

## DAFTAR ISI

|  |      |
|--|------|
| HALAMAN PENGESAHAN PEMBIMBING .....                          | ii   |
| HALAMAN PENGESAHAN PENGUJI.....                              | iii  |
| PERNYATAAN KARYA ASLI TUGAS AKHIR.....                       | iv   |
| PERNYATAAN BEBAS PLAGIASI.....                               | v    |
| ABSTRAK .....  | vi   |
| ABSTRACT .....   | vii  |
| KATA PENGANTAR.....  | viii |
| DAFTAR ISI .....   | ix   |
| DAFTAR TABEL .....   | xii  |
| DAFTAR GAMBAR.....   | xiii |
| DAFTAR PERSAMAAN .....                                       | xiv  |
| BAB I PENDAHULUAN .....                                      | 1    |
| 1.1 Latar Belakang.....                                      | 1    |
| 1.2 Rumusan Masalah.....                                     | 2    |
| 1.3 Batasan Masalah .....                                    | 2    |
| 1.4 Tujuan Penelitian .....                                  | 3    |
| 1.5 Manfaat Penelitian .....                                 | 3    |
| 1.6 Tahapan Penelitian.....                                  | 3    |
| 1.6.1 Metodologi Penelitian .....                            | 3    |
| 1.6.2 Metodologi Pengembangan Sistem.....                    | 4    |
| 1.6.3 Sistematika Penulisan .....                            | 5    |
| BAB II TINJAUAN LITERATUR .....                              | 6    |
| 2.1 Analisis Sentimen .....                                  | 6    |
| 2.2 BPJS Kesehatan .....                                     | 6    |
| 2.3 Twitter.....   | 6    |
| 2.4 Instagram .....  | 7    |
| 2.5 Preprocessing .....                                      | 7    |
| 2.5.1 Case Folding.....                                      | 7    |
| 2.5.2 Remove Emoji .....                                     | 7    |
| 2.5.3 Data Cleaning .....                                    | 7    |
| 2.5.4 Remove Repeation Character.....                        | 7    |
| 2.5.5 Tokenization .....                                     | 8    |
| 2.5.6 Word Normalization .....                               | 8    |
| 2.5.7 Stemming.....  | 8    |
| 2.5.8 Stopwords Removal.....                                 | 8    |
| 2.6 Term Frequency-Inverse Document Frequency (TF-IDF) ..... | 9    |

|   |   |           |
|---|---|-----------|
| 2.7   | Naïve Bayes Classifier.....   | 9         |
| 2.8   | K-Nearest Neighbor.....   | 11        |
| 2.9   | Confussion Matrix.....  | 11        |
| 2.10  | Penelitian Sebelumnya.....  | 13        |
| <b>BAB III METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM.....</b> |   | <b>21</b> |
| 3.1   | Metodologi Penelitan.....   | 21        |
| 3.1.1   | Pengambilan Data.....   | 22        |
| 3.1.2   | Pelabelan Data.....   | 23        |
| 3.1.3   | Preprocessing.....  | 24        |
| 3.1.4   | Flowchart Metode Naïve Bayes Classifier Menggunakan Skenario<br>Preprocessing I.....  | 33        |
| 3.1.5   | Flowchart Metode Naïve Bayes Classifier Menggunakan Skenario<br>Preprocessing II..... | 34        |
| 3.1.6   | Flowchart Metode K-Nearest Neighbor Menggunakan Skenario<br>Preprocessing I.....      | 35        |
| 3.1.7   | Flowchart Metode K-Nearest Neighbor Menggunakan Skenario<br>Preprocessing II.....     | 36        |
| 3.1.8   | Pembobotan Kata.....  | 36        |
| 3.1.9   | Klasifikasi Naïve Bayes Classifier.....   | 40        |
| 3.1.10  | Klasifikasi K-Nearest Neighbor.....   | 43        |
| 3.1.11  | Pengujian Confussion Matrix.....  | 47        |
| 3.2   | Metodologi Pengembangan Sistem.....   | 48        |
| 3.2.1   | Analisis Kebutuhan Sistem.....  | 48        |
| 3.2.2   | Perancangan Arsitektur Sistem.....  | 49        |
| 3.2.3   | Perancangan Proses.....   | 50        |
| 3.2.4   | Perancangan Desain Antarmuka.....   | 52        |
| 3.2.5   | Pengujian Sistem.....   | 55        |
| 3.2.6   | Operation and Maintenance.....  | 56        |
| <b>BAB IV HASIL DAN PEMBAHASAN.....</b>                           |   | <b>57</b> |
| 4.1   | Hasil Penelitian.....   | 57        |
| 4.1.1   | Pengambilan Data.....   | 57        |
| 4.1.2   | Preprocessing.....  | 58        |
| 4.1.3   | Split Data.....   | 61        |
| 4.1.4   | Pembobotan Kata dengan TF-IDF.....  | 61        |
| 4.1.5   | Klasifikasi Naïve Bayes Classifier.....   | 61        |
| 4.1.6   | Klasifikasi K-Nearest Neighbor.....   | 62        |

