

## **TUGAS AKHIR**

### **ANALISIS SENTIMEN KOMENTAR *INSTAGRAM REELS* TENTANG PENANGANAN BANJIR DI SUMATRA MENGUNAKAN PEMBOBOTAN TF-IDF DAN METODE *MULTINOMIAL NAIVE BAYES***



**FARREL OCTANOV RADISTHA GABRIELLE**

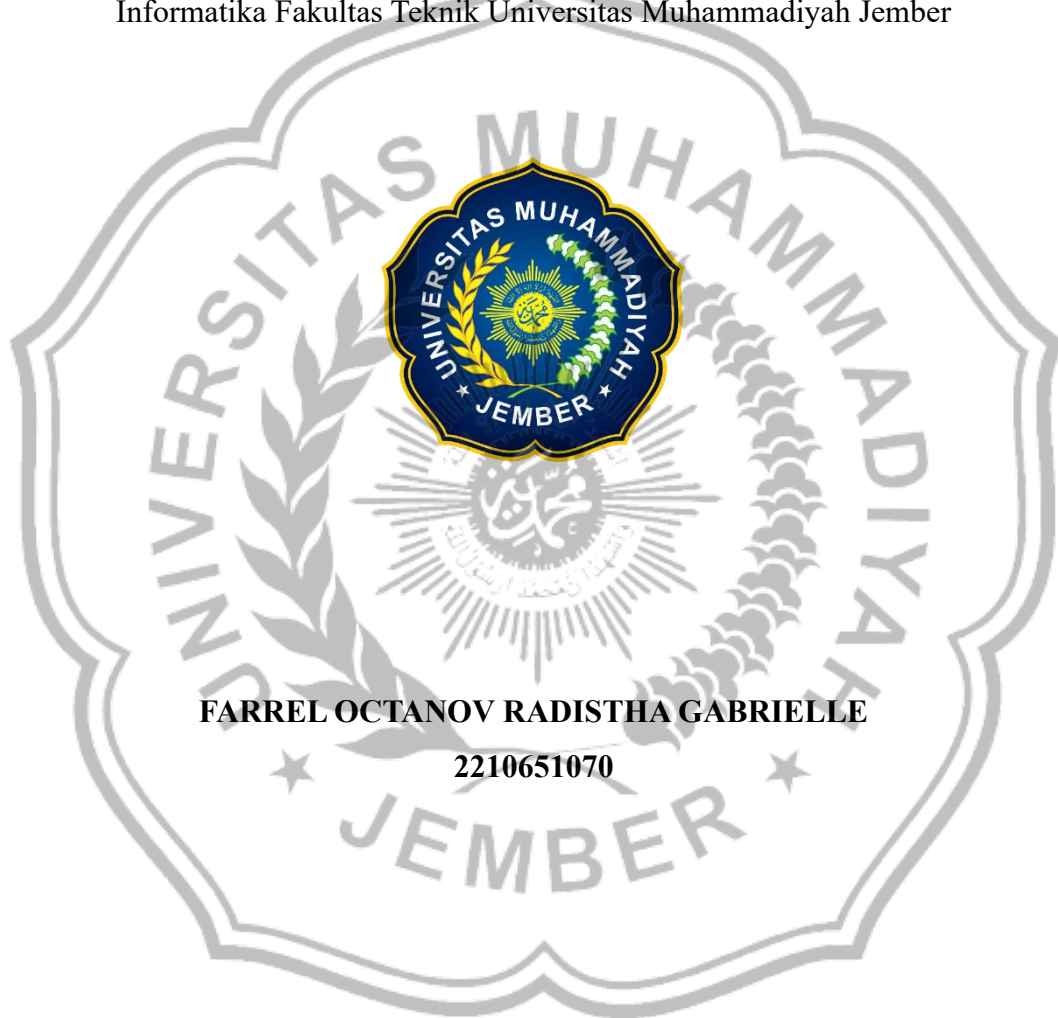
**2210651070**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS MUHAMMADIYAH JEMBER  
2026**

