

Optimalisasi Model SciBERT dengan Attention-BiLSTM-CRF untuk Pengenalan Entitas Penyakit dalam Teks Biomedis

Tahta Arya Pamungkas, Abu Salam*

Fakultas Ilmu Komputer, Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: ¹111202113900@mhs.dinus.ac.id.com, ^{2,*}abu.salam@dsn.dinus.ac.id

Email Penulis Korespondensi: abu.salam@dsn.dinus.ac.id

Submitted: 03/05/2025; Accepted: 31/05/2025; Published: 01/06/2025

Abstrak—Penelitian ini bertujuan untuk meningkatkan kinerja pengenalan entitas medis dalam teks biomedis dengan memodifikasi model SciBERT menggunakan Attention-BiLSTM-CRF. Meskipun SciBERT, yang berbasis pada arsitektur BERT dan dilatih dengan data teks biomedis, sudah efektif dalam pengenalan entitas medis, model ini masih memiliki keterbatasan dalam menangani entitas medis yang kompleks, terutama entitas bertumpuk (nested entities). Sebagai solusi, penelitian ini mengintegrasikan komponen Attention, BiLSTM, dan CRF ke dalam model SciBERT untuk memperbaiki akurasi pengenalan entitas. Hasil eksperimen menunjukkan bahwa model SciBERT + Attention-BiLSTM-CRF mengungguli model SciBERT pada semua metrik evaluasi utama. Precision meningkat sebesar 1.7% (dari 0.8221 menjadi 0.8364), Recall meningkat 2.9% (dari 0.8537 menjadi 0.8768), dan F1-Score meningkat 2.1% (dari 0.8372 menjadi 0.8554). Peningkatan ini menunjukkan bahwa modifikasi ini secara signifikan memperbaiki kemampuan model dalam mengenali entitas medis, terutama yang lebih kompleks, dalam teks biomedis. Penambahan Attention dan BiLSTM memperkaya pemahaman konteks, sementara CRF memastikan konsistensi antar label entitas. Dengan hasil ini, pendekatan ini dapat memberikan kontribusi besar bagi sistem otomatis dalam pemrosesan data medis.

Kata Kunci: Named Entity Recognition; SciBERT; Attention; BiLSTM; CRF; BC5CDR

Abstract—This research aims to improve the performance of medical entity recognition in biomedical text by modifying the SciBERT model with Attention-BiLSTM-CRF. Although SciBERT, based on the BERT architecture and trained on biomedical text data, has proven effective in entity recognition, it still has limitations in handling complex medical entities, especially nested entities. As a solution, this research integrates Attention, BiLSTM, and CRF components into the SciBERT model to enhance entity recognition accuracy. Experimental results show that the SciBERT + Attention-BiLSTM-CRF model outperforms the SciBERT model across all key evaluation metrics. Precision improved by 1.7% (from 0.8221 to 0.8364), Recall increased by 2.9% (from 0.8537 to 0.8768), and F1-Score increased by 2.1% (from 0.8372 to 0.8554). These improvements demonstrate that this modification significantly enhances the model's ability to recognize more complex medical entities in biomedical text. The addition of Attention and BiLSTM enriches contextual understanding, while CRF ensures consistency across entity labels. These results indicate that this approach could significantly contribute to automated systems in processing medical data.

Keywords: Named Entity Recognition; SciBERT; Attention; BiLSTM; CRF; BC5CDR

1. PENDAHULUAN

Kemajuan teknologi dalam bidang kecerdasan buatan (AI) telah membawa perubahan besar di banyak sektor, salah satunya di dunia medis[1], [2]. Seiring dengan pesatnya perkembangan data medis, terutama yang ada dalam bentuk teks, tantangan utama yang dihadapi adalah bagaimana cara mengelola dan menganalisis data yang begitu besar dan kompleks ini[3]. Data medis yang banyak tersimpan dalam format teks tidak terstruktur, seperti yang ada pada Rekam Medis Elektronik (EHR), menjadi salah satu hambatan terbesar dalam memprosesnya secara otomatis[3], [4]. Salah satu solusi yang bisa membantu adalah Named Entity Recognition (NER), sebuah teknik dalam Natural Language Processing (NLP) yang digunakan untuk mengenali dan mengklasifikasikan entitas penting dalam teks, seperti nama penyakit, obat-obatan, prosedur medis, dan entitas lainnya yang relevan[5], [6].

Penerapan teknologi NER di bidang medis dapat mengotomatisasi proses ekstraksi informasi medis dari teks yang tidak terstruktur. Ini sangat penting, terutama dalam membantu pengelolaan data medis yang lebih efisien dan mendukung pengambilan keputusan klinis berbasis bukti. Namun, dalam penerapannya, NER menghadapi tantangan besar dalam bidang medis, terutama terkait dengan kerumitan dan variasi istilah medis yang ada. Istilah medis sering kali sangat teknis dan memiliki banyak variasi, seperti singkatan atau sinonim yang bisa membingungkan model-model NLP standar[5], [7].

Untuk mengatasi masalah ini, diperlukan model NLP yang dirancang khusus untuk memahami teks medis dengan baik. Salah satu model yang cukup menjanjikan adalah SciBERT, sebuah varian dari model BERT (Bidirectional Encoder Representations from Transformers) [8] yang dilatih menggunakan teks ilmiah, termasuk teks-teks medis. Dengan menggunakan SciBERT, model dapat lebih memahami konteks medis serta istilah teknis yang sering ditemukan dalam teks biomedis. Ini menjadikan SciBERT lebih unggul dibandingkan dengan BERT standar, yang awalnya hanya dilatih menggunakan teks dari domain umum[9].

Namun, meskipun SciBERT menunjukkan hasil yang cukup baik dalam mengenali entitas medis, masih terdapat beberapa keterbatasan. Salah satunya adalah ketidakmampuannya dalam mengatasi entitas medis yang lebih rumit, terutama dalam hal entitas yang saling bergantung atau nested entities (entitas bertumpuk)[10]. Sebagai contoh, ketika ada penyakit dengan nama panjang atau penyakit yang terstruktur secara bertahap, model ini sering kesulitan untuk mengidentifikasi dan mengklasifikasikan entitas tersebut dengan akurat. Selain itu, masalah terkait hubungan

kompleks antar entitas medis, seperti kombinasi obat dan efek sampingnya, atau penyakit dengan gejala yang tumpang tindih, juga menjadi tantangan bagi SciBERT murni.

Untuk itu, diperlukan pendekatan yang lebih canggih untuk mengatasi masalah ini, salah satunya adalah dengan menambahkan teknik-teknik lanjutan seperti Attention-BiLSTM-CRF[11]. Pendekatan ini menggabungkan kekuatan BiLSTM (Bidirectional Long Short-Term Memory) yang efektif dalam memproses urutan teks dengan kemampuan CRF (Conditional Random Field)[12] untuk meningkatkan konsistensi prediksi entitas dalam urutan token yang saling bergantung. BiLSTM sendiri memiliki keunggulan dalam memproses konteks kata dari dua arah (dari depan ke belakang dan sebaliknya), yang memungkinkan model untuk menangkap hubungan kata yang lebih kompleks dan kontekstual. Sedangkan CRF digunakan untuk memastikan prediksi label yang lebih akurat, khususnya pada entitas yang saling tumpang tindih[12].

