

Analisis Perbandingan Model CNN dan IndoBERT Dalam Sentimen Berita Politik Indonesia

Chandra Ramadhan^{1*}, Vihi Atina², Hanifah Permatasari³

¹Sistem Informasi/Fakultas Ilmu
Komputer
Universitas Duta Bangsa
^{1*}210101011@mhs.udb.ac.id

² Teknologi Rekayasa Perangkat
Lunak/Fakultas Ilmu Komputer
Universitas Duta Bangsa
²vihi_atina@udb.ac.id

³ Sistem Informasi/Fakultas Ilmu
Komputer
Universitas Duta Bangsa
³hanifah_permatasari@udb.ac.id

Abstrak— Penelitian ini membandingkan kinerja dua model *deep learning*, *Convolutional Neural Network (CNN)* dan *IndoBERT*, dalam mengklasifikasikan sentimen berita politik berbahasa Indonesia. Data berasal dari portal berita *online* dengan topik politik seperti kebijakan publik, kinerja menteri, dan konflik legislatif. Label sentimen ditentukan dengan pendekatan berbasis leksikon (positif, negatif, netral). Proses pemodelan meliputi prapemrosesan teks dan pelatihan model CNN menggunakan *embedding layer* serta *Conv1D*, sementara *IndoBERT* menggunakan arsitektur *transformer* yang telah dilatih sebelumnya. Evaluasi menggunakan akurasi, presisi, *recall*, dan *f1-score* menunjukkan *IndoBERT* unggul dengan akurasi 92,93%, terutama pada klasifikasi kelas netral. CNN memiliki akurasi 89,13%, namun kesulitan dalam mengklasifikasikan sentimen netral. Hasil ini menunjukkan keunggulan *IndoBERT* dalam klasifikasi sentimen berita politik dan potensi model berbasis *transformer* dalam pemrosesan bahasa alami Indonesia (*Indonesian Natural Language Processing*).
Kata kunci— CNN, *IndoBERT*, Indonesian NLP, Berita Politik, Analisis Sentimen.

Abstract— This study compares the performance of two deep learning models, *Convolutional Neural Network (CNN)* and *IndoBERT*, in classifying the sentiment of Indonesian political news. The data is sourced from online news portals, covering political topics such as public policy, ministerial performance, and legislative conflicts. Sentiment labels are determined using a lexicon-based approach (positive, negative, neutral). The modeling process includes text preprocessing, with the CNN model trained using an embedding layer and *Conv1D*, while *IndoBERT* employs a pre-trained transformer architecture. Evaluation metrics such as accuracy, precision, recall, and *f1-score* reveal that *IndoBERT* outperforms CNN, achieving a validation accuracy of 92.93%, particularly excelling in the neutral sentiment classification. CNN achieves 89.13% accuracy but struggles with accurately classifying neutral sentiment. These results demonstrate the superiority of *IndoBERT* for political news sentiment classification and highlight the potential of transformer-based models in Indonesian natural language processing.
Keywords— CNN, *IndoBERT*, Indonesian NLP, Political News, Sentiment Analysis.

I. PENDAHULUAN

Cara masyarakat sekarang memperoleh dan merespons informasi, terutama berita politik, telah berubah secara signifikan akibat kemajuan teknologi informasi dan komunikasi [1]. Opini publik tentang berbagai topik nasional, termasuk kebijakan pemerintah, efektivitas lembaga negara, dan laporan kasus korupsi, kini sebagian besar terbentuk dan didapat dari media seperti portal berita *online* [2]. Kecepatan penyebaran informasi menimbulkan kesulitan khusus dalam memahami opini publik secara sistematis [3].

Salah satu metode berguna untuk mengukur tren opini publik tentang konten berita adalah analisis sentimen berbasis teks [4]. Penelitian ini dapat memberikan gambaran umum tentang bagaimana publik memandang suatu topik dengan mengelompokkan pesan ke dalam kategori positif, negatif, atau netral.

Pada penelitian sebelumnya telah menunjukkan efektivitas berbagai metode dalam tugas analisis sentimen. Made dkk (2024) menggunakan model LSTM untuk menganalisis sentimen berita pariwisata berbahasa Inggris dan mencapai akurasi 81,36%, serta mengidentifikasi topik tersembunyi menggunakan metode LDA [5]. Lalu Susanto dkk (2024) menerapkan algoritma *K-Nearest Neighbor* untuk mengklasifikasikan sentimen berita terkait *Bitcoin* dan memperoleh akurasi 74,40% [6]. Sementara itu, Simatupang dan Nursinta (2024) menggunakan metode *Naive Bayes* untuk mendeteksi sentimen dalam berita hoaks lowongan kerja, dengan akurasi sebesar 81% [7].

Convolutional Neural Network (CNN) dan *IndoBERT* adalah dua metode berbasis *deep learning* untuk analisis sentimen teks Indonesia yang dibandingkan dalam penelitian ini. CNN

dikenal karena kemampuannya mengidentifikasi pola lokal dalam data teks, seperti kalimat atau kata kunci yang menunjukkan sentimen. Di sisi lain, IndoBERT adalah model *transformer* yang dilatih hanya pada korpus bahasa Indonesia, sehingga dapat memahami konteks kalimat secara penuh [8].