|                     |                           |    |
|---------------------|---------------------------|----|
| 4.1.7               | Implementasi Sistem ..... | 62 |
| 4.1.8               | Pengujian Sistem .....    | 67 |
| 4.1.9               | Pengujian Model.....      | 67 |
| 4.2                 | Pembahasan .....          | 73 |
| BAB V PENUTUP ..... |                           | 74 |
| 5.1                 | Kesimpulan .....          | 74 |
| 5.2                 | Saran .....               | 74 |
| DAFTAR PUSTAKA..... |                           | 75 |

## DAFTAR TABEL

|   |    |
|---|----|
| Tabel 2.1 Confusion Matrix .....  | 11 |
| Tabel 2.2 <i>State of The Art</i> .....   | 18 |
| Tabel 3.1 Hasil label manual .....  | 23 |
| Tabel 3.2 Contoh Penggunaan Case Folding.....   | 24 |
| Tabel 3.3 Contoh Penggunaan Remove Emoji.....   | 26 |
| Tabel 3.4 Contoh Penggunaan Data Cleaning .....                                       | 27 |
| Tabel 3.5 Contoh Penggunaan Remove Repetition Character .....                         | 28 |
| Tabel 3.6 Contoh Penggunaan Tokenization.....   | 29 |
| Tabel 3.7 Contoh Penggunaan Word Normalization.....                                   | 30 |
| Tabel 3.8 Contoh Penggunaan Stemming .....  | 31 |
| Tabel 3.9 Contoh Penggunaan Stopwords Removal .....                                   | 32 |
| Tabel 3.10 Contoh Hasil Preprocessing .....   | 38 |
| Tabel 3.11 Contoh Perhitungan TF .....  | 38 |
| Tabel 3.12 Contoh Hasil TF-IDF .....  | 39 |
| Tabel 3.13 Contoh Kasus Klasifikasi K-Nearest Neighbor.....                           | 44 |
| Tabel 3.14 Hasil TF-IDF Contoh Kasus Klasifikasi K-Nearest Neighbor.....              | 45 |
| Tabel 3.15 Hasil Perhitungan Euclidean Distance .....                                 | 46 |
| Tabel 3.16 Hasil Perhitungan Euclidean Distance .....                                 | 46 |
| Tabel 3.17 Pengujian Confusion Matrix Naïve Bayes Classifier Skenario I .....         | 47 |
| Tabel 3.18 Pengujian <i>Confusion Matrix Naïve Bayes Classifier</i> Skenario II ..... | 47 |
| Tabel 3.19 Pengujian <i>Confusion Matrix K-Nearest Neighbor</i> Skenario I.....       | 48 |
| Tabel 3.20 Pengujian <i>Confusion Matrix K-Nearest Neighbor</i> Skenario II .....     | 48 |
| Tabel 3.21 Kebutuhan Perangkat Keras .....  | 49 |
| Tabel 3.22 Kebutuhan Perangkat Lunak .....  | 49 |
| Tabel 4.1 Hasil Pengujian Sistem.....   | 67 |
| Tabel 4.2 Confusion Matrix Model Naïve Bayes Classifier Skenario I.....               | 68 |
| Tabel 4.3 <i>Confusion Matrix Model Naïve Bayes Classifier</i> Skenario II.....       | 69 |
| Tabel 4.4 <i>Confusion Matrix Model K-Nearest Neighbor</i> Skenario I .....           | 70 |
| Tabel 4.5 <i>Confusion Matrix Model K-Nearest Neighbor</i> Skenario II.....           | 71 |
| Tabel 4.6 Hasil pengujian.....  | 72 |