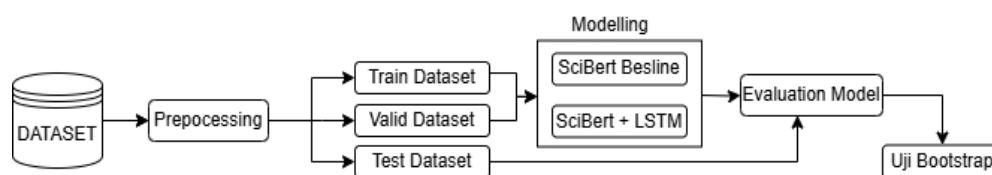
Selain itu, mekanisme perhatian (attention) yang ditambahkan pada model ini memberikan fokus lebih pada kata-kata atau bagian teks yang lebih penting, yang pada gilirannya dapat meningkatkan akurasi model dalam mengidentifikasi entitas yang lebih kompleks dan saling terhubung[13]. Dengan kombinasi ini, diharapkan model dapat lebih efektif dalam mengenali entitas medis yang lebih rumit dan menghasilkan prediksi yang lebih stabil. Pendekatan ini sangat cocok diterapkan pada dataset medis seperti BC5CDR[14], yang berfokus pada entitas chemical dan disease dalam teks medis, di mana sering ditemukan entitas yang tumpang tindih dan hubungan antar kata yang rumit. BC5CDR juga sering menghadapi tantangan terkait ketidakseimbangan data, dan dengan menggunakan model Attention-BiLSTM-CRF, diharapkan kinerjanya dalam mendeteksi dan mengklasifikasikan entitas medis bisa lebih akurat dan konsisten[11].

BC5CDR adalah salah satu dataset yang sering digunakan dalam penelitian NER untuk entitas medis, khususnya dalam pengenalan penyakit dan zat kimia dalam teks biomedis[15]. Dataset ini mengandung anotasi yang sangat detail mengenai nama-nama penyakit dan zat kimia yang terdapat dalam literatur medis. Keberagaman dan kedalaman data dalam BC5CDR sangat membantu dalam melatih dan menguji model-model NER, khususnya dalam konteks medis. Dengan menggunakan dataset ini, model SciBERT[16] yang dimodifikasi dengan Attention-BiLSTM-CRF diharapkan dapat menghasilkan akurasi yang lebih tinggi dalam mengenali entitas medis, baik yang sederhana maupun yang lebih kompleks[17].

Penelitian ini bertujuan untuk mengembangkan dan menguji model NER berbasis SciBERT yang dimodifikasi dengan Attention-BiLSTM-CRF untuk meningkatkan kinerja dalam pengenalan entitas penyakit dalam teks biomedis. Dengan menggunakan dataset BC5CDR, diharapkan penelitian ini dapat memberikan kontribusi yang signifikan dalam pengembangan teknologi NER yang lebih akurat dan efisien dalam pengelolaan data medis. Hasil dari penelitian ini juga diharapkan dapat membuka peluang untuk penerapan lebih lanjut dalam sistem kesehatan berbasis teknologi NLP, yang akan mempercepat pengambilan keputusan klinis dan pengelolaan informasi medis secara lebih optimal.

2. METODOLOGI PENELITIAN

Dalam penelitian ini, metode yang digunakan merupakan serangkaian langkah sistematis yang dirancang oleh peneliti. Rangkaian tahapan tersebut ditampilkan pada Gambar 1. Penelitian diawali dengan membagi dataset BC5CDR menjadi tiga bagian, yaitu train dataset, validation dataset, dan test dataset. Seluruh data kemudian diproses melalui tahap preprocessing untuk memastikan data siap digunakan dalam proses pemodelan. Pada tahap pemodelan, dilakukan pelatihan dengan dua pendekatan. Pendekatan pertama menggunakan model SciBERT baseline tanpa modifikasi tambahan, sedangkan pendekatan kedua mengintegrasikan SciBERT dengan lapisan attention, Bidirectional LSTM (BiLSTM), dan Conditional Random Fields (CRF) untuk meningkatkan akurasi prediksi entitas. Setelah proses pelatihan selesai, kedua model dievaluasi, dan perbandingan kinerja dilakukan menggunakan uji bootstrap[18], [19] untuk menentukan model yang memberikan hasil terbaik.



Gambar 1. Alur Penelitian

2.1 Dataset

Dataset yang digunakan dalam penelitian ini adalah BC5CDR (Biomedical Chemical and Disease Relation), yang merupakan dataset berfokus pada pengenalan entitas dalam teks biomedis, khususnya entitas yang berhubungan dengan bahan kimia (Chemical) dan penyakit (Disease). Dataset ini berisi teks-teks medis yang diambil dari publikasi-publikasi di PubMed, yang terdiri dari abstrak-artikel ilmiah yang telah dianotasi dengan format BIO tagging. Format BIO tagging ini digunakan untuk menandai entitas dalam teks dengan tiga kategori utama: B (begin), I (inside), dan O (outside), yang memudahkan pemisahan antara entitas yang saling terkait.

Tabel 1. Label Dataset

Label	Keterangan
O	Token yang bukan entitas (Outside)
B-Chemical	Awal dari entitas yang berjenis bahan kimia (Chemical)
B-Disease	Awal dari entitas yang berjenis penyakit (Disease)
I-Disease	Bagian dalam entitas penyakit (Disease)
I-Chemical	Bagian dalam entitas bahan kimia (Chemical)

Seperti yang terlihat pada Tabel 1, dataset BC5CDR terdiri dari lima label utama untuk analisis Named Entity Recognition (NER), yang masing-masing memiliki peran penting dalam identifikasi entitas. Pada Gambar 1 dan Tabel 2 dataset ini juga di bagi menjadi tiga bagian yaitu Train Dataset yang digunakan untuk melatih model dalam mengenali dan mengidentifikasi entitas dalam teks biomedis, Valid Dataset yang digunakan untuk mengevaluasi model selama proses pelatihan, dan Test Dataset untuk mengevaluasi kemampuan model dalam mengenali entitas pada data yang belum pernah dilihat sebelumnya.

Tabel 2. Entitas Label

Label	Train	Validation	Test
O	99324	98681	106194
B-Chemical	5101	5270	5282
B-Disease	4033	4094	4321
I-Disease	2208	2119	2204
I-Chemical	519	484	354

Distribusi kemunculan masing-masing label dapat dilihat lebih rinci pada Tabel 2, yang menunjukkan frekuensi kemunculan setiap label dalam dataset. Hal ini memberikan gambaran mengenai seberapa banyak setiap entitas baik itu penyakit maupun bahan kimia ditemui dalam dataset yang terdiri dari teks-teks medis tersebut.

Proses pembagian dataset ini dilakukan dengan teknik random split, yang memastikan distribusi data yang merata antara ketiga bagian dataset. Pembagian ini bertujuan untuk menghindari bias dalam pelatihan dan evaluasi model, sehingga model dapat diuji dengan cara yang adil dan dapat menghasilkan hasil yang lebih representatif.