16.691 data artikel berita politik Indonesia yang didapat dari hasil *web scraping* portal berita *online* digunakan dalam penelitian ini mencakup berbagai topik atau *keyword*, termasuk sidang parlemen, kebijakan publik, kinerja menteri, dan transparansi pemerintah, membentuk *dataset* yang digunakan. Setiap artikel berita dikategorikan ke dalam salah satu dari tiga kelompok sentimen menggunakan kategorisasi leksikal otomatis.

Selain memberikan wawasan tentang pemilihan model terbaik untuk analisis sentimen dalam konteks bahasa Indonesia, penelitian ini berusaha mengevaluasi dan membandingkan kinerja CNN dan IndoBERT dalam mengkategorikan sentimen berita politik Indonesia.

II. METODOLOGI PENELITIAN

Penelitian ini mengikuti pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang terdiri dari enam tahap utama, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Namun, dalam konteks penelitian ini, metode CRISP-DM diterapkan hanya sampai tahap kelima, yaitu *Evaluation*, karena fokus penelitian berada pada proses pemodelan dan analisis performa model klasifikasi sentimen:

A. *Business Understanding*

Tahap ini bertujuan untuk memahami tujuan utama dari penelitian, permasalahan yang ingin diselesaikan, dan manfaat yang diharapkan dari hasil analisis [9].

B. *Data Understanding*

Pada tahap ini dilakukan proses identifikasi, eksplorasi, dan penilaian terhadap data yang akan digunakan. Tujuannya adalah untuk memahami struktur, kualitas, dan potensi penggunaan data dalam analisis lebih lanjut [10].

C. *Data Preparation*

Merupakan tahap untuk menyiapkan data agar layak digunakan dalam proses pemodelan. Ini mencakup proses pembersihan data, transformasi, seleksi fitur, dan pelabelan jika diperlukan [11].

D. *Data Modeling*

Pada tahap ini dilakukan pemilihan algoritma dan pembangunan model prediktif berdasarkan data yang telah disiapkan. Setiap model memiliki pendekatan dan parameter yang berbeda sesuai karakteristik data [12].

E. *Evaluation*

Tahap ini bertujuan untuk mengevaluasi performa model yang telah dibangun, memastikan bahwa hasil yang diperoleh telah sesuai dengan tujuan awal, serta mengukur keakuratan dan efektivitas model dalam konteks data yang digunakan [13].

III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari setiap tahapan proses berdasarkan pendekatan CRISP-DM yang telah diterapkan, dimulai dari pemahaman masalah hingga evaluasi performa model klasifikasi sentimen.

A. *Business Understanding*

Masalah pokok dalam penelitian ini adalah kurangnya teknik yang efisien untuk memahami bagaimana masyarakat umum memandang berita politik Indonesia yang secara luas dibagikan secara *online*. Untuk memfasilitasi analisis opini publik yang luas secara otomatis, diperlukan metode berbasis pemodelan teks yang dapat secara andal mengkategorikan sentimen, mengingat kompleksitas bahasa dan konteks dalam berita politik.

B. *Data Understanding*

Data yang digunakan adalah berita politik berbahasa Indonesia dari portal seperti Detiknews. Data dikumpulkan melalui *web scraping* menggunakan 10 kata kunci topik utama seperti "Kinerja Menteri", "Predisen", "Program Pemerintah", "Kebijakan Publik", "RUU DPR", "Korupsi", "Rapat Paripurna", "Transparansi Pemerintah", "Kebebasan

Berpendapat”, dan “Sidang DPR”. Total data yang berhasil dikumpulkan berjumlah 16.691 baris, yang masing-masing memuat teks isi berita, judul, dan tanggal.

C. Data Preparation

Tahapan *preprocessing* dilakukan agar data siap untuk pemodelan. Proses yang diterapkan mencakup beberapa proses seperti normalisasi, *stopword removal*, *labeling*, *tokenizing*, *label encoding*, dan *split data*.

Proses pertama yaitu normalisasi teks, pada tahap ini data berita atau variabel *content* diubah menjadi huruf kecil atau *lowercase*, lalu semua angka, simbol atau karakter *non-alfabet*, dan spasi lebih akan dihapus. Proses ini bisa dilihat gambarannya pada tabel 1.

Tabel 1. Sampel Data Sebelum dan Sesudah Normasilasi

Sebelum	Sesudah
Ketua DPR RI Puan Maharani meminta Pemerintah terus memantau dan memastikan keselamatan WNI yang berada di wilayah konflik di Iran dan Israel. "Keselamatan harus menjadi prioritas," imbuh Puan. Ia juga mengecam agresi militer Israel terhadap Iran karena dianggap melanggar hukum internasional.....	ketua dpr ri puan maharani meminta pemerintah terus memantau memastikan keselamatan wni berada wilayah konflik iran israel keselamatan harus menjadi prioritas imbuh puan mengecam agresi militer israel terhadap iran dianggap melanggar hukum internasional.....

