

**EKSTRAKSI INFORMASI PADA TEKS BERITA *GAME*
BERBAHASA INDONESIA DENGAN INDOBERT**

TESIS

**Oleh:
FIRMANSYAH REKSO WIBOWO
NIM. 230605210004**



**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**EKSTRAKSI INFORMASI PADA TEKS BERITA *GAME*
BERBAHASA INDONESIA DENGAN INDOBERT**

TESIS

Diajukan kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)

Oleh:
FIRMANSYAH REKSO WIBOWO
NIM. 230605210004

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025

HALAMAN PERSETUJUAN

EKSTRAKSI INFORMASI PADA TEKS BERITA GAME BERBAHASA INDONESIA DENGAN INDOBERT

TESIS

Oleh:
FIRMANSYAH REKSO WIBOWO
NIM. 230605210004

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 12 November 2025

Pembimbing I,



Dr. Zainal Abidin, M.Kom
NIP. 19760613 200501 1 004

Pembimbing II,



Dr. Ririen Kusumawati, S.Si, M.Kom
NIP. 19720309 200501 2 002

Mengetahui,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

HALAMAN PENGESAHAN

EKSTRAKSI INFORMASI PADA TEKS BERITA *GAME* BERBAHASA INDONESIA DENGAN INDOBERT

TESIS

Oleh:
FIRMANSYAH REKSO WIBOWO
NIM. 230605210004

Telah Dipertahankan di Depan Dewan Pengaji Tesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)

Tanggal: 12 DESEMBER 2025

Pengaji I : Dr. Ir. Yunifa Miftachul Arif, M.T
NIP. 19830616 201101 1 004

()

Pengaji II : Dr. Irwan Budi Santoso M. Kom
NIP. 19770103 201101 1 004

()

Pembimbing I : Dr. Zainal Abidin, M.Kom
NIP. 19760613 200501 1 004

()

Pembimbing II : Dr. Ririen Kusumawati, S.Si, M.Kom
NIP. 19720309 200501 2 002

()

Mengetahui dan Mengesahkan,
Ketua Program Studi Magister Informatika
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Firmansyah Rekso Wibowo
NIM : 230605210004
Program Studi : Magister Informatika
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Tesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan tesis ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 19 Desember 2025
Yang membuat pernyataan,



FIRMANSYAH REKSO WIBOWO
NIM. 230605210003

MOTTO

“Belajarlah seolah hidup selamanya, beramallah seolah mati esok hari”

HALAMAN PERSEMBAHAN

Dengan segenap rasa syukur *Alhamdulillahi rabbil 'alamin*, saya persembahkan karya Tesis ini kepada:

1. Orang-orang terkasih dalam hidup saya: ibuku Tutik Wahyuningtyas, adik Sinta, mbak Fifit, mas Bhekti, serta keponakan Sila dan Jena. Juga untuk ayahanda tercinta (Alm. Muhammad Rodhi). Terima kasih atas doa dan kasih sayang yang selalu menjadi kekuatan terbesar bagi saya..
2. Bapak dan Ibu pembimbing: Dr. Zainal Abidin, M.Kom dan Dr. Ririen Kusumawati, S.Si., M.Kom, serta para penguji: Dr. Ir. Yunifa Miftachul Arif, M.T dan Dr. Irwan Budi Santoso, M.Kom. Terima kasih atas bimbingan, ilmu, dan arahan yang telah membentuk pemahaman akademik saya hingga penelitian ini dapat terselesaikan dengan baik.
3. Teman-teman seperjuangan di grup WhatsApp Tesis Angkatan 8. Terima kasih atas dukungan, diskusi, dan semangat yang selalu menguatkan di tengah proses yang panjang ini.
4. Para sahabat yang mewarnai perjalanan hidup saya: sohib Naditira, sohib SMA Bryan dan Kempong, *support partner* saya, M.A.F, serta rekan-rekan band Toxictoast dan Getz!. Terima kasih atas kebersamaan, tawa, musik, dan cerita yang memberi warna sepanjang perjalanan ini.

Semoga Allah senantiasa membimbing kita ke jalan yang lurus dan memudahkan setiap ikhtiar dalam mewujudkan cita-cita. *Aamiin ya Rabbal 'Alamiin.*

KATA PENGANTAR

Assalamualaikum warahmatullahi wabarakatuh.

Puji syukur penulis panjatkan ke hadirat Allah Swt. atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tesis ini dengan baik dan tepat waktu. Tesis yang berjudul “Ekstraksi Informasi Pada Teks Berita *Game* Berbahasa Indonesia Dengan IndoBERT” ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Magister Informatika.

Dalam proses penyusunan tesis ini, penulis menyadari bahwa banyak pihak yang telah memberikan bantuan, bimbingan, doa, serta dukungan baik secara moral maupun material. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Dr. Zainal Abidin, M.Kom dan Dr. Ririen Kusumawati, S.Si., M.Kom selaku Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan wawasan yang sangat berharga selama proses penelitian hingga tesis ini dapat terselesaikan dengan baik.
2. Dr. Ir. Yunifa Miftachul Arif, M.T dan Dr. Irwan Budi Santoso, M.Kom selaku Dosen Pengaji yang telah memberikan banyak masukan, kritik, dan saran konstruktif untuk penyempurnaan penelitian ini.
3. Seluruh dosen dan jajaran staf Program Studi Magister Informatika yang telah memberikan ilmu, bantuan, dan dukungan selama masa studi penulis.
4. Keluarga tercinta: ibuku Tutik Wahyuningtyas, adik Sinta, mbak Fifit, mas Bhekti, keponakan Sila dan Jena, serta ayahanda tercinta (Alm. Muhammad Rodhi). Terima kasih atas doa, kasih sayang, dan dukungan yang menjadi kekuatan terbesar bagi penulis selama menyelesaikan tesis ini.

5. Teman-teman seperjuangan di grup WhatsApp Tesis Angkatan 8 yang telah menjadi rekan diskusi dan sumber semangat dalam proses panjang penulisan tesis ini.
6. Sahabat-sahabat terbaik: sohib Naditira, sohib SMA Bryan dan Kempong, serta rekan band Toxictoast dan Getz!. Terima kasih atas kebersamaan, cerita, dan dukungan di sela-sela perjalanan akademik ini.
7. *Support partner* penulis, M.A.F, atas doa, dukungan, dan kekuatan yang selalu diberikan dalam setiap langkah.
8. Serta seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan kontribusi, bantuan, dan dukungan selama penelitian dan penyusunan tesis ini.

Akhir kata, penulis menyadari bahwa tesis ini masih memiliki banyak kekurangan. Kritik dan saran yang membangun sangat penulis harapkan untuk penyempurnaan karya selanjutnya. Semoga tesis ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan bagi para pembaca.

Wassalamualaikum warahmatullahi wabarakatuh.

Malang, 20 Desember 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiii
ABSTRAK	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Pernyataan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Batasan Masalah	4
1.5 Manfaat Penelitian	4
BAB II KAJIAN PUSTAKA.....	5
2.1 Entitas Penting Industri <i>Game</i> Indonesia.....	5
2.2 Analisis Literatur Penelitian Terkait <i>Game</i>	7
2.3 Pemanfaatan Teks Berita <i>Online</i> pada Industri <i>Game</i>	9
2.4 Model Bahasa Berbasis Transformer untuk Ekstraksi Informasi	9
2.5 Kerangka Teori	11
BAB III METODOLOGI PENELITIAN	17
3.1 Kerangka Konsep.....	17
3.2 Prosedur Penelitian	18
3.3 Persiapan Data dan Konstruksi Dataset	19
3.3.1 Pengumpulan Data (Data Collection)	19
3.3.2 Pra-pemrosesan Artikel	22
3.3.3 Kurasi Kalimat	23
3.3.4 Anotasi Kalimat.....	25
3.3.5 Pembagian Dataset	26
3.4 Desain Sistem dan Model	27
3.4.1 Arsitektur Transformer.....	28
3.4.2 BERT.....	33
3.4.3 IndoBERT	34
3.4.4 mBERT (Multilingual BERT).....	36
3.4.5 XLM-RoBERTa	38
3.5 Pelatihan dan Evaluasi Model.....	41
3.5.1 Pelatihan Model.....	41
3.5.2 Evaluasi Model.....	42
BAB IV ANOTASI DATASET	44
4.1 Tahap Akuisisi dan Pra-pemrosesan Data Mentah	44
4.1.1 Pengumpulan Artikel dari Sumber Berita	45

4.1.2 Pra-pemrosesan Teks Artikel	46
4.2 Seleksi dan Kurasi Kalimat Artikel	46
4.3 Skema Anotasi Entitas	48
4.3.1 Definisi Tipe Entitas.....	49
4.3.2 Skema Penandaan IOB2.....	50
4.3.3 Perangkat Anotasi	51
4.4 Pelaksanaan dan Validasi Proses Anotasi	52
4.4.1 Ambiguitas Entitas	53
4.5 Karakteristik dan Statistik Final Dataset GameNER-ID.....	54
4.5.1 Statistik Pembagian Dataset.....	55
4.5.2 Distribusi Label dan Entitas	56
4.5.3 Hasil Anotasi	58
BAB V MODEL INDOBERT	63
5.1 Arsitektur Model	63
5.2 Konfigurasi Pelatihan dan Hyperparameter	64
5.3 Metrik Evaluasi Kerja	66
5.4 Hasil Pelatihan Model.....	67
5.5 Analisa Kesalahan dan Perilaku Model	70
BAB VI MODEL NER BERBASIS TRANSFORMER	82
6.1 Model Transformer	82
6.2 Hasil Eksperimen dan Analisis Kinerja	83
6.2.1 Hasil Komparatif Kinerja Model.....	84
6.2.2 Analisis Perbandingan Hasil	86
6.3 Analisis Kesalahan Model pada Tugas NER Domain <i>Game</i>	88
6.3.1 Entitas Berkinerja Tinggi	88
6.3.2 Entitas Bermasalah	90
6.3.3 Pola Kesalahan pada Model Berkinerja Tinggi.....	91
6.3.4 Pola Kesalahan pada Model Berkinerja Rendah	94
BAB VII PEMBAHASAN	99
7.1 GameNER-ID di antara Dataset NER Indonesia	99
7.2 Implikasi Penelitian.....	102
7.3 Keterbatasan Penelitian.....	103
7.4 Arah Penelitian Selanjutnya.....	104
7.5 Integrasi Nilai Islam dalam Validasi Informasi	105
BAB VIII KESIMPULAN	108
8.1 Kesimpulan	108
8.2 Saran.....	109
DAFTAR PUSTAKA	112
LAMPIRAN.....	119

DAFTAR GAMBAR

Gambar 2.1 Kerangka Teori.....	12
Gambar 3.1 Prosedur Penelitian.....	19
Gambar 3.2 Desain Sistem <i>Crawling</i> Data	20
Gambar 3.3 Desain Sistem.....	28
Gambar 3.4 Arsitektur Model Transformer (Vaswani dkk., 2017).....	29
Gambar 3.5 SDPA dan MHA (Vaswani dkk., 2017).....	30
Gambar 3.6 Prosedur Pra-pelatihan dan Fine-tuning BERT (Devlin dkk., 2019)	33
Gambar 3.7 Representasi Input BERT (Devlin dkk., 2019)	34
Gambar 3.8 Hubungan antara Kinerja <i>Zero-Shot</i> mBERT (Wu & Dredze (2019))	37
Gambar 3.9 Ilustrasi Pra-pelatihan XLM-R (Lample & Conneau, 2019).....	39
Gambar 3.10 Wikipedia dan <i>Common Crawl</i> (Conneau dkk., 2019).	40
Gambar 4.1 Proses Anotasi Pada Label Studio.....	51
Gambar 4.2 Distribusi Label Entitas pada Set Data Latih, Validasi, dan Uji	56
Gambar 5.1 Tampilan Antarmuka Google Colab	65
Gambar 5.2 Proses Pelatihan Dataset GameNER-ID	67
Gambar 5.3 Metrik Kinerja Per Entitas pada Dataset Uji.....	70
Gambar 5.4 Confusion Matrix GameNER-ID (Test Set).....	78
Gambar 6.1 Visualisasi Grafik Perbandingan <i>F1-Score</i> Untuk Setiap Model.....	85
Gambar 6.2 Perbandingan Skor F1 antar Model NER Bahasa Indonesia.....	87
Gambar 6.3 Presisi, <i>Recall</i> dan <i>F1-Score</i> per Entitas	88
Gambar 6.4 Jumlah Data dengan <i>F1-Score</i> per Entitas	89
Gambar 6.5 Performa per Tag (<i>True Positive</i> , Presisi <i>Recall</i>).....	90
Gambar 6.6 <i>Confusion Matrix</i> XLM-RoBERTa-large	92
Gambar 6.7 <i>Confusion Matrix</i> IndoBERT-large-p2	93
Gambar 6.8 <i>Confusion Matrix</i> IndoBERT-large-p2	94
Gambar 6.9 <i>Confusion Matrix</i> XLM-RoBERTa-base	95
Gambar 6.10 <i>Confusion Matrix</i> IndoBERT-uncased.....	97
Gambar 7.1 Perbandingan Jumlah Entitas	101
Gambar 7.2 Perbandingan Jumlah Kalimat	101
Gambar 7.3 Perbandingan <i>F1-Score</i> Terbaik.....	102

DAFTAR TABEL

Tabel 2.1 Literatur <i>Review</i> Terkait Tantangan Analisa Industri <i>Game</i>	6
Tabel 2.2 Literatur <i>Review</i> Penelitian Terkait <i>Game</i>	8
Tabel 2.3 Performa Setiap Model Kerangka Teori	13
Tabel 3.1 Data Artikel Berita <i>Online Game</i>	21
Tabel 3.2 Langkah-Langkah Pra-pemrosesan Artikel Berita.....	23
Tabel 3.3 Contoh Proses Kurasi Kalimat.....	24
Tabel 3.4 Contoh Anotasi dengan Format IOB2	25
Tabel 3.5 Pembagian dan Fungsi Dataset	26
Tabel 3.6 Perbandingan Tipe Lapisan (Vaswani dkk., 2017)	32
Tabel 3.7 Hasil Model pada POS Tagging dan NER (Koto dkk., 2020)	35
Tabel 3.8 mBERT Terhadap Baseline (Wu & Dredze, 2019)	37
Tabel 3.9 Perbandingan Kinerja Lintas-Bahasa XNLI	40
Tabel 3.10 Konfigurasi Hyperparameter untuk Pelatihan Model	41
Tabel 3.11 Pembagian dan Fungsi Dataset	42
Tabel 4.1 Ringkasan Sumber Data Awal	45
Tabel 4.2 Pembersihan Data Artikel (Pra-pemrosesan).....	46
Tabel 4.3 Kriteria Kurasi Kalimat Pada Artikel Terkait	47
Tabel 4.4 Jumlah Data (Artikel) Setelah Kurasi	47
Tabel 4.5 Nama dan Deskripsi Entitas dalam GameNER-ID	49
Tabel 4.6 Anotasi dalam Format XML dan IOB2	50
Tabel 4.7 Hasil Validasi Anotasi untuk Entitas Ambigu	54
Tabel 4.8 Hasil Pembagian Dataset Menjadi Data Latih, Validasi dan Uji	56
Tabel 4.9 Distribusi Entitas pada Setiap Bagian Dataset.....	57
Tabel 4.10 Distribusi Label Rinci per Split dengan Format IOB2	58
Tabel 4.11 Hasil Anotasi pada Teks Berita Umum.....	58
Tabel 5.1 Parameter Baseline Model	64
Tabel 5.2 Konfigurasi <i>Hyperparameter</i> untuk Pelatihan Model	64
Tabel 5.3 Kasus Pada Perhitungan Metrik Evaluasi NER	66
Tabel 5.4 Metrik Hasil Pelatihan <i>Tain</i> Dataset GameNER-ID	68
Tabel 5.5 Metrik Kinerja Per Entitas pada Dataset Uji.....	69
Tabel 5.6 Hasil Pelabelan Dataset Train	70
Tabel 5.7 Kesalahan Pada Pelabelan Data Uji	75
Tabel 5.8 Rangkuman Kinerja Per Kelas	79
Tabel 6.1 Kelompok Model yang Digunakan dalam Eksperimen	82
Tabel 6.2 Hasil Perbandingan Kinerja Model pada Dataset Uji GameNER-ID ...	84
Tabel 6.3 Hasil Perbandingan Ukuran Model dan Kinerja	86
Tabel 6.4 Metrik Performa Per Kelas	89
Tabel 6.5 Metrik Performa Per Kelas (Problematik)	91
Tabel 6.6 Kesalahan Prediksi Paling Banyak	91
Tabel 6.7 Kesalahan pada Model Berkinerja Tinggi	92
Tabel 6.8 Kesalahan pada Model Berkinerja Rendah (XLM-RoBERTa-base)....	96
Tabel 6.9 Kesalahan pada Model Berkinerja Rendah (IndoBERT-uncased)	98
Tabel 7.1 Posisi GameNER-ID dengan Dataset NER Indonesia	100
Tabel 7.2 Perbandingan Jenis Entitas di Seluruh Dataset NER Indonesia	100

ABSTRAK

Wibowo, Firmansyah Rekso. 2025. **Ekstraksi Informasi pada Teks Berita Game Berbahasa Indonesia dengan IndoBERT**. Tesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Zainal Abidin, M.Kom (II) Dr. Ririen Kusumawati, S.Si, M.Kom.

Kata kunci: *Information Extraction, Named Entity Recognition (NER), Natural Language Processing (NLP), Transformer, IndoBERT, Berita Game, GameNER-ID*

Industri *game* Indonesia yang berkembang pesat menghasilkan volume besar teks berita daring yang tidak terstruktur, namun pemanfaatannya sebagai sumber informasi strategis masih terbatas akibat ketiadaan sumber daya *Natural Language Processing* (NLP) khusus domain serta rendahnya kemampuan model eksisting dalam menangani terminologi teknis dan fenomena bahasa campuran (*code-switching*). Penelitian ini membangun GameNER-ID, *dataset Named Entity Recognition* (NER) pertama untuk domain industri *game* berbahasa Indonesia, serta mengevaluasi efektivitas model Transformer dalam melakukan ekstraksi informasi dari teks tersebut. Proses konstruksi *dataset* mencakup pengumpulan artikel berita, standardisasi format, kurasi manual, dan anotasi 3.351 kalimat menggunakan sepuluh tipe entitas domain-spesifik dengan skema IOB2, sehingga menghasilkan korpus terstruktur yang representatif dan layak digunakan sebagai tolok ukur pengembangan NLP pada sektor *game*. Evaluasi performa model dilakukan melalui *fine-tuning* beberapa rsitektur Transformer dan pengukuran metrik *Precision*, *Recall*, serta *F1-Score*. Hasil eksperimen menunjukkan bahwa IndoBERT mengekstraksi informasi dengan baik melalui *F1-Score* 0.8175. Analisis komparatif menunjukkan bahwa model multilingual XLM-RoBERTa-large mencapai kinerja lebih tinggi dengan *F1-Score* 0.8386. Temuan ini menegaskan bahwa pendekatan lintas-bahasa lebih efektif dalam menangani karakteristik linguistik teks *game* di Indonesia, termasuk *loanwords* dan *code-switching*, sehingga memberikan kemampuan ekstraksi informasi yang lebih akurat. Kontribusi ini menunjukkan bahwa pemanfaatan teks berita *online* dapat dioptimalkan melalui penyediaan *dataset* terstandar dan penerapan model *multilingual* sebagai *baseline* yang andal untuk tugas ekstraksi informasi menggunakan NER pada domain industri *game*.

ABSTRACT

Wibowo, Firmansyah Rekso. 2025. **Information Extraction on Indonesian Game News Texts using IndoBERT**. Thesis. Department of Informatics Engineering Faculty of Science and Technology Maulana Malik Ibrahim State Islamic University of Malang. Advisors: (I) Dr. Zainal Abidin, M.Kom; (II) Dr. Ririen Kusumawati, S.Si, M.Kom.

Indonesia's rapidly growing game industry produces a large volume of unstructured online news text, yet its use as a strategic information source remains limited due to the absence of domain-specific Natural Language Processing (NLP) resources and the insufficient capability of existing models to handle technical terminology and code-switching phenomena. This study constructs GameNER-ID, the first Named Entity Recognition (NER) dataset for the Indonesian game industry, and evaluates the effectiveness of Transformer-based models in extracting information from such texts. The dataset construction process involves collecting news articles, standardizing formats, conducting manual curation, and annotating 3,351 sentences using ten domain-specific entity types with the IOB2 scheme, resulting in a structured and representative corpus suitable as a benchmark for NLP development in the game sector. Model performance is assessed through fine-tuning several Transformer architectures and measuring Precision, Recall, and F1-Score. Experimental results show that IndoBERT performs well in extracting information, achieving an F1-Score of 0.8175. Comparative analysis reveals that the multilingual XLM-RoBERTa-large model attains higher performance with an F1-Score of 0.8386. These findings highlight that cross-lingual approaches are more effective in handling the linguistic characteristics of Indonesian game-related texts, including loanwords and code-switching, thereby providing more accurate information extraction. This contribution demonstrates that the utilization of online news text can be optimized through the provision of standardized datasets and the application of multilingual models as reliable baselines for NER-based information extraction in the game industry domain.

Keywords: *Information Extraction, Named Entity Recognition (NER), Natural Language Processing (NLP), Transformer, IndoBERT, Berita Game, GameNER-ID*

مستخلص البحث

ويووو، فيرمان شاه ركسو. 2025. استخراج المعلومات من نصوص الأخبار المتعلقة بالألعاب باللغة الإندونيسية باستخدام IndoBERT. رسالة الماجستير. قسم المعلومات، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: د. زين العابدين، الماجستير؛ المشرف الثاني: د. ريرين كوسوموازي، الماجستير.

الكلمات الرئيسية: استخراج معلومات، تعرف على كيانات مسماة (NER)، معالجة لغة طبيعية (NLP)، *Transformer*، GameNER-ID، IndoBERT، أخبار ألعاب.

تنبع صناعة الألعاب في إندونيسيا، التي تشهد نمواً سريعاً، حجماً كبيراً من نصوص الأخبار على الإنترنت غير المنظمة، ومع ذلك لا يزال استخدام هذه النصوص كمصدر للمعلومات الاستراتيجية محدوداً بسبب نقص الموارد الخاصة بمعالجة اللغة الطبيعية (NLP) في مجال محمد، بالإضافة إلى ضعف قدرة النماذج الحالية على التعامل مع المصطلحات التقنية والظواهر اللغوية المختلطة (تبديل الشيفرة). تهدف هذه الرسالة إلى بناء GameNER-ID، أول مجموعة بيانات للتعرف على الكيانات المسماة (NER) في مجال صناعة الألعاب باللغة الإندونيسية، وكذلك تقييم فعالية نماذج *Transformer* في استخراج المعلومات من هذه النصوص. تشمل عملية بناء مجموعة البيانات جمع مقالات الأخبار، وتوحيد الصياغة، والمراجعة اليدوية، وتعليم 3,351 جملة باستخدام عشرة أنواع من الكيانات الخاصة بالألعاب وفق مخطط IOB2 ، مما يؤدي إلى إنشاء مجموعة بيانات منظمة تمثل النموذج ومتاسبة للاستخدام كمقاييس لتطوير معالجة اللغة الطبيعية في قطاع الألعاب . تم تقييم أداء النموذج من خلال ضبط دقيق لعدة هيكل *Transformer* وقياس مقاييس الدقة (*Precision*) والاسترجاع (*Recall*) وكذلك معدل F1. أظهرت نتائج التجارب أن نموذج IndoBERT يستخرج المعلومات بشكل جيد حيث حقق معدل F1 يبلغ 0.8175 . أظهر التحليل المقارن أن نموذج *XLM-RoBERTa-large* متعدد اللغات يحقق أداء أعلى بمعدل F1 يبلغ 0.8386 . أكدت هذه النتائج أن منهج متعدد اللغات أكثر فعالية في التعامل مع الخصائص اللغوية لنصوص الألعاب في إندونيسيا، بما في ذلك الكلمات المستعارة وتبديل الرموز (*code-switching*)، مما يوفر قدرة أعلى على استخراج المعلومات بدقة. يوضح هذا الإسهام أن الاستفادة من نصوص الأخبار عبر الإنترنت يمكن تحسينها من خلال توفير مجموعة بيانات معيارية وتطبيق نموذج متعدد اللغات كأساس موثوق لهمة استخراج المعلومات باستخدام NER في مجال صناعة الألعاب.

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Industri *game* Indonesia mengalami pertumbuhan pesat dan kini menempati posisi strategis sebagai salah satu pasar terbesar di Asia Tenggara dan peringkat ke-16 di dunia (Ramadani dkk., 2025). Didorong oleh dukungan pemerintah, peningkatan jumlah pengembang lokal, dan ekspansi ekosistem esports, sektor ini telah menjadi salah satu pilar ekonomi kreatif nasional (Goh dkk., 2023; Gunawan dkk., 2021). Dinamika industri yang cepat ini tercermin dalam volume besar berita daring, yang menjadi sumber informasi strategis bagi pemangku kepentingan seperti pengembang, penerbit, dan investor (Nieborg, 2021). Namun, sifat data teksual yang tidak terstruktur menyulitkan proses ekstraksi wawasan kunci secara efisien, menyebabkan hilangnya peluang strategis (missed information) (Yang dkk., 2022).

Meskipun penelitian *Natural Language Processing* (NLP) pada domain *game* di Indonesia telah ada, fokusnya masih sangat terbatas. Sebagian besar studi menitikberatkan pada analisis sentimen dan pemodelan topik dari ulasan pengguna di platform seperti Google Play Store atau Steam (Ababil & Setiawan, 2022; Aditya dkk., 2023a). Pendekatan ini, meskipun berguna untuk memahami persepsi pemain, belum mampu menangkap informasi terstruktur mengenai ekosistem industri, seperti nama perusahaan yang merilis *game*, platform distribusi yang digunakan, atau acara esports yang sedang tren. Keterbatasan ini mengindikasikan adanya kesenjangan riset (*research gap*) yang signifikan dalam pemanfaatan data teks berita untuk analisis industri secara komprehensif.

Untuk mengatasi kesenjangan ini, pendekatan Ekstraksi Informasi menggunakan *Named Entity Recognition* (NER) menjadi sangat relevan. NER adalah tugas fundamental untuk mengubah teks tidak terstruktur menjadi data terstruktur dengan mengidentifikasi entitas penting seperti nama organisasi, produk, dan peristiwa (Grishman, 1997). Penerapan NER pada berita *game* memungkinkan pemantauan tren pasar, analisis kompetitor, dan identifikasi perkembangan teknologi secara otomatis.

Tantangan utama dalam penerapan NER untuk Bahasa Indonesia adalah ketersediaan sumber daya. Meskipun model-model NER yang ada seperti NER UI dan NERGrit (Wilie dkk., 2020) telah dikembangkan, model tersebut dilatih pada korpus domain umum (berita umum, Wikipedia). Akibatnya, model-model ini kesulitan mengenali entitas-entitas spesifik dalam domain *game* yang sarat dengan terminologi teknis, nama produk, dan jargon berbahasa campuran (*code-switching*) (Liu dkk., 2021). Ketiadaan dataset NER beranotasi yang berfokus pada industri *game* Indonesia menjadi hambatan kritis yang menghambat kemajuan riset dan aplikasi NLP di sektor ini.

Penelitian ini mengusulkan dua kontribusi utama. Pertama, membangun GameNER-ID, sebuah dataset NER pertama yang dianotasi secara manual untuk domain industri *game* berbahasa Indonesia, yang mencakup beberapa entitas spesifik domain. Kedua, melatih dan mengevaluasi secara komparatif kinerja model Transformer monolingual (IndoBERT) dan multilingual (mBERT, XLM-RoBERTa) pada dataset GameNER-ID. Tujuannya adalah untuk menetapkan sebuah *baseline* performa yang kuat dan mengidentifikasi arsitektur model yang paling efektif untuk domain ini. Dengan merilis dataset dan hasil evaluasi ini,

penelitian bertujuan menyediakan sumber daya fundamental untuk mendorong pengembangan aplikasi NLP canggih yang dapat mendukung pertumbuhan ekosistem industri *game* di Indonesia.

Di era digital yang penuh dengan informasi dan berita, kemampuan untuk memilah fakta dan informasi yang valid menjadi kian penting. Di sinilah nilai-nilai luhur Islam hadir sebagai panduan, salah satunya melalui perintah untuk bertabayyun dalam Surah Al-Hujurat ayat 6. Ayat ini menjadi landasan penting dalam membangun masyarakat yang harmonis dan saling menghormati, di mana informasi yang beredar didasarkan pada fakta dan bukan informasi palsu. Berikut ini adalah bacaan surah Al-Hujurat ayat 6:

يَأَيُّهَا الَّذِينَ ءَامَنُوا إِنْ جَاءَكُمْ فَاسِقٌ بِنَبَإٍ فَتَبَيَّنُوا أَنْ تُصِيبُوا قَوْمًا بِجَهَلٍ فَتُصْبِحُوا عَلَىٰ مَا
فَعَلْتُمْ نَدِيمِينَ

"Wahai orang-orang yang beriman, jika datang kepada kamu orang fasik membawa suatu berita, maka periksalah dengan teliti agar kamu tidak menimpa suatu musibah kepada suatu kaum tanpa mengetahui keadaannya yang menyebabkan kamu menyesal atas perbuatanmu itu." (Qs. al-Hujurat/49:6)

Surah Al-Hujurat ayat 6 secara khusus mengajarkan kita tentang pentingnya ber-*tabayyun* untuk mencari kebenaran informasi sebelum kita merespon atau meneruskannya. Ayat ini mengajarkan umat Islam tentang kehati-hatian dalam menerima informasi, dan dengan ber-*tabayyun*, kita terhindar dari menjadi penyebar hoaks dan fitnah yang dapat menjerumuskan diri sendiri dan orang lain.

1.2 Pernyataan Masalah

Bagaimana kinerja model IndoBERT dalam melakukan ekstraksi informasi pada teks berita *game* berbahasa Indonesia apabila diukur menggunakan metrik *Precision*, *Recall*, dan *F1-Score*?

1.3 Tujuan Penelitian

Mengevaluasi kinerja model IndoBERT dalam melakukan ekstraksi informasi pada teks berita *game* berbahasa Indonesia berdasarkan metrik *Precision*, *Recall*, dan *F1-Score*.

1.4 Batasan Masalah

1. Data penelitian diambil dari artikel berita *online* Gamebrott, Detik, CNN dan Kompas.
2. Artikel berita online yang diambil berdasarkan topik seputar industri *game* berbahasa Indonesia.

1.5 Manfaat Penelitian

1. Industri *game*, memberikan informasi terkait berbagai entitas penting di dalam industri seperti nama *game*, *developer*, *publisher*, pemerintah, pemain *game* dan lainnya.
2. Komunitas NLP, Menyumbangkan dataset GameNER-ID sebagai dataset publik baru dan memberikan hasil *benchmarking* yang menjadi acuan untuk pengembangan aplikasi NLP di domain *game* Indonesia.

BAB II

KAJIAN PUSTAKA

Bab ini menyajikan tinjauan pustaka yang komprehensif untuk membangun landasan penelitian. Pembahasan dimulai dengan mengidentifikasi konteks dan entitas-entitas kunci yang membentuk ekosistem industri *game* di Indonesia. Selanjutnya, bab ini akan menganalisis literatur penelitian yang ada untuk menyoroti kesenjangan riset (*research gap*) dalam pemanfaatan *Natural Language Processing* (NLP) pada domain ini. Terakhir, akan diuraikan justifikasi pemilihan sumber data, teknologi, serta kerangka teori yang relevan untuk menjawab kesenjangan tersebut, sekaligus memposisikan kontribusi penelitian ini dalam lanskap studi yang lebih luas.

2.1 Entitas Penting Industri *Game* Indonesia

Industri *game* Indonesia berkembang sebagai pasar potensial dengan tingkat pertumbuhan tinggi, menempati posisi sebagai pasar terbesar di ASEAN dan peringkat ke-17 di dunia (Gunawan dkk., 2021; Hudrasyah dkk., 2019). Pertumbuhan tersebut memperluas ekosistem industri dan mendorong keterlibatan berbagai aktor yang berperan dalam proses produksi, distribusi, dan konsumsi *game* di tingkat nasional maupun global (Goh dkk., 2023).

Berbagai penelitian menunjukkan bahwa ekosistem industri *game* Indonesia dibentuk oleh sejumlah entitas penting yang saling berinteraksi, meliputi pengembang (*developer*) sebagai produsen utama, penerbit (*publisher*) sebagai pengelola distribusi dan promosi, investor sebagai penyedia sumber pendanaan, serta pemerintah sebagai regulator dan fasilitator kebijakan. Selain itu, media

berperan dalam memperkuat visibilitas dan persepsi publik terhadap produk *game*, sedangkan pemain (*gamer*) menjadi pusat aktivitas ekonomi sekaligus penentu arah inovasi. Lembaga pendidikan dan institusi pendukung juga memiliki kontribusi dalam penyediaan sumber daya manusia dan pengembangan kompetensi teknis.

Penelitian-penelitian terdahulu secara konsisten menyoroti keberadaan dan keterkaitan entitas-entitas tersebut dalam membangun ekosistem industri *game* yang berkelanjutan. Rangkuman berbagai studi yang membahas peran dan relasi antarentitas ditampilkan pada Tabel 2.1.

Tabel 2.1 Literatur *Review* Terkait Tantangan Analisa Industri *Game*

No	Peneliti	Topik	Entitas Terkait
1	Gunawan dkk. (2021)	Potensi industri <i>game</i> Indonesia	Pengembang, Pemerintah, Investor
2	Hudrasyah dkk. (2019)	Ekosistem industri <i>game</i> global	Pengembang, Pemerintah, Platform
3	Fajri (2012)	Dukungan kebijakan pemerintah	Pemerintah, Pengembang
4	Al Rosyid dkk., (2024)	Potensi industri animasi dan <i>game</i> Surabaya	Pengembang, Pemerintah, Investor
5	Wibowo & Lubis (2013)	Peran media <i>online</i> dalam promosi	Media, Pengembang, Pemasar
6	Akbar & Asmara (2022)	Perbedaan SDM pada perusahaan korporasi dan <i>indie</i>	Pengembang Korporasi, <i>Indie</i> , Pendidikan, Pemerintah
7	Suciastih (2024)	Pemetaan pelaku dan strategi pengembangan	Pengembang, Pemerintah, Investor,
8	Goh dkk. (2023)	Tinjauan literatur industri <i>game</i> secara global	Pengembang, Penerbit, Platform, Pemain

Identifikasi terhadap entitas utama dalam industri *game* seperti pengembang, penerbit, dan lembaga pemerintah (Tabel 2.1) menegaskan peran krusial masing-masing aktor dalam ekosistem *game*. Meskipun keberadaan entitas-entitas ini diakui sangat penting, muncul pertanyaan mengenai sejauh mana penelitian ilmiah, khususnya di bidang NLP, telah berfokus pada analisis interaksi dan dinamika antar entitas tersebut. Tahapan berikutnya akan meninjau literatur penelitian terkini untuk memahami fokus utama dan mengidentifikasi area yang belum terjamah.

2.2 Analisis Literatur Penelitian Terkait *Game*

Kajian akademik di Indonesia menunjukkan dominasi pendekatan *machine learning* yang berfokus pada analisis sentimen dan pemodelan topik dari ulasan pengguna. Pendekatan ini dimanfaatkan untuk menilai persepsi pemain terhadap kualitas, performa, dan pengalaman bermain berbagai judul populer seperti *Bus Simulator Indonesia*, *Mobile Legends*, dan *Genshin Impact* (Alfawas dkk., 2024; Hafidz Syauqie dkk., 2024; Safrudin dkk., 2024; Utami dkk., 2023). Analisis tersebut menyoroti keterkaitan antara opini pengguna dan reputasi *game*, namun sebagian besar masih menitikberatkan pada aspek teknis tanpa menautkan hasil terhadap dinamika industri secara menyeluruh.

Pendekatan *topic modeling* juga berkembang untuk mengidentifikasi tema dominan dalam ulasan pemain. Ababil & Setiawan (2022) memanfaatkan LDA untuk menemukan topik terkait performa dan kendala teknis dalam *Wild Rift*, sementara model berbasis representasi kontekstual seperti BERT dan BERTopic menghasilkan pemetaan tema yang lebih mendalam terhadap persepsi pengguna (Priyatna, 2023; Ramadhan & Siswoyo, 2024).

Walaupun berbagai penelitian tersebut memperkaya pemahaman terhadap perilaku dan pengalaman pemain, ruang lingkupnya masih terbatas pada analisis konten ulasan. Sebagian besar studi belum menautkan hasilnya dengan faktor ekonomi, kebijakan, maupun perkembangan ekosistem industri *game* nasional (Fajri, 2012). Keterbatasan ini menandai tantangan utama dalam upaya membangun pemahaman holistik terhadap sektor *game* Indonesia yang semakin dinamis. Tabel 2.2 menampilkan ringkasan penelitian terkait yang memperlihatkan variasi metode,

sumber data, serta fokus utama yang cenderung berpusat pada persepsi pengguna dibandingkan konteks industri yang lebih luas.

Tabel 2.2 Literatur *Review* Penelitian Terkait *Game*

No	Peneliti	Metode	Data	Fokus Utama	Keterbatasan
1	Alfawas dkk. (2024)	Naive Bayes TF-IDF	Ulasan Pemain	Sentimen pada BUSSID & Sky	Tidak mengaitkan hasil dengan industri
2	Hafidz Syauqie dkk. (2024)	Naive Bayes TF-IDF	Ulasan Pemain	Sentimen BUSSID	Fokus terbatas pada sentimen
3	Nurdy (2024)	Naive Bayes KDD	Ulasan Pemain	Sentimen Stumble Guys	Tidak membahas konteks industri
4	Utami dkk. (2023)	SVM	Ulasan Pemain	Sentimen Mobile Legends	Fokus pada fitur <i>game</i>
5	Safrudin dkk. (2024)	Naive Bayes & SVM	Ulasan Pemain	Sentimen Genshin Impact	Tidak mengaitkan dengan industri
6	Ababil & Setiawan (2022)	LDA	Ulasan Pemain	Pemodelan topik Wild Rift	Tidak membahas konteks industri
7	Priyatna, (2023) Ramadhan & Siswoyo (2024)	BERT, BERTopic	Ulasan Pemain	Sentimen & topik berbasis embedding	Terbatas pada analisis konten
8	Aditya dkk. (2023b)	Naive Bayes	Ulasan Pemain	Sentimen Subway Surf	Fokus pada sentimen saja
9	Sihotang (2020)	Multinomial Naive Bayes	Ulasan Pemain	Klasifikasi rating <i>game</i>	Tidak mengaitkan analisis dengan industri
10	Alkindi & Nasution (2024)	SVM & Naive Bayes	Ulasan Pemain	Sentimen Roblox	Terfokus pada ulasan pengguna
11	Firdaus (2022)	SVM	Ulasan Pemain	Aspect-based Sentiment Analysis	Tidak membahas konteks industri
12	Febriansyah & Qoiriah (2024)	Spectral Clustering PCA	Ulasan Pemain (Steam)	Pengelompokan ulasan Elden Ring	Tidak menilai dampak terhadap industri
13	Rahman dkk. (2021)	Naive Bayes	Ulasan Pemain	Sentimen Hogwarts Mystery	Tidak menyoroti kontribusi ekonomi
14	Almaudy & Kalifia, (2022)	TF-IDF & Naive Bayes	Ulasan Pemain	Analisis aspek pengalaman pengguna	Tidak mengaitkan hasil dengan industri
15	AR dkk. (2023); Juhaeni dkk, (2023); Khasanah dkk. (2023); Kusuma & Fadiana, (2024); Mulyani, (2023); Salfina dkk. (2023); Ulya dkk. (2023)	Pengembangan <i>Game</i> Edukasi	-	Pengembangan <i>game</i> edukasi	Fokus pada desain pengembangan <i>game</i> , bukan ekosistem industri

Minimnya penelitian yang menyoroti entitas, relasi, dan terminologi dalam konteks industri *game* menunjukkan kebutuhan akan pendekatan analisis teks yang

lebih terstruktur. Keterbatasan tersebut membuka ruang bagi penerapan *Named Entity Recognition (NER)* sebagai metode untuk mengekstraksi entitas kunci dan membangun representasi informasi yang lebih kontekstual terhadap ekosistem industri *game* di Indonesia.

2.3 Pemanfaatan Teks Berita *Online* pada Industri *Game*

Pertumbuhan industri *game* di Indonesia menghasilkan volume data tekstual dalam jumlah besar setiap hari. Artikel berita, ulasan permainan, forum diskusi, dan media sosial menjadi sumber utama informasi yang menggambarkan dinamika industri tersebut (Fazrian dkk., 2024). Indonesia tercatat memiliki sekitar 47.000 media massa, dengan 43.300 di antaranya berbentuk media *online* (Miko & Jafar, 2018). Kondisi tersebut menempatkan teks berita daring sebagai penghasil *big data* yang signifikan (Fauziah dkk., 2021) sekaligus menyediakan sumber informasi potensial untuk analisis berbasis *Natural Language Processing* (NLP).

