

Fine-tuning Model BERT untuk Named Entity Recognition

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

Achmad Fariz Rizky Yanuar
NIM: 09021382126138

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2025

HALAMAN PENGESAHAN

SKRIPSI

Fine-tuning Model BERT Untuk Named Entity Recognition

Sebagai salah satu syarat untuk penyelesaian studi di
Program Studi S1 Teknik Informatika

Oleh:

ACHMAD FARIZ RIZKY YANUAR

09021382126138

Pembimbing I	: <u>Novi Yusliani, S.Kom., M.T.</u> NIP. 198211082012122001
Pembimbing II	: <u>Desty Rodiah, S.Kom., M.T.</u> NIP. 198912212020122011

Mengetahui

Ketua Jurusan Teknik Informatika



Hadipurnawan Satria, Ph.D
198004182020121001

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jumat tanggal 23 Mei 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Achmad Fariz Rizky Yanuar

NIM : 09021382126138

Judul : Fine-tuning Model BERT untuk Named Entity Recognition
dan dinyatakan LULUS.

1. Ketua Penguji

Hadipurnawan Satria, M.Sc., Ph.D.

NIP. 198004182020121001

2. Penguji I

Dr. Firdaus, S.T., M.Kom.

NIP. 197801212008121003

3. Pembimbing I

Novi Yusliani, S.Kom., M.T.

NIP. 198211082012122001

4. Pembimbing II

Desty Rodiah, S.Kom., M.T.

NIP. 198912212020122011

Mengetahui,

Ketua Jurusan Teknik Informatika

Hadipurnawan Satria, M.Sc., Ph.D.

NIP. 198004182020121001

HALAMAN PERNYATAAN

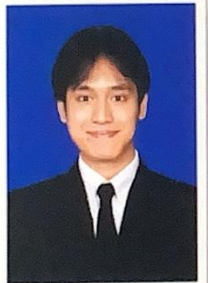
Yang bertanda tangan di bawah ini:

Nama : Achmad Fariz Rizky Yanuar
NIM : 09021382126138
Program Studi : Teknik Informatika
Judul : *Fine-tuning Model BERT untuk Named Entity Recognition*

Hasil Pengecekan Software iThenticate/Turnitin: 6%

Menyatakan bahwa laporan proyek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.



Palembang, 3 Juni 2025

Penulis,



Achmad Fariz Rizky Yanuar
NIM 09021382126138

MOTTO DAN PERSEMBAHAN

“WHAT A PRIVILEGE IT IS TO STUDY AND IMPROVE ONESELF”

Karya Tulis Ini Dipersembahkan Kepada

- Allah SWT
- Kedua Orang tua dan Keluargaku
- Teman-teman Seperjuangan
- Dosen Pembimbing
- Fakultas Ilmu Komputer
- Universitas Sriwijaya
- Dan diriku sendiri, Fariz

ABSTRACT

A sentence can contain various named entities with important meanings, such as names of people, locations, organizations, and time expressions. However, extracting this information manually requires significant time and resources. Named Entity Recognition (NER) offers an automated solution that improves the efficiency of this task. One method for developing an NER system is by using BERT, a transformer-based model that has proven effective in a wide range of natural language processing tasks. The BERT-base-cased model was fine-tuned using several hyperparameter configurations to identify the best combination based on the F1-score. The evaluation results showed that the best configuration was achieved with a learning rate of $5e-5$, a batch size of 32, and 2 epochs, yielding an F1-score of 0.836765.

Keywords: Named Entity Recognition, BERT, Fine-tuning, F1-Score, Natural Language Processing.

ABSTRAK

Sebuah kalimat dapat mengandung berbagai entitas bernama yang memiliki makna penting, seperti nama orang, lokasi, organisasi, dan waktu, namun proses ekstraksi informasi ini secara manual memerlukan waktu dan sumber daya yang besar. *Named Entity Recognition* (NER) menjadi solusi otomatis yang dapat meningkatkan efisiensi dalam tugas ini. Salah satu metode untuk mengembangkan sistem NER adalah dengan menggunakan BERT, model berbasis transformer yang telah terbukti efektif dalam berbagai tugas pemrosesan bahasa alami. Model BERT-base-cased di-*fine-tune* dengan berbagai konfigurasi *hyperparameter* untuk menemukan kombinasi terbaik berdasarkan *F1-score*. Hasil evaluasi menunjukkan bahwa konfigurasi terbaik diperoleh dengan *learning rate* $5e-5$, *batch size* 32, dan 2 *epoch*, menghasilkan *F1-score* sebesar 0.836765.