Setelah proses normalisasi dilakukan, tahap berikutnya adalah *stopword removal*, yaitu menghapus kata-kata umum yang tidak memiliki makna signifikan dalam konteks sentimen, seperti "yang", "dan", "di", "untuk", dan sebagainya. Penghapusan ini bertujuan untuk mereduksi dimensi teks dan memfokuskan model pada kata-kata bermakna penting.

Stopword dihapus menggunakan *library Sastrawi*, yang telah disesuaikan untuk Bahasa Indonesia. Proses ini menghasilkan teks yang lebih ringkas namun tetap merepresentasikan inti informasi dari kalimat aslinya. Proses ini bisa dilihat pada gambarannya pada tabel 2.

Tabel 2. Sampel Data Sebelum dan Sesudah *Stopword Removal*

Sebelum	Sesudah
ketua dpr ri puan maharani meminta pemerintah terus memantau memastikan	ketua dpr puan maharani meminta pemerintah memantau memastikan keselamatan wni

keselamatan wni berada wilayah konflik iran israel keselamatan prioritas imbuh puan mengecam agresi militer israel terhadap iran dianggap melanggar hukum internasional.....

Setelah teks dibersihkan dari *stopword*, proses selanjutnya adalah pemberian label sentimen secara otomatis menggunakan pendekatan berbasis leksikon (*lexicon-based*). Pendekatan ini dilakukan dengan mencocokkan kata-kata dalam teks terhadap dua daftar kata kunci, yaitu:

- a. Kata positif: seperti apresiasi, dukungan, berhasil, transparan, optimal, dll.
- b. Kata negatif: seperti korupsi, kecurangan, kerugian, pelanggaran, protes, dll.

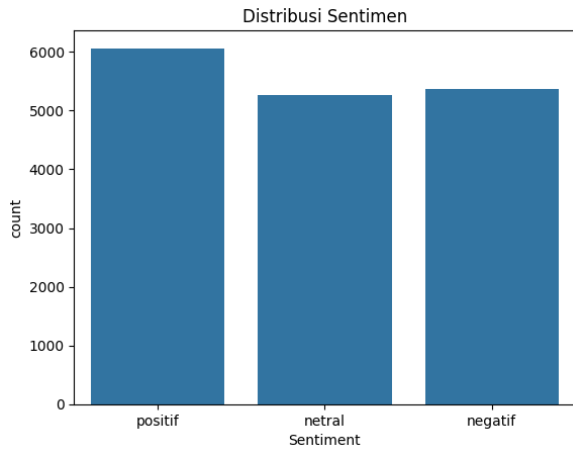
Sistem akan menghitung jumlah kemunculan kata positif dan negatif dalam setiap teks. Skema penentuan label dilakukan sebagai berikut:

- a. Jika jumlah kata positif > negatif → label = positif
- b. Jika jumlah kata negatif > positif → label = negatif
- c. Jika jumlah keduanya sama atau nol → label = netral

Tabel 3. Sampel Data Hasil Labeling

Sampel Data Hasil Stopword Removal	Label
ketua dpr puan maharani meminta pemerintah memantau memastikan keselamatan wni wilayah konflik iran israel keselamatan prioritas imbuh puan mengecam agresi militer israel iran dianggap melanggar hukum internasional.....	positif

Setelah diberi label berikut adalah distribusi datanya yaitu label positif sebanyak 6062 data, label negatif sebanyak 5367 data, dan label netral sebanyak 5262 data.



Gambar 1. Distribusi Hasil Labeling

Setelah teks diberi label sentimen, langkah berikutnya adalah tokenisasi, yaitu memecah teks menjadi unit-unit kata atau token. Tokenisasi memungkinkan sistem mengubah teks menjadi bentuk yang dapat diproses oleh model *deep learning*. Proses ini menghasilkan daftar token dalam bentuk *array string*.

Tabel 4. Sampel Data Hasil Tokenisasi

Sampel Data Hasil Stopword Removal	Hasil Tokenisasi
ketua dpr puan maharani meminta pemerintah memantau memastikan keselamatan wni wilayah konflik iran israel keselamatan prioritas imbuh puan mengecam agresi militer israel iran dianggap melanggar hukum internasional.....	['ketua', 'dpr', 'puan', 'maharani', 'meminta', 'pemerintah', 'memantau', 'memastikan', 'keselamatan', 'wni', 'wilayah', 'konflik', 'iran', 'israel', 'keselamatan', 'prioritas', 'imbuh', 'puan', 'mengecam', 'agresi', 'militer', 'israel', 'iran', 'dianggap', 'melanggar', 'hukum', 'internasional', ...]

Untuk keperluan model CNN, token selanjutnya dikonversi ke bentuk numerik menggunakan *Keras Tokenizer*, lalu dilakukan *padding* untuk menyamakan panjang input teks. *Padding* diperlukan karena model membutuhkan input dengan panjang tetap. Panjang maksimum teks ditetapkan sebesar 350 token, dan *padding* dilakukan di bagian akhir (*post-padding*).