## **TUGAS AKHIR**

### **ANALISIS SENTIMEN KOMENTAR *INSTAGRAM REELS* TENTANG PENANGANAN BANJIR DI SUMATRA MENGUNAKAN PEMBOBOTAN TF-IDF DAN METODE *MULTINOMIAL NAIVE BAYES***

Diajukan sebagai salah satu syarat untuk kelulusan Strata Satu (S-1) Prodi Teknik  
Informatika Fakultas Teknik Universitas Muhammadiyah Jember



**FARREL OCTANOV RADISTHA GABRIELLE**

**2210651070**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS MUHAMMADIYAH JEMBER  
2026**

## HALAMAN PERSETUJUAN

### **ANALISIS SENTIMEN KOMENTAR *INSTAGRAM REELS* TENTANG PENANGANAN BANJIR DI SUMATRA MENGGUNAKAN PEMBOBOTAN TF-IDF DAN METODE *MULTINOMIAL NAIVE BAYES***

**Farrel Octanov Radistha Gabrielle**

**2210651070**

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom) di Universitas Muhammadiyah Jember

Disetujui oleh,

**Dosen Pembimbing I**

**Dosen Pembimbing II**

  
**DUDI IRAWAN S.T., M.Kom**  
**NIDN. 0730037703**

  
**GINANJAR ABDURRAHMAN S.Si., M.Pd.**  
**NIDN. 0714078704**

**HALAMAN PENGESAHAN**

**ANALISIS SENTIMEN KOMENTAR *INSTAGRAM REELS* TENTANG  
PENANGANAN BANJIR DI SUMATRA MENGGUNAKAN  
PEMBOBOTAN TF-IDF DAN METODE *MULTINOMIAL NAIVE BAYES***

**Farrel Octanov Radistha Gabrielle**

**2210651070**

Telah mempertanggungjawabkan Laporan Tugas Akhir pada sidang Tugas Akhir  
Tanggal 28 Agustus 2026 sebagai salah satu syarat kelulusan dan mendapatkan  
gelar Sarjana Komputer (S.Kom)

Disetujui oleh,

**Dosen Penguji I**



**Zainul Arifin S.Si., M.Kom**  
**NIDN. 0714078102**

**Dosen Pembimbing 1**



**Dudi Irawan S.T., M.Kom**  
**NIDN. 0730037703**

**Dosen Penguji II**



**Lintang Setyo Kurniawati S.Kom., M.Pd**  
**NIDN. 0724089204**

**Dosen Pembimbing II**



**Ginanjar Abdurrahman S.Si., M.Pd.**  
**NIDN. 0714078704**

**Mengesahkan,  
Dekan Fakultas Teknik**



**Prof. Dr. H. M. Mantar, S.T., M.T., IPM.**  
**NIDN. 0010067301**

**Mengetahui,  
Ketua Program Studi  
Teknik Informatika**



**Rosita Yanuanti, S.Kom., M.Cs.**  
**NIDN. 0629018601**

## MOTTO

“Sesungguhnya Allah menyukai orang-orang yang bekerja dengan tekun dan sungguh-sungguh.”

**(QS. At-Taubah: 105)**

“Sebaik-baik manusia adalah yang paling bermanfaat bagi manusia lainnya.”

**(Hadis Riwayat Ahmad)**

*“Knowledge is a seed that grows through patience, dedication, and endless learning.”*

**(Nahida)**

“Tidak ada mimpi yang gagal, yang ada hanya mimpi yang tertunda. Cuma sekiranya teman-teman merasa gagal dalam mencapai mimpi, jangan khawatir, mimpi-mimpi lain bisa diciptakan”

**(Windah Basudara)**

“Mungkin.. Dari banyaknya manusia di dunia ini, yang paling pantas untuk dinasihati adalah diri kita sendiri”

**(Penulis)**

## HALAMAN PERNYATAAN ORISINALITAS

Saya yang bertanda tangan di bawah ini:

Nama : Farrel Octanov R.G.  
NIM : 2210651070  
Program Studi : Teknik Informatika  
Judul : Analisis Sentimen Komentar *Instagram Reels*  
Tentang Penanganan Banjir Di Sumatra  
Menggunakan Pembobotan TF-IDF Dan Metode  
*Multinomial Naive Bayes*

Dengan ini menyatakan bahwa karya ilmiah yang saya susun merupakan hasil pemikiran, penelitian, dan penyusunan saya sendiri. Setiap sumber informasi yang saya gunakan atau kutip dari karya pihak lain, baik berupa publikasi ilmiah, artikel, buku, maupun sumber lainnya, telah dicantumkan dalam daftar pustaka sesuai dengan kaidah penulisan ilmiah.

Dalam proses penyusunan Tugas Akhir ini, saya memanfaatkan bantuan teknologi kecerdasan buatan (*\*Artificial Intelligence\*/AI*), seperti *ChatGPT*, *Grammarly*, atau alat AI lainnya, hanya untuk keperluan sebagai berikut:

1. Membantu penyusunan tata bahasa, parafrasa, dan penyempurnaan redaksi.
2. Memberikan referensi ide awal atau kerangka pemikiran yang kemudian dikembangkan dan disusun kembali secara mandiri.
3. Membantu memeriksa konsistensi format, ejaan, dan tata penulisan.

Saya menyatakan bahwa tidak terdapat bagian dalam karya ilmiah ini yang dibuat sepenuhnya oleh AI tanpa adanya proses pemikiran dan pengembangan dari saya sendiri. Seluruh isi dan kebenaran karya ilmiah ini menjadi tanggung jawab saya sebagai penyusun.

Apabila di kemudian hari ditemukan bahwa pernyataan ini tidak benar, saya bersedia menerima konsekuensi sesuai dengan peraturan yang berlaku di Universitas Muhammadiyah Jember.

Jember, September 2026  
Yang membuat pernyataan,



Farrel Octanov R.G.  
2210651070

## HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah Subhaanahu wa Ta'ala atas segala rahmat, kemudahan, dan pertolongan-Nya, serta rasa terima kasih kepada semua pihak yang telah memberikan doa, dukungan, dan bimbingan, karya skripsi ini penulis persembahkan kepada.

1. Allah Subhaanahu wa Ta'ala, atas segala limpahan rahmat, hidayah, dan kekuatan yang diberikan sehingga dapat menyelesaikan penyusunan skripsi ini dengan baik.
2. Kedua orang tua tercinta, Bapak Taufik dan Ibu Fenny yang senantiasa memberikan doa, kasih sayang, dukungan, dan semangat tanpa henti selama proses pendidikan hingga penyelesaian skripsi ini.
3. Keluarga besar Penulis yang selalu memberi dukungan kepada penulis, serta Nenek Suratmi dan Uti Lilik Sudarini yang selalu memberikan doa kepada penulis.
4. Seluruh dosen Program Studi Teknik Informatika yang telah memberikan ilmu, pengalaman, dan bimbingan selama masa perkuliahan.
5. Bapak Dudi Irawan, S.T., M.Kom. dan Bapak Ginanjar Abdurrahman, S.Si., M.Pd. selaku dosen pembimbing yang telah memberikan arahan, bimbingan, serta masukan selama proses penyusunan skripsi ini.
6. Bapak Zainul Arifin, S.Si., M.Kom. dan Ibu Lintang Setyo Kurniawati, S.Kom., M.Pd. selaku dosen penguji yang telah memberikan kritik dan saran yang membangun untuk penyempurnaan skripsi ini.
7. Sahabat saya, Bayu F.P., yang selalu memberikan dukungan, semangat, serta menghadirkan tawa dan kebahagiaan.
8. Teman-teman TI Kelas B 22 yang telah memberikan dukungan, kebersamaan, dan motivasi selama masa perkuliahan hingga penyelesaian skripsi.
9. Ibu Helmi Erfani, S.Pd. selaku ahli bahasa dan validator sentimen yang telah membantu proses validasi pelabelan sentimen pada penelitian ini.