Selain itu, dataset BC5CDR yang telah tersedia secara publik memberikan standar evaluasi yang lebih objektif. Dengan menggunakan dataset yang telah dikenal luas di kalangan peneliti biomedis, model yang dihasilkan dapat dibandingkan secara langsung dengan penelitian-penelitian sebelumnya dan dipastikan memiliki kemampuan generalisasi yang baik.

2.2 Preprocessing

Preprocessing adalah tahap penting dalam persiapan data untuk pengenalan entitas (NER). Pada tahap ini, data mentah yang berupa teks harus dipersiapkan sehingga model dapat memprosesnya dengan efektif dan akurat. Proses preprocessing membantu menghilangkan kebisingan atau informasi yang tidak relevan dalam teks, serta mengubahnya menjadi format yang sesuai untuk model[20]. Berikut adalah langkah-langkah yang dilakukan dalam tahap preprocessing:

- Pembersihan Teks**, bertujuan untuk menghilangkan elemen-elemen yang tidak relevan, sehingga model dapat fokus pada informasi penting. Hal ini dilakukan dengan menghapus tanda baca yang tidak memiliki makna penting dalam konteks entitas penyakit dan kimia, serta menghapus karakter khusus dan simbol seperti angka, tanda kurung, atau simbol lainnya yang tidak membantu pemahaman konteks. Selain itu, teks yang berisi metadata atau informasi tidak terkait seperti nomor halaman, header, dan catatan kaki juga dihapus agar model hanya memproses teks yang relevan dan berguna dalam mengenali entitas.
- Case Folding**, tahap ini mengubah seluruh teks menjadi huruf kecil untuk memastikan konsistensi dalam pengolahan teks dan mengurangi variasi token yang sebenarnya merepresentasikan kata yang sama (misalnya, "Diabetes" dan "diabetes" dianggap sama). Langkah ini membantu model dalam mengenali entitas tanpa terganggu oleh variasi huruf kapital.
- Tokenisasi**, setelah proses pembersihan tahap berikutnya adalah tokenisasi, yaitu proses memecah teks menjadi unit-unit terkecil yang disebut token. Tokenisasi ini dilakukan untuk memungkinkan model memproses teks dengan lebih efisien[21]. Pada tahap ini, teks dibagi menjadi kata-kata terpisah berdasarkan spasi dan tanda baca. Selain itu, subword tokenization digunakan untuk menangani kata-kata yang belum pernah dilihat model sebelumnya, dengan memecahnya menjadi bagian yang lebih kecil (misalnya, "naloxone" menjadi "na", "lo", "xone"). Pendekatan ini juga mempertimbangkan entitas dalam teks, memastikan entitas seperti "diabetes mellitus" tetap utuh meskipun terdiri dari dua kata, agar keterkaitannya tetap terjaga untuk pelabelan yang akurat.
- Labeling**, pada tahap ini setiap token yang berhubungan dengan entitas penyakit atau kimia diberi label sesuai dengan jenis entitasnya menggunakan skema BIO (Begin, Inside, Outside)[22]. Skema ini membantu menandai setiap token yang merupakan bagian dari entitas atau bukan. Token yang berada di awal entitas diberi label B-X (misalnya, B-Disease untuk entitas penyakit), token yang berada di dalam entitas diberi label I-X, dan token yang bukan bagian dari entitas diberi label O (Outside). Contohnya, dalam kalimat "aloxone adalah obat untuk overdosis

penyakit diabetes mellitus", "naloxone" diberi label B-Chemical karena itu adalah awal dari entitas kimia, sedangkan "diabetes" diberi label B-Disease dan "mellitus" diberi label I-Disease karena mereka membentuk entitas penyakit yang sama.

2.3 SciBERT

Setelah tahap Preprocessing selesai, langkah selanjutnya adalah Modeling, di mana kita menggunakan Baseline SciBERT sebagai model dasar untuk Named Entity Recognition (NER) pada teks biomedis[16]. SciBERT, yang merupakan pengembangan dari model BERT, telah dilatih sebelumnya pada teks ilmiah dan biomedis. Hal ini membuat SciBERT sangat cocok untuk NER yang memerlukan pemahaman mendalam mengenai konteks medis, seperti Chemical (zat kimia) dan Disease (penyakit). Dalam penelitian ini, Baseline SciBERT digunakan untuk membangun dasar model pengenalan entitas medis dalam teks biomedis.

2.4 att-BiLSTM-CRF

Setelah tahap Modeling dengan Baseline SciBERT selesai, langkah berikutnya adalah melakukan modifikasi pada model dengan menambahkan Attention-BiLSTM-CRF[11], [17]. Modifikasi ini bertujuan untuk memperbaiki kemampuan model dalam mengenali entitas medis yang lebih kompleks dan tidak sering muncul dalam teks biomedis. Dalam hal ini, kita menambahkan tiga komponen utama untuk meningkatkan akurasi model, yaitu Attention, BiLSTM, dan CRF.

2.5 Evaluasi

Tahap evaluasi dalam penelitian ini terdiri dari dua langkah utama. Langkah pertama adalah pengujian model, yang melibatkan pengujian kedua model, baik baseline model maupun model yang telah dilatih, menggunakan data uji yang belum pernah dilihat sebelumnya. Pengujian ini bertujuan untuk mengukur sejauh mana model dapat menggeneralisasi dan memberikan prediksi yang akurat pada data yang baru. Hasil dari pengujian ini berupa prediksi entitas yang akan digunakan untuk tahap evaluasi kinerja model.

Langkah kedua adalah evaluasi kinerja model, yang dilakukan menggunakan tiga metrik utama: Precision, Recall, F1-Score, dan Accuracy, yang sangat relevan dalam Named Entity Recognition (NER)[23], terutama ketika bekerja dengan dataset yang tidak seimbang. Precision mengukur akurasi model dalam mendeteksi entitas relevan, Recall menilai seberapa banyak entitas relevan yang berhasil ditemukan, dan F1-Score memberikan gambaran menyeluruh mengenai keseimbangan antara kedua metrik tersebut. Precision mengukur proporsi entitas yang diprediksi oleh model yang benar-benar relevan, dibandingkan dengan jumlah total entitas yang diprediksi relevan oleh model(1).

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

Recall mengukur seberapa banyak entitas relevan yang berhasil ditemukan oleh model dari total entitas relevan yang ada(2).

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

F1-Score menggabungkan kedua metrik tersebut untuk memberikan gambaran yang lebih komprehensif mengenai keseimbangan antara precision dan recall(3). Mengingat pentingnya keseimbangan dalam NER, F1-Score menjadi metrik utama yang digunakan dalam penelitian ini.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

Accuracy mengukur seberapa banyak prediksi yang benar dilakukan oleh model dibandingkan dengan total prediksi yang dilakukan. Accuracy memberikan gambaran umum tentang seberapa efektif model dalam memprediksi entitas secara keseluruhan, baik positif maupun negative(4).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + False\ Positives + True\ Negatives + False\ Negatives} \quad (4)$$

Selanjutnya, untuk memperoleh gambaran umum tentang kinerja model, dihitung rata-rata (mean) dari nilai Precision, Recall, dan F1-Score untuk setiap model dan label yang diuji. Rata-rata ini memberikan pandangan keseluruhan tentang kinerja model di seluruh entitas yang dianalisis(5).

$$Rata - rata = \frac{\sum_{i=1}^n x_i}{n} \quad (5)$$

Selain itu, standar deviasi digunakan untuk mengukur konsistensi hasil evaluasi model. Nilai standar deviasi yang rendah menunjukkan bahwa model menghasilkan hasil yang lebih konsisten, sementara standar deviasi yang tinggi menandakan adanya variasi besar dalam kinerja model di seluruh label(6).

