingkat akurasi tinggi dianggap lebih baik dalam melakukan prediksi. Rumus perhitungan akurasi adalah sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- b. *Recall* adalah salah satu perhitungan keakuratan prediksi yang digunakan sebagai ukuran tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. Rumus perhitungan *recall* adalah sebagai berikut:

$$Recall = \frac{TP}{TP + FN}$$

- c. *Precision* adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban system. Rumus perhitungan *Precision* adalah sebagai berikut”:

$$Precision = \frac{TP}{TP + FP}$$

- d. Menurut (Duei Putri dkk., 2022), “*F1-Score* adalah perbandingan antara *precision* dan *recall* yang digunakan untuk menentukan akurasi data yang diuji. Rumus perhitungan *F1-Score* adalah sebagai berikut”:

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

### 2.1.17. Jupyter Notebook

Dikutip dari (Jupyter, 2024) “*Jupyter Notebook* adalah aplikasi *web* asli untuk membuat dan berbagi dokumen komputasi. Ini menawarkan pengalaman yang sederhana, efisien, dan berpusat pada dokumen.”

### 2.1.18. Streamlit

Dikutip dari (Streamlit Docs, 2024) “*Streamlit* adalah *library/pustaka Python* yang sumbernya itu terbuka (*open source*) yang memudahkan pembuatan dan berbagi, mengkostumisasi aplikasi *web* untuk *machine learning* dan *data science*.”

Hanya dalam beberapa menit, dapat membuat dan menerapkan aplikasi data yang canggih.”

### 2.1.19. Google Colab

Dikutip dari (*Research Google, 2024*) “*Google Colab* adalah layanan *Notebook Jupyter* yang dihosting yang tidak memerlukan penyiapan untuk digunakan dan menyediakan akses gratis ke sumber daya komputasi, termasuk GPU dan TPU. Colab sangat cocok untuk pembelajaran mesin, ilmu data, dan pendidikan.”

## 2.2. Kajian Penelitian Terdahulu

No	Judul	Penulis	Tahun terbit	Klasifikasi
1	Analisis Sentimen Pada Kasus Positif Covid-19 Berdasarkan Pemberitaan Media Di Indonesia Menggunakan Indobert	Widiya Nurfitri dan Andry Chowanda	2024	Jurnal Ilmiah Komputer, Vol. 20 No. 1, Februari 2024, e-ISSN: 2685-0877, p-ISSN: 0216-3284, Sinta 4
<p>Metode: BERT (<i>Bidirectional Encoder Representations from Transformers</i>)</p> <p>Hasil: Ukuran batch 16 menunjukkan kinerja model yang stabil, dengan skor F1 antara 80,10% hingga 80,53%. Sementara itu, ukuran batch 32 menghasilkan variasi skor F1, akurasi, dan <i>loss</i>. Meskipun terkadang</p>				

memberikan hasil yang lebih baik, ukuran batch 16 juga mampu menghasilkan kinerja yang unggul dalam beberapa situasi. Peningkatan jumlah epoch tidak selalu berbanding lurus dengan peningkatan kinerja. Meskipun skor F1 dan akurasi meningkat pada beberapa kasus (misalnya, dari epoch 5 ke 10), penurunan juga terjadi pada kasus lain (misalnya, dari epoch 10 ke 15). Epoch 10 pada ukuran batch 16 menghasilkan skor F1 dan akurasi yang tinggi, sementara epoch 5 pada ukuran batch 32 menunjukkan hasil yang baik. Secara keseluruhan, model mencapai kinerja yang baik dengan skor F1 dan akurasi di atas 0,80 dan 0,81. *Loss* cenderung meningkat seiring bertambahnya epoch, namun perlu diingat bahwa *loss* bukanlah satu-satunya tolok ukur kinerja model.