Semoga skripsi ini dapat memberikan manfaat dan menjadi langkah awal menuju perjalanan yang lebih baik.

## KATA PENGANTAR

Penulis mengucapkan puji syukur ke hadirat Allah Subhanahu wa Ta'ala atas limpahan rahmat dan karunia-Nya sehingga proposal skripsi dengan judul “**Analisis Sentimen Komentar *Instagram Reels* tentang Penanganan Banjir di Sumatra Menggunakan Pembobotan TF-IDF dan Metode *Multinomial Naive Bayes***” dapat disusun.

Tugas Akhir ini disusun sebagai langkah awal dalam pelaksanaan penelitian yang berfokus pada pemanfaatan *platform* media sosial berbasis kecerdasan buatan. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap penanganan bencana banjir di Sumatra melalui data yang diperoleh dari *Instagram Reels*, serta merancang penerapan algoritma dalam proses klasifikasi sentimen.

Dalam penyusunannya, penulis berupaya menyajikan latar belakang, rumusan masalah, tujuan penelitian, serta metodologi yang akan digunakan secara sistematis dan informatif. Penulis menyadari bahwa dalam proses penyusunan proposal ini masih terdapat keterbatasan, baik dari segi referensi, perumusan metode, maupun teknis penulisan.

Jember, 2026

Penulis

## DAFTAR ISI

|                                                                 |             |
|-----------------------------------------------------------------|-------------|
| <b>TUGAS AKHIR.....</b>                                         | <b>i</b>    |
| <b>HALAMAN PERSETUJUAN .....</b>                                | <b>ii</b>   |
| <b>HALAMAN PENGESAHAN.....</b>                                  | <b>iii</b>  |
| <b>MOTTO .....</b>                                              | <b>iv</b>   |
| <b>HALAMAN PERNYATAAN ORISINALITAS.....</b>                     | <b>v</b>    |
| <b>HALAMAN PERSEMBAHAN .....</b>                                | <b>vi</b>   |
| <b>ABSTRAK .....</b>                                            | <b>vii</b>  |
| <b>ABSTRACT .....</b>                                           | <b>viii</b> |
| <b>KATA PENGANTAR.....</b>                                      | <b>ix</b>   |
| <b>DAFTAR ISI.....</b>                                          | <b>x</b>    |
| <b>DAFTAR GAMBAR.....</b>                                       | <b>xiv</b>  |
| <b>DAFTAR TABEL .....</b>                                       | <b>xv</b>   |
| <b>BAB I PENDAHULUAN.....</b>                                   | <b>1</b>    |
| 1.1    Latar Belakang .....                                     | 1           |
| 1.2    Rumusan Masalah .....                                    | 4           |
| 1.3    Tujuan Penelitian.....                                   | 4           |
| 1.4    Manfaat penelitian.....                                  | 4           |
| 1.5    Batasan Penelitian .....                                 | 5           |
| <b>BAB II TINJAUAN PUSTAKA.....</b>                             | <b>6</b>    |
| 2.1    Analisis Sentimen .....                                  | 6           |
| 2.2    Media Sosial Sebagai Sumber Data .....                   | 6           |
| 2.3    Flowchart .....                                          | 7           |
| 2.4    Python .....                                             | 8           |
| 2.5    Preprocessing .....                                      | 8           |
| 2.5.1    Text Cleaning.....                                     | 9           |
| 2.5.2    Case Folding .....                                     | 9           |
| 2.5.3    Tokenizing.....                                        | 9           |
| 2.5.4    Normalization.....                                     | 10          |
| 2.5.5    Stopword Removal.....                                  | 10          |
| 2.5.6    Stemming .....                                         | 10          |
| 2.6    TF-IDF (Term Frequency-Inverse Document Frequency) ..... | 10          |
| 2.6.1    Term Frequency.....                                    | 11          |
| 2.6.2    Inverse Document Frequency.....                        | 11          |

