IMPLEMENTASI DAN EVALUASI MACHINE LEARNING APPROACH UNTUK PREDIKSI KATA BERDASARKAN RANGKAIAN HURUF HASIL DETEKSI SEMAPHORE

TUGAS AKHIR
SKRIPSI

I GUSTI MADE ARIMBAWA PUTRA
41522010074

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026

lib.mercubuana.ac.id

IMPLEMENTASI DAN EVALUASI MACHINE LEARNING APPROACH UNTUK PREDIKSI KATA BERDASARKAN RANGKAIAN HURUF HASIL DETEKSI SEMAPHORE

TUGAS AKHIR
SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

I GUSTI MADE ARIMBAWA PUTRA
41522010074

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026

lib.mercubuana.ac.id

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama	:	I GUSTI MADE ARIMBAWA PUTRA
NIM	:	41522010074
Fakultas/Program Studi	:	Fakultas Ilmu Komputer / Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:
"Implementasi Dan Evaluasi Machine Learning Approach Untuk Prediksi Kata Berdasarkan Rangkaian Huruf Hasil Deteksi Semaphore" adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apa pun dan tidak melanggar hukum atau hak pihak mana pun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

UNIVERSITAS
MERCU BUANA

Jakarta, 18 Agustus 2026



I Gusti Made Arimbawa Putra

lib.mercubuana.ac.id

UNIVERSITAS
MERCU BUANA
AKREDITASI UNGGUL

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name	: I Gusti Made Arimbawa Putra
NIM /Student ID Number	: 41522010074
Program Studi /Study Program	: Teknik Informatika

Dengan Judul Tugas Akhir
/The title:

"IMPLEMENTASI DAN EVALUASI MACHINE LEARNING APPROACH UNTUK PREDIKSI KATA BERDASARKAN RANGKAIAN HURUF HASIL DETEKSI SEMAPHORE"

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:
/Similarity checks have been carried out with the Turnitin system on the date:

25 Juli 2026

dengan nilai persentase sebesar :
/with a percentage value of:

11%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana.
/declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana

File hasil cek similarity turnitin:
/Turnitin similarity report file

https://drive.google.com/file/d/1A8FbrO6tE7GV_MRGc5cXIXG0w_MrHGmi/view?usp=drive_link

Jakarta, 25 Juli 2026
Admin Turnitin Fasilkom UMB

Agung Prawoto

Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

Fakultas Ilmu Komputer
KAMPUS MENARA BHAKTI
Jl. Raya Meruya Selatan No. 1 Kembangan, Jakarta Barat 11650
Telp. 021-5840816 (Hunting), Psw : 5700 Fax. 021-5840813
http://www.mercubuana.ac.id, e-mail : fasilkom@mercubuana.ac.id

lib.mercubuana.ac.id

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama	:	I GUSTI MADE ARIMBAWA PUTRA
NIM	:	41522010074
Fakultas/Program Studi	:	Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir	:	Implementasi Dan Evaluasi Machine Learning Approach Untuk Prediksi Kata Berdasarkan Rangkaian Huruf Hasil Deteksi Semaphore

Telah berhasil dipertahankan pada sidang tanggal 18 Agustus 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing

Ilham Nugraha, S.Kom ,M.Sc

NIDN/NUPTK: 0307098904

Jakarta, 18 Agustus 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer	Ketua Program Studi Teknik Informatika
Dr. Bambang Jokonowo, S.Si., MTI	Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0320037002	NIDN/NUPTK: 0225067701

lib.mercubuana.ac.id

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Bapak Ilham Nugraha, S.Kom ,M.Sc selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana.
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 18 Agustus 2026

I Gusti Made A. P.

lib.mercubuana.ac.id

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama	:	I GUSTI MADE ARIMBAWA PUTRA
NIM	:	41522010074
Fakultas/Program Studi	:	Fakultas Ilmu Komputer / Teknik Informatika
Judul Laporan Skripsi	:	Implementasi Dan Evaluasi Machine Learning Approach Untuk Prediksi Kata Berdasarkan Rangkaian Huruf Hasil Deteksi Semaphore

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS MERCU BUANA

Jakarta, 18 Agustus 2026

Yang menyatakan,



I Gusti Made Arimbawa Putra

lib.mercubuana.ac.id

IMPLEMENTASI DAN EVALUASI MACHINE LEARNING APPROACH UNTUK PREDIKSI KATA BERDASARKAN RANGKAIAN HURUF HASIL DETEKSI SEMAPHORE

