

**DESAIN SMALL LANGUAGE MODEL DALAM
PENYUSUNAN SOAL MULTIPLE CHOICE DI DESA PETE
TIGARAKSA**



LAPORAN MBKM PENELITIAN

**Josua Rivo Muin
00000072776**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**DESAIN SMALL LANGUAGE MODEL DALAM
PENYUSUNAN SOAL MULTIPLE CHOICE DI DESA PETE
TIGARAKSA**



LAPORAN MBKM PENELITIAN

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S.Kom.)

Josua Rivo Muin

00000072776

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Josua Rivo Muin

Nomor Induk Mahasiswa : 00000072776

Program studi : Sistem Informasi

Laporan MBKM Penelitian dengan judul:

DESAIN SMALL LANGUAGE MODEL DALAM PENYUSUNAN SOAL MULTIPLE CHOICE DI DESA PETE TIGARAKSA

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan MBKM, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk laporan MBKM yang telah saya tempuh.

Tangerang, 19 Desember 2025



Josua Rivo Muin

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Josua Rivo Muin
Nomor Induk Mahasiswa : 00000072776
Program Studi : Sistem Informasi
Jenjang : Sarjana
Jenis Karya : Laporan MBKM Penelitian
Judul Karya Ilmiah : DESAIN SMALL LANGUAGE MODEL DALAM
PENYUSUNAN SOAL MULTIPLE CHOICE DI
DESA PETE TIGARAKSA

Menyatakan dengan sesungguhnya bahwa saya bersedia* (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 06 Januari 2026



Josua Rivo Muin

* Pilih salah satu opsi tanpa harus menghapus opsi yang tidak dipilih

** Jika tidak bisa membuktikan LoA jurnal/ HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN

KATA PENGANTAR

Saya mengucapkan syukur atas terselesaikannya penulisan penelitian berjudul: “DESAIN SMALL LANGUAGE MODEL DALAM PENYUSUNAN SOAL MULTIPLE CHOICE DI DESA PETE TIGARAKSA”. Penelitian ini saya susun sebagai salah satu persyaratan untuk memperoleh gelar Sarjana Sistem Informasi di Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Saya menyadari bahwa proses perkuliahan hingga penyusunan penelitian ini tidak akan dapat saya selesaikan tanpa bantuan, dukungan, dan bimbingan dari banyak pihak. Karena itu, dengan sepenuh hati saya ingin menyampaikan ucapan terima kasih kepada:

1. Bapak Ir. Andrey Andoko, M.Sc., Ph.D., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Bapak Johan Setiawan. S.Kom., M.M., MBA, sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Bapak Ahmad Faza, S.Kom., M.T.I., sebagai Pembimbing kedua yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
6. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
7. Teman-teman terdekat yang senantiasa memberikan semangat, bantuan, dan kebersamaan selama proses penyusunan laporan.
8. Seseorang yang pernah hadir dan memberikan dukungan serta pelajaran berharga dalam perjalanan hidup saya.

Saya berharap hasil penelitian dalam laporan ini dapat menjadi sumber inspirasi serta memberikan manfaat yang berarti bagi pembaca, lembaga pendidikan, maupun dunia industri. Semoga karya ini mampu mendorong lahirnya ide-ide baru dan inovasi yang dapat berkontribusi pada perkembangan teknologi *Small Language Model*, penelitian selanjutnya, serta berbagai bidang lain yang bermanfaat bagi masyarakat dan nilai-nilai yang kita junjung bersama.

Tangerang, 18 Desember 2025



Josua Rivo Muin



DESAIN SMALL LANGUAGE MODEL DALAM PENYUSUNAN SOAL MULTIPLE CHOICE DI DESA PETE TIGARAKSA

Josua Rivo Muin

ABSTRAK

Penelitian ini mengembangkan generator soal *Multiple Choice Questions* berbasis *Small Language Model* untuk materi budidaya cabai di Desa Pete, Tigaraksa, yang telah berhasil dideploy pada website resmi desa. Dengan memanfaatkan transkrip tiga video tutorial budidaya cabai total 27 menit sebagai dataset utama, tiga model *SLM GPT-2 Small*, *ELECTRA-Small*, dan *ALBERT-Base-v2* di-fine-tune menggunakan pipeline CRISP-DM dengan parameter ringan *5 epoch*, *batch size 2*, *learning rate 5e-5* agar dapat berjalan pada perangkat *low-resource*. Evaluasi menggunakan *BERTScore* menunjukkan peningkatan signifikan setelah fine-tuning: GPT-2 mencapai F1-score rata-rata 0,8311, ELECTRA-Small 0,7586, dan ALBERT-Base-v2 0,7423. Aplikasi web berbasis Laravel 11 dengan backend *FastAPI* dan database *MySQL* telah terintegrasi penuh pada situs desa melalui tombol floating, memungkinkan siswa SD, guru, dan warga mengakses ribuan soal MCQs secara interaktif dan offline. Penelitian ini membuktikan bahwa SLM kecil yang di-fine-tune pada dataset domain sangat terbatas tetap mampu menghasilkan soal berkualitas tinggi, sekaligus menjadi pionir penerapan AI edukasi berbasis kearifan lokal yang benar-benar hidup dan bermanfaat bagi masyarakat desa.