Tabel 5. Sampel Data Hasil *Padding*

Sampel Data Hasil Stopword Removal	Hasil <i>Padding</i>
['ketua', 'dpr', ..., 'iran']	[3, 12, 24, 87, 51, ..., 0, 0, 0, ...]

Setelah setiap teks diberi label sentimen, tokenisasi dan padding, langkah berikutnya adalah melakukan *label encoding*, yaitu

mengubah label dalam bentuk teks (positif, negatif, netral) menjadi format numerik yang dapat dikenali oleh algoritma *machine learning*. *Encoding* dilakukan dengan skema sebagai berikut:

Tabel 6. Skema *Label Encoding*

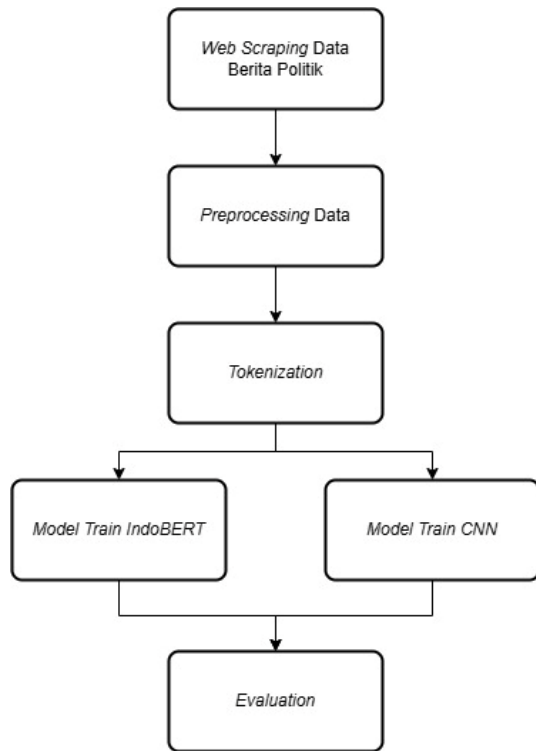
Label	Encode
Negatif	0
Positif	1
Netral	2

Setelah seluruh data selesai diproses dan diberi label numerik, tahap selanjutnya adalah membagi dataset menjadi data latih dan data uji. Tujuannya adalah untuk melatih model menggunakan sebagian data, lalu menguji kinerjanya menggunakan data yang tidak pernah dilihat sebelumnya.

Pembagian dilakukan menggunakan metode *stratified split* dengan rasio 8:2 yaitu 80% data untuk pelatihan (*training*) dan 20% data untuk pengujian (*testing*). Penggunaan *stratified split* memastikan bahwa distribusi label sentimen (negatif, netral, positif) tetap seimbang di kedua *subset*, sehingga hasil evaluasi model menjadi lebih valid dan representatif.

D. Data Modeling

Penelitian ini menggunakan dua pendekatan model *deep learning* untuk klasifikasi sentimen, yaitu *Convolutional Neural Network (CNN)* dan *IndoBERT*, untuk mengetahui perbedaan performa keduanya dalam mengolah teks berbahasa Indonesia. Proses CNN dan *IndoBERT* bisa dilihat pada gambar 2.



Gambar 2. Alur Proses CNN dan IndoBERT

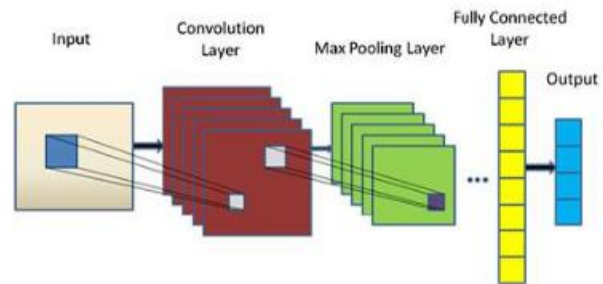
a. Convolutional Neural Network (CNN)

Model Convolutional Neural Network (CNN) dirancang untuk mengenali pola lokal dalam teks melalui operasi konvolusi satu dimensi. Dalam penelitian ini, CNN digunakan sebagai salah satu metode untuk melakukan klasifikasi sentimen terhadap berita politik Indonesia ke dalam tiga kategori: positif, negatif, dan netral.

Arsitektur model terdiri dari beberapa lapisan. Pertama, teks yang telah melalui tahap tokenisasi dikonversi menjadi representasi vektor melalui *Embedding layer* berdimensi 128. Selanjutnya, *Conv1D layer* dengan 128 *filter* dan ukuran *kernel* 5 digunakan untuk mengekstraksi fitur penting dari urutan kata dalam teks. Output dari lapisan konvolusi kemudian diproses oleh *MaxPooling1D layer* untuk mereduksi dimensi dan mempertahankan fitur dominan.

