

**EKSTRAKSI INFORMASI DAN KLASIFIKASI BERITA PEMERINTAHAN
DAERAH MENGGUNAKAN FINE-TUNING INDOBERT**

Arfiana Maulidiyah¹, Mustafa²

^{1,2}Universitas Islam Sultan Agung Semarang

¹arfianamaulidyh@gmail.com

²cakmustafa@unissula.ac.id

Received: 01-04- 2026

Revised: 15-04-2026

Approved: 27-04-2026

ABSTRACT

Pemanfaatan kebijakan berbasis riset (*Evidence-Based Policy*) semakin penting dalam mendukung pengambilan keputusan pemerintah daerah, khususnya pada sektor pendidikan, kesehatan, dan ketenagakerjaan di Provinsi Jawa Tengah. Tingginya volume berita daring yang dipublikasikan setiap hari menjadi tantangan bagi Organisasi Perangkat Daerah (OPD) dalam memantau dan menganalisis informasi secara manual. Penelitian ini bertujuan mengembangkan sistem ekstraksi informasi dan klasifikasi berita pemerintahan daerah secara otomatis menggunakan fine-tuning IndoBERT. Dataset terdiri dari 1.025 artikel berita yang dikumpulkan melalui web scraping dari empat media daring, kemudian diseleksi menjadi 377 data berlabel setelah melalui preprocessing, ekstraksi entitas Named Entity Recognition (NER), dan pelabelan semi-otomatis. Model IndoBERT di-fine-tuning untuk mengklasifikasikan berita ke dalam tiga kategori OPD: Dinas Pendidikan, Dinas Kesehatan, dan Dinas Ketenagakerjaan. Hasil evaluasi menunjukkan akurasi sebesar 88,16%, dengan precision macro 0,88, recall macro 0,87, dan F1-score macro 0,87 pada data pengujian. Sistem ini diimplementasikan dalam dashboard Streamlit interaktif yang membantu OPD mengidentifikasi isu strategis secara cepat dan akurat, mendukung tata kelola pemerintahan berbasis data.

Kata kunci: IndoBERT, Ekstraksi, Klasifikasi, NER, NLP

PENDAHULUAN

Pemerintah Provinsi Jawa Tengah menetapkan arah pembangunan daerah melalui Rencana Pembangunan Jangka Menengah Daerah (RPJMD) dengan visi terwujudnya Jawa Tengah yang mandiri, sejahtera, dan bebas korupsi [1]. Tiga sektor strategis yang menjadi fokus utama pembangunan sumber daya manusia adalah pendidikan, kesehatan, dan ketenagakerjaan [2]. Ketiga sektor ini memiliki peran krusial dalam membentuk sumber daya manusia yang berkualitas, berdaya saing, dan mampu beradaptasi terhadap dinamika ekonomi serta perkembangan industri yang semakin pesat [3].

Dalam rangka mengoptimalkan pembangunan pada ketiga sektor tersebut, Badan Riset dan Inovasi Daerah (BRIDA) Jawa Tengah mendorong penerapan Evidence-Based Policy (EBP), yakni perumusan kebijakan publik yang berpijak pada data empiris, hasil penelitian, serta analisis sistematis sebagai dasar pengambilan keputusan [4]. Pendekatan ini bertujuan menghasilkan kebijakan yang lebih akuntabel, responsif terhadap permasalahan aktual, dan selaras dengan kebutuhan masyarakat di lapangan [5].

Media berita daring merupakan salah satu sumber informasi strategis yang mencerminkan dinamika sosial, permasalahan publik, dan respons masyarakat terhadap kebijakan pemerintah. Media seperti Tribun Jateng, Suara Merdeka, Radar Semarang, dan Kompas.com secara konsisten menyajikan pemberitaan terkait aktivitas Organisasi Perangkat Daerah (OPD), mulai dari

program kerja, capaian kebijakan, hingga kritik masyarakat [5].

Namun, volume berita yang dipublikasikan setiap hari menimbulkan tantangan bagi OPD dalam memantau dan menganalisis informasi secara manual [6]. Tanpa dukungan sistem otomatis yang memadai, proses pemilahan dan penilaian relevansi berita memerlukan waktu serta tenaga yang besar, dan berpotensi menimbulkan ketidakkonsistenan dalam pengambilan keputusan [7]. Di sisi lain, pemerintah daerah tidak hanya perlu mengetahui relevansi berita terhadap OPD tertentu, tetapi juga perlu memahami kecenderungan sentimen pemberitaan—apakah bersifat positif, netral, atau negative agar dapat mengidentifikasi isu yang memerlukan perhatian dan tindakan segera.

Kemajuan teknologi Natural Language Processing (NLP) dan deep learning memungkinkan pengolahan teks berskala besar secara otomatis dan efisien [8]. Arsitektur Transformer, khususnya BERT (Bidirectional Encoder Representations from Transformers), terbukti unggul dalam memahami konteks bahasa secara bidireksional [9]. IndoBERT, varian BERT yang dilatih secara khusus pada korpus bahasa Indonesia berskala besar, mampu menangkap karakteristik linguistik bahasa Indonesia secara lebih akurat dibandingkan model multilingual umum [10]. Melalui proses fine-tuning, IndoBERT dapat diadaptasi untuk tugas-tugas spesifik seperti klasifikasi teks dan ekstraksi informasi pada domain berita pemerintahan [11].

Berbagai penelitian terdahulu telah memanfaatkan IndoBERT untuk tugas-tugas NLP berbahasa Indonesia. Simanjuntak et al. [12] memanfaatkan IndoBERT-base-p1 dengan tiga metode hyperparameter tuning—Bayesian optimization, grid search, dan random search—untuk deteksi berita hoaks, menghasilkan F1-score terbaik sebesar 91,56% menggunakan Bayesian optimization. Kunaefi et al. [13] mengombinasikan IndoBERT dengan fungsi Focal Loss untuk mengatasi ketidakseimbangan data pada klasifikasi berita hoaks [14], mencapai akurasi 97,2% yang melampaui pendekatan Cross-Entropy Loss standar. Tobing et al. [6] mengembangkan model deteksi hoaks politik berbasis fine-tuning IndoBERT dengan dataset 20.928 artikel dari media besar Indonesia, menghasilkan akurasi 94,1% dan ROC AUC 0,991. Iskandar et al. [15] mengeksplorasi kombinasi IndoBERT Embedding dan Long Short-Term Memory (LSTM) untuk klasifikasi multiclass tweet berbahasa Indonesia, menghasilkan F1-score 98,90%. Budi dan Suryono [16] melakukan kajian sistematis (Systematic Literature Review) tentang penerapan metode NER pada dataset bahasa Indonesia, memberikan landasan teoritis yang komprehensif bagi implementasi ekstraksi entitas bernama dalam teks berita pemerintahan.