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \tag{6}$$

Evaluasi ini memungkinkan penilaian objektif terhadap kinerja setiap model yang diuji dan memberikan gambaran yang lebih jelas tentang seberapa efektif model dalam mengidentifikasi entitas medis, terutama dalam menghadapi dataset yang tidak seimbang.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dan analisis dari proses pelatihan serta evaluasi model baseline SciBERT dan model Modified SciBERT yang telah dimodifikasi dengan penambahan Attention, Bidirectional Long Short-Term Memory (BiLSTM), dan Conditional Random Field (CRF). Pada tahap awal, dilakukan fine-tuning terhadap model baseline SciBERT dengan pengaturan parameter yang disesuaikan untuk membangun fondasi yang kuat dalam mengenali entitas medis pada teks biomedis.

Selanjutnya, evaluasi performa dilakukan dengan membandingkan hasil kedua model menggunakan tiga metrik evaluasi utama, yaitu Precision, Recall, dan F1-Score. Evaluasi ini difokuskan pada kemampuan model dalam mengenali lima jenis label entitas, yakni Outside (O), Beginning-Chemical (B-Chemical), Beginning-Disease (B-Disease), Inside-Disease (I-Disease), dan Inside-Chemical (I-Chemical). Analisis terhadap hasil evaluasi ini bertujuan untuk menilai efektivitas modifikasi model dalam meningkatkan kinerja Named Entity Recognition (NER) pada domain biomedis.

3.1 Training Baseline SciBERT

Pelatihan model Baseline SciBERT dilakukan dengan menggunakan TrainingArguments yang disesuaikan. Berikut adalah konfigurasi Parameter yang digunakan untuk fine-tuning pada model SciBERT:

Tabel 3. Parameter SciBERT

Parameters	Nilai	Deskripsi
learning_rate	(2e-5)	Nilai 2e-5 umum digunakan untuk fine-tuning model SciBERT .
batch_size	(16)	Batch size 16 dipilih untuk menjaga kestabilan pelatihan tanpa membebani memori GPU.
epochs	(35)	35 epoch dipilih untuk memastikan model cukup belajar dari data.
weight_decay	(0.05)	Nilai 0.05 digunakan untuk menjaga model tetap umum dan tidak menghafal data.
Optimizer	(AdamW)	Digunakan karena menyesuaikan learning rate
Early Stopping	(3)	Pelatihan dihentikan setelah 3 epoch berturut-turut tanpa perbaikan
dropout rate	(0.15)	diterapkan untuk mencegah overfitting

Dengan pengaturan Parameter yang disesuaikan ini, model SciBERT diharapkan dapat melakukan pelatihan yang efektif, mengenali entitas medis secara akurat, dan menghindari masalah seperti overfitting pada data yang terbatas. Pelatihan ini akan menjadi dasar yang kuat untuk modifikasi model lebih lanjut dengan Attention-BiLSTM-CRF untuk meningkatkan kinerjanya lebih lanjut.

3.2 Modified

Dalam upaya meningkatkan performa model dalam mengenali entitas medis dari teks biomedis, modifikasi terhadap model baseline SciBERT menjadi hal yang krusial. Modifikasi ini dilakukan dengan menambahkan mekanisme Attention, BiLSTM (Bidirectional Long Short-Term Memory), serta lapisan CRF (Conditional Random Field). Attention berfungsi untuk memungkinkan model memfokuskan perhatian pada bagian teks yang relevan, khususnya entitas medis yang tersembunyi dalam kalimat panjang. BiLSTM memperkaya pemahaman model terhadap konteks sekuensial dari dua arah, sementara CRF memastikan konsistensi prediksi antar label untuk Named Entity Recognition (NER).

Model SciBERT + Attention-BiLSTM-CRF dimulai dengan pemrosesan teks menggunakan SciBERT, yang menghasilkan representasi vektor untuk setiap token dalam input. Output dari SciBERT kemudian diteruskan ke lapisan BiLSTM yang memproses data sekuensial dari dua arah (kiri ke kanan dan kanan ke kiri) untuk menangkap hubungan yang lebih mendalam antar token dalam teks. BiLSTM ini terdiri dari 2 lapisan, masing-masing dengan 256 unit per lapisan, yang memungkinkan model memahami konteks sekuensial secara lebih kompleks. Setelah melewati BiLSTM, output diteruskan ke lapisan Attention, yang memberikan bobot pada token-token yang relevan dalam kalimat. Mekanisme Attention ini diterapkan dengan menggunakan dua lapisan Linear yang disertai dengan fungsi aktivasi ReLU, diikuti oleh lapisan Dropout untuk regularisasi. Attention memungkinkan model lebih fokus pada kata atau token yang penting dalam konteks pengenalan entitas medis. Setelah token-token yang relevan diberi bobot melalui Attention, output tersebut diteruskan ke lapisan Conditional Random Field (CRF). CRF digunakan untuk memastikan konsistensi dalam prediksi label antar token yang berurutan. Dengan cara ini, CRF membantu model

memprediksi entitas dengan cara yang lebih logis dan koheren, menghindari label yang tidak sesuai secara struktur, khususnya pada entitas medis yang saling bertumpuk atau bergantung.

Dengan pengaturan parameter yang optimal seperti learning rate, batch size, jumlah epoch, weight decay, dan warmup steps, pelatihan model Modified SciBERT diharapkan berlangsung lebih stabil, efisien, dan akurat. Langkah ini merupakan bagian penting dari proses modifikasi model untuk menghasilkan performa terbaik dalam ekstraksi informasi medis, yang pada akhirnya diukur melalui metrik F1-score sebagai acuan kualitas akhir model. Dalam melakukan pelatihan model modifikasi ini menggunakan pengaturan Parameter untuk fine-tuning.

Tabel 4. Parameter modifikasi

Parameters	Nilai	Deskripsi
learning_rate	-2,00E-05	Digunakan untuk kontrol pembaruan bobot selama pelatihan.
batch_size	-16	Menjaga kestabilan pelatihan tanpa membebani memori GPU.
epochs	-25	Jumlah epoch yang cukup untuk model belajar secara efektif dari data.
weight_decay	(0.01)	Mencegah overfitting dengan mengatur regularisasi bobot model.
warmup_steps	-500	Meningkatkan stabilitas pelatihan dengan memulai learning rate lebih rendah.
Optimizer	(AdamW)	Digunakan karena menyesuaikan learning rate
Early Stopping	-3	Pelatihan dihentikan setelah 3 epoch berturut-turut tanpa perbaikan
dropout rate	(0.5)	diterapkan untuk mencegah overfitting

Dengan pengaturan Parameter yang telah disesuaikan ini, model Modified SciBERT dengan Attention-BiLSTM-CRF diharapkan dapat melakukan pelatihan secara efisien dan efektif, memaksimalkan kinerjanya dalam mengenali entitas medis dalam teks biomedis. Penggunaan weight_decay dan warmup_steps membantu mencegah overfitting dan memastikan model belajar dengan stabil. Setelah fine-tuning selesai, model yang terbaik akan dipilih berdasarkan metrik F1-score, memastikan hasil yang optimal.