Kata Kunci: *Named Entity Recognition*, BERT, *Fine-tuning*, *F1-Score*, Pemrosesan Bahasa Alami.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “*Fine-tuning Model BERT untuk Named Entity Recognition*”, sebagai salah satu syarat untuk memperoleh gelar Sarjana (S1) pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Sriwijaya.

Proses penyusunan skripsi ini tidak terlepas dari berbagai tantangan dan hambatan. Namun, berkat kerja keras, doa, serta dukungan dan bantuan dari berbagai pihak, penulis dapat menyelesaikannya dengan baik dan tepat waktu. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT, yang telah memberikan kekuatan, kesabaran, dan kemudahan dalam proses penyusunan skripsi ini.
2. Orang tua, kakak, adik dan keluarga besar yang selalu memberikan dukungan, kasih sayang, dan doa tanpa henti.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. Selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Hadipurnawan Satria, Ph.D. Selaku Ketua Jurusan Teknik Informatika Universitas Sriwijaya.
5. Ibu Novi Yusliani. S.Kom. M.T. dan Ibu Desty Rodiah S.Kom. M.T. Selaku dosen pembimbing yang telah membimbing dengan sabar dan jelas serta memberikan saran yang membangun selama penyusunan skripsi ini.

6. Bapak Dr. M. Fachrurrozi, S.Si., M.T. Selaku dosen pembimbing akademik, yang telah membimbing akademik selama masa studi perkuliahan.
7. Seluruh dosen, staff, dan pegawai Fakultas Ilmu Komputer Universitas Sriwijaya yang telah banyak membantu selama perkuliahan. viii
8. Teman-teman Fariz (Mise, Diva, Rani, Bintang, Nabzah, Ais, Aul, Ufa, Bibil, Hakim, Nico, Rizki, Fariz, Rio, Brilliant, Felix, Mayori, Ariel, Kamilah, Miranty, dan Alpian) yang telah menemani serta mewarnai hari-hari penulis selama proses penyusunan skripsi.
9. Semua pihak yang telah memberikan dukungan, semangat, dan bantuan kepada penulis dalam menyelesaikan skripsi ini, yang tidak dapat disebutkan satu per satu..

Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dikarenakan keterbatasan pengetahuan dan pengalaman. Oleh karena itu, kritik dan saran yang diberikan sangat penulis harapkan untuk perbaikan kedepannya.

Palembang, 3 Juni 2025

Penulis,

Achmad Fariz Rizky Yanuar
NIM 09021382126138

DAFTAR ISI

HALAMAN PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF.....	iii
HALAMAN PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN.....	v
<i>ABSTRACT</i>	vi
ABSTRAK	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
I. BAB I PENDAHULUAN.....	1
1.1 Pendahuluan	1
1.2 Latar Belakang Masalah.....	1
1.3 Rumusan Masalah	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	4
1.6 Batasan Masalah.....	5
1.5 Sistematika Penulisan	6
1.6 Kesimpulan	7
II. BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 <i>Natural Language Processing</i> (NLP)	II-1
2.2.2 <i>Named Entity Recognition</i> (NER).....	II-1
2.2.3 Pre-Trained Language Model (PLM).....	II-3
2.2.4 <i>Bidirectional Encoder Representations from Transformers</i> (BERT) II-4	
2.2.5 <i>Confusion Matrix</i>	II-8
2.2.6 <i>Confidence Score</i>	II-10

2.2.7	<i>Rational Unified Process (RUP)</i>	II-11
2.2.8	Penelitian Lain yang Relevan.....	II-13
III.	BAB III METODOLOGI PENELITIAN	III-1
3.1	Pendahuluan	III-1
3.2	Pengumpulan Data	III-1
3.3	Tahapan Penelitian	III-5
3.3.1	Menentukan Kerangka Kerja Penelitian	III-6
3.3.2	Menentukan Kriteria Pengujian	III-8
3.3.3	Format Data Pengujian.....	III-8
3.3.4	Menentukan Alat Bantu Penelitian.....	III-11
3.3.5	Pengujian Penelitian.....	III-12
3.3.6	Analisis dan Menarik Kesimpulan Penelitian	III-13
3.4	Metode Pengembangan Perangkat Lunak.....	III-13
3.4.1	Fase Insepsi	III-13
3.4.2	Fase Elaborasi	III-14
3.4.3	Fase Konstruksi.....	III-14
IV.	BAB IV PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1	Pendahuluan	IV-1
4.2	Fase Insepsi	IV-1
4.2.1	Pemodelan Bisnis	IV-1
4.2.2	Kebutuhan Sistem	IV-2
4.2.3	Analisis dan Perancangan	IV-3
4.2.4	Implementasi	IV-5
4.3	Fase Elaborasi	IV-10
4.3.1	Pemodelan Bisnis	IV-11
4.3.2	Kebutuhan	IV-13
4.3.3	Analisis dan Perancangan	IV-14
4.4	Fase Konstruksi.....	IV-21
4.4.1	Kebutuhan Sistem	IV-21
4.4.2	Implementasi	IV-23
4.5	Fase Transisi.....	IV-28
4.5.1	Pemodelan Bisnis	IV-28