|                                            |                                                         |           |
|--------------------------------------------|---------------------------------------------------------|-----------|
| 2.7                                        | <i>Data Split</i> .....                                 | 12        |
| 2.8                                        | <i>Unseen data</i> .....                                | 12        |
| 2.9                                        | <i>Random Over Sampling (ROS)</i> .....                 | 12        |
| 2.10                                       | Metode Klasifikasi <i>Multinomial Naïve Bayes</i> ..... | 14        |
| 2.11                                       | <i>Confusion Matrix</i> .....                           | 17        |
| 2.12                                       | Penelitian Terdahulu.....                               | 18        |
| <b>BAB III METODOLOGI PENELITIAN .....</b> |                                                         | <b>23</b> |
| 3.1                                        | Tahapan Penelitian. ....                                | 23        |
| 3.2                                        | Pengumpulan Data .....                                  | 25        |
| 3.2.1                                      | Sumber data.....                                        | 25        |
| 3.2.2                                      | Tahapan Pengambilan Data.....                           | 25        |
| 3.2.3                                      | Pelabelan Data.....                                     | 26        |
| 3.3                                        | <i>Praprocessing Data</i> .....                         | 26        |
| 3.3.1                                      | <i>Text Cleaning</i> .....                              | 26        |
| 3.3.2                                      | <i>Case Folding</i> .....                               | 28        |
| 3.3.3                                      | <i>Tokenizing</i> .....                                 | 29        |
| 3.3.4                                      | <i>Normalization</i> .....                              | 30        |
| 3.3.5                                      | <i>Stopword Removal</i> .....                           | 32        |
| 3.3.6                                      | <i>Stemming</i> .....                                   | 33        |
| 3.4                                        | Pembobotan TF-IDF. ....                                 | 34        |
| 3.5                                        | <i>Data Split</i> .....                                 | 37        |
| 3.6                                        | <i>Random Over Sampling</i> . ....                      | 37        |
| 3.7                                        | Metode Klasifikasi <i>Multinomial Naïve Bayes</i> ..... | 40        |
| 3.8                                        | <i>Evaluation</i> .....                                 | 44        |
| 3.8.1                                      | <i>Accuracy</i> .....                                   | 45        |
| 3.8.2                                      | <i>Precision</i> .....                                  | 45        |
| 3.8.3                                      | <i>Recall</i> .....                                     | 45        |
| 3.8.4                                      | <i>F1-Score</i> .....                                   | 46        |
| <b>BAB IV HASIL DAN PEMBAHASAN .....</b>   |                                                         | <b>47</b> |
| 4.1                                        | Deskripsi Dataset .....                                 | 47        |
| 4.1.1                                      | Proses Pengumpulan Data.....                            | 47        |
| 4.1.2                                      | Struktur dan Distribusi Kelas Sentimen.....             | 48        |
| 4.1.3                                      | Analisis Dataset.....                                   | 49        |
| 4.2                                        | Hasil <i>Preprocessing Data</i> .....                   | 50        |
| 4.2.1                                      | <i>Text Cleaning</i> .....                              | 50        |
| 4.2.2                                      | <i>Case Folding</i> .....                               | 50        |

