KLASIFIKASI TEKS DEPRESI PADA MEDIA SOSIAL MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN FASTTEXT EMBEDDING

Skripsi

Oleh:

AZKI MUNTAZUL HAMDI

NPM : 2155201001040

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS MERDEKA PASURUAN

TAHUN 2026

KLASIFIKASI TEKS DEPRESI PADA MEDIA SOSIAL MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN FASTTEXT EMBEDDING

Skripsi

Oleh:

AZKI MUNTAZUL HAMDI

NPM : 2155201001040

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS MERDEKA PASURUAN

TAHUN 2026

KLASIFIKASI TEKS DEPRESI PADA MEDIA SOSIAL MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN FASTTEXT EMBEDDING

SKRIPSI

Diajukan sebagai salah satu persyaratan

Dalam menyelesaikan program Sarjana Komputer (S.Kom.)

Oleh:

AZKI MUNTAZUL HAMDI

NPM : 2155201001040

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS MERDEKA PASURUAN

TAHUN 2026

PERSETUJUAN PEMBIMBING SKRIPSI

KLASIFIKASI TEKS DEPRESI PADA MEDIA SOSIAL MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN FASTTEXT EMBEDDING

Nama	:	**Azki Muntazul Hamdi**
NPM	:	**2155201001040**
Program Studi	:	**Teknik Informatika**

Pasuruan, 3 Agustus 2026
Telah diperiksa dan disetujui oleh :

Pembimbing I	**Pembimbing II**
Moch. Zoqi Sarwani, S.pd., M.Kom.	**Anang Aris Widodo, S.Kom., M.T.**
NIDN. 0716049101	**NIDN. 0702038102**

PERSETUJUAN DAN PENGESAHAN SKRIPSI

Srikpsi Oleh Azki Muntazul Hamdi ini

Telah dipertahankan didepan dewan penguji

Pada Tanggal 03 Agustus 2026

Ketua Penguji

(Dian Ahkam Sani, S.kom.,M.Kom.)
NIDN. 0724079202

Penguji I

(Dr.Muhammad Misdram,S.Kom.,M.Kom.)
NIDN. 0717046704

Penguji II

(Rudi Hariyanto, S.Kom., M.T.)
NIDN. 0718018401

Mengesahkan,
Dekan Fakultas Teknologi Informasi

Muhammad Mahrus Ali, S.Kom.,M.Kom.
NIDN.0721129003

Mengetahui,

Ketua Program Studi Informatika

Rudi Hariyanto, S.Kom., M.T.
NIDN. 0718018401

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama	: Azki Muntazul Hamdi
NPM	:2155201001040
Jurusan/Program Studi	: Teknik Informatika
Fakultas/Program	: Teknologi Informasi

Menyatakan dengan sesungguhnya bahwa skripsi yang saya tulis ini benar-benar tulisan saya,dan bukan merupakan plagiasi baik sebagian atau seluruhnya. Apabila dikemudian hari terbuktiatau dapat dibuktikan bahwa skripsi ini hasil plagiasi, baik Sebagian atau seluruhnya, makasaya bersedia menerima sanksi atas perbuatan tersebut sesuai dengan ketentuan yang berlaku.

Pasuruan, Agustus 2026

Yang membuat pernyataan



AZKI MUNTAZUL HAMDI

2155201001040

ABSTRAK

Hamdi, Azki, Muntazul. 2025. *Klasifikasi Teks Depresi Pada Media Sosial Menggunakan Convolutional Neural Network (CNN) Dengan Fasttext Embedding*. Program Studi Informatika, Fakultas Teknologi Informasi Universitas Merdeka Pasuruan.

Pembimbing : (I) Moh. Zoqi Sarwani, S.Pd., M.Kom. (II) Anang Aris Widodo, S.Kom.,M.T.

Kata Kunci : *Depresi, Media Sosial,* Natural Language Processing; Convolutional Neural, Network, Fasttext Embedding.

