

**PENERAPAN METODE M2M100 DENGAN
PARAMETER-EFFICIENT FINE-TUNING PADA
TRANSLITERASI TEKS ARAB KE TEKS LATIN INDONESIA**

Skripsi

Diajukan Untuk Memenuhi Salah Satu Syarat
Sarjana Strata-1 Teknologi Informasi



OLEH:

SITI AINUR RAHMAWATI

NIM. 2210817220029

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, APRIL 2026**

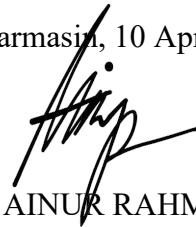
PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Siti Ainur Rahmawati
NIM : 2210817220029
Fakultas : Teknik
Prodi : Teknologi Informasi
Judul : Penerapan Metode M2M100 dengan *Parameter-Efficient Fine-tuning* pada Transliterasi Teks Arab ke Teks Latin Indonesia
Pembimbing Utama : Andreyan Rizky Baskara, S.Kom., M.Kom.

Dengan ini saya menyatakan bahwa dalam Skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya, juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.

Banjarmasin, 10 April 2026



SITI AINUR RAHMAWATI

NIM. 2210817220029

LEMBAR PENGESAHAN

SKRIPSI PROGRAM STUDI S-1 TEKNOLOGI INFORMASI

Penerapan Metode M2M100 dengan *Parameter-Efficient Fine-tuning* pada

Transliterasi Teks Arab ke Teks Latin Indonesia

Oleh

Siti Ainur Rahmawati (2210817220029)

Telah dipertahankan di depan Tim Penguji pada 28 April 2026 dan dinyatakan

LULUS

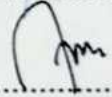
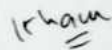
Komite Penguji :

Ketua : Ir. Eka Setya Wijaya, S.T., M.Kom.
NIP. 198205082008011010

Anggota 1 : Irham Maulani Abdul Gani, S.Kom., M.Kom.
NIP. 199710312025061009

Anggota 2 : Dr. Ir. Yuslena Sari, S.Kom., M.Kom
NIP. 198411202015042002

Pembimbing : Andreyan Rizky Baskara, S.Kom., M.Kom.
Utama NIP. 199307032019031011



Banjarbaru, 13 MAY 2026

Diketahui dan disahkan oleh:

Wakil Dekan Bidang Akademik
Fakultas Teknik ULM,


Dr. Mahmud, S.T., M.T.
NIP. 197401071998021001

Koordinator Program Studi
S-1 Teknologi Informasi,


Andreyan Rizky Baskara, S.Kom., M.Kom.
NIP. 199307032019031011



ABSTRAK

Transliterasi Arab–Latin merupakan solusi penting untuk meningkatkan aksesibilitas masyarakat Indonesia terhadap teks Arab. Namun, standar transliterasi berdasarkan Surat Keputusan Bersama Menteri Agama dan Menteri Pendidikan dan Kebudayaan RI Nomor 158 Tahun 1987 belum diterapkan secara luas pada aplikasi digital, sementara sistem transliterasi otomatis yang ada masih memiliki kualitas transliterasi rendah dan terbatasnya data pelatihan. Penelitian ini bertujuan mengembangkan sistem transliterasi Arab–Latin yang akurat dan efisien sesuai standar SKB 1987 dengan menerapkan metode *Parameter-Efficient Fine-tuning* pada model M2M100 418M, model terjemahan multibahasa. Dua teknik PEFT yaitu LoRA dan Adapter dievaluasi secara komparatif pada tiga skenario pembagian data menggunakan dataset paralel Al-Qur'an berjumlah 6.052 pasang kalimat. Kualitas transliterasi dinilai menggunakan metrik BLEU, CharBLEU, dan ChrF, sedangkan efisiensi komputasi diukur dari ukuran model dan waktu inferensi. Hasil pengujian menunjukkan bahwa metode Adapter dengan bottleneck 128 mencapai skor BLEU tertinggi 64,48 pada skenario split 90:5:5, melampaui kinerja model *baseline* M2M100 418M maupun model M2M100 1.2B yang lebih besar. Sementara itu, LoRA dengan rank 64 menghasilkan model paling ringkas berukuran sekitar 1.036 MB, menyusut sekitar 44 persen dari *baseline*, dan menunjukkan efisiensi komputasi terbaik. Temuan ini menegaskan bahwa metode PEFT, khususnya LoRA, mampu menyeimbangkan kualitas transliterasi dan efisiensi komputasi secara optimal, sehingga lebih berpotensi diimplementasikan pada perangkat bergerak dan aplikasi mobile dibandingkan metode lain yang diuji.