4.5.2	Kebutuhan	IV-28
4.5.3	Analisis dan Perancangan	IV-29
4.6	Kesimpulan	IV-32
V.	BAB V HASIL DAN ANALISIS	V-1
5.1	Pendahuluan	V-1
5.2	Hasil Penelitian	V-1
5.2.1	Konfigurasi Percobaan	V-1
5.2.2	Hasil Pengujian Konfigurasi I	V-3
5.2.3	Hasil Pengujian Konfigurasi II	V-6
5.2.4	Hasil Pengujian Konfigurasi III	V-9
5.2.5	Hasil Pengujian Konfigurasi IV	V-12
5.2.6	Hasil Pengujian Konfigurasi V	V-15
5.2.7	Hasil Pengujian Konfigurasi VI	V-18
5.2.8	Hasil Pengujian Konfigurasi VII	V-21
5.2.9	Hasil Pengujian Konfigurasi VIII	V-23
5.2.10	Hasil Pengujian Konfigurasi IX	V-26
5.2.11	Hasil Pengujian Konfigurasi X	V-28
5.2.12	Hasil Pengujian Konfigurasi XI	V-31
5.2.13	Hasil Pengujian Konfigurasi XII	V-33
5.2.14	Hasil Pengujian Konfigurasi XIII	V-36
5.2.15	Hasil Pengujian Konfigurasi XIV	V-38
5.2.16	Hasil Pengujian Konfigurasi XV	V-41
5.2.17	Hasil Pengujian Konfigurasi XVI	V-44
5.2.18	Hasil Pengujian Konfigurasi XVII	V-47
5.2.19	Hasil Pengujian Konfigurasi XVIII	V-50
5.3	Analisis Hasil Penelitian	V-53
5.4	Kesimpulan	V-56
VI.	BAB VI KESIMPULAN DAN SARAN.....	VI-1
6.1	Pendahuluan	VI-1
6.2	Kesimpulan	VI-1
6.3	Saran.....	VI-3
DAFTAR PUSTAKA.....		x