Depresi merupakan gangguan suasana hati persisten yang secara signifikan memengaruhi fungsi fungsional individu dan prevalensinya kian mengkhawatirkan pada kelompok remaja (Strawberry Generation). Ekspresi tekanan psikologis ini kini banyak bermigrasi ke media sosial dalam bentuk unggahan teks tidak terstruktur yang sarat akan istilah non-formal, singkatan, dan salah ketik (typo). Karakteristik bahasa tersebut memicu kendala Out-of-Vocabulary (OOV) pada model klasifikasi tradisional. Penelitian ini mengusulkan integrasi word embedding FastText berbasis informasi tingkat sub-kata (subword information) untuk memperkuat representasi fitur, yang dikombinasikan dengan arsitektur Convolutional Neural Network (CNN) sebagai pengklasifikasi utama. Menggunakan dataset sekunder "Student Depression Text" sebanyak 7.489 entri dari Kaggle, ketidakseimbangan kelas diatasi secara eksklusif pada data latih menggunakan teknik Random Oversampling. Hasil eksperimen menunjukkan bahwa model yang diusulkan meraih akurasi keseluruhan sebesar 95%, dengan nilai Precision, Recall, dan F1-Score untuk kelas depresi masing-masing sebesar 90%, 80%, dan 84%. Integrasi CNN dan pre-trained FastText terbukti efektif menjadi solusi jalan tengah yang andal dan efisien dalam menangani kompleksitas bahasa digital untuk deteksi dini kesehatan mental.

ABSTRACT

Hamdi, Azki, Muntazul. 2025. Depression Text Classification on Social Media Using Convolutional Neural Network (CNN) with FastText Embedding**.** Informatics Study Program, Faculty of Information Technology, Merdeka University of Pasuruan

Supervisors**:** (I) Moh. Zoqi Sarwani, S.Pd., M.Kom. (II) Anang Aris Widodo, S.Kom., M.T.

Keywords: Depression, Social Media, Natural Language Processing; Convolutional Neural, Network, FastText Embedding.

Depression is a persistent mood disorder that significantly affects an individual's functioning, with its prevalence becoming increasingly concerning among adolescents, particularly the so-called "Strawberry Generation." Expressions of psychological distress have increasingly migrated to social media in the form of unstructured text posts containing non-formal terms, abbreviations, and typographical errors (typos). These linguistic characteristics present challenges related to the Out-of-Vocabulary (OOV) problem in traditional classification models. This study proposes the integration of FastText word embeddings based on subword information to strengthen feature representation, combined with a Convolutional Neural Network (CNN) architecture as the primary classifier. Using the secondary dataset "Student Depression Text," consisting of 7,489 entries obtained from Kaggle, class imbalance was addressed exclusively in the training data using the Random Oversampling technique. The experimental results show that the proposed model achieved an overall accuracy of 95%, with Precision, Recall, and F1-Score values for the depression class of 90%, 80%, and 84%, respectively. The integration of CNN and pre-trained FastText proved effective as a reliable and efficient middle-ground solution for addressing the complexity of digital language in the early detection of mental health conditions.

KATA PENGANTAR

Puji syukur kehadirat Tuhan Yang Maha Esa atas segala rahmat, karunia, dan kasih-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Klasifikasi Teks Depresi Pada Media Sosial Menggunakan Convolutional Neural Network (CNN) Dengan FastText Embedding" dengan baik dan tepat pada waktunya.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Teknologi Informasi, Universitas Merdeka Pasuruan. Penelitian ini membahas penerapan metode Natural Language Processing (NLP), khususnya integrasi FastText Embedding dan Convolutional Neural Network (CNN), dalam melakukan klasifikasi teks terkait depresi pada media sosial.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa penyelesaian penelitian tidak terlepas dari bantuan, bimbingan, dukungan, serta doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada :

1. Tuhan Yang Maha Esa yang telah memberikan kesehatan, kekuatan, dan kemudahan kepada penulis selama proses penyusunan skripsi.
2. Bapak Dr. M. Misdram, S.Kom., M.Kom. selaku Rektor Universitas Merdeka Pasuruan.
3. Bapak Muhammad Mahrus Ali,S.Kom.M.Kom. selaku Dekan Fakultas Teknologi Informasi.
4. Bapak Rudi Hariyanto, S.Kom.,M.T. selaku ketua Program Studi Teknik Informatika.
5. Bapak Moh. Zoqi Sarwani, S.Pd., M.Kom. dan Bapak Anang Aris Widodo, S.Kom.,M.T., selaku dosen pembimbing yang telah memberikan arahan, masukan, selama proses penyusunan skripsi.
6. Seluruh dosen dan staf akademi di Fakultas Teknologi Informasi yang telah memberikan ilmu dan pengalaman selama masa perkuliahan..
7. Orang tua tercinta, ayah, ibu, dan seluruh saudara-saudara saya yang selalu memberikan doa, dukungan, serta semangat tanpa henti kepada saya.
8. Untuk Annisa Marshanda Rahmania selaku pacar saya yang senantiasa memberikan doa, semangat dan sekaligus dukungan kepada saya.