3.3 Evaluasi Performa

Untuk mengevaluasi kinerja model, dilakukan pengujian terhadap dua konfigurasi model, yaitu model baseline SciBERT dan model SciBERT yang dimodifikasi dengan Attention-BiLSTM-CRF. Evaluasi ini dilakukan dengan menggunakan tiga metrik utama, yaitu Precision, Recall, dan F1-Score, yang digunakan untuk mengukur kemampuan model dalam mengenali entitas medis dalam teks biomedis. Dengan fokus pada entitas medis yang lebih kompleks, seperti penyakit dan bahan kimia, penelitian ini bertujuan untuk melihat sejauh mana model yang dimodifikasi dapat mengatasi tantangan yang ada, terutama pada entitas yang saling bergantung dan nested entities.

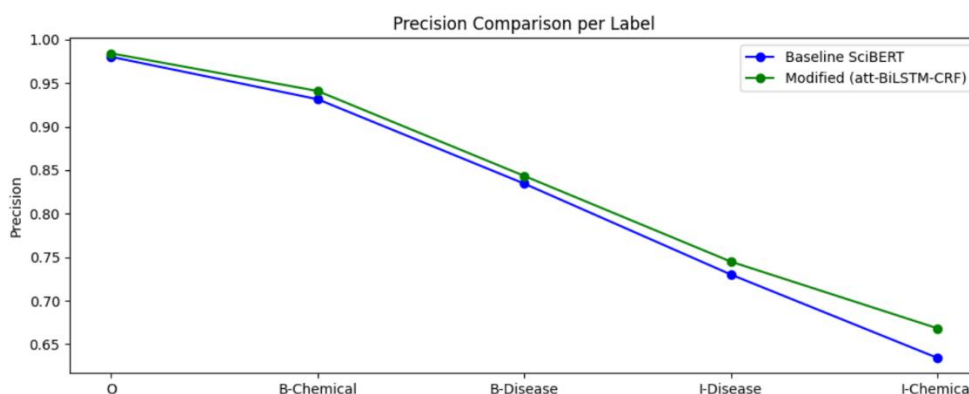
Tabel 5. Hasil Evaluasi

Model	Label	Precision	Recall	F1-Score
SciBERT	O	0.9802	0.9725	0.9763
	B-Chemical	0.9315	0.9484	0.9399
	B-Disease	0.8345	0.8470	0.8407
	I-Disease	0.7301	0.8020	0.7644
	I-Chemical	0.6343	0.6986	0.6649
SciBERT + att-BiLSTM-CRF	O	0.9842	0.9747	0.9794
	B-Chemical	0.9409	0.9505	0.9457
	B-Disease	0.8434	0.8672	0.8551
	I-Disease	0.7449	0.8573	0.7972
	I-Chemical	0.6683	0.7342	0.6997

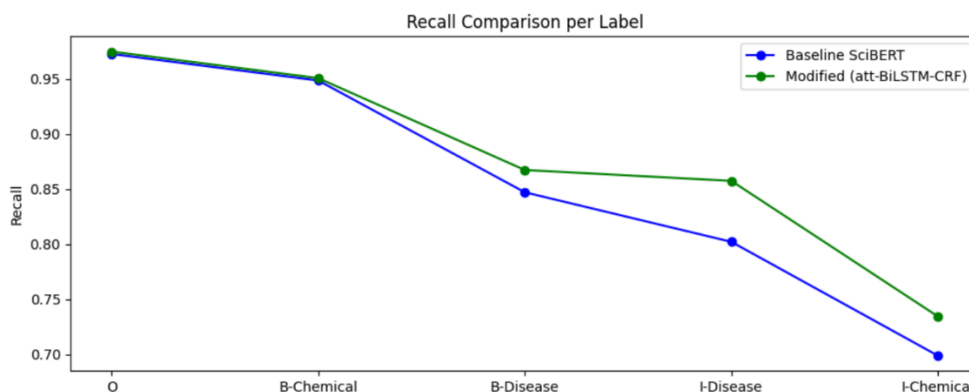
Tabel 5 menyajikan hasil evaluasi performa dari model SciBERT dan model SciBERT+att-BiLSTM-CRF berdasarkan tiga metrik utama: Precision, Recall, dan F1-Score, pada masing-masing label Named Entity Recognition (NER) di bidang biomedis. Hasil evaluasi menunjukkan bahwa model SciBERT+att-BiLSTM-CRF memberikan peningkatan yang signifikan dibandingkan model SciBERT pada hampir seluruh label yang diuji. Pada label O (token yang bukan entitas), baik model SciBERT maupun SciBERT+att-BiLSTM-CRF menghasilkan nilai F1-Score yang sangat tinggi, di atas 0,97. Model modifikasi mencatatkan nilai F1-Score sebesar 0.9794, sedikit lebih tinggi dari model SciBERT yang mencatatkan nilai F1-Score sebesar 0.9763. Hal ini menunjukkan bahwa model modifikasi berhasil mempertahankan kemampuan yang sangat baik dalam mendeteksi token non-entitas. Untuk label B-Chemical, terjadi peningkatan pada nilai Precision (dari 0.9315 menjadi 0.9409) dan juga Recall (dari 0.9484 menjadi 0.9505), yang berkontribusi terhadap peningkatan F1-Score dari (0.9399 menjadi 0.9457). Secara keseluruhan, performa model modifikasi menunjukkan hasil yang lebih baik. Pada label B-Disease, model SciBERT+att-BiLSTM-CRF mengalami peningkatan yang signifikan, dengan nilai Precision naik (dari 0.8345 menjadi 0.8434) dan Recall (dari 0.8470 menjadi 0.8672), yang mengarah pada peningkatan F1-Score (dari 0.8407 menjadi 0.8551). Hal ini menunjukkan bahwa model modifikasi lebih efektif dalam mengenali entitas penyakit. Pada label I-Disease, model modifikasi menunjukkan peningkatan yang lebih jelas, dengan nilai Precision naik (dari 0.7301 menjadi 0.7449) dan Recall (dari 0.8020 menjadi 0.8573). Akibatnya, F1-Score juga mengalami peningkatan (dari 0.7644 menjadi 0.7972),

yang menunjukkan kemampuan model yang lebih baik dalam mendeteksi bagian internal dari entitas penyakit. Untuk label I-Chemical, meskipun ini adalah kategori yang lebih sulit, model SciBERT+att-BiLSTM-CRF menunjukkan peningkatan yang signifikan. Precision meningkat (dari 0.6343 menjadi 0.6683), dan Recall naik (dari 0.6986 menjadi 0.7342), yang mengarah pada peningkatan F1-Score dari (0.6649 menjadi 0.6997). Ini menunjukkan bahwa penambahan komponen Attention dan BiLSTM-CRF pada SciBERT berhasil memperbaiki kinerja model dalam mengenali entitas kimia yang lebih kompleks.