Penelitian lain yang relevan juga dilakukan oleh Juarto [11] yang mengklasifikasikan berita berbahasa Indonesia menggunakan IndoBERT dengan akurasi 87,3%. Merdiansah dan Ridha [17] menerapkan IndoBERT untuk analisis sentimen pengguna media sosial terkait kebijakan publik. Nuryadi et al. [18] mengimplementasikan fine-tuning IndoBERT pada analisis sentimen layanan pemerintah dengan F1-score 91,2%. Ramdani [19] membuktikan keunggulan IndoBERT dibandingkan model Transformer lainnya untuk klasifikasi teks berbahasa Indonesia. Federico et al. [20] mengkombinasikan IndoBERT dengan CNN untuk meningkatkan akurasi klasifikasi berita.

Penelitian-penelitian tersebut menunjukkan efektivitas IndoBERT untuk berbagai tugas NLP berbahasa Indonesia, namun umumnya terbatas pada satu

tugas tunggal seperti deteksi hoaks atau klasifikasi media sosial. Belum ada penelitian yang secara khusus mengintegrasikan tiga tugas sekaligus yaitu ekstraksi entitas NER, klasifikasi OPD multiclass, dan analisis sentiment pada domain berita pemerintahan daerah dalam satu sistem terpadu.

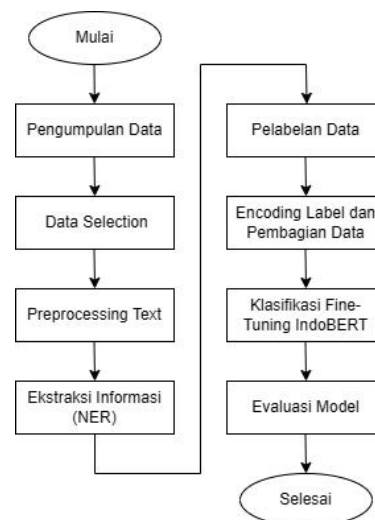
Penelitian ini dilakukan untuk mengatasi permasalahan tersebut, dengan mengembangkan sistem berbasis fine-tuning IndoBERT yang mampu:

1. Merancang model *fine-tuning* IndoBERT untuk melakukan ekstraksi informasi secara otomatis dari berita pemerintahan daerah untuk memperoleh entitas dan konteks yang relevan,
2. Mengembangkan model *fine-tuning* IndoBERT untuk mengklasifikasikan berita pemerintahan daerah berdasarkan kategori Organisasi Perangkat Daerah (OPD), sehingga setiap berita dapat dikelompokkan sesuai dengan instansi yang relevan,
3. Menganalisis dan mengevaluasi kinerja model *fine-tuning* IndoBERT dalam melakukan ekstraksi dan klasifikasi berita pemerintahan daerah berdasarkan OPD menggunakan metrik evaluasi standar.

Seluruh hasil divisualisasikan dalam dashboard Streamlit interaktif yang dapat diakses oleh OPD untuk mendukung penerapan kebijakan berbasis data di Provinsi Jawa Tengah.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan supervised learning dengan fine-tuning model IndoBERT untuk tugas klasifikasi teks. Alur penelitian dirancang secara sistematis yang terdiri dari delapan tahapan utama yaitu : 1. pengumpulan data, 2. data selection, 3. preprocessing data, 4. ekstraksi informasi menggunakan NER, 5. pelabelan data, 6. encoding label dan pembagian data, 7. klasifikasi berita menggunakan fine-tuning IndoBERT, serta 8. evaluasi dan pengujian model.



Gambar 1. Flowchart Sistem

Penelitian ini terdiri dari beberapa tahapan utama, dimana setiap tahapan menghasilkan output yang menjadi masukan (input) bagi tahap berikutnya, sehingga alur data penelitian berjalan secara berurutan dan terstruktur.

A. Pengumpulan Data

Data dikumpulkan melalui teknik web scraping menggunakan library Selenium untuk otomatisasi navigasi browser dan BeautifulSoup4 untuk ekstraksi konten HTML. Scraping dilakukan pada empat media berita daring yang secara konsisten menyajikan pemberitaan terkait pemerintahan Provinsi Jawa Tengah, yaitu Tribun Jateng, Suara Merdeka, Kompas.com, dan Radar Semarang. Keywords yang digunakan mencakup nama OPD target dan istilah pemerintahan daerah, seperti "dinas kesehatan", "dinas pendidikan", "dinas ketenagakerjaan", "pemkot semarang", dan "pemprov jateng". Proses ini menghasilkan 1.025 artikel berita dalam rentang periode satu tahun. Setiap artikel mencakup kolom url, judul berita, konten lengkap, tanggal publikasi, dan nama sumber media.

B. Seleksi Data

Tahap seleksi data bertujuan memastikan kualitas dan relevansi dataset sebelum masuk ke tahap pemodelan. Proses seleksi mencakup empat langkah: 1. pemeriksaan missing value pada seluruh kolom, hasilnya tidak ditemukan nilai kosong pada dataset awal; 2. pengecekan duplikasi berdasarkan URL dan kemiripan judul, tidak ditemukan duplikasi; 3. filter rentang waktu untuk mempertahankan hanya berita dalam periode satu tahun, sehingga 385 berita di luar rentang waktu dihapus; dan 4. filter relevansi OPD menggunakan kata kunci domain, sehingga 5 berita tidak relevan dihapus. Dari 1.025 berita awal, proses ini menghasilkan 635 berita bersih yang siap untuk preprocessing.

C. Preprocessing Data

Preprocessing dilakukan untuk meningkatkan kualitas dan konsistensi teks sebelum digunakan dalam pemodelan. Tahapan preprocessing meliputi: 1. case folding meliputi konversi seluruh teks menjadi huruf kecil untuk standarisasi; 2. text cleaning yaitu penghapusan HTML tags, URL, karakter khusus, angka, dan whitespace berlebih; 3. tokenisasi menggunakan NLTK word_tokenize yang telah disesuaikan untuk bahasa Indonesia; 4. stopword removal menggunakan library Sastrawi untuk menghapus kata-kata umum yang tidak memiliki makna signifikan; 5. stemming menggunakan Sastrawi Stemmer untuk mengubah kata berimbuhan ke bentuk kata dasar; dan 6. normalisasi kata tidak baku ke bentuk baku. Teks hasil preprocessing disimpan dalam kolom baru pada dataset untuk digunakan pada tahap berikutnya.

D. Ekstraksi Informasi Menggunakan NER

Ekstraksi entitas bernama (Named Entity Recognition) dilakukan menggunakan model pretrained 'cahya/bert-base-indonesian-NER' dari Hugging Face Transformers. Model ini mengidentifikasi empat jenis entitas yang relevan dengan domain berita pemerintahan: Person (PER) untuk nama pejabat atau tokoh, Organization (ORG) untuk nama instansi atau lembaga, Location (LOC) untuk nama wilayah atau daerah, dan Program (PROG) untuk nama program atau kebijakan pemerintah. Proses ekstraksi dilakukan terhadap seluruh konten berita, kemudian sistem mengambil

maksimal 100 kata yang mengandung entitas hasil ekstraksi sebagai ringkasan informatif. Hasil ekstraksi disimpan sebagai kolom tambahan dalam dataset untuk memperkaya informasi dan membantu proses klasifikasi.