2.4 Model Bahasa Berbasis Transformer untuk Ekstraksi Informasi

Kemajuan teknologi pemrosesan bahasa alami mengubah secara fundamental pendekatan terhadap ekstraksi informasi. Arsitektur Transformer muncul sebagai fondasi utama karena kemampuannya memahami konteks kalimat secara menyeluruh melalui mekanisme *self-attention* (Vaswani dkk., 2017). Mekanisme ini menilai keterkaitan antar-token dalam satu kalimat sehingga menghasilkan representasi makna yang lebih akurat dibandingkan metode konvensional seperti *Recurrent Neural Network* (RNN) dan *Long Short-Term Memory* (LSTM). Kemunculan Transformer melahirkan berbagai model pralatih seperti BERT, mBERT, dan XLM-RoBERTa yang terbukti efektif dalam tugas *Named Entity Recognition* (NER).

Model Multilingual BERT (mBERT) (Wu & Dredze, 2019) memanfaatkan konteks dua arah (*bidirectional*) untuk memahami hubungan antar kata dalam kalimat (Devlin dkk., 2019). Model ini dilatih pada korpus Wikipedia dari 104 bahasa, termasuk Bahasa Indonesia. Keunggulan utamanya terletak pada kemampuan *cross-lingual transfer* secara *zero-shot*, yaitu ketika model yang dilatih pada data berbahasa Inggris mampu mengenali entitas dalam bahasa lain tanpa pelatihan ulang. (Wu & Dredze, 2019) menunjukkan bahwa mBERT sangat kompetitif untuk tugas NER lintas bahasa. Namun, penerapan mBERT pada domain *game* menghadapi keterbatasan kosakata karena sumber pelatihannya yang berasal dari Wikipedia. Istilah teknis dan jargon khas dunia *game* umumnya tidak tercakup, sehingga menurunkan akurasi dalam mengenali entitas spesifik domain.

Model XLM-RoBERTa (XLM-R) dikembangkan untuk mengatasi keterbatasan tersebut dengan memanfaatkan korpus yang lebih besar dari *Common Crawl* untuk 100 bahasa (Conneau dkk., 2019). Skala data yang luas meningkatkan kemampuan XLM-R dalam memahami bahasa dengan sumber daya rendah. Pada *benchmark* IndoNLU, XLM-R menunjukkan performa kuat dan kerap melampaui mBERT. Meski demikian, model ini tetap menghadapi tantangan pada domain spesifik seperti *game*. Korpus *Common Crawl* yang beragam tidak menjamin cakupan memadai terhadap istilah dinamis dan jargon berbahasa campuran yang umum digunakan dalam komunitas gaming.

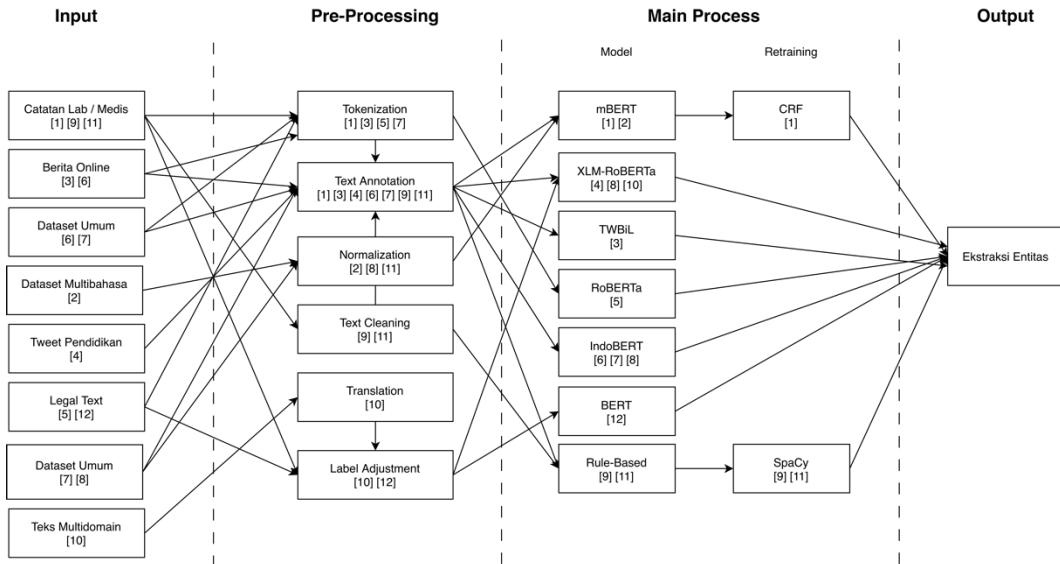
Untuk menangkap konteks lokal secara lebih mendalam, IndoBERT hadir sebagai model yang dilatih secara khusus pada korpus masif berbahasa Indonesia (Indo4B). Pendekatan monolingual memungkinkan IndoBERT memahami nuansa, morfologi, dan struktur kalimat Bahasa Indonesia secara lebih presisi dibandingkan

model multibahasa. Pada tugas NER dalam *benchmark* IndoNLU, IndoBERT, terutama varian IndoBERT-Large menunjukkan performa *state-of-the-art*, melampaui mBERT dan XLM-R pada sejumlah dataset. Meskipun demikian, kendala utama tetap muncul ketika diterapkan pada domain *game*. Korpus Indo4B yang bersumber dari berita, media sosial, dan blog umum belum mencakup secara memadai kosakata dan entitas khas dari dunia *game*.

Secara keseluruhan, ketiga model Transformer tersebut menunjukkan potensi tinggi untuk ekstraksi informasi dalam Bahasa Indonesia, namun efektivitasnya sangat bergantung pada kesesuaian antara korpus pelatihan dan konteks domain aplikasi. Kondisi ini menegaskan perlunya adaptasi model melalui *fine-tuning* dan pembangunan dataset khusus domain agar sistem NER mampu mengenali entitas secara akurat. Kajian berikut membahas landasan teoretis dan hasil penelitian terdahulu yang menyoroti pentingnya adaptasi domain dan pengembangan sumber daya bahasa untuk mendukung performa model NER secara optimal.

2.5 Kerangka Teori

Kerangka teori ini menyintesis berbagai pendekatan relevan dan memetakan alur konseptual dari sumber data hingga arsitektur model. Gambar 2.1 menggambarkan lanskap penelitian *Named Entity Recognition* (NER) lintas domain serta adopsi praktik terbaik dari studi terdahulu untuk diterapkan pada domain industri *game*.



(1) Polignano dkk, (2021); (2) Mo dkk. (2024); (3) Zainuddin, Mudassir & Tahir (2025); (4) Ullah dkk. (2025); (5) Abdullah dkk. (2024); (6) Istiqomah & Novika (2025); (7) Koto dkk. (2020); (8) Wilie dkk. (2020); (9) Tarcar dkk. (2020); (10) Khairunnisa, Chen & Komachi (2024); (11) Kamal & Hassanain (2025); (12) Oralbekova dkk. (2025).

Gambar 2.1 Kerangka Teori

Efektivitas model NER sangat bergantung pada kemampuan menyesuaikan diri terhadap karakteristik domain. Beragam studi menunjukkan bahwa adaptasi domain menjadi faktor utama dalam peningkatan kinerja model berbasis Transformer. (Polignano dkk., 2021) mengkaji NER pada laporan ultrasonografi pediatrik berbahasa Spanyol dan menemukan bahwa model BERT multibahasa yang di-*fine-tuning* menggunakan data medis mencapai akurasi tertinggi meskipun ukuran dataset terbatas. Relevansi data pelatihan terbukti memiliki pengaruh lebih besar dibandingkan jumlah data.

Tarcar dkk. (2019) memperkuat temuan tersebut dengan menunjukkan bahwa transfer learning dan pre-training pada korpus kesehatan meningkatkan kinerja NER secara signifikan meskipun data anotasi terbatas. Model dengan 50% data spesifik domain mampu melampaui model umum yang dilatih dengan 100% data, menandakan pentingnya adaptasi domain dalam proses pelatihan.

Tabel 2.3 Performa Setiap Model Kerangka Teori

No	Peneliti	Dataset	Domain	Metode / Model	<i>F1-Score</i>
1	(Polignano dkk., 2021)	513 laporan ultrasonografi pediatrik (Spanyol)	Kesehatan	mBERT + CRF	0.6963
2	Mo dkk. (2024)	XTREME (40 bahasa) & CoNLL (4 bahasa)	Multibahasa	MCL-NER + mBERT	70.4 (XTREME) / 80.5 (CoNLL)
3	Zainuddin, Mudassir & Tahir (2025)	5.000 kalimat berita daring (Detik.com)	Berita	TWBiL (Transformer + Word2Vec + Attention + Bi-LSTM)	85.11
4	Ullah dkk. (2025)	EDU-NER-2025 (30.000 tweet pendidikan, Urdu)	Pendidikan	XLM-RoBERTa	98.00
5	Abdullah dkk. (2024)	1.500 kalimat hukum (Kurdi Sorani)	Hukum	RoBERTa	92.9
6	Istiqomah & Novika (2025)	1.137 berita kesehatan pascabencana (Indonesia)	Kesehatan	IndoBERT fine-tuned (vs mBERT, XLM-R)	88.26
7	Koto dkk. (2020)	NER UI & NER UGM	Umum	IndoBERT	90.1 / 74.9
8	Wilie dkk. (2020)	NERGrit & NERP (Indo4B corpus)	Umum	IndoBERTLARGE, XLM-RLARGE	79.25 (NERP) / 79.09 (NERGrit)
9	Tarcar dkk. (2020)	4.212 kalimat EHR (Inggris)	Kesehatan	spaCy + pre-training domain spesifik	0.785
10	Khairunnisa dkk. (2024)	IDCrossNER (5 domain: politik, sains, musik, literatur, AI)	Multi-domain	XLM-RLARGE + DoSEA	73.08
11	Kamal & Hassanain, (2025)	3.000 laporan mikrobiologi sintetis	Kesehatan	Rule-based, spaCy, DistilBERT (QA-based)	1.00 (rule-based & spaCy) / 0.68-0.80 (QA)
12	Oralbekova dkk. (2025)	KazNERD (Kazakh legal corpus)	Hukum	BERT, GPT, LSTM (perbandingan klasifikasi)	92.02 (akurasi)

Selanjutnya, Kamal & Hassanain (2025) membandingkan pendekatan berbasis aturan, statistik, dan Transformer pada laporan mikrobiologi. Model Transformer yang *di-fine-tuning* secara spesifik pada domain mencapai akurasi tertinggi karena mampu menangkap konteks terminologi kompleks. Sebaliknya, model umum mengalami kesulitan menafsirkan ambiguitas kontekstual. Alur hasil

tersebut menegaskan bahwa penyesuaian model terhadap karakteristik linguistik dan semantik domain menjadi kunci utama keberhasilan NER.

Keterbatasan sumber daya bahasa menjadi tantangan utama dalam pengembangan NLP lintas bahasa. Ullah dkk. (2025) menyoroti isu ini pada bahasa Urdu melalui pengenalan dataset EDU-NER-2025 di domain pendidikan. Model Transformer multibahasa seperti XLM-RoBERTa menunjukkan kemampuan unggul dalam menangani kompleksitas bahasa rendah sumber daya dengan akurasi tinggi, melampaui metode tradisional. Abdullah dkk. (2024) memperkuat temuan tersebut melalui pembangunan korpus baru berbahasa Kurdi dan *fine-tuning* pada RoBERTa, yang menetapkan tolok ukur baru bagi NER Kurdi. Upaya tersebut menegaskan bahwa pembuatan dataset dan adaptasi model pralatih menjadi strategi efektif untuk meningkatkan performa NLP pada bahasa rendah sumber daya.

Mo dkk. (2024) menambahkan dimensi baru melalui kerangka *Multi-view Contrastive Learning* (MCL-NER) yang meningkatkan kinerja *Cross-Lingual NER* dengan memanfaatkan transfer pengetahuan dari bahasa sumber ke bahasa target. Pendekatan ini menunjukkan efisiensi tinggi pada kondisi data terbatas. Di sisi lain, Oralbekova dkk. (2025) membuktikan efektivitas arsitektur BERT dalam memahami konteks hukum berbahasa Kazakh, memperkuat posisi Transformer sebagai pilihan utama untuk bahasa yang kekurangan sumber daya.

Bahasa Indonesia tergolong bahasa rendah sumber daya meskipun memiliki jumlah penutur besar. Koto dkk. (2020) memperkenalkan IndoLEM dan IndoBERT, yang menjadi *benchmark* dan model monolingual berskala besar pertama untuk Bahasa Indonesia. IndoBERT secara konsisten mengungguli model multibahasa pada berbagai tugas, termasuk NER. Wilie dkk. (2020) memperluas

fondasi tersebut melalui IndoNLU, *benchmark* komprehensif yang mencakup dataset Indo4B serta model IndoBERT. Kedua penelitian ini menyediakan infrastruktur penting yang mendorong kemajuan penelitian NLP berbahasa Indonesia.

Kemajuan berikutnya terlihat pada penerapan NER di berbagai domain spesifik. Zainuddin dkk. (2025) mengembangkan model hibrida TWBiL untuk NER pada berita daring dan mencatat peningkatan kinerja signifikan meskipun dengan keterbatasan data dan komputasi. Istiqomah & Novika (2025) menunjukkan bahwa IndoBERT yang di-*fine-tuning* pada data medis mengenali istilah lokal dan singkatan lebih akurat dibandingkan model multibahasa. Sementara itu, Khairunnisa dkk. (2024) memperkenalkan dataset IDCrossNER yang mencakup lima domain, membuktikan bahwa *cross-domain transfer learning* dan augmentasi data mampu meningkatkan kinerja pada skenario rendah sumber daya.

Rangkaian penelitian tersebut menegaskan tiga prinsip utama. Pertama, NER membutuhkan dataset dan model yang disesuaikan dengan karakteristik domain untuk mencapai akurasi tinggi. Kedua, bahasa rendah sumber daya seperti Bahasa Indonesia memerlukan strategi khusus melalui pengembangan korpus dan adaptasi model. Ketiga, meskipun riset NER di Indonesia telah berkembang pesat pada domain berita dan kesehatan, masih terdapat kekosongan besar pada sektor lain yang relevan secara sosial dan ekonomi.

Industri *game* menjadi contoh nyata dari sektor dengan pertumbuhan cepat namun minim kajian NLP. Kajian pustaka menunjukkan bahwa penelitian di bidang ini masih berfokus pada analisis sentimen, sementara pemetaan entitas seperti perusahaan, platform, dan peristiwa dalam ekosistem *game* belum mendapat

perhatian memadai. Kondisi tersebut menegaskan perlunya pendekatan sistematis untuk ekstraksi informasi berbasis *Named Entity Recognition* (NER) menggunakan model Transformer pada korpus berita *game* daring berbahasa Indonesia.

BAB III

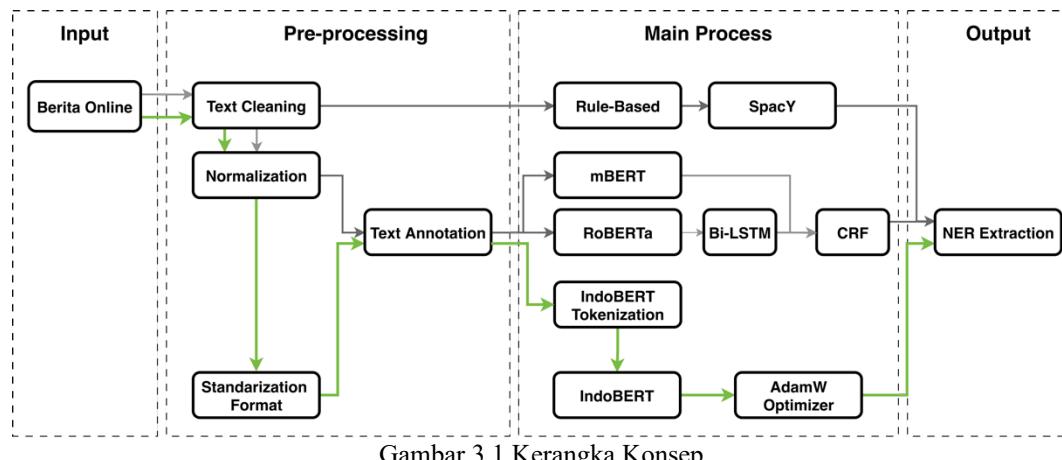
METODOLOGI PENELITIAN

Bab ini menguraikan metode penelitian yang dilakukan dalam penelitian. Pembahasan mencakup kerangka konsep, prosedur penelitian secara keseluruhan, mulai dari tahap pengumpulan dan persiapan data, skema anotasi, hingga pengukuran metrik evaluasi yang digunakan untuk mengukur kinerja model.

3.1 Kerangka Konsep

Kerangka konsep pada penelitian ini akan digambarkan pada Gambar 3.1.

Tahapan terdiridari *input*, *preprocessing*, *main process*, dan *output*.



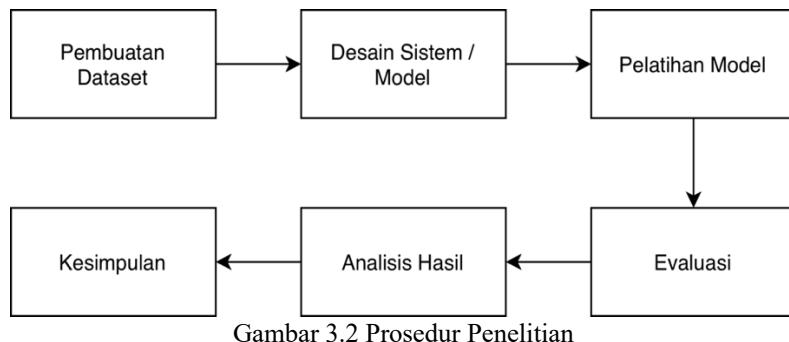
Gambar 3.1 Kerangka Konsep

Kerangka konsep pada Gambar 3.1 menggambarkan perbedaan alur kerja antara pendekatan konvensional (alur abu-abu) dan metode usulan (alur hijau) dalam penyelesaian tugas ekstraksi entitas (*Named Entity Recognition*) pada berita daring. Pada jalur konvensional, proses berlangsung melalui tahapan *preprocessing* yang panjang, mencakup *text cleaning* dan normalisasi sebelum masuk ke tahap anotasi. Arsitektur model yang digunakan masih bergantung pada pipeline terpisah, seperti kombinasi *fine-tuning* mBERT dan CRF (*Conditional Random Field*) (Polignano dkk., 2021), IndoBERT/Bi-GRU/CRF (Dave & Chowanda, 2024), serta

RoBERTa/Bi-LSTM/CRF (Y. Liu & Liu, 2024) atau metode berbasis aturan melalui SpaCy (Kamal & Hassanain, 2025), sehingga memerlukan rekayasa fitur tambahan dan bergantung pada beberapa modul eksternal. Sebaliknya, jalur usulan menekankan proses yang lebih ringkas secara *end-to-end*. Alur preprocessing dilanjutkan sesuai standardisasi format artikel untuk kebutuhan NER, dilanjutkan oleh anotasi, tokenisasi otomatis melalui IndoBERT *Tokenization*, serta pelatihan model seperti IndoBERT, mBERT atau XLM-RoBERTa yang dioptimalkan langsung menggunakan AdamW Optimizer. Integrasi ini mengurangi ketergantungan pada komponen eksternal seperti CRF, menyederhanakan kompleksitas arsitektur, dan mengarahkan keseluruhan proses menuju NER *Extraction* dengan lebih ringkas dan efisien.

3.2 Prosedur Penelitian

Prosedur penelitian merupakan tahapan-tahapan yang dilakukan di dalam penelitian. Penelitian ini dilakukan melalui enam tahapan utama sebagaimana ditunjukkan pada Gambar 3.1. Tahap pertama adalah pembuatan dataset, yang mencakup pengumpulan artikel berita, pra-pemrosesan teks, kurasi kalimat, anotasi entitas, dan pembagian data. Selanjutnya desain sistem dan model, yang mendefinisikan arsitektur model NER berbasis Transformer serta alur pemrosesan data. Tahap ketiga adalah pelatihan model, di mana model dilatih menggunakan dataset yang telah dibangun. Selanjutnya, dilakukan evaluasi untuk menilai kinerja model berdasarkan metrik *precision*, *recall*, dan *F1-score*. Hasil evaluasi kemudian dianalisis pada tahap analisis hasil guna mengidentifikasi pengaruh *domain shift* dan pola kesalahan model.



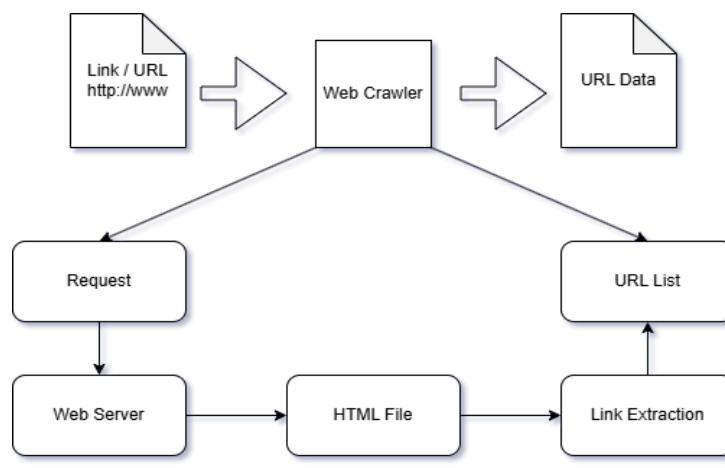
3.3 Persiapan Data dan Konstruksi Dataset

Persiapan data dan konstruksi dataset merupakan fase awal dalam penelitian ini yang bertujuan untuk mengubah data teks mentah tidak terstruktur menjadi sumber daya terstruktur dan berlabel berkualitas tinggi. Proses ini dirancang secara sistematis yang terbagi ke dalam lima tahapan utama: (1) pengumpulan data mentah dari berbagai sumber berita daring; (2) pra-pemrosesan untuk membersihkan dan menstandarkan format teks; (3) kurasi manual untuk menyeleksi kalimat-kalimat yang relevan dengan domain *game*; (4) anotasi entitas untuk memberikan label pada informasi spesifik; dan (5) pembagian dataset menjadi set data latih, validasi, dan uji. Setiap tahapan dilakukan dengan cermat untuk memastikan bahwa dataset final yang dihasilkan, GameNER-ID, memiliki konsistensi, relevansi, dan kualitas yang memadai untuk melatih dan mengevaluasi model *Named Entity Recognition* (NER) secara efektif.

3.3.1 Pengumpulan Data (Data Collection)

Pengumpulan data memegang peranan penting dalam proses pengolahan dan analisa selanjutnya. Proses pengumpulan data harus dikerjakan dengan teknik yang tepat sehingga data dapat diperoleh secara akurat. Selain itu saat ini data memiliki jumlah yang begitu besar sehingga teknik yang tepat sangat mempengaruhi perolehan data. Selanjutnya sumber berita *online* yang digunakan

merupakan sumber yang digunakan dalam penelitian seperti *Gamebrott* (Imballo Zaki Hasibuan & Triase, 2022) , *Detik*, *Kompas* (Rahmawati dkk., 2022), *CNN* (Laras dkk., 2024). Namun, karena banyaknya jumlah artikel berita yang terdapat pada sumber berita tersebut membuat penelitian ini tidak bisa mendapatkan secara manual. Keterbatasan tersebut membuat penelitian ini menggunakan sistem *crawling* untuk website tersebut atau yang biasa disebut dengan *web crawling*.



Gambar 3.3 Desain Sistem *Crawling* Data

Web Crawling merupakan program otomatis yang menjelajahi laman web secara berkala dengan mengikuti tautan yang ditemukan di setiap halaman. Proses dimulai dari sekumpulan halaman awal yang disebut *seed*, kemudian crawler menelusuri setiap tautan yang terdapat di dalamnya (Bifulco dkk., 2021). Alur proses crawling data yang digunakan tergambar pada Gambar 3.2 sebagai rancangan sistem pengumpulan data berita.

Desain sistem crawling tersusun dalam beberapa tahap terstruktur. Tahap awal menerima masukan berupa URL dari pengguna sebagai sumber utama artikel berita yang akan dikumpulkan. Web crawler kemudian bekerja otomatis mengunjungi setiap URL yang ditentukan. Implementasi sistem menggunakan

pustaka Scrapy pada Python karena mendukung interaksi otomatis dengan laman web.

Proses dimulai dari URL awal (*seed*) yang biasanya berupa halaman utama situs berita atau kategori khusus terkait *game*. Crawler mengirimkan *request* ke server situs target, kemudian server merespons dengan mengirimkan berkas HTML. Setelah menerima berkas tersebut, crawler mengekstraksi tautan baru di dalamnya, khususnya yang menuju artikel berita *game* lain.

Tabel 3.1 Data Artikel Berita *Online Game*

No	URL	Judul	Konten
1	https://gamebrott.com/berita/...	AI Diklaim Tak Picu PHK di Industri Gaming Lokal, Malah Beri Kerjaan	Pengembang <i>game</i> Agate menyebut pemanfaatan kecerdasan buatan (AI) di industri permainan lokal tak memicu pemutusan...
2	https://gamebrott.com/games-features/d...	Cerita Remaja 13 Tahun Pecahkan Rekor Tetris, Bikin <i>Game</i> Jadi Crash	Remaja Berusia 13 tahun bernama Willis Gibson menjadi yang pertama memenangkan <i>game</i> Tetris versi original besutan Nintendo...
3	https://www.detik.com/jabar/berita/d-7316...	Bukan Cuma <i>Game</i> ! Ternyata Ini Faktor Lain Penyebab Anak Agresif	Psikolog Veronica Adesla menjelaskan sejumlah faktor selain <i>Game</i> ...
4	https://inet.detik.com/games-news/d-7314...	Kode Redeem Ojol The <i>Game</i> 28 April 2024, Dapat Tiga Hadiah Menarik	Daftar Isi Kode Redeem Ojol The <i>Game</i> Terbaru Cara Klaim Kode Redeem Ojol The <i>Game</i> Daftar Isi Kode Redeem Ojol ...
5	https://www.detik.com/bali/berita/d-7312...	AGI soal Perpres <i>Game Online</i> : Tidak Akan Diblokir Jika Sudah Ada Batasan Usia	Ketua Umum Asosiasi <i>Game</i> Indonesia (AGI) Cipto Adiguno buka suara terkait rencana pemerintah ...
...
1703	https://gamebrott.com/call-of-duty-mobile	Call of Duty <i>Mobile</i> menjadi salah satu <i>game</i> paling banyak...	Call of Duty <i>Mobile</i> , salah satu <i>game</i> FPS yang diluncurkan pertama kali pada beberapa...
1704	https://gamebrott.com/game...	<i>Game</i> Paling Banyak Didownload di Playstore...	Meskipun kurang populer di negara barat dan sering sekali dikritik namun <i>game</i> garena...

Tautan baru ditambahkan ke daftar URL yang akan diproses secara berulang dengan mekanisme yang sama. Setiap URL dikunjungi, dianalisis, dan diekstraksi untuk menemukan tautan tambahan. Siklus ini terus berlangsung hingga sistem

memperoleh kumpulan besar artikel berita *game* dari berbagai situs. Dengan mekanisme ini, proses pengumpulan data berlangsung otomatis dan efisien tanpa intervensi manual.

3.3.2 Pra-pemrosesan Artikel

Tahap pra-pemrosesan berfungsi menyaring dan menata teks agar siap digunakan pada proses anotasi dan pelatihan model NLP. Seluruh artikel hasil akuisisi diproses melalui serangkaian operasi pembersihan untuk menghilangkan elemen mengganggu sekaligus menstandarkan format penulisan. Tujuannya menciptakan korpus yang bersih, seragam, dan mudah dibaca, sehingga setiap token memiliki representasi yang konsisten di seluruh dokumen.

Kerapian teks dijaga melalui normalisasi spasi yang menghapus tab, baris kosong, dan jarak ganda antar-kata. Langkah ini memastikan setiap kalimat tersusun secara proporsional tanpa ketidakteraturan pemisahan antarparagraf. Pembersihan karakter khusus dilakukan untuk menyingkirkan simbol, tag sumber, atau penanda lokasi seperti “– Jakarta” yang tidak memberikan nilai semantik pada isi artikel. Keberadaan karakter semacam ini sering mengganggu tokenisasi serta menurunkan akurasi pengenalan entitas pada tahap anotasi.

Konsistensi format diperkuat melalui standardisasi tanda baca dan penyeragaman format numerik. Koreksi tanda baca dilakukan dengan memperbaiki jarak di dalam tanda kurung dan kutipan ganda serta menambahkan spasi yang benar setelah koma atau titik. Penulisan angka, mata uang, persentase, dan desimal distandarkan agar seluruh representasi numerik mengikuti pola yang sama, misalnya “Rp 5,000,000” diubah menjadi “Rp5,000,000”. Penyeragaman semacam

ini mencegah variasi penulisan yang dapat menimbulkan ambiguitas selama pemrosesan linguistik.

Tabel 3.2 Langkah-Langkah Pra-pemrosesan Artikel Berita

Kategori Langkah	Deskripsi Tindakan	Contoh Transformasi
Normalisasi Spasi	Menghapus spasi, tab, dan baris kosong berlebih untuk merapikan struktur teks.	"Game ini populer" → "Game ini populer"
Pembersihan Karakter Khusus	Menghilangkan simbol, tag, atau penanda sumber berita yang tidak relevan.	"- Jakarta, CNN —" → "CNN —"
Standardisasi Tanda Baca	Memperbaiki jarak di dalam tanda kurung dan kutipan serta menambahkan spasi yang benar.	"Game(RPG)" → "Game (RPG)"
Penyeragaman Format Numerik	Menstandarkan penulisan mata uang, angka, desimal, dan persentase.	"Rp 5,000,000" → "Rp5,000,000"
Pelestarian Format Penting	Menjaga agar domain dan URL tetap utuh tanpa diubah selama pembersihan.	"HUB.ID" → "HUB.ID"

Kestabilan struktur teks dijaga dengan mempertahankan format penting yang berpotensi berisi informasi entitas, seperti domain situs (contoh: *hub.id*) atau alamat URL. Pemeliharaan elemen tersebut memastikan sistem anotasi tetap mengenali pola entitas berbasis simbol titik atau garis bawah tanpa kehilangan konteks aslinya. Integrasi seluruh langkah pra-pemrosesan menghasilkan kumpulan teks dengan kualitas struktural tinggi, siap untuk kurasi dan anotasi lebih lanjut.

3.3.3 Kurasi Kalimat

Proses kurasi berfokus pada penyaringan kalimat hasil pra-pemrosesan agar hanya kalimat yang relevan dan informatif terhadap konteks domain *game* yang dipertahankan. Proses ini dilakukan secara manual oleh anotator berpengalaman yang memahami terminologi dan praktik industri *game*. Setiap keputusan seleksi didasarkan pada pemahaman konseptual terhadap istilah dan konteks penggunaannya dalam ekosistem permainan digital. Tujuannya memastikan setiap kalimat dalam korpus mencerminkan karakteristik linguistik dan semantik khas

domain *game*, seperti penyebutan nama gim, perusahaan pengembang, atau istilah teknis yang umum digunakan dalam media dan komunitas pemain.

Tabel 3.3 Contoh Proses Kurasi Kalimat

No.	Kalimat Sampel	Entitas <i>Game</i> ?	Alasan
1	Garena akan merilis update terbaru untuk <i>Free Fire</i> .	Ya	Mengandung entitas COMPANY (<i>Garena</i>) dan GAME (<i>Free Fire</i>).
2	Industri <i>game</i> menunjukkan pertumbuhan yang pesat.	Tidak	Kalimat bersifat umum dan tidak mengandung entitas spesifik.
3	Turnamen ini akan digelar di <i>Jakarta Convention Center</i> .	Ya	Mengandung entitas LOCATION (<i>Jakarta Convention Center</i>) yang relevan dalam konteks acara <i>game</i> .
4	<i>Square Enix</i> mengumumkan perilisan <i>Final Fantasy XVI</i> pada tahun depan.	Ya	Terdapat entitas COMPANY (<i>Square Enix</i>) dan GAME (<i>Final Fantasy XVI</i>).
5	Pemain mengeluhkan bug grafis setelah pembaruan versi terbaru.	Tidak	Tidak terdapat entitas bernama atau istilah spesifik dalam domain <i>game</i> .
6	<i>miHoYo</i> merilis trailer baru untuk <i>Honkai: Star Rail</i> .	Ya	Mengandung entitas COMPANY (<i>miHoYo</i>) dan GAME (<i>Honkai: Star Rail</i>).
7	Kompetisi <i>MPL Indonesia Season 12</i> menampilkan performa luar biasa dari <i>RRQ</i> .	Ya	Terdapat entitas EVENT (<i>MPL Indonesia Season 12</i>) dan ORGANIZATION (<i>RRQ</i>).
8	<i>Steam</i> mengadakan diskon besar-besaran untuk <i>game indie</i> .	Ya	Entitas PLATFORM (<i>Steam</i>) relevan dalam konteks distribusi <i>game</i> .
9	Teknologi <i>Unreal Engine 5</i> memberikan peningkatan signifikan pada pencahayaan.	Ya	Mengandung entitas TECH (<i>Unreal Engine 5</i>).
10	Pemerintah mendorong pengembangan industri kreatif digital di Indonesia.	Tidak	Konteks terlalu umum dan tidak spesifik pada domain <i>game</i> .

Kriteria seleksi diterapkan secara ketat untuk menjaga relevansi korpus.

Setiap kalimat dievaluasi berdasarkan keberadaan entitas domain-spesifik sesuai skema entitas yang mencakup GAME, COMPANY, PLATFORM, EVENT, dan TECH. Kategori tersebut menggambarkan tren utama dalam industri *game* (Goh dkk., 2023; Politowski dkk., 2021; Rykała, 2020). Kalimat yang mengandung sedikitnya satu entitas relevan diterima, sedangkan kalimat bersifat umum atau tidak memuat entitas spesifik dieliminasi.

Proses evaluasi menghasilkan korpus yang padat dan terfokus karena seluruh kalimat diverifikasi oleh pakar industri yang memahami konteks istilah

secara mendalam. Hasil kurasi tercantum pada Tabel 3.3 yang menampilkan contoh seleksi kalimat beserta alasan penerimaan atau penolakannya. Struktur tabel menunjukkan kesesuaian antara keputusan anotator dan kriteria konseptual yang telah ditetapkan, memastikan hanya kalimat dengan nilai informatif tinggi yang dilanjutkan ke tahap anotasi berikutnya.

3.3.4 Anotasi Kalimat

Tahap anotasi mengubah kumpulan kalimat terpilih hasil kurasi menjadi dataset beranotasi yang terstruktur. Setiap kalimat diberi label secara manual pada token yang mengandung entitas penting berdasarkan pedoman yang disusun untuk menjaga konsistensi dan validitas antar anotator.

Proses anotasi berlandaskan tiga pilar metodologis. Pilar pertama mencakup definisi tipe entitas yang dirumuskan secara eksplisit agar anotator memiliki panduan seragam dalam memberi label. Skema ini memuat sepuluh kategori, meliputi entitas umum PERSON, LOCATION, dan ORGANIZATION (Budi dkk., 2005), serta entitas generik DATE/TIME dan QUANTITY (Fachri, 2014). Lima kategori tambahan yang bersifat spesifik pada domain *game* memperluas cakupan anotasi, yaitu GAME (misalnya *DreadOut*), COMPANY (misalnya *Agate*), EVENT (misalnya *IGDX*), PLATFORM (misalnya *PlayStation*), dan TECH (misalnya *Unity Engine*). Kategori ini mencerminkan karakteristik industri *game* dan membantu model mengenali konteks linguistik khas teks berita *game* (Goh dkk., 2023; Politowski dkk., 2021; Rykała, 2020).

Tabel 3.4 Contoh Anotasi dengan Format IOB2

Token	Label IOB2	Keterangan
Moonton	B-COMPANY	Awal dari entitas COMPANY.
merilis	O	Di luar entitas.
Mobile	B-GAME	Awal dari entitas Game.
Legends	I-GAME	Bagian dalam dari entitas Game yang sama.

di	O	Di luar entitas.
Manila	B-LOCATION	Awal dari entitas Lokasi.

Pilar kedua menggunakan skema penandaan IOB2 (*Inside*, *Outside*, *Beginning*) (Sang & Veenstra, 1999). Awalan B- menandai awal entitas, I- menunjukkan bagian dalam entitas, dan O menandai token di luar entitas. Pilar ketiga menerapkan platform open-source *Label Studio* (Heartex, 2025) untuk memfasilitasi anotasi melalui antarmuka visual yang efisien dan mudah digunakan. Tahap ini menghasilkan dataset yang relevan secara tematik dan andal sebagai dasar pelatihan model NER di domain *game*. Contoh penerapan format IOB2 disajikan pada Tabel 3.4.

3.3.5 Pembagian Dataset

Dataset beranotasi yang telah melalui proses validasi kemudian dibagi menjadi tiga subset yang berdiri secara independen: data latih (*training set*), data validasi (*validation set*), dan data uji (*test set*). Pembagian ini berfungsi untuk menjamin bahwa model dilatih dan dievaluasi secara objektif tanpa kebocoran informasi antara tahap pelatihan dan pengujian.

Tabel 3.5 Pembagian dan Fungsi Dataset

Jenis Dataset	Proporsi	Fungsi Utama
Data Latih (<i>Train</i>)	60%	Melatih model agar mampu mempelajari pola dan fitur linguistik dari data.
Data Validasi (<i>Validation</i>)	20%	Menyelaraskan hyperparameter serta memantau potensi overfitting selama proses pelatihan.
Data Uji (<i>Test</i>)	20%	Menilai performa akhir model secara objektif pada data yang belum pernah digunakan sebelumnya.

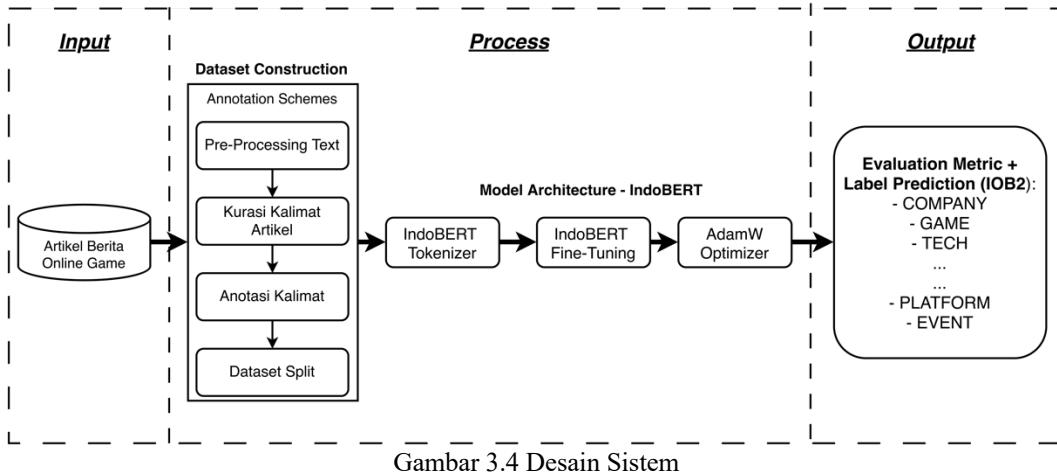
Rasio pembagian ditetapkan sebesar 60:20:20, yang merepresentasikan praktik umum dalam pengembangan model berbasis pembelajaran mesin (Muraina, 2022). Proporsi ini dinilai ideal untuk dataset berukuran menengah karena mampu menyediakan data yang memadai bagi proses pelatihan, penyetelan

hyperparameter, serta evaluasi akhir. Pembagian semacam ini juga membantu menyeimbangkan kebutuhan antara stabilitas model dan keandalan metrik evaluasi.

Masing-masing subset memiliki fungsi spesifik dalam siklus pengembangan model. Data latih berperan sebagai fondasi pembelajaran model untuk mengenali pola dan fitur linguistik dari teks. Data validasi digunakan selama pelatihan guna menilai kinerja model terhadap data yang belum dilihat serta membantu proses penyesuaian parameter agar menghindari overfitting. Data uji berfungsi sebagai acuan akhir untuk mengevaluasi kemampuan generalisasi model pada data yang benar-benar baru. Rincian pembagian dan fungsi setiap subset disajikan pada Tabel 3.5.

3.4 Desain Sistem dan Model

Desain penelitian ini diilustrasikan dalam Gambar 3.4 menunjukkan alur pemrosesan dari teks artikel *game* hingga menghasilkan prediksi entitas NER. Pada tahap preprocessing, teks diubah menjadi representasi numerik melalui AutoTokenizer yang memetakan kata menjadi token ID serta attention mask, kemudian Label Aligner menyesuaikan label anotasi agar sesuai dengan jumlah token, termasuk pemberian nilai -100 pada token yang tidak dihitung dalam loss. Selanjutnya, pada model layer, XLM-RoBERTa-Large menerima token ID dan mengubahnya menjadi contextual embeddings berukuran 1024 dimensi, yang kemudian melalui dropout untuk mencegah overfitting sebelum diproses oleh classification head linear yang menghasilkan probabilitas label untuk setiap token.