Kata kunci: Adapter, LoRA, M2M100, *Parameter-Efficient Fine-tuning*, Transliterasi Arab-Latin

ABSTRACT

Arabic–Latin transliteration plays a crucial role in improving Indonesian public access to Arabic texts. However, the transliteration standard established by the Joint Decree of the Indonesian Ministry of Religious Affairs and Ministry of Education and Culture No. 158 of 1987 has not been widely adopted in digital applications, while existing automatic transliteration systems still suffer from low transliteration quality and limited training data. This study aims to develop an accurate and efficient Arabic–Latin transliteration system compliant with the 1987 standard by applying Parameter-Efficient Fine-tuning methods to the M2M100 418M multilingual Machine Translation model. Two PEFT techniques, LoRA and Adapter, were evaluated comparatively across three data split scenarios using a parallel Quranic dataset of 6,052 sentence pairs. Transliteration quality was assessed using BLEU, Char-BLEU, and ChrF metrics, while computational efficiency was measured by model size and inference time. The Adapter method with a bottleneck size of 128 achieved the highest transliteration quality with a BLEU score of 64.48 on the 90:5:5 data split, surpassing both the M2M100 418M baseline and the larger M2M100 1.2B model. Meanwhile, LoRA with rank 64 produced the most compact model at approximately 1,037 MB, representing a 44% reduction from the baseline, and achieved the highest preference score in the TOPSIS multi-criteria analysis. The findings demonstrate that PEFT methods, particularly LoRA, offer an optimal balance between transliteration quality and computational efficiency, making them more suitable for deployment on mobile devices compared to other methods examined in this study.

Keywords: Adapter, LoRA, M2M100, Parameter-Efficient Fine-tuning, Arabic–Latin transliteration

KATA PENGANTAR

Puji dan syukur senantiasa kami panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat, kesehatan, dan kekuatan, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Metode M2M100 dengan *Parameter-Efficient Fine-tuning* pada Transliterasi Teks Arab ke Teks Latin Indonesia” ini dengan baik. Shalawat serta salam semoga tercurah kepada junjungan kita, Nabi Besar Muhammad SAW, yang telah membawa kita dari zaman kegelapan menuju cahaya keilmuan.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata-1 (S1) pada Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Lambung Mangkurat, Banjarmasin. Penulis menyadari bahwa penyelesaian skripsi ini tidak lepas dari dukungan dan bantuan berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penulis menyampaikan penghargaan kepada:

1. Bapak Dekan Fakultas Teknik Universitas Lambung Mangkurat, yang telah memberikan izin dan dukungan dalam proses penyusunan skripsi ini.
2. Bapak Andreyan Rizky Baskara, S.Kom., M.Kom., selaku Koordinator Program Studi Teknologi Informasi sekaligus Pembimbing Utama, yang telah memberikan bimbingan, arahan, serta motivasi dengan penuh kesabaran selama penyusunan tugas akhir ini.
3. Orang tua dan keluarga, yang selalu memberikan doa, dukungan moral, dan materiil tanpa henti.
4. Teman-teman seperjuangan di Program Studi Teknologi Informasi, yang telah menjadi tempat berbagi ilmu dan pengalaman selama masa perkuliahan.
5. Seluruh pihak yang telah membantu dan memberikan dukungan dalam penyelesaian skripsi ini, yang tidak dapat penulis sebutkan satu per satu.

Skripsi ini terdiri dari beberapa bab dengan sistematika penulisan sebagai berikut:

- BAB I: Pendahuluan, yang mencakup latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, dan manfaat penelitian.
- BAB II: Tinjauan Pustaka, berisi landasan teori, penelitian terkait, dan

kerangka pemikiran.

- BAB III: Metodologi Penelitian, meliputi alat dan bahan, lokasi penelitian, alur penelitian.
- BAB IV: Hasil dan Pembahasan, yang memaparkan hasil penelitian serta analisisnya.
- BAB V: Kesimpulan dan Saran, yang berisi kesimpulan dari hasil penelitian serta saran untuk pengembangan lebih lanjut.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun dari berbagai pihak sangat penulis harapkan untuk perbaikan di masa mendatang. Semoga skripsi ini dapat bermanfaat bagi pengembangan ilmu pengetahuan dan menjadi referensi bagi para pembaca.