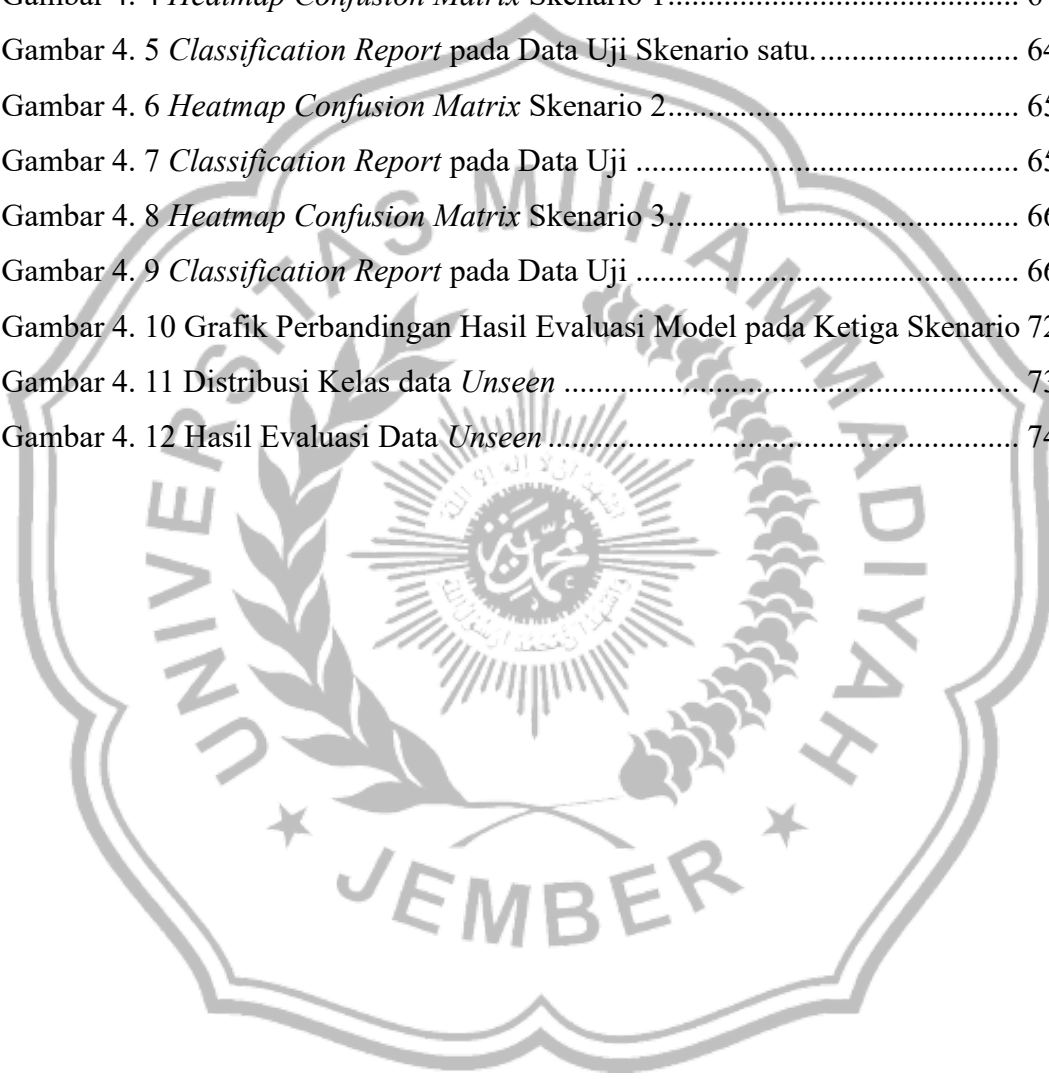
|                                         |                                            |           |
|-----------------------------------------|--------------------------------------------|-----------|
| 4.2.3                                   | <i>Tokenizing</i> .....                    | 51        |
| 4.2.4                                   | <i>Normalization</i> .....                 | 51        |
| 4.2.5                                   | <i>Stopword Removal</i> .....              | 52        |
| 4.2.6                                   | <i>Stemming</i> .....                      | 52        |
| 4.2.7                                   | Hasil Akhir <i>Preprocessing</i> .....     | 53        |
| 4.3                                     | Pembobotan TF-IDF .....                    | 54        |
| 4.3.1                                   | Hasil Representasi TF-IDF .....            | 54        |
| 4.3.2                                   | Distribusi Fitur .....                     | 55        |
| 4.4                                     | Pembagian Data .....                       | 56        |
| 4.4.1                                   | Hasil Pembagian Data Latih, dan Uji.....   | 56        |
| 4.4.2                                   | Distribusi Kelas per Split.....            | 57        |
| 4.5                                     | <i>Random Over Sampling</i> .....          | 58        |
| 4.5.1                                   | Distribusi Sebelum ROS .....               | 59        |
| 4.5.2                                   | Distribusi Setelah ROS .....               | 59        |
| 4.6                                     | <i>Multinomial Naïve Bayes</i> .....       | 60        |
| 4.6.1                                   | Proses Pelatihan .....                     | 61        |
| 4.6.2                                   | Hasil Pengujian .....                      | 61        |
| 4.7                                     | Evaluasi Model.....                        | 63        |
| 4.7.1                                   | <i>Confusion Matrix</i> .....              | 64        |
| 4.7.2                                   | <i>Accuracy</i> .....                      | 67        |
| 4.7.3                                   | <i>Precision</i> .....                     | 67        |
| 4.7.4                                   | <i>Recall</i> .....                        | 68        |
| 4.7.5                                   | <i>F1-Score</i> .....                      | 69        |
| 4.8                                     | Perbandingan Skenario.....                 | 70        |
| 4.8.1                                   | Skenario 1 – <i>Split data</i> 70:30 ..... | 71        |
| 4.8.2                                   | Skenario 2 – <i>Split data</i> 80:20 ..... | 71        |
| 4.8.3                                   | Skenario 3 – <i>Split data</i> 90:10 ..... | 71        |
| 4.8.4                                   | Analisis Perbandingan Ketiga Skenario..... | 71        |
| 4.9                                     | Pengujian Data <i>Unseen</i> .....         | 72        |
| 4.9.1                                   | Karakteristik Data <i>Unseen</i> .....     | 72        |
| 4.9.2                                   | Hasil Evaluasi Data <i>Unseen</i> .....    | 73        |
| 4.10                                    | Pembahasan.....                            | 74        |
| 4.10.1                                  | Kecenderungan Sentimen Masyarakat.....     | 74        |
| 4.10.2                                  | Kinerja Model Secara Keseluruhan .....     | 75        |
| <b>BAB V KESIMPULAN DAN SARAN .....</b> |                                            | <b>78</b> |
| 5.1                                     | Kesimpulan .....                           | 78        |