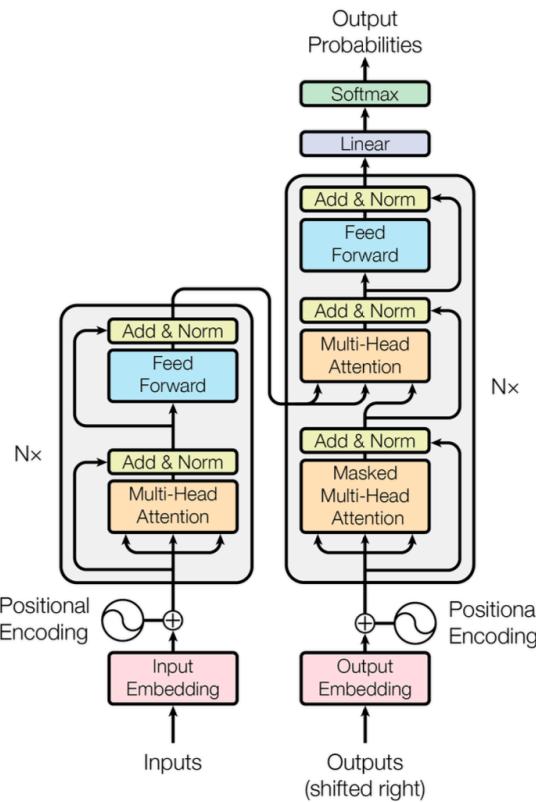
Pada tahap training, dilakukan forward pass untuk menghasilkan prediksi, perhitungan loss menggunakan Cross Entropy hanya pada label valid, serta backward pass dengan AdamW untuk memperbarui bobot model. Terakhir, evaluasi dilakukan menggunakan metrik *seqeval* yang menghitung *precision*, *recall*, dan *F1-score* berdasarkan kesesuaian prediksi entitas secara utuh, sehingga model dinilai tidak hanya mengenali token terpisah tetapi juga entitas pada kalimat.

3.4.1 Arsitektur Transformer

Model-model yang digunakan dalam penelitian ini didasarkan pada arsitektur Transformer, yang pertama kali diperkenalkan oleh Vaswani dkk (2017). Arsitektur ini merevolusi pemrosesan sekuensial dengan sepenuhnya meninggalkan mekanisme rekurensi (seperti pada RNN/LSTM) dan sebagai gantinya mengandalkan mekanisme atensi (*attention*) untuk menangkap dependensi global antara input dan output.

Secara fundamental, Transformer mempertahankan struktur *encoder-decoder* yang umum pada model sekuensial. *Encoder* bertugas memetakan sekuensial input dari representasi simbol (misalnya, token kata) (x_1, \dots, x_n) ke dalam

sekuensial representasi kontinu $z = (z_1, \dots, z_n)$. Berdasarkan z , *decoder* kemudian menghasilkan sekuensial output (y_1, \dots, y_m) token demi token.



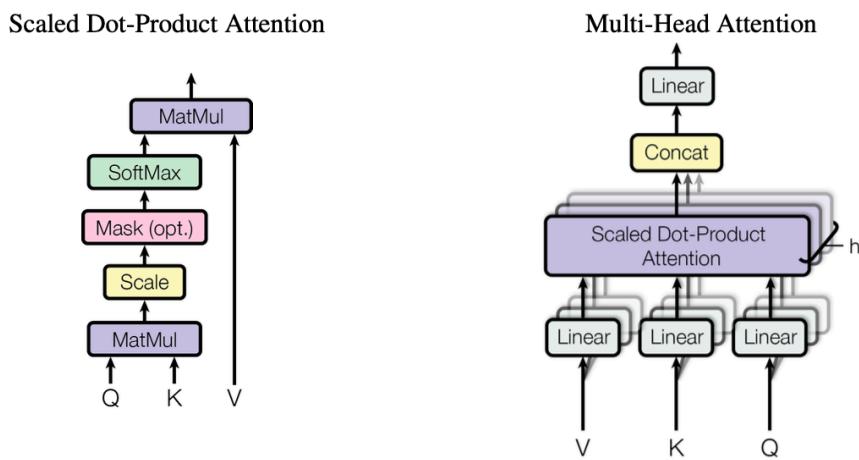
Gambar 3.5 Arsitektur Model Transformer (Vaswani dkk., 2017)

Seperti yang ditunjukkan pada Gambar 3.5, baik *encoder* maupun *decoder* tersusun dari tumpukan (*stack*) beberapa lapisan identik. Struktur encoder pada arsitektur Transformer terdiri atas enam lapisan identik yang disusun secara berurutan. Setiap lapisan memiliki dua sublapisan utama yang berperan penting dalam pemrosesan representasi konteks. Sublapisan pertama adalah *Multi-Head Self-Attention*, yaitu mekanisme yang memungkinkan setiap posisi dalam urutan input memperhatikan seluruh posisi lain pada lapisan sebelumnya. Mekanisme ini menjadi inti dari kemampuan Transformer dalam menangkap hubungan kontekstual antar-token secara global.

Sublapisan kedua adalah *Position-wise Fully Connected Feed-Forward Network*, berupa jaringan saraf sederhana yang terdiri atas dua transformasi linear

dengan fungsi aktivasi ReLU di antaranya untuk memperkuat non-linearitas dalam representasi data. Kedua sublapisan ini menerapkan koneksi residual yang diikuti oleh normalisasi lapisan (*Layer Normalization*). Dengan demikian, keluaran dari setiap sublapisan dirumuskan sebagai *LayerNorm* ($x + \text{Sublayer}(x)$), di mana *Sublayer*(x) merupakan fungsi yang dijalankan oleh sublapisan tersebut. Struktur ini menjaga stabilitas gradien selama pelatihan sekaligus mempertahankan informasi dari lapisan sebelumnya.

Transformer memanfaatkan mekanisme atensi yang disebut *Scaled Dot-Product Attention*. Atensi dapat diartikan sebagai fungsi yang memetakan sebuah *query* dan sekumpulan pasangan *key-value* ke sebuah output. Outputnya adalah jumlah tertimbang dari *values*, di mana bobot untuk setiap *value* dihitung berdasarkan fungsi kompatibilitas dari *query* dengan *key* yang bersangkutan.



Gambar 3.6 SDPA dan MHA (Vaswani dkk., 2017)

Scaled Dot-Product Attention sendiri inputnya terdiri dari *queries* dan *keys* dengan dimensi d_k , serta *values* dengan dimensi d_v . Perhitungan atensi dilakukan dengan mengambil produk titik (*dot product*) dari *query* dengan semua *keys*, membaginya dengan $\sqrt{d_k}$, lalu menerapkan fungsi softmax untuk mendapatkan

bobot pada *values*. Rumusnya ditunjukkan pada persamaan 3.1. di mana Q , K , dan V adalah matriks yang berisi kumpulan *queries*, *keys*, dan *values*.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3.1)$$

Transformer tidak hanya melakukan satu fungsi atensi, melainkan memproyeksikan *queries*, *keys*, dan *values* sebanyak h kali melalui proyeksi linear yang berbeda dan dapat dipelajari. Setiap hasil proyeksi menjalankan mekanisme atensi secara paralel, menghasilkan h keluaran yang merepresentasikan berbagai perspektif perhatian. Seluruh keluaran tersebut kemudian digabungkan (*concatenated*) dan diproyeksikan kembali melalui transformasi linear untuk memperoleh representasi akhir. Pendekatan ini memungkinkan model untuk secara simultan menangkap beragam jenis informasi dari berbagai subruang representasi serta posisi token yang berbeda, sehingga memperkaya pemahaman konteks dalam setiap lapisan. di mana $head_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ dan W adalah matriks parameter proyeksi yang dipelajari. Dalam paper aslinya, digunakan $h = 8$ kepala atensi paralel. Rumusnya ditunjukkan pada persamaan 3.2.

$$\text{MultiHead}(Q, K, V) = \text{Concat}(head_1, \dots, head_n)W^0 \quad (3.2)$$

Arsitektur Transformer tidak mengandung mekanisme rekurensi maupun konvolusi, sehingga informasi mengenai urutan token dalam sekuens harus disisipkan secara eksplisit. Mekanisme penyisipan tersebut dilakukan melalui Positional Encoding, yang memiliki dimensi sama dengan *embedding* input dan ditambahkan langsung ke dalam representasi token. Pendekatan ini memungkinkan model mempertahankan informasi posisi tanpa ketergantungan pada urutan pemrosesan. Vaswani dkk., (2017) merancang Positional

Encoding menggunakan fungsi sinus dan kosinus dengan frekuensi berbeda agar setiap posisi memiliki pola unik yang membantu model mengenali hubungan relatif antar-token secara efektif di mana pos adalah posisi token dan i adalah dimensi dari *embedding*. Rumusnya ditunjukkan pada persamaan 3.3.

$$\begin{aligned} PE_{(pos,2i)} &= \sin(pos / 10000^{2i / d_{model}}) \\ PE_{(pos,2i+1)} &= \cos(pos / 10000^{2i / d_{model}}) \end{aligned} \quad (3.3)$$

Penggunaan arsitektur berbasis atensi seperti Transformer didasari oleh keunggulannya dibandingkan model sekuensial tradisional seperti RNN. Arsitektur ini menawarkan efisiensi komputasi lebih tinggi, kemampuan paralelisasi yang lebih luas, serta jarak dependensi antar-token yang lebih pendek. Tabel 3.6 merangkum perbandingan ketiga aspek tersebut, yaitu kompleksitas komputasi per lapisan, jumlah operasi sekuensial yang dapat diparalelkan, dan panjang jalur maksimum antara dua posisi dalam jaringan.

Tabel 3.6 Perbandingan Tipe Lapisan (Vaswani dkk., 2017)

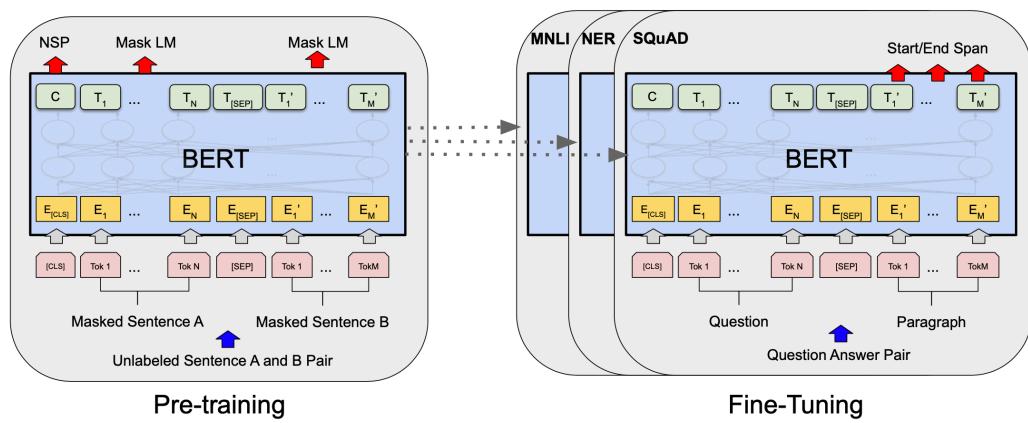
Tipe Lapisan	Kompleksitas per Lapisan	Operasi Sekuensial Minimum	Panjang Jalur Maksimum
Self-Attention	$O(n^2 \cdot d)$	$O(1)$	$O(1)$
Recurrent (RNN)	$O(n \cdot d^2)$	$O(n)$	$O(n)$
Convolutional (CNN)	$O(k \cdot n \cdot d^2)$	$O(1)$	$O(\log_k(n))$

Pada tabel 3.6 menunjukkan bahwa n merupakan panjang sekuens, d dimensi representasi, dan k ukuran kernel konvolusi. Hasil pada tabel memperlihatkan keunggulan signifikan lapisan *Self-Attention* yang memiliki panjang jalur maksimum konstan ($O(1)$) antara dua token mana pun dalam sekuens. Struktur tersebut memudahkan model mempelajari dependensi jarak jauh yang menjadi kendala utama pada RNN dengan panjang jalur ($O(n)$). Jumlah operasi sekuensial yang minimal ($O(1)$) juga memungkinkan paralelisasi

komputasi secara masif, sehingga proses pelatihan berlangsung lebih cepat dan efisien dibandingkan arsitektur RNN yang bersifat sekvensial.

3.4.2 BERT

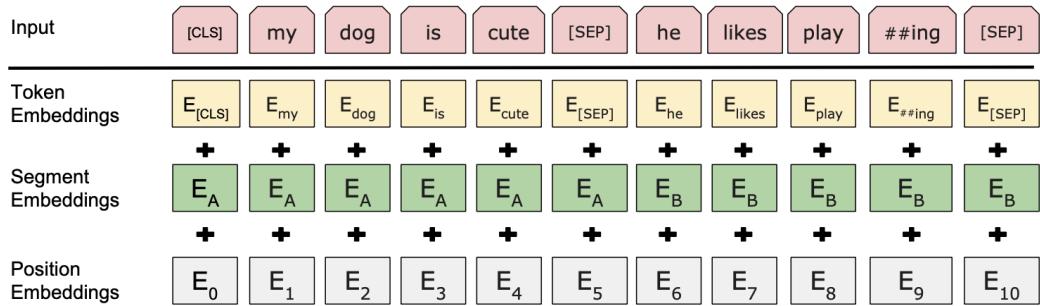
BERT (Devlin dkk., 2019) adalah model representasi bahasa yang dibangun di atas arsitektur *encoder* Transformer. Inovasi fundamental BERT adalah kemampuannya untuk melakukan pra-pelatihan (*pre-training*) representasi yang bidireksional secara mendalam dari teks tanpa label. Berbeda dengan model-model sebelumnya yang memproses teks secara satu arah (kiri-ke-kanan atau kanan-ke-kiri), BERT menggunakan seluruh konteks—baik dari kiri maupun kanan—secara bersamaan di semua lapisannya untuk membangun representasi kata yang kaya akan konteks. Kerangka kerja BERT terdiri dari dua tahap utama yaitu pra-pelatihan dan *fine-tuning*.



Gambar 3.7 Prosedur Pra-pelatihan dan Fine-tuning BERT (Devlin dkk., 2019)

BERT membangun representasi input yang mampu menangani satu kalimat maupun pasangan kalimat dalam satu urutan token guna mendukung berbagai tugas NLP. Setiap token direpresentasikan melalui penjumlahan tiga jenis *embedding*: *Token Embedding* yang mewakili bentuk kata menggunakan metode *WordPiece*, *Segment Embedding* yang menandai perbedaan antara kalimat pertama (A) dan kalimat kedua (B), serta *Position Embedding* yang menyisipkan informasi

posisi token sebagaimana *Positional Encoding* pada Transformer. Kombinasi ketiga komponen ini memungkinkan BERT memahami konteks antar-kalimat dan hubungan semantik secara mendalam. Selain itu, BERT memperkenalkan dua token spesial, [CLS] yang disisipkan di awal setiap sekuensial dan digunakan untuk tugas klasifikasi tingkat sekuensial, serta [SEP] yang digunakan untuk memisahkan kalimat.



Gambar 3.8 Representasi Input BERT (Devlin dkk., 2019)

Prosedur pra-pelatihan dan pelatihan lanjutan (*fine-tuning*) pada BERT menggunakan arsitektur yang sama, kecuali pada lapisan keluaran yang disesuaikan dengan tugas spesifik. Parameter hasil pra-pelatihan digunakan kembali sebagai inisialisasi untuk berbagai tugas *downstream*, kemudian seluruh parameter disesuaikan selama proses *fine-tuning*. Setiap masukan diawali dengan token khusus [CLS] sebagai representasi keseluruhan sekuens dan dipisahkan dengan token [SEP] yang berfungsi sebagai pemisah, misalnya antara pertanyaan dan jawaban.

3.4.3 IndoBERT

IndoBERT merupakan model representasi bahasa berbasis arsitektur Transformer yang dirancang secara khusus untuk bahasa Indonesia. Model ini dikembangkan oleh Koto dkk. (2020) sebagai solusi terhadap keterbatasan sumber daya dan model NLP yang berfokus pada bahasa Indonesia. Struktur IndoBERT

mengadaptasi arsitektur BERT-Base (uncased) (Devlin dkk., 2019), dengan 12 lapisan tersembunyi, dimensi vektor sebesar 768, 12 attention heads, serta lapisan feed-forward berukuran 3.072 dimensi.

Proses pra-pelatihan dilakukan dari awal menggunakan metode masked language modeling pada korpus besar berbahasa Indonesia yang terdiri atas sekitar 220 juta kata (Koto dkk., 2021). Sumber data tersebut meliputi tiga kategori utama, yaitu Wikipedia Indonesia, artikel berita dari portal seperti Kompas, Tempo, dan Liputan 6, serta *Indonesian Web Corpus*. Ketiga sumber ini dipilih untuk merepresentasikan variasi bahasa tulis formal, jurnalistik, dan umum dalam konteks bahasa Indonesia.

Korpus pelatihan tersebut disusun menggunakan metode tokenisasi *WordPiece* dengan total ukuran kosakata sebanyak 31.923 token. Pada tugas *Named Entity Recognition* (NER), IndoBERT menunjukkan hasil terbaik dibandingkan model multibahasa seperti mBERT maupun model serumpun seperti MALAYBERT. Hasil uji pada dataset NER dari Universitas Gadjah Mada (UGM) dan Universitas Indonesia (UI) menunjukkan peningkatan skor F1 yang signifikan (Tabel 3.7).

Tabel 3.7 Hasil Model pada POS Tagging dan NER (Koto dkk., 2020)

Model	POS Tagging (Acc)	NER UGM (F1)	NER UI (F1)
BiLSTM-CRF	95.4	70.9	82.2
mBERT	96.8	71.6	82.2
MALAYBERT	96.8	73.2	87.4
IndoBERT	96.8	74.9	90.1

Keunggulan hasil tersebut memperkuat posisi IndoBERT sebagai model representasi bahasa yang mampu menyesuaikan diri dengan karakteristik linguistik bahasa Indonesia secara mendalam, sekaligus menjadi fondasi yang solid untuk penerapan dalam domain berita *game*.

3.4.4 mBERT (Multilingual BERT)

mBERT (multilingual BERT) adalah model Transformer yang dilatih pada korpus teks dari 104 bahasa secara bersamaan, termasuk bahasa Indonesia. Dalam penelitian ini, mBERT digunakan sebagai salah satu model pembanding (*baseline*) yang kuat untuk mengevaluasi kinerja model monolingual IndoBERT. Model ini dikembangkan oleh Google dengan arsitektur BERT-Base (*cased*), yang memiliki 12 lapisan Transformer, 768 dimensi tersembunyi (*hidden size*), dan 12 *attention heads*.

Secara teknis, mBERT mengikuti arsitektur Transformer standar. Setiap blok Transformer terdiri dari dua sub-lapisan utama: *Multi-Head Self-Attention* (MHSA) dan sebuah jaringan *Feed-Forward* (FF). Menurut Wu & Dredze (2019), proses transformasi representasi pada setiap blok dapat dapat dilihat pada persamaan 3.4.

$$h^{i+1} = \text{Skip}(\text{FF}, \text{Skip}(\text{MHSA}, h^i)) \quad (3.4)$$

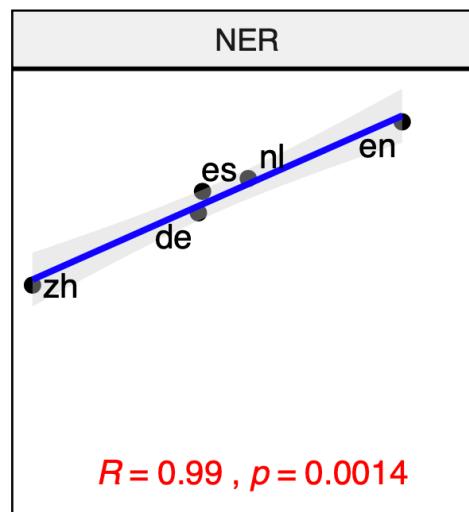
yang mana berdasarkan persamaan 3.4, fungsi '*Skip*' didefinisikan pada persamaan 3.5.

$$\text{Skip}(f, h) = \text{LayerNorm}(h + \text{Dropout}(f(h))) \quad (3.5)$$

Di sini, h^i adalah output dari blok ke- i , *LayerNorm* adalah normalisasi lapisan, dan *Dropout* digunakan untuk regularisasi. Salah satu kunci dari kemampuan mBERT adalah penggunaan satu kosakata bersama (*shared vocabulary*) yang dibangun dengan algoritma *WordPiece* untuk semua bahasa.

Kekuatan utama mBERT terletak pada kemampuannya untuk melakukan *zero-shot cross-lingual transfer*. Artinya, model yang dilatih pada tugas tertentu dalam satu bahasa (misalnya, NER dalam bahasa Inggris) dapat diterapkan

secara langsung pada bahasa target (seperti bahasa Indonesia) tanpa perlu *fine-tuning* pada data bahasa target. Kemampuan ini muncul secara implisit karena model memetakan token atau sub-kata yang serupa (seperti nama, angka, atau kata serumpun) dari berbagai bahasa ke dalam ruang representasi yang sama. Efektivitas transfer ini, sebagaimana ditunjukkan oleh Wu & Dredze, (2019), memiliki korelasi positif yang kuat dengan tumpang tindih sub-kata antara bahasa sumber dan target. Seperti diilustrasikan pada Gambar 3.9, semakin tinggi persentase tumpang tindih kosakata *WordPiece* antara bahasa target dengan bahasa sumber, maka kinerja model pada tugas NER cenderung semakin baik.



Gambar 3.9 Hubungan antara Kinerja *Zero-Shot* mBERT (Wu & Dredze (2019)

Justifikasi penggunaan mBERT sebagai model pembanding didukung oleh kinerjanya yang terbukti unggul pada tugas *Named Entity Recognition* (NER) lintas bahasa. Hasil penelitian Wu & Dredze, (2019) menunjukkan bahwa mBERT secara signifikan mengungguli *baseline state-of-the-art* pada masanya (Xie dkk., 2018).

Tabel 3.8 mBERT Terhadap Baseline (Wu & Dredze, 2019)

Bahasa Target	Baseline (Xie dkk., 2018) <i>F1-Score</i>	Kinerja mBERT <i>F1-Score</i>
Belanda (nl)	71.25	77.57
Spanyol (es)	72.37	74.96
Jerman (de)	57.76	69.56
Rata-rata	67.13	74.03

Data pada tabel 3.8 menunjukkan bahwa mBERT memiliki kemampuan representasi antarbahasa yang efektif dalam mengenali entitas, bahkan tanpa pelatihan spesifik pada bahasa target. Keunggulan ini menjadikan mBERT model pembanding yang valid dan kompetitif terhadap IndoBERT.

Perbandingan antara IndoBERT dan mBERT memberikan dasar evaluasi yang jelas terhadap perbedaan antara model monolingual dan multilingual. Analisis hasil kinerja keduanya memungkinkan identifikasi sejauh mana spesialisasi pada satu bahasa memberikan keunggulan dibandingkan pendekatan generalis dalam konteks tugas NER pada teks berita *game* berbahasa Indonesia.

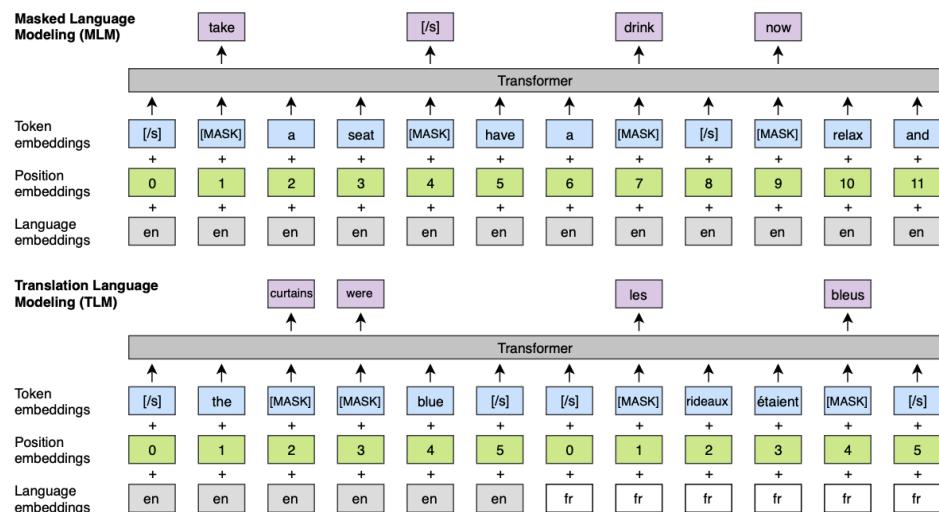
3.4.5 XLM-RoBERTa

XLM-RoBERTa (XLM-R) merupakan model representasi bahasa lintas-bahasa berbasis arsitektur Transformer yang berfokus pada penskalaan masif. Conneau dkk. (2019) mengembangkan model ini berdasarkan fondasi metodologis XLM (Lample & Conneau, 2019) dan membuktikan bahwa performa *state-of-the-art* dapat dicapai hanya dengan objektif *Masked Language Modeling* (MLM) ketika pelatihan dilakukan pada skala data dan kapasitas model yang jauh lebih besar.

Proses pra-pelatihan XLM-R menggunakan objektif MLM tanpa supervisi yang diadaptasi untuk konteks multibahasa. Model dilatih untuk memprediksi token yang disembunyikan dari konteks kalimat sekitarnya, serupa dengan BERT. Agar mampu menangani banyak bahasa, proses pelatihan memanfaatkan mekanisme pengambilan sampel teks dari berbagai bahasa menggunakan distribusi probabilitas yang menyeimbangkan bahasa sumber daya tinggi dan rendah. Distribusi ini dijelaskan pada Persamaan (3.6).

$$q_i = \frac{p_i^\alpha}{\sum_{j=1}^N p_j^\alpha} \text{ dengan } p_i = \frac{n_i}{\sum_{k=1}^N n_k} \quad (3.6)$$

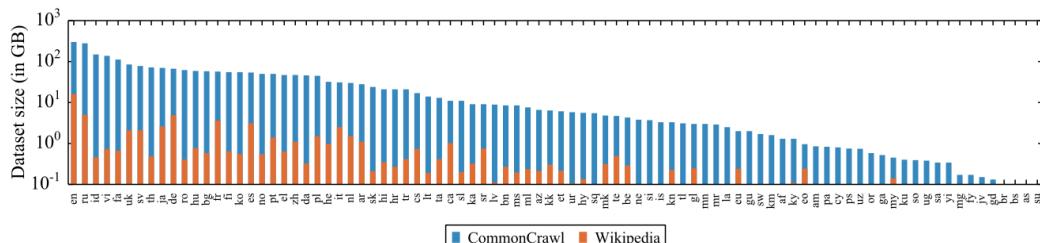
Di sini, n_i adalah jumlah kalimat dalam korpus bahasa ke- i , dan parameter α (diatur ke 0.3 untuk XLM-R) digunakan untuk meningkatkan frekuensi pengambilan sampel dari bahasa sumber daya rendah, mencegahnya "tenggelam" oleh bahasa sumber daya tinggi seperti Inggris. Meskipun model XLM asli juga memperkenalkan objektif *Translation Language Modeling* (TLM) yang memanfaatkan data paralel untuk mendorong penyelarasan lintas-bahasa (seperti diilustrasikan pada Gambar 3.10), XLM-R secara strategis hanya menggunakan objektif MLM. Keputusan ini didasari oleh hipotesis bahwa dengan data monolingual yang cukup besar dan kapasitas model yang memadai, penyelarasan representasi lintas-bahasa yang kuat dapat muncul secara implisit tanpa memerlukan data terjemahan yang seringkali sulit didapat.



Gambar 3.10 Ilustrasi Pra-pelatihan XLM-R (Lample & Conneau, 2019)

Inovasi utama XLM-R terletak pada tiga dimensi penskalaan ekstrem. Dari sisi skala data, model ini dilatih pada 100 bahasa menggunakan dataset *Common Crawl* (CC-100) yang telah difilter, mencakup lebih dari 2,5 terabyte teks. Jumlah ini meningkat beberapa kali lipat dibandingkan model terdahulu yang hanya

mengandalkan Wikipedia. Gambar 3.11 memperlihatkan perbedaan skala data antara Wikipedia dan *Common Crawl*, yang menunjukkan peningkatan signifikan bagi bahasa sumber daya rendah.



Gambar 3.11 Wikipedia dan *Common Crawl* (Conneau dkk., 2019).

Dari sisi kapasitas model, XLM-R menggunakan konfigurasi Transformer Large dengan kosakata bersama sebesar 250.000 token yang dibangun menggunakan Sentence Piece. Kosakata yang besar memungkinkan efisiensi dalam menangani berbagai bahasa sekaligus, serta mengurangi efek *curse of multilinguality*, penurunan performa akibat distribusi bahasa yang terlalu luas dibandingkan kapasitas model.

Tabel 3.9 Perbandingan Kinerja Lintas-Bahasa XNLI

Model	Dataset Pra-pelatihan	Jumlah Bahasa	Rata-rata Akurasi (%)
mBERT	Wikipedia	104	66.3
XLM (MLM)	Wikipedia	100	71.3
XLM (MLM+TLM)	Wikipedia + Data Paralel	15	75.1
XLM-R (Base)	<i>Common Crawl</i> (CC-100)	100	76.2
XLM-R (Large)	<i>Common Crawl</i> (CC-100)	100	80.9

Kinerja XLM-R melampaui seluruh model multibahasa sebelumnya pada berbagai tugas, termasuk *Named Entity Recognition* (NER) dan *Cross-lingual Natural Language Inference* (XNLI). Hasil evaluasi pada *benchmark* XNLI menunjukkan peningkatan akurasi signifikan dibandingkan mBERT dan XLM, seperti yang ditunjukkan pada tabel 3.9.

Keunggulan kinerja tersebut membuktikan efektivitas XLM-R sebagai baseline multibahasa yang kuat. Skala data masif, kapasitas besar, serta fokus pada

objektif MLM menjadikannya model yang relevan untuk mengevaluasi sejauh mana pendekatan multibahasa mampu menandingi model monolingual seperti IndoBERT dalam tugas ekstraksi entitas pada teks berita *game* berbahasa Indonesia.

3.5 Pelatihan dan Evaluasi Model

Pelatihan dan evaluasi model berfokus pada proses pelatihan model *Transformer* yang telah dipilih dan evaluasi kinerjanya secara kuantitatif. Proses pelatihan dilakukan dengan mengikuti standar dari *benchmark* IndoNLU untuk memastikan perbandingan yang adil dan relevan dengan penelitian sebelumnya di bidang pemrosesan bahasa alami untuk Bahasa Indonesia. Selanjutnya, model yang telah dilatih dievaluasi menggunakan metrik standar untuk mengukur kemampuannya dalam tugas yang diberikan.

3.5.1 Pelatihan Model

Pelatihan model dilakukan dengan proses *fine-tuning* terhadap model *pre-trained* pada dataset yang telah disiapkan. Pengaturan dalam proses ini mengacu pada kerangka kerja yang digunakan dalam *benchmark* IndoNLU (Wilie dkk., 2020) untuk menjaga konsistensi dan komparabilitas hasil. Parameter-parameter kunci yang digunakan dalam proses pelatihan dirinci dalam tabel 3.10.

Tabel 3.10 Konfigurasi Hyperparameter untuk Pelatihan Model

Hyperparameter	Nilai	Deskripsi
<i>Batch Size</i>	16	Jumlah sampel data yang diproses dalam setiap iterasi pelatihan. Ukuran ini dipilih untuk menyeimbangkan efisiensi komputasi dan stabilitas gradien.
<i>Jumlah Epoch</i>	4	Jumlah total siklus di mana model akan melihat keseluruhan data latih. Empat epoch dianggap cukup untuk konvergensi model pada tugas <i>fine-tuning</i> .
<i>Optimizer</i>	AdamW	Algoritma optimasi yang merupakan varian dari Adam dan dirancang khusus untuk arsitektur <i>Transformer</i> , yang efektif dalam menangani <i>weight decay</i> .
<i>Learning Rate</i>	2e-5	Laju pembelajaran yang digunakan untuk menyesuaikan bobot model. Nilai ini umum digunakan untuk proses <i>fine-tuning</i> model <i>Transformer</i> agar tidak mengubah bobot <i>pre-trained</i> secara drastis.

3.5.2 Evaluasi Model

Penelitian ini menetapkan baseline kinerja pada GameNER-ID melalui proses *fine-tuning* beberapa model berbasis Transformer yang telah dilatih sebelumnya. Pendekatan ini mengikuti metodologi yang digunakan dalam *benchmark* standar seperti IndoNLU (Wilie dkk., 2020) dan IndoLEM (Koto dkk., 2021). Tahap evaluasi model menggunakan metrik standar tingkat entitas berupa *Precision*, *Recall*, dan *F1-Score*. Evaluasi ketat ini memerlukan prediksi yang benar untuk kedua aspek yaitu span (batasan) entitas dan tipe spesifiknya (Tabel 3.6).

Tabel 3.11 Pembagian dan Fungsi Dataset

Komponen Evaluasi	Deskripsi
True Positives (TP)	Entitas yang diprediksi yang cocok sempurna dengan entitas <i>ground-truth</i> , baik dalam <i>span</i> maupun tipe entitas.
False Positives (FP)	Entitas yang diprediksi tetapi tidak memiliki kecocokan dengan entitas <i>ground-truth</i> manapun.
False Negatives (FN)	Entitas <i>ground-truth</i> yang tidak berhasil diprediksi oleh model.

Precision mengkuantifikasi fraksi entitas yang diprediksi yang benar. Rumusnya disajikan dalam persamaan 3.7.

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \quad (3.7)$$

Recall mengukur proporsi dari semua entitas *ground-truth* yang berhasil diidentifikasi model dengan benar. Perhitungannya ditunjukkan dalam rumus 3.8.

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \quad (3.8)$$

F1-Score merupakan rata-rata harmonik dari *Precision* dan *Recall*. Metrik ini memberikan ukuran tunggal yang seimbang dan sangat berguna untuk tugas yang tidak seimbang seperti NER, karena mempertimbangkan trade-off antara kedua metrik. Rumusnya ditampilkan dalam persamaan 3.9.

$$F1\ Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (3.9)$$

Dharviyanti & Wilantika (2024) menyatakan bahwa akurasi tidak digunakan dalam evaluasi model NER karena tugas NER lebih sesuai dievaluasi menggunakan metrik *Precision*, *Recall*, dan *F1-Score*. Ketidakseimbangan kelas pada data NER di mana sebagian besar token berada pada label O (*non-entity*) menyebabkan akurasi dapat tampil tinggi meskipun model hanya memprediksi label O tanpa benar-benar mengenali entitas lain. Meskipun akurasi tidak digunakan, indikator tersebut tetap ditampilkan, dan rumus akurasi berada pada Rumus 3.10.

$$Accuration = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.10)$$

BAB IV

ANOTASI DATASET

Proses pembuatan dataset GameNER-ID terbagi ke dalam empat fase utama. Tahap akuisisi dan pra-pemrosesan mengumpulkan ribuan artikel berita dari berbagai sumber, lalu membersihkan teks dari noise membuat format menjadi standar agar bisa diolah. Tahap kurasi dan seleksi manual menyaring data mentah menjadi kalimat yang relevan dengan domain *game*. Tahap anotasi entitas memberi label secara manual pada kalimat terpilih menggunakan skema IOB2 dengan sepuluh entitas khusus industri *game*, yang didukung perangkat lunak untuk proses anotasi. Tahap validasi dan penanganan ambiguitas menyelesaikan konflik antar anotator agar label konsisten di seluruh dataset. Analisis statistik dan karakteristik final melengkapi proses ini dengan rincian pembagian data latih, validasi, dan uji, termasuk distribusi label entitas serta contoh anotasi yang menunjukkan struktur dan kualitas data.

4.1 Tahap Akuisisi dan Pra-pemrosesan Data Mentah

Tahap akuisisi data merupakan dasar dari pembuatan dataset GameNER-ID, yang mencakup dua langkah krusial yaitu pengumpulan data teks mentah dari sumber-sumber yang relevan dan pembersihan data tersebut melalui tahap pra-pemrosesan. Tujuan dari tahap ini adalah untuk menghasilkan kumpulan teks yang bersih, konsisten, dan berkualitas sebagai bahan dasar untuk proses kurasi dan anotasi manual pada tahap selanjutnya.

4.1.1 Pengumpulan Artikel dari Sumber Berita

Pemilihan sumber data dilakukan secara strategis untuk memastikan cakupan informasi yang luas dan relevan dengan industri *game* di Indonesia. Sumber data diambil dari CNN Indonesia, Kompas, dan Detik karena ketiganya merupakan portal berita umum yang paling banyak diakses di Indonesia (Pangerang, 2024). Pemanfaatan ketiga media tersebut memungkinkan identifikasi representasi industri *game* dalam diskursus berita nasional yang lebih luas.

Gamebrott dimasukkan sebagai sumber spesialis yang berfokus secara eksklusif pada berita *game* (Adiputra, 2019). Kehadiran media ini memberikan data dengan terminologi, entitas, dan konteks yang lebih mendalam serta spesifik terhadap industri *game*. Proses pengumpulan data awal dari keempat sumber data dirangkum didalam Tabel 4.1.

Tabel 4.1 Ringkasan Sumber Data Awal

Data Source	Number of Articles
CNN	3,676
Detik	1,311
<i>Gamebrott</i>	6,892
Kompas	4,829
Total Articles	16,708

Tabel 4.1 menunjukkan total 16.708 artikel yang berhasil terkumpul sebagai dataset mentah. Proses *web crawling* menggunakan pustaka Python BeautifulSoup (Richardson, 2007) dan Selenium (Sharma, 2019).

Untuk memastikan setiap artikel yang dikumpulkan memiliki relevansi tinggi dengan domain yang diteliti, proses *web crawling* memanfaatkan kata kunci pencarian "*game*". Penggunaan kata kunci ini berfungsi sebagai filter awal untuk menyaring konten yang secara eksplisit membahas topik permainan, sehingga mengoptimalkan efisiensi pada tahap kurasi manual selanjutnya.

Tabel 4.2 Pembersihan Data Artikel (Pra-pemrosesan)

Kategori	Deskripsi	Contoh
Normalisasi Spasi	Menghapus spasi/tab/baris berlebih, memangkas spasi di akhir	" Hello world\n" → "Hello world"
Karakter Khusus Awal	Menghapus simbol atau sumber di awal teks	"– Ministry..." → "Ministry..."
Tanda Kurung/Garis Miring	Memperbaiki spasi dalam tanda kurung dan dekat garis miring	"(AI / intelligence)" → "(AI/intelligence)"
Tanda Kutip	Memperbaiki tanda kutip ganda, spasi, dan tanda baca	"""Hello""world"" → ""Hello"" world
Format Mata Uang	Menyeragamkan penulisan mata uang	"Rp 5,000,000" → "Rp5,000,000"
Tanda Baca	Menghapus spasi sebelum titik/koma, menambahkan spasi setelahnya	"Hello ,world ." → "Hello, world."
Domain/URL	Memperbaiki spasi pada TLD, menjaga format URL	"website. com" → "website.com"
Format Angka	Menyeragamkan angka, persentase, dan desimal	"5, 5 percent" → "5.5 percent"
Format Khusus	Menjaga format waktu & entitas khusus	"HUB. ID" → "HUB.ID"

4.1.2 Pra-pemrosesan Teks Artikel

Tahap pra-pemrosesan dilakukan setelah artikel mentah terkumpul. Proses ini bertujuan membersihkan teks dari *noise* atau elemen yang tidak relevan sekaligus menstandarkan formatnya. Langkah tersebut penting untuk memudahkan proses anotasi manual serta meningkatkan konsistensi hasil (Vijayarani dkk., 2015). Rincian kategori pembersihan beserta contohnya tercantum pada Tabel 4.2.

Penerapan langkah-langkah pembersihan pada Tabel 4.2 menghasilkan teks yang lebih rapi, seragam, dan siap untuk proses anotasi manual. Hasil pra-pemrosesan ini juga memastikan setiap artikel memiliki konsistensi struktural yang diperlukan untuk melatih model NLP secara efektif.

4.2 Seleksi dan Kurasi Kalimat Artikel

Seleksi dan kurasi bertujuan mengurangi volume data mentah secara signifikan agar proses anotasi manual lebih terkelola dan terfokus. Penyaringan ribuan artikel memastikan hanya kalimat yang paling informatif dan relevan dengan

domain *game* yang diproses lebih lanjut. Langkah ini penting untuk menghasilkan dataset yang padat informasi sekaligus berkualitas tinggi.

Tabel 4.3 Kriteria Kurasi Kalimat Pada Artikel Terkait

Kategori Entitas	Ketentuan Seleksi	Contoh Kalimat
Judul <i>Game</i> (GAME)	Kalimat harus menyebutkan nama <i>game</i> secara eksplisit.	"PUBG Mobile berhasil meraih popularitas global."
Perusahaan (COMPANY)	Kalimat menyebutkan nama <i>developer</i> atau <i>publisher game</i> .	"Garena merilis update terbaru untuk Free Fire."
Platform (PLATFORM)	Kalimat mengandung informasi tentang platform tempat <i>game</i> dirilis/berjalan.	"Game ini tersedia di PlayStation 5 dan PC."
Acara/Turnamen (EVENT)	Kalimat menyebutkan acara atau kompetisi terkait <i>game/esports</i> .	"Piala Presiden Esports 2023 diikuti oleh tim internasional."
Teknologi (TECH)	Kalimat menyinggung teknologi yang digunakan dalam <i>game</i> .	"Game tersebut memanfaatkan teknologi Unreal Engine 5."
Lokasi (LOCATION)	Kalimat menyebutkan lokasi yang relevan dengan industri <i>game</i> .	"Turnamen digelar di Jakarta Convention Center."
Waktu (TIME)	Kalimat menyebutkan waktu spesifik terkait peristiwa/ <i>game</i> .	"Update dirilis pada Januari 2023."
Kuantitas (QUANTITY)	Kalimat berisi angka yang relevan dengan <i>game</i> , seperti jumlah investasi, hadiah, atau pendapatan.	"Lebih dari 10 juta pemain aktif setiap hari."