Setelah itu, hasilnya diratakan menggunakan *Flatten layer* dan diteruskan ke *Dense layer* dengan 64 *neuron* beraktivasi *ReLU*. Untuk mencegah *overfitting*, diterapkan *Dropout layer* sebesar 0.5. Akhirnya, *Output layer* menggunakan fungsi aktivasi *softmax* dengan 3 *neuron* untuk melakukan klasifikasi multi-kelas

[14]. Untuk gambaran arsitektur umum CNN bisa dilihat pada gambar 3.



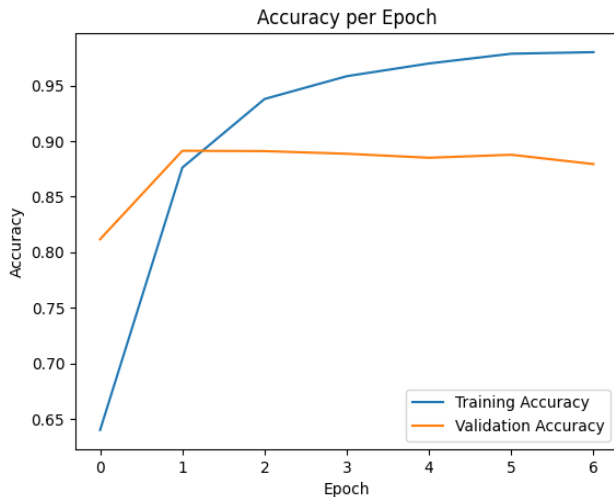
Gambar 3. Arsitektur CNN

Setelah seluruh data selesai diproses dan diberi label numerik, tahap selanjutnya adalah membagi *dataset* menjadi data latih dan data uji. Tujuannya adalah untuk melatih model menggunakan sebagian data, lalu menguji kinerjanya menggunakan data yang tidak pernah dilihat sebelumnya.

Model dikompilasi menggunakan fungsi *loss sparse_categorical_crossentropy*, *optimizer Adam* dengan learning rate 0.0005, dan metrik evaluasi akurasi. Proses pelatihan dilakukan selama maksimal 15 *epoch* dengan *batch size* 32. Untuk menjaga generalisasi model, digunakan teknik *EarlyStopping* dengan nilai *patience* = 5. Berdasarkan hasil pelatihan, model menunjukkan peningkatan performa signifikan dalam beberapa *epoch* pertama, dengan akurasi validasi terbaik sebesar 89.13% yang dicapai pada *epoch* ke-2. Pelatihan berhenti pada *epoch* ke-7 dikarenakan tidak ada peningkatan dari *validation_loss*. Hasil pelatihan bisa dilihat pada tabel 7 dan visualisasi grafik pelatihan bisa dilihat pada gambar 4.

Tabel 7. Hasil Training Model CNN

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	63.98%	81.16%	0.777	0.4686
2	87.61%	89.13%	0.3546	0.3171
3	93.79%	89.10%	0.1788	0.3275
4	95.84%	88.86%	0.1155	0.4096
5	96.99%	88.50%	0.0771	0.5118
6	97.87%	88.77%	0.0649	0.5687
7	98.01%	87.93%	0.0537	0.6469

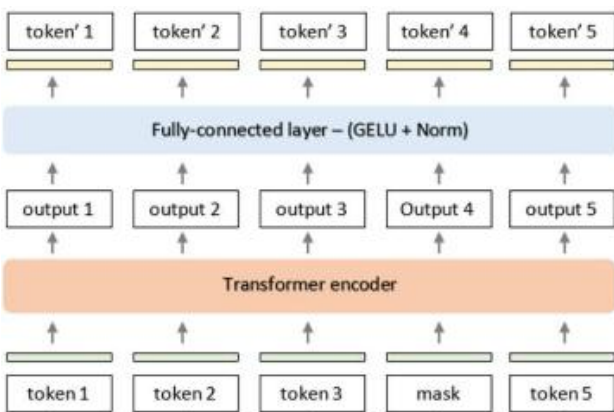


Gambar 4. Grafik Pelatihan Model CNN

b. IndoBERT

Model kedua yang digunakan dalam penelitian ini adalah IndoBERT, yaitu model *pre-trained* berbasis arsitektur *transformer* yang dikembangkan khusus untuk Bahasa Indonesia. IndoBERT mampu memahami konteks kata dalam kalimat secara menyeluruh melalui mekanisme *self-attention* dua arah, sehingga cocok digunakan untuk tugas klasifikasi sentimen terhadap teks yang kompleks seperti berita politik.

Penelitian ini, memakai model *indobenchmark/indobert-base-p1* dari *Huggingface* sebagai dasar, yang kemudian di-*fine-tune* pada *dataset* berita politik berlabel sentimen. Untuk gambaran arsitektur indoBERT bisa dilihat pada gambar 5.



