

## DAFTAR ISI

|                                         |     |
|-----------------------------------------|-----|
| ABSTRAK .....                           | i   |
| ABSTRACT .....                          | ii  |
| KATA PENGANTAR .....                    | iii |
| DAFTAR ISI .....                        | vi  |
| DAFTAR GAMBAR .....                     | x   |
| DAFTAR TABEL .....                      | xii |
| DAFTAR LAMPIRAN .....                   | xiv |
| DAFTAR SIMBOL .....                     | xv  |
| BAB 1 PENDAHULUAN .....                 | 1   |
| 1.1 Latar Belakang Masalah .....        | 1   |
| 1.2 Identifikasi Masalah .....          | 2   |
| 1.3 Maksud dan Tujuan .....             | 2   |
| 1.4 Batasan Masalah .....               | 2   |
| 1.5 Metode Penelitian .....             | 3   |
| 1.5.1 Identifikasi Masalah .....        | 4   |
| 1.5.2 Studi Literatur .....             | 5   |
| 1.5.3 Pengumpulan Data .....            | 5   |
| 1.5.4 Pembangunan Perangkat Lunak ..... | 5   |
| 1.5.5 Pengujian .....                   | 6   |
| 1.5.6 Kesimpulan Penelitian .....       | 6   |
| 1.6 Sistematika Penulisan .....         | 6   |
| BAB 2 LANDASAN TEORI .....              | 9   |
| 2.1 Dokumen Karya Tulis Ilmiah .....    | 9   |
| 2.2 Algoritma .....                     | 11  |
| 2.3 Ekstraksi Informasi .....           | 12  |
| 2.4 Hidden Markov Model .....           | 16  |
| 2.4.1 Observasi State .....             | 17  |
| 2.4.2 Hidden State .....                | 17  |
| 2.4.3 Probabilitas Awal ( $\pi$ ) ..... | 18  |

|                                      |                                              |    |
|--------------------------------------|----------------------------------------------|----|
| 2.4.4                                | Probabilitas Transisi (A) .....              | 18 |
| 2.4.5                                | Probabilitas Emisi (B) .....                 | 18 |
| 2.5                                  | Algoritma Viterbi .....                      | 19 |
| 2.5.1                                | Inisialisasi .....                           | 19 |
| 2.5.2                                | Rekursi .....                                | 19 |
| 2.5.3                                | Terminasi .....                              | 20 |
| 2.5.4                                | Backtracking .....                           | 20 |
| 2.6                                  | Natural Language Processing (NLP) .....      | 21 |
| 2.7                                  | Pemodelan Sistem .....                       | 21 |
| 2.7.1                                | Blok Diagram .....                           | 21 |
| 2.7.2                                | DFD (Data Flow Diagram) .....                | 22 |
| 2.7.3                                | Diagram Konteks .....                        | 22 |
| 2.8                                  | Tokenisasi .....                             | 23 |
| 2.9                                  | Ekstraksi Fitur .....                        | 23 |
| 2.10                                 | <i>PDF</i> .....                             | 25 |
| 2.11                                 | <i>TIFF (Tagged Image File Format)</i> ..... | 25 |
| 2.12                                 | <i>TXT</i> .....                             | 25 |
| 2.13                                 | <i>Regular Expression</i> .....              | 26 |
| 2.14                                 | <i>CSV</i> .....                             | 26 |
| 2.15                                 | Perancangan Sistem .....                     | 26 |
| 2.16                                 | Bahasa Pemrograman .....                     | 26 |
| 2.16.1                               | <i>PHP (Personal Home Page)</i> .....        | 26 |
| 2.17                                 | Pengujian Akurasi .....                      | 27 |
| 2.17.1                               | Nilai Akurasi .....                          | 27 |
| 2.18                                 | Perangkat Lunak Pendukung .....              | 27 |
| 2.18.1                               | <i>XAMPP</i> .....                           | 28 |
| 2.18.2                               | <i>Visual Studio Code</i> .....              | 28 |
| 2.18.3                               | <i>Imagemagick</i> .....                     | 29 |
| 2.18.4                               | <i>Tesseract</i> .....                       | 29 |
| BAB 3 ANALISIS DAN PERANCANGAN ..... |                                              | 31 |
| 3.1                                  | Analisis Masalah .....                       | 31 |