Banjarmasin, 10 April 2026

Penulis,

SITI AINUR RAHMAWATI

DAFTAR ISI

PERNYATAAN.....	i
LEMBAR PENGESAHAN	ii
PERSETUJUAN MAJU SIDANG SKRIPSI	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN.....	xv
DAFTAR RIWAYAT HIDUP.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Penelitian Terkait	7
2.1.1 <i>Phrase Based and Neural Network Translation for Text Transliteration from Arabic to Indonesian</i>	7
2.1.2 <i>Comparative Analysis of Neural Translation Models based on Transformers Architecture</i>	7
2.1.3 <i>Sinhala Transliteration: A Comparative Analysis Between Rule-based</i>	

<i>and Seq2Seq Approaches</i>	8
2.1.4 <i>Low-Resource Transliteration for Roman-Urdu and Urdu Using Transformer-Based Models</i>	9
2.1.5 <i>An Effective Machine Learning Approach for English-Arabic Transliteration</i>	9
2.1.6 Penerapan <i>Recurrent Neural Network</i> (RNN) Pada Transliterasi Teks Arab ke Teks Latin Indonesia berbasis <i>Optical Character Recognition</i> (OCR)	10
2.1.7 <i>Automatic Romanization of Arabic Bibliographic Records</i>	10
2.1.8 <i>When does Parameter-Efficient Transfer Learning Work for Machine Translation?</i>	11
2.1.9 <i>Toward Low-Resource Languages Machine Translation: A Language-Specific Fine-tuning With LoRA for Specialized Large Language Models</i> ...	11
2.2 Landasan Teori	16
2.2.1 Karakteristik Bahasa Arab	16
2.2.2 Karakteristik Ortografi Arab	18
2.2.3 Transliterasi Arab-Latin	20
2.2.4 M2M100	25
2.2.5 <i>Parameter-Efficient Fine-tuning</i> (PEFT)	27
2.2.6 BLEU	29
2.2.7 CharBLEU	30
2.2.8 ChrF	31
2.2.9 <i>Technique for Order of Preference by Similarity to Ideal Solution</i>	

(TOPSIS)	31
2.3 Kerangka Pemikiran.....	34
BAB III METODOLOGI PENELITIAN.....	36
3.1 Alat dan Bahan.....	36
3.1.1 Alat Penelitian.....	36
3.1.2 Bahan Penelitian.....	36
3.2 Alur Penelitian	37
3.2.1 Identifikasi Masalah	37
3.2.2 Studi Literatur	37
3.2.3 Pengumpulan Data	37
3.2.4 Akuisisi Data.....	38
3.2.5 Eksperimen Model	38
3.2.6 Analisis Hasil	43
BAB IV HASIL DAN PEMBAHASAN	45
4.1 Pengumpulan Data	45
4.2 Akuisisi Data.....	45
4.3 Konfigurasi Eksperimen	46
4.3.1 Pembagian Dataset.....	46
4.3.2 <i>Preprocessing</i> Dataset.....	47
4.3.3 Model dan Strategi <i>Fine-tuning</i>	47
4.3.4 Konfigurasi Hyperparameter.....	49
4.3.5 Infrastruktur dan Lingkungan Eksperimen	50

4.4	Evaluasi Hasil Pengujian Model	51
4.5	Analisis Hasil	56
4.5.1	Interpretasi Kualitas Transliterasi Berdasarkan Metrik	56
4.5.2	Analisis Efisiensi dan <i>Trade-off</i>	57
4.5.3	Analisis Kualitatif dan Pola Kesalahan Karakter	58
4.6	Pemilihan Model terbaik	64
4.7	Implementasi Model pada Sistem	66
BAB V KESIMPULAN DAN SARAN.....		70
5.1	Kesimpulan	70
5.2	Saran.....	71
DAFTAR PUSTAKA		72
LAMPIRAN.....		83