E. Pelabelan Data

Pelabelan dilakukan secara semi-otomatis menggunakan Label Studio, platform open-source untuk anotasi data. Setiap artikel diberi dua jenis label secara bersamaan. Pertama, label OPD ditentukan berdasarkan frekuensi kemunculan kata kunci per kategori pada judul dan konten berita; jika terjadi skor yang sama antar kategori, digunakan prioritas hierarki sesuai fokus RPJMD: Dinas Kesehatan > Dinas Pendidikan > Dinas Ketenagakerjaan. Kedua, label sentimen ditentukan menggunakan formula Sentiment Score = $(n_{\text{positif}} - n_{\text{negatif}}) / (n_{\text{positif}} + n_{\text{negatif}} + n_{\text{netral}})$, dengan threshold positif jika skor > 0,2, netral jika $-0,2 \leq \text{skor} \leq 0,2$, dan negatif jika skor < -0,2. Setelah pelabelan otomatis, dilakukan verifikasi manual untuk memastikan konsistensi label. Dari 635 berita bersih, sebanyak 258 berita yang tidak secara jelas termasuk dalam salah satu kategori OPD dihapus, sehingga diperoleh 377 berita berlabel valid.

F. Encoding Label dan Pembagian Dataset

Label kategorikal OPD dikonversi menjadi format numerik menggunakan label encoding: Dinas Pendidikan = 0, Dinas Kesehatan = 1, Dinas Ketenagakerjaan = 2. Dataset kemudian dibagi dengan rasio 80:20, menghasilkan 301 data pelatihan dan 76 data pengujian. Pembagian ini dilakukan secara stratified untuk mempertahankan proporsi distribusi kelas pada masing-masing subset.

G. Arsitektur dan Konfigurasi Fine-Tuning IndoBERT

Model yang digunakan adalah indobenchmark/indobert-base-p2 yang dimuat menggunakan AutoModelForSequenceClassification dari library Transformers dengan num_labels=3. Teks berita ditokenisasi menggunakan AutoTokenizer IndoBERT dengan parameter max_length=512 token, padding='max_length', dan truncation=True. Token khusus [CLS] pada awal urutan digunakan sebagai representasi keseluruhan dokumen untuk tugas klasifikasi.

Tabel 1. Arsitektur Model Fine-Tuning IndoBERT

Komponen	Konfigurasi	Fungsi
IndoBERT Base	12 Transformer Layers	Encoder kontekstual bahasa Indonesia
Hidden Size	768 unit	Dimensi representasi 95ector konteks
Dropout	0,1	Regularisasi untuk mencegah overfitting
Linear Classifier	768 → 3 output	Pemetaan ke 3 kelas OPD

Komponen	Konfigurasi	Fungsi
Fungsi Aktivasi	Softmax	Menghasilkan probabilitas per kelas

Proses pelatihan menggunakan fungsi Cross-Entropy Loss sebagai objektif optimasi. Model checkpoint terbaik disimpan berdasarkan validation loss terendah dengan mekanisme early stopping untuk menghindari overfitting. Konfigurasi hyperparameter yang digunakan disajikan pada Tabel 2.

Tabel 2. Hyperparameter Fine-Tuning IndoBERT

Hyperparameter	Nilai	Keterangan
Learning Rate	2×10^{-5}	Kecepatan pembaruan bobot model
Batch Size	8	Jumlah sampel per iterasi pelatihan
Epochs	3	Jumlah iterasi pelatihan keseluruhan
Optimizer	AdamW	Optimasi dengan weight decay bawaan
Weight Decay	0,01	Regularisasi L2 untuk cegah overfitting
Warmup Steps	500	Penyesuaian awal learning rate
Max Token Length	512	Panjang token maksimum per input

H. Evaluasi Model

Evaluasi kinerja model klasifikasi dilakukan menggunakan empat metrik standar: 1. Accuracy—persentase prediksi benar dari total prediksi; 2. Precision—proporsi prediksi positif yang benar; 3. Recall—proporsi instance positif aktual yang berhasil dideteksi; dan 4. F1-Score—harmonic mean dari precision dan recall. Metrik dihitung per kelas dan diagregasi menggunakan macro average (rata-rata aritmatika per kelas) dan weighted average (rata-rata tertimbang berdasarkan jumlah instance per kelas). Visualisasi hasil evaluasi ditampilkan dalam bentuk confusion matrix dan classification report.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Pengumpulan dan Seleksi Data

Proses web scraping berhasil mengumpulkan 1.025 artikel berita dari empat sumber media. Distribusi perolehan data per sumber media disajikan pada Tabel 3.

Tabel 3. Distribusi Data Hasil Scraping

Sumber Media	Jumlah Berita	Persentase
TribunJateng.com	413	40,3%
SuaraMerdeka.com	402	39,2%

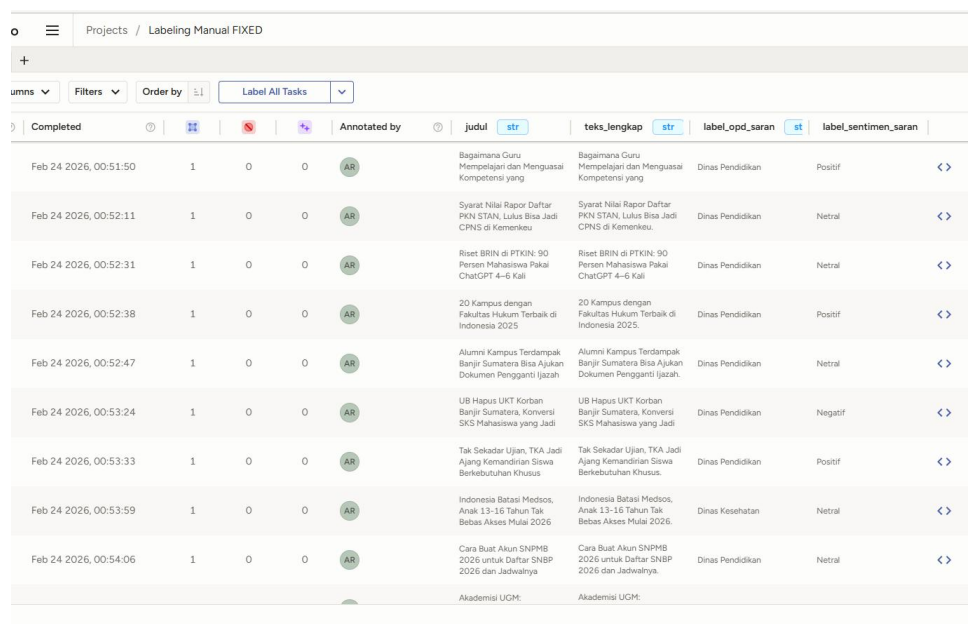
Sumber Media	Jumlah Berita	Persentase
Kompas.com	158	15,4%
RadarSemarang.com	52	5,1%
Total	1.025	100%

Pada tahap seleksi data, tidak ditemukan missing value maupun duplikasi pada seluruh 1.025 data awal. Setelah filter rentang waktu 1 tahun dan filter relevansi OPD, diperoleh 635 berita bersih. Setelah pelabelan dan verifikasi manual menggunakan Label Studio, sebanyak 377 berita berlabel valid digunakan dalam proses pelatihan model.