2	Analisis Sentimen Terhadap Game Genshin Impact Menggunakan BERT	Ryo Kusnadi, dkk.	2021	RABIT, Vol. 6 No. 2, Juli 2021, ISSN CETAK: 2477-2062, ISSN ONLINE: 2502-891X, Sinta 4
---	---	-------------------	------	--

Metode: BERT (*Bidirectional Encoder Representations from Transformers*)

Hasil: Evaluasi performa model analisis sentimen berbasis BERT dilakukan menggunakan 12.000 data. Data ini diklasifikasikan berdasarkan kategori (positif, negatif, dan netral) dan sentimennya. Hasil pengujian menunjukkan bahwa model lebih baik dalam mengidentifikasi sentimen positif, dengan nilai

<p>presisi tertinggi 0,86%, recall 0,78%, dan skor F1 0,82%. Hal ini menandakan kemampuan model dalam mengklasifikasikan sentimen secara akurat.</p>				
3	<p>Sentiment Analysis of Indonesian Youtube Reviews About Lesbian, Guy, Bisexual and Transgender (LGBT) using IndoBERT Fine Tuning</p>	<p>Teddy Oswari, dkk.</p>	2024	<p>LONTAR KOMPUTER, Vol. 15 No. 1, April 2024, E-ISSN: 2541-5832, P-ISSN: 2088-1541, Sinta 2</p>
<p>Metode: BERT (<i>Bidirectional Encoder Representations from Transformers</i>)</p> <p>Hasil: Penelitian ini mengeksplorasi analisis sentimen komentar YouTube tentang LGBT menggunakan metode IndoBERT dan memperoleh hasil yang positif. Terungkap bahwa pandangan mayoritas masyarakat Indonesia cenderung tidak mendukung LGBT. Model IndoBERT mampu mengklasifikasikan komentar pro dan kontra dengan akurasi 74%, serta menghasilkan presisi 86%, recall 78%, dan skor F1 82% untuk sentimen positif.</p>				
4	<p>Analisis Sentimen Review Hotel Menggunakan</p>	<p>Vidya Chandradev, dkk.</p>	2023	<p>Jurnal Buana Informatika, Vol. 14 No. 2, November</p>

	Metode Deep Learning BERT			2023, ISSN: 2087-2534, Sinta 3
<p>Metode: BERT (<i>Bidirectional Encoder Representations from Transformers</i>)</p> <p>Hasil: Model BERT yang telah melalui tahap <i>pre-trained</i> dan <i>fine-tuning</i> menunjukkan hasil analisis sentimen yang akurat. Evaluasi performa model <i>fine-tuning</i> SmallBERT yang dilatih dengan 515 ribu data ulasan hotel selama 5 <i>epoch</i> menunjukkan hasil yang memuaskan. Saat diuji menggunakan data berlabel manual, SmallBERT mencapai akurasi 91,40%, presisi 90,51%, <i>recall</i> 90,51%, dan skor F1 90,51%.</p>				
5	Implementasi Fine-Tuning BERT untuk Analisis Sentimen terhadap Review Aplikasi PUBG Mobile di Google Play Store	Alex Sander Prasetya Braja dan Achmad Kodar	2022	JIMP, Vol. 7 No. 3, Desember 2022, E-ISSN: 2503-1945, P-ISSN: 2502-5716, Sinta 4
<p>Metode: BERT (<i>Bidirectional Encoder Representations from Transformers</i>)</p> <p>Hasil: Model BERT yang telah di-<i>pre-train</i> dapat ditingkatkan kemampuannya melalui <i>fine-tuning</i> untuk <i>transfer learning</i>, sehingga memberikan pemahaman yang lebih mendalam. Metode pelabelan data mempengaruhi hasil akhir. Pelabelan berbasis TextBlob menghasilkan akurasi yang signifikan lebih tinggi daripada pelabelan berbasis skor. Model IndoBERT BASE efektif dalam memprediksi sentimen. Performa terbaik dicapai oleh model IndoBERT BASE</p>				

dengan akurasi 94%, menggunakan *hyperparameter learning rate* 0,00002, *batch size* 32, 5 *epoch*, dan waktu pelatihan 12 menit.

