

Universitas  
**Esa Unggul**


Universi  
**Esa**

## DAFTAR ISI

|   |      |
|---|------|
| HALAMAN PERNYATAAN KEASLIAN .....               | ii   |
| HALAMAN PENGESAHAN TUGAS AKHIR.....             | iii  |
| HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH..... | iv   |
| KATA PENGANTAR .....                            | v    |
| ABSTRAK .....                                   | vi   |
| ABSTRACT .....                                  | vii  |
| DAFTAR ISI .....                                | viii |
| DAFTAR GAMBAR.....                              | xii  |
| DAFTAR TABEL .....                              | xiii |
| BAB I.....                                      | 1    |
| PENDAHULUAN .....                               | 1    |
| 1.1 Latar Belakang .....                        | 1    |
| 1.2 Identifikasi Masalah.....                   | 4    |
| 1.3 Tujuan Tugas Akhir.....                     | 4    |
| 1.4 Manfaat Tugas Akhir.....                    | 4    |
| 1.5 Batasan Masalah.....                        | 5    |
| 1.6 Lingkup Tugas Akhir.....                    | 5    |
| 1.7 Kerangka Berpikir .....                     | 6    |
| 1.8 Sistematika Penulisan .....                 | 7    |
| BAB II.....                                     | 8    |
| LANDASAN TEORI.....                             | 8    |

|                        |  |    |
|------------------------|--|----|
| 2.1                    | Analisis Sentimen.....                             | 8  |
| 2.2                    | Python.....  | 9  |
| 2.3                    | Natural Language ToolKit (NLTK).....               | 9  |
| 2.4                    | Crawling Data.....                                 | 10 |
| 2.5                    | Text Mining.....                                   | 11 |
| 2.6                    | Preprocessing.....                                 | 11 |
| 2.7                    | Sastrawi / PySastrawi.....                         | 13 |
| 2.8                    | Frequency based.....                               | 13 |
| 2.9                    | Random forest.....                                 | 14 |
| 2.10                   | Metode Naïve Bayes.....                            | 15 |
| 2.11                   | Lexicon Based.....                                 | 17 |
| 2.12                   | Twitter.....                                       | 18 |
| 2.13                   | TikTok Shop.....                                   | 19 |
| 2.14                   | Metrik Evaluasi.....                               | 20 |
| 2.15                   | Studi Literatur.....                               | 21 |
| BAB III.....           |  | 25 |
| METODE PENELITIAN..... |  | 25 |
| 3.1                    | Tahapan Penelitian.....                            | 25 |
| 3.2                    | Analisa Permasalahan.....                          | 26 |
| 3.3                    | Teknik Pengumpulan Data.....                       | 26 |
| 3.4                    | Preprocessing.....                                 | 27 |
| 3.5                    | Lexicon Based.....                                 | 30 |
| 3.6                    | Membangun Model Naïve Bayes dan Random Forest..... | 30 |
| 3.7                    | Visualisasi.....                                   | 31 |
| 3.8                    | Evaluasi.....                                      | 31 |
| 3.9                    | Solusi yang Diusulkan.....                         | 32 |
| 3.10                   | Alat Bantu Penelitian.....                         | 32 |
| 3.11                   | Rencana Penelitian.....                            | 33 |
| 3.1.1                  | Tempat Penelitian.....                             | 33 |
| 3.1.2                  | Waktu Penelitian.....                              | 33 |
| BAB IV.....            |  | 35 |

|  |    |
|--|----|
| PEMBAHASAN.....  | 35 |
| 4.1    Crawlling Data .....                                      | 35 |
| 4.2    Pre-Processing.....                                       | 37 |
| 4.2.1    Case Folding.....                                       | 38 |
| 4.2.2    Cleaning .....  | 38 |
| 4.2.3    Tokenize.....   | 41 |
| 4.2.4    Normalization.....                                      | 41 |
| 4.2.5    Stopwords.....  | 43 |
| 4.2.1.    Stemming .....   | 45 |
| 4.3    Labeling .....  | 47 |
| 4.3.1    Menyiapkan Kamus.....                                   | 47 |
| 4.3.2    Proses Analisis Sentimen.....                           | 47 |
| 4.4    Visualiasi.....   | 49 |
| 4.4.1    Diagram.....  | 50 |
| 4.4.2    Wordcloud.....  | 50 |
| 4.5    Modeling.....   | 52 |
| 4.5.1    Algoritma Naïve Bayes.....                              | 52 |
| 4.5.2.    Algoritma Random Forest.....                           | 59 |
| 4.4.1.....   | 63 |
| 4.6    Evaluasi.....   | 66 |
| 4.6.1.    Confusion Matrix Naïve Bayes – TF IDF.....             | 66 |
| 4.6.2.    Confusion Matrix Naïve Bayes – Count Vectorize.....    | 68 |
| 4.6.3.    Confusion Matrix Random Forest – TF IDF .....          | 71 |
| 4.6.4.    Confusion Matrix Random Forest – Count Vectorizer..... | 74 |
| BAB V.....   | 78 |
| KESIMPULAN DAN SARAN .....                                       | 78 |
| 5.1.    Kesimpulan .....   | 78 |
| 5.2.    Saran .....  | 78 |
| DAFTAR PUSTAKA.....  | 80 |
| LAMPIRAN.....  | 84 |