Pelabelan Data

Pelabelan data dilakukan secara semi otomatis menggunakan Label Studio, yaitu platform *open-source* yang digunakan untuk anotasi data. Dari total 635 data berita yang telah melalui tahap pembersihan (*cleaning*), dilakukan proses pelabelan manual dan diperoleh sebanyak 377 artikel berita yang relevan dengan tiga kategori OPD. Sementara itu, sebanyak 258 artikel berita dinyatakan tidak relevan karena tidak termasuk ke dalam ketiga kategori tersebut, sehingga data tersebut dihapus dan tidak digunakan pada tahap selanjutnya.

Setiap artikel berita diberikan dua jenis label, yaitu label kategori OPD dan label sentimen berita. Setelah proses pelabelan selesai, dilakukan pembersihan ulang untuk memastikan konsistensi dan kualitas data yang telah diberi label.



Gambar 2. Tampilan Proses Labeling di Label Studio

Pada Gambar 2. ditampilkan proses *labeling* manual menggunakan Label Studio terhadap data berita yang telah melalui tahap pembersihan sebelumnya. Hasil pelabelan pada Label Studio kemudian diekspor dalam format CSV untuk digunakan pada proses pengolahan dan analisis data selanjutnya.

Distribusi Label Dataset

Distribusi label OPD pada 377 data berlabel disajikan pada Tabel 4. Dinas Ketenagakerjaan memiliki jumlah data terbanyak, mencerminkan tingginya intensitas pemberitaan terkait isu ketenagakerjaan dan pasar kerja di Jawa Tengah.

Tabel 4. Distribusi Label OPD

Kategori OPD	Jumlah Data	Persentase
Dinas Ketenagakerjaan	165	43,8%
Dinas Kesehatan	130	34,5%
Dinas Pendidikan	82	21,8%
Total	377	100%

Distribusi label sentimen pada 377 data berlabel disajikan pada Tabel 5. Dominasi sentimen positif mencerminkan kecenderungan pemberitaan pemerintahan daerah yang lebih banyak menyoroti program dan capaian positif kebijakan pemerintah dibandingkan kritik atau permasalahan.

Tabel 5. Distribusi Label Sentimen

Sentimen	Jumlah Data	Persentase
Positif	177	46,95%
Netral	150	39,79%
Negatif	50	13,26%
Total	377	100%

Encoding Label dan Pembagian Dataset

Label kategorikal kemudian dikonversi menjadi bentuk numerik menggunakan metode label encoding untuk memudahkan proses pengolahan data oleh model machine learning. Proses ini dilakukan dengan memberikan representasi angka pada setiap kategori label OPD.

Tabel 6. Encoding Label OPD

Label OPD	Nilai Encoding
Dinas Pendidikan	0
Dinas Kesehatan	1
Dinas Ketenagakerjaan	2

Tahapan selanjutnya adalah memisahkan data training dan data testing dengan tujuan untuk menghindari overfitting yang akan mengganggu kemampuan generalisasi model IndoBERT. Pembagian data ini juga diperlukan untuk memastikan model yang sudah dibuat dapat mengikuti dengan baik dan memberikan hasil prediksi yang akurat.

Tabel 7. Pembagian Dataset

Jenis Data	Jumlah	Persentase
Data Training	301	80%
Data Testing	76	20%
Total	377	100%

Berdasarkan Tabel 7, dataset dibagi menggunakan rasio 80:20, yaitu sebesar 80% digunakan sebagai data *training* dan 20% digunakan sebagai data *testing*. Dari total 377 data yang tersedia, sebanyak 301 data digunakan untuk proses pelatihan model dan 76 data digunakan untuk proses pengujian model. Rasio pembagian ini umum digunakan dalam penelitian *machine learning* karena dapat memberikan keseimbangan antara proses pembelajaran model dan evaluasi performa model secara optimal.

Pembuatan Model Klasifikasi *Fine-Tuning* IndoBERT

Tahapan ini merupakan proses pembangunan model klasifikasi menggunakan metode fine-tuning pada model IndoBERT. Data yang sebelumnya telah dipersiapkan digunakan sebagai masukan untuk proses analisis klasifikasi. Model IndoBERT dirancang untuk melakukan klasifikasi multiclass terhadap tiga kategori OPD.

Pembuatan model dilakukan menggunakan library Hugging Face Transformers dengan memanfaatkan fungsi `AutoModelForSequenceClassification` dan base model `indobenchmark/indobert-base-p2`. Model ini kemudian ditambahkan `classification layer` untuk menghasilkan prediksi kategori berdasarkan teks berita yang diberikan.

Tabel 8. Arsitektur Model Fine-tuning IndoBERT

Layer	Konfigurasi	Deskripsi
IndoBERT Base	12 Transformer Layers	Pre-trained encoder bahasa Indonesia
Hidden Size	768 units	Dimensi representasi konteks
Dropout	0.1	Mencegah overfitting
Linear Classifier	768 → 3 output	Klasifikasi ke 3 kategori OPD
Softmax	3 classes	Probabilitas per kategori

Berdasarkan Tabel 8, model menggunakan arsitektur IndoBERT base yang terdiri dari 12 transformer layers dengan hidden size sebesar 768. Model juga dilengkapi dengan dropout layer sebesar 0,1 untuk mengurangi risiko overfitting selama proses pelatihan. Pada bagian akhir model terdapat linear classifier yang mengubah representasi hidden state berukuran 768 menjadi tiga kelas keluaran sesuai jumlah kategori OPD. Selanjutnya, fungsi softmax digunakan untuk menghasilkan probabilitas prediksi pada masing-masing kategori.

Tahapan berikutnya adalah menentukan hyperparameter yang digunakan selama proses fine-tuning model IndoBERT.

Tabel 9. Hyperparameter Fine-Tuning IndoBERT

Hyperparameter	Nilai	Keterangan
Learning Rate	2e-5	Kecepatan pembelajaran model
Batch Size	8	Jumlah sampel per batch
Epochs	3	Jumlah iterasi pelatihan
Optimizer	AdamW	Optimasi dengan weight decay
Weight Decay	0.01	Regularisasi L2
Warmup Steps	500	Pemanasan learning rate

Berdasarkan Tabel 4.13, proses pelatihan model menggunakan optimizer AdamW untuk melakukan pembaruan parameter model secara efisien dengan penerapan weight decay sebagai bentuk regularisasi. Nilai learning rate sebesar 2e-5 dipilih karena merupakan nilai yang umum digunakan dan cukup optimal dalam proses fine-tuning model berbasis BERT. Jumlah epoch sebanyak 3 digunakan agar model dapat mencapai konvergensi tanpa menyebabkan overfitting, terutama karena jumlah data training yang relatif terbatas. Selain itu, batch size sebesar 8 dipilih untuk menyesuaikan kapasitas memori GPU yang tersedia selama proses pelatihan model.