DAFTAR GAMBAR

Gambar 2.1. Ilustrasi Arsitektur Encoder-Decoder Transformer.....	26
Gambar 2.2. Ilustrasi LoRA	27
Gambar 2.3. Ilustrasi Adapter	28
Gambar 2.4. Kerangka Pemikiran.....	34
Gambar 3.1. Alur Penelitian.....	37
Gambar 3.2. Alur Eksperimen Model	39
Gambar 4.1. Tampilan Dataset.....	46
Gambar 4.2. Struktur Direktori Model Hasil Penyimpanan LoRA	58
Gambar 4.3. Struktur Direktori Model Hasil Penyimpanan Adapter.....	58
Gambar 4.4. <i>Boxplot</i> Panjang Teks Berdasarkan Jumlah Model yang Benar (<i>Correct Count</i>).....	60
Gambar 4.5. Tampilan Input Teks Arab	67
Gambar 4.6. Tampilan Hasil Transliterasi	67
Gambar 4.7. Tampilan Input Teks Arab	68
Gambar 4.8. Tampilan Hasil OCR.....	68
Gambar 4.9. Tampilan Hasil Transliterasi dari Hasil OCR	68

DAFTAR TABEL

Tabel 2.1. Perbandingan Hasil Transliterasi Sinhala	8
Tabel 2.2. Ilustrasi Transliterasi Urdu ke Roman-Urdu.....	9
Tabel 2.3. Penelitian Terkait	13
Tabel 2.4. Perbandingan Praktik Pembacaan Teks Arab Umum dan Al-Qur'an..	17
Tabel 2.5 Perbedaan Penulisan Pada Rasm Utsmani dan Imla'i	19
Tabel 2.6. Huruf Konsonan Pada Bahasa Arab.....	20
Tabel 2.7. Huruf Vokal Pada Bahasa Arab	22
Tabel 2.8. Huruf Vokal Panjang Pada Bahasa Arab	22
Tabel 2.9. Perbandingan Transliterasi Al-Qur'an Kemenag dan Transliterasi SKB 1987.....	24
Tabel 3.1. Alat Penelitian.....	36
Tabel 3.2. Bahan Penelitian	36
Tabel 3.3. Sampel Dataset Transliterasi.....	38
Tabel 3.4. Skenario Pembagian Dataset.....	40
Tabel 3.5. Ilustrasi Tokenisasi pada Model M2M100	40
Tabel 3.6. Ilustrasi Padding pada Batch Sekuens	41
Tabel 3.7. Interpretasi Skor BLEU	43
Tabel 3.8. Interpretasi Skor ChrF.....	44
Tabel 4.1. Hasil Pengumpulan Data.....	45
Tabel 4.2. Ukuran Pembagian Dataset.....	47
Tabel 4.3. Konfigurasi Hyperparameter.....	49
Tabel 4.4. Variasi Rank LoRA.....	49
Tabel 4.5. Variasi Bottleneck Adapter.....	50
Tabel 4.6. Library dan Dependencies	50

Tabel 4.7. Hasil Pengujian pada Split Data 70:15:15	52
Tabel 4.8. Hasil Pengujian pada Split Data 75:15:10	53
Tabel 4.9. Hasil Pengujian pada Split Data 80:10:10	54
Tabel 4.10. Hasil Pengujian pada Split Data 90:5:5	55
Tabel 4.11. Jumlah Prediksi Benar dan Salah pada Model Representatif	57
Tabel 4.12. Distribusi Data Uji Berdasarkan Jumlah Model yang Berhasil Memprediksi dengan Benar	59
Tabel 4.13. Perbandingan Hasil Transliterasi Metode Tanpa dan Dengan Harakat	59
Tabel 4.14. Perbandingan Kesalahan Hasil Transliterasi.....	61
Tabel 4.15. Perbandingan Hasil Transliterasi yang Benar	62
Tabel 4.16. Bobot Kriteria Model Terbaik.....	64
Tabel 4.17. Hasil Perhitungan TOPSIS.....	65
Tabel 4.18. Hasil Pengujian Aplikasi pada Teks Arab Tanpa Harakat.....	69

DAFTAR LAMPIRAN

Lampiran 1. Tinjauan Transliterasi Pada Aplikasi Al-Qur'an Digital	83
Lampiran 2. Tinjauan Transliterasi Pada Aplikasi Terjemahan.....	84
Lampiran 3. Perbandingan Performa Penelitian Sebelumnya antara <i>Full Fine-tuning</i> dan <i>Parameter-Efficient Fine-tuning</i>	85
Lampiran 4. Kode Pelatihan Model M2M100	88
Lampiran 5. Kode Pelatihan Model M2M100 dengan Adapter.....	91
Lampiran 6. Kode Pelatihan Model M2M100 dengan LoRA.....	96
Lampiran 7. Kode Aplikasi Transliterasi	99
Lampiran 8. Lembar Konsultasi.....	103