|                            |    |
|----------------------------|----|
| DAFTAR RIWAYAT HIDUP ..... | 85 |
| Data Pribadi .....         | 85 |
| Pendidikan Formal .....    | 85 |
| Riwayat Pekerjaan.....     | 85 |

## DAFTAR GAMBAR

|  |    |
|--|----|
| Gambar 1 Kerangka Berpikir 1.....                          | 6  |
| Gambar 2. 1 Matrikulasi .....                              | 20 |
| Gambar 3. 1 Tahapan Penelitian.....                        | 25 |
| Gambar 3. 2 Alur Pengumpulan Data.....                     | 27 |
| Gambar 4. 1 Auth_Token Twitter.....                        | 35 |
| Gambar 4. 2 Input Auth_Token Twitter .....                 | 36 |
| Gambar 4. 3 Hasil Crawling Data.....                       | 36 |
| Gambar 4. 4 Proses Crawling Data Twitter .....             | 36 |
| Gambar 4. 5 Kolom full_text yang akan diproses.....        | 37 |
| Gambar 4. 6 Hasil Cleaning.....                            | 40 |
| Gambar 4. 7 Proses Cleaning.....                           | 40 |
| Gambar 4. 8 Proses Tokenize .....                          | 41 |
| Gambar 4. 9 kamus kata baku.....                           | 42 |
| Gambar 4. 10 kode membaca kamus kata baku pada folder..... | 42 |
| Gambar 4. 11 Proses Normalisasi.....                       | 42 |
| Gambar 4. 12 kode normalisasi .....                        | 43 |
| Gambar 4. 13 Pengaplikasian pada dokumen .....             | 43 |
| Gambar 4. 14 Import NLTK .....                             | 43 |
| Gambar 4. 15 Stopwords.....                                | 44 |
| Gambar 4. 16 List stopwords tambahan.....                  | 44 |
| Gambar 4. 17 Library Naive Bayes dan TF-IDF .....          | 53 |
| Gambar 4. 18 Pembagian Dataset .....                       | 54 |
| Gambar 4. 19 Hasil Pembagian Dataset.....                  | 54 |
| Gambar 4. 20 Kode Ekstraksi Fitur TF-IDF .....             | 55 |
| Gambar 4. 21 Pemilihan kode naive bayes .....              | 55 |
| Gambar 4. 22 Proses prediksi data test .....               | 56 |
| Gambar 4. 25 Matrix Naive Bayes - TF IDF .....             | 66 |
| Gambar 4. 26 Library Naive Bayes - Count Vectorizer.....   | 56 |
| Gambar 4. 27 Pembagian Dataset .....                       | 57 |
| Gambar 4. 28 Hasil Pembagian Dataset.....                  | 57 |
| Gambar 4. 29 Ekstraksi Fitur Count Vectorizer.....         | 58 |
| Gambar 4. 30 Pemilihan Modul Multinomial.....              | 58 |
| Gambar 4. 31 Proses Prediksi data testing.....             | 59 |
| Gambar 4. 34 Matrix Naive Bayes - Count Vectorizer.....    | 68 |

**DAFTAR TABEL**

|                                 |    |
|---------------------------------|----|
| Tabel 1 Studi Literatur .....   | 24 |
| Tabel 2 Proses Cleaning.....    | 28 |
| Tabel 3 Proses CaseFolding..... | 28 |
| Tabel 4 Proses StopWord.....    | 29 |
| Tabel 5 Proses Tokenisasi.....  | 29 |
| Tabel 6 Proses Stemming.....    | 30 |