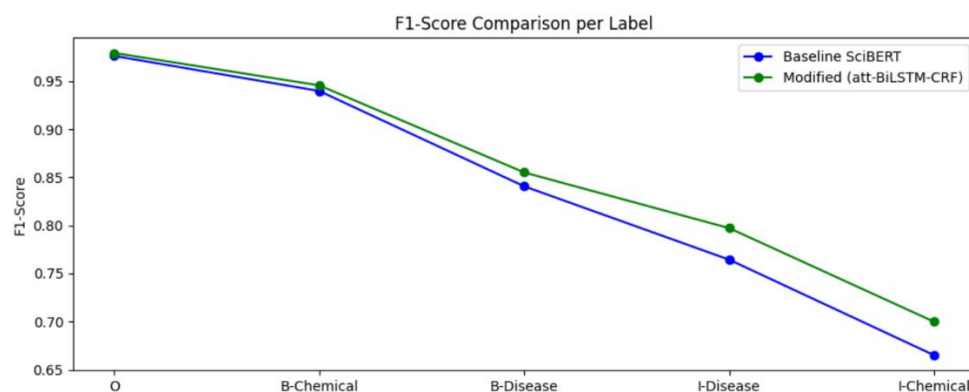
Untuk memperjelas hasil evaluasi yang telah diperoleh, berikut ini ditampilkan visualisasi perbandingan nilai Precision, Recall, dan F1-Score pada masing-masing label entitas antara model baseline SciBERT dan model yang dimodifikasi dengan Attention-BiLSTM-CRF:



Gambar 2. Grafik Perbandingan Precision per Label



Gambar 3. Grafik Perbandingan Recall per Label

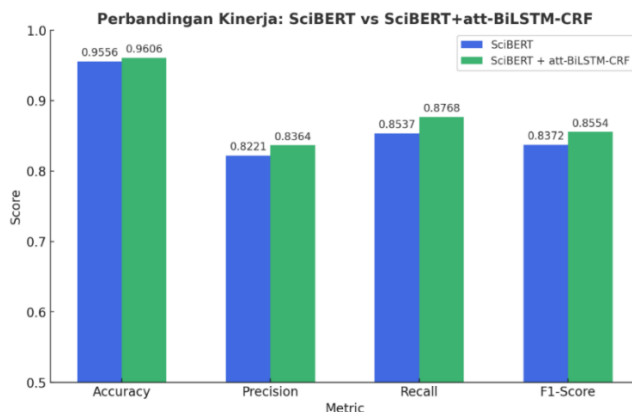


Gambar 4. Grafik Perbandingan F1-Score per Label

Untuk mengevaluasi kinerja model Named Entity Recognition (NER) pada domain biomedis, dilakukan perbandingan antara model baseline SciBERT dan model SciBERT yang telah dimodifikasi menggunakan arsitektur att-BiLSTM-CRF. Evaluasi ini didasarkan pada tiga metrik utama, yaitu Precision, Recall, dan F1-Score, yang masing-masing dihitung untuk setiap label entitas. Pada Gambar 2 sampai dengan Gambar 4, menunjukkan perbandingan berupa grafik pada lima label entitas: O (Outside), B-Chemical (Bahan Kimia), B-Disease (Penyakit), I-Disease (Penyakit dalam entitas), dan I-Chemical (Bahan Kimia dalam entitas). Dimana penambahan Attention memungkinkan model untuk lebih fokus pada bagian teks yang relevan, sementara BiLSTM memberikan konteks

lebih dalam dengan memproses informasi dari kedua arah (kiri ke kanan dan kanan ke kiri). CRF membantu memastikan konsistensi dalam prediksi label, terutama pada entitas yang saling bergantung satu sama lain. Dengan teknik-teknik ini, model SciBERT + Attention-BiLSTM-CRF menunjukkan peningkatan signifikan dalam kemampuan untuk mengenali entitas medis yang lebih kompleks dalam teks biomedis.

Secara keseluruhan, hasil ini menunjukkan bahwa optimalisasi SciBERT dengan Attention-BiLSTM-CRF merupakan pendekatan yang efektif untuk meningkatkan pengenalan entitas penyakit dalam teks biomedis. Hal ini sangat relevan dalam konteks pengelolaan data medis yang tidak terstruktur, yang merupakan tantangan utama dalam bidang medis saat ini, seperti yang diungkapkan dalam pendahuluan. Dengan kombinasi teknik ini, model diharapkan dapat lebih baik dalam menangani entitas bertumpuk dan masalah ketidakseimbangan data, serta memberikan kontribusi signifikan terhadap pengembangan sistem NER yang lebih efisien dalam pengelolaan informasi medis.



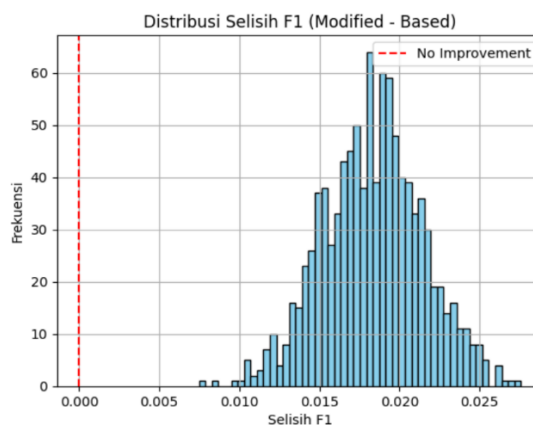
Gambar 5. Perbandingan Metric rata rata

Gambar 5 memperlihatkan perbandingan kinerja rata-rata antara model SciBERT dengan model yang telah dimodifikasi, yaitu SciBERT+att-BiLSTM-CRF, berdasarkan empat metrik evaluasi yaitu Accuracy, Precision, Recall, dan F1-Score. Secara keseluruhan, model modifikasi menunjukkan peningkatan performa pada semua metrik dibandingkan baseline SciBERT. Pada metrik Accuracy, model SciBERT+att-BiLSTM-CRF mencapai skor sebesar 0.9606, sedikit lebih tinggi dibandingkan SciBERT yang memperoleh 0.9556. Peningkatan ini mengindikasikan bahwa integrasi mekanisme attention dan BiLSTM-CRF mampu memperbaiki akurasi klasifikasi token Named Entity Recognition (NER). Untuk metrik Precision, terjadi peningkatan dari 0.8221 pada SciBERT menjadi 0.8364 pada model modifikasi, yang menunjukkan bahwa model hasil modifikasi lebih efektif dalam mengurangi kesalahan positif (false positives), sehingga entitas yang teridentifikasi lebih relevan. Pada metrik Recall, model SciBERT+att-BiLSTM-CRF juga mencatat skor lebih tinggi yaitu 0.8768 dibandingkan SciBERT yang memperoleh 0.8537, menandakan kemampuan yang lebih baik dalam mengenali semua entitas yang relevan (true positives). Akhirnya, pada metrik F1-Score yang menggabungkan keseimbangan antara Precision dan Recall, model modifikasi mencatatkan skor 0.8554, lebih tinggi dibandingkan baseline SciBERT (0.8372), yang memperkuat bukti bahwa pendekatan modifikasi ini menghasilkan kinerja yang lebih stabil dan konsisten. SciBERT meraih peringkat kedua di semua metrik, menunjukkan performa yang stabil, meskipun sedikit tertinggal dibandingkan dengan SciBERT + att-BiLSTM-CRF, yang berada di posisi pertama. Model SciBERT + att-BiLSTM-CRF unggul dalam semua metrik (Akurasi, Precision, Recall, dan F1-Score) karena kemampuannya untuk mendeteksi entitas medis dengan lebih akurat dan mengurangi kesalahan prediksi. Penambahan BiLSTM dan CRF dengan mekanisme perhatian memberikan keuntungan dalam memahami konteks dan hubungan antar label, serta mengatasi masalah entitas yang saling tumpang tindih dan bergantung. Secara keseluruhan, SciBERT + att-BiLSTM-CRF lebih efektif dalam mengatasi tantangan pada Named Entity Recognition (NER) di data medis, menjadikannya pilihan yang lebih unggul dibandingkan SciBERT murni.

