



**PERBANDINGAN KINERJA METODE XLM DAN XLM-ROBERTA
DALAM IDENTIFIKASI LAPORAN BENCANA ALAM DI MEDIA SOSIAL**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

Oleh

MAIMUNAH

NIM 1911016220019

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JANUARI 2026**

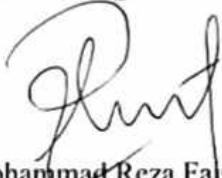
SKRIPSI

PERBANDINGAN KINERJA METODE XLM DAN XLM ROBERTA DALAM IDENTIFIKASI LAPORAN BENCANA ALAM DI MEDIA SOSIAL.

Oleh:
MAIMUNAH
NIM 1911016220019

Telah dipertahankan di depan Dosen Penguji pada Tanggal 19 Januari 2026
Susunan Dosen Penguji:

Pembimbing I



Mohamad Reza Faisal, S.T., M.T., Ph.D.
NIP. 197612202008121001

Dosen Penguji I



Andi Farmadi, M.T.
NIP. 197307252008011006

Pembimbing II



Triando Hamonangan Saragih, S.Kom., M.Kom.
NIP. 199308242019031012

Dosen Penguji II

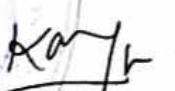


Rudy Hertoho, S.Kom., M.Kom.
NIP. 198809252022031003

Banjarbaru, 23 Januari 2026

Koordinator Program Studi Ilmu Komputer




Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diberikan orang lain, kecuali yang secara tertulis diacu di dalam naskah ini dan disebutkan dalam Daftar Pustaka

Banjarbaru, 21 Januari 2026



MAIMUNAH
NIM. 1911016220019

ABSTRAK

PERBANDINGAN KINERJA METODE XLM DAN XLM ROBERTA DALAM IDENTIFIKASI LAPORAN BENCANA ALAM DI MEDIA SOSIAL

(Oleh: Maimunah; Pembimbing: Mohammad Reza Faisal, S.T., M.T., PhD, Triando Hamonangan Saragih, S.Kom, M.Kom, Andi Farmadi, M.T, Rudy Herteno S.Kom, M.Kom; 2026; 47 halaman)

Bencana alam sering terjadi di Indonesia dan media sosial seperti X (Twitter) menjadi sumber informasi real-time. Penelitian ini bertujuan mengidentifikasi laporan bencana alam dari tweet berbahasa Indonesia dengan membandingkan kinerja model *Cross-lingual Language Model* (XLM) dan XLM-RoBERTa (XLM-R). Dataset yang digunakan adalah Natural Disaster on Twitter (Indonesia) yang terdiri dari 3.000 tweet untuk setiap bencana (banjir, gempa, kebakaran) dengan tiga kelas: eyewitness, noneyewitness, dan dontknow. Setelah preprocessing, data dibagi dengan rasio 80:10:10 untuk pelatihan, validasi, dan pengujian. Model di-*fine-tune* menggunakan hyperparameter yang sama (learning rate $5e-6$, batch size 16, max length 128). Hasil evaluasi menunjukkan bahwa XLM-R secara konsisten mengungguli XLM dengan akurasi rata-rata 84,33% dibandingkan 78,11%. Akurasi tertinggi dicapai pada kategori kebakaran (XLM: 85,67%, XLM-R: 88,67%). Hasil ini menunjukkan bahwa XLM-R lebih efektif dalam mengklasifikasikan laporan bencana dari media sosial berbahasa Indonesia karena arsitektur dan corpus pretraining yang lebih besar. Implikasi penelitian ini adalah model XLM-R dapat diintegrasikan ke dalam sistem pemantauan bencana berbasis media sosial untuk respon yang lebih cepat dan akurat.

Kata kunci: Bencana Alam, Klasifikasi Teks, XLM, XLM-RoBERTa, Media Sosial, Bahasa Indonesia

ABSTRACT

PERFORMANCE COMPARISON OF XLM AND XLM-ROBERTA FOR IDENTIFYING NATURAL DISASTER REPORTS ON SOCIAL MEDIA

(By: Maimunah; Supervisor: Mohammad Reza Faisal, S.T., M.T., PhD, Triando Hamonangan Saragih, S.Kom, M.Kom, Andi Farmadi, M.T, Rudy Herteno S.Kom, M.Kom; 2026; 47 pages)

Natural disasters frequently occur in Indonesia, and social media platforms such as X (formerly Twitter) have become an important source of real-time information. This study aims to identify natural disaster reports from Indonesian-language tweets by comparing the performance of Cross-lingual Language Model (XLM) and XLM-RoBERTa (XLM-R). The dataset used is Natural Disaster on Twitter (Indonesia), comprising 3,000 tweets for each disaster type (flood, earthquake, and fire) labeled into three classes: eyewitness, non-eyewitness, and dontknow. The data were split 80:10:10 for training, validation, and testing. Both models were fine-tuned using the same hyperparameters (learning rate $5e-6$, batch size 16, maximum sequence length 128). The evaluation results show that XLM-R consistently outperforms XLM, with an average accuracy of 84.33% compared with 78.11%. The highest accuracy for both models is achieved in the fire category, while the earthquake category is the most challenging. Confusion-matrix analysis indicates that XLM-R is more effective in predicting the eyewitness and non-eyewitness classes and reduces misclassification of the dontknow class. The findings support integrating XLM-R into Indonesian social-media-based disaster monitoring systems to enable faster and more accurate response.