Evaluasi Model Klasifikasi IndoBERT

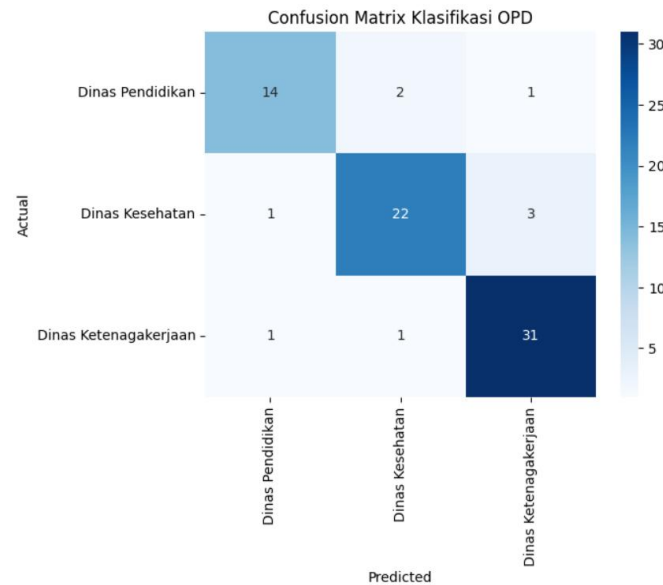
Model fine-tuning IndoBERT mencapai akurasi keseluruhan sebesar 88,16% pada 76 data pengujian. Hasil evaluasi lengkap per kategori OPD disajikan pada Tabel 6. Nilai eval loss sebesar 0,1866 mengindikasikan konvergensi pelatihan yang stabil tanpa tanda-tanda overfitting yang signifikan.

Tabel 10. Hasil Evaluasi Model Klasifikasi IndoBERT

Kategori OPD	Precision	Recall	F1-Score	Support
Dinas Pendidikan	0,88	0,82	0,85	17
Dinas Kesehatan	0,88	0,85	0,86	26
Dinas Ketenagakerjaan	0,89	0,94	0,91	33
Accuracy	-	-	0,88	76
Macro Average	0,88	0,87	0,87	76
Weighted Average	0,88	0,88	0,88	76

Hasil evaluasi per kelas menunjukkan kinerja yang konsisten di ketiga kategori OPD. Dinas Ketenagakerjaan memperoleh F1-score tertinggi sebesar 0,91 dengan recall 0,94, yang berarti model mampu mendeteksi 31 dari 33 berita ketenagakerjaan secara benar. Dinas Kesehatan memperoleh precision 0,88 dengan 22 dari 26 berita diklasifikasikan tepat. Dinas Pendidikan memperoleh F1-score 0,85 dengan precision dan recall yang seimbang.

Analisis confusion matrix pada Gambar 2 memberikan gambaran detail tentang pola kesalahan klasifikasi model. Dari 76 data pengujian, model melakukan 67 prediksi benar (diagonal utama) dan 9 kesalahan prediksi.



Gambar 3. Confusion Matrix Klasifikasi OPD

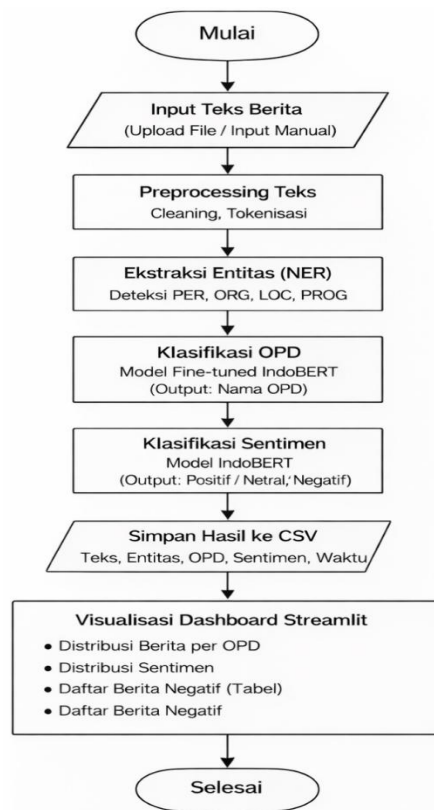
Berdasarkan confusion matrix, kesalahan klasifikasi terbanyak terjadi pada Dinas Kesehatan, di mana 3 berita kesehatan diprediksi sebagai Dinas Ketenagakerjaan dan 1 berita diprediksi sebagai Dinas Pendidikan. Pada kategori Dinas Pendidikan, 2 berita diprediksi sebagai Dinas Kesehatan dan 1 berita sebagai Dinas Ketenagakerjaan. Pola kesalahan ini umumnya terjadi pada berita lintas-sektoral yang membahas lebih dari satu domain OPD secara bersamaan, misalnya berita tentang tenaga kesehatan yang juga menyentuh aspek vokasi.

Analisis Sentimen per OPD

Analisis distribusi per OPD disajikan pada Tabel 8. Dinas Kesehatan memiliki proporsi tertinggi sebesar 21,5%, jauh di atas rata-rata keseluruhan (13,26%). Hal ini mengindikasikan banyaknya pemberitaan kritis terkait layanan kesehatan, distribusi tenaga kesehatan, dan akses fasilitas kesehatan di wilayah Jawa Tengah. Sebaliknya, Dinas Ketenagakerjaan memiliki proporsi terendah (7,3%), yang mencerminkan respons positif masyarakat terhadap program pelatihan kerja dan penyerapan tenaga kerja.

Tabel 11. Distribusi Sentimen per Kategori OPD

Kategori OPD	Positif	Netral	Negatif	Total
Dinas Kesehatan	65 (50,0%)	37 (28,5%)	28 (21,5%)	130
Dinas Ketenagakerjaan	75 (45,5%)	78 (47,3%)	12 (7,3%)	165
Dinas Pendidikan	37 (45,1%)	35 (42,7%)	10 (12,2%)	82
Total	177 (46,95%)	150 (39,79%)	50 (13,26%)	377

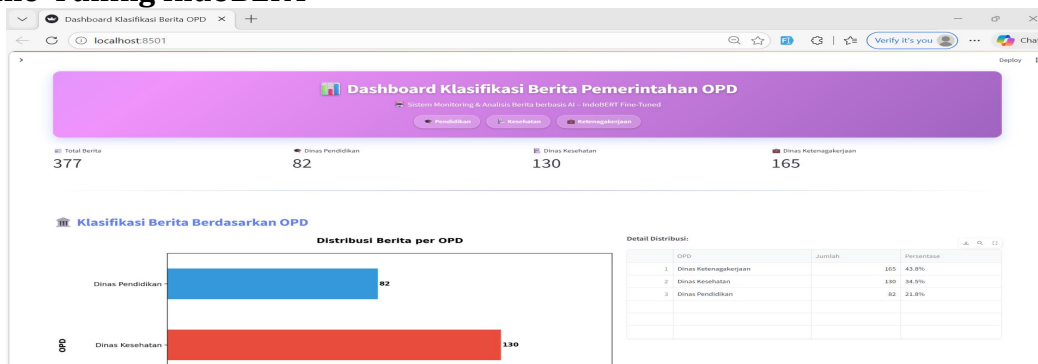


Gambar 4. Flowchart Alur Aplikasi

Gambar 4 menunjukkan alur proses aplikasi dashboard klasifikasi berita berbasis Streamlit. Proses dimulai dari input data berita hasil web scraping yang kemudian dilakukan tahap pre-processing, seperti case folding, cleaning, tokenisasi, dan stemming untuk menghasilkan teks yang lebih bersih dan terstandarisasi. Setelah itu, sistem melakukan ekstraksi Named Entity Recognition (NER) untuk mengidentifikasi entitas penting dalam berita.