Tiga anotator dengan pengalaman profesional lebih dari lima tahun di industri *game* melaksanakan proses kurasi secara manual. Setiap kalimat dari artikel yang telah melalui pra-pemrosesan dievaluasi secara cermat. Kriteria seleksi menekankan bahwa kalimat harus memuat setidaknya satu entitas yang relevan dengan dunia *game* (Goh dkk., 2023; Politowski dkk., 2021; Rykała, 2020). Kriteria dalam menentukan proses seleksi ditunjukkan pada Tabel 4.3.

Tabel 4.4 Jumlah Data (Artikel) Setelah Kurasi

Source	Artikel	Kalimat	Tokens (Kata)
Kompas	574	1,347	28,505
CNN	253	421	9,342
Detik	232	450	10,491
Gamebrott	632	1,420	35,938
Total	1,691	3,638	84,276

Proses seleksi yang ketat mereduksi jumlah data dari 16.708 artikel menjadi 1.691 artikel pilihan dengan total 3.638 kalimat siap anotasi. Proses kurasi yang

melibatkan evaluasi manual terhadap ribuan kalimat ini merupakan tahap yang penting. Pelaksanaannya oleh tiga anotator berpengalaman membutuhkan waktu sekitar 1,5 minggu untuk diselesaikan. Durasi ini mencerminkan tingkat ketelitian yang diterapkan untuk memastikan bahwa hanya kalimat-kalimat paling penting yang lolos ke tahap anotasi. Hasil ringkasan seleksi tercantum pada Tabel 4.4.

Hasil kurasi pada Tabel 4.4 menghasilkan kumpulan kalimat yang ringkas, relevan, dan representatif terhadap industri *game* di Indonesia sesuai dengan ketentuan dari anotator. Dataset yang telah terpilih selanjutnya siap untuk melalui tahap anotasi manual, yang akan memberikan label entitas secara konsisten dan terstruktur.

4.3 Skema Anotasi Entitas

Setelah kalimat-kalimat yang relevan berhasil dikurasi, tahap selanjutnya adalah mendefinisikan kerangka kerja (*framework*) untuk proses anotasi entitas itu sendiri. Sub-bab ini secara fundamental menetapkan "aturan main" dalam pelabelan data, yang bertujuan untuk menangkap informasi spesifik dari domain *game* secara sistematis. Pembahasan akan mencakup tiga pilar utama. Pertama, definisi dari sepuluh tipe entitas yang telah ditetapkan. Kedua, akan diuraikan skema penandaan IOB2 (*Inside, Outside, Beginning*), sebuah metode teknis yang memastikan setiap token dalam kalimat diberi label, sehingga mampu menangani entitas yang terdiri dari satu atau beberapa kata. Terakhir, akan diperkenalkan perangkat lunak anotasi yang digunakan, yaitu Label Studio, yang menyediakan antarmuka visual untuk memfasilitasi proses pelabelan secara manual.

4.3.1 Definisi Tipe Entitas

Sebanyak sepuluh tipe entitas ditetapkan berdasarkan perannya yang penting dalam ekosistem industri *game* (Goh dkk., 2023; Politowski dkk., 2021; Rykała, 2020). Definisi dan deskripsi setiap tipe entitas ditunjukkan pada Tabel 4.5.

Setiap tipe entitas memiliki fungsi yang berbeda. COMPANY mencakup pengembang, penerbit, atau perusahaan teknologi *game*, misalnya Tencent, Electronic Arts, dan Anantarupa Studios. EVENT merujuk pada kompetisi atau pameran, seperti *Gamescom* dan Piala Presiden Esports. GAME mencatat judul spesifik, seperti *Valorant* atau *A Space for the Unbound*. LOCATION merepresentasikan lokasi geografis nyata maupun fiktional, misalnya Tokyo atau Jakarta.

Tabel 4.5 Nama dan Deskripsi Entitas dalam GameNER-ID

Tag NER	Deskripsi
COMPANY	Nama perusahaan <i>game</i>
EVENT	Nama acara atau turnamen <i>game</i>
GAME	Nama video <i>game</i>
LOCATION	Nama lokasi (nyata maupun fiktional)
ORGANIZATION	Nama asosiasi atau organisasi
PERSON	Nama individu
PLATFORM	Nama platform <i>game</i>
QUANTITY	Nilai numerik
TECH	Nama teknologi <i>game</i>
TIME	Referensi waktu

Entitas ORGANIZATION meliputi asosiasi atau lembaga pemerintah, seperti Asosiasi *Game* Indonesia atau Kominfo. PERSON menandai individu, misalnya Shigeru Miyamoto atau Hideo Kojima. PLATFORM mencakup perangkat keras maupun layanan digital, seperti *Nintendo Switch* atau *Google Play Store*. QUANTITY menandai nilai numerik, misalnya 5 juta kopi atau diskon 50%. TECH meliputi istilah teknis, genre, atau mesin *game*, seperti *Unreal Engine 5* atau

Virtual Reality (VR). Terakhir, TIME menandai referensi waktu spesifik, seperti 15 November 2023 atau kuartal ketiga 2024.

Pemisahan entitas COMPANY, TECH, dan PLATFORM dari kategori umum ORGANIZATION bersifat krusial. Pemisahan ini memungkinkan analisis granular, misalnya membedakan Ubisoft sebagai COMPANY, PlayStation sebagai PLATFORM, dan Unity Engine sebagai TECH. Detail semacam ini penting untuk analisis tren pasar maupun pemetaan ekosistem industri *game*.

4.3.2 Skema Penandaan IOB2

Proses anotasi menggunakan skema IOB2 (*Inside, Outside, Beginning*) untuk menandai setiap token dalam kalimat. Skema ini terbukti efektif pada tugas *sequence labeling* seperti NER (Sang & Veenstra, 1999). Setiap token diberi label sesuai posisinya terhadap entitas, seperti B-XXX yang menandai token pertama (*Beginning*) dari sebuah entitas bertipe XXX. Selanjutnya I-XXX menandai token di dalam (*Inside*) sebuah entitas, selain token pertama dan yangterakhir, O menandai token di luar (*Outside*) entitas manapun.

Tabel 4.6 Anotasi dalam Format XML dan IOB2

Format	Teks Teranotasi
XML	<COMPANY>Moonton</COMPANY> <O>mengumumkan</O> <GAME>Mobile Legends</GAME> <EVENT>M5 World Championship</EVENT> <O>di</O> <LOCATION>Manila</LOCATION>
IOB2	Moonton/B-COMPANY mengumumkan/O Mobile/B-GAME Legends/I- GAME M5/B-EVENT World/I-EVENT Championship/I- EVENT di/O Manila/B-LOCATION

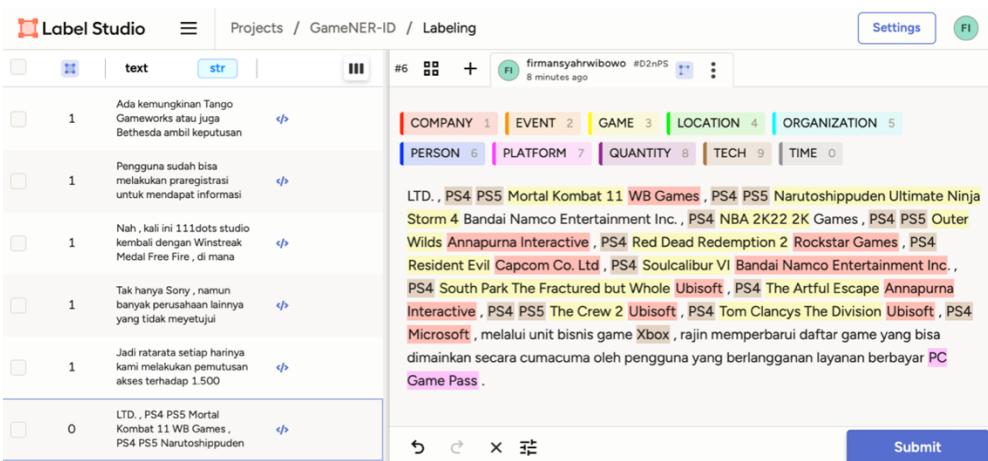
Tabel 4.6 berikut memperlihatkan bagaimana skema IOB2 diaplikasikan pada kalimat dengan entitas ganda. Entitas *Mobile Legends* ditandai sebagai B-GAME untuk kata *Mobile* dan I-GAME untuk kata *Legends*. Penerapan skema IOB2 yang konsisten di seluruh dataset ini memastikan bahwa struktur anotasi

bersifat seragam dan dapat diandalkan, yang merupakan ciri dari dataset berkualitas.

4.3.3 Perangkat Anotasi

Tim anotator menggunakan platform open-source Label Studio (Heartex, 2025) untuk melaksanakan seluruh proses anotasi manual. Alat tersebut menyediakan antarmuka visual yang intuitif sehingga anotator dapat menyorot satu atau beberapa token dalam sebuah kalimat, lalu menetapkan tag entitas yang sesuai seperti COMPANY, GAME, atau PLATFORM.

Gambar 4.1 memperlihatkan tampilan Label Studio saat anotator memilih kata atau frasa, kemudian menetapkan label entitas dari daftar yang tersedia. Antarmuka ini mempercepat proses kerja dan menjaga akurasi pelabelan di seluruh dataset.



Gambar 4.1 Proses Anotasi Pada Label Studio.

Gambar 4.1 memperlihatkan ilustrasi visual dari alur kerja anotasi. Panel kiri menampilkan daftar kalimat yang perlu dianotasi sehingga anotator dapat dengan mudah menavigasi antar kalimat dalam tugas mereka. Panel kanan berfungsi sebagai area kerja utama tempat kalimat yang sedang diproses ditampilkan, sementara anotator dapat secara interaktif menyorot kata atau frasa

yang relevan. Pada bagian atas area kerja, palet label menyediakan serangkaian entitas seperti COMPANY, EVENT, atau GAME yang diberi kode warna. Setelah menyorot teks, anotator cukup mengklik label yang sesuai untuk menerapkannya.

Label Studio menyediakan mekanisme pelabelan yang jelas dan mudah dipahami oleh anotator. Sistem ini menampilkan entitas dengan kode warna sehingga setiap kata atau frasa yang relevan dapat diberi label secara konsisten. Contoh konkret terlihat pada gambar, di mana frasa seperti *WB Games* dan *Ubisoft* mendapat label COMPANY dengan warna merah muda, judul seperti *Mortal Kombat 11* mendapat label GAME dengan warna kuning, dan platform seperti *PS4*, *Xbox*, serta *PC Game Pass* mendapat label PLATFORM dengan warna ungu. Antarmuka ini tidak hanya mempercepat proses anotasi tetapi juga membantu menjaga akurasi pelabelan di seluruh dataset.

4.4 Pelaksanaan dan Validasi Proses Anotasi

Fase anotasi manual untuk 3.638 kalimat yang telah dikurasi berlangsung selama kurang lebih 4,5 minggu. Dengan demikian, total waktu yang dibutuhkan dari tahap kurasi hingga selesainya anotasi adalah sekitar 6 minggu. Durasi yang signifikan ini merefleksikan kompleksitas dan tantangan yang ditemukan dalam proses pelabelan, terutama dalam menangani ambiguitas entitas yang pada akhirnya membentuk kualitas dari dataset final.

Validasi merupakan bagian integral dari proses anotasi. Salah satu tantangan terbesar yang teridentifikasi selama validasi adalah sifat polisemik dari terminologi industri *game*, di mana banyak istilah memiliki makna ganda tergantung pada konteks kalimat. Untuk mengatasi hal ini, sebuah protokol resolusi konflik yang ketat diterapkan guna menjamin konsistensi internal dataset.

4.4.1 Ambiguitas Entitas

Aturan penyelesaian konflik ketika proses anotasi dirancang untuk mengatasi situasi ketika para anotator memberikan label yang berbeda pada entitas yang sama. Dalam praktiknya, proses ini melibatkan dua pendekatan utama:

- 1 Pertama, para anotator menerapkan suara mayoritas untuk menentukan hasil akhir. Jika sebagian besar anotator memberi label tertentu pada sebuah entitas, maka label tersebut ditetapkan sebagai label final. Cara ini untuk menyelesaikan sebagian besar kasus perbedaan label karena keputusan akhir langsung mencerminkan kecenderungan terbesar.
- 2 Kedua, apabila tidak ada suara mayoritas yang jelas, seperti ketika jumlah suara terbagi rata, para anotator melanjutkan dengan diskusi bersama. Dalam sesi diskusi ini, mereka menelaah konteks kalimat secara lebih mendalam untuk memahami makna sebenarnya dari entitas yang bersangkutan. Proses analisis ini memungkinkan anotator mencapai kesepakatan yang dapat dipertanggungjawabkan, sehingga hasil anotasi tetap konsisten dan selaras dengan pedoman yang telah ditetapkan.

Penerapan aturan penyelesaian konflik ini memastikan bahwa setiap label yang dihasilkan tidak hanya berasal dari keputusan teknis semata, tetapi juga dari pemahaman konteks. Dengan demikian, kualitas dataset tetap terjaga karena semua entitas yang ambigu telah melalui proses validasi yang hati-hati. Beberapa contoh penerapan aturan ini pada kasus entitas yang sulit diinterpretasikan dapat dilihat pada Tabel 4.7.

Seperti yang ditunjukkan pada tabel, entitas seperti "Sega" diselesaikan sebagai COMPANY berdasarkan suara mayoritas, karena konteksnya lebih sering

merujuk pada perusahaan di balik konsol tersebut. Sebaliknya, "Unity" ditetapkan sebagai TECH, karena konteks kalimat ("dikembangkan menggunakan...") secara eksplisit menunjuk pada teknologi mesin *game*, bukan perusahaannya. Sementara itu, ambiguitas pada "Piala Presiden Esports" diselesaikan melalui diskusi deliberatif yang menyimpulkan bahwa istilah tersebut merujuk pada sebuah EVENT (acara spesifik), bukan ORGANIZATION (panitia penyelenggara). Penyelarasan yang sistematis dan cermat ini menjadi kunci untuk menjaga kualitas dan keandalan dataset GameNER-ID.

Tabel 4.7 Hasil Validasi Anotasi untuk Entitas Ambigu

Entitas Ambigu	Konteks Kalimat	Usulan Anotator	Label	Distribusi Suara (Hipotesis)	Label Final & Justifikasi
Sega	"...mengenang konsol legendaris dari Sega ."	COMPANY, PLATFORM		COMPANY: 9 PLATFORM: 3	COMPANY: Resolusi berdasarkan suara mayoritas, sesuai protokol yang dijelaskan dalam jurnal.
Unity	" <i>Game</i> ini dikembangkan menggunakan Unity ."	TECH, COMPANY		TECH: 8 COMPANY: 2	TECH: Konteks kalimat jelas merujuk pada mesin <i>game</i> , divalidasi oleh suara mayoritas.
Piala Presiden Esports	" Piala Presiden Esports akan segera dimulai."	EVENT, ORGANIZATION		EVENT: 5 ORGANIZATION: 4	EVENT: Diputuskan melalui diskusi deliberatif; merujuk pada sebuah turnamen (acara) spesifik, bukan panitia penyelenggaranya.

4.5 Karakteristik dan Statistik Final Dataset GameNER-ID

Setelah melalui seluruh tahapan, mulai dari akuisisi data, pra-pemrosesan, kurasi manual, hingga anotasi entitas, dataset GameNER-ID telah selesai dibangun. Sub-bab ini akan menyajikan karakteristik dan statistik akhir dari dataset yang dihasilkan, yang mencakup total 3.351 kalimat beranotasi. Analisis ini penting untuk memahami komposisi, distribusi, dan kualitas dataset sebelum digunakan

sebagai landasan untuk melatih dan mengevaluasi model *Named Entity Recognition* (NER).

4.5.1 Statistik Pembagian Dataset

Dataset GameNER-ID yang terdiri dari 3.351 kalimat beranotasi kemudian dibagi menjadi tiga bagian fungsional: data latih (*train*), data validasi (*validation*), dan data uji (*test*). Pembagian ini menggunakan rasio standar 60:20:20. Rasio ini dipilih karena mengikuti praktik yang direkomendasikan dalam pengembangan model, terutama untuk dataset berukuran sedang (Muraina, 2022).

Pembagian dataset dilakukan untuk memastikan model dapat dilatih, divalidasi, dan diuji secara seimbang. Bagian data latih (60%) menyediakan jumlah data yang cukup representatif sehingga model dapat mempelajari pola serta karakteristik entitas dalam domain *game*. Bagian data validasi (20%) berfungsi selama proses pelatihan untuk melakukan penyetelan *hyperparameter* sekaligus mencegah terjadinya *overfitting* tanpa membocorkan informasi dari data uji. Bagian data uji (20%) digunakan pada tahap akhir guna memberikan evaluasi yang objektif dan andal terhadap performa model yang telah dilatih. Ukuran yang memadai pada data validasi dan data uji memastikan evaluasi performa model tetap kuat serta dapat dipercaya. Rincian jumlah kalimat dan token untuk setiap bagian tercantum pada Tabel 4.8.

Pembagian ini menghasilkan dataset yang seimbang sekaligus representatif untuk mendukung pelatihan dan evaluasi model. Dengan struktur yang terkontrol, GameNER-ID dapat digunakan secara optimal untuk menguji kemampuan model dalam mengenali entitas pada domain *game*. Hasil ini juga memberikan dasar yang

baik untuk analisis lebih lanjut mengenai distribusi label dan entitas pada langkah selanjutnya.

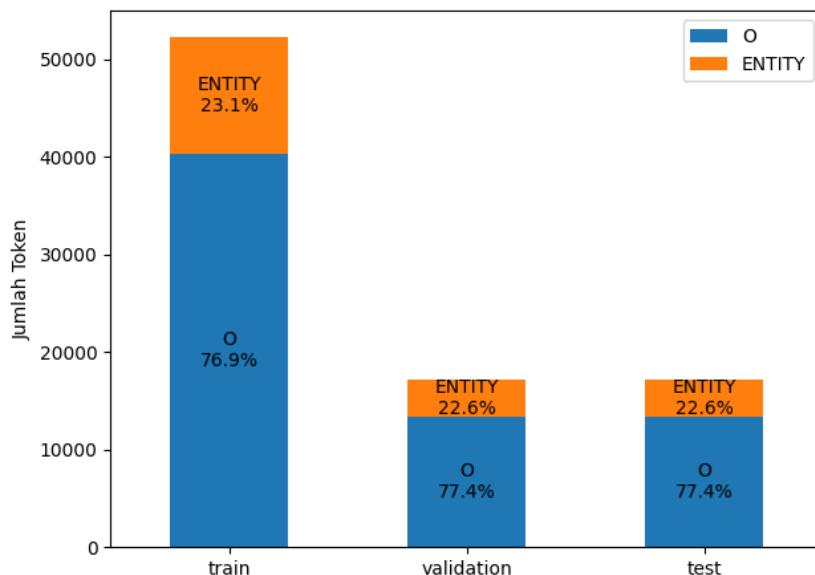
Tabel 4.8 Hasil Pembagian Dataset Menjadi Data Latih, Validasi dan Uji

Jenis Dataset	Jumlah Kalimat	Total Token (Kata)
Train	2,010	52,342
Validation	670	17,232
Test	671	17,214
Total	3,351	86,788

4.5.2 Distribusi Label dan Entitas

Analisis distribusi label entitas dilakukan untuk memastikan setiap bagian dataset (latih, validasi, dan uji) memiliki karakteristik serupa dan tidak bias.

Gambar 4.2 menyajikan perbandingan persentase token entitas (ENTITY) dengan token non-entitas (label 'O'). Visualisasi tersebut menunjukkan bahwa rasio token entitas dan non-entitas konsisten di ketiga bagian dataset, yaitu sekitar 23% token entitas dan 77% token non-entitas.



Gambar 4.2 Distribusi Label Entitas pada Set Data Latih, Validasi, dan Uji

Tabel 4.9 dan Tabel 4.10 menyajikan distribusi yang lebih rinci berdasarkan jenis entitas. Tabel 4.9 memperlihatkan distribusi agregat untuk sepuluh tipe entitas yang telah didefinisikan. Data tersebut menunjukkan bahwa entitas GAME muncul

dengan frekuensi tertinggi, diikuti oleh PLATFORM, COMPANY, TIME, dan TECH. Dominasi entitas-entitas ini mencerminkan karakteristik domain industri *game*, yang secara umum berfokus pada judul permainan, platform distribusi atau perangkat keras, perusahaan pengembang, serta referensi waktu dan teknologi yang terkait.

Tabel 4.9 Distribusi Entitas pada Setiap Bagian Dataset

Label	Train	Validation	Test
O	40,268	13,344	13,327
COMPANY	1,316	480	409
EVENT	436	155	199
GAME	4,743	1,438	1,437
LOCATION	316	95	124
ORGANIZATION	344	137	127
PERSON	772	255	229
PLATFORM	1,330	384	399
QUANTITY	557	144	212
TECH	1,012	371	322
TIME	1,067	319	347

Selanjutnya, Tabel 4.10 memberikan rincian yang lebih granular dengan memisahkan label berdasarkan skema penandaan IOB2 (*Inside*, *Outside*, *Beginning*). Tabel ini mengonfirmasi temuan dari distribusi agregat dan menunjukkan jumlah kemunculan untuk setiap tag B-(awal entitas) dan I- (di dalam entitas). Data ini krusial untuk analisis performa model pada level token yang lebih mendalam.

Analisis statistik yang telah dilakukan menegaskan kualitas sekaligus konsistensi dataset GameNER-ID yang telah dibangun. Distribusi label yang cukup seimbang antara token entitas dan non-entitas di seluruh bagian dataset (latih, validasi, dan uji) memastikan evaluasi model berlangsung tanpa bias. Frekuensi kemunculan entitas yang didominasi oleh GAME, PLATFORM, dan COMPANY juga menunjukkan bahwa dataset ini merepresentasikan diskursus utama dalam industri *game* Indonesia secara akurat. Konsistensi proporsi entitas di setiap bagian

menjadikan GameNER-ID sebagai *benchmark* yang andal untuk mengukur performa model NER secara objektif. Setelah memperoleh gambaran statistik dataset, bagian selanjutnya menyajikan contoh nyata hasil anotasi untuk memperlihatkan struktur data final secara lebih konkret.

Tabel 4.10 Distribusi Label Rinci per Split dengan Format IOB2

Tag	Train	Validation	Test
O	40,268	13,344	13,327
B-COMPANY	975	349	301
I-COMPANY	341	131	108
B-EVENT	142	46	57
I-EVENT	294	109	142
B-GAME	2,004	626	624
I-GAME	2,739	812	813
B-LOCATION	262	77	105
I-LOCATION	54	18	19
B-ORGANIZATION	344	137	127
I-ORGANIZATION	181	110	82
B-PERSON	517	160	154
I-PERSON	255	95	75
B-PLATFORM	891	256	270
I-PLATFORM	439	128	129
B-QUANTITY	266	72	99
I-QUANTITY	291	72	113
B-TECH	658	241	209
I-TECH	354	130	113
B-TIME	481	140	153
I-TIME	586	179	194

4.5.3 Hasil Anotasi

Sebagai ilustrasi hasil akhir proses anotasi, bagian ini menyajikan beberapa hasil kalimat dari dataset GameNER-ID yang telah diberi label sesuai dengan format IOB2. Tabel 4.11 menunjukkan bagaimana skema anotasi yang telah dijelaskan sebelumnya diterapkan dalam kalimat nyata untuk mengidentifikasi berbagai entitas spesifik domain.

Tabel 4.11 Hasil Anotasi pada Teks Berita Umum

No.	Kalimat dan Anotasi (Format IOB2)
1.	... PT(B-ORGANIZATION) Pertamina(I-ORGANIZATION) ((I-ORGANIZATION)Persero(I-ORGANIZATION))(I-ORGANIZATION) ... di kawasan Pantura(B-LOCATION) seperti Cirebon(B-LOCATION), Kuningan(B-LOCATION) ...

No.	Kalimat dan Anotasi (Format IOB2)
2.	Bouygues(B-ORGANIZATION) S.A(I-ORGANIZATION) .(I-ORGANIZATION), ... paruh pertama 1989(B-TIME), ... hingga 188(B-QUANTITY) juta(I-QUANTITY)francs(I-QUANTITY) ... \$(B-QUANTITY) 30.2(I-QUANTITY) juta(I-QUANTITY) ... dari 65(B-QUANTITY) juta(I-QUANTITY) francs(I-QUANTITY)...
3.	" Angel(B-PERSON) ... " kata Hajrul(B-PERSON) .
4.	Mereka meneriakkan nama Jokowi(B-PERSON) sambil mengibarkan bendera PDI - P.
5.	... di DPR(B-ORGANIZATION) atau di luar DPR(B-ORGANIZATION) ...
6.	... di Sriwijaya(B-LOCATION) Promotion(I-LOCATION) Center(I-LOCATION)Palembang(I-LOCATION), Sumatera(B-LOCATION) Selatan(I-LOCATION), Minggu(B-TIME) dini(I-TIME) hari(I-TIME), ... kontingen Propinsi DKI(B-LOCATION) Jakarta(I-LOCATION) ... juara umum adalah Propinsi Jawa(B-LOCATION) Tengah(I-LOCATION).
7.	... banyak kader PPP(B-ORGANIZATION) untuk diusung sebagai capres .
8.	... nama Kalla(B-PERSON) sering disebut akan didaulat ...
9.	... pengurus PAN(B-ORGANIZATION) kabupaten atau kota se(B-LOCATION) -(I-LOCATION) Sulawesi(I-LOCATION) Selatan(I-LOCATION), Zulkifli(B-PERSON) menyatakan akan ...
10.	... patung Alex(B-PERSON) dan Frans(B-PERSON) ... rumah adat Minahasa(B-LOCATION) ...
11.	... " ujar Benny(B-PERSON) dalam rapat , Jakarta(B-TIME) ,(I-TIME) Rabu(I-TIME) ((I-TIME) 4/2/2015(I-TIME))(I-TIME).
12.	... berangkat ke Afrika(B-LOCATION) menanyakan kepada saya ...
13.	... Direktur Pemasaran dan Niaga Pertamina(B-ORGANIZATION) Hanung(B-LOCATION) Budya(I-LOCATION) jelang Raker dengan Komisi VII di Gedung(B-LOCATION) DPR(I-LOCATION) -(I-LOCATION) RI(I-LOCATION), Senin(B-TIME) ((I-TIME) 25/8/2014(I-TIME))(I-TIME).
14.	... Dinas Pertanian dan Peternakan Kabupaten Kutai(B-LOCATION) Timur(I-LOCATION) ini .
15.	Terdata juga 42(B-QUANTITY) titik(I-QUANTITY) konflik(I-QUANTITY)sosial(I-QUANTITY) dan Indonesia(B-LOCATION) berada pada ...
16.	Partai(B-ORGANIZATION) Golongan(I-ORGANIZATION) Karya(I-ORGANIZATION) ... di Dewan(B-ORGANIZATION) Perwakilan(I-ORGANIZATION) Rakyat(I-ORGANIZATION), pekan lalu .
17.	... kinerja Jokowi(B-PERSON) sebagai Gubernur DKI Jakarta masih berantakan .
18.	Rodney(B-PERSON) Lyle(I-PERSON) Petersen(I-PERSON) (30(B-QUANTITY)) ... kata kantor berita AAP(B-ORGANIZATION) .
19.	Menurut Shallicar(B-PERSON), ... menewaskan Ghul(B-PERSON) ... di Gaza(B-LOCATION) City(I-LOCATION) yang ditumpangi Ghul(B-PERSON) .
20.	Masa inkubasi berlangsung dari 2(B-QUANTITY) -(I-QUANTITY) 21(I-QUANTITY) hari(I-QUANTITY) .
21.	Namun , NASA(B-ORGANIZATION) butuh bantuan ... katalog berisi 1,8(B-QUANTITY) juta(I-QUANTITY) foto(I-QUANTITY) .

No.	Kalimat dan Anotasi (Format IOB2)
22.	... Kepala BEI(B-ORGANIZATION) Medan(B-LOCATION), Muhammad(B-PERSON) Pintor(I-PERSON), ... pada September(B-TIME) 2014(I-TIME) .
23.	... bergerak ke arah barat laut dengan kecepatan 15(B-QUANTITY) kilometer(I-QUANTITY) per(I-QUANTITY) jam(I-QUANTITY) .
24.	... PPP(B-ORGANIZATION) siap untuk mendukung pemerintahan Jokowi - JK untuk periode 2014(B-TIME) -(I-TIME) 2019(I-TIME) .
25.	... keputusan Dewan Kehormatan Demokrat , Nazaruddin(B-PERSON) dicopot dari jabatan ...
26.	... upaya Angkasa(B-ORGANIZATION) Pura(I-ORGANIZATION) I(I-ORGANIZATION) dalam meningkatkan pendapatan ...
27.	... pihak Marriot(B-ORGANIZATION) pun belum bisa dihubungi .
28.	... mendukung penuh aturan KPU(B-ORGANIZATION) yang menyatakan jika ...
29.	... serangan tadi sekitar(B-TIME) pukul(I-TIME) 01.30(I-TIME) ... (04.30(B-TIME) WIB(I-TIME))(I-TIME) Minggu(I-TIME), ... antara Israel(B-LOCATION)dengan gerakan(B-ORGANIZATION) Islam(I-ORGANIZATION) Hamas(I-ORGANIZATION) di dalam dan sekitar(B-LOCATION) Gaza(I-LOCATION).
...	...
...	...
84276.	Firman(B-PERSON) meyakini ada perbuatan melawan hukum yang dilakukan KPU(B-ORGANIZATION).

Frasa panjang seperti "*PT Pertamina (Persero)*" atau "*Sriwijaya Promotion Center Palembang*" harus ditandai sebagai satu entitas utuh. Proses anotasi dimulai dengan memberi label B- (misalnya, B-ORGANIZATION) pada token pertama, seperti *PT* atau *Sriwijaya*. Selanjutnya, setiap kata, simbol, atau singkatan yang masih termasuk bagian dari nama resmi diberi label I-. Dengan cara ini, keseluruhan nama entitas tetap terjaga tanpa terpotong, sehingga maknanya tidak hilang.

Angka selalu diproses berdasarkan konteksnya. Pada frasa "*188 juta francs*", anotasi menggabungkan angka, satuan, dan keterangan jumlah sebagai satu entitas kuantitas (QUANTITY). Token *188* diberi label B-QUANTITY, sedangkan *juta* dan *francs* ditandai sebagai I-QUANTITY. Hal yang sama berlaku untuk rentang waktu seperti "*2 – 21 hari*", di mana seluruh frasa diperlakukan

sebagai satu entitas durasi. Pendekatan ini memastikan bahwa nilai numerik dan satuannya tetap konsisten sebagai satu informasi.

Kalimat dengan banyak entitas berdekatan, seperti pada contoh No. 19, menuntut ketelitian tinggi. Dalam frasa "*Menurut Shalicar, ... menewaskan Ghul ... di Gaza City ... ditumpangi Ghul*", setiap entitas harus diidentifikasi secara terpisah. *Shalicar* diberi label B-PERSON, *Ghul* pertama sebagai B-PERSON lain, *Gaza City* sebagai LOCATION, dan *Ghul* kedua kembali diberi label PERSON. Pemisahan ini penting untuk menjaga kejelasan struktur entitas di dalam kalimat yang kompleks.

Kasus ambiguitas menuntut pemahaman konteks yang lebih dalam. Pada contoh No. 13, frasa "...*Direktur Pemasaran dan Niaga Pertamina Hanung Budy*..." menunjukkan bahwa *Hanung Budy* adalah nama orang (PERSON), bukan lokasi. Penentuan ini dilakukan karena posisi nama mengikuti jabatan, sehingga lebih tepat dianggap sebagai individu. Pada contoh No. 32, frasa "*StatesWest Airlines, Phoenix, Ariz.*" secara keseluruhan diberi label ORGANIZATION, meskipun *Phoenix* dan *Ariz.* biasanya berfungsi sebagai nama lokasi. Keputusan ini mengikuti aturan yang disepakati agar anotasi tetap konsisten dengan konteks kalimat.

Entitas sederhana yang hanya terdiri dari satu token dapat ditandai dengan cepat. Nama seperti *Angel* (PERSON), *Jokowi* (PERSON), atau *PPP* (ORGANIZATION) langsung diberi label B- sesuai kategori yang tepat. Kalimat sederhana seperti ini mempercepat proses anotasi karena tidak memerlukan interpretasi mendalam. Proses anotasi ini mengubah teks mentah yang tidak terstruktur menjadi data berlabel.

Setelah melalui serangkaian tahapan yang sistematis mulai dari akuisisi, pra-pemrosesan, kurasi, hingga anotasi dan validasi dataset, GameNER-ID berhasil dibangun sebagai sumber daya yang konsisten dan berkualitas tinggi. Dengan tersedianya dataset ini, langkah selanjutnya adalah menggunakan untuk melatih dan mengevaluasi model *Named Entity Recognition* (NER). Oleh karena itu, bab berikutnya akan memfokuskan pembahasan pada penerapan model *baseline*, yaitu IndoBERT, yang akan *di-fine-tunemenggunakan dataset GameNER-ID untuk mengukur kinerja dasarnya dalam mengenali entitas pada domain game.*

BAB V

MODEL INDOBERT

Bab ini memberikan pembahasan terhadap hasil eksperimen yang telah dilakukan. Pembahasan difokuskan pada interpretasi kinerja model IndoBERT yang di *fine-tune* pada dataset GameNER-ID. Analisis ini tidak hanya melihat metrik evaluasi secara kuantitatif, tetapi juga mengontekstualisasikannya dengan temuan dari penelitian asli yang memperkenalkan IndoBERT, yaitu oleh Koto et al. (2020) dalam jurnal IndoLEM. Pembahasan akan mencakup evaluasi kinerja model secara keseluruhan, analisis performa per-tipe entitas, serta analisis kesalahan kualitatif untuk memahami kekuatan dan kelemahan model secara komprehensif.

5.1 Arsitektur Model

Pemilihan arsitektur model merupakan langkah fundamental yang menentukan arah dan potensi keberhasilan penelitian ini. Seperti yang dirangkum pada Tabel 5.1, model yang diadopsi sebagai baseline adalah indobert-uncased. Keputusan ini didasarkan pada beberapa pertimbangan strategis yang krusial. Pertama, indobert-uncased adalah model berbasis arsitektur Transformer yang secara inheren unggul dalam menangkap dependensi kontekstual jarak jauh dalam teks melalui mekanisme self-attention. Kemampuan ini sangat vital untuk tugas NER, di mana makna sebuah kata (dan apakah kata tersebut merupakan entitas) sangat bergantung pada kata-kata di sekitarnya.

Model indobert ini bersifat Monolingual. Artinya, seluruh kapasitasnya yang terdiri dari 124.5 juta parameter, didedikasikan untuk mempelajari representasi Bahasa Indonesia. Hal ini memberikan keunggulan signifikan

dibandingkan model multilingual (seperti mBERT) karena indobert-uncased memiliki pemahaman yang lebih mendalam dan bernuansa terhadap tata bahasa, kosakata, dan bahkan entitas yang umum dalam konteks lokal Indonesia. Pilihan ini selaras dengan temuan (Koto dkk., 2020), yang menunjukkan bahwa IndoBERT secara konsisten mengungguli model multilingual pada berbagai tugas NLP Bahasa Indonesia. Dengan demikian, penggunaan indobert-uncased tidak hanya mengikuti praktik terbaik, tetapi juga membangun di atas fondasi yang telah terbukti efektif untuk ekosistem bahasa target.

Tabel 5.1 Parameter Baseline Model

Model	Ukuran Model (#Params)	Tipe Bahasa
indobert-uncased	124.5M	Monolingual

5.2 Konfigurasi Pelatihan dan Hyperparameter

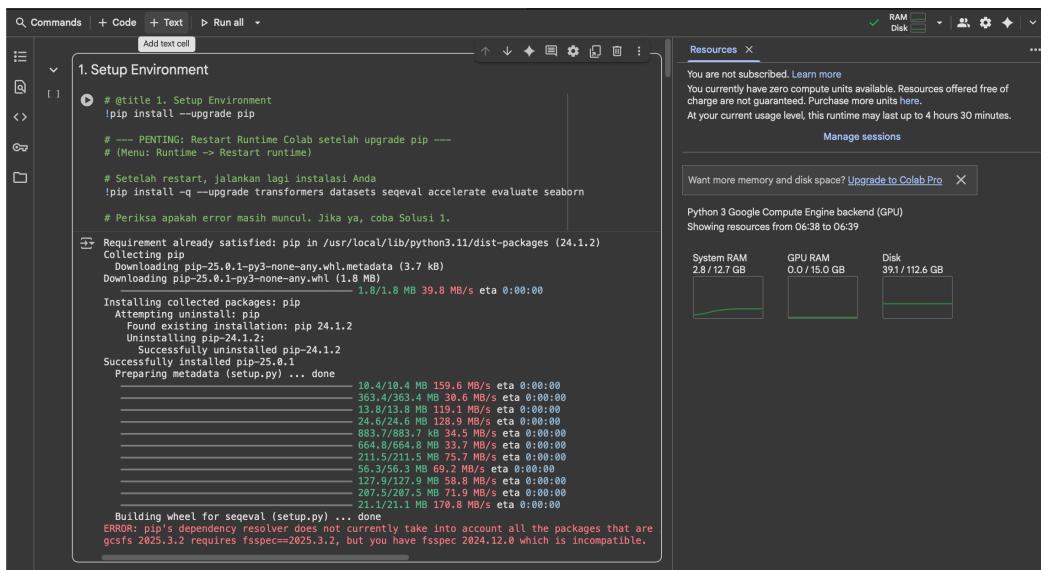
Model dilatih (*fine-tuned*) pada dataset GameNER-ID dengan konfigurasi *hyperparameter* mengikuti praktik yang terbukti efektif dalam penelitian sejenis, khususnya yang dilaporkan pada tolak ukur IndoNLU (Wilie dkk., 2020). Konfigurasi tersebut dipilih untuk menyeimbangkan efisiensi waktu pelatihan tanpa mengalami overfitting. Parameter yang digunakan selama fase pelatihan dirangkum dalam Tabel 5.2.

Tabel 5.2 Konfigurasi *Hyperparameter* untuk Pelatihan Model

Hyperparameter	Nilai	Deskripsi
Batch Size	16	Jumlah sampel data yang diproses dalam setiap iterasi pelatihan.
Jumlah Epoch	4	Jumlah total siklus ketika model melihat keseluruhan data pelatihan.
Optimizer	AdamW	Algoritma optimasi yang dirancang untuk arsitektur Transformer.
Learning Rate	2e-5	Laju pembelajaran yang digunakan untuk menyesuaikan bobot model pra-terlatih.

Seluruh proses eksperimen dijalankan melalui pemanfaatan GPU virtual T4 pada platform Google Colab (Gambar 5.1). Perangkat keras virtual tersebut

berperan penting dalam mengoptimalkan operasi matriks berskala besar yang menjadi inti dari mekanisme *self-attention* pada arsitektur Transformer. Dari sisi perangkat lunak, keseluruhan *pipeline* mulai dari pemuatan data, tokenisasi, hingga siklus pelatihan dan evaluasi diimplementasikan menggunakan ekosistem Hugging Face (Jain, 2022). Pustaka Transformers dari ekosistem ini menyediakan akses ke arsitektur model pra-terlatih dan *pipeline* pelatihan tingkat tinggi yang menyederhanakan proses *fine-tuning*.



The image shows a screenshot of the Google Colab interface. On the left, a code cell titled '1. Setup Environment' contains Python code for upgrading pip and installing the transformers library. The code output shows the upgrade of pip from 24.1.2 to 25.0.1, followed by the successful installation of transformers. A warning message at the bottom indicates compatibility issues with fsspec. On the right, a 'Resources' panel shows a GPU Compute Engine backend (GPU) with 0.0/15.0 GB of RAM and 39.1/112.6 GB of disk space. The interface includes standard Colab navigation and search bars at the top.

Gambar 5.1 Tampilan Antarmuka Google Colab

Untuk memastikan kompatibilitas penuh dan alur kerja yang efisien dari tahap persiapan data hingga pelatihan, dataset GameNER-ID secara spesifik distrukturkan dalam format Hugging Face *DatasetDict*. Penggunaan format ini menjamin integrasi yang mulus dengan arsitektur Transformer modern yang disediakan oleh pustaka tersebut (Jain, 2022). Seluruh arsitektur perangkat lunak ini berjalan di atas *framework deep learning* PyTorch. Secara keseluruhan, konfigurasi dalam Tabel 5.2 membentuk konfigurasi *fine-tuning* yang kokoh dan telah teruji sebelumnya. Tujuannya adalah memaksimalkan transfer pengetahuan dari IndoBERT ke tugas NER dalam domain video *game*.