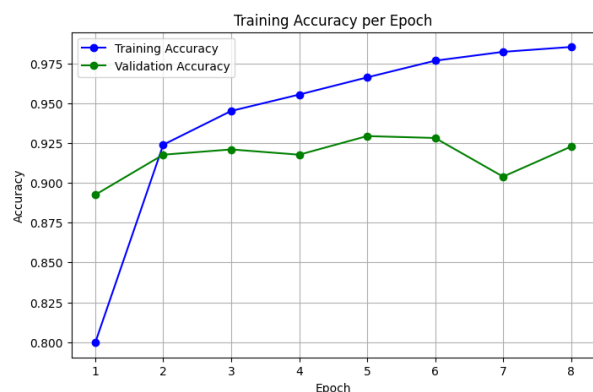
Gambar 5. Arsitektur IndoBERT

Implementasi IndoBERT diawali dengan proses tokenisasi menggunakan *tokenizer indobenchmark/indobert-base-p1*, dengan *padding* dan *truncation* hingga maksimal 350 token per teks. *Dataset* kemudian diubah ke format *DatasetDict* dari *Huggingface* dan dikonversi menjadi *tf.data.Dataset* untuk efisiensi pelatihan. Model klasifikasi dibangun di atas struktur IndoBERT dengan tiga output node untuk sentimen negatif, netral, dan positif, menggunakan fungsi aktivasi *softmax*.

Model dikompilasi menggunakan *loss function Sparse Categorical Crossentropy (from_logits=True)*, *optimizer Adam (learning rate 2e-5)*, *batch size 32*, dan dilatih selama maksimal 10 *epoch* dengan *EarlyStopping (patience = 3)*, tetapi pada penelitian ini pelatihan berhenti di *epoch* ke-8 karena tidak ada peningkatan di *validation_loss*. Hasil pelatihan menunjukkan performa yang stabil, dengan akurasi validasi terbaik sebesar 92.93% pada *epoch* ke-5, yang mengindikasikan bahwa model mampu mengenali pola sentimen secara efektif tanpa mengalami *overfitting* [15]. Hasil pelatihan bisa dilihat pada tabel 8 dan visualisasi grafik pelatihan bisa dilihat pada gambar 6.

Tabel 7. Hasil Training Model IndoBERT

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	79.97%	89.25%	0,4827	0,3256
2	92.39%	91.76%	0,2158	0,2231
3	94.51%	92.09%	0,1609	0,2271
4	95.54%	91.76%	0,125	0,2447
5	96.61%	92.93%	0,0977	0,2207
6	97.66%	92.81%	0,0656	0,2407
7	98.22%	90.39%	0,0524	0,3658
8	98.52%	92.27%	0,0398	0,2986



Gambar 6. Grafik Pelatihan Model IndoBERT

E. Evaluation

Pada tahap ini, performa kedua model dievaluasi dengan menggunakan classification report yang mencakup metrik *precision*, *recall*, *f1-score*, dan *support* untuk masing-masing kelas sentimen. Selain itu, juga digunakan *confusion matrix* yang menyajikan informasi mengenai distribusi prediksi model terhadap label yang sebenarnya, sehingga dapat diketahui sejauh mana model mampu membedakan kelas secara tepat dan identifikasi kesalahan klasifikasi antar kelas. Untuk hasil dari *classification report* bisa dilihat pada gambar 7 dan 8, sedangkan untuk *confusion matrix* bisa dilihat pada gambar 9 dan 10.

	precision	recall	f1-score	support
Negatif	0.89	0.95	0.92	1074
Netral	0.93	0.78	0.85	1052
Positif	0.87	0.94	0.90	1213
accuracy			0.89	3339
macro avg	0.89	0.89	0.89	3339
weighted avg	0.89	0.89	0.89	3339

Gambar 7. Classification Report Model CNN

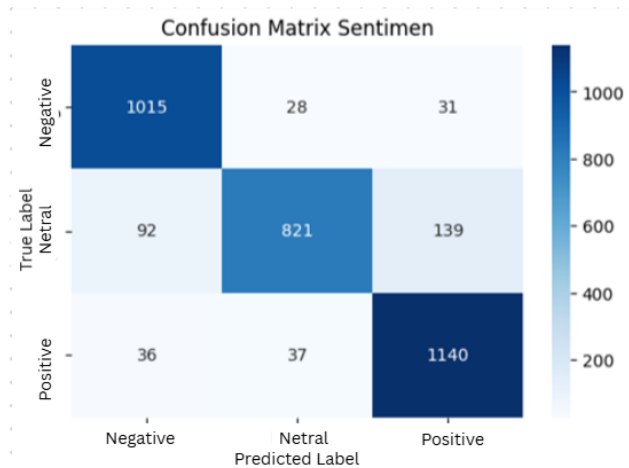
	precision	recall	f1-score	support
Negatif	0.92	0.95	0.94	1074
Netral	0.92	0.90	0.91	1052
Positif	0.94	0.93	0.94	1213
accuracy			0.93	3339
macro avg	0.93	0.93	0.93	3339
weighted avg	0.93	0.93	0.93	3339