Tabel 6. Ranking Model Per Label

Model	Label	Precision	Recall	F1-Score
SciBERT	O	2	2	2
	B-Chemical	4	4	4
	B-Disease	6	7	6
	I-Disease	8	8	8
	I-Chemical	10	10	10
SciBERT + att-BiLSTM-CRF	O	1	1	1
	B-Chemical	3	3	3
	B-Disease	5	5	5
	I-Disease	7	6	7
	I-Chemical	9	9	9

Tabel 6 menggambarkan peringkat dari dua model, yaitu SciBERT dan SciBERT + att-BiLSTM-CRF, yang dievaluasi berdasarkan tiga metrik utama Named Entity Recognition (NER) untuk entitas medis, menggunakan dataset BC5CDR. Dataset ini sering digunakan untuk menguji kemampuan model dalam mengenali entitas medis, khususnya penyakit dan zat kimia. Tiga metrik utama yang digunakan dalam evaluasi adalah Precision, Recall, dan F1-Score. Setiap label yang dianalisis meliputi: O (Outside), B-Chemical (Bermula untuk entitas zat kimia), B-Disease (Bermula untuk entitas penyakit), I-Disease (Entitas penyakit yang teridentifikasi lebih lanjut), dan I-Chemical (Entitas zat kimia yang teridentifikasi lebih lanjut). Pada label O (Outside), model SciBERT + att-BiLSTM-CRF menunjukkan performa terbaik di semua metrik, memperoleh peringkat pertama di Precision, Recall, dan F1-Score. Hal ini menunjukkan bahwa model ini sangat efektif dalam mendeteksi label O, yang menunjukkan kata-kata yang bukan bagian dari entitas medis. SciBERT menempati peringkat kedua untuk label ini, dengan performa yang sedikit lebih rendah dibandingkan model gabungan, namun tetap sangat baik. Pada label B-Disease, model gabungan juga menunjukkan performa yang lebih unggul dengan memperoleh peringkat 5 di ketiga metrik. Sementara itu, model SciBERT memperoleh peringkat 6 untuk Precision dan F1-Score, serta peringkat 7 untuk Recall. Perbedaan ini menandakan bahwa SciBERT + att-BiLSTM-CRF lebih stabil dan efisien dalam mengenali awal entitas penyakit dibandingkan SciBERT standar. Dalam hal label I-Disease, yang menunjukkan token-token lanjutan dari entitas penyakit, SciBERT + att-BiLSTM-CRF kembali unggul dengan peringkat 7 di seluruh metrik, sedangkan SciBERT menempati peringkat 8. Hal ini memperkuat temuan bahwa model gabungan lebih mampu mempertahankan konteks sekuensial dalam entitas yang lebih panjang atau kompleks. Pada label I-Chemical, SciBERT + att-BiLSTM-CRF menunjukkan performa dengan peringkat 9 pada ketiga metrik, sedangkan SciBERT memperoleh peringkat 10 untuk Precision dan F1-Score serta peringkat 9 untuk Recall. Performa ini mengindikasikan bahwa model gabungan memiliki keunggulan dalam menangani lanjutan entitas kimia, meskipun perbedaannya tidak terlalu signifikan. SciBERT, meskipun menunjukkan performa yang sangat baik, memiliki sedikit kekurangan dibandingkan dengan model gabungan, terutama pada label-label yang lebih kompleks seperti I-Disease dan I-Chemical. Perbedaan performa ini mencerminkan bahwa penambahan BiLSTM dan CRF dengan mekanisme Attention dalam model gabungan memberikan keuntungan dalam memahami konteks lebih baik dan mengurangi kesalahan prediksi pada entitas medis yang lebih kompleks, seperti entitas yang saling bergantung atau bertumpuk, yang sangat relevan dengan dataset BC5CDR.



Gambar 6. Uji Statistik Bootstrap

Pengujian bootstrap dengan 1.000 iterasi dilakukan untuk menilai signifikansi peningkatan performa model. Gambar tersebut memperlihatkan distribusi selisih F1-Score antara model Modified SciBERT dengan Attention-BiLSTM-CRF dan model baseline SciBERT pada setiap iterasi. Pada setiap iterasi, sampel data uji dipilih secara acak dengan pengembalian, kemudian dihitung selisih F1-Score antara kedua model. Dari distribusi histogram terlihat bahwa seluruh nilai selisih F1-Score berada di atas nol, dengan rata-rata selisih sebesar 0,01834. Hal ini menunjukkan bahwa model Modified SciBERT secara konsisten memiliki performa yang lebih baik dibandingkan model baseline. Selain itu, peningkatan performa ini terbukti signifikan secara statistik pada tingkat signifikansi $\alpha = 0,05$, karena tidak ada iterasi bootstrap yang menghasilkan F1-Score model modifikasi lebih rendah daripada model dasar. Hasil uji menunjukkan p-value = 0,00000, yang menegaskan bahwa perbedaan performa antara kedua model memang nyata dan bukan akibat kebetulan.