5.3 Metrik Evaluasi Kerja

Kinerja model dievaluasi menggunakan metrik *Precision*, *Recall*, dan *F1-Score*, yang dihitung pada level entitas. Pemilihan metrik ini didasarkan pada kemampuannya untuk memberikan gambaran yang holistik, sementara metrik umum seperti Akurasi (*Accuracy*) secara sengaja tidak digunakan. Alasan utama pengabaian Akurasi adalah karena sifat dataset NER yang secara inheren tidak seimbang (*imbalanced*). Mayoritas token dalam korpus adalah token non-entitas (dilabeli 'O' atau '*Outside*'), sedangkan token yang menjadi bagian dari entitas hanya merupakan sebagian kecil. Jika Akurasi digunakan, sebuah model dengan hanya memprediksi label 'O' untuk semua token dapat mencapai skor Akurasi yang sangat tinggi, namun model tersebut sama sekali tidak berguna karena gagal mengidentifikasi satu pun entitas. Hal ini menjadikan Akurasi sebagai metrik yang menyesatkan dan tidak informatif untuk tugas ini.

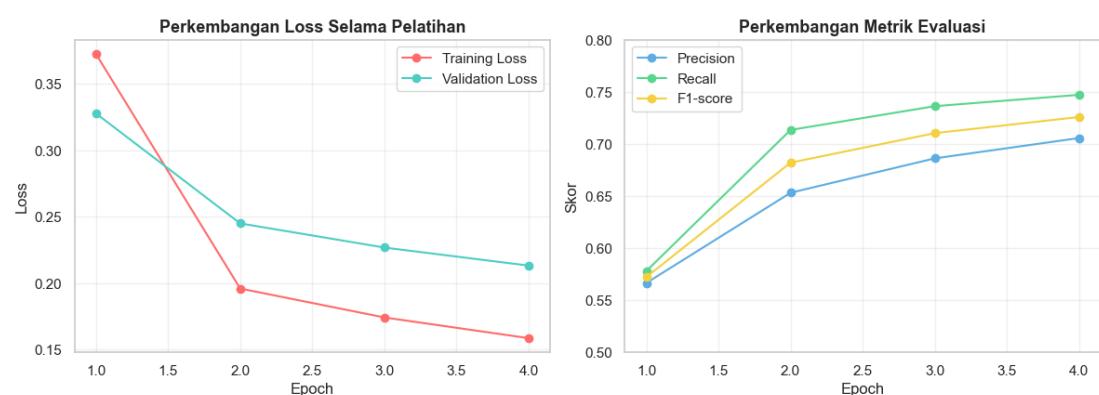
Tabel 5.3 Kasus Pada Perhitungan Metrik Evaluasi NER

Kasus	Ground Truth	Prediksi Model	Hasil Evaluasi
Identifikasi Sempurna (<i>True Positive</i>)	"Event [M5]_{EVENT} diadakan di [Manila]_{LOCATION}."	"Event [M5]_{EV} ENT</sub> diadakan di [Manila]_{LOCATION}."	TP=2, FP=0, FN=0 <i>Model mengidentifikasi semua entitas dengan benar.</i>
Kegagalan Deteksi (<i>False Negative</i>)	"Game [DreadOut]_{GAME} dibuat oleh [Digital Happiness]_{COMPANY}."	"Game DreadOut dibuat oleh [Digital Happiness]_{CO} MPANY</sub>."	TP=1, FP=0, FN=1 <i>Model melewati entitas "DreadOut".</i>
Prediksi Salah (<i>False Positive</i>)	"Update terbaru dirilis untuk [PlayStation 5]_{PLATFORM}."	"[Update terbaru]_{EVENT} dirilis untuk [PlayStation 5]_{PLATFORM}."	TP=1, FP=1, FN=0 <i>Model secara keliru menganggap "Update terbaru" sebagai entitas.</i>
Kesalahan Klasifikasi	"Gim ini dibuat oleh [Agate]_{COMPANY}."	"Gim ini dibuat oleh [Agate]_{ORGANIZATION}."	TP=0, FP=1, FN=1 <i>Span benar tapi label salah, dihitung sebagai FP (prediksi salah) dan FN (entitas COMPANY terlewat).</i>

Tabel 5.2 menyajikan analisa yang memperlihatkan cara kerja metrik evaluasi secara konkret. Perhitungan *True Positives* (TP), *False Positives* (FP), dan *False Negatives* (FN) dilakukan dengan membandingkan anotasi *ground-truth* dan prediksi model. Selain itu mengilustrasikan bahwa setiap kesalahan, baik itu melewatkannya entitas (FN) maupun membuat prediksi yang salah (FP), akan memberikan penalti pada skor akhir. Nilai TP, FP, dan FN yang diakumulasi dari seluruh data uji inilah yang kemudian dimasukkan ke dalam Persamaan (3.1), (3.2), dan (3.3) untuk menghasilkan skor *Precision*, *Recall*, dan *F1-Score* secara keseluruhan untuk setiap model. Dengan demikian, *F1-Score* menjadi representasi yang paling adil untuk mengukur performa model secara komprehensif

5.4 Hasil Pelatihan Model

Pada sub-bab ini selanjutnya akan disajikan analisis kuantitatif terhadap hasil dari proses *fine-tuning* model IndoBERT pada dataset GameNER-ID. Analisis disusun secara sistematis, dimulai dari evaluasi proses pelatihan untuk memastikan stabilitas model, kemudian menilai kinerja agregat pada data uji untuk mengukur efektivitas keseluruhan, dan diakhiri dengan deagregasi performa per kelas entitas untuk mengidentifikasi faktor yang mempengaruhi performa model.



Gambar 5.2 Proses Pelatihan Dataset GameNER-ID

Penelitian memvalidasi proses pelatihan sebagai langkah krusial agar hasil akhir bukan artefak dari pelatihan yang tidak stabil atau prematur. Gambar 5.2 menampilkan metrik kinerja pada *validation set* di setiap akhir epoch, termasuk Training Loss, Validation Loss, *Precision*, *Recall*, dan F1-score.

Analisis Tabel 5.4 mengungkap dua temuan utama. Pertama, *Training Loss* dan *Validation Loss* menurun secara monoton sepanjang empat epoch. Hasil ini menunjukkan model yang baik. Penurunan konsisten pada Validation Loss mengindikasikan model mampu melakukan generalisasi pada data yang tidak terlihat sebelumnya tanpa menunjukkan gejala overfitting.

Tabel 5.4 Metrik Hasil Pelatihan *Tain* Dataset GameNER-ID

Metric	Value
Epoch	4.0
Eval Accuracy	0.9373
Eval F1	0.7097
Eval Precision	0.6873
Eval Recal	0.7336
Eval Loss	0.2117
Eval Runtime	0:00:02.79

Kedua, F1-score pada validation set meningkat progresif dari 0,5723 pada epoch pertama hingga 0,7259 pada epoch keempat. Peningkatan ini menegaskan efektivitas proses *fine-tuning*, di mana setiap iterasi memperbaiki kemampuan diskriminatif model. Berdasarkan temuan ini, model hasil epoch keempat dipilih sebagai model final untuk evaluasi lebih lanjut pada dataset uji, karena merepresentasikan performa optimal yang dicapai selama pelatihan.

Temuan utama dari evaluasi Tabel 5.5 adalah model mencapai F1-score sebesar 0.7097. Metrik ini, yang juga direfleksikan oleh *Weighted avg* (0.70) dan *Micro avg* (0.71), mengindikasikan efektivitas model secara keseluruhan yang solid. Namun, perhatian khusus perlu diberikan pada perbedaan antara metrik *Weighted avg* dan *Macro avg*. *Macro avg* (0.63) yang jauh lebih rendah

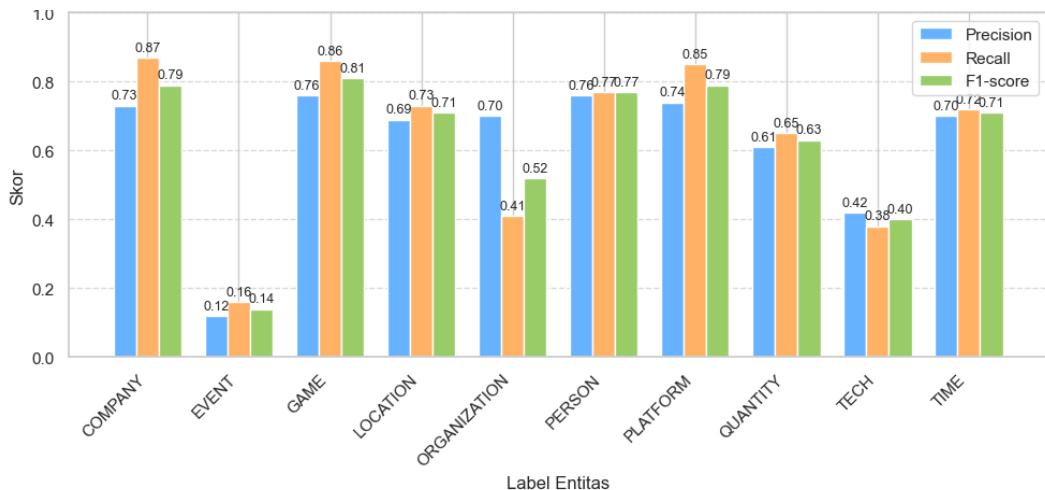
menunjukkan adanya perbedaan besar dalam kinerja antar-kategori entitas. Kesenjangan ini menunjukkan bahwa nilai keseluruhan yang tinggi terutama ditopang oleh kinerja baik pada kategori mayoritas, sedangkan kategori minoritas cenderung memiliki kinerja lebih rendah. Kondisi ini menuntut analisis lebih lanjut pada tingkat per kategori.

Tabel 5.5 Metrik Kinerja Per Entitas pada Dataset Uji

Label	Precision	Recall	F1-score	Accuracy	Support
COMPANY	0.73	0.87	0.79	0.9906	301
EVENT	0.12	0.16	0.14	0.9911	57
GAME	0.76	0.86	0.81	0.9795	624
LOCATION	0.69	0.73	0.71	0.9968	105
ORGANIZATION	0.70	0.41	0.52	0.9912	127
PERSON	0.76	0.77	0.77	0.9953	157
PLATFORM	0.74	0.85	0.79	0.9911	270
QUANTITY	0.61	0.65	0.63	0.9956	100
TECH	0.42	0.38	0.40	0.9836	209
TIME	0.70	0.72	0.71	0.9923	156
Micro avg	0.69	0.73	0.71	0.8800	2106
Macro avg	0.63	0.64	0.63	0.9920	2106
Weighted avg	0.68	0.73	0.70	0.9680	2106

Akurasi pada tugas *Named Entity Recognition* (NER) cenderung menunjukkan nilai yang sangat tinggi, bahkan ketika kinerja model terhadap entitas tertentu sebenarnya kurang baik. Hal ini terjadi karena distribusi token dalam korpus sangat tidak seimbang. Sebagian besar token merupakan kelas O (bukan entitas), sementara token entitas hanya menempati porsi kecil dari keseluruhan data. Dalam perhitungan akurasi, seluruh token non-entitas yang diprediksi dengan benar dihitung sebagai *true negative* (TN), sehingga jumlah TN menjadi sangat besar dan secara otomatis mendorong nilai akurasi mendekati 1, meskipun nilai *true positive* (TP) kecil atau *false negative* (FN) tinggi pada entitas tertentu. Oleh sebab itu, akurasi dianggap tidak representatif dan tidak digunakan sebagai metrik utama dalam evaluasi model NER. Metrik seperti *Precision*, *Recall*, dan *F1-score* jauh lebih informatif karena mampu menangkap kemampuan model dalam mengenali

entitas secara lebih akurat dan seimbang pada kelas yang jarang muncul. Untuk menginvestigasi hipotesis heterogenitas performa, dilakukan analisis deagregat pada setiap kelas entitas seperti yang terperinci dalam Gambar 5.3.



Gambar 5.3 Metrik Kinerja Per Entitas pada Dataset Uji

5.5 Analisa Kesalahan dan Perilaku Model

Penelitian ini melengkapi analisis kuantitatif dengan melakukan investigasi terhadap hasil prediksi model. Analisis ini bertujuan mendiagnosis pola kesalahan sistematis sekaligus memahami mekanisme perilaku model secara lebih mendalam. Tabel 5.6 beserta tabel-tabel ringkasan berikutnya menyajikan sampel prediksi yang menjadi dasar investigasi. Observasi terhadap data tersebut mengungkap beberapa kategori utama perilaku model. Pola kesalahan yang paling sering muncul adalah kesalahan klasifikasi, yaitu kondisi ketika model berhasil mendeteksi span teks sebagai entitas tetapi gagal menetapkan label kategori yang benar.

Tabel 5.6 Hasil Pelabelan Dataset Train

No	Komponen	Isi
1	Teks	UCL pic.twitter.com9oDw67wEhs <i>Game</i> Elden Ring mendapat popularitas yang tinggi karena dinilai berhasil <i>gameplay</i> yang menegangkan dan visual yang memukau .
	Label Asli	[('UCL', 'B-EVENT'), ('pic.twitter.com9oDw67wEhs', 'O'), ('Game', 'O'), ('Elden', 'B-GAME'), ('Ring', 'I-GAME'), ('mendapat', 'O'), ('popularitas', 'O'), ('yang', 'O'), ('tinggi', 'O'), ('karena', 'O'), ('dinilai', 'O'), ('berhasil', 'O'), ('gameplay', 'O'), ('yang', 'O'), ('menegangkan', 'O'), ('dan', 'O'), ('visual', 'O'), ('yang', 'O'), ('memukau', 'O'), ('.', 'O')]

No	Komponen	Isi
	Prediksi Pipeline	[('uc', 'B-TECH', 0.2723), ('##l', 'I-PLATFORM', 0.2090), ('pic', 'I-PLATFORM', 0.4342), ('twitter', 'B-PLATFORM', 0.8776), ('com', 'I-PLATFORM', 0.3622), ('##od', 'I-GAME', 0.5335), ('##w', 'I-GAME', 0.7011), ('##6', 'I-GAME', 0.8503), ('##7', 'I-GAME', 0.8578), ('##weh', 'I-GAME', 0.8497), ('##s', 'I-GAME', 0.7362), ('el', 'B-GAME', 0.9940), ('##den', 'I-GAME', 0.9897), ('ring', 'I-GAME', 0.9942)]
2	Teks	Sehingga , nantinya pemain bisa langsung melanjutkan petualangannya saat Final Fantasy XVI resmi meluncur 22 Juni nanti .
	Label Asli	[('Sehingga', 'O'), ('.', 'O'), ('nantinya', 'O'), ('pemain', 'O'), ('bisa', 'O'), ('langsung', 'O'), ('melanjutkan', 'O'), ('petualangannya', 'O'), ('saat', 'O'), ('Final', 'B-GAME'), ('Fantasy', 'I-GAME'), ('XVI', 'I-GAME'), ('resmi', 'O'), ('meluncur', 'O'), ('22', 'B-TIME'), ('Juni', 'I-TIME'), ('nant', 'O'), ('.', 'O')]
	Prediksi Pipeline	[('final', 'B-GAME', 0.9902), ('fantasy', 'I-GAME', 0.9920), ('xvi', 'I-GAME', 0.9935), ('22', 'B-TIME', 0.9527), ('juni', 'I-TIME', 0.9515)]
3	Teks	Keinginan ini pun semakin diperkuat oleh seorang fans di forum Reddit , di mana ia membuat karakter rendering Tyson sebagai fighter Tekken 8 dengan bantuan AI .
	Label Asli	[('Keinginan', 'O'), ('ini', 'O'), ('pun', 'O'), ('semakin', 'O'), ('diperkuat', 'O'), ('oleh', 'O'), ('seorang', 'O'), ('fans', 'O'), ('di', 'O'), ('forum', 'O'), ('Reddit', 'B-PLATFORM'), ('.', 'O'), ('di', 'O'), ('mana', 'O'), ('ia', 'O'), ('membuat', 'O'), ('karakter', 'O'), ('rendering', 'O'), ('Tyson', 'B-PERSON'), ('sebagai', 'O'), ('fighter', 'O'), ('Tekken', 'B-GAME'), ('8', 'I-GAME'), ('dengan', 'O'), ('bantuan', 'O'), ('AI', 'B-TECH'), ('.', 'O')]
	Prediksi Pipeline	[('redd', 'B-PLATFORM', 0.9233), ('##it', 'I-PLATFORM', 0.6390), ('ren', 'B-PERSON', 0.7711), ('##ing', 'I-PERSON', 0.4584), ('ty', 'B-PERSON', 0.6051), ('##son', 'I-PERSON', 0.9104), ('tek', 'B-GAME', 0.8294), ('##ken', 'I-GAME', 0.9845), ('8', 'I-GAME', 0.9875), ('ai', 'B-PERSON', 0.3977)]
4	Teks	Banyak pihak menyebutkan bahwa Microsoft akan merilis konsol Xbox generasi terbaru pada tahun 2028 mendatang .
	Label Asli	[('Banyak', 'O'), ('pihak', 'O'), ('menyebutkan', 'O'), ('bahwa', 'O'), ('Microsoft', 'B-COMPANY'), ('akan', 'O'), ('merilis', 'O'), ('konsol', 'O'), ('Xbox', 'B-PLATFORM'), ('generasi', 'O'), ('terbaru', 'O'), ('pada', 'O'), ('tahun', 'O'), ('2028', 'B-TIME'), ('mendatang', 'O'), ('.', 'O')]
	Prediksi Pipeline	[('microsoft', 'B-COMPANY', 0.9981), ('x', 'B-PLATFORM', 0.9274), ('##box', 'I-PLATFORM', 0.9932), ('2028', 'B-TIME', 0.9126)]
5	Teks	Sony Interactive Entertainment mengumumkan bahwa PlayStation 5 telah terjual sebanyak 40 juta unit di seluruh dunia .
	Label Asli	[('Sony', 'B-COMPANY'), ('Interactive', 'I-COMPANY'), ('Entertainment', 'I-COMPANY'), ('mengumumkan', 'O'), ('bahwa', 'O'), ('PlayStation', 'B-PLATFORM'), ('5', 'I-PLATFORM'), ('telah', 'O'), ('terjual', 'O'), ('sebanyak', 'O'), ('40', 'B-QUANTITY'), ('juta', 'I-QUANTITY'), ('unit', 'O'), ('di', 'O'), ('seluruh', 'O'), ('dunia', 'O'), ('.', 'O')]
	Prediksi Pipeline	[('sony', 'B-COMPANY', 0.9964), ('interactive', 'I-COMPANY', 0.9951), ('entertainment', 'I-COMPANY', 0.9948), ('play', 'B-PLATFORM', 0.9911), ('##station', 'I-PLATFORM', 0.9936), ('5', 'I-PLATFORM', 0.9872), ('40', 'B-QUANTITY', 0.9712), ('juta', 'I-QUANTITY', 0.9750)]
6	Teks	Nintendo Switch OLED dirilis pada Oktober 2021 dengan peningkatan layar dan kapasitas penyimpanan .
	Label Asli	[('Nintendo', 'B-COMPANY'), ('Switch', 'B-PLATFORM'), ('OLED', 'I-PLATFORM'), ('dirilis', 'O'), ('pada', 'O'), ('Oktober', 'B-TIME'), ('2021', 'I-TIME'), ('dengan', 'O'), ('peningkatan', 'O'), ('layar', 'O'), ('dan', 'O'), ('kapasitas', 'O'), ('penyimpanan', 'O'), ('.', 'O')]
	Prediksi Pipeline	[('nintendo', 'B-COMPANY', 0.9970), ('switch', 'B-PLATFORM', 0.9948), ('ole', 'I-PLATFORM', 0.8732), ('##d', 'I-PLATFORM', 0.9157), ('oktober', 'B-TIME', 0.9681), ('2021', 'I-TIME', 0.9599)]

No	Komponen	Isi
7	Teks	Rockstar Games mengumumkan trailer perdana Grand Theft Auto VI yang akan dirilis pada Desember 2025 .
	Label Asli	[('Rockstar', 'B-COMPANY'), ('Games', 'I-COMPANY'), ('mengumumkan', 'O'), ('trailer', 'O'), ('perdana', 'O'), ('Grand', 'B-GAME'), ('Theft', 'I-GAME'), ('Auto', 'I-GAME'), ('VI', 'I-GAME'), ('yang', 'O'), ('akan', 'O'), ('dirilis', 'O'), ('pada', 'O'), ('Desember', 'B-TIME'), ('2025', 'I-TIME'), ('.', 'O')]
	Prediksi Pipeline	[('rock', 'B-COMPANY', 0.9455), ('##star', 'I-COMPANY', 0.9810), ('games', 'I-COMPANY', 0.9968), ('grand', 'B-GAME', 0.9915), ('theft', 'I-GAME', 0.9937), ('auto', 'I-GAME', 0.9941), ('vi', 'I-GAME', 0.9950), ('desember', 'B-TIME', 0.9716), ('2025', 'I-TIME', 0.9604)]
8	Teks	Electronic Arts resmi memperkenalkan EA Sports FC sebagai pengganti franchise FIFA mulai tahun 2023 .
	Label Asli	[('Electronic', 'B-COMPANY'), ('Arts', 'I-COMPANY'), ('resmi', 'O'), ('memperkenalkan', 'O'), ('EA', 'B-COMPANY'), ('Sports', 'I-COMPANY'), ('FC', 'I-COMPANY'), ('sebagai', 'O'), ('pengganti', 'O'), ('franchise', 'O'), ('FIFA', 'B-GAME'), ('mulai', 'O'), ('tahun', 'O'), ('2023', 'B-TIME'), ('.', 'O')]
	Prediksi Pipeline	[('electronic', 'B-COMPANY', 0.9959), ('arts', 'I-COMPANY', 0.9967), ('ea', 'B-COMPANY', 0.9932), ('sports', 'I-COMPANY', 0.9918), ('fc', 'I-COMPANY', 0.9875), ('fifa', 'B-GAME', 0.9972), ('2023', 'B-TIME', 0.9619)]
9	Teks	Capcom mengumumkan Resident Evil 9 akan dirilis pada tahun 2026 dengan peningkatan grafis signifikan .
	Label Asli	[('Capcom', 'B-COMPANY'), ('mengumumkan', 'O'), ('Resident', 'B-GAME'), ('Evil', 'I-GAME'), ('9', 'I-GAME'), ('akan', 'O'), ('dirilis', 'O'), ('pada', 'O'), ('tahun', 'O'), ('2026', 'B-TIME'), ('dengan', 'O'), ('peningkatan', 'O'), ('grafis', 'O'), ('signifikan', 'O'), ('.', 'O')]
	Prediksi Pipeline	[('cap', 'B-COMPANY', 0.8207), ('##com', 'I-COMPANY', 0.9451), ('resident', 'B-GAME', 0.9938), ('evil', 'I-GAME', 0.9949), ('9', 'I-GAME', 0.9878), ('2026', 'B-TIME', 0.9542)]
10	Teks	Ubisoft berencana merilis Assassin's Creed Mirage pada paruh kedua 2023 .
	Label Asli	[('Ubisoft', 'B-COMPANY'), ('berencana', 'O'), ('merilis', 'O'), ('Assassin's', 'B-GAME'), ('Creed', 'I-GAME'), ('Mirage', 'I-GAME'), ('pada', 'O'), ('paruh', 'O'), ('kedua', 'O'), ('2023', 'B-TIME'), ('.', 'O')]
	Prediksi Pipeline	[('ubisoft', 'B-COMPANY', 0.9960), ('assassin', 'B-GAME', 0.9924), ('##', 'I-GAME', 0.6335), ('##', 'I-GAME', 0.7029), ('creed', 'I-GAME', 0.9940), ('mirage', 'I-GAME', 0.9931), ('2023', 'B-TIME', 0.9567)]
11	Teks	Di game Ragnarok Frontier , pemain dapat mengendalikan 5 karakter sekaligus , dan dengan fitur AutoHuntnya akan sangat memudahkan pemain dalam melakukan leveling maupun farming dari semua karakter yang dimiliki selama 24 Jam NONSTOP tanpa harus menyita banyak waktu bermain online , ujar Harry Choi , President of Gravity Game Link dalam keterangan yang diterima detikINET .
	Label Asli	[('Di', 'O'), ('game', 'O'), ('Ragnarok', 'B-GAME'), ('Frontier', 'I-GAME'), ('.', 'O'), ('pemain', 'O'), ('dapat', 'O'), ('mengendalikan', 'O'), ('5', 'O'), ('karakter', 'O'), ('sekaligus', 'O'), ('.', 'O'), ('dan', 'O'), ('dengan', 'O'), ('fitur', 'O'), ('AutoHuntnya', 'O'), ('akan', 'O'), ('sangat', 'O'), ('memudahkan', 'O'), ('pemain', 'O'), ('dalam', 'O'), ('melakukan', 'O'), ('leveling', 'O'), ('maupun', 'O'), ('farming', 'O'), ('dari', 'O'), ('semua', 'O'), ('karakter', 'O'), ('yang', 'O'), ('dimiliki', 'O'), ('selama', 'O'), ('24', 'O'), ('Jam', 'O'), ('NONSTOP', 'O'), ('tanpa', 'O'), ('harus', 'O'), ('menyita', 'O'), ('banyak', 'O'), ('waktu', 'O'), ('bermain', 'O'), ('online', 'O'), ('.', 'O'), ('ujar', 'O'), ('Harry', 'B-PERSON'), ('Choi', 'I-PERSON'), ('.', 'O'), ('President', 'O'), ('of', 'O'), ('Gravity', 'B-COMPANY'), ('Game', 'I-COMPANY'), ('Link', 'I-COMPANY'), ('dalam', 'O'), ('keterangan', 'O'), ('yang', 'O'), ('diterima', 'O'), ('detikINET', 'B-ORGANIZATION'), ('.', 'O')]

No	Komponen	Isi
	Prediksi Pipeline	[('di', 'O', 0.9983), ('game', 'O', 0.9981), ('ragnarok', 'B-GAME', 0.9934), ('frontier', 'I-GAME', 0.9919), ('', 'O', 0.9992), ('pemain', 'O', 0.9978), ('dapat', 'O', 0.9979), ('mengendalikan', 'O', 0.9980), ('5', 'O', 0.9944), ('karakter', 'O', 0.9982), ('sekaligus', 'O', 0.9975), ('', 'O', 0.9990), ('dan', 'O', 0.9985), ('dengan', 'O', 0.9987), ('fitur', 'O', 0.9982), ('autohunt', 'O', 0.8762), ('##nya', 'O', 0.8813), ('akan', 'O', 0.9986), ('sangat', 'O', 0.9983), ('memudahkan', 'O', 0.9980), ('pemain', 'O', 0.9976), ('dalam', 'O', 0.9979), ('melakukan', 'O', 0.9978), ('leveling', 'O', 0.8722), ('maupun', 'O', 0.9969), ('farming', 'O', 0.8899), ('dari', 'O', 0.9967), ('semua', 'O', 0.9970), ('karakter', 'O', 0.9978), ('yang', 'O', 0.9975), ('dimiliki', 'O', 0.9972), ('selama', 'O', 0.9970), ('24', 'O', 0.9912), ('jam', 'O', 0.9950), ('nonstop', 'O', 0.9422), ('tanpa', 'O', 0.9969), ('harus', 'O', 0.9967), ('menyita', 'O', 0.9965), ('banyak', 'O', 0.9967), ('waktu', 'O', 0.9964), ('bermain', 'O', 0.9962), ('online', 'O', 0.9951), ('', 'O', 0.9989), ('ujar', 'O', 0.9963), ('harry', 'B-PERSON', 0.9911), ('choi', 'I-PERSON', 0.9845), ('', 'O', 0.9987), ('president', 'O', 0.9941), ('of', 'O', 0.9940), ('gravity', 'B-COMPANY', 0.9812), ('game', 'I-COMPANY', 0.9634), ('link', 'I-COMPANY', 0.9599), ('dalam', 'O', 0.9961), ('keterangan', 'O', 0.9962), ('yang', 'O', 0.9965), ('diterima', 'O', 0.9960), ('detik', 'B-ORGANIZATION', 0.9238), ('##inet', 'I-ORGANIZATION', 0.4292), ('', 'O', 0.9993)]
12	Teks	Melalui akun Twitter resmi Counter Strike , pihak Valve merespon pelonjakan pemain dengan mem posting klip dari pegulat Randy Savage yang sedang menganggukan kepala , tentu dimaksud sebagai penggambaran rasa kagum dari tercetaknya rekor baru .
	Label Asli	[('Melalui', 'O'), ('akun', 'O'), ('Twitter', 'B-PLATFORM'), ('resmi', 'O'), ('Counter', 'B-GAME'), ('Strike', 'I-GAME'), ('', 'O'), ('pihak', 'O'), ('Valve', 'B-COMPANY'), ('merespon', 'O'), ('pelonjakan', 'O'), ('pemain', 'O'), ('dengan', 'O'), ('mem', 'O'), ('posting', 'O'), ('klip', 'O'), ('dari', 'O'), ('pegulat', 'O'), ('Randy', 'B-PERSON'), ('Savage', 'I-PERSON'), ('yang', 'O'), ('sedang', 'O'), ('menganggukan', 'O'), ('kepala', 'O'), ('', 'O'), ('tentu', 'O'), ('dimaksud', 'O'), ('sebagai', 'O'), ('penggambaran', 'O'), ('rasa', 'O'), ('kagum', 'O'), ('dari', 'O'), ('tercetaknya', 'O'), ('rekor', 'O'), ('baru', 'O'), ('', 'O')]
	Prediksi Pipeline	[('melalui', 'O', 0.9982), ('akun', 'O', 0.9985), ('twitter', 'B-PLATFORM', 0.9807), ('resmi', 'O', 0.9981), ('counter', 'B-GAME', 0.9679), ('strike', 'I-GAME', 0.9772), ('', 'O', 0.9991), ('pihak', 'O', 0.9980), ('val', 'B-COMPANY', 0.9812), ('##ve', 'I-COMPANY', 0.8901), ('merespon', 'O', 0.9974), ('pelonjakan', 'O', 0.9973), ('pemain', 'O', 0.9970), ('dengan', 'O', 0.9969), ('mem', 'O', 0.9967), ('posting', 'O', 0.9966), ('klip', 'O', 0.9964), ('dari', 'O', 0.9961), ('pegulat', 'O', 0.9959), ('randy', 'B-PERSON', 0.8985), ('sav', 'I-PERSON', 0.8828), ('##age', 'I-PERSON', 0.7069), ('yang', 'O', 0.9957), ('sedang', 'O', 0.9958), ('menganggukan', 'O', 0.9955), ('kepala', 'O', 0.9952), ('', 'O', 0.9989), ('tentu', 'O', 0.9956), ('dimaksud', 'O', 0.9954), ('sebagai', 'O', 0.9953), ('penggambaran', 'O', 0.9952), ('rasa', 'O', 0.9951), ('kagum', 'O', 0.9950), ('dari', 'O', 0.9949), ('tercetaknya', 'O', 0.9947), ('rekor', 'O', 0.9946), ('baru', 'O', 0.9944), ('', 'O', 0.9992)]
13	Teks	Saya dan tunangan saya harus membatalkan pernikahan kami karena COVID19 , jadi sahabatsahabat saya memberikan kejutan pernikahan di Animal Crossing , ujar Ashmush .
	Label Asli	[('Saya', 'O'), ('dan', 'O'), ('tunangan', 'O'), ('saya', 'O'), ('harus', 'O'), ('membatalkan', 'O'), ('pernikahan', 'O'), ('kami', 'O'), ('karena', 'O'), ('COVID19', 'B-EVENT'), ('', 'O'), ('jadi', 'O'), ('sahabatsahabat', 'O'), ('saya', 'O'), ('memberikan', 'O'), ('kejutan', 'O'), ('pernikahan', 'O'), ('di', 'O'), ('Animal', 'B-GAME'), ('Crossing', 'I-GAME'), ('', 'O'), ('ujar', 'O'), ('Ashmush', 'B-PERSON'), ('', 'O')]
	Prediksi Pipeline	[('saya', 'O', 0.9983), ('dan', 'O', 0.9981), ('tunangan', 'O', 0.9979), ('saya', 'O', 0.9978), ('harus', 'O', 0.9976), ('membatalkan', 'O', 0.9975), ('pernikahan', 'O', 0.9974), ('kami', 'O', 0.9972), ('karena', 'O', 0.9970),

No	Komponen	Isi
		('cov', 'B-PERSON', 0.6641), ('##id', 'I-PERSON', 0.3359), ('##19', 'I-PERSON', 0.2596), (',', 'O', 0.9988), ('jadi', 'O', 0.9969), ('sahabat', 'O', 0.9968), ('##s', 'O', 0.9967), ('##aha', 'O', 0.9966), ('##bat', 'O', 0.9965), ('saya', 'O', 0.9964), ('memberikan', 'O', 0.9963), ('kejutan', 'O', 0.9962), ('pernikahan', 'O', 0.9960), ('di', 'O', 0.9959), ('anim', 'B-GAME', 0.9199), ('##al', 'I-GAME', 0.9471), ('crossing', 'I-GAME', 0.9098), (',', 'O', 0.9987), ('ujar', 'O', 0.9957), ('ash', 'B-PERSON', 0.9511), ('##mus', 'I-PERSON', 0.8543), ('##h', 'I-PERSON', 0.9233), (',', 'O', 0.9991)]
15	Teks	Game besutan Mihoyo ini berhasil memikat banyak pemain karena kualitas grafisnya yang setara dengan game konsol .
	Label Asli	[('Game', 'O'), ('besutan', 'O'), ('Mihoyo', 'B-COMPANY'), ('ini', 'O'), ('berhasil', 'O'), ('memikat', 'O'), ('banyak', 'O'), ('pemain', 'O'), ('karena', 'O'), ('kualitas', 'O'), ('grafisnya', 'O'), ('yang', 'O'), ('setara', 'O'), ('dengan', 'O'), ('game', 'O'), ('konsol', 'O'), (',', 'O')]
	Prediksi Pipeline	[('game', 'O', 0.9984), ('besutan', 'O', 0.9981), ('mi', 'B-COMPANY', 0.8897), ('##hoyo', 'I-COMPANY', 0.8777), ('ini', 'O', 0.9980), ('berhasil', 'O', 0.9978), ('memikat', 'O', 0.9977), ('banyak', 'O', 0.9975), ('pemain', 'O', 0.9973), ('karena', 'O', 0.9971), ('kualitas', 'O', 0.9969), ('grafis', 'O', 0.9967), ('##nya', 'O', 0.9966), ('yang', 'O', 0.9965), ('setara', 'O', 0.9963), ('dengan', 'O', 0.9961), ('game', 'O', 0.9959), ('konsol', 'O', 0.9958), (',', 'O', 0.9992)]
15	Teks	Turnamen Mobile Legends tingkat Asia Tenggara digelar secara daring karena pandemi .
	Label Asli	[('Turnamen', 'O'), ('Mobile', 'B-GAME'), ('Legends', 'I-GAME'), ('tingkat', 'O'), ('Asia', 'B-LOCATION'), ('Tenggara', 'I-LOCATION'), ('digelar', 'O'), ('secara', 'O'), ('daring', 'O'), ('karena', 'O'), ('pandemi', 'B-EVENT'), (',', 'O')]
	Prediksi Pipeline	[('turnamen', 'O', 0.9983), ('mobile', 'B-GAME', 0.9823), ('legends', 'I-GAME', 0.9761), ('tingkat', 'O', 0.9979), ('asia', 'B-LOCATION', 0.9810), ('tenggara', 'I-LOCATION', 0.9788), ('digelar', 'O', 0.9977), ('secara', 'O', 0.9976), ('daring', 'O', 0.9975), ('karena', 'O', 0.9973), ('pandemi', 'B-EVENT', 0.8792), (',', 'O', 0.9991)]
16	Teks	Sony mengumumkan perilisan PlayStation 5 yang akan bersaing dengan Xbox Series X dari Microsoft .
	Label Asli	[('Sony', 'B-COMPANY'), ('mengumumkan', 'O'), ('perilisan', 'O'), ('PlayStation', 'B-CONSOLE'), ('5', 'I-CONSOLE'), ('yang', 'O'), ('akan', 'O'), ('bersaing', 'O'), ('dengan', 'O'), ('Xbox', 'B-CONSOLE'), ('Series', 'I-CONSOLE'), ('X', 'I-CONSOLE'), ('dari', 'O'), ('Microsoft', 'B-COMPANY'), (',', 'O')]
	Prediksi Pipeline	[('sony', 'B-COMPANY', 0.9833), ('mengumumkan', 'O', 0.9979), ('perilisan', 'O', 0.9977), ('playstation', 'B-CONSOLE', 0.9777), ('5', 'I-CONSOLE', 0.9722), ('yang', 'O', 0.9975), ('akan', 'O', 0.9974), ('bersaing', 'O', 0.9972), ('dengan', 'O', 0.9971), ('xbox', 'B-CONSOLE', 0.9751), ('series', 'I-CONSOLE', 0.9708), ('X', 'I-CONSOLE', 0.9681), ('dari', 'O', 0.9968), ('microsoft', 'B-COMPANY', 0.9822), (',', 'O', 0.9990)]
17	Teks	Blizzard Entertainment mengumumkan ekspansi terbaru World of Warcraft di acara BlizzCon .
	Label Asli	[('Blizzard', 'B-COMPANY'), ('Entertainment', 'I-COMPANY'), ('mengumumkan', 'O'), ('ekspansi', 'O'), ('terbaru', 'O'), ('World', 'B-GAME'), ('of', 'I-GAME'), ('Warcraft', 'I-GAME'), ('di', 'O'), ('acara', 'O'), ('BlizzCon', 'B-EVENT'), (',', 'O')]
	Prediksi Pipeline	[('blizzard', 'B-COMPANY', 0.9799), ('entertainment', 'I-COMPANY', 0.9765), ('mengumumkan', 'O', 0.9978), ('ekspansi', 'O', 0.9977), ('terbaru', 'O', 0.9975), ('world', 'B-GAME', 0.9733), ('of', 'I-GAME', 0.9701), ('warcraft', 'I-GAME', 0.9699), ('di', 'O', 0.9972), ('acara', 'O', 0.9970), ('blizz', 'B-EVENT', 0.8755), ('##con', 'I-EVENT', 0.8622), (',', 'O', 0.9992)]

No	Komponen	Isi
18	Teks	Nintendo Switch menjadi konsol terlaris pada tahun 2020 menurut laporan NPD Group .
	Label Asli	[('Nintendo', 'B-COMPANY'), ('Switch', 'B-CONSOLE'), ('menjadi', 'O'), ('konsol', 'O'), ('terlaris', 'O'), ('pada', 'O'), ('tahun', 'O'), ('2020', 'O'), ('menurut', 'O'), ('laporan', 'O'), ('NPD', 'B-ORGANIZATION'), ('Group', 'I-ORGANIZATION'), ('!', 'O')]
	Prediksi Pipeline	[('nintendo', 'B-COMPANY', 0.9828), ('switch', 'B-CONSOLE', 0.9733), ('menjadi', 'O', 0.9978), ('konsol', 'O', 0.9976), ('terlaris', 'O', 0.9975), ('pada', 'O', 0.9974), ('tahun', 'O', 0.9972), ('2020', 'O', 0.9971), ('menurut', 'O', 0.9969), ('laporan', 'O', 0.9967), ('npd', 'B-ORGANIZATION', 0.9022), ('group', 'I-ORGANIZATION', 0.8966), ('!', 'O', 0.9991)]
19	Teks	Epic Games meluncurkan Fortnite Chapter 2 Season 5 dengan berbagai fitur baru .
	Label Asli	[('Epic', 'B-COMPANY'), ('Games', 'I-COMPANY'), ('meluncurkan', 'O'), ('Fortnite', 'B-GAME'), ('Chapter', 'I-GAME'), ('2', 'I-GAME'), ('Season', 'I-GAME'), ('5', 'I-GAME'), ('dengan', 'O'), ('berbagai', 'O'), ('fitur', 'O'), ('baru', 'O'), ('!', 'O')]
	Prediksi Pipeline	[('epic', 'B-COMPANY', 0.9811), ('games', 'I-COMPANY', 0.9766), ('meluncurkan', 'O', 0.9977), ('fortnite', 'B-GAME', 0.9721), ('chapter', 'I-GAME', 0.9688), ('2', 'I-GAME', 0.9644), ('season', 'I-GAME', 0.9633), ('5', 'I-GAME', 0.9620), ('dengan', 'O', 0.9975), ('berbagai', 'O', 0.9973), ('fitur', 'O', 0.9971), ('baru', 'O', 0.9969), ('!', 'O', 0.9992)]
20	Teks	Steam menggelar diskon besar pada musim panas untuk berbagai judul game populer .
	Label Asli	[('Steam', 'B-PLATFORM'), ('menggelar', 'O'), ('diskon', 'O'), ('besar', 'O'), ('pada', 'O'), ('musim', 'O'), ('panas', 'O'), ('untuk', 'O'), ('berbagai', 'O'), ('judul', 'O'), ('game', 'O'), ('populer', 'O'), ('!', 'O')]
	Prediksi Pipeline	[('steam', 'B-PLATFORM', 0.9787), ('menggelar', 'O', 0.9978), ('diskon', 'O', 0.9976), ('besar', 'O', 0.9975), ('pada', 'O', 0.9973), ('musim', 'O', 0.9971), ('panas', 'O', 0.9969), ('untuk', 'O', 0.9967), ('berbagai', 'O', 0.9966), ('judul', 'O', 0.9965), ('game', 'O', 0.9963), ('populer', 'O', 0.9962), ('!', 'O', 0.9991)]