Keywords: Natural Disaster, Text Classification, XLM, XLM-RoBERTa, Social Media, Indonesian Language

PRAKATA

Assalamualaikum Warahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Identifikasi Laporan Bencana Alam dari Media Sosial Menggunakan XLM dan XLM-RoBERTa” untuk memenuhi syarat dalam menyelesaikan pendidikan program S-1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa pula penulis panjatkan shalawat dan salam ke hadirat Rasulullah Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga yaumul qiyamah.

Pada lembar ini penulis ingin menyampaikan ucapan terima kasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

1. Keluarga terutama orang tua yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian skripsi ini.
2. Bapak Mohammad Reza Faisal, S.T., M.T., Ph.D. selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran penyelesaian skripsi ini.
3. Bapak Triando Hamonangan Saragih, S.Kom., M.Kom. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
5. Teman-teman keluarga Ilmu Komputer angkatan 2019, terima kasih untuk canda, tawa, perjuangan yang sudah dilewati bersama, untuk semua kenangan manis yang telah terukir selama ini. Senang bisa menjadi salah satu bagian dari kehidupan kalian.
6. Ucapan terimakasih kepada Khadijah, Yama, Ica, Fadia, Hartati, Molika, Ridha, Exzy, dan teman-teman lainnya yang telah kebersamai melewati lika liku perkuliahan serta memberikan dukungan dalam proses pengerjaan skripsi.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, maka saran dan kritik dari semua pihak sangat diharapkan demi penyempurnaan pada penelitian selanjutnya. Diharapkan skripsi ini dapat bermanfaat bagi semua pihak dan dapat menambah pengetahuan kita semua.

Banjarbaru, 21 Januari 2026

A handwritten signature in black ink, appearing to be 'Muh', with a horizontal line underneath.

Penulis

DAFTAR ISI

SKRIPSI	ii
PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA	vi
Penulis DAFTAR ISI	vii
BAB I PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 RUMUSAN MASALAH	4
1.3 TUJUAN PENELITIAN	4
1.4 MANFAAT PENELITIAN	5
1.5 BATASAN MASALAH.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Kajian terdahulu	7
2.2. Landasan Teori.....	12
BAB III	23
METODE PENELITIAN	23
3.1. Alat Penelitian	23
3.2. Bahan Penelitian.....	23
3.3. Prosedur Penelitian.....	24
BAB IV	31
HASIL DAN PEMBAHASAN	31
BAB V PENUTUP	46
DAFTAR PUSTAKA	48
LAMPIRAN	
RIWAYAT HIDUP	

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	9
Tabel 2. Kelas Eyewitness pada dataset Gempa	25
Tabel 3. Kelas Eyewitness pada dataset Kebakaran	25
Tabel 4. Kelas Eyewitness pada dataset Banjir	26
Tabel 5. Kelas Noneyewitness Gempa	26
Tabel 6. Kelas Noneyewitness Kebakaran	26
Tabel 7. Kelas Noneyewitness Banjir	27
Tabel 8. Kelas dontknow Gempa	27
Tabel 9 Kelas dontknow Kebakaran	28
Tabel 10. Kelas dontknow Banjir	28
Tabel 11. Dataset Bencana Alam	31
Tabel 12. Class label	32
Tabel 13. Hasil Tokenisasi dan Konversi ke Input Ids	33
Tabel 14. Distribusi Data Latih, Validasi, dan Uji	34
Tabel 15. Parameter Konfigurasi Model	35
Tabel 16. Hasil Akurasi Model	35
Tabel 17. Hasil Presisi Model	36
Tabel 18. Hasil Recall Model	36
Tabel 19. Hasil F1-Score Model	37
Tabel 20. Confusion Matrix Model XLM pada Dataset Banjir	37
Tabel 21. Confusion Matrix Model XLM pada Dataset Gempa	38
Tabel 22. Confusion Matrix Model XLM pada Dataset Kebakaran	38
Tabel 23. Confusion Matrix Model XLM-RoBERTa pada Dataset Banjir	39
Tabel 24. Confusion Matrix Model XLM-RoBERTa pada Dataset Gempa	39
Tabel 25. Confusion Matrix Model XLM-RoBERTa pada Dataset Kebakaran	39
Tabel 26. Hasil Uji Paired t-test Perbandingan Akurasi XLM dan XLM-RoBERTa	41

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Ilustrasi Mekanisme MLM dan TLM pada XLM	17
Gambar 2. Arsitektur Model XLM-RoBERTa untuk Tugas Klasifikasi Teks	18
Gambar 3 <i>Confusion Matrix</i> Multikelas (Tharwat, 2018)	20
Gambar 4. Prosedur Penelitian	24
Gambar 3. Perbandingan Rata-rata Setiap Evaluasi Model pada XLM dan XLM-RoBERTa	41
Gambar 4. Perbandingan Rata-rata Evaluasi Model XLM dan XLM-RoBERTa pada Setiap Kategori Bencana	41