|                                               |            |           |
|-----------------------------------------------|------------|-----------|
| 5.2                                           | Saran..... | 79        |
| <b>DAFTAR PUSTAKA.....</b>                    |            | <b>81</b> |
| <b>LAMPIRAN.....</b>                          |            | <b>85</b> |
| Lampiran 1. Dokumentasi Data Penelitian ..... |            | 85        |



## DAFTAR GAMBAR

|                                                                            |    |
|----------------------------------------------------------------------------|----|
| Gambar 3. 1 <i>Flowchart</i> Proses Penelitian Analisis Sentimen.....      | 24 |
| Gambar 4. 1 Perbandingan Skenario Pembagian dataset.....                   | 57 |
| Gambar 4. 2 Distribusi Sebelum dan sesudah ROS .....                       | 60 |
| Gambar 4. 3 Grafik akurasi data latih .....                                | 61 |
| Gambar 4. 4 <i>Heatmap Confusion Matrix</i> Skenario 1.....                | 64 |
| Gambar 4. 5 <i>Classification Report</i> pada Data Uji Skenario satu.....  | 64 |
| Gambar 4. 6 <i>Heatmap Confusion Matrix</i> Skenario 2.....                | 65 |
| Gambar 4. 7 <i>Classification Report</i> pada Data Uji .....               | 65 |
| Gambar 4. 8 <i>Heatmap Confusion Matrix</i> Skenario 3.....                | 66 |
| Gambar 4. 9 <i>Classification Report</i> pada Data Uji .....               | 66 |
| Gambar 4. 10 Grafik Perbandingan Hasil Evaluasi Model pada Ketiga Skenario | 72 |
| Gambar 4. 11 Distribusi Kelas data <i>Unseen</i> .....                     | 73 |
| Gambar 4. 12 Hasil Evaluasi Data <i>Unseen</i> .....                       | 74 |



