

## DAFTAR ISI

|  |             |
|--|-------------|
| <b>HALAMAN PENGESAHAN .....</b>                | <b>i</b>    |
| <b>HALAMAN PERNYATAAN.....</b>                 | <b>ii</b>   |
| <b>PRAKATA .....</b>                           | <b>iii</b>  |
| <b>DAFTAR ISI.....</b>                         | <b>iv</b>   |
| <b>DAFTAR TABEL .....</b>                      | <b>vii</b>  |
| <b>DAFTAR GAMBAR.....</b>                      | <b>viii</b> |
| <b>INTISARI .....</b>                          | <b>x</b>    |
| <b>ABSTRACT .....</b>                          | <b>xi</b>   |
| <b>BAB I PENDAHULUAN.....</b>                  | <b>1</b>    |
| 1.1 Latar Belakang .....                       | 1           |
| 1.2 Rumusan Masalah .....                      | 5           |
| 1.3 Batasan Masalah.....                       | 5           |
| 1.4 Tujuan Penelitian.....                     | 6           |
| 1.5 Manfaat Penelitian.....                    | 6           |
| 1.6 Metodologi Penelitian .....                | 6           |
| 1.7 Sistematika Penulisan.....                 | 7           |
| <b>BAB II TINJAUAN PUSTAKA.....</b>            | <b>9</b>    |
| <b>BAB III LANDASAN TEORI.....</b>             | <b>16</b>   |
| 3.1. Teks Preprocessing.....                   | 16          |
| 3.2. Algoritma Stemming Nazief & Adriani ..... | 17          |
| 3.3. Algoritma Boyer Moore .....               | 19          |
| 3.4. Pembobotan TF-IDF.....                    | 24          |
| 3.5. Cosine Similarity .....                   | 26          |
| 3.6. Euclidean Distance .....                  | 27          |
| 3.7. Evaluasi Sistem .....                     | 28          |
| <b>BAB IV ANALISIS DAN PERANCANGAN .....</b>   | <b>30</b>   |
| 4.1. Analisis Permasalahan.....                | 30          |
| 4.2. Rancangan Umum Sistem .....               | 31          |
| 4.3. Model.....                                | 33          |
| 4.4. Pengumpulan Data .....                    | 34          |
| 4.4.1. Analisis Dataset Artikel Islam .....    | 34          |



|  |            |
|--|------------|
| 4.4.2. Analisis Dataset Tema .....               | 39         |
| 4.4.3. Analisis Data Artikel Uji .....           | 45         |
| 4.5. Teks Preprocessing .....                    | 48         |
| 4.5.1. Case Folding .....                        | 50         |
| 4.5.2. Tokenizing .....                          | 52         |
| 4.5.3. Filtering .....                           | 53         |
| 4.5.4. Stemming .....                            | 55         |
| 4.6. Filter Data .....                           | 58         |
| 4.6.1. Ekstrak Tema .....                        | 58         |
| 4.6.2. Filtering Artikel .....                   | 61         |
| 4.7. Pembobotan TF-IDF .....                     | 63         |
| 4.8. Uji Similaritas .....                       | 69         |
| 4.8.1. Cosine Similarity .....                   | 70         |
| 4.8.2. Euclidean Distance .....                  | 71         |
| 4.9. Evaluasi .....                              | 73         |
| <b>BAB V IMPLEMENTASI .....</b>                  | <b>76</b>  |
| 5.1. Spesifikasi Hardware dan Software .....     | 76         |
| 5.2. Data .....                                  | 77         |
| 5.2.1. Dataset Artikel Islam .....               | 78         |
| 5.2.2. Data Tema .....                           | 79         |
| 5.2.3. Data Artikel Uji .....                    | 81         |
| 5.3. Text Preprocessing .....                    | 82         |
| 5.4. Ekstrak Tema .....                          | 86         |
| 5.5. Filtering Artikel .....                     | 89         |
| 5.6. Pembobotan TF-IDF .....                     | 93         |
| 5.7. Uji Similaritas .....                       | 102        |
| 5.7.1. Cosine Similarity .....                   | 103        |
| 5.7.2. Euclidean Distance .....                  | 106        |
| 5.8. Evaluasi Precision, Recall & Akurasi .....  | 110        |
| <b>BAB VI HASIL DAN PEMBAHASAN .....</b>         | <b>119</b> |
| 6.1. Hasil Pengujian .....                       | 119        |
| 6.1.1. Pengujian Boyer Moore (Filter Data) ..... | 120        |



|  |            |
|--|------------|
| 6.1.2. Pengujian Cosine Similarity .....                       | 123        |
| 6.1.3. Pengujian Euclidean Distance .....                      | 128        |
| 6.2. Perbandingan Uji Nilai Similaritas.....                   | 133        |
| 6.3. Perbandingan Kinerja Filter Data & Tanpa Filter Data..... | 135        |
| 6.4. Evaluasi .....  | 138        |
| 6.4.1. Nilai Prediksi .....                                    | 141        |
| 6.4.2. Confusion Matrix .....                                  | 143        |
| 6.4.3. Precision, Recall, Accuracy & F1-score .....            | 148        |
| <b>BAB VII KESIMPULAN .....</b>                                | <b>153</b> |
| 7.1. Kesimpulan.....   | 153        |
| 7.2. Saran .....   | 154        |
| <b>DAFTAR PUSTAKA .....</b>                                    | <b>155</b> |