Gambar 8. Classification Report Model IndoBERT

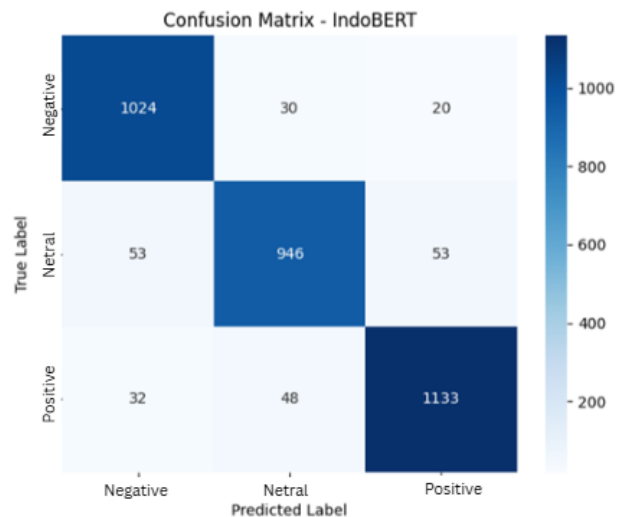
Model CNN menunjukkan akurasi keseluruhan sebesar 89%, dengan skor f1 tertinggi pada kelas negatif (0.92) dan skor terendah pada kelas netral (0.85). Meskipun precision pada kelas netral tergolong tinggi (0.93), *recall*-nya jauh lebih rendah (0.78), menandakan bahwa model CNN kesulitan menangkap semua data netral dengan baik. Secara umum, *macro average* dan *weighted average* dari *precision*, *recall*, dan *f1-score* semuanya berada di angka 0.89.

Sementara itu, model IndoBERT menunjukkan performa yang lebih unggul dengan akurasi keseluruhan sebesar 93%. *Precision*, *recall*, dan *f1-score* relatif seimbang dan tinggi di semua

kelas, dengan nilai *f1-score* tertinggi pada kelas positif (0.94) dan nilai terendah tetap tinggi pada kelas netral (0.91). *Recall* untuk kelas netral juga meningkat signifikan dari 0.78 (pada CNN) menjadi 0.90, menunjukkan kemampuan IndoBERT dalam mengenali semua jenis sentimen dengan lebih merata. Rata-rata makro dan rata-rata tertimbang untuk *precision*, *recall*, dan *f1-score* mencapai 0.93, lebih tinggi dari CNN.



Gambar 9. Confusion Matrix Model CNN



Gambar 10. Confusion Matrix Model IndoBERT

Berdasarkan *confusion matrix*, model CNN menunjukkan performa yang cukup baik dalam mengklasifikasikan kelas negatif dan positif, namun mengalami kesulitan dalam membedakan kelas netral. Dari 1.052 data netral, hanya 821 yang berhasil diklasifikasikan dengan benar, sementara sisanya keliru diprediksi sebagai

negatif (92) atau positif (139). Hal ini mengindikasikan bahwa CNN cenderung tidak stabil dalam menangani sentimen yang bersifat netral, mungkin karena keterbatasannya dalam menangkap konteks kalimat yang ambigu.

Sebaliknya, model IndoBERT menunjukkan kinerja yang lebih konsisten dan seimbang di semua kelas. Untuk kelas netral, model berhasil mengklasifikasikan 946 data dengan benar dari total 1.052, dengan jumlah kesalahan yang relatif seimbang ke kelas negatif (53) dan positif (53). Selain itu, akurasi klasifikasi pada kelas negatif dan positif juga sedikit lebih tinggi dibandingkan CNN. Performa ini mencerminkan keunggulan IndoBERT dalam memahami konteks semantik dan gramatikal dalam teks berita berbahasa Indonesia, berkat arsitektur *transformer* yang mampu memproses hubungan antar kata secara dua arah.

IV. KESIMPULAN

Penelitian ini membandingkan model CNN dan IndoBERT dalam klasifikasi sentimen berita politik berbahasa Indonesia. Hasil evaluasi menunjukkan bahwa IndoBERT memiliki performa yang lebih unggul, dengan akurasi validasi tertinggi sebesar 92.93% dan metrik evaluasi yang stabil di semua kelas. Sementara itu, CNN mencatat akurasi 89.13%, namun kurang akurat dalam mengenali sentimen netral. Dari hasil tersebut, IndoBERT dinilai lebih efektif dan andal dalam memahami konteks teks politik. Penelitian ini menunjukkan bahwa model berbasis *transformer* sangat cocok untuk tugas analisis sentimen berbahasa Indonesia.