Pada kasus 1, entitas 'UCL' yang seharusnya berlabel B-EVENT (merujuk pada *event* UEFA Champions League) salah diklasifikasikan sebagai B-TECH atau PLATFORM. Hipotesis untuk kesalahan ini adalah model mengalami ambiguitas semantik yang dipengaruhi oleh konteks domain. Meski-

Tabel 5.7 Kesalahan Pada Pelabelan Data Uji

No	Teks (Ringkas)	Label Asli	Prediksi Pipeline	Catatan
1	UCL ... Elden Ring ...	UCL → EVENT, Elden Ring → GAME	UCL → TECH/PLATFORM(salah), Elden Ring → GAME (benar)	Salah pada UCL
2	... Final Fantasy XVI ... 22 Juni ...	Final Fantasy XVI → GAME, 22 Juni → TIME	Final Fantasy XVI → GAME, 22 Juni → TIME	Semua benar
3	... Reddit ... Tyson ... Tekken 8 ... AI ...	Reddit → PLATFORM, Tyson → PERSON, Tekken 8 → GAME, AI → TECH	Reddit → PLATFORM(benar), Tyson → salah pecah token, Tekken 8 → GAME(benar), AI → salah (PERSON)	Tyson & AI salah

No	Teks (Ringkas)	Label Asli	Prediksi Pipeline	Catatan
4	... Cage ... Kojima Productions ... Death Stranding 2 ...	Cage → PERSON , Kojima Productions → COMPANY , Death Stranding 2 → GAME	Cage → PERSON (benar), Kojima Productions → COMPANY (benar), Death Stranding 2 → GAME (benar)	Hampir semua benar
5	... beberapa dekade ... Tetris ... perangkat lunak ...	beberapa dekade → TIME , Tetris → GAME , perangkat lunak → TECH	TIME → benar, Tetris → GAME (benar), perangkat lunak → tidak terbaca	TECH hilang
6	... Mizobe ... Pocketpair ...	Mizobe → PERSON , Pocketpair → COMPANY	Mizobe → salah (COMPANY), Pocketpair → salah (GAME)	Kebalik
7	... Free Fire ... Venom Let There Be Carnage ... Venom ...	Free Fire → GAME , Venom LTB Carnage → GAME , Venom → GAME , skin → TECH	Free Fire → GAME , Venom LTB Carnage → GAME , Venom → GAME , skin → hilang	TECH hilang
8	... SpiderMan ... Peter ... Miles ...	SpiderMan → GAME , Peter → PERSON , Miles → PERSON	SpiderMan → GAME , Peter → PERSON , Miles → PERSON	Semua benar
9	... Ainrun ... FromSoftware ... dua tahun lalu ...	Ainrun → PERSON , FromSoftware → COMPANY , dua tahun lalu → TIME	Ainrun → salah (GAME), FromSoftware → COMPANY (benar), dua tahun lalu → TIME (benar)	Ainrun salah
10	... Overwatch 2 ... Blizzard ... Overwatch ...	Overwatch 2 → GAME , Blizzard → COMPANY , Overwatch → GAME	Semua sesuai	Benar

pun model mengenali 'UCL' sebagai *proper noun*, konteks sekitarnya yang didominasi oleh istilah teknologi dan *game* ("pic.twitter.com", "Game Elden Ring") kemungkinan besar memberikan bobot yang lebih tinggi pada kelas TECH atau PLATFORM yang lebih sering muncul dalam data pelatihan. Hal ini menunjukkan bahwa representasi model untuk kelas EVENT yang langka kurang diskriminatif dibandingkan kelas-kelas mayoritas. Contoh lain terlihat pada kasus 3, di mana 'AI' (seharusnya B-TECH) salah diklasifikasikan sebagai B-PERSON. Kesalahan ini menyoroti keterbatasan model dalam membedakan antara konsep teknis dan entitas lain ketika kata tersebut (seperti "AI" atau nama-nama

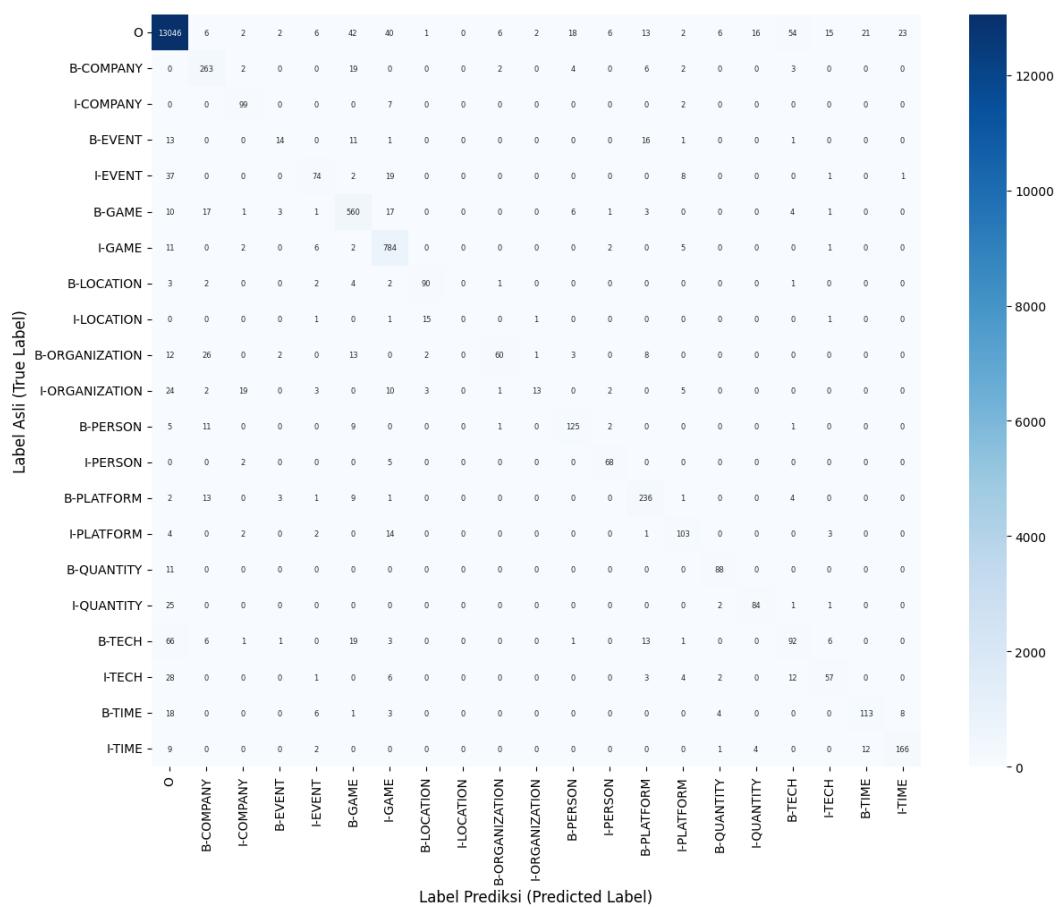
lain) dapat merujuk ke beberapa tipe entitas tergantung pada konteks. Kategori kesalahan signifikan lainnya berkaitan dengan ketidakmampuan model untuk secara konsisten menentukan batas awal dan akhir sebuah entitas.

Kesalahan klasifikasi yang muncul pada model IndoBERT merupakan artefak langsung dari mekanisme tokenisasi sub-kata (*WordPiece*). Tokenisasi ini memecah entitas nama diri seperti “*Tyson*” dan nama *game* seperti “*Tekken*” menjadi beberapa token (*ty*, *##son* dan *tek*, *##ken*). Model sering memberikan label benar pada masing-masing sub-token, Namun, pemecahan kata tersebut menimbulkan kesulitan saat penggabungan kembali pada tahap setelah pemrosesan. Akibatnya, entitas sering terdeteksi secara parsial atau tidak utuh. Fenomena ini menunjukkan adanya tantangan dalam penggunaan model berbasis sub-kata untuk tugas NER yang sensitif terhadap batas kata.

Model IndoBERT juga gagal mendeteksi entitas tertentu dan justru mengklasifikasikannya sebagai non-entitas (*O*). Fenomena ini berkorelasi langsung dengan nilai *recall* yang rendah pada beberapa kelas, sebagaimana dibahas dalam Sub-bab 5.3.3. Analisis kuantitatif mencatat bahwa *recall* untuk label B-TECH hanya mencapai 0,38. Nilai tersebut berarti model melewatkannya lebih dari 60% entitas teknologi yang sebenarnya muncul dalam data. Pola kesalahan model menunjukkan bahwa petunjuk konteks di sekitar beberapa entitas masih lemah, terutama pada entitas yang ambigu atau jarang muncul. Model kesulitan membedakan entitas tersebut dari kata umum dalam kalimat sehingga terjadi salah deteksi.

Analisis kualitatif menegaskan kapabilitas model dalam mengenali entitas yang terdefinisi jelas dan berfrekuensi tinggi. Pada kasus 2, entitas multi-kata

seperti “Final Fantasy XVI” berhasil diidentifikasi dengan tepat dari segi batas maupun kategorinya. Pada kasus 4, entitas “Microsoft” (B-COMPANY), “Xbox” (B-PLATFORM), dan “2028” (B-TIME) terdeteksi dengan skor probabilitas yang sangat tinggi. Analisis kualitatif secara keseluruhan menyimpulkan bahwa kesalahan model tidak terjadi secara acak, melainkan mengikuti pola yang dapat diidentifikasi. Tantangan utama terletak pada penanganan ambiguitas semantik, keterbatasan data pada kelas minoritas, serta artefak yang muncul akibat tokenisasi sub-kata.



Gambar 5.4 Confusion Matrix GameNER-ID (Test Set)

Matriks pada Gambar 5.4 memperlihatkan distribusi hasil prediksi model pada sumbu mendatar dan label asli pada sumbu tegak, sehingga tampak jelas

bagaimana model sering mencampuradukkan kategori entitas. Visualisasi ini menunjukkan dominasi diagonal utama.

Tabel 5.8 Rangkuman Kinerja Per Kelas

Kelas	Jumlah Asli (True)	Jumlah Prediksi	Prediksi Benar (TP)	Recall	Precision	Accuracy
O	13,327	13,324	13,046	0.9789	0.9791	0.9675
B-COMPANY	301	346	263	0.8738	0.7601	0.9930
I-COMPANY	108	130	99	0.9167	0.7615	0.9977
B-EVENT	57	25	14	0.2456	0.5600	0.9969
I-EVENT	142	105	74	0.5211	0.7048	0.9942
B-GAME	624	691	560	0.8974	0.8104	0.9901
I-GAME	813	913	784	0.9643	0.8587	0.9863
B-LOCATION	105	111	90	0.8571	0.8108	0.9942
I-LOCATION	19	0	0	0.0000	0.0000	0.9989
B-ORGANIZATION	127	71	60	0.4724	0.8451	0.9950
I-ORGANIZATION	82	17	13	0.1585	0.7647	0.9984
B-PERSON	154	157	125	0.8117	0.7962	0.9923
I-PERSON	75	81	68	0.9067	0.8395	0.9959
B-PLATFORM	270	299	236	0.8741	0.7893	0.9910
I-PLATFORM	129	134	103	0.7984	0.7687	0.9923
B-QUANTITY	99	103	88	0.8889	0.8544	0.9942
I-QUANTITY	113	104	84	0.7434	0.8077	0.9920
B-TECH	209	173	92	0.4402	0.5318	0.9881
I-TECH	113	86	57	0.5044	0.6628	0.9934
B-TIME	153	146	113	0.7386	0.7740	0.9930
I-TIME	194	198	166	0.8557	0.8384	0.9907

Tabel 5.8 menyajikan data numerik yang mendasari *confusion matrix*, merangkum jumlah prediksi benar (*True Positives* - TP) dan menghubungkannya dengan metrik *Recall* dan *Precision* untuk setiap kelas. Analisis menunjukkan adanya ketidakseimbangan antara *recall* dan *precision* pada beberapa kelas. Nilai tertinggi terkonsentrasi pada diagonal dari kiri atas ke kanan bawah, yang berarti mayoritas prediksi model benar. Konsentrasi kuat pada sel seperti (O, O), (B-GAME, B-GAME), dan (I-GAME, I-GAME) menegaskan bahwa model bekerja baik pada kategori mayoritas. Analisis juga menemukan kebocoran prediksi ke label ‘O’. Banyak entitas yang seharusnya terdeteksi justru diprediksi sebagai bukan entitas. Misalnya, baris B-TECH menampilkan 66 contoh yang jatuh ke

kolom ‘O’. Pola ini memperkuat temuan sebelumnya bahwa model memiliki daya tangkap rendah pada entitas teknologi.

Kasus serupa meskipun lebih kecil juga tampak pada baris B-ORGANIZATION dan I-ORGANIZATION. Selain itu, kebingungan antar-kategori juga muncul meskipun tidak sebesar kebocoran ke ‘O’. Pada baris B-COMPANY, misalnya, terdapat 19 prediksi yang salah masuk ke kolom B-ORGANIZATION. Hal ini mendukung dugaan sebelumnya bahwa ada tumpang tindih makna antara COMPANY dan ORGANIZATION. Kelas B-ORGANIZATION misalnya, memiliki *precision* yang sangat tinggi (0,8451) tetapi *recall* yang rendah (0,4724). Pola ini berarti prediksi model hampir selalu benar ketika ia mengenali entitas sebagai B-ORGANIZATION, tetapi model bersifat terlalu hati-hati sehingga melewatkannya lebih dari separuh entitas yang sebenarnya ada. Perilaku ini menggambarkan bahwa model hanya mampu mendeteksi entitas ORGANIZATION yang paling jelas dan tidak ambigu.

Data juga menegaskan kelemahan model secara kuantitatif. Nilai *recall* untuk B-EVENT (0,2456) dan B-TECH (0,4402) sangat rendah, yang memperlihatkan besarnya jumlah entitas yang tidak terdeteksi. Angka ini langsung menguatkan pola salah deteksi (false negatives) yang sebelumnya terlihat pada *confusion matrix*. Di sisi lain, model menunjukkan kekuatan pada entitas multi-kata. Performa I-GAME dengan *recall* 0,9643 membuktikan bahwa setelah model berhasil mengenali awal nama *game* (B-GAME), ia sangat efektif menangkap token berikutnya. Kemampuan ini penting untuk mengekstrak nama *game* yang panjang secara lengkap dan akurat.

Nilai akurasi per kelas pada Tabel 5.8 tampak sangat tinggi di hampir seluruh label, termasuk kelas dengan kinerja lemah seperti B-TECH dan I-TECH. Hal ini bukan menunjukkan bahwa model benar-benar unggul dalam mengenali setiap entitas, tetapi lebih disebabkan oleh besarnya jumlah token non-entitas (kelas O) yang dihitung sebagai *true negative* (TN) dalam perhitungan akurasi. Karena proporsi token entitas relatif kecil dibanding total token, TN menjadi sangat dominan dan otomatis mendorong akurasi mendekati 1 meskipun *true positive* rendah atau *false negative* tinggi pada label tertentu. Dengan demikian, akurasi tidak dapat dijadikan indikator utama untuk menilai kualitas model NER. Evaluasi yang lebih bermakna tercermin pada nilai *Precision*, *Recall*, dan *F1-score*, yang lebih sensitif terhadap performa model dalam mengenali entitas yang sebenarnya penting namun berjumlah lebih sedikit.

Analisis mendalam terhadap model IndoBERT telah memberikan gambaran yang jelas mengenai kinerja *baseline* serta pola-pola kesalahan yang sistematis, terutama dalam menangani istilah asing dan entitas minoritas. Temuan ini menggarisbawahi keterbatasan model monolingual dan mendorong perlunya eksplorasi arsitektur yang lebih kuat. Untuk itu, bab selanjutnya akan memperluas eksperimen dengan melakukan evaluasi komparatif terhadap berbagai model berbasis Transformer, mencakup varian monolingual dan multilingual dengan ukuran yang berbeda, guna mengidentifikasi arsitektur yang paling efektif dalam mengatasi kompleksitas linguistik pada dataset GameNER-ID.

BAB VI

MODEL NER BERBASIS TRANSFORMER

Bab ini menyajikan alur eksperimen secara terstruktur. Perancangan dimulai dengan pemilihan arsitektur model, penetapan konfigurasi hyperparameter, serta penentuan metrik evaluasi. Tahap berikutnya menampilkan hasil kinerja dari setiap model untuk mengidentifikasi arsitektur yang paling efektif. Analisis kemudian menelusuri hubungan antara model, seperti ukuran dan cakupan bahasa (monolingual vs. multilingual), dengan performa yang dihasilkan. Bagian akhir membahas analisis kesalahan (*error analysis*) pada model dengan hasil terbaik dan terburuk untuk memetakan kekuatan dan kelemahan masing-masing dalam mengenali entitas pada domain *game*.

6.1 Model Transformer

Pemilihan model Transformer bertujuan mengevaluasi kinerja secara menyeluruh melalui pengujian arsitektur dengan karakteristik yang berbeda. Model berbasis Transformer yang telah melalui tahap pra-pelatihan (*pre-trained*) dipilih karena popularitas dan performa unggulnya pada berbagai tugas NLP bahasa Indonesia. Dua kelompok utama model yang digunakan dalam eksperimen ditunjukkan pada Tabel 6.1.

Tabel 6.1 Kelompok Model yang Digunakan dalam Eksperimen

Kelompok Model	Varian yang Diuji
Monolingual (Indonesia)	indobert-uncased, indobert-base-p1, indobert-base-p2, indobert-large-p1, indobert-large-p2
Multilingual	bert-multilingual-cased, xlm-roberta-base, xlm-roberta-large

Pemilihan ragam model pada Tabel 6.1 mengikuti strategi evaluasi komprehensif yang mengacu pada tolak ukur standar NLP bahasa Indonesia seperti

IndoNLU dan IndoLEM. Model monolingual, khususnya varian IndoBERT, berfungsi sebagai baseline karena dirancang dan dilatih secara khusus menggunakan korpus bahasa Indonesia. Penggunaan model ini memungkinkan pengukuran kemampuan arsitektur yang dioptimalkan untuk bahasa Indonesia dalam mengenali terminologi khas pada domain *game*.

Domain *game* menampilkan banyak kata serapan (*loanwords*) dan fenomena alih kode (*code-switching*) dari bahasa Inggris, misalnya pada entitas *PlayStation*, *Square Enix*, atau *Unity Engine*. Kondisi tersebut berpotensi menurunkan akurasi model monolingual karena sebagian istilah dianggap sebagai token *out-of-vocabulary* (OOV). Sebagai pembanding, model multilingual yang telah melalui pra-pelatihan menggunakan korpus besar dari berbagai bahasa, termasuk bahasa Inggris, memiliki kemampuan lebih baik dalam memahami serta mengidentifikasi entitas lintas-bahasa.

Perbandingan antara kedua kelompok model memberikan dasar analisis yang lebih dalam untuk menentukan arsitektur paling efektif dalam menangani kompleksitas kata pada domain *game*. Hubungan antara performa model monolingual dan multilingual memperlihatkan sejauh mana *domain shift* memengaruhi kemampuan model dalam mengenali entitas pada teks berbahasa Indonesia yang bercampur istilah asing.

6.2 Hasil Eksperimen dan Analisis Kinerja

Evaluasi model dilakukan secara sistematis untuk menyoroti temuan utama dari hasil eksperimen. Analisis mencakup tiga tahap yang saling berkaitan. Tahap pertama menyajikan hasil komparatif untuk seluruh arsitektur yang diuji guna menilai performa kuantitatif setiap model. Tahap berikutnya berfokus pada

identifikasi tren serta faktor penentu kinerja, termasuk pengaruh ukuran model dan cakupan bahasa (monolingual vs. multilingual). Tahap akhir menempatkan hasil eksperimen dalam konteks pengembangan sumber daya NER berbahasa Indonesia agar kontribusi dan daya saing dataset dapat terlihat secara jelas.

6.2.1 Hasil Komparatif Kinerja Model

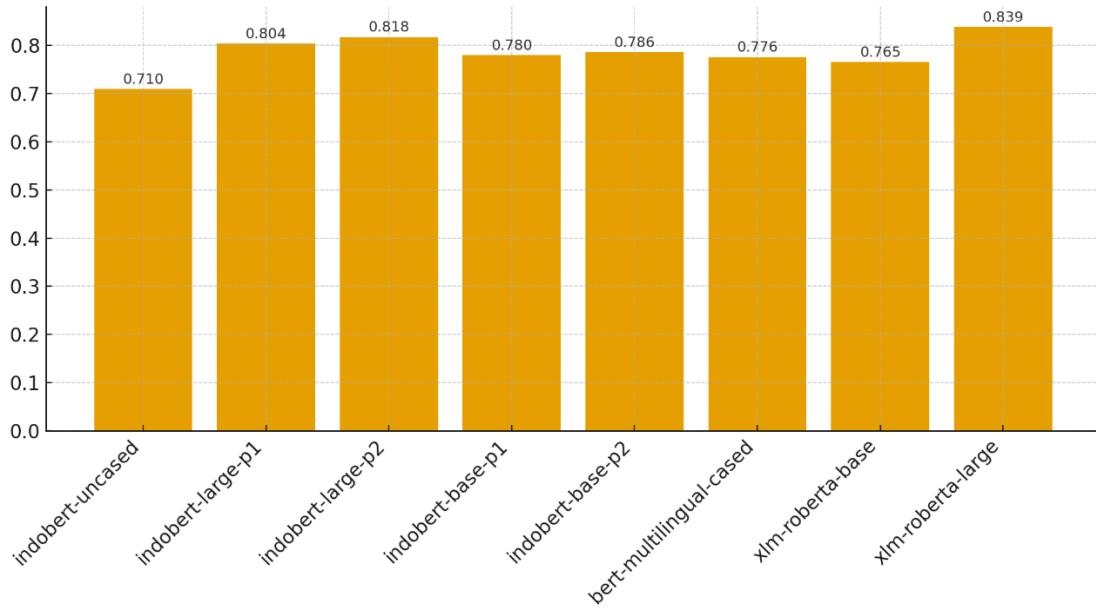
Seluruh skenario eksperimen menghasilkan data kuantitatif yang menunjukkan performa model setelah proses *fine-tuning* pada dataset uji GameNER-ID. Tabel 6.2 menampilkan perbandingan nilai *F1-Score*, *Precision*, *Recall*, dan *Loss* dari masing-masing arsitektur, yang berfungsi sebagai dasar analisis kinerja berikutnya.

Tabel 6.2 Hasil Perbandingan Kinerja Model pada Dataset Uji GameNER-ID

Model	F1	Prec.	Rec.	Acc.	Loss
xlm-roberta-large	0.8386	0.8276	0.8500	0.9640	0.1325
indobert-large-p2	0.8175	0.8089	0.8262	0.9569	0.1462
indobert-large-p1	0.8036	0.7905	0.8172	0.9577	0.1487
indobert-base-p2	0.7860	0.7711	0.8015	0.9508	0.1688
indobert-base-p1	0.7796	0.7600	0.8001	0.9489	0.1733
bert-multilingual-cased	0.7756	0.7616	0.7901	0.9518	0.1627
xlm-roberta-base	0.7651	0.7415	0.7901	0.9471	0.1721
indobert-uncased	0.7097	0.6873	0.7336	0.9373	0.2117

Seluruh model mencapai *F1-Score* di atas 0.70, dengan arsitektur yang lebih kuat secara konsisten mencatat skor antara 0.76 hingga 0.84. Hasil tersebut menunjukkan bahwa anotasi dalam GameNER-ID memiliki konsistensi tinggi dan dapat dipelajari secara efektif oleh model Transformer. Model multilingual XLM-RoBERTa-large mencatat performa tertinggi dengan *F1-Score* 0.8386, *Precision* 0.8276, dan *Recall* 0.8500. Kombinasi nilai tersebut menandakan kemampuan model dalam menjaga keseimbangan antara akurasi prediksi dan cakupan entitas yang dikenali. Sebagai pembanding, IndoBERT-uncased yang berfungsi sebagai

baseline dasar menghasilkan *F1-Score* 0.7097, sehingga menetapkan batas bawah kinerja seluruh eksperimen (Gambar 6.1).



Gambar 6.1 Visualisasi Grafik Perbandingan *F1-Score* Untuk Setiap Model

Analisis terhadap pola kinerja memperlihatkan pengelompokan yang jelas berdasarkan ukuran arsitektur. Model berukuran besar seperti XLM-RoBERTa-large dan varian IndoBERT-large (p1 dan p2) membentuk kelompok unggul dengan *F1-Score* di atas 0.80. Model berukuran dasar, termasuk IndoBERT-base (p1 dan p2), BERT-multilingual-*cased*, dan XLM-RoBERTa-base, menampilkan performa yang stabil di kisaran 0.76–0.79. Keseragaman skor pada kelompok ini menunjukkan bahwa GameNER-ID tidak bias terhadap arsitektur tertentu, sehingga dapat berfungsi sebagai tolok ukur yang adil untuk evaluasi lintas-model. Kinerja tersebut menegaskan posisi XLM-RoBERTa-large sebagai arsitektur paling efektif dalam mengenali entitas pada domain *game* sekaligus memvalidasi kualitas GameNER-ID sebagai *benchmark* baru untuk NER berbahasa Indonesia.

6.2.2 Analisis Perbandingan Hasil

Perbandingan performa antar-model pada dataset GameNER-ID mengungkap hubungan yang kuat antara ukuran arsitektur, cakupan bahasa, dan kemampuan model dalam mengenali entitas. Tabel 6.3 merangkum hasil perbandingan berdasarkan jumlah parameter dan tipe bahasa.

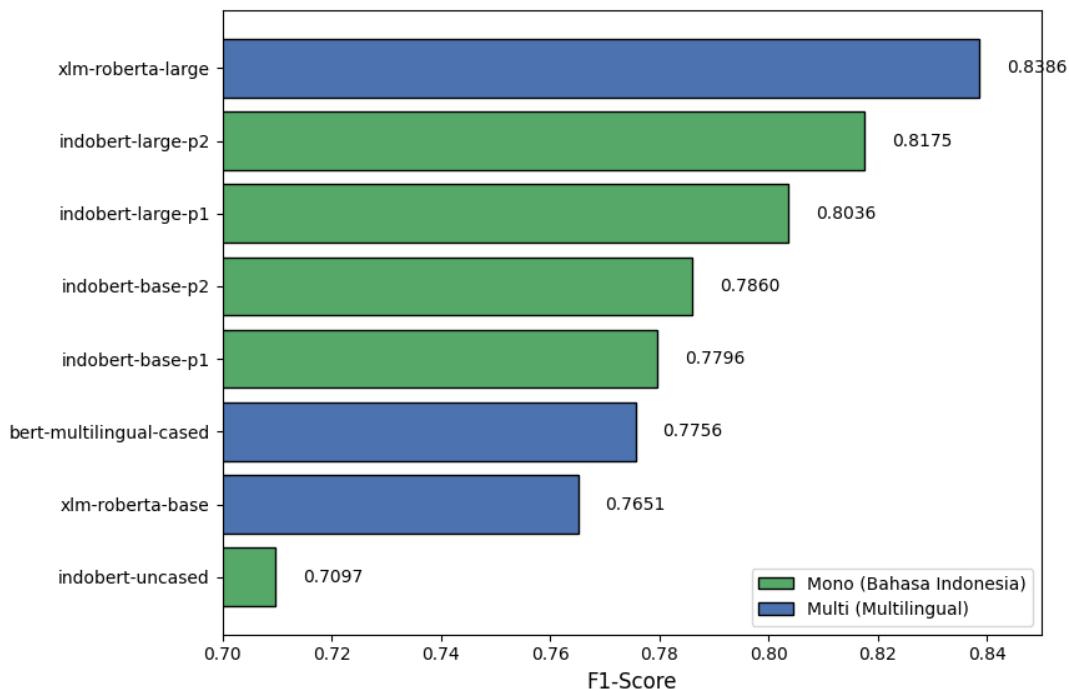
Tabel 6.3 Hasil Perbandingan Ukuran Model dan Kinerja

Model	Ukuran Model (Parameters)	Tipe Bahasa	F1-Score
xlm-roberta-large	561.0M	Multi	0.8386
indobert-large-p2	335.2M	Mono	0.8175
indobert-large-p1	335.2M	Mono	0.8036
indobert-base-p2	124.5M	Mono	0.7860
indobert-base-p1	124.5M	Mono	0.7796
bert-multilingual-cased	167.4M	Multi	0.7756
xlm-roberta-base	278.7M	Multi	0.7651
indobert-uncased	124.5M	Mono	0.7097

Ukuran model berbanding lurus dengan performa yang dicapai. Arsitektur berukuran besar seperti IndoBERT-large dan XLM-RoBERTa-large menghasilkan *F1-Score* antara 0.80 hingga 0.84, sedangkan model berukuran dasar seperti IndoBERT-base dan XLM-RoBERTa-base mencatat skor 0.76–0.79. Peningkatan kapasitas parameter memungkinkan pemahaman yang lebih mendalam terhadap variasi bahasa dan terminologi spesifik dalam teks *game*. Model berukuran besar lebih mampu mengenali entitas dengan konteks kompleks, sedangkan model dasar tetap menunjukkan stabilitas pada tingkat performa menengah.

Keunggulan model multilingual terlihat jelas pada hasil tertinggi yang dicapai XLM-RoBERTa-large dengan *F1-Score* 0.8386, melampaui IndoBERT-large-p2 sebagai model monolingual terbaik. Perbedaan tersebut berkaitan dengan karakteristik dataset GameNER-ID yang banyak memuat istilah serapan dari bahasa Inggris seperti *PlayStation*, *Square Enix*, dan *Unreal Engine*. Korpus pelatihan

multibahasa yang digunakan pada XLM-RoBERTa memungkinkan model memahami token lintas bahasa secara lebih efektif. Sebaliknya, model monolingual seperti IndoBERT cenderung memperlakukan istilah asing sebagai token out-of-vocabulary (OOV), yang menurunkan akurasi klasifikasi entitas.



Gambar 6.2 Perbandingan Skor F1 antar Model NER Bahasa Indonesia

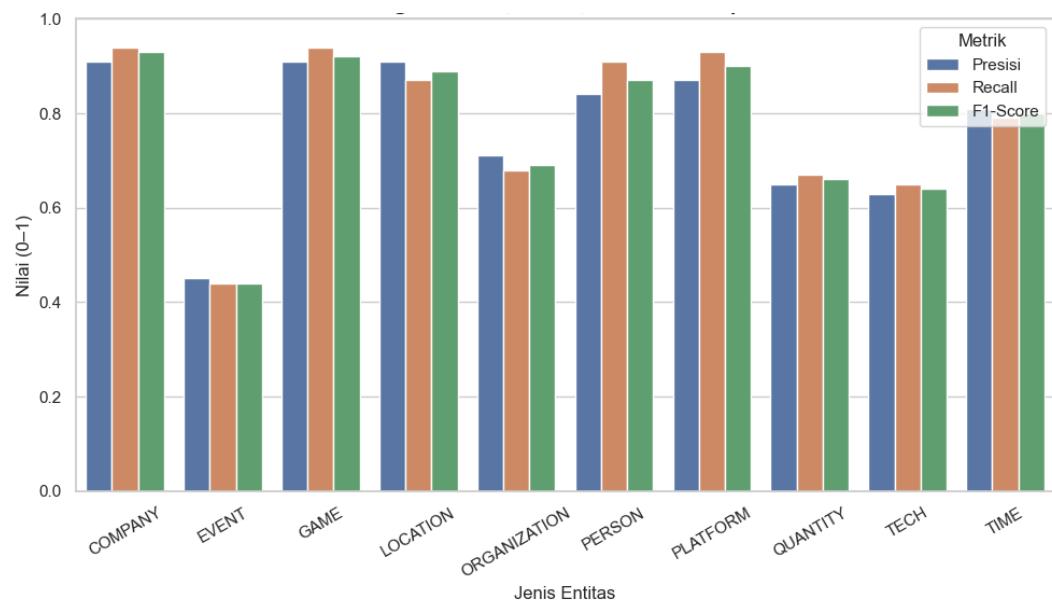
Perbandingan keseluruhan menunjukkan bahwa arsitektur besar dan multilingual memberikan hasil terbaik untuk pengenalan entitas pada domain *game*. Kapasitas model yang lebih tinggi meningkatkan kemampuan dalam menangkap nuansa bahasa dan istilah teknis, sementara cakupan multibahasa memperluas pemahaman terhadap entitas yang berakar dari bahasa asing. Kombinasi kedua faktor tersebut menjadikan model multilingual besar seperti XLM-RoBERTa-large pilihan paling efektif untuk menangani *domain shift* pada dataset GameNER-ID.

6.3 Analisis Kesalahan Model pada Tugas NER Domain *Game*

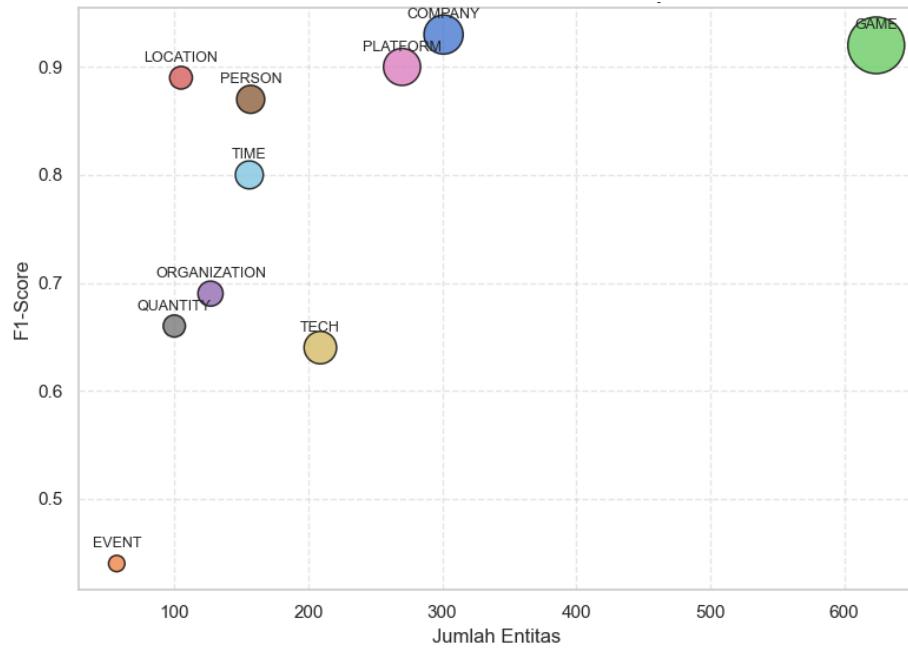
Analisis kesalahan dilakukan untuk memahami tantangan spesifik dalam tugas NER pada domain *game*. Evaluasi difokuskan pada model dengan performa tertinggi, yaitu XLM-RoBERTa-large, guna mengidentifikasi batas kemampuannya. Perbandingan confusion matrix antara XLM-RoBERTa-large, IndoBERT-large-p2, IndoBERT-uncased, dan XLM-RoBERTa-base membantu menyoroti pola kelemahan fundamental antararsitektur. Tujuan utama analisis ini adalah menemukan pola kesalahan sistematis yang dapat menjadi dasar peningkatan model.

6.3.1 Entitas Berkinerja Tinggi

Entitas COMPANY, GAME, dan PLATFORM menunjukkan performa tertinggi yang ditunjukkan pada gambar 6.3 dan 6.4.



Gambar 6.3 Presisi, Recall dan F1-Score per Entitas



Gambar 6.4 Jumlah Data dengan *F1-Score* per Entitas

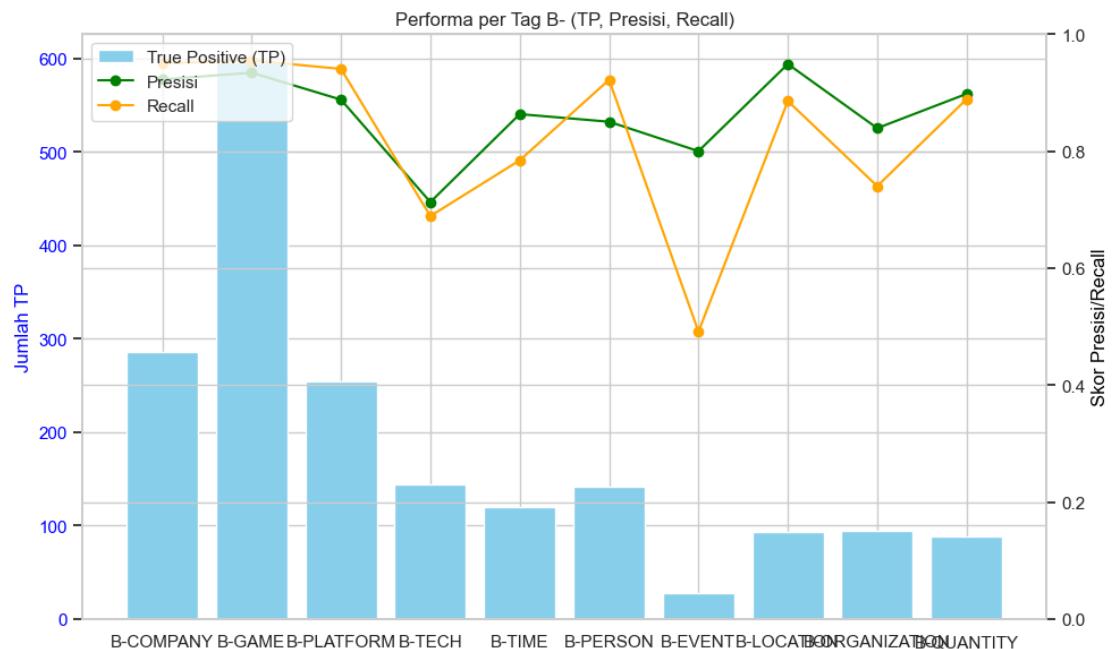
Hasil per tag pada Tabel 6.4 memperkuat temuan entitas yang berkinerja tinggi. Baris B-COMPANY menampilkan presisi dan *Recall* tinggi dengan 286 TP dari 301 contoh. Baris B-GAME menunjukkan kesalahan relatif kecil dengan 596 TP dan hanya 28 FN. Baris B-PLATFORM memiliki 254 TP dengan kesalahan minor pada token yang mirip nama platform. Pola ini membuktikan bahwa kesalahan pada entitas berkinerja tinggi bersifat terlokalisir dan umumnya disebabkan variasi penulisan, bukan kegagalan deteksi mendasar.

Tabel 6.4 Metrik Performa Per Kelas

Tag Entitas	Jumlah	Diprediksi	TP	Presisi	Recall
O	13327	13356	13146	0.9843	0.9864
B-COMPANY	301	310	286	0.9226	0.9502
I-COMPANY	108	113	102	0.9027	0.9444
B-EVENT	57	35	28	0.8000	0.4912
I-EVENT	142	133	119	0.8947	0.8380
B-GAME	624	638	596	0.9342	0.9551
I-GAME	813	833	797	0.9568	0.9803
B-LOCATION	105	98	93	0.9490	0.8857
I-LOCATION	19	22	16	0.7273	0.8421
B-ORGANIZATION	127	112	94	0.8393	0.7402
I-ORGANIZATION	82	67	51	0.7612	0.6220

Tag Entitas	Jumlah	Diprediksi	TP	Presisi	Recall
B-PERSON	154	167	142	0.8503	0.9221
I-PERSON	75	80	74	0.9250	0.9867
B-PLATFORM	270	286	254	0.8881	0.9407
I-PLATFORM	129	132	116	0.8788	0.8992
B-QUANTITY	99	98	88	0.8980	0.8889
I-QUANTITY	113	107	89	0.8318	0.7876
B-TECH	209	202	144	0.7129	0.6890
I-TECH	113	103	73	0.7087	0.6460
B-TIME	153	139	120	0.8633	0.7843
I-TIME	194	183	167	0.9126	0.8608

Grafik pada gambar 6.5 menunjukkan bahwa entitas B-GAME, B-COMPANY, dan B-PLATFORM memiliki performa paling konsisten dengan jumlah *True Positive (TP)* tinggi serta nilai Presisi dan *Recall* di atas 0.9. Hal ini menandakan bahwa model mampu mengenali entitas-entitas tersebut dengan sangat baik dan stabil.



Gambar 6.5 Performa per Tag (*True Positive, Presisi Recall*)

6.3.2 Entitas Bermasalah

Pola kesalahan utama meliputi dua kategori. Pertama, kegagalan mendeteksi token awal (B-) menyebabkan false negative tinggi pada tag TECH dan

EVENT. Faktor penyebabnya termasuk variasi penamaan besar, kemunculan entitas langka, dan tokenisasi subword yang memecah istilah teknis. Kedua, kebingungan semantik muncul antara kategori yang beririsan, seperti ORGANIZATION dan COMPANY, akibat konteks linguistik yang tidak cukup membedakan asosiasi komersial dari non-komersial (Tabel 6.5).