Kata kunci: Budidaya Cabai, Edukasi Desa, *Multiple Choice Questions*, *Small Language Model*

SMALL LANGUAGE MODEL DESIGN IN COMPILING MULTIPLE CHOICE QUESTIONS IN PETE TIGARAKSA VILLAGE

Josua Rivo Muin

ABSTRACT (English)

This research develops a Multiple Choice Questions generator based on Small Language Models for chili cultivation material in Pete Village, Tigaraksa, which has been successfully deployed on the official village website. Using only transcripts from three chili cultivation tutorial videos total 27 minutes as the primary dataset, three SLM models GPT-2 Small, ELECTRA-Small, and ALBERT-Base-v2 were fine-tuned using the CRISP-DM pipeline with lightweight parameters 5 epochs, batch size 2, learning rate 5e-5 to run on low-resource devices. BERTScore evaluation showed significant improvement post fine-tuning: GPT-2 achieved an average F1-score of 0.8311, ELECTRA-Small 0.7586, and ALBERT-Base-v2 0.7423. The web application, built with Laravel 11, FastAPI backend, and MySQL database, has been fully integrated into the village website via a floating button, enabling elementary students, teachers, and villagers to access thousands of interactive MCQs both online and offline. This research proves that small SLMs fine-tuned on extremely limited domain datasets can still produce high-quality questions and becomes a pioneer in implementing local-wisdom-based AI education that truly lives and benefits rural communities.

Keywords: Chili Cultivation, Rural Education, Multiple Choice Questions, Small Language Model

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iii
KATA PENGANTAR	iv
ABSTRAK	vi
<i>ABSTRACT (English)</i>	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan Penelitian	5
1.4. Urgensi Penelitian	5
1.5. Luaran Penelitian	6
1.6. Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	8
2.1 Penelitian Terdahulu	8
2.2 Tinjauan Teori	15
2.2.1 Multiple Choice	15
2.2.2 Desa Wisata	15
2.2.3 Artificial Intelligence	16
2.2.4 Fine Tuning	17
2.3 Framework/Algoritma yang digunakan	18
2.3.1 Crisp DM	18
2.3.2 Small Language Model	20
2.3.3 GPT 2	20
2.3.4 Electra-small	21
2.3.5 Albert-Base-V2	21
2.4 Tools/software yang digunakan	22

2.4.1	Jupyter Notebook	22
2.4.2	Python	23
2.4.3	Visual Studio Code	23
2.4.4	Laragon	24
BAB III METODE PENELITIAN		25
3.1.	Metode Penelitian	25
3.1.1	Perbandingan Metode	25
3.2	Tahapan Penelitian	27
3.2.1	Business Understanding	29
3.2.2	Data Understanding	29
3.2.3	Data Preparation	31
3.2.4	Modelling	31
3.2.5	Evaluation	32
3.2.6	Deployment	33
3.3	Teknik Pengumpulan Data	34
3.3.1	Wawancara	34
3.3.2	Observasi	35
3.3.3	Studi Pustaka	35
3.4	Teknik Analisis Data	36
3.5	Keterbatasan Model SLM	37
BAB IV HASIL DAN PEMBAHASAN		39
4.1	Business Understanding	39
4.2	Data Understanding	40
4.3	Data Preparation	46
4.4	Modelling	61
4.4.1	GPT 2 Small	62
4.4.2	Electra-Small Discriminator	64
4.4.3	Albert-Base-V2	67
4.4.4	Fine Tuning GPT 2	70
4.4.5	Fine Tuning Electra	72
4.4.6	Fine Tuning Albert	74
4.5	Evaluation	76

4.6 Deployment	79
BAB V SIMPULAN SARAN	90
5.1 Simpulan	90
5.2 Saran	91
5.2.1 Saran untuk Pengelola Desa:	91
5.2.2 Saran untuk Guru/Sekolah:	91
5.2.3 Saran untuk Peneliti Berikutnya:	91
5.3 Kontribusi Utama Penelitian	92
DAFTAR PUSTAKA	93
LAMPIRAN	100