LAMPIRAN.....	xiii
----------------------	-------------

DAFTAR GAMBAR

Gambar II-1. Arsitektur model Transformer (Vaswani, 2017).....	II-3
Gambar II-2. <i>Overall</i> prosedur <i>pre-training</i> dan <i>fine-tuning</i> untuk BERT.....	II-5
Gambar II-3. Prosedur fine-tuning BERT untuk NER (Devlin et al., 2019)	II-6
Gambar II-4. Arsitektur RUP (Sudarma et al., 2021).....	II-12
Gambar III-1 Tahapan Penelitian	III-6
Gambar III-2 Kerangka Kerja Sistem	III-7
Gambar IV-1. Use Case Melatih Model NER.....	IV-6
Gambar IV-2. Use Case Mengekstraksi Entitas Teks Masukan Menjadi Pertanyaan.....	IV-6
Gambar IV-3. Rancangan Antarmuka Halaman Awal Sistem NER.....	IV-12
Gambar IV-4. Rancangan Antarmuka Halaman Hasil NER	IV-13
Gambar IV-5. Diagram Aktivitas Sistem Pelatihan Model NER.	IV-15
Gambar IV-6. Diagram Aktivitas Sistem NER untuk <i>Question Generation</i>	IV-16
Gambar IV-7. Diagram Alur Sistem Pelatihan Model NER.	IV-18
Gambar IV-8. Diagram Alur Sistem NER untuk <i>Question Generation</i>	IV-20
Gambar IV-9. Diagram Kelas Sistem Pelatihan Model NER.	IV-22
Gambar IV-10. Diagram Kelas Sistem NER untuk <i>Question Generation</i>	IV-23
Gambar IV-11. Implementasi Antarmuka Halaman Utama Sistem NER.	IV-26
Gambar IV-12. Implementasi Antarmuka Hasil Pemodelan Sistem NER.....	IV-27
Gambar V-1. Grafik Pengujian Konfigurasi I	V-4
Gambar V-2. <i>Confusion Matrix</i> Konfigurasi I	V-5
Gambar V-3. Grafik Pengujian Konfigurasi II.....	V-6
Gambar V-4. <i>Confusion Matrix</i> Konfigurasi II.....	V-8
Gambar V-5. Grafik Pengujian Konfigurasi III	V-9
Gambar V-6. <i>Confusion Matrix</i> Konfigurasi III.....	V-11
Gambar V-7. Grafik Pengujian Konfigurasi IV	V-12
Gambar V-8. <i>Confusion Matrix</i> Konfigurasi IV	V-14
Gambar V-9. Grafik Pengujian Konfigurasi V.....	V-15
Gambar V-10. <i>Confusion Matrix</i> Konfigurasi V	V-17
Gambar V-11. Grafik Pengujian Konfigurasi VI.....	V-19
Gambar V-12. <i>Confusion Matrix</i> Konfigurasi VI	V-20
Gambar V-13. Grafik Pengujian Konfigurasi VII	V-21
Gambar V-14. <i>Confusion Matrix</i> Konfigurasi VII	V-22
Gambar V-15. Grafik Pengujian Konfigurasi VIII.....	V-24
Gambar V-16. <i>Confusion Matrix</i> Konfigurasi VIII.....	V-25
Gambar V-17. Grafik Pengujian Konfigurasi IX	V-26
Gambar V-18. <i>Confusion Matrix</i> Konfigurasi IX	V-27
Gambar V-19. Grafik Pengujian Konfigurasi X.....	V-29
Gambar V-20. <i>Confusion Matrix</i> Konfigurasi X.....	V-30
Gambar V-21. Grafik Pengujian Konfigurasi XI	V-31
Gambar V-22. <i>Confusion Matrix</i> Konfigurasi XI	V-32

Gambar V-23. Grafik Pengujian Konfigurasi XII	V-34
Gambar V-24. <i>Confusion Matrix</i> Konfigurasi XII	V-35
Gambar V-25. Grafik Pengujian Konfigurasi XIII.....	V-36
Gambar V-26. <i>Confusion Matrix</i> Konfigurasi XIII.....	V-37
Gambar V-27. Grafik Pengujian Konfigurasi XIV	V-39
Gambar V-28. <i>Confusion Matrix</i> Konfigurasi XIV	V-40
Gambar V-29. Grafik Pengujian Konfigurasi XV	V-41
Gambar V-30. <i>Confusion Matrix</i> Konfigurasi XV	V-43
Gambar V-31. Grafik Pengujian Konfigurasi XVI	V-44
Gambar V-32. <i>Confusion Matrix</i> Konfigurasi XVI	V-46
Gambar V-33. Grafik Pengujian Konfigurasi XVII	V-47
Gambar V-34. <i>Confusion Matrix</i> Konfigurasi XVII	V-49
Gambar V-35. Grafik Pengujian Konfigurasi XVIII.....	V-50
Gambar V-36. <i>Confusion Matrix</i> Konfigurasi XVIII.....	V-52
Gambar V-37. Perbandingan Hasil Konfigurasi Pengujian.....	V-54