## DAFTAR TABEL

|                                                                                     |    |
|-------------------------------------------------------------------------------------|----|
| Tabel 2. 1 Contoh <i>Confusion Matrix</i> Klasifikasi Sentimen Tiga Kelas .....     | 17 |
| Tabel 2. 2 Penelitian Terdahulu .....                                               | 18 |
| Tabel 3. 1 Contoh Hasil Proses <i>Text Cleaning</i> .....                           | 27 |
| Tabel 3. 2 Contoh Hasil Proses <i>Case Folding</i> .....                            | 28 |
| Tabel 3. 3 Contoh Hasil Proses <i>Tokenizing</i> .....                              | 29 |
| Tabel 3. 4 Contoh Hasil Proses <i>Normalization</i> .....                           | 31 |
| Tabel 3. 5 Contoh Hasil Proses <i>Stopword Removal</i> .....                        | 32 |
| Tabel 3. 6 Contoh Hasil Proses <i>Stemming</i> .....                                | 33 |
| Tabel 3. 7 Hasil Pembobotan <i>Term Frequency</i> (TF) .....                        | 35 |
| Tabel 3. 8 Hasil Perhitungan <i>Inverse Document Frequency</i> (IDF) .....          | 36 |
| Tabel 3. 9 Skenario <i>Split Data</i> .....                                         | 37 |
| Tabel 3. 10 <i>Confusion Matrix</i> Hasil Evaluasi Model .....                      | 43 |
| Tabel 4. 1 Distribusi Label Sentimen Hasil Pelabelan <i>IndoBERT</i> dan Ahli ..... | 49 |
| Tabel 4. 2 Hasil <i>Preprocessing</i> Data Komentar .....                           | 53 |
| Tabel 4. 3 Parameter Ekstraksi Fitur TF-IDF .....                                   | 54 |
| Tabel 4. 4 Hasil Pembobotan TF-IDF .....                                            | 55 |
| Tabel 4. 5 Perbandingan Skenario Pembagian Dataset .....                            | 56 |
| Tabel 4. 6 <i>Split Data</i> Skenario 1 .....                                       | 58 |
| Tabel 4. 7 <i>Split Data</i> Skenario 2 .....                                       | 58 |
| Tabel 4. 8 <i>Split Data</i> Skenario 3 .....                                       | 58 |
| Tabel 4. 9 Distribusi Sentimen Sebelum ROS .....                                    | 59 |
| Tabel 4. 10 Distribusi Data Setelah ROS .....                                       | 59 |
| Tabel 4. 11 Hasil Akurasi Pelatihan pada Tiga Skenario .....                        | 61 |
| Tabel 4. 12 Hasil Akurasi Pengujian pada Tiga Skenario .....                        | 61 |
| Tabel 4. 13 Hasil Prediksi Model pada Skenario 1 (70:30) .....                      | 62 |
| Tabel 4. 14 Hasil Prediksi Model pada Skenario 2 (80:20) .....                      | 62 |
| Tabel 4. 15 Hasil Prediksi Model pada Skenario 3 (90:10) .....                      | 63 |
| Tabel 4. 16 Hasil Evaluasi Model Berdasarkan Skenario Pembagian Data .....          | 63 |
| Tabel 4. 17 <i>Confusion Matrix</i> Skenario 1 .....                                | 64 |
| Tabel 4. 18 <i>Confusion Matrix</i> Skenario 2 .....                                | 65 |

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| Tabel 4. 19 <i>Confusion Matrix</i> Skenario 3 .....                  | 66 |
| Tabel 4. 20 Ringkasan Hasil Evaluasi Model pada Ketiga Skenario ..... | 71 |
| Tabel 4. 21 Distribusi Kelas Data <i>Unseen</i> .....                 | 73 |
| Tabel 4. 22 Hasil Evaluasi Data <i>Unseen</i> .....                   | 73 |