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	8
Tabel 3. 1 Tabel Perbandingan Metode	25
Tabel 3. 2 Tools.....	36
Tabel 4. 1 Tabel Evaluasi Sebelum Fine Tuning	77
Tabel 4. 2 Hasil Evaluasi Setelah Fine Tuning	78



DAFTAR GAMBAR

Gambar 2. 1 Kerangka CRISP-DM [32].....	19
Gambar 3. 1 Alur Penelitian.....	28
Gambar 3. 2 Alur Data Understanding	30
Gambar 3. 3 Alur Data Preparation.....	31
Gambar 3. 4 Alur Modelling.....	32
Gambar 3. 5 Alur Evaluation	33
Gambar 4. 1 Video Pertama	41
Gambar 4. 2 Video Kedua.....	42
Gambar 4. 3 Video Ketiga	42
Gambar 4. 4 Hasil Transkrip 1.....	43
Gambar 4. 5 Hasil Transkrip 2.....	44
Gambar 4. 6 Hasil Transkrip 3.....	44
Gambar 4. 7 Code Transkrip 1	46
Gambar 4. 8 Output Transkrip 1	47
Gambar 4. 9 Code Transkrip 2.....	47
Gambar 4. 10 Output Transkrip 2	48
Gambar 4. 11 Code Transkrip 3.....	49
Gambar 4. 12 Output Transkrip 3	50
Gambar 4. 13 Stopwords Kata Tidak Relevan.....	50
Gambar 4. 14 Menghapus Karakter Tidak Perlu	51
Gambar 4. 15 Menyaring Kalimat	51
Gambar 4. 16 Terapkan Data Cleaned ke Dataframe	52
Gambar 4. 17 Code Output Data Cleaned.....	53
Gambar 4. 18 Output Data Cleaned	54
Gambar 4. 19 Code Menggabungkan Jadi Paragraf.....	55
Gambar 4. 20 Output Paragraf	56
Gambar 4. 21 Code Ringkasan Paragraf	56
Gambar 4. 22 Output Ringkasan Paragraf	57
Gambar 4. 23 Segmentasi Konteks	58
Gambar 4. 24 Output Segmentasi Konteks	59
Gambar 4. 25 Code Knowlegde Graph.....	60
Gambar 4. 26 Knowlegde Graph	61
Gambar 4. 27 Code GPT 2.....	62
Gambar 4. 28 Output MCQs GPT 2	63
Gambar 4. 29 Hasil Evaluasi GPT 2	64
Gambar 4. 30 Code Electra-Small Discriminator	65
Gambar 4. 31 Output MCQs Electra.....	66
Gambar 4. 32 Hasil Evaluasi Electra	67
Gambar 4. 33 Code MCQs Albert	68
Gambar 4. 34 Output MCQs Albert.....	69

Gambar 4. 35 Evaluasi Albert.....	70
Gambar 4. 36 Fine Tuning GPT-2	71
Gambar 4. 37 Evaluasi Fine Tuning GPT-2.....	72
Gambar 4. 38 Code Fine Tuning Electra	73
Gambar 4. 39 Evaluasi Fine Tuning Electra	74
Gambar 4. 40 Code Fine Tuning Albert.....	75
Gambar 4. 41 Evaluasi Fine Tuning Albert	75
Gambar 4. 42 Code Evaluasi.....	77
Gambar 4. 43 Membuat Tabel di Kecamatan_Tigaraksa.....	80
Gambar 4. 44 Database Kecamatan_Tigaraksa	81
Gambar 4. 45 Code Button Ke Page MCQs	81
Gambar 4. 46 Blade Untuk Generate Soal MCQs	82
Gambar 4. 47 Membuat Routes	82
Gambar 4. 48 Membuat Controller Edukasi	83
Gambar 4. 49 Code Fast API MCQs.....	84
Gambar 4. 50 Integrasi Button pada Web Desa Pete	85
Gambar 4. 51 Halaman Awal Kategori.....	85
Gambar 4. 52 Kategori Mode Flora	86
Gambar 4. 53 Generate Soal Otomatis.....	87
Gambar 4. 54 Tampilan Soal	87

DAFTAR LAMPIRAN

Lampiran A Surat Pengantar MBKM	100
Lampiran B Kartu MBKM.....	101
Lampiran C Daily Task Penelitian.....	118
Lampiran D Lembar Verifikasi Laporan MBKM Penelitian.....	119
Lampiran E Lembar Penerimaan MBKM.....	120
Lampiran F Konseling Meeting	121
Lampiran G Lampiran pengecekan hasil Turnitin	122