DAFTAR TABEL

Tabel II-1. Visualisasi <i>Confusion Matrix</i>	II-8
Tabel III-1. Sampel Data yang Digunakan.....	III-3
Tabel III-2 Rancangan Tabel Percobaan Hyperparameter	III-9
Tabel III-3 Tabel <i>Confusion Matrix</i>	III-10
Tabel III-4 Keterangan Entitas Bernama dari <i>Confusion Matrix</i>	III-11
Tabel III-5 Alat Bantu Penelitian	III-12
Tabel III-6 Tabel Hasil Evaluasi Model	III-13
Tabel III-7 Tabel Manajemen Proyek Penelitian.....	III-16
Tabel IV-1 Kebutuhan Fungsional	IV-2
Tabel IV-2 Kebutuhan Non-Fungsional	IV-2
Tabel IV-3. Tabel Definisi Aktor	IV-7
Tabel IV-4. Skenario Use Case Melatih Model NER.....	IV-7
Tabel IV-5. Skenario Use Case Mengekstraksi Entitas Teks Masukan Menjadi Pertanyaan	IV-9
Tabel V-1. Tabel Konfigurasi Percobaan.....	V-2
Tabel V-2. Tabel Hasil Evaluasi Model Konfigurasi I	V-6
Tabel V-3. Tabel Hasil Evaluasi Model Konfigurasi II.....	V-9
Tabel V-4. Tabel Hasil Evaluasi Model Konfigurasi III.....	V-12
Tabel V-5. Tabel Hasil Evaluasi Model Konfigurasi IV	V-15
Tabel V-6. Tabel Hasil Evaluasi Model Konfigurasi V.	V-18
Tabel V-7. Tabel Hasil Evaluasi Model Konfigurasi VI	V-21
Tabel V-8. Tabel Hasil Evaluasi Model Konfigurasi VII	V-23
Tabel V-9. Tabel Hasil Evaluasi Model Konfigurasi VIII.....	V-26
Tabel V-10. Tabel Hasil Evaluasi Model Konfigurasi IX	V-28
Tabel V-11. Tabel Hasil Evaluasi Model Konfigurasi X.....	V-31
Tabel V-12. Tabel Hasil Evaluasi Model Konfigurasi XI	V-33
Tabel V-13. Tabel Hasil Evaluasi Model Konfigurasi XII	V-36
Tabel V-14. Tabel Hasil Evaluasi Model Konfigurasi XIII.....	V-38
Tabel V-15. Tabel Hasil Evaluasi Model Konfigurasi XIV	V-41
Tabel V-16. Tabel Hasil Evaluasi Model Konfigurasi XV	V-44
Tabel V-17. Tabel Hasil Evaluasi Model Konfigurasi XVI	V-47
Tabel V-18. Tabel Hasil Evaluasi Model Konfigurasi XVII	V-50
Tabel V-19. Tabel Hasil Evaluasi Model Konfigurasi XVIII.....	V-53

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini akan menguraikan elemen-elemen utama yang menjadi landasan penelitian ini. Pembahasan akan mencakup beberapa aspek kunci, di antaranya latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, dan manfaat penelitian. Uraian mengenai aspek-aspek fundamental tersebut akan berfungsi sebagai kerangka dasar yang mengarahkan keseluruhan proses penelitian ini.

1.2 Latar Belakang Masalah

Sebuah kalimat dapat memiliki berbagai entitas bernama yang memiliki makna penting, seperti nama orang, lokasi, organisasi, waktu, dan lain sebagainya. Pemrosesan ekstraksi informasi-informasi ini secara manual dapat memakan waktu dan sumber daya. Informasi entitas yang terdapat dalam kalimat, teks, atau dokumen dapat digunakan untuk mengembangkan berbagai fitur kecerdasan buatan, terutama di bidang pemrosesan bahasa alami.

Named Entity Recognition (NER) adalah sebuah metode untuk mengidentifikasi, mengklasifikasikan, dan memisahkan entitas bernama ke dalam kelompok berdasarkan kategori yang telah ditentukan sebelumnya (Shishehgarkhaneh et al., 2024). Pengembangan sistem NER yang akurat menjadi kebutuhan utama mengingat perannya sebagai langkah awal dalam berbagai aplikasi pengolahan bahasa alami (Koroteev, 2021).

Pengembangan sistem NER dapat dilakukan dengan cara melakukan *fine-tuning* model *pre-trained*. Model *pre-trained* adalah model generik yang telah dilatih menggunakan dataset berskala besar untuk menyelesaikan tugas-tugas umum. *Bidirectional Encoder Representations from Transformers* (BERT), model *pre-trained* dengan performa yang baik pada berbagai tugas umum, dirancang dengan tujuan agar dapat dilakukan *fine-tuning* lebih lanjut untuk menyelesaikan tugas-tugas spesifik (M. L. Wang, 2024). Model ini pertama kali dikembangkan oleh peneliti dari Google pada tahun 2018. (Geetha & Karthika Renuka, 2021).

BERT juga dapat digunakan untuk tugas *Named Entity Recognition*. Model ini dirancang untuk mempelajari representasi teks secara dua arah (*bidirectional*), sehingga dapat menjadi dasar bagi berbagai model pembelajaran mesin di tahap selanjutnya (Koroteev, 2021). Salah satu contoh model yang dapat di-*fine-tuning* berdasarkan studi sebelumnya adalah BERT-base-cased. Sebagai model yang *case-sensitive*, BERT-base-cased menunjukkan performa yang lebih baik pada teks yang ditulis dengan baik (*well-written*) dan memberikan informasi tambahan dari penggunaan huruf kapital (Dao & Aizawa, 2023).

