UNIVERSITAS
GADJAH MADA
Prediksi jumlah likes pada unggahan gambar dengan caption multibahasa di Instagram menggunakan mBERT dan InceptionResnetV2
Muhammad Cahaya Saputra, Dr. Agfianto Eko Putra, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN UTAMA
HALAMAN PENGESAHAN iii
PERNYATAAN BEBAS PLAGIASI iv
PRAKATA v
DAFTAR ISI vi
DAFTAR GAMBAR ix
DAFTAR TABEL xii
INTISARI xiii
ABSTRACT xiv
BAB I PENDAHULUAN 1
1.1 Latar Belakang Masalah 1
1.2 Rumusan Masalah 3
1.3 Batasan Masalah 3
1.4 Tujuan 3
1.5 Manfaat 4
BAB II TINJAUAN PUSTAKA 5
2.1. Prediksi Engagement Media Sosial 5
2.2. Pendekatan Unimodal berbasis Machine Learning 5
2.3. Pendekatan Multimodal berbasis konten Visual dan Teks 6
2.4. Integrasi Fitur Metadata Sosial dan Struktur jaringan 7
2.5. Attention Mechanism dan Permodelan Representasi Kontekstual 8
2.6. Prediksi Engagement dalam konteks multibahasa 9
2.7. Lingkup dan kontribusi penelitian 10
BAB III LANDASAN TEORI 15
3.1 Sosial Media *Engagement* 15
3.2 Natural Language Preprocessing 17
3.3 Bidirectional Encoder Representations from Transformers (BERT) 18
3.3.1. Representasi Input pada BERT 19
3.3.2. Mekanisme Self-attention 20
3.3.3. Pretraining BERT 22
3.3.4. Multilingual BERT 24

UNIVERSITAS
GADJAH MADA
Prediksi jumlah likes pada unggahan gambar dengan caption multibahasa di Instagram menggunakan mBERT dan InceptionResnetV2
Muhammad Cahaya Saputra, Dr. Agfianto Eko Putra, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.4 Computer Vision 24
3.5 Convolutional Neural Network 26
3.6 Inception-ResNet V2 29
3.7 Mekanisme Proyeksi 31
3.8 Loss Function 32
3.8.1. Main Loss (Mean Square Error) 33
3.8.2. Auxiliary Loss 33
3.8.3. Total Loss 34
3.9. Metrik Evaluasi 35
3.9.1. Mean Absolute Error (MAE) 35
3.9.2. Root Mean Square Error (RMSE) 35
3.9.3. Koefisien Determinasi (R^2) 36
BAB IV METODOLOGI PENELITIAN 37
4.1 Alur Rencana Penelitian 37
4.1.1 Dataset 39
4.1.2 Eksplorasi analisis data 40
4.1.3 Data Splitting 41
4.1.4 Prapemrosesan Data 42
4.1.5 Feature Engineering 42
4.1.6 Modeling 47
4.1.7 Eksperimen dan pelatihan model 55
4.2 Strategi Pelatihan dan Eksperimen 55
4.3 Alat & Bahan 61
BAB V IMPLEMENTASI 62
5.1. Analisa & Eksplorasi data 62
5.2. Prapemrosesan Data 74
5.3. Ekstraksi Fitur gambar dan *caption* 82
5.4. Implementasi Model 86
5.5. Evaluasi Peforma 91

UNIVERSITAS
GADJAH MADA
Prediksi jumlah likes pada unggahan gambar dengan caption multibahasa di Instagram menggunakan mBERT dan InceptionResnetV2
Muhammad Cahaya Saputra, Dr. Agfianto Eko Putra, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

BAB VI HASIL DAN PEMBAHASAN 96
6.1. Dataset Overview 96
6.2. Eksplorasi Analisis data 98
6.2.1. Analisis Univariate 98
6.2.2. Analisis korelasi data numerik 100
6.2.3. Analisis *Caption* 101
6.2.4. Analisis Gambar 105
6.3. Perbaikan data numerik 106
6.3.1. Data Cleaning 106
6.3.2. Data Transform 107
6.4. Analisis Bivariat 110
6.5. Hasil Feature engineering 112
6.5.1. *Engagement* Rate Komentar 112
6.5.2. Rasio Class komentar 113
6.6. *Hasil Feature Selection metadata* 114
6.7. Permodelan 115
6.7.1. Pembagian Dataset 115
6.7.2. Hasil Eksperimen model *Multimodal* 116
6.7.3. Hyperparameter Tunning 119
6.8. Studi Ablasi dan Evaluasi 120
6.9. Hasil pengujian *inference* model 123
6.9.1. Pengujian pada data latih 123
6.9.2. *Inference* pada unggahan asli 126
6.9.3. Implikasi dan Keterbatasan 130
BAB VII KESIMPULAN DAN SARAN 132
7.1. Kesimpulan 132
7.2. Saran 133
DAFTAR PUSTAKA 134
LAMPIRAN