I GUSTI MADE ARIMBAWA PUTRA

ABSTRAK

Komunikasi menggunakan sandi semaphore merupakan salah satu bentuk komunikasi visual yang merepresentasikan huruf melalui posisi bendera. Dalam implementasi sistem penerjemahan semaphore berbasis komputer, hasil deteksi huruf sering kali bersifat parsial dan tidak sempurna, sehingga diperlukan mekanisme prediksi kata untuk meningkatkan akurasi dan kegunaan sistem. Penelitian ini bertujuan untuk mengimplementasikan dan mengevaluasi pendekatan *machine learning* dalam memprediksi kata berdasarkan rangkaian huruf hasil deteksi semaphore. Empat pendekatan digunakan dalam penelitian ini, yaitu *Long Short-Term Memory* (LSTM), *Gated Recurrent Unit* (GRU), *Hybrid N-Gram + LSTM*, dan *Hybrid N-Gram + GRU*. Dataset yang digunakan terdiri dari dua jenis kosakata, yaitu kosakata Bahasa Indonesia dan kosakata Bahasa Inggris. Dataset Bahasa Indonesia digunakan sebagai dataset utama, sedangkan dataset Bahasa Inggris digunakan sebagai dataset pembanding untuk menguji konsistensi performa model pada karakteristik bahasa yang berbeda. Kedua dataset melalui proses *preprocessing* yang sama, meliputi normalisasi huruf sampai pembentukan pasangan *prefix* dan target karakter. Model neural dilatih pada tingkat karakter atau *character-level*, sedangkan pendekatan *hybrid* mengombinasikan probabilitas model neural dengan model bahasa statistik N-Gram. Evaluasi dilakukan menggunakan metrik berbasis *Top-K*, yaitu *Top-K Accuracy*, *Precision@K*, *Recall@K*, dan *F1@K*, yang lebih relevan untuk sistem prediksi kata berbasis *prefix*. Hasil penelitian menunjukkan bahwa pendekatan *hybrid* secara konsisten mengungguli model neural murni pada kedua dataset. Pada dataset Bahasa Indonesia, model *Hybrid N-Gram + LSTM* memperoleh performa terbaik dengan *Top-5 Accuracy* sebesar 0,8700, sedangkan pada dataset Bahasa Inggris, model *Hybrid N-Gram + GRU* memperoleh performa terbaik dengan *Top-5 Accuracy* sebesar 0,9912. Model neural murni LSTM dan GRU memperoleh performa yang lebih rendah. Berdasarkan hasil tersebut, dapat disimpulkan bahwa pendekatan *Hybrid N-Gram + Neural Network* merupakan solusi yang efektif dan robust untuk sistem prediksi kata berbasis rangkaian huruf hasil deteksi semaphore. Integrasi N-Gram membantu meningkatkan validitas linguistik dan konsistensi prediksi, sedangkan model neural berperan dalam mempelajari pola urutan karakter. Dengan demikian, pendekatan *hybrid* mampu meningkatkan akurasi prediksi kata pada dataset Bahasa Indonesia maupun Bahasa Inggris.

Kata Kunci: semaphore, prediksi kata, Bahasa Indonesia, Bahasa Inggris, LSTM, GRU, N-Gram, machine learning, Top-K evaluation.

lib.mercubuana.ac.id

IMPLEMENTATION AND EVALUATION OF A MACHINE LEARNING APPROACH FOR WORD PREDICTION BASED ON LETTER SEQUENCES GENERATED FROM SEMAPHORE DETECTION

I GUSTI MADE ARIMBAWA PUTRA

ABSTRACT

Semaphore communication is a visual communication method that represents letters through specific flag positions. In computer-based semaphore translation systems, detected letter sequences are often incomplete or noisy, which requires a word prediction mechanism to improve system accuracy and usability. This study aims to implement and evaluate machine learning approaches for word prediction based on sequences of letters obtained from semaphore detection. Four approaches are investigated in this study, namely Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Hybrid N-Gram + LSTM, and Hybrid N-Gram + GRU. The datasets used in this research consist of two vocabularies, namely Indonesian vocabulary and English vocabulary. The Indonesian dataset is used as the main dataset, while the English dataset is used as a comparative dataset to evaluate the consistency of model performance on vocabularies with different linguistic characteristics. Both datasets undergo the same preprocessing stages, including character normalization, removal of non-alphabetic characters, removal of numbers, word-length filtering, and prefix-target character mapping. Neural models are trained at the character level, while the hybrid approach combines neural model probabilities with a statistical N-Gram language model. Model evaluation is conducted using Top-K-based metrics, including Top-K Accuracy, Precision@K, Recall@K, and F1@K, which are more suitable for prefix-based word prediction tasks. The experimental results show that hybrid approaches consistently outperform standalone neural models on both datasets. On the Indonesian dataset, the Hybrid N-Gram + LSTM model achieves the best performance with a Top-5 Accuracy of 0.8700, while on the English dataset, the Hybrid N-Gram + GRU model achieves the highest performance with a Top-5 Accuracy of 0.9912. Standalone LSTM and GRU models obtain lower performance, with Top-5 Accuracy ranging from 0.40 to 0.47. Based on these results, it can be concluded that the Hybrid N-Gram + Neural Network approach is an effective and robust solution for word prediction systems based on semaphore letter sequences. The integration of N-Gram improves linguistic validity and prediction consistency, while neural models contribute to learning sequential character patterns. Therefore, the hybrid approach is able to improve word prediction accuracy on both Indonesian and English datasets.