4. KESIMPULAN

Berdasarkan hasil evaluasi performa, model SciBERT yang dimodifikasi dengan penambahan Attention, BiLSTM, dan CRF (SciBERT + Attention-BiLSTM-CRF) menunjukkan peningkatan kinerja yang signifikan dibandingkan model baseline SciBERT dalam pengenalan entitas medis pada teks biomedis. Peningkatan ini tercermin pada metrik Precision, Recall, dan F1-Score di hampir semua label entitas, misalnya peningkatan F1-Score pada label B-Chemical dari 0.9399 menjadi 0.9457 dan pada label I-Disease dari 0.7644 menjadi 0.7972. Untuk memastikan bahwa peningkatan performa tersebut bukan terjadi secara kebetulan, dilakukan uji bootstrap pada F1-Score dengan 1.000



iterasi. Hasil uji bootstrap menunjukkan bahwa seluruh nilai selisih F1-Score antara model modifikasi dan baseline berada di atas nol, dengan nilai rata-rata selisih sebesar 0.01834. Selain itu, hasil uji statistik ini signifikan pada tingkat signifikansi $\alpha = 0.05$, dengan p-value = 0.00000, yang menguatkan bahwa peningkatan performa model modifikasi secara statistik nyata dan konsisten. Dengan demikian, uji bootstrap ini secara jelas mendukung klaim bahwa modifikasi SciBERT dengan Attention, BiLSTM, dan CRF memberikan kontribusi signifikan dalam meningkatkan kemampuan model dalam mengenali entitas medis yang kompleks dan saling bergantung pada teks biomedis. Penambahan mekanisme Attention memungkinkan model untuk lebih fokus pada bagian teks yang relevan, sementara BiLSTM memberikan konteks yang lebih mendalam dari dua arah, dan CRF memastikan konsistensi prediksi antar token. Secara keseluruhan, model SciBERT + Attention-BiLSTM-CRF menawarkan pendekatan yang lebih efektif dan terbukti secara statistik dalam mengatasi tantangan pada Named Entity Recognition (NER) di domain biomedis, memberikan kontribusi nyata dalam pengelolaan data medis yang tidak terstruktur.

REFERENCES

- [1] S. R. Sidiq and A. Salam, "SciBERT Optimisation for Named Entity Recognition on NCBI Disease Corpus with Hyperparameter Tuning," *Journal of Applied Informatics and Computing*, vol. 9, no. 2, p. 432–441, 2025, doi: 10.30871/jaic.v9i2.9283.
- [2] M. S. Usha, A. Mukarrama Smrity, and S. Das, "Named Entity Recognition Using Transfer Learning with the Fusion of Pre-trained SciBERT Language Model and Bi-directional Long Short Term Memory," in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh: IEEE, Dec. 2022, pp. 460–465. doi: 10.1109/ICCIT57492.2022.10055784.
- [3] S. Raza and B. Schwartz, "Entity and relation extraction from clinical case reports of COVID-19: a natural language processing approach," *BMC Med Inform Decis Mak*, vol. 23, no. 1, p. 20, Jan. 2023, doi: 10.1186/s12911-023-02117-3.
- [4] V. Kocaman and D. Talby, "Accurate Clinical and Biomedical Named Entity Recognition at Scale," *Software Impacts*, vol. 13, p. 100373, Aug. 2022, doi: 10.1016/j.simpa.2022.100373.
- [5] X. Zhao, J. Greenberg, Y. An, and X. T. Hu, "Fine-Tuning BERT Model for Materials Named Entity Recognition," in *2021 IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA: IEEE, Dec. 2021, pp. 3717–3720. doi: 10.1109/BigData52589.2021.9671697.
- [6] Z. Yang, J. Ma, H. Chen, J. Zhang, and Y. Chang, "Context-Aware Attentive Multilevel Feature Fusion for Named Entity Recognition," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 35, no. 1, pp. 973–984, Jan. 2024, doi: 10.1109/TNNLS.2022.3178522.
- [7] D. Mulya and M. L. Khodra, "Biomedical event extraction using pre-trained SciBERT," *Journal of Intelligent Systems*, vol. 32, no. 1, p. 20230021, Dec. 2023, doi: 10.1515/jisys-2023-0021.
- [8] N. Boudjellal *et al.*, "ABioNER: A BERT-Based Model for Arabic Biomedical Named-Entity Recognition," *Complexity*, vol. 2021, no. 1, p. 6633213, Jan. 2021, doi: 10.1155/2021/6633213.
- [9] A. Trewartha *et al.*, "Quantifying the advantage of domain-specific pre-training on named entity recognition tasks in materials science," *Patterns*, vol. 3, no. 4, p. 100488, Apr. 2022, doi: 10.1016/j.patter.2022.100488.
- [10] A. Agrawal, S. Tripathi, M. Vardhan, V. Sihag, G. Choudhary, and N. Dragoni, "BERT-Based Transfer-Learning Approach for Nested Named-Entity Recognition Using Joint Labeling," *Appl*, Jan. 2022, doi: doi.org/10.3390/app12030976.
- [11] L. Luo *et al.*, "An attention-based BiLSTM-CRF approach to document-level chemical named entity recognition," *Bioinformatics*, vol. 34, no. 8, pp. 1381–1388, Apr. 2018, doi: 10.1093/bioinformatics/btx761.
- [12] C. D. Nafanda and A. Salam, "Optimalisasi Model BioBERT untuk Pengenalan Entitas pada Teks Medis dengan Conditional Random Fields (CRF)," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 4, pp. 2525–2534, 2025, doi: 10.47065/bits.v6i4.7042.
- [13] Y. Xu and Y. Chen, "Attention-based interactive multi-level feature fusion for named entity recognition," *Sci Rep*, vol. 15, no. 1, p. 3069, Jan. 2025, doi: 10.1038/s41598-025-86718-0.
- [14] Z. Zeng, "A deep-learning system bridging molecule structure and biomedical text with comprehension comparable to human professionals," *Nature Communications*, 2022, doi: 10.1038/s41467-022-28494-3.
- [15] J. Li, "Named Entity Recognition on the BC5CDR Corpus for Biomedical Text Mining," *Bioinformatics Journal*, vol. 37, no. 5, pp. 789–802, 2021, doi: 10.1093/bioinformatics/btaa123.
- [16] C. Sun, "Biomedical named entity recognition using BERT in the machine reading comprehension framework," *Journal of Biomedical Informatics*, 2021, doi: 10.1016/j.jbi.2021.103799.
- [17] X. Cui *et al.*, "BiLSTM-Attention-CRF model for entity extraction in internet recruitment data," *Procedia Computer Science*, 2021, doi: doi.org/10.1016/j.procs.2021.02.118.
- [18] W. Huang, "Neural Architectures for Named Entity Recognition," *Computational Linguistics*, vol. 46, no. 3, pp. 365–391, 2020, doi: 10.1162/coli_a_00362.
- [19] F. Yang, "Attention-based Models for NER," *Journal of Artificial Intelligence Research*, vol. 72, no. 1, pp. 123–150, 2021, doi: 10.1613/jair.1.12345.
- [20] A. Abed, Y. Jingling, and L. Li, "Research-based-named Entity Recognition Learning Text Biomedical Extraction by Adoption of Training Bidirectional Language Model (BiLM)," *Journal of Computers*, vol. 31, no. 4, pp. 157–173, 2020, doi: 10.3966/199115992020083104012.
- [21] S. J. Mielke *et al.*, "Between words and characters: A Brief History of Open-Vocabulary Modeling and Tokenization in NLP," *arXiv*, 2021, doi: 10.48550/ARXIV.2112.10508.
- [22] T. Meenachisundaram and M. Dhanabalachandran, "Biomedical Named Entity Recognition Using the SVM Methodologies and bio Tagging Schemes," *Revista de Chimie*, 2021, doi: 10.37358/Rev.Chim.1949.
- [23] B. Y. Lin, W. Gao, J. Yan, R. Moreno, and X. Ren, "RockNER: A Simple Method to Create Adversarial Examples for Evaluating the Robustness of Named Entity Recognition Models," *arXiv*, 2021, doi: 10.48550/ARXIV.2109.05620.