Fine-tuning untuk tugas spesifik dapat dilakukan dengan berbagai pendekatan, baik dengan menyesuaikan arsitektur model maupun dengan mengubah *hyperparameters*, seperti *learning rate* atau pemilihan algoritma optimasi terbaik (Sebbaq & Faddouli, 2022). Untuk *fine-tuning* BERT, dataset standar yang sering digunakan dalam NER adalah CoNLL2003. Popularitas dataset ini memungkinkan hasil penelitian dibandingkan dengan banyak studi sebelumnya yang juga menggunakannya. Dataset CoNLL2003 mencakup empat jenis entitas,

yaitu PER (*person*), LOC (*location*), ORG (*organization*), dan MISC (*miscellaneous*) (Dao & Aizawa, 2023).

Pada tahun 2021, sebuah penelitian tentang *fine-tuning* BERT base menggunakan dataset CoNLL2003 menunjukkan hasil yang baik. Penelitian oleh Li Bing (2021) mencatat bahwa meskipun hasil yang dicapai sudah baik, masih ada peluang untuk meningkatkan performa model melalui *hyperparameter-tuning* lebih lanjut, karena optimasi yang dilakukan pada tahap pertama (menggunakan CoNLL2003) belum sepenuhnya maksimal (Li, 2021).

Dataset lain yang dapat dipertimbangkan adalah dataset *Groningen Meaning Bank* (GMB). Dataset ini berisi 10.102 dokumen dengan lebih dari 1 juta token. Perbedaan terbesar dibandingkan dengan dataset CoNLL-2003 adalah ukurannya yang jauh lebih besar dan lebih baru. Selain itu, GMB juga mencakup lebih banyak tag NER, seperti TIM (*time indicators*), EVE (*events*), ART (*artifacts*), GPE (*geopolitical entities*), dan NAT (*natural objects*) (Van Lit, 2022).

Berdasarkan uraian di atas, penelitian ini berfokus pada penerapan *hyperparameter-tuning* pada model BERT base-cased dengan menggunakan dataset *Groningen Meaning Bank* (GMB). Dengan memanfaatkan dataset yang *well-written* dan terstruktur dengan baik seperti CoNLL2003, penelitian ini bertujuan untuk mengevaluasi kinerja performa model.

1.3 Rumusan Masalah

Berdasarkan latar belakang tersebut maka didapatlah rumusan masalah sebagai berikut:

1. Bagaimana mengembangkan model NER terbaik berdasarkan hasil dari *fine-tuning* model BERT-base?
2. Bagaimana kinerja model NER berdasarkan hasil *fine-tuning* model BERT-base menggunakan metrik *F1-score*?

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, maka terdapat tujuan:

1. Menghasilkan model NER terbaik berdasarkan hasil dari *fine-tuning* model BERT-base.
2. Mengetahui kinerja model NER berdasarkan hasil *fine-tuning* model BERT-base menggunakan metrik *F1-score*?

1.5 Manfaat Penelitian

1. Penelitian ini memberikan kontribusi pada pengembangan dan sistem NER khususnya yang memanfaatkan *fine tuning* model *pre-trained* BERT-base-cased.
2. Penelitian ini dapat digunakan sebagai referensi yang dapat digunakan dalam mengembangkan pendekatan NER dengan arsitektur dan metode lainnya.
3. Hasil dari penelitian ini digunakan sebagai subsistem dalam pengembangan sistem *question generation*, khususnya untuk tipe soal *fill in the blanks*, berdasarkan entitas yang telah diekstraksi oleh model NER.

1.6 Batasan Masalah

Adapun beberapa batasan masalah yang menjadi batas atau limitasi dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini difokuskan pada penggunaan dan pengembangan model sistem NER berbasis BERT-base-cased dengan melakukan *fine-tuning*, serta tidak mencakup penggunaan model bahasa lainnya.
2. Model sistem NER yang dikembangkan hanya difokuskan pada bahasa Inggris, sehingga hasil penelitian ini belum mencakup penerapan pada bahasa lain.
3. Evaluasi dilakukan pada dataset NER yang terbatas hanya pada bahasa Inggris dengan kategori entitas tertentu, seperti GEO (*geographical entity*), ORG (*organization*), PER (*person*), GPE (*geopolitical entities*), TIM (*time indicators*), ART (*artifacts*), EVE (*events*), dan NAT (*natural objects*).
4. Penelitian kinerja model akan dilakukan dengan menggunakan metrik standar seperti *accuracy*, *recall*, *precision*, dan *F1 score*.
5. Penelitian ini tidak menerapkan strategi khusus untuk mengatasi ketidakseimbangan distribusi entitas dalam dataset, seperti teknik *resampling*, penyesuaian bobot kelas, dsb. Oleh karena itu, model berpotensi menunjukkan bias terhadap entitas mayoritas, yang berdampak pada menurunnya kinerja dalam mendeteksi entitas dengan jumlah kemunculan yang lebih sedikit