REFERENSI

- [1] N. Hariyani and A. Y. Rahmawati, "Media Sosial dan Aplikasi Kecerdasan Buatan Dalam Preferensi dan Partisipasi Pemilih Di Pemilihan Umum," *Jurnal Intelektual Administrasi Publik dan Ilmu Komunikasi*, vol. 11, no. 2, pp. 102–111, 2024, doi: doi.org/10.55499/intelektual.v11i02.1210.
- [2] C. Rizma Riendani, I. S. Raghda Atshila, A. Abhinaya, A. Ridlo Abdillah, and B. Dzaky Mufadhol, "Pengaruh Algoritma Media Sosial terhadap Selektivitas Konsumsi Berita Politik pada Generasi Z di Indonesia," *Jurnal Pustaka Cendekia Hukum dan Ilmu Sosial*, vol. 2, no. 3, pp. 224–228, 2024, doi: doi.org/10.70292/pchukumsosial.v2i3.68.
- [3] A. Turmudi Zy, A. Nugroho, A. Rivaldi, and I. Afriantoro, "Analisis Sentimen Terhadap Pembobolan Data pada Twitter dengan Algoritma Naive Bayes," *Jurnal Teknologi Informatika dan Komputer*, vol. 8, no. 2, pp. 202–213, Sep. 2022, doi: [10.37012/jtik.v8i2.1240](https://doi.org/10.37012/jtik.v8i2.1240).
- [4] M. Khadapi *et al.*, "Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi Twitter tentang Pemilu 2024," *JUKI : Jurnal Komputer dan Informatika*, vol. 6, no. 2, pp. 130–137, 2024.
- [5] N. Made, K. Sedana, I. Nyoman, S. Wijaya, I. Ketut, and R. Artana, "Analisis Sentimen Berbahasa Inggris Dengan Metode LSTM Studi Kasus Berita Online Pariwisata Bali," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 11, no. 6, pp. 1325–1334, 2024, doi: [10.25126/jtiik.2024118792](https://doi.org/10.25126/jtiik.2024118792).
- [6] H. Susanto, A. Setyanto, and A. H. Muhammad, "Analisis Sentimen Berita terhadap Bitcoin dengan Metode Klasifikasi K-Nearest Neighbor," *Jurnal Teknik Informatika*, vol. 4, no. 2, p. 2024, 2024, [Online]. Available: <https://en.wikipedia.org/wiki/2023#References>
- [7] D. S. Simatupang and Siti Nursinta, "Analisis Sentimen Terhadap Berita Hoaks Lowongan Kerja Dengan Metode Naive Bayes," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 5, no. 2, pp. 474–482, Sep. 2024, doi: [10.37859/coscitech.v5i2.7719](https://doi.org/10.37859/coscitech.v5i2.7719).
- [8] M. Napiyah, R. D. Astuti, and E. K. Pratama, "Komparasi Algoritma Machine Learning untuk Klasifikasi Gejala Corona virus Disease 19 (Covid-19)," *Computer Science (CO-SCIENCE)*, vol. 3, no. 2, pp. 78–83, 2023, doi: doi.org/10.31294/coscience.v3i2.1984.
- [9] Rahman Kadafi Abdul and Septi Rizqiah Alifia, "Analisis Sentimen Pengguna Twitter Terhadap Vaksin Booster Menggunakan Metode Naive Bayes Dan K-Nearest Neighbor," *Jurnal Sains dan Teknologi*, vol. 4, no. 1, pp. 1–5, 2022.
- [10] P. A. Nugroho, N. Suchahyo, and I. Kurniati, "Sentimen Analisis pada Sosial Media Twitter untuk Menilai Respon Masyarakat terhadap Seleksi Kartu Prakerja," *Jurnal Teknologi Informatika dan Komputer*, vol. 9, no. 1, pp. 72–83, Mar. 2023, doi: [10.37012/jtik.v9i1.862](https://doi.org/10.37012/jtik.v9i1.862).
- [11] N. Firdausy, I. Yuadi, and I. Puspitasari, "Analisis Sentimen Evaluasi Reaksi E-Learning Menggunakan Algoritma Naive Bayes Support Vector Machine dan Deep Learning Sentiment analysis of e-learning reaction evaluations using Naive Bayes, Support Vector Machine, and Deep Learning algorithms," *Techno.COM*, vol. 22, no. 3, pp. 677–689, 2023.
- [12] N. Suchahyo *et al.*, "SWADHARMA (JRIS) ANALISIS SENTIMEN MASYARAKAT TERHADAP UU CIPTA KERJA PADA MEDIA SOSIAL TWITTER," *JURNAL REKAYASA INFORMASI SWADHARMA (JRIS)*, vol. 2, no. 1, pp. 63–70, 2022.
- [13] A. Musthafa, T. Harmini, A. Rafiq, and N. Marantika, "Pemanfaatan Machine Learning dalam Menganalisis Sentimen Terhadap Program TAPERA di Platform Digital X," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 2, pp. 587–597, Mar. 2025, doi: [10.57152/malcom.v5i2.1801](https://doi.org/10.57152/malcom.v5i2.1801).
- [14] S. Ahmad, A. M. Ridwan, and G. D. Setyawan, "ANALISIS SENTIMEN PRODUCT TOOLS & HOME MENGGUNAKAN METODE CNN DAN LSTM," *TEKNOKOM*, vol. 6, no. 2, pp. 133–140, Aug. 2023, doi: [10.31943/teknokom.v6i2.154](https://doi.org/10.31943/teknokom.v6i2.154).
- [15] J. Khatib Sulaiman, L. Septian, T. Aljauza, and C. Juliane, "Analisis Sentimen Putusan Mahkamah

Konstitusi Terhadap Batas Usia Capres Dan Cawapres Menggunakan IndoBERT,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 6, pp. 4428–4439, 2023.