Tabel 6.5 Metrik Performa Per Kelas (Problematik)

Tag Entitas	Jumlah Sebenarnya	TP	FP	FN	Presisi	Recall
B-EVENT	57	28	7	29	0.8000	0.4912
I-EVENT	142	119	14	23	0.8947	0.8380
B-ORGANIZATION	127	94	18	33	0.8393	0.7402
I-ORGANIZATION	82	51	16	31	0.7612	0.6220
B-TECH	209	144	58	65	0.7129	0.6890
I-TECH	113	73	30	40	0.7087	0.6460
B-QUANTITY	99	88	10	11	0.8980	0.8889
I-QUANTITY	113	89	18	24	0.8318	0.7876

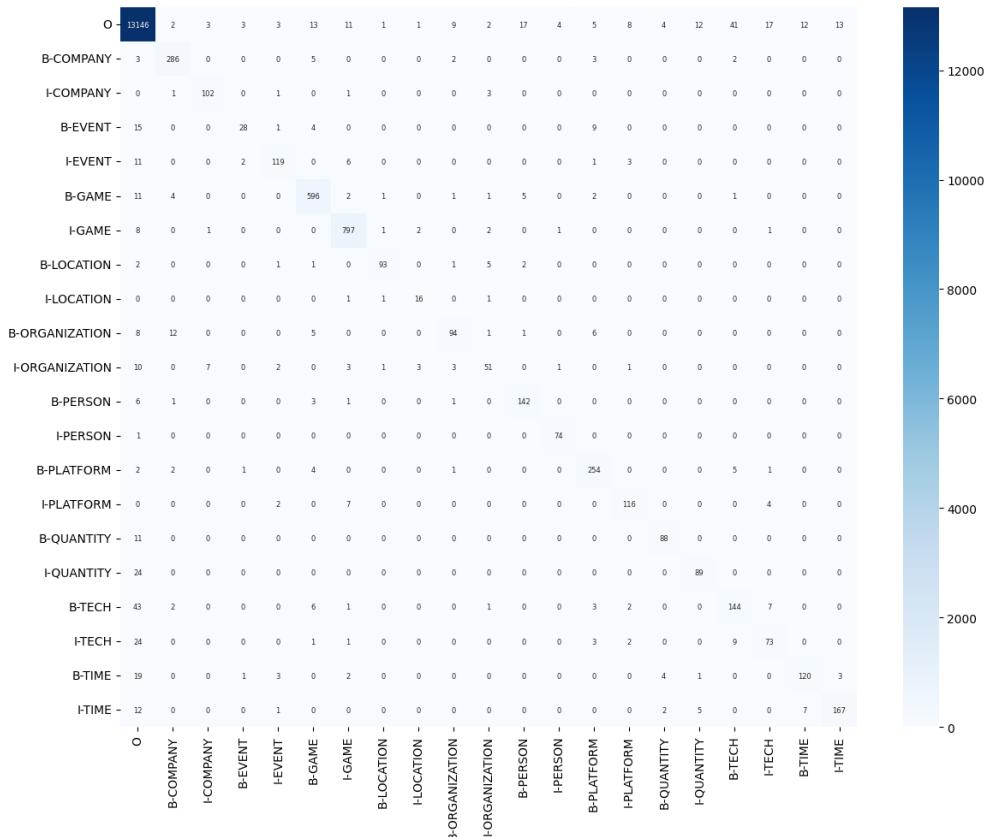
Tabel 6.6 memperlihatkan bahwa sebagian besar kesalahan TECH dan EVENT berakhir sebagai label O, sementara kesalahan ORGANIZATION sering bergeser menjadi COMPANY. Rekomendasi perbaikan meliputi penambahan data anotasi untuk entitas langka, augmentasi variasi penamaan, penggunaan leksikon domain, serta penyesuaian tokenisasi agar istilah teknis tetap utuh.

Tabel 6.6 Kesalahan Prediksi Paling Banyak

Pasangan Kebingungan	Jumlah Kasus
B-TECH → O	43
B-EVENT → O	15
B-ORGANIZATION → B-COMPANY	12
B-GAME <i>false negative</i> total	28
B-GAME <i>false positive</i> total	42

6.3.3 Pola Kesalahan pada Model Berkinerja Tinggi

Analisis mendalam pada model berkinerja tinggi didasarkan pada matriks kebingungan yang ditunjukkan pada Gambar 6.3. Model XLM-RoBERTa-large memperlihatkan akurasi sangat tinggi dengan dominasi nilai diagonal. Meskipun demikian, sejumlah kesalahan muncul pada kasus-kasus yang bersifat semantik.



Gambar 6.6 Confusion Matrix XLM-RoBERTa-large

Kesalahan pertama terjadi pada kebingungan antara ORGANIZATION dan COMPANY. Matriks pada Gambar 6.3 menunjukkan terdapat 12 kasus B-ORGANIZATION yang diprediksi sebagai B-COMPANY. Dalam kalimat “*IESPA selaku organisasi esports mendukung penuh acara tersebut*”, entitas IESPA yang seharusnya dikenali sebagai ORGANIZATION justru diprediksi sebagai COMPANY. Kesalahan ini membuktikan bahwa model mampu mengenali entitas sebagai lembaga, tetapi tidak selalu berhasil membedakan antara organisasi asosiasi dan perusahaan komersial.

Tabel 6.7 Kesalahan pada Model Berkinerja Tinggi

Jenis Kesalahan	Jumlah Kasus	Contoh Kalimat	Prediksi Salah	Seharusnya
ORGANIZATION diprediksi COMPANY	12	IESPA selaku organisasi esports mendukung penuh acara tersebut	IESPA → COMPANY	IESPA → ORGANIZATION

Jenis Kesalahan	Jumlah Kasus	Contoh Kalimat	Prediksi Salah	Seharusnya
TECH tidak terdeteksi (O)	43	<i>Game</i> ini dibuat menggunakan Unreal Engine 5	Unreal Engine 5 → O	Unreal Engine 5 → TECH
EVENT tidak terdeteksi (O)	15	Banyak developer lokal hadir di IGDX 2023	IGDX 2023 → O	IGDX 2023 → EVENT
<i>Loanwords</i> berbahasa Inggris (O)	–	<i>Game</i> baru dari Square Enix sangat ditunggu	Square Enix → O	Square Enix → COMPANY

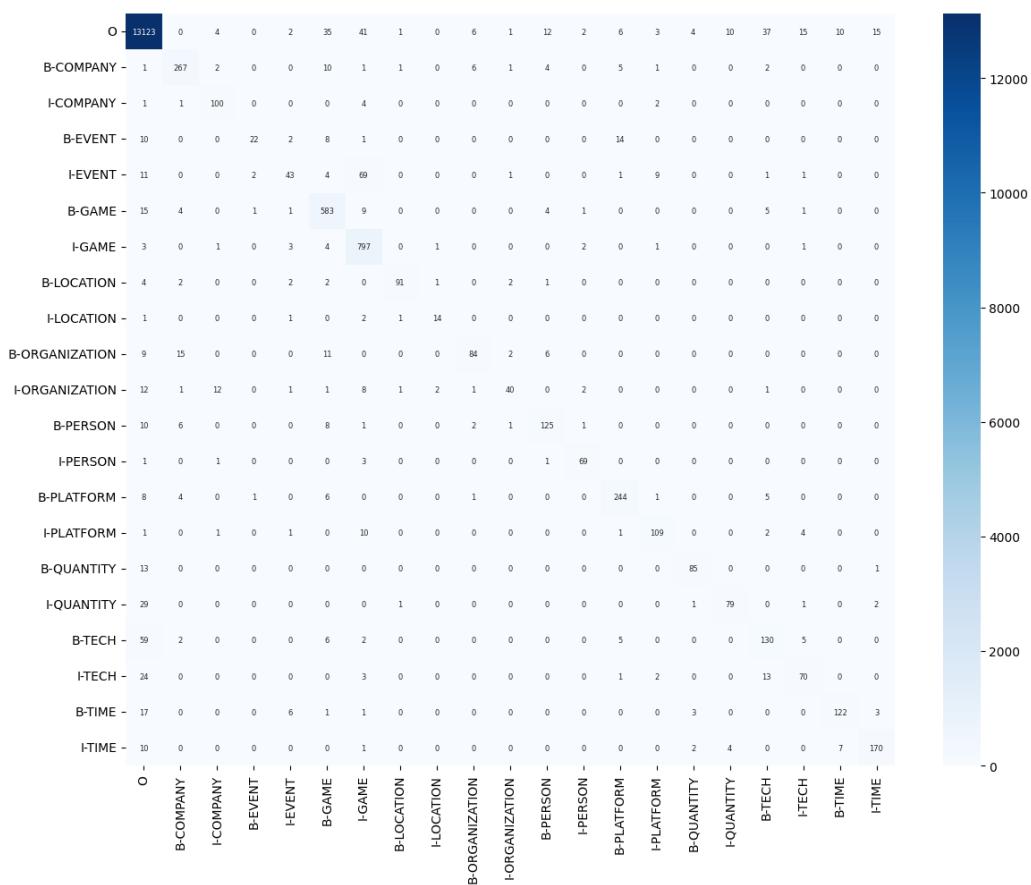
Kesalahan kedua muncul pada deteksi entitas yang bersifat ambigu. Matriks pada Gambar 1 memperlihatkan bahwa terdapat 43 entitas B-TECH yang salah diprediksi sebagai O dan 15 entitas B-EVENT yang juga tidak terdeteksi. Dalam kalimat “*Game ini dibuat menggunakan Unreal Engine 5*”, model gagal mengenali Unreal Engine 5 sebagai TECH dan mengklasifikasikannya sebagai O. Pada contoh lain, “*Banyak developer lokal hadir di IGDX 2023*”, model melewatkannya IGDX 2023 sebagai EVENT dan juga mengklasifikasikannya sebagai O (Tabel 6.7). Visualisasi kesalahan tersebut ditunjukkan pada gambar 6.7.



Gambar 6.7 Confusion Matrix IndoBERT-large-p2

Model IndoBERT-large-p2 pada Gambar 6.5 menunjukkan pola kesalahan yang mirip dengan XLM-RoBERTa-large, tetapi kelemahannya lebih jelas pada

kasus *loanwords*. Kinerja F1 yang lebih rendah memperlihatkan indikasi bahwa model ini lebih sering gagal mengenali entitas berbahasa Inggris. Dalam kalimat “*Game baru dari Square Enix sangat ditunggu*”, model ini berisiko lebih tinggi untuk salah mengklasifikasikan Square Enix sebagai O dibandingkan XLM-RoBERTa-large. Hal ini terjadi karena keterbatasan kosakata monolingual yang dimiliki IndoBERT. Untuk memperjelas perbedaan kasus kesalahan, Tabel 6.8 merangkum contoh-contohnya yang ditemukan pada model berkinerja tinggi.



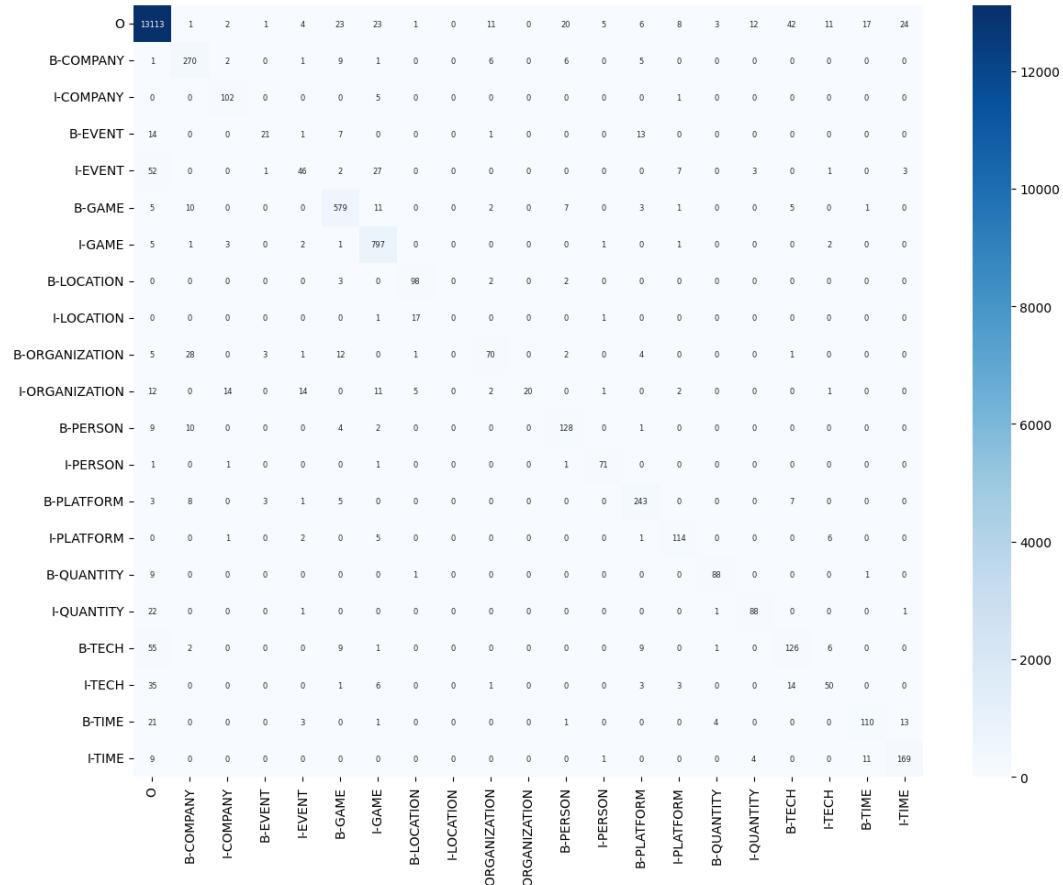
Gambar 6.8 Confusion Matrix IndoBERT-large-p2

6.3.4 Pola Kesalahan pada Model Berkinerja Rendah

Model berkinerja rendah tidak hanya menghasilkan kesalahan klasifikasi yang halus, tetapi juga menunjukkan kelemahan yang lebih fundamental. Matriks kebingungan memperlihatkan pola kesalahan yang tersebar luas, yang

mencerminkan kegagalan dalam memahami konteks serta fitur linguistik dasar yang penting bagi tugas *Named Entity Recognition* (NER).

Model XLM-RoBERTa-base dengan kapasitas yang lebih kecil, mengalami kesulitan dalam menggeneralisasi pola dari data latih (Gambar 6.9). Kesalahan yang dihasilkan tidak hanya merupakan versi lebih buruk dari model large,-



Gambar 6.9 Confusion Matrix XLM-RoBERTa-base

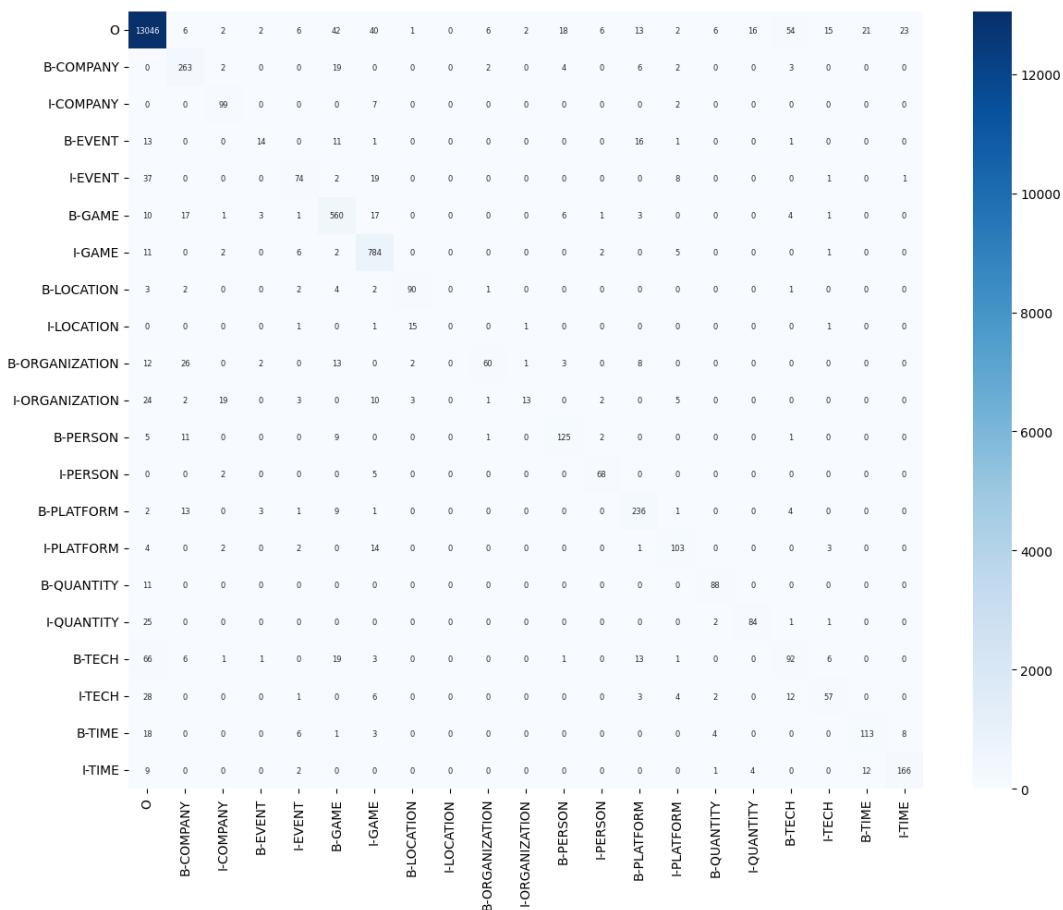
tetapi juga memperlihatkan bentuk kesalahan baru yang lebih mendasar. Jumlah kesalahan antara ORGANIZATION dan COMPANY meningkat dari 12 pada model large menjadi sekitar 20 hingga 25 pada model base. Selain itu, muncul kebingungan tambahan pada kategori yang seharusnya lebih mudah dipisahkan, seperti GAME dan TECH yang menimbulkan 10 hingga 15 kesalahan. Dalam kalimat “*Turnamen Free Fire diadakan di Jakarta*”, entitas Free Fire yang

seharusnya diklasifikasikan sebagai GAME justru diprediksi sebagai EVENT. Pada kalimat “*Unity menjadi mesin game pilihan developer indie*”, entitas Unity yang berfungsi sebagai TECH sering kali salah dikenali sebagai COMPANY karena model lebih mengaitkannya dengan entitas komersial (Tabel 6.10).

Tabel 6.8 Kesalahan pada Model Berkinerja Rendah (XLM-RoBERTa-base)

Jenis Kesalahan	Jumlah Kasus	Contoh Kalimat	Prediksi Salah	Seharusnya
ORGANIZATION vs COMPANY	20–25	IESPA selaku organisasi esports mendukung penuh acara	IESPA → COMPANY	IESPA → ORGANIZATION
GAME vs EVENT	10–15	Turnamen Free Fire diadakan di Jakarta	Free Fire → EVENT	Free Fire → GAME
TECH vs COMPANY	10–15	Unity menjadi mesin game pilihan developer indie	Unity → COMPANY	Unity → TECH
Kegagalan deteksi B-TECH	60–70	Game ini dibuat dengan CryEngine	CryEngine → O	CryEngine → TECH
Kegagalan deteksi B-EVENT	25–30	Banyak developer hadir di BEKRAF Game Prime	BEKRAF Game Prime → O	BEKRAF Game Prime → EVENT

Kesalahan model base juga terlihat pada tingkat kegagalan deteksi yang lebih tinggi dibandingkan model large. Jumlah entitas B-TECH yang tidak terdeteksi meningkat dari 43 pada model large menjadi 60 hingga 70 pada model base. Hal serupa terjadi pada entitas B-EVENT yang meningkat dari 15 menjadi sekitar 25 hingga 30 kasus. Model ini cenderung kurang percaya diri dalam mengidentifikasi entitas yang jarang muncul atau memiliki nama yang kompleks. Istilah teknologi seperti *CryEngine* atau *Godot Engine* hampir selalu dilabeli sebagai O. Nama acara seperti *BEKRAF Game Prime* atau *COMPFEST* juga sering kali tidak dikenali sebagai EVENT (Tabel 6.8).



Gambar 6.10 Confusion Matrix IndoBERT-uncased

Model IndoBERT-uncased menjadi model dengan performa terburuk karena hilangnya informasi kapitalisasi. Kehilangan fitur ini sangat merugikan NER, sebab huruf kapital berfungsi sebagai penanda utama nama diri. Lebih dari separuh entitas nama diri dalam data uji gagal terdeteksi dan diklasifikasikan sebagai O. Dalam kalimat “*Budi dari Agate merilis game baru*” yang diproses sebagai “*budi dari agate merilis game baru*”, model gagal mengenali Budi sebagai PERSON dan Agate sebagai COMPANY. Kesalahan serupa muncul dalam kalimat “*Komunitas Valorant di Jakarta sangat besar*” yang diproses sebagai “*komunitas valorant di jakarta sangat besar*”, sehingga Valorant tidak dikenali sebagai GAME dan Jakarta tidak dikenali sebagai LOCATION.

Model *uncased* juga memperlihatkan ambiguitas fatal pada nama entitas yang berupa kata benda umum. Nama *game* seperti *Journey* selalu diproses sebagai kata *journey* (perjalanan) dan dilabeli sebagai O (Tabel 6.9). Hal yang sama terjadi pada *game Control* yang diproses sebagai kata *control* (kendali). Nama perusahaan seperti *Unity* juga diperlakukan sebagai kata *unity* (persatuan) dan tidak dikenali sebagai entitas.

Tabel 6.9 Kesalahan pada Model Berkinerja Rendah (IndoBERT-uncased)

Jenis Kesalahan	Jumlah Kasus (Perkiraan)	Contoh Kalimat	Prediksi Salah	Seharusnya
Kehilangan kapitalisasi	> 50% entitas gagal	budi dari agate merilis <i>game</i> baru	budi → O, agate → O	budi → PERSON, agate → COMPANY
GAME & LOCATION terabaikan	> 50% entitas gagal	komunitas valorant di jakarta sangat besar	valorant → O, jakarta → O	valorant → GAME, jakarta → LOCATION
Ambiguitas dengan kata benda umum	> 50% entitas gagal	Journey adalah <i>game</i> indie populer	Journey → O	Journey → GAME
Ambiguitas dengan kata umum	> 50% entitas gagal	Unity merilis engine versi terbaru	Unity → O	Unity → TECH

Analisis terhadap seluruh hasil eksperimen menunjukkan bahwa perbedaan kinerja antar-model terutama dipengaruhi oleh variasi cakupan bahasa dan kapasitas arsitektur. Model multilingual dengan ukuran besar memperlihatkan ketahanan yang lebih baik terhadap fenomena *domain shift* yang muncul akibat percampuran istilah bahasa Inggris dalam korpus domain *game*. Sebaliknya, model monolingual menghadapi kesulitan dalam mengenali istilah serapan dan entitas lintas bahasa karena keterbatasan kosakata pra-pelatihannya. Pola kesalahan ini menegaskan pentingnya kemampuan lintas-bahasa dalam menangani data dengan kompleksitas linguistik tinggi.

BAB VII

PEMBAHASAN

Eksperimen pada dataset GameNER-ID menunjukkan bahwa model Transformer, khususnya arsitektur multilingual berukuran besar seperti XLM-RoBERTa-large, mencapai kinerja terbaik dengan *F1-Score* sebesar 0.8386. Pencapaian ini menandakan bahwa model dengan cakupan lintas-bahasa mampu memahami konteks terminologi industri *game* secara lebih efektif dibandingkan model monolingual. Hasil tersebut juga memperkuat relevansi penggunaan arsitektur Transformer dalam menangani fenomena linguistik kompleks yang melibatkan *code-switching* dan *loanwords* dari bahasa Inggris. Untuk memahami kontribusi temuan ini secara menyeluruh, pembahasan berikut menguraikan posisi strategis dataset GameNER-ID di antara dataset NER berbahasa Indonesia, menelaah implikasi hasil eksperimen, serta membahas keterbatasan dan arah pengembangan di masa mendatang..

7.1 GameNER-ID di antara Dataset NER Indonesia

Dataset GameNER-ID menempati posisi penting di antara sumber daya NER berbahasa Indonesia lainnya. Berdasarkan Tabel 7.1, GameNER-ID berisi 3.351 kalimat dengan 10 tipe entitas, menjadikannya dataset berukuran moderat namun dengan keragaman entitas yang jauh lebih tinggi dibandingkan dataset seperti NERGrit (2.090 kalimat, 3 entitas) dan NERP (8.400 kalimat, 5 entitas). Tingginya variasi tipe entitas ini menunjukkan bahwa GameNER-ID dirancang untuk mencakup aspek semantik yang lebih luas, sesuai dengan

kompleksitas teks pada domain *game* yang mengandung istilah teknis dan nama-nama produk komersial.

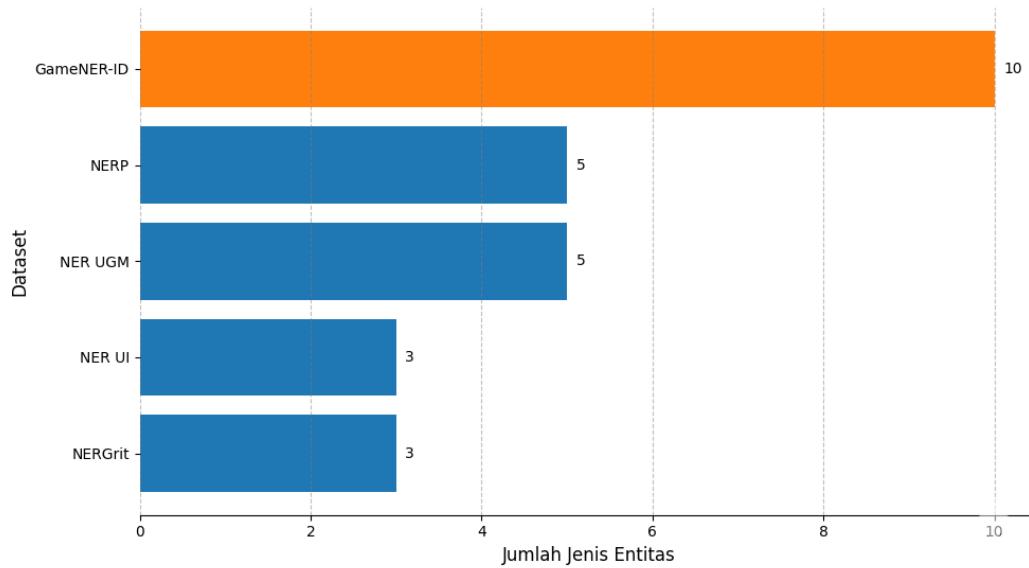
Tabel 7.1 Posisi GameNER-ID dengan Dataset NER Indonesia

Dataset	Sumber	Kalimat	Entitas	F1-Score Terbaik (Model)	Fitur Utama
GameNER-ID	Berita <i>Game</i>	3.351	10	0.8386 (XLM-RoBERTa-large)	COMPANY, EVENT, GAME, LOCATION, ORGANIZATION, PERSON, PLATFORM, QUANTITY, TECH, TIME
NERP	Berita <i>Online</i>	8.400	5	0.7925 (IndoBERT-large-p1)	PER, LOC, EVT, IND, FNB
NERGrit	Wiki	2.090	3	0.7909 (XLM-RoBERTa-base)	PERSON, PLACE, ORG
NER UGM	Tugas	2.343	5	0.7490 (IndoBERT-uncased)	PER, ORG, LOC, TIME, QTY
NER UI	Berita Umum	2.125	3	0.9010 (IndoBERT-uncased)	PER, LOC, ORG

Perbandingan yang ditampilkan pada tabel 7.2 selanjutnya memperlihatkan keunikan GameNER-ID dalam menghadirkan tujuh entitas unik/baru yang tidak ditemukan pada dataset lain, seperti GAME, COMPANY, PLATFORM, TECH, QUANTITY, EVENT, dan TIME.

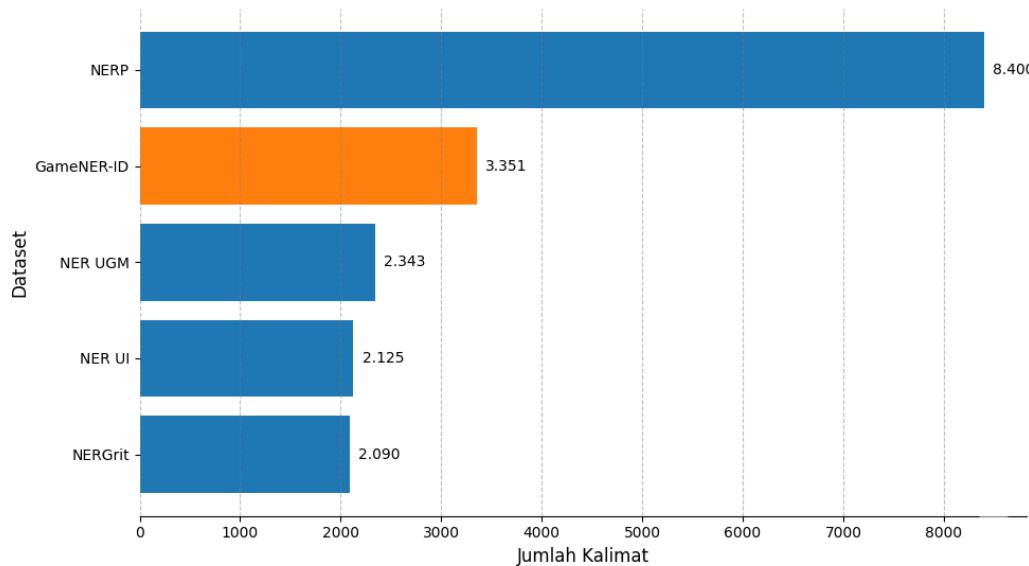
Tabel 7.2 Perbandingan Jenis Entitas di Seluruh Dataset NER Indonesia

Dataset	Jenis Entitas	Entitas Unik
GameNER-ID	COMPANY, EVENT, GAME, LOCATION, ORGANIZATION, PERSON, PLATFORM, QUANTITY, TECH, TIME	COMPANY, EVENT, GAME, PLATFORM, QUANTITY, TECH, TIME (7)
NERP	PER, LOC, EVT, IND, FNB	EVT, IND, FNB (3)
NERGrit	PERSON, PLACE, ORG	-
NER UGM	PER, ORG, LOC, TIME, QTY	TIME, QTY (2)
NER UI	PER, LOC, ORG	-



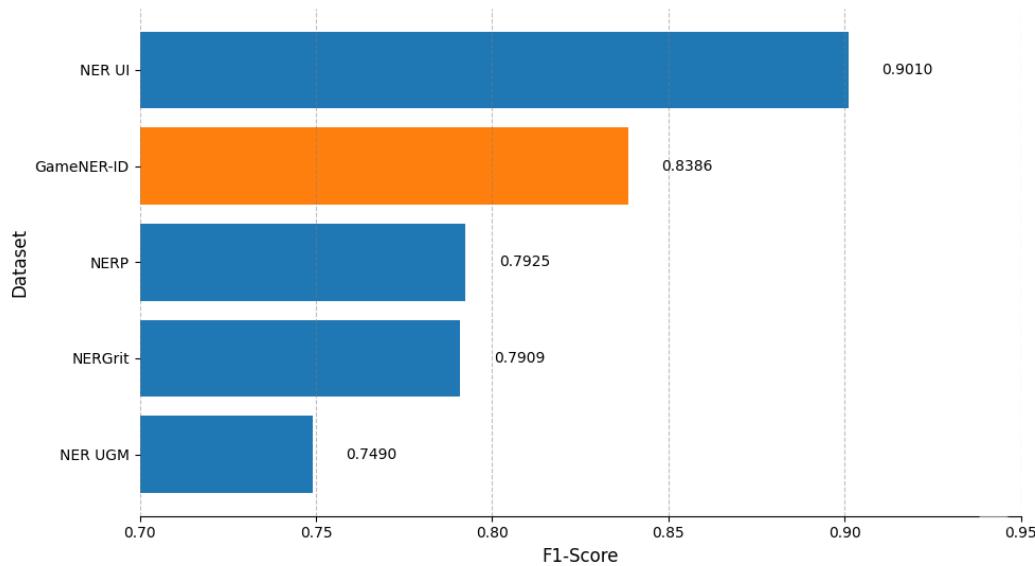
Gambar 7.1 Perbandingan Jumlah Entitas

Grafik yang ditunjukkan gambar 7.1 memperlihatkan keragaman entitas di setiap dataset. Semakin banyak jenis entitas, semakin kompleks tugas pengenalan entitas yang bisa dilakukan.



Gambar 7.2 Perbandingan Jumlah Kalimat

Grafik pada gambar 7.2 menunjukkan bahwa NERP adalah dataset terbesar. Namun, GameNER-ID memiliki ukuran yang signifikan (3.351 kalimat), menjadikannya dataset yang cukup besar dan representatif untuk domainnya.



Gambar 7.3 Perbandingan *F1-Score* Terbaik

GameNER-ID mencapai *F1-Score* kompetitif sebesar 0.8386 dan menempati posisi kedua setelah NER UI. Meskipun perbandingan ini tidak sepenuhnya setara karena perbedaan model yang digunakan pada masing-masing dataset, capaian tersebut tetap menunjukkan kualitas anotasi yang tinggi dan kemampuan dataset ini menghasilkan performa terbaik di konteksnya.

7.2 Implikasi Penelitian

Hasil eksperimen memberikan beberapa implikasi penting terhadap pengembangan sumber daya dan model NLP di Indonesia. Pertama, keberhasilan model XLM-RoBERTa menunjukkan bahwa pemrosesan teks pada domain modern seperti industri *game* memerlukan kapabilitas lintas-bahasa yang kuat. Fenomena *loanwords* dan *code-switching* yang dominan dalam teks *game* menjadi tantangan utama bagi model monolingual, sehingga pendekatan multilingual terbukti lebih adaptif. Kedua, kehadiran GameNER-ID memperluas cakupan penelitian NER di Indonesia dengan menyediakan dataset beranotasi tinggi pada domain yang sebelumnya belum terwakili. Dataset ini dapat digunakan untuk

melatih model yang lebih akurat dan kontekstual terhadap bahasa campuran yang sering muncul dalam industri kreatif digital.

Kontribusi GameNER-ID membuka peluang riset lanjutan di bidang analisis teks *game*, seperti analisis sentimen pada ulasan *game*, ekstraksi informasi produk, dan pemantauan tren industri. Penggunaan dataset ini juga berpotensi memperkuat riset-riset terkait information retrieval, text classification, dan entity linking yang memanfaatkan konteks domain spesifik. Dengan demikian, hasil penelitian ini tidak hanya memperkaya sumber daya linguistik, tetapi juga memperluas penerapan NLP dalam sektor ekonomi digital Indonesia.

7.3 Keterbatasan Penelitian

Beberapa keterbatasan perlu diakui agar hasil penelitian ini dapat dipahami secara proporsional. Ukuran dataset GameNER-ID masih tergolong moderat dibandingkan dataset NER berskala besar di bahasa lain, seperti CoNLL atau OntoNotes. Kondisi ini dapat membatasi kemampuan generalisasi model, terutama terhadap istilah baru atau variasi penulisan yang belum terwakili dalam data pelatihan. Selain itu, eksperimen hanya berfokus pada model berbasis BERT dan RoBERTa, tanpa melibatkan arsitektur generatif seperti T5 atau LLM instruction-based yang kini banyak digunakan dalam tugas NER modern. Variasi pendekatan seperti prompt-based learning atau adapter-tuning juga belum diuji dalam konteks ini.

Keterbatasan lain terletak pada cakupan domain yang masih berfokus pada teks berita. Walaupun berita menyediakan struktur kalimat yang formal dan informatif, banyak istilah domain *game* muncul secara alami pada forum, media sosial, dan blog, yang memiliki karakter bahasa lebih informal. Hal ini menandakan

perlunya perluasan korpus agar representasi bahasa pada domain *game* menjadi lebih lengkap.

7.4 Arah Penelitian Selanjutnya

Beberapa arah pengembangan dapat ditempuh untuk memperkuat hasil yang telah dicapai. Perluasan dataset GameNER-ID dengan menambahkan sumber data dari forum komunitas, media sosial, dan ulasan pengguna akan meningkatkan keragaman gaya bahasa serta cakupan terminologi. Diversifikasi sumber tersebut memperkaya konteks linguistik dan membantu model mengenali entitas dalam variasi bahasa yang lebih alami. Selain itu, pengembangan model yang lebih efisien dan ringan perlu dilakukan agar hasil dapat diterapkan pada sistem dengan keterbatasan sumber daya komputasi, seperti aplikasi daring atau perangkat mobile. Upaya tersebut membuka peluang penerapan teknologi NER pada skenario dunia nyata dengan jangkauan pengguna yang lebih luas.

GameNER-ID juga dapat dimanfaatkan untuk tugas lanjutan NLP, seperti Relation Extraction, guna mengidentifikasi relasi antarentitas, misalnya hubungan antara COMPANY dan GAME yang dikembangkannya. Pengembangan ini memperluas fungsi GameNER-ID dari sekadar dataset pengenalan entitas menjadi fondasi bagi riset lanjutan dalam pemahaman hubungan semantik antarentitas pada domain spesifik bahasa Indonesia. Dengan demikian, pengembangan lanjutan berpotensi memperkuat ekosistem NLP lokal melalui sumber daya yang lebih adaptif, terukur, dan aplikatif.

Seluruh pembahasan dalam bab ini menjelaskan kontribusi utama GameNER-ID, mulai dari posisi strategisnya sebagai dataset domain *game*, implikasi hasil eksperimen terhadap model Transformer, hingga keterbatasan dan

arah pengembangan di masa mendatang. Rangkaian analisis tersebut menegaskan pentingnya pengembangan berkelanjutan untuk meningkatkan kualitas representasi entitas dalam teks berbahasa Indonesia.

7.5 Integrasi Nilai Islam dalam Validasi Informasi

Hasil eksperimen menunjukkan bahwa model terbaik, yaitu XLM-RoBERTa-large, mencapai *F1-Score* sebesar 0.8386. Nilai ini menggambarkan performa yang kuat, meskipun masih terdapat sekitar 16% kesalahan prediksi yang menunjukkan bahwa sistem tetap memiliki batasan. Keberhasilan dan keterbatasan tersebut menegaskan bahwa proses verifikasi tetap menjadi aspek penting dalam pengolahan informasi otomatis. Konteks ini memiliki kesesuaian dengan prinsip Tabayyun dalam ajaran Islam, khususnya sebagaimana disebutkan dalam Qs. Al-Hujurat ayat 6:

يَا يَاهَا الَّذِينَ ءَامَنُرَا إِنْ جَاءَكُمْ فَاسِقٌ بِنَبِيٍّ فَتَبَيَّنُوا أَنْ تُصِيبُوْا قَوْمًا بِجَهَلٍ فَتُصْبِحُوْا عَلَىٰ مَا
فَعَلْنُمْ نَدِمِينَ

"Wahai orang-orang yang beriman, jika datang kepada kamu orang fasik membawa suatu berita, maka periksalah dengan teliti agar kamu tidak menimpa suatu musibah kepada suatu kaum tanpa mengetahui keadaannya yang menyebabkan kamu menyesal atas perbuatanmu itu" (Qs. al-Hujurat/49:6).

Makna ayat tersebut semakin menjadi jelas ketika diperkuat oleh penjelasan tafsir. Tafsir Wajiz menjelaskan bahwa Allah memberikan tuntunan kepada kaum beriman mengenai etika menerima informasi dari sesama manusia, termasuk dari orang fasik. Tafsir ini menguraikan bahwa sikap tergesa-gesa dalam menerima berita berpotensi menimbulkan tindakan yang salah dan berakhir pada penyesalan. Proses meneliti atau memverifikasi berita ditekankan sebagai upaya untuk menghindari kebodohan dan kecerobohan dalam bertindak. Penjelasan tersebut

menegaskan bahwa kehati-hatian merupakan langkah dasar sebelum memutuskan sesuatu yang berdampak pada pihak lain, sehingga prinsip Tabayyun memiliki relevansi langsung terhadap proses penyaringan informasi dalam penelitian ini.

Penjelasan Tafsir Tahlili memperkuat pesan tersebut dengan menegaskan bahwa berita dari orang fasik tidak boleh diterima sebelum diperiksa dengan seksama. Tafsir ini menyatakan bahwa seseorang yang meremehkan kefasikannya cenderung tidak memperdulikan kebenaran informasi yang disampaikannya, sehingga potensi munculnya berita bohong atau tidak akurat semakin besar. Tindakan yang diambil berdasarkan informasi yang tidak diteliti akan menimbulkan penyesalan, bahkan dapat menyebabkan kerugian jiwa dan harta. Penjelasan tersebut menggarisbawahi bahwa penelitian terhadap kebenaran informasi merupakan pedoman penting bagi setiap mukmin agar tidak terjerumus dalam dampak negatif yang sebenarnya dapat dihindari. Pemahaman mendalam dari kedua tafsir ini memberikan landasan moral yang kuat bagi prinsip verifikasi dalam sistem ekstraksi informasi.