Keywords: semaphore, *word prediction*, Indonesian, English, LSTM, GRU, N-Gram, *machine learning*, Top-K evaluation.

lib.mercubuana.ac.id

DAFTAR ISI

HALAMAN SAMPUL **0**

HALAMAN JUDUL ..... **i**

HALAMAN PERNYATAAN KARYA SENDIRI ..... **ii**

HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN ..... **iii**

HALAMAN PENGESAHAN ..... **iv**

KATA PENGANTAR ..... **v**

ABSTRAK ..... **vii**

ABSTRACT ..... **viii**

DAFTAR ISI ..... **ix**

DAFTAR TABEL ..... **xii**

DAFTAR GAMBAR ..... **xiii**

DAFTAR LAMPIRAN ..... **xiv**

BAB I PENDAHULUAN **1**

1.1 Latar Belakang 1

1.2 Perumusan Masalah 3

1.3 Batasan Masalah 4

1.4 Tujuan Penelitian 5

1.5 Manfaat Penelitian 6

1.6 Sistematika Penulisan 7

BAB II TINJAUAN PUSTAKA **9**

2.1 Teori Utama 9

2.1.1 Prediksi Kata Berbasis Urutan Karakter (Character-Level Word Prediction) 9

2.1.2 Model Statistik N-Gram dalam Prediksi Teks 10

2.1.3 Long Short-Term Memory (LSTM) untuk Pemodelan Urutan Karakter 11

2.1.4 Gated Recurrent Unit (GRU) sebagai Alternatif LSTM 11

2.1.5 Algoritma Long Short-Term Memory (LSTM) 12

2.2 Teori Pendukung 12

2.2.1 Representasi Teks dan Embedding dalam NLP 12

2.2.2 Model Bahasa Pralatih dan Representasi Kontekstual 13

2.2.3 Deep Learning dalam Pemrosesan Bahasa Alami 13

2.2.4 Evaluasi Model Bahasa dan Metrik Kinerja 14

2.2.5 Ensemble dan Hybrid Approach dalam NLP 14

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

2.3 Penelitian Terdahulu 17
2.4 Gap Penelitian 23

BAB III METODE PENELITIAN 26
3.1 Pendekatan Penelitian 26
3.2 Desain Penelitian 27
3.2.1 Identifikasi Masalah dan Perumusan Tujuan 29
3.2.2 Pengumpulan dan Persiapan Data 29
3.2.3 Perancangan dan Implementasi Model 30
3.2.4 Pelatihan dan Pengujian Model 31
3.2.5 Evaluasi dan Analisis Hasil 32
3.3 Subjek Penelitian 32
3.4 Instrumen Penelitian 34
3.5 Teknik Pengumpulan Data 37
3.6 Analisis Data 41
3.6.1 Preprocessing dan Normalisasi Data 41
3.6.2 Pembentukan Data Pelatihan (Prefix–Target Mapping) 43
3.6.3 Pelatihan Model Neural 43
3.6.4 Implementasi Model Hybrid (N-Gram + Neural Network) 44
3.6.5 Evaluasi Model 45
3.7 Prosedur Penelitian 46
3.8 Evaluasi Hasil Penelitian 50