|                                        |                                                                 |     |
|----------------------------------------|-----------------------------------------------------------------|-----|
| 3.2                                    | Analisis Sistem.....                                            | 33  |
| 3.3                                    | Analisis Data Masukan .....                                     | 35  |
| 3.3.1                                  | Analisis Data Masukan ( <i>Preprocessing</i> Data Latih).....   | 36  |
| 3.3.2                                  | Analisis Data Masukan ( <i>Preprocessing</i> Data Testing)..... | 46  |
| 3.4                                    | Analisis Proses .....                                           | 50  |
| 3.4.1                                  | <i>Preprocessing</i> .....                                      | 50  |
| 3.4.2                                  | Tokenisasi .....                                                | 51  |
| 3.4.3                                  | Ekstraksi Fitur .....                                           | 56  |
| 3.5                                    | <i>Hidden Markov Model</i> .....                                | 68  |
| 3.5.1                                  | Kondisi Awal ( $\pi$ ).....                                     | 68  |
| 3.5.2                                  | Probabilitas Transisi (A) .....                                 | 68  |
| 3.5.3                                  | Probabilitas Emisi (B).....                                     | 71  |
| 3.6                                    | Algoritma Viterbi.....                                          | 73  |
| 3.6.1                                  | Decoding Pengujian .....                                        | 73  |
| 3.6.1.1                                | Inisialisasi.....                                               | 73  |
| 3.6.1.2                                | Rekursi .....                                                   | 76  |
| 3.6.1.3                                | Terminasi.....                                                  | 79  |
| 3.6.1.4                                | Backtracking .....                                              | 81  |
| 3.7                                    | Analisis Rencana Pengujian.....                                 | 83  |
| 3.7.1                                  | Rencana Pengujian Klasifikasi Token-Kelas .....                 | 83  |
| 3.8                                    | Analisis Kebutuhan Non Fungsional .....                         | 90  |
| 3.8.1                                  | Analisis Kebutuhan Pengguna .....                               | 91  |
| 3.8.2                                  | Analisis Kebutuhan Perangkat Keras.....                         | 91  |
| 3.8.3                                  | Analisis Perangkat Lunak .....                                  | 91  |
| 3.9                                    | Analisis Kebutuhan Fungsional .....                             | 92  |
| 3.9.1                                  | Diagram Konteks .....                                           | 92  |
| 3.9.2                                  | <i>DFD (Data Flow Diagram)</i> .....                            | 92  |
| 3.10                                   | Spesifikasi Proses.....                                         | 98  |
| 3.11                                   | Perancangan Antarmuka .....                                     | 105 |
| 3.12                                   | Jaringan Semantik .....                                         | 122 |
| BAB 4 IMPLEMENTASI DAN PENGUJIAN ..... |                                                                 | 125 |

|                                 |                                   |     |
|---------------------------------|-----------------------------------|-----|
| 4.1                             | Implementasi Sistem .....         | 125 |
| 4.1.1                           | Implementasi Perangkat Keras..... | 125 |
| 4.1.2                           | Implementasi Perangkat Lunak..... | 125 |
| 4.1.3                           | Implementasi Antarmuka .....      | 126 |
| 4.2                             | Pengujian Sistem.....             | 127 |
| 4.2.1                           | Skenario Pengujian.....           | 127 |
| 4.2.2                           | Pengujian.....                    | 128 |
| 4.2.3                           | Analisis Hasil Pengujian .....    | 129 |
| BAB 5 KESIMPULAN DAN SARAN..... |                                   | 131 |
| 5.1                             | Kesimpulan .....                  | 131 |
| 5.2                             | Saran.....                        | 131 |
| DAFTAR PUSTAKA .....            |                                   | 133 |