## DAFTAR GAMBAR

|   |    |
|---|----|
| Gambar 3.1 Langkah Penelitian .....                                     | 21 |
| Gambar 3.2 Flowchart Scraping Data Twitter.....                         | 22 |
| Gambar 3.3 Flowchart Scraping Data Instagram .....                      | 23 |
| Gambar 3.4 Flowchart Case Folding .....                                 | 25 |
| Gambar 3.5 Flowchart Remove Emoji .....                                 | 26 |
| Gambar 3.6 Flowchart Data Cleaning .....                                | 27 |
| Gambar 3.7 Flowchart Remove Repetition Character.....                   | 28 |
| Gambar 3.8 Flowchart Tokenization .....                                 | 29 |
| Gambar 3.9 Flowchart Word Normalization.....                            | 30 |
| Gambar 3.10 Flowchart Stemming.....                                     | 31 |
| Gambar 3.11 Flowchart Stopwords Removal .....                           | 32 |
| Gambar 3.12 Flowchart NBC Dengan Skenario Preprocessing I .....         | 33 |
| Gambar 3.13 Flowchart NBC Dengan Skenario Preprocessing II.....         | 34 |
| Gambar 3.14 Flowchart KNN Dengan Skenario Preprocessing I.....          | 35 |
| Gambar 3.15 Flowchart KNN Dengan Skenario Preprocessing II.....         | 36 |
| Gambar 3.16 Flowchart TF-IDF.....                                       | 37 |
| Gambar 3.17 Flowchart Data Latih Naïve Bayes Classifier .....           | 40 |
| Gambar 3.18 Flowchart Data Uji Naïve Bayes Classifier.....              | 42 |
| Gambar 3.19 Flowchart K-Nearest Neighbor .....                          | 43 |
| Gambar 3.20 Perancangan Arsitektur.....                                 | 50 |
| Gambar 3.21 DFD Level 0 .....   | 50 |
| Gambar 3.22 DFD Level 1 .....   | 51 |
| Gambar 3.23 DFD Level 2 .....   | 52 |
| Gambar 3.24 Halaman Beranda.....  | 52 |
| Gambar 3.25 Halaman Dataset Hasil Scrapping .....                       | 53 |
| Gambar 3.26 Halaman Dataset Hasil Preprosesing.....                     | 53 |
| Gambar 3.27 Halaman Input Komentar NBC & KNN .....                      | 54 |
| Gambar 3.28 Halaman Hasil Klasifikasi NBC & KNN .....                   | 54 |
| Gambar 3.29 Halaman Presentase Hasil Pengujian .....                    | 55 |
| Gambar 3.30 Halaman Grafik Hasil Pengujian .....                        | 55 |
| Gambar 4.1 Halaman Dashboard.....                                       | 63 |
| Gambar 4.2 Halaman Dataset Hasil Scrapping .....                        | 63 |
| Gambar 4.3 Halaman Dataset Hasil Preprocesssing .....                   | 64 |
| Gambar 4.4 Input Komentar Klasifikasi NBC & KNN .....                   | 64 |
| Gambar 4.5 Halaman Hasil Klasifikasi NBC & KNN .....                    | 65 |
| Gambar 4.6 Halaman Hasil Klasifikasi NBC & KNN (lanjutan).....          | 65 |
| Gambar 4.7 Halaman Penambahan Kamus Word Normalization (lanjutan) ..... | 66 |
| Gambar 4.8 Halaman Presentase Hasil Pengujian.....                      | 66 |
| Gambar 4.9 Halaman Grafik Hasil Pengujian .....                         | 67 |

## DAFTAR PERSAMAAN

|   |    |
|---|----|
| Persamaan 2.1 Bobot Term Frequency.....                       | 9  |
| Persamaan 2.2 Inverse Document Frequency.....                 | 9  |
| Persamaan 2.3 TF-IDF .....                                    | 9  |
| Persamaan 2.4 <i>Prosterior</i> .....                         | 10 |
| Persamaan 2.5 Klasifikasi <i>Naïve Bayes Classifier</i> ..... | 10 |
| Persamaan 2.6 Nilai <i>Evidence</i> .....                     | 10 |
| Persamaan 2.7 Nilai <i>Prior</i> .....                        | 10 |
| Persamaan 2.8 Nilai <i>Prosterior</i> .....                   | 10 |
| Persamaan 2.9 Metode <i>Euclidean Distance</i> .....          | 11 |
| Persamaan 2.10 Akurasi .....                                  | 12 |
| Persamaan 2.11 Presisi Kelas Positif.....                     | 12 |
| Persamaan 2.12 Presisi Kelas Negatif .....                    | 12 |
| Persamaan 2.13 Presisi Kelas Netral .....                     | 12 |
| Persamaan 2.14 Presisi .....                                  | 12 |
| Persamaan 2.15 Recall Kelas Positif .....                     | 13 |
| Persamaan 2.16 Recall Kelas Negatif .....                     | 13 |
| Persamaan 2.17 Recall Kelas Netral.....                       | 13 |
| Persamaan 2.18 Recall.....                                    | 13 |