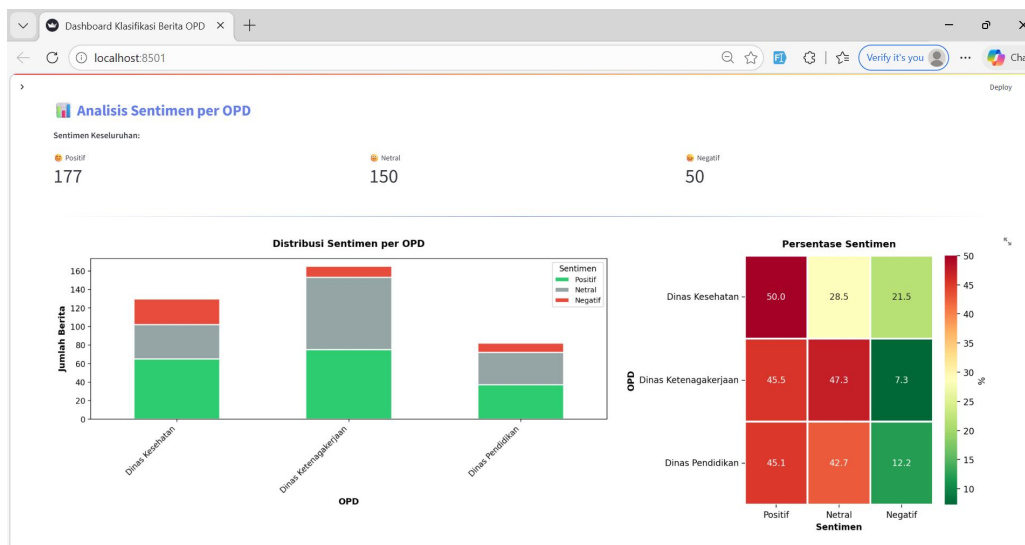
Selanjutnya, berita diproses menggunakan model IndoBERT untuk klasifikasi OPD dan analisis sentimen. Hasil klasifikasi kemudian disimpan dalam file CSV terstruktur dan ditampilkan melalui dashboard interaktif Streamlit yang memudahkan pengguna dalam melihat, memfilter, dan menganalisis informasi berita.

Implementasi Dashboard Ekstraksi dan Klasifikasi Berita Menggunakan Fine-Tuning IndoBERT



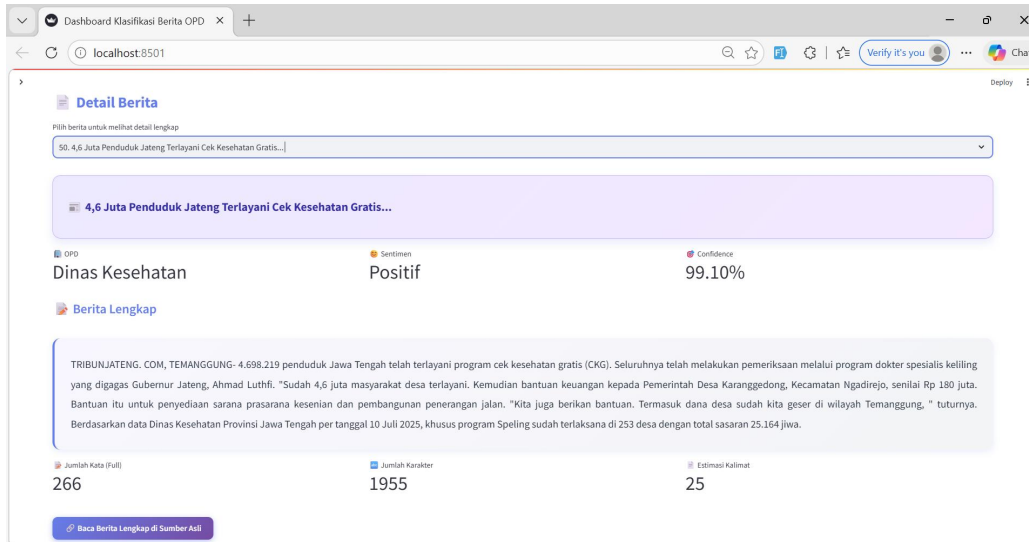
Gambar 5. Dashboard Klasifikasi Berita Pemerintahan

Gambar 5. merupakan implementasi desain halaman utama dashboard klasifikasi berita OPD. Pada bagian atas ditampilkan metrik ringkasan berupa kartu-kartu informatif yang menunjukkan Total Berita sebanyak 377, Dinas Pendidikan sebanyak 82, Dinas Kesehatan sebanyak 130, dan Dinas Ketenagakerjaan sebanyak 165. Di bawahnya terdapat visualisasi berupa horizontal *bar chart* yang menampilkan distribusi berita per OPD secara grafis. Di sisi kanan *chart* ditampilkan tabel detail distribusi yang menunjukkan urutan dari tertinggi yaitu Dinas Ketenagakerjaan (165, 43,8%), Dinas Kesehatan (130, 34,5%), dan Dinas Pendidikan (82, 21,8%).