Relevansi ajaran Tabayyun dengan penelitian ini tampak dalam tujuan pengembangan model *Named Entity Recognition* (NER) yang berfokus pada kemampuan mengekstraksi fakta secara akurat dari teks berita. Sistem yang memiliki akurasi tinggi akan semakin mampu membantu pengguna dalam mengenali informasi penting, seperti pelaku, lokasi, dan peristiwa, sehingga proses verifikasi kebenaran dapat dilakukan dengan lebih mudah. Peran model ini menjadi bentuk implementasi nilai Tabayyun dalam ranah teknologi, yaitu upaya menyaring kebenaran dari tumpukan data yang tidak terstruktur.

Analisis kesalahan yang ditemukan pada model memberikan pengingat bahwa teknologi tidak dapat sepenuhnya menggantikan kehati-hatian manusia. Kesalahan seperti salah mengidentifikasi nama seseorang sebagai lokasi menunjukkan bahwa hasil otomatis memerlukan verifikasi tambahan. Proses identifikasi error ini menjadi wujud nyata dari prinsip Tabayyun yang mengajarkan agar sebuah informasi tidak diterima begitu saja tanpa penelitian yang memadai. Keterkaitan antara ajaran agama dan hasil penelitian ini menegaskan bahwa akurasi teknologi dan etika kehati-hatian harus berjalan berdampingan untuk mencegah penyebaran informasi yang salah dan mendorong penggunaan teknologi secara bertanggung jawab.

BAB VIII

KESIMPULAN

Bab ini menyajikan rangkuman akhir dari seluruh rangkaian penelitian yang telah dilakukan. Bagian pertama, kesimpulan, merangkum temuan-temuan utama untuk menjawab tujuan penelitian secara lugas. Bagian kedua, saran, menguraikan rekomendasi untuk pengembangan penelitian di masa depan serta potensi penerapan praktis dari hasil penelitian ini.

8.1 Kesimpulan

Berdasarkan hasil analisis, eksperimen, dan pembahasan yang telah dilakukan, beberapa poin yang dapat disimpulkan dari penelitian ini adalah sebagai berikut:

1. Penelitian berhasil memanfaatkan korpus berita *online* dengan menghasilkan GameNER-ID sebagai sumber daya baru yang representatif untuk tugas *Named Entity Recognition* (NER) pada domain *game*. Dataset ini mencakup 3.351 kalimat terkuras dengan sepuluh tipe entitas relevan (GAME, COMPANY, PLATFORM, TECH, DATE, QUANTITY, LOCATION, PERSON, EVENT, ORGANIZATION) serta memiliki struktur anotasi yang konsisten dan cakupan terminologi yang luas. Hasil ini menunjukkan bahwa teks berita *online* yang sebelumnya tidak terstruktur dapat diubah menjadi sumber informasi terstruktur yang valid dan layak dijadikan tolok ukur pengembangan NLP pada sektor kreatif digital di Indonesia..
2. Pemanfaatan model IndoBERT untuk melakukan ekstraksi informasi, eksperimen menunjukkan bahwa model tersebut mampu menjalankan tugas

dengan baik melalui capaian *F1-Score* 0.8175. Namun, evaluasi komparatif menemukan bahwa model multilingual XLM-RoBERTa-large memberikan kinerja lebih tinggi dengan *F1-Score* 0.8386. Keunggulan ini menegaskan efektivitas pendekatan lintas-bahasa dalam menangani karakteristik linguistik pada teks *game*, seperti code-switching dan penggunaan istilah teknis bahasa Inggris yang dominan, sehingga ekstraksi informasi pada industri *game* di Indonesia paling efektif dilakukan menggunakan model multilingual.

8.2 Saran

Berdasarkan hasil kesimpulan di atas, terdapat beberapa saran yang dapat dijadikan acuan untuk pengembangan penelitian dan penerapan praktis di masa mendatang, yaitu sebagai berikut:

1. Perluasan dan diversifikasi sumber data memperkuat kualitas dan cakupan GameNER-ID. Penambahan data dari forum komunitas seperti Kaskus, media sosial seperti Twitter dan Facebook Gaming, serta ulasan pengguna di Steam dan Google Play Store meningkatkan keragaman gaya bahasa dan terminologi. Langkah ini memperluas konteks linguistik yang dipahami model sekaligus menyiapkannya menghadapi variasi komunikasi nyata dalam ekosistem industri *game*.
2. Augmentasi dan perbaikan distribusi entitas diperlukan untuk menyeimbangkan representasi kelas yang kurang terwakili. Teknik augmentasi terarah seperti penggantian sinonim atau deep back-translation dengan penyaringan kontekstual efektif meningkatkan ketahanan model terhadap variasi ekspresi. Penerapan strategi ini memperkuat kemampuan generalisasi dan stabilitas kinerja pada domain dengan sumber daya terbatas.

3. Eksplorasi arsitektur dan strategi *fine-tuning* modern membuka peluang peningkatan kinerja dan efisiensi. Model generatif seperti T5 atau Large Language Models (LLMs) berbasis instruksi menawarkan pemahaman semantik yang lebih mendalam. Pendekatan prompt-based learning atau adapter-tuning memungkinkan penyesuaian cepat terhadap karakteristik linguistik domain *game* tanpa memerlukan sumber daya komputasi besar.
4. Pendekatan few-shot learning memberikan arah adaptasi cepat terhadap dinamika industri *game* yang terus berubah. Kemunculan entitas baru seperti judul *game*, perusahaan, dan teknologi menuntut model beradaptasi dengan data terbatas. Kerangka seperti FewNER mampu mengoptimalkan pembelajaran dari sedikit contoh sehingga mempercepat pembaruan model tanpa pelatihan ulang berskala besar.
5. Ekstensi menuju tugas NLP lanjutan memperluas fungsi GameNER-ID dari sekadar dataset anotasi menjadi sumber daya linguistik terstruktur. Implementasi Relation Extraction memungkinkan pemetaan hubungan antarentitas, misalnya antara COMPANY dan GAME, atau antara GAME dan PLATFORM. Pendekatan ini membangun jejaring pengetahuan yang merepresentasikan struktur industri *game* secara semantik dan informatif.
6. Aplikasi praktis menegaskan relevansi dan efisiensi hasil pengembangan. Model berbasis GameNER-ID dapat dimanfaatkan untuk pemantauan tren industri *game*, ekstraksi informasi produk otomatis, serta analisis sentimen kontekstual terhadap ulasan pemain. Efisiensi model menjadi prioritas agar sistem dapat beroperasi pada perangkat dengan sumber daya terbatas.

Integrasi antara efisiensi, adaptabilitas, dan relevansi konteks memastikan penerapan NLP di sektor *game* berlangsung efektif dan berkelanjutan.

Secara umum, saran-saran tersebut diharapkan dapat menjadi dasar bagi pengembangan penelitian selanjutnya dalam memperluas cakupan GameNER-ID serta meningkatkan kapabilitas model NER di ranah bahasa Indonesia. Penguatan sinergi antara kualitas data, inovasi arsitektur model, dan penerapan praktis akan mendorong kemajuan teknologi NLP di sektor industri kreatif Indonesia.

DAFTAR PUSTAKA

- Ababil, M.H. dan G.J.B. Setiawan. 2022. Topic Modelling pada Ulasan *Game* Online Wildrift Menggunakan Latent Dirichlet Allocation (LDA). *Jurnal Pendidikan dan Konseling (JPDK)*. 4(6): 6185–6196.
- Abdullah, A.A. dkk. 2024. NER-RoBERTa: Fine-Tuning RoBERTa for *Named Entity Recognition (NER)* within low-resource languages. *arXiv preprint arXiv:2412.15252*. <https://doi.org/10.48550/arXiv.2412.15252>.
- Adiputra, W.M. 2019. Jurnalisme *Game* dan Membaurnya Karakter Konten Media: Studi Kasus the Lazy Monday, Gamebrott dan Playverse. *Journal Communication Spectrum*. 8(2).
- Aditya, P., A. Azzahra dan A. Wijaya. 2023. Analisis Sentimen Pemain Subway Surf Melalui Metode Naive Bayes Menurut Ulasan Play Store. *Jurnal Ilmiah Sistem Informasi*. 3(2): 267–275.
- Akbar, M. dan I.J. Asmara. 2022. Worker in Video *Game* Industry. *Journal of Games, Game Art, and Gamification*. 7(1): 22–26.
- Al Rosyid, M.T., I. Tohar dan M. Faisal. 2024. Penerapan Pendekatan Biofilik Pada Perancangan Fasilitas Industri Animasi dan *Game* di Surabaya. *Kohesi: Jurnal Sains dan Teknologi*. 4(9): 41–50.
- Alfawas, T.I., A. Rahim dan R. Rudiman. 2024. Penerapan Fitur Ekstraksi TF-IDF untuk Analisis Sentimen Ulasan *Game* Bus Simulator Indonesia dengan Algoritma Naive Bayes. *Innovative: Journal Of Social Science Research*. 4(5): 3177–3193.
- Almaudy, R. dan A.D. Kalifia. 2022. Analisis Sentimen Ulasan Aplikasi *Game* AOV di Play Store Menggunakan Rapidminer dan Algoritma Naïve Bayes. *Jurnal Ilmiah Komputasi*. 21(4): 493-502.
- AR, M.M. dkk. 2023. Pelatihan *Game* Edukasi Berbasis Sdgs Dalam Upaya Membentuk Karakter Melalui Penguatan Projek Profil Pelajar Pancasila Bagi Guru Mi Ziyadatul Ulum Desa Kambangan Barat. *Jurnal Abdimas Bina Bangsa*. 4(1): 416–424.
- Bifulco, I. dkk. 2021. An intelligent system for focused crawling from Big Data sources. *Expert Systems with Applications*. 184: 115560.
- Budi, I. dkk. 2005. Named entity recognition for the Indonesian language: combining contextual, morphological and part-of-speech features into a knowledge engineering approach. *Discovery Science: 8th International Conference*. Singapura: Springer.

- Conneau, A. dkk. 2019. Unsupervised Cross-lingual Representation Learning at Scale. *CoRR*. abs/1911.02116. <http://arxiv.org/abs/1911.02116> .
- Dave, E. dan A. Chowanda. 2024. IPerFEX-2023: Indonesian personal financial entity extraction using indoBERT-BiGRU-CRF model. *Journal of Big Data*. 11(1): 139.
- Devlin, J. dkk. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North*. Minneapolis: Association for Computational Linguistics.
- Dharviyanti, N.A.D. dan N. Wilantika. 2024. Rule-Based NER for Crime Information Extraction through Online News Site. *2024 International Conference on Information Technology Research and Innovation (ICITRI)*. Jakarta: IEEE.
- Fachri, M. 2014. Pengenalan entitas bernama pada teks bahasa Indonesia menggunakan Hidden Markov Model. Skripsi Tidak Diterbitkan. Yogyakarta: Universitas Gadjah Mada.
- Fajri, C. 2012. Tantangan Industri Kreatif-Game Online di Indonesia. *Jurnal Aspikom*. 1(5): 443–454.
- Fauziah, Y. dkk. 2021. A Systematic Literature Review of 5W1H in Manufacturing and Services Industries. *Jurnal Titra*. 9(2): 37-44.
- Fazrian, V., T. Suprapti dan R. Narasati. 2024. Penerapan Algoritma Naive Bayes Terhadap Analisis Sentimen Aplikasi Game Multiplayer Online Battle Arena. *JATI (Jurnal Mahasiswa Teknik Informatika)*. 8(1): 1005–1012.
- Febriansyah, T.Y. dan A. Qoiriah. 2024. Pengelompokan Ulasan Video Game Elden Ring pada Platform Steam dengan Metode Spectral Clustering Menggunakan Principal Component Analysis Berbasis Similarity Gaussian. *Journal of Informatics and Computer Science (JINACS)*. 5(3): 297–306.
- Firdaus, M.R. 2022. Analisis Sentimen Berbasis Aspek terhadap Data Ulasan Pengguna Pada Game Roblox Dengan Metode Support Vector Machine. Skripsi Tidak Diterbitkan. Bandung: Universitas Komputer Indonesia.
- Goh, E., O. Al-Tabbaa dan Z. Khan. 2023. Unravelling the complexity of the Video Game Industry: An integrative framework and future research directions. *Telematics and Informatics Reports*. 12: 100100.
- Grishman, R. 1997. Information extraction: Techniques and challenges. *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology*. Frascati: Springer.

- Gunawan, A., A. Hidayatullah dan A. Hidayat. 2021. Pengembangan E-Sport dan Industri Gaming Menggunakan Analisis SWOT. *Jurnal Syntax Transformation*. 2(04): 409–421.
- Hasibuan, M.I.Z. dan T. Triase. 2022. Implementasi Sistem Database Nosql Secara Realtime Menggunakan Firebase Realtime Database Pada Aplikasi Ourticle. *SIBATIK JURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*. 2(1): 1–24.
- Heartex. 2025. Label Studio: Open source data labeling platform. <https://labelstud.io>
- Hudrasyah, H. dkk. 2019. Marketing Strategy for Game Developer Based on Micro and Macro Environment in Indonesia. *Journal of Global Business and Social Entrepreneurship (GBSE)*. 5(14): 78–92.
- Istiqomah, N. dan F. Novika. 2025. Comparative Performance of IndoBERT and IndoLEM Baseline Models for Post-Disaster Health Information Extraction from Indonesian Online News. *Journal of Computer Science and Informatics Engineering*. 4(3): 158–174.
- Jain, S.M. 2022. Hugging face. Dalam *Introduction to transformers for NLP: With the hugging face library and models to solve problems*. Hlm. 51–67. [S.l.]: Springer.
- Juhaeni, J. dkk. 2023. Pengembangan Media Game Edukasi dalam Meningkatkan Hasil Belajar Matematika Kelas III Siswa Madrasah Ibtidaiyah. *Journal of Instructional and Development Researches*. 3(2): 58–66.
- Kamal, M. dan O. Hassanain. 2025. Automated Antibiogram Extraction from Unstructured Microbiology Reports: A Comparative Performance and Efficiency Analysis of Domain-Specific Named Entity Recognition (NER) Pipelines. *medRxiv*. 2023–2025.
- Khairunnisa, S.O., Z. Chen dan M. Komachi. 2024. Improving Domain-Specific NER in the Indonesian Language Through Domain Transfer and Data Augmentation. *Journal of Advanced Computational Intelligence and Intelligent Informatics*. 28(6): 1299–1312.
- Khasanah, L.A.I.U., I.E. Ningrum dan M.M. Huda. 2023. Pengembangan Game Edukasi Berbasis kearifan Lokal Berorientasi dalam Peningkatan Kemampuan Membaca Permulaan di Sekolah Dasar. *Jurnal Basicedu*. 7(1): 760–770.
- Koto, F., J.H. Lau dan T. Baldwin. 2021. IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization. *Proceedings of the 2021 Conference on*

Empirical Methods in Natural Language Processing. Online: Association for Computational Linguistics.

- Koto, F. dkk. 2020. IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *Proceedings of the 28th International Conference on Computational Linguistics.* Barcelona: International Committee on Computational Linguistics.
- Kusuma, E. dan M. Fadiana. 2024. Pemanfaatan *Game* Edukasi Wordwall untuk Meningkatkan Hasil Belajar Bahasa Indonesia Siswa Kelas V Sekolah Dasar. *Jurnal Basicedu.* 8(2): 1566–1573.
- Lample, G. dan A. Conneau. 2019. Cross-lingual language model pretraining. *arXiv preprint arXiv:1901.07291.* <https://arxiv.org/abs/1901.07291>
- Laras, N.G., C. Darmawan dan J. Jufrizal. 2024. Analisis Framing Pemberitaan PONPUS Al-Zaytun Pada Media TvOneNews.Com Dan CNN Indonesia. *Social Science and Contemporary Issues Journal.* 2(2): 426–430.
- Liu, Y. dan M. Liu. 2024. Research on *Named Entity Recognition* of Traditional Chinese Medicine Text Based on RoBERTa-BiLSTM-CRF. *2024 9th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS).* IEEE.
- Liu, Z. dkk. 2021. NER-BERT: A Pre-trained Model for Low-Resource Entity Tagging. *arXiv preprint arXiv:2112.00405.* <http://arxiv.org/abs/2112.00405>
- Miko, E. dan M.S. Jafar. 2018. Punya 47.000 media massa, Indonesia terbanyak di dunia. *ANTARA.* <https://www.antaranews.com/berita/684461/punya-47000-media-massa-indonesia-terbanyak-di-dunia>.
- Mo, Y. dkk. 2024. MCL-NER: cross-lingual named entity recognition via multi-view contrastive learning. *Proceedings of the AAAI conference on artificial intelligence.* 38(17): 18789–18797.
- Mulyani, S. 2023. Pengembangan Media Pembelajaran Interaktif Berbasis *Game* Edukasi untuk Meningkatkan Literasi dan Numerasi pada Siswa. *Jurnal Pendidikan dan Media Pembelajaran.* 2(03): 5–10.
- Muraina, I. 2022. Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts. *7th international Mardin Artuklu scientific research conference.* Mardin: Artuklu University.

- Nieborg, D.B. 2021. How to study *game* publishers: Activision Blizzard's corporate history. Dalam *Game production studies*. Amsterdam: Amsterdam University Press.
- Nurdy, A.H. 2024. Analisis sentimen ulasan *game* Stumble Guys pada Playstore menggunakan algoritma Naïve Bayes.
- Oralbekova, D. dkk. 2025. Using Kazakh NER Datasets for Multiclass Classification in The Legal Domain: A Comparative Study of BERT, GPT, and LSTM Models. *Proceedings of the 2025 14th International Conference on Software and Computer Applications*. 307–312.
- Pangerang, C.Z. 2024. Analisis framing berita silent majority pada media online Kompas.com, CNN Indonesia.com, dan Detik.com. Skripsi Tidak Diterbitkan. Jakarta: Universitas Bakrie.
- Polignano, M. dkk. 2021. Comparing Transformer-based NER approaches for analysing textual medical diagnoses. *CLEF (Working Notes)*. 818–833.
- Politowski, C. dkk. 2021. Game industry problems: An extensive analysis of the gray literature. *Information and Software Technology*. 134: 106538.
- Priyatna, G.G. 2023. Pemodelan Topik Terkait Ulasan Video *Game* dengan Genre Battle Royale Menggunakan Metode Bertopic dengan Fitur Guided Topic Modelling. Skripsi Tidak Diterbitkan. Jakarta: Fakultas Sains dan Teknologi UIN Syarif Hidayatullah.
- Rahman, P.A.A., E. Ermatita dan H.N. Irmarda. 2021. Analisis Sentimen Ulasan *Game* Harry Potter: Hogwarts Mystery Pada Situs Google Play Menggunakan Naïve Bayes Classifier. *Prosiding SEINASI-KESI*. 4(1): 38–47.
- Rahmawati, A., A. Alamsyah dan A. Romadhony. 2022. Hoax News Detection Analysis using IndoBERT Deep Learning Methodology. *2022 10th International Conference on Information and Communication Technology (ICoICT)*. IEEE.
- Ramadani, R. dkk. 2025. Reformulation of legal policy to support Indonesia's online gaming industry sector. *2nd International Conference Changing of Business Law (ICOCLB 2024)*. 403–410.
- Ramadhan, M.F. dan B. Siswoyo. 2024. Mengenal Model BERT dan Implementasinya untuk Analisis Sentimen Ulasan *Game*. *Prosiding SISFOTEK*. 8(1): 395–398.
- Richardson, L. 2007. *Beautiful soup documentation*. April.
- Rykała, P. 2020. The growth of the gaming industry in the context of creative industries. *Biblioteka Regionalisty*. 20: 124–136.

- Safrudin, M., M. Martanto dan U. Hayati. 2024. Perbandingan Kinerja Naïve Bayes Dan Support Vector Machine Untuk Klasifikasi Sentimen Ulasan *Game* Genshin Impact. *JATI (Jurnal Mahasiswa Teknik Informatika)*. 8(3): 3182–3188.
- Salfina, R. dkk. 2023. Perancangan *Game* Edukasi Untuk Mata Pelajaran Bahasa Jepang Kelas X Menggunakan Aplikasi Educandy Di SMKN 4 Payakumbuh. *Journal of Information System and Education Development*. 1(1): 1–7.
- Sang, E.F. dan J. Veenstra. 1999. Representing text chunks. *arXiv preprint cs/9907006* . <https://arxiv.org/abs/cs/9907006>.
- Sharma, P.R. 2019. *Selenium with Python-A Beginner's Guide: Get started with Selenium using Python as a programming language*. New Delhi: BPB Publications.
- Sihotang, S.G. 2020. Klasifikasi rating berdasarkan pemain menggunakan multinomial Naïve Bayes. Skripsi Tidak Diterbitkan. Medan: Universitas Sumatera Utara.
- Suciasih, L. 2024. Analisis Strategi Penguatan Industri Gim Online Dengan Pendekatan Klaster Industri. Tesis Tidak Diterbitkan. Jakarta: Universitas Mercu Buana.
- Syauqie, H., A.S. Nunka dan M.A.R. Fahrozi. 2024. Analisis Sentimen Ulasan Pengguna Aplikasi *Game* Sky Children Of The Light Menggunakan Metode Algoritma Naive Bayes (Studi Kasus Tingkat Kepuasan Pengguna). *Merkurius: Jurnal Riset Sistem Informasi dan Teknik Informatika*. 2(4): 65–75.
- Tarcar, A.K. dkk. 2019. Healthcare NER models using language model pretraining. *arXiv preprint arXiv:1910.11241*. <https://arxiv.org/abs/1910.11241>
- Ullah, F. dkk. 2025. EDU-NER-2025: *Named Entity Recognition* in Urdu Educational Texts using XLM-RoBERTa with X (formerly Twitter). *arXiv preprint arXiv:2504.18142*. <https://arxiv.org/abs/2504.18142>
- Ulya, S., H. Hapidin dan Z. Akbar. 2023. SIGANA Banjir: *Game* Edukasi Kesiapsiagaan Bencana Banjir Untuk Anak Usia 5-6 Tahun. *Murhum: Jurnal Pendidikan Anak Usia Dini*. 4(2): 151–164.
- Utami, M.A.A.T., P. Silvianti dan M. Masjkur. 2023. Algoritme Support Vector Machine untuk Analisis Sentimen Berbasis Aspek Ulasan *Game* Online Mobile Legends: Bang-Bang. *Xplore: Journal of Statistics*. 12(1): 63–77.

- Vaswani, A. dkk. 2017. Attention is all you need. *Advances in neural information processing systems*. 30.
- Vijayarani, S. dkk. 2015. Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*. 5(1): 7–16.
- Wibowo, A. dan I. Lubis. 2013. Pengaruh Lima Dimensi Berita Pada Media Online Khusus Game Terhadap Brand Awareness Produk Nintendo 3Ds. *Jurnal Kajian Komunikasi*. 1(2): 155–162.
- Wilie, B. dkk. 2020. IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*. Suzhou: Association for Computational Linguistics.
- Wu, S. dan M. Dredze. 2019. Beto, bentz, becas: The surprising cross-lingual effectiveness of BERT. *arXiv preprint arXiv:1904.09077*. <https://arxiv.org/abs/1904.09077>
- Xie, J. dkk. 2018. Neural cross-lingual named entity recognition with minimal resources. *arXiv preprint arXiv:1808.09861*. <https://arxiv.org/abs/1808.09861>
- Yang, Y. dkk. 2022. A Survey of Information Extraction Based on Deep Learning. *Applied Sciences*. 12(19): 9691.
- Zainuddin, Z. dkk. 2025. Entity Extraction in Indonesian Online News Using *Named Entity Recognition* (NER) with Hybrid Method Transformer, Word2Vec, Attention and Bi-LSTM. *JOIV: International Journal on Informatics Visualization*. 9(3): 964–973.

LAMPIRAN

Tabel 9.1 Tabel Dataset GameNER-ID (*dibaca menyamping, lalu kebawah*)

Token	Entitas	Token	Entitas	Token	Entitas
Ada	O	kemungkinan	O	Tango	B-COMPANY
Gameworks	I-COMPANY	atau	O	juga	O
Bethesda	B-COMPANY	ambil	O	keputusan	O
ini	O	karena	O	potensi	O
ekspansi	O	baru	O	.	O
Pengguna	O	sudah	O	bisa	O
melakukan	O	praregistrasi	O	untuk	O
mendapat	O	informasi	O	terkini	O
seputar	O	Project	B-GAME	Mugen	I-GAME
.	O	Nah	O	,	O
kali	O	ini	O	111dots	B-COMPANY
studio	I-COMPANY	kembali	O	dengan	O
Winstreak	O	Medal	O	Free	O
Fire	O	,	O	di	O
mana	O	kamu	O	dapat	O
melihat	O	tulisan	O	reward	O
atau	O	achievement	O	keren	O
.	O	Tak	O	hanya	O
Sony	B-COMPANY	,	O	namun	O
banyak	O	perusahaan	O	lainnya	O
yang	O	tidak	O	meyetujui	O
keputusan	O	Microsoft	B-COMPANY	lantaran	O
diangap	O	memonopoli	O	pasar	O
,	O	terutama	O	dengan	O
game	O	COD	B-GAME	yang	O
menjadi	O	favorit	O	para	O
gamer	O	.	O	Jadi	O
ratarata	O	setiap	O	harinya	O
kami	O	melakukan	O	pemutusan	O
akses	O	terhadap	O	1.500	B-QUANTITY
sampai	I-QUANTITY	2.000	I-QUANTITY	situs	O
dan	O	puluhan	B-QUANTITY	aplikasi	O
termasuk	O	aplikasi	O	game	O
terkait	O	perjudian	O	online	O
yang	O	serupa	O	dengan	O
Higgs	B-GAME	Domino	I-GAME	Island	I-GAME
,	O	kata	O	Budi	B-PERSON
.	O	LTD	O	,	O
PS4	B-PLATFORM	PS5	B-PLATFORM	Mortal	B-GAME
Kombat	I-GAME	11	I-GAME	WB	B-COMPANY
Games	I-COMPANY	.	O	PS4	B-PLATFORM
PS5	B-PLATFORM	Narutoshippuden	B-GAME	Ultimate	I-GAME
Ninja	I-GAME	Storm	I-GAME	4	I-GAME
Bandai	B-COMPANY	Namco	I-COMPANY	Entertainment	I-COMPANY
Inc.	I-COMPANY	,	O	PS4	B-PLATFORM
NBA	B-GAME	2K22	I-GAME	2K	B-COMPANY
Games	I-COMPANY	,	O	PS4	B-PLATFORM
PS5	B-PLATFORM	Outer	B-GAME	Wilds	I-GAME
Annapurna	B-COMPANY	Interactive	I-COMPANY	,	O
PS4	B-PLATFORM	Red	B-GAME	Dead	I-GAME
Redemption	I-GAME	2	I-GAME	Rockstar	B-COMPANY
Games	I-COMPANY	,	O	PS4	B-PLATFORM
Resident	B-GAME	Evil	I-GAME	Capcom	B-COMPANY
Co.	I-COMPANY	Ltd	I-COMPANY	,	O
PS4	B-PLATFORM	Soulcalibur	B-GAME	VI	I-GAME
Bandai	B-COMPANY	Namco	I-COMPANY	Entertainment	I-COMPANY
Inc.	I-COMPANY	,	O	PS4	B-PLATFORM
South	B-GAME	Park	I-GAME	The	I-GAME
Fractured	I-GAME	but	I-GAME	Whole	I-GAME
Ubisoft	B-COMPANY	,	O	PS4	B-PLATFORM
The	B-GAME	Artful	I-GAME	Escape	I-GAME
Annapurna	B-COMPANY	Interactive	I-COMPANY	,	O
PS4	B-PLATFORM	PS5	B-PLATFORM	The	B-GAME
Crew	I-GAME	2	I-GAME	Ubisoft	B-COMPANY
.	O	PS4	B-PLATFORM	Tom	B-GAME
Clancys	I-GAME	The	I-GAME	Division	I-GAME

Token	Entitas	Token	Entitas	Token	Entitas
Ubisoft	B-COMPANY	,	O	O	Microsoft
,	O	melalui	O	unit	O
bisnis	O	game	O	Xbox	B-PLATFORM
,	O	rajin	O	memperbarui	O
daftar	O	game	O	yang	O
bisa	O	dimainkan	O	secara	O
cumacuma	O	oleh	O	pengguna	O
yang	O	berlangganan	O	layanan	O
berbayar	O	PC	B-PLATFORM	Game	I-PLATFORM
Pass	I-PLATFORM	,	O	HoYoverse	B-COMPANY
sebagai	O	developer	O	dari	O
Honkai	B-GAME	Star	I-GAME	Rail	I-GAME
pastinya	O	akan	O	terus	O
rilis	O	karakter	O	baru	O
melalui	O	setiap	O	updatenya	O
,	O	Untuk	O	memulainya	O
,	O	EA	B-COMPANY	melakukan	O
uji	B-TECH	beta	I-TECH	Battlefield	B-GAME
dalam	O	lingkup	O	kecil	O
dahulu	O	sebelum	O	nanti	O
resmi	O	dirilis	O	,	O
Kalian	O	bisa	O	lihat	O
postingannya	O	di	O	bawah	O
ini	O	Jika	O	kalian	O
pemain	O	game	O	Genshin	B-GAME
Impact	I-GAME	,	O	pasti	O
ingat	O	dengan	O	insiden	O
dimana	O	banyak	O	Primogem	O
pemain	O	menjadi	O	minus	O
,	O	Sebelumnya	O	,	O
ada	O	pula	O	rumor	O
mengenai	O	jadwal	O	rilis	O
Switch	B-PLATFORM	2	I-PLATFORM	di	O
mana	O	manufaktor	O	SoC	O
asal	O	Cina	B-LOCATION	yang	O
juga	O	merupakan	O	salah	O
satu	O	partner	O	Nintendo	B-COMPANY
yaitu	O	PixArt	B-COMPANY	Imaging	I-COMPANY
,	O	mengisyaratkan	O	bahwa	O
Nintendo	B-COMPANY	akan	O	merilis	O
konsol	B-PLATFORM	tersebut	O	di	O
2024	B-TIME	melalui	O	laporan	O
finansial	O	mereka	O	,	O
Tak	O	terkecuali	O	dengan	O
sekuelnya	O	Remnant	B-GAME	2	I-GAME
yang	O	baru	O	dirilis	O
sekitar	O	seminggu	B-TIME	ini	I-TIME
pun	O	turut	O	sukses	O
mencuri	O	hati	O	para	O
fans	O	,	O	Terakhir	O
,	O	Mobile	B-GAME	Legend	I-GAME
adalah	O	game	O	yang	O
disertai	O	dengan	O	microtransaction	B-TECH
atau	O	pembelian	O	menggunakan	O
uang	O	asli	O	,	O
Karena	O	adanya	O	banyak	O
kemungkinan	O	yang	O	dapat	O
terjadi	O	di	O	dalam	O
game	O	,	O	The	B-GAME
Quarry	I-GAME	menawarkan	O	waktu	O
bermain	O	hingga	O	710	B-QUANTITY
jam	I-QUANTITY	lamanya	O	,	O
SpiderMan	B-GAME	2	I-GAME	Standard	I-GAME
Edition	I-GAME	Rp	B-QUANTITY	1.029.000	I-QUANTITY
SpiderMan	B-GAME	2	I-GAME	Digital	I-GAME
Deluxe	I-GAME	Edition	I-GAME	Rp	B-QUANTITY
1.179.000	I-QUANTITY	Edisi	O	paling	O
mahal	O	yaitu	O	Collectors	O
Edition	O	,	O	tidak	O
dapat	O	dibeli	O	melalui	O
PlayStation	B-PLATFORM	Store	I-PLATFORM	,	O
Oleh	O	sebab	O	itu	O

Token	Entitas	Token	Entitas	Token	Entitas
,	O	pemasukan	O	Agate	B-COMPANY
dikatakan	O	lebih	O	banyak	O
berasal	O	dari	O	<i>gamer</i>	O
di	O	luar	O	Indonesia	B-LOCATION
.	O	Sumber	O	Variety	B-ORGANIZATION
menyebut	O	casting	O	bahkan	O
menggunakan	O	penggalan	O	cerita	O
dari	O	gim	O	The	B-GAME
Last	I-GAME	of	I-GAME	Us	I-GAME
Part	I-GAME	II	I-GAME	sebagai	O
materi	O	audisi	O	.	O
The	B-GAME	Legend	I-GAME	of	I-GAME
Zelda	I-GAME	Tears	I-GAME	of	I-GAME
the	I-GAME	Kingdom	I-GAME	dengan	O
sempurna	O	mengkombinasikan	O	elemen	O
eksplorasi	O	dari	O	pendahulunya	O
dengan	O	mekanisme	O	Ultrahand	B-TECH
yang	O	baru	O	.	O
Berikut	O	adalah	O	tim	O
tim	O	Riyadh	B-EVENT	Masters	I-EVENT
2023	I-EVENT	Dota	I-EVENT	2	I-EVENT
yang	O	akan	O	bertanding	O
nanti	O	berdasarkan	O	group	O
mereka	O	.	O	Kami	O
mohon	O	maaf	O	atas	O
ketidaktahuan	O	kami	O	mengenai	O
hal	O	ini	O	dan	O
menghargai	O	bahwa	O	hal	O
ini	O	telah	O	menjadi	O
perhatian	O	kami	O	sehingga	O
kami	O	dapat	O	mengatasinya	O
secepatnya	O	,	O	ujar	O
Nightmare	B-COMPANY	Forge	I-COMPANY	Games	I-COMPANY
di	O	Twitter	B-PLATFORM	.	O
Nah	O	dalam	O	The	B-GAME
Last	I-GAME	of	I-GAME	Us	I-GAME
,	O	Ellie	B-PERSON	akan	O
melakukan	O	perjalanan	O	dengan	O
teman	O	terbaiknya	O	.	O
Sosok	O	Fatui	O	Harbingers	O
telah	O	diungkapkan	O	melalui	O
sebuah	O	video	O	teaser	O
A	O	Winters	O	Night	O
Lazzo	O	,	O	dimana	O
pada	O	akhirnya	O	penggemar	O
diberi	O	tahu	O	bagaimana	O
good	O	lookingnya	O	setiap	O
anggota	O	Harbinger	O	yang	O
terkenal	O	sebagai	O	villain	O
di	O	Genshin	B-GAME	Impact	I-GAME
ini	O	.	O	For	O
further	O	information	O	and	O
other	O	inquiries	O	,	O
you	O	can	O	contact	O
us	O	via	O	Marvels	B-GAME
SpidermanMan	I-GAME	2	I-GAME	sangat	O
ditunggutunggu	O	baik	O	oleh	O
gamers	O	pemilik	O	konsol	B-PLATFORM
PS5	I-PLATFORM	maupun	O	fans	O
dari	O	SpiderMan	B-GAME	itu	O
sendiri	O	.	O	Angka	O
ini	O	merupakan	O	penjualan	O
terbesar	O	diantara	O	konsol	B-PLATFORM
Nintendo	I-PLATFORM	lainnya	O	dengan	O
Nintendo	B-PLATFORM	DS	I-PLATFORM	berada	O
di	O	peringkat	O	kedua	O
948,7	B-QUANTITY	juta	I-QUANTITY	copy	I-QUANTITY
terjual	O	.	O	Judul	O
game	O	lainnya	O	yakni	O
Coffee	B-GAME	Talk	I-GAME	,	O
masuk	O	dalam	O	nominasi	O

Token	Entitas	Token	Entitas	Token	Entitas
Best	O	Narration	O	karena	O
memiliki	O	jalan	O	cerita	O
yang	O	bermacammacam	O	.	O
Football	B-GAME	Manager	I-GAME	2024	I-GAME
dikemas	O	dengan	O	visual	O
dan	O	pencahayaan	O	yang	O
lebih	O	ciamik	O	,	O
sehingga	O	laga	O	di	O
lapangan	O	sepak	O	bola	O
terlihat	O	lebih	O	realistik	O
.	O	Namun	O	,	O
perlu	O	diingatkan	O	kembali	O
bahwa	O	semu	O	informasi	O
ini	O	masihlah	O	bersifat	O
rumor	O	sehingga	O	belum	O
bisa	O	dipastikan	O	kebenarannya	O
sampai	O	ada	O	konfirmasi	O
resmi	O	dari	O	HoYoverse	B-COMPANY
.	O	Saat	B-TIME	ini	I-TIME
Fortnite	B-GAME	sendiri	O	resmi	O
mencapai	O	Chapter	B-EVENT	4	I-EVENT
Season	I-EVENT	3	I-EVENT	yang	O
baru	O	saja	O	diumumkan	O
melalui	O	Summer	B-EVENT	Game	I-EVENT
Fest	I-EVENT	2023	I-EVENT	.	O
Game	O	dijadwalkan	O	rilis	O
tanggal	O	20	B-TIME	September	I-TIME
2023	I-TIME	di	O	PS5	B-PLATFORM
,	O	Nintendo	B-PLATFORM	Switch	I-PLATFORM
,	O	Xbox	B-PLATFORM	Series	I-PLATFORM
XS	I-PLATFORM	,	O	dan	O
PC	B-PLATFORM	via	O	Steam	B-PLATFORM
.	O	Dengan	O	kesuksesan	O
dan	O	popularitas	O	yang	O
masif	O	,	O	tak	O
heran	O	jika	O	hingga	O
kini	O	masih	O	ada	O
banyak	O	gamer	O	yang	O
memainkan	O	game	O	garapan	O
Bethesda	B-COMPANY	itu	O	.	O
Hal	O	ini	O	membuat	O
Liu	B-PERSON	Wei	I-PERSON	dan	O
tim	O	developer	O	Genshin	B-GAME
khawatir	O	dengan	O	apa	O
yang	O	telah	O	mereka	O
kerjakan	O	.	O	Sehingga	O
,	O	kami	O	mempunyai	O
misi	O	untuk	O	menciptakan	O
dunia	O	epik	O	untuk	O
penggemar	O	Black	B-GAME	Panther	I-GAME
yang	O	ingin	O	menjelajahi	O
Wakanda	B-LOCATION	,	O	imbuhnya	O
sebagaimana	O	dikutip	O	KompasTekno	B-ORGANIZATION
dari	O	blog	O	resmi	O
EA	B-COMPANY	,	O	Selasa	B-TIME
11/7/2023	I-TIME	.	O	Sedangkan	O
untuk	O	gameplay	O	,	O
Ubisoft	B-COMPANY	punya	O	banyak	O
pilihan	O	bagaimana	O	kita	O
akan	O	memainkan	O	kedua	O
karakter	O	tersebut	O	dengan	O
merujuk	O	pada	O	game	O
AC	B-GAME	yang	O	memiliki	O
lebih	O	dari	O	satu	O
karakter	O	playable	O	.	O
Dengan	O	avatar	O	karakter	O
yang	O	terlihat	O	pada	O
dunia	O	arcade	B-PLATFORM	ini	O
,	O	mode	O	Arcade	B-TECH
Quest	I-TECH	sediki	O	terlihat	O
seperti	O	mode	O	Battle	O

Token	Entitas	Token	Entitas	Token	Entitas
Hub	O	yang	O	ada	O
pada	O	game	O	Street	B-GAME
Fighter	I-GAME	6	I-GAME	dimana	O
pemain	O	menggunakan	O	karakter	O
mereka	O	sendiri	O	berjalan	O
di	O	sebuah	O	Hub	O
.	O	Menurut	O	penjelasannya	O
.	O	ia	O	menghabiskan	O
sebagian	O	besar	O	waktunya	O
di	O	Stardew	B-GAME	Valley	I-GAME
dengan	O	menanam	O	berbagai	O
jenis	O	mixed	O	seeds	O
.	O	Itu	O	karena	O
artist	O	Lowlight	B-PERSON	yang	O
membuat	O	desain	O	karakter	O
Kaltsit	O	tersebut	O	tidak	O
lain	O	dan	O	tidak	O
bukan	O	adalah	O	Hai	B-PERSON
Mao	I-PERSON	selaku	O	Founder	O
.	O	Vice	O	President	O
and	O	the	O	Head	O
Producer	O	Arknights	B-GAME	.	O
Baca	O	pula	O	informasi	O
Gamebrott	B-ORGANIZATION	lainnya	O	tentang	O
Starfield	B-GAME	beserta	O	dengan	O
kabarkabar	O	menarik	O	lainnya	O
seputar	O	dunia	O	video	O
game	O	dari	O	saya	O
.	O	Muhammad	B-PERSON	Maulana	I-PERSON
.	O	Peter	B-PERSON	dikisahkan	O
belum	O	andal	O	dan	O
cenderung	O	ceroboh	O	menjadi	O
SpiderMan	B-GAME	karena	O	belum	O
memiliki	O	banyak	O	pengalaman	O
.	O	Steam	B-PLATFORM	No	B-GAME
Mans	I-GAME	Sky	I-GAME	Hal	O
ini	O	dimungkinkan	O	karena	O
planet	O	di	O	No	B-GAME
Mans	I-GAME	Sky	I-GAME	bersifat	O
procedurally	O	generated	O	,	O
yang	O	artinya	O	data	O
dalam	O	game	O	diciptakan	O
dengan	O	algoritma	B-TECH	,	O
bukan	O	secara	O	manual	O
.	O	Bagi	O	pemain	O
yang	O	berminat	O	,	O
bisa	O	menanti	O	info	O
selengkapnya	O	melalui	O	media	O
sosial	O	Honkai	B-GAME	Star	I-GAME
Rail	I-GAME	di	O	tautan	O
berikut	O	ini	O	.	O