BAB IV EKSPERIMEN DAN HASIL 53
4.1. Deskripsi Umum implementasi 53
4.2. Hasil Pelatihan Model Neural (LSTM dan GRU) 56
4.2.1 Konfigurasi Pelatihan Model Neural 56
4.2.2 Proses Pelatihan Model LSTM 57
4.2.3 Proses Pelatihan Model GRU 59
4.2.4 Analisis Perbandingan LSTM dan GRU 60
4.2.5 Implikasi terhadap Tahap Selanjutnya 61
4.3. Hasil Implementasi Model Hybrid (N-Gram + Neural Network) 61
4.3.1 Latar Belakang Pendekatan Hybrid 62
4.3.2 Model N-Gram Karakter 62
4.3.3 Integrasi N-Gram dan Model Neural 63
4.3.4 Implementasi Model Hybrid 64
4.3.5 Variasi Model Hybrid yang Digunakan 64
4.3.6 Peran Model Hybrid dalam Sistem Prediksi Semaphore 65
4.3.7 Kesimpulan Eksperimen 65
4.4. Evaluasi Kuantitatif Model (Top-K Metrics) 65
4.4.1 Tujuan Evaluasi Kuantitatif 66
4.4.2 Skema Pengujian 66
4.4.3 Metrik Evaluasi 66

lib.mercubuana.ac.id

4.4.4 Hasil Evaluasi Kuantitatif 67
4.4.5 Analisis Hasil Kuantitatif 68
4.5. Analisis Kualitatif Hasil Prediksi 69
4.5.1 Tujuan Analisis Kualitatif 69
4.5.2 Perbandingan Hasil Prediksi Berdasarkan Prefix 69
4.5.3 Konsistensi Prediksi Top-K 72
4.5.4 Implikasi terhadap Sistem Prediksi Semaphore 72
4.5.5 Ringkasan Analisis Kualitatif 72
4.6. Pembahasan Hasil Penelitian 73
4.6.1 Keterkaitan Hasil dengan Tujuan Penelitian 73
4.6.2 Pembahasan Performa Model Neural Murni (LSTM dan GRU) 73
4.6.3 Dampak Integrasi N-Gram terhadap Kinerja Model 74
4.6.4 Perbandingan Hybrid N-Gram + LSTM dan Hybrid N-Gram + GRU 74
4.6.5 Perbandingan Performa Dataset Bahasa Indonesia dan Bahasa Inggris 75
4.6.6 Relevansi Metrik Top-K terhadap Sistem Semaphore 75
4.6.7 Implikasi terhadap Implementasi Sistem Nyata 76
4.6.8 Keterbatasan Penelitian 76
4.6.9 Ringkasan Pembahasan 76

BAB V Kesimpulan dan saran 78
5.1 Kesimpulan 78
5.2 Saran 81

DAFTAR PUSTAKA 84

LAMPIRAN 90
Lampiran 1 Kartu Asistensi 90
Lampiran *2* Curiculum Vitae 91
Lampiran *3* Surat Pernyataan & Pengalihan HAKI 92
Lampiran *4* Surat Keterangan Lulus 94

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait 17
Tabel 3. 1 Hyperparameter 31
Tabel 3. 2 Komponen Hardware 35
Tabel 3. 3 Komponen Software 35
Tabel 3. 4 Dataset Kosakata 37
Tabel 3. 5 Dataset Kosakata 38
Tabel 3. 6 Hyperparameter 44
Tabel 4. 1 Hyperparameter 55
Tabel 4. 2 Konfigurasi Hyperparameter Model Neural 57
Tabel 4. 3 Perbandingan Hasil Pelatihan Model Neural pada Dataset Bahasa Indonesia dan Bahasa Inggris 60
Tabel 4. 4 Hasil Evaluasi Dataset Bahasa Indonesia 67
Tabel 4. 5 Hasil Evaluasi Dataset Bahasa Inggris 67

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR GAMBAR

Gambar 3. 1 Desain penelitian 28
Gambar 3. 2 Flowchart Teknik Pengumpulan dan Persiapan Data 40
Gambar 3. 3 Distribusi Panjang Kata dalam Dataset Bahasa Indonesia Setelah preprocessing 42
Gambar 3. 4 Distribusi Panjang Kata dalam Dataset Bahasa Inggris Setelah preprocessing 42
Gambar 4. 1 Hasil Uji Prefix LSTM dan GRU (Bahasa Indonesia)..... 70
Gambar 4. 2 Hasil Uji Prefix LSTM dan GRU (Bahasa Inggris)..... 70
Gambar 4. 3 Hasil Uji Prefix Hybrid N-Gram + LSTM / GRU (Bahasa Indonesia) 71
Gambar 4. 4Hasil Uji Prefix Hybrid N-Gram + LSTM / GRU (Bahasa Inggris). 71

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi 90
Lampiran 2 Curiculum Vitae 91
Lampiran 3 Surat Pernyataan & Pengalihan HAKI 92
Lampiran 4 Surat Keterangan Lulus 94

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id