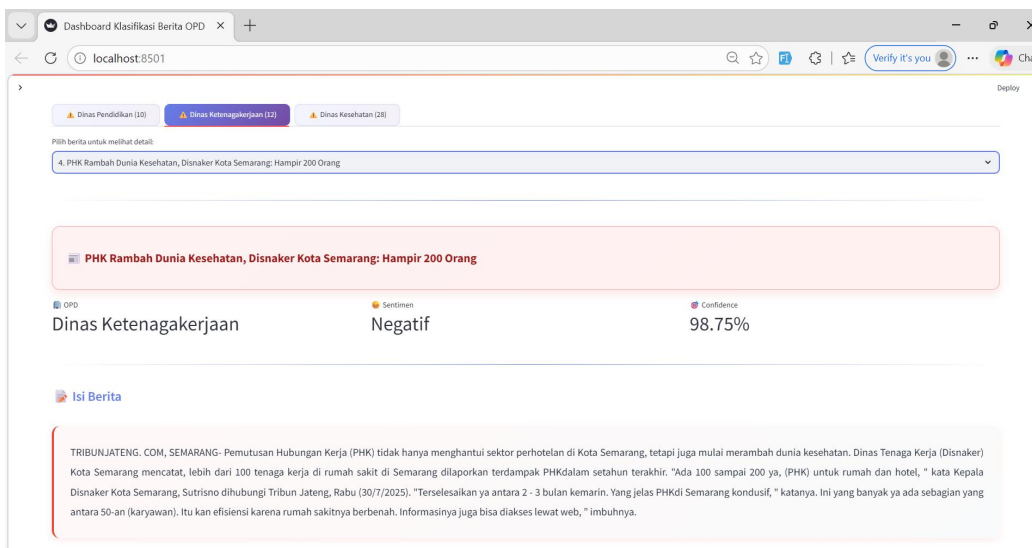
Gambar 6. Hasil Analisis Sentiment

Pada Gambar 6. ditampilkan halaman Analisis Sentimen per OPD yang memuat ringkasan jumlah sentimen, yaitu positif sebanyak 177, netral 150, dan negatif 50. Visualisasi terdiri dari stacked bar chart untuk distribusi sentimen tiap OPD dan heatmap persentase sentimen. Hasil menunjukkan bahwa Dinas Kesehatan memiliki persentase sentimen positif tertinggi sebesar 50,0%, sedangkan Dinas Ketenagakerjaan memiliki sentimen negatif terendah sebesar 7,3%. Secara keseluruhan, mayoritas pemberitaan terkait OPD didominasi oleh sentimen positif dan netral, sehingga visualisasi ini membantu memahami persepsi publik terhadap kinerja OPD.



Gambar 7. Detail Berita

Pada Gambar 7. ditampilkan halaman Detail Berita yang memungkinkan pengguna memilih berita tertentu untuk melihat informasi lengkap. Halaman ini menampilkan judul berita, hasil prediksi kategori OPD, sentimen, confidence score, serta isi berita lengkap hasil ekstraksi NER. Selain itu, ditampilkan juga statistik teks seperti jumlah kata, jumlah karakter, dan estimasi jumlah kalimat pada berita.



Gambar 8. Fokus Berita Negatif

Pada Gambar 8. ditampilkan halaman Fokus Berita Negatif yang digunakan untuk memantau berita dengan sentimen negatif sebagai bentuk early warning bagi pemerintah daerah. Halaman ini menampilkan jumlah berita negatif sebanyak 50 berita beserta distribusinya pada masing-masing OPD. Pengguna juga dapat memilih berita tertentu untuk melihat detail seperti kategori OPD, sentimen, confidence score, dan isi berita.

Konten berita yang ditampilkan merupakan hasil ekstraksi informasi menggunakan metode Named Entity Recognition (NER), sehingga hanya

menampilkan informasi penting dan entitas utama seperti nama, organisasi, lokasi, dan program. Pendekatan ini membantu pengguna memahami inti permasalahan secara cepat tanpa harus membaca keseluruhan berita.

Hasil penelitian ini menunjukkan bahwa pendekatan transfer learning melalui fine-tuning IndoBERT efektif untuk klasifikasi berita pemerintahan berbahasa Indonesia. Akurasi 88,16% yang dicapai dengan hanya 301 data pelatihan membuktikan kekuatan representasi kontekstual yang telah dipelajari model selama fase pretraining pada korpus bahasa Indonesia berskala besar. Hasil ini konsisten dengan temuan penelitian-penelitian terdahulu yang menggunakan IndoBERT untuk tugas klasifikasi teks berbahasa Indonesia [12][13][6][15].

Dibandingkan dengan penelitian Tobing et al. [6] yang mencapai akurasi 94,1% pada deteksi hoaks politik dengan 2.604 data seimbang, penelitian ini menggunakan dataset yang jauh lebih kecil (377 data) namun menangani masalah yang lebih kompleks—klasifikasi multiclass tiga OPD dengan distribusi kelas yang tidak seimbang, ditambah dengan dimensi analisis sentimen dan ekstraksi entitas. Perbedaan akurasi sekitar 6% tersebut dapat dijelaskan oleh dua faktor utama: 1. ketidakseimbangan distribusi kelas OPD (43,8% : 34,5% : 21,8%), dan 2. ambiguitas inheren pada berita lintas-sektoral yang melibatkan lebih dari satu domain OPD. Meskipun demikian, akurasi 88,16% yang dicapai sudah melampaui target minimum yang ditetapkan dalam penelitian ini (85%) dan menunjukkan kinerja yang memuaskan untuk aplikasi praktis.

Nilai precision macro 0,88 yang seimbang di ketiga kelas menunjukkan bahwa model tidak memiliki bias prediksi yang signifikan terhadap kelas mayoritas. Ini merupakan indikator penting bahwa model dapat diandalkan untuk digunakan dalam kondisi data dunia nyata yang memiliki distribusi tidak seimbang. Recall macro 0,87 menunjukkan bahwa model secara konsisten berhasil mengidentifikasi sebagian besar berita pada setiap kategori OPD.

Dari perspektif implementasi, dashboard Streamlit yang dikembangkan memberikan nilai praktis yang signifikan bagi OPD. Fitur visualisasi distribusi berita per OPD memungkinkan pemantauan tren pemberitaan secara real-time. Fitur fokus berita negatif berfungsi sebagai early warning system yang memungkinkan pemangku kebijakan mengidentifikasi isu kritis secara proaktif sebelum berkembang menjadi krisis komunikasi. Fitur filter dan eksplorasi data memungkinkan analisis mendalam berdasarkan kombinasi kategori OPD, sentimen, sumber media, dan rentang waktu. Hal ini sejalan dengan tujuan penerapan Evidence-Based Policy yang menekankan pentingnya data terstruktur sebagai dasar pengambilan keputusan [5].

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, dataset yang digunakan terbatas pada 377 sampel berlabel dari empat sumber media, sehingga representativitas terhadap seluruh pemberitaan pemerintahan daerah Jawa Tengah masih dapat ditingkatkan. Dan kedua, kategori OPD dibatasi pada tiga dinas saja, sementara dalam praktiknya terdapat puluhan OPD yang beroperasi di

tingkat provinsi maupun kabupaten atau kota. Keterbatasan-keterbatasan ini menjadi arah pengembangan yang jelas untuk penelitian selanjutnya.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, diperoleh beberapa kesimpulan sebagai berikut:

1. Model IndoBERT yang dilakukan fine-tuning untuk tugas Named Entity Recognition (NER) mampu melakukan ekstraksi informasi secara otomatis dari berita pemerintahan daerah. Model berhasil mengidentifikasi entitas penting seperti organisasi (organization), lokasi (location), pejabat (person), dan program (program), sehingga informasi dalam berita dapat dipahami secara lebih terstruktur.
2. Model IndoBERT mampu mengklasifikasikan berita pemerintahan daerah ke dalam kategori OPD dengan sangat baik, dengan memperoleh nilai akurasi sebesar 88,16%. Hasil tersebut menunjukkan bahwa pendekatan transfer learning menggunakan IndoBERT efektif diterapkan pada tugas klasifikasi teks berbahasa Indonesia.
3. Berdasarkan hasil confusion matrix, model mengalami kesalahan klasifikasi pada 9 dari 76 data testing. Kesalahan tersebut umumnya terjadi pada berita yang bersifat lintas sektoral dan melibatkan lebih dari satu OPD. Meskipun demikian, model tetap menunjukkan performa yang baik dalam menangani kompleksitas data berita pemerintahan daerah.
4. Hasil evaluasi menunjukkan bahwa kategori Dinas Ketenagakerjaan memiliki nilai recall tertinggi sebesar 100%, sedangkan kategori Dinas Kesehatan memperoleh nilai precision sebesar 100%. Hal ini menunjukkan bahwa setiap kategori OPD memiliki karakteristik linguistik yang berbeda sehingga memengaruhi performa klasifikasi model.
5. Hasil analisis sentimen menunjukkan bahwa sebanyak 65% berita memiliki sentimen positif. Selain itu, Dinas Kesehatan menjadi kategori OPD dengan jumlah sentimen negatif terbanyak, yaitu sebanyak 28 dari total 50 berita negatif. Temuan ini dapat memberikan insight penting dalam mengidentifikasi isu prioritas pada pemerintahan daerah.

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran untuk pengembangan penelitian selanjutnya. Penelitian berikutnya diharapkan dapat memperluas dataset training dengan menambahkan variasi topik berita, sumber media, serta rentang waktu yang lebih panjang agar model memiliki kemampuan generalisasi yang lebih baik. Selain itu, penggunaan model yang lebih ringan seperti DistilBERT dapat dieksplorasi untuk mengurangi computational cost tanpa menurunkan performa model secara signifikan sehingga lebih efisien pada proses deployment.

Proses fine-tuning model juga dapat dikembangkan lebih lanjut melalui optimasi hyperparameter seperti learning rate, batch size, jumlah epoch, dan dropout rate untuk memperoleh konfigurasi model yang lebih optimal sesuai karakteristik data berita pemerintahan daerah. Selain itu, kategori OPD yang digunakan dalam klasifikasi dapat diperluas dengan menambahkan instansi lain seperti Dinas Perhubungan, Dinas Sosial, dan Bappeda agar sistem yang dikembangkan menjadi lebih komprehensif dan relevan dengan kebutuhan pemerintah daerah.

DAFTAR PUSTAKA

- [1] D. A. N. Berkesinambungan, "Rencana pembangunan jangka menengah nasional 2020-2024," 2024.
- [2] A. Davina, M. Putri, N. Sulistianingsih, and R. Rismayati, "JTIM : Jurnal Teknologi Informasi dan Multimedia Pengaruh Teknik Representasi Teks Bag-of-Words dan TF-IDF," vol. 7, no. 4, pp. 675–688, 2025.
- [3] N. Hafidhigbal, *IMPLEMENTASI MODEL INDOBERT UNTUK KLASIFIKASI JENIS TINDAK PIDANA KORUPSI KLASIFIKASI JENIS TINDAK PIDANA KORUPSI*. 2024.
- [4] M. Deagama, S. Antariksa, A. Sugiharto, and B. Surarso, "Jurnal resti," vol. 9, no. 4, pp. 754–764, 2025.
- [5] S. Sagita, "Peran Indeks Transformasi Digital Nasional sebagai Instrumen Evidence Based Policy di Indonesia Pendahuluan," vol. 5, no. 2, pp. 47–59, 2025.
- [6] C. J. L. Tobing, I. G. N. L. Wijayakusuma, L. Putu, I. Harini, and U. Udayana, "Deteksi Berita Hoax Politik Menggunakan Fine-Tuning IndoBERT," vol. 9, no. 2, pp. 354–360, 2025.
- [7] I. A. Oktariansyah, F. R. Umbara, and F. Kasyidi, "Klasifikasi Sentimen Untuk Mengetahui Kecenderungan Politik Pengguna X Pada Calon Presiden Indonesia 2024 Menggunakan Metode IndoBert," vol. 6, no. 2, pp. 636–648, 2024.
- [8] A. S. Rizky and E. Y. Hidayat, "Emotion Classification in Indonesian Text Using IndoBERT," vol. x, no. x, 2022.
- [9] H. F. Karim and A. P. Wibowo, "Kinerja Metode Fine-Tuning IndoBERT untuk Klasifikasi Emosi Multi-Kelas pada Teks Informal Bahasa Indonesia," vol. 6, no. 1, pp. 63–74, 2025, doi: 10.47065/bulletincsr.v6i1.850.
- [10] F. Koto and A. Rahimi, "Machine Translated by Google IndoLEM dan IndoBERT : Dataset Benchmark dan Dataset Pra-Pelatihan Model Bahasa untuk NLP Indonesia 1Universitas Melbourne 2Universitas Queensland Timothy Baldwin1 Machine Translated by Google," pp. 8–13, 2020.
- [11] B. Juarto, "INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Indonesian News Classification Using IndoBert," pp. 0–2, 2023.
- [12] A. Simanjuntak *et al.*, "Studi dan Analisis Hyperparameter Tuning IndoBERT Dalam Pendeteksian Berita Palsu," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 13, no. 1, pp. 60–67, 2024.
- [13] A. Kunaefi, Z. Abidin, and R. Kusumawati, "Klasifikasi berita hoaks bahasa indonesia menggunakan indobert," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 2, pp. 1706–1714, 2025.
- [14] C. J. L. Tobing, I. G. N. L. Wijayakusuma, L. Putu, I. Harini, and U. Udayana, "Detection of Political Hoax News Using Fine-Tuning IndoBERT," vol. 9, no. 2, pp. 354–360, 2025.
- [15] T. Iskandar, Z. Maulana, and A. F. Bukhori, "Model Klasifikasi Berbasis Multiclass Classification dengan Kombinasi Indobert Embedding dan Long Short-Term Memory untuk Tweet Berbahasa Indonesia (Classification Model Based on Multiclass Classification with a Combination of Indobert Embedding and Long Short-Term Memory for Indonesian-language Tweets)," vol. 1, no. 1, pp. 1–28, 2022.

- [16] I. Budi and R. R. Suryono, “Application of named entity recognition method for Indonesian datasets : a review,” vol. 12, no. 2, pp. 969–978, 2023, doi: 10.11591/eei.v12i2.4529.
- [17] R. Merdiansah and A. A. Ridha, “Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT,” vol. 7, pp. 221–228, 2024.
- [18] D. Nuryadi *et al.*, “FINE TUNING INDOBERT UNTUK ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI TIKET . COM DI GOOGLE PLAY STORE,” vol. 9, no. 2, pp. 3577–3583, 2025.
- [19] D. Ramdani, M. Irfan, and N. Lukman, “Perbandingan Kinerja IndoBERT , IndoRoBERTa dan NusaBERT dalam Analisis Sentimen Isu LGBT di Media Sosial X,” vol. 8, no. 2, pp. 91–107, 2025.
- [20] J. Federico, I. D. Made, and B. Atmaja, “Klasifikasi Berita Berdasarkan Kategori Menggunakan Convolutional Neural Network dengan IndoBERT,” vol. 3, pp. 899–906, 2025.