9. Sahabat dekat dan semua teman-teman saya yang senantiasa memberikan doa dan semangat kepada saya.
10. Semua pihak yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan dan jauh dari kesempurnaan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari berbagai pihak sebagai bahan perbaikan dan penyempurnaan di masa yang akan datang.

Akhir kata, penulis berharap semoga skripsi ini dapat memberikan manfaat bagi pembaca, peneliti selanjutnya, serta dapat menjadi salah satu referensi dalam pengembangan penerapan *Natural Language Processing*, khususnya dalam klasifikasi teks dan deteksi dini terkait kesehatan mental pada media sosial.

Pasuruan, Agustus 2026

Penulis

DAFTAR ISI

ABSTRAK ..... **ii**
ABSTRACT ..... **iii**
KATA PENGANTAR ..... **iv**
DAFTAR ISI ..... **vi**
DAFTAR TABEL ..... **viii**
DAFTAR GAMBAR ..... **ix**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang **1**
1.2 Perumusan Masalah **4**
1.3 Batasan Masalah **4**
1.4 Tujuan Penelitian **5**
1.5 Manfaat Penelitian **5**
BAB II TINJAUAN PUSTAKA **7**
2.1 Tinjauan Penelitian **7**
2.2 Landasan Teori **12**
2.2.1 Kesehatan Mental dan Depresi **12**
2.2.2 Natural Language Processing (NLP) **13**
2.2.3 Text Preprocessing **14**
2.2.4 Word Embedding FastText **15**
2.2.5 Convolutional Neural Network (CNN) **17**
2.2.6 Evaluasi Model **19**
BAB III METODOLOGI PENELITIAN **22**
3.1 Alur Penelitian **22**
3.2 Perancangan System **25**
3.2.1 Dataset **26**

3.2.2 Preprocessing 30
3.2.3 Data Splitting 32
3.2.4 Oversampling 33
3.2.5 FastText Word Embedding 34
3.2.6 Convolutional Neural Network 36
3.2.7 Model Evaluation 38
3.3 Teknik Analisa Data 40
3.3.1 Analisis Eksplorasi Data 40
3.3.2 Analisis Performa Klasifikasi 41
3.3.3 Analisis Kesalahan (Error Analysis) 41
3.3.4 Analisis Stabilitas dan Konvergensi Model 41
BAB IV HASIL DAN PEMBAHASAN 43
4.1 Implementasi 43
4.2 Evaluasi Performa Klasifikasi Teks Depresi 43
4.3 Error Analysis Confusion Matrix 45
4.4 Analisis Stabilitas dan Konvergensi Pelatihan Model 46
4.4.1 Analisis Kurva Akurasi (Model Accuracy) 47
4.4.2 Analisis Kurva Kerugian (Model Loss) 48
4.5 Pembahasan Integrasi CNN dan FastText Word Embedding 49
BAB V PENUTUP 50
5.1 Kesimpulan 50
5.2 Saran 51
DAFTAR PUSTAKA 52

DAFTAR TABEL

Tabel 3. 1 Dataset Student Depression Text 26

Tabel 4. 1 Parameter dan Konfigurasi Model CNN-FastText 43
Tabel 4. 2 Classification Report 44

DAFTAR GAMBAR

Gambar 2. 1. Representasi FastText 16
Gambar 2. 2 Convolutional Neural Network 17
Gambar 2. 3 Confusion Matrix 19

Gambar 3. 1 Alur Penelitian 22
Gambar 3. 2 Perancangan System..... 26
Gambar 3. 3 Distribusi Usia..... 27
Gambar 3. 4 Distribusi Gender 28
Gambar 3. 5 Distribusi Label..... 29

Gambar 4. 1 Confussion Matrix..... 45
Gambar 4. 2 Grafik Hasil Model Accuracy 47
Gambar 4. 3 Grafik Hasil Model Loss 48