1.5 Sistematika Penulisan

Sistematika penulisan tugas akhir mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yaitu sebagai berikut:

BAB I. PENDAHULUAN

Bab ini akan mengulas mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, serta rencana sistematika penelitian yang akan menjadi fokus utama dalam penelitian ini.

BAB II. KAJIAN LITERATUR

Bab ini membahas landasan teori yang digunakan dalam penelitian, seperti definisi metode dan arsitektur BERT-base-cased, *fine-tuning*, termasuk definisi *Named Entity Recognition* (NER), serta beberapa literatur yang relevan dengan penelitian ini.

BAB III. METODOLOGI PENELITIAN

Bab ini akan menjelaskan rangkaian proses yang ditempuh selama penelitian, meliputi tahap pengumpulan data, analisis data, serta perancangan perangkat lunak. Setiap tahap akan diuraikan berdasarkan kerangka kerja yang telah direncanakan.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini akan membahas tentang proses perancangan perangkat lunak yang akan dibangun, mulai dari analisis kebutuhan hingga tahap pengujian untuk menilai hasil pengembangan perangkat lunak. I-7

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini menyajikan hasil dan analisis dari penelitian yang telah dilaksanakan sesuai dengan langkah dan metode yang telah direncanakan sebelumnya. Analisis ini digunakan sebagai landasan untuk menarik kesimpulan dari penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Bab ini membahas kesimpulan yang diambil berdasarkan uraian dalam bab sebelumnya serta saran yang diberikan berdasarkan penelitian yang telah dilakukan.

1.6 Kesimpulan

Pada bab ini telah dijelaskan latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penelitian yang akan dijadikan sebagai pokok pikiran penelitian *Fine-tuning Model Bidirectional Encoder Representations from Transformers (BERT)* untuk *Named Entity Recognition (NER)*.

DAFTAR PUSTAKA

- Abilio, R., Coelho, G. P., & da Silva, A. E. A. (2024). Evaluating Named Entity Recognition: Comparative Analysis of Mono-and Multilingual Transformer Models on Brazilian Corporate Earnings Call Transcriptions. *ArXiv Preprint ArXiv:2403.12212*.
- Arkhipov, M., Trofimova, M., Kuratov, Y., & Sorokin, A. (2019). Tuning multilingual transformers for language-specific named entity recognition. *Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing*, 89–93. <https://aclanthology.org/W19-3712.pdf>
- Chen, W., Lin, F., Li, G., & Liu, B. (2024). A survey of automatic sarcasm detection: Fundamental theories, formulation, datasets, detection methods, and opportunities. *Neurocomputing*, 578, 127428. <https://doi.org/10.1016/j.neucom.2024.127428>
- Dao, T. A., & Aizawa, A. (2023). Evaluating the Effect of Letter Case on Named Entity Recognition Performance. *International Conference on Applications of Natural Language to Information Systems*, 588–598. https://link.springer.com/chapter/10.1007/978-3-031-35320-8_45
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NaacL-HLT, 1*, 2. <https://arxiv.org/pdf/1810.04805>
- Donnelly, L. F., Grzeszczuk, R., & Guimaraes, C. V. (2022). Use of natural language processing (NLP) in evaluation of radiology reports: an update on applications and technology advances. *Seminars in Ultrasound, CT and MRI*, 43(2), 176–181. <https://doi.org/10.1053/j.sult.2022.02.007>
- Geetha, M. P., & Karthika Renuka, D. (2021). Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. *International Journal of Intelligent Networks*, 2, 64–69. <https://doi.org/10.1016/j.ijin.2021.06.005>
- Gonsior, J. , F. C. , M. S. , R. A. , T. M. , L. W. (2022). To Softmax, or not to Softmax: that is the question when applying active learning for transformer models. *CoRR*. <https://fis.tu-dresden.de/portal/files/58461417/2210.03005v1.pdf>
- Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: an overview. *ArXiv Preprint ArXiv:2008.05756*. <https://arxiv.org/pdf/2008.05756>
- Koroteev, M. V. (2021). BERT: a review of applications in natural language processing and understanding. *ArXiv Preprint ArXiv:2103.11943*.

- Lauriola, I., Lavelli, A., & Aioli, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, 470, 443–456. <https://doi.org/10.1016/j.neucom.2021.05.103>
- Li, B. (2021). *Named Entity Recognition in the Style of Object Detection*. <http://arxiv.org/abs/2101.11122>
- Liu, S., & Ritter, A. (2022). Do CoNLL-2003 named entity taggers still work well in 2023? *ArXiv Preprint ArXiv:2212.09747*. <https://arxiv.org/pdf/2212.09747>
- Liu, Z., Jiang, F., Hu, Y., Shi, C., & Fung, P. (2021). NER-BERT: a pre-trained model for low-resource entity tagging. *ArXiv Preprint ArXiv:2112.00405*.
- Min, B., Ross, H., Sulem, E., Veyseh, A. P. Ben, Nguyen, T. H., Sainz, O., Agirre, E., Heintz, I., & Roth, D. (2023). Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Computing Surveys*, 56(2), 1–40. <https://dl.acm.org/doi/abs/10.1145/3605943>
- Sathyanarayanan, S., & Tantri, B. R. (2024). Confusion matrix-based performance evaluation metrics. *Afr. J. Biomed. Res*, 27(4), 4023–4031. https://www.researchgate.net/profile/Sathyanarayanan-Swaminathan/publication/386347454_Confusion_Matrix-Based_Performance_Evaluation_Metrics/links/674e9e453d17281c7df6b440/Confusion-Matrix-Based-Performance-Evaluation-Metrics.pdf
- Sebbaq, H., & Faddouli, N. E. El. (2022). Fine-Tuned BERT Model for Large Scale and Cognitive Classification of MOOCs. *International Review of Research in Open and Distributed Learning*, 23(2), 170–190. <https://doi.org/10.19173/irrodl.v23i2.6023>
- Shishehgharkhaneh, M. B., Moehler, R. C., Fang, Y., Hijazi, A. A., & Aboutorab, H. (2024). Transformer-Based Named Entity Recognition in Construction Supply Chain Risk Management in Australia. *IEEE Access*, 12, 41829–41851. <https://doi.org/10.1109/ACCESS.2024.3377232>
- Siregar, S. W., & Putri, R. A. (2024). Information System for Data Collection on Building Tools and Materials Using the Rational Unified Process (RUP) Method. *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, 12(2), 455–466. <http://jurnal.unismabekasi.ac.id/index.php/piksel/article/view/9951/3318>
- Situmeang, S. (2022). Impact of text preprocessing on named entity recognition based on conditional random field in Indonesian text. *Jurnal Mantik*, 6(1), 423–430. <https://iocscience.org/ejournal/index.php/mantik/article/view/2297>
- Sudarma, M., Ariyani, S., & Wicaksana, P. A. (2021). Implementation of the rational unified process (RUP) model in design planning of sales order

- management system. *INTENSIF: Jurnal Ilmiah Penelitian Dan Penerapan Teknologi Sistem Informasi*, 5(2), 249–265.
<https://ojs.unpkediri.ac.id/index.php/intensif/article/view/15543/2191>
- Van Lit, W. (2022). *Conspiracy Tweet Classification with Named Entities*.
<https://arno.uvt.nl/show.cgi?fid=161213>
- Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.1706.03762>
- Wang, H., Li, J., Wu, H., Hovy, E., & Sun, Y. (2023). Pre-trained language models and their applications. *Engineering*, 25, 51–65.
<https://doi.org/10.1016/j.eng.2022.04.024>
- Wang, M. L. (2024). *Fine-Tuning BERT for Sentiment Analysis*.
<https://escholarship.org/uc/item/5kr0p5m7>
- Wang, S., Sun, X., Li, X., Ouyang, R., Wu, F., Zhang, T., Li, J., & Wang, G. (2023). Gpt-ner: Named entity recognition via large language models. *ArXiv Preprint ArXiv:2304.10428*. <https://arxiv.org/pdf/2304.10428>
- Zhang, T., Wu, F., Katiyar, A., Weinberger, K. Q., & Artzi, Y. (2020). Revisiting few-sample BERT fine-tuning. *ArXiv Preprint ArXiv:2006.05987*.
<https://arxiv.org/pdf/2006.05987>