

Daftar Isi

| | |
|---|-----------|
| ABSTRAK..... | I |
| ABSTRACT | II |
| LEMBAR PERSEMBAHAN | III |
| KATA PENGANTAR | IV |
| DAFTAR ISI | V |
| DAFTAR GAMBAR | VII |
| DAFTAR TABEL | VIII |
| DAFTAR ISTILAH | IX |
| 1. PENDAHULUAN..... | 1 |
| 1.1. LATAR BELAKANG..... | 1 |
| 1.2. PERUMUSAN MASALAH | 2 |
| 1.3. TUJUAN | 2 |
| 1.4. METODOLOGI PENYELESAIAN MASALAH..... | 2 |
| 2. DASAR TEORI | 4 |
| 2.1. WEB MINING..... | 4 |
| 2.2. KONSEP WRAPPER INDUCTION | 5 |
| 2.2.1. <i>Wrapper</i> | 5 |
| 2.2.2. <i>Wrapper Induction</i> | 6 |
| 2.2.3. <i>LR Wrapper</i> | 8 |
| 2.2.4. <i>Contoh Hasil Ekstraksi Teks dai Halaman Web Komersial</i> | 12 |
| 2.3. KINERJA SISTEM..... | 13 |
| 2.3.1. <i>Recall dan Precision</i> | 13 |
| 2.3.2. <i>F-Measure</i> | 13 |
| 3. ANALISIS PERANCANGAN DAN IMPLEMENTASI | 15 |
| 3.1. GAMBARAN UMUM SISTEM..... | 15 |
| 3.2. DESKRIPSI DAN ANALISIS SISTEM..... | 15 |
| 3.2.1. <i>Penentuan Calon Delimiters</i> | 16 |
| 3.2.1.1. Penentuan Calon Kandidat | 16 |
| 3.2.1.2. Penentuan Kandidat | 16 |
| 3.2.1.3. Validitas Calon Kandidat | 16 |
| 3.2.2. <i>Pencocokan Kandidat</i> | 16 |
| 3.2.3. <i>Proses Ekstraksi</i> | 17 |
| 3.2.4. <i>Perhitungan Akurasi</i> | 17 |
| 3.3. PERANCANGAN SISTEM | 17 |
| 3.3.1. <i>Diagram Aliran Data</i> | 17 |
| 3.3.1.1. Diagram Konteks..... | 17 |
| 3.3.1.2. Diagram Aliran Data Level 1 | 17 |
| 3.3.1.3. Diagram Aliran Data Level 2 | 18 |
| 3.3.1.3.1. Diagram Aliran Data Level 2 Proses 2 | 18 |
| 3.3.1.3.2. Diagram Aliran Data Level 2 Proses 3 | 18 |
| 3.3.1.3.3. Diagram Aliran Data Level 2 Proses 4 | 18 |
| 3.3.2. <i>Kamus Data</i> | 18 |
| 3.3.3. <i>Spesifikasi Proses</i> | 19 |
| 3.3.4. <i>Tabel Terkait</i> | 22 |
| 3.4. IMPLEMENTASI | 23 |

| | | |
|-----------|---|-----------|
| 3.4.1 | <i>Lingkungan Implementasi</i> | 23 |
| 3.4.2 | <i>Spesifikasi Perangkat Keras</i> | 23 |
| 3.4.3 | <i>Spesifikasi Perangkat Lunak</i> | 23 |
| 4. | PENGUJIAN DAN ANALISIS | 24 |
| 4.1 | DESKRIPSI SISTEM | 24 |
| 4.2 | PENGUJIAN | 24 |
| 4.2.1. | <i>Data Uji yang Digunakan</i> | 24 |
| 4.2.2. | <i>Skenario Pengujian</i> | 25 |
| 4.2.3. | <i>Relevance Judgement</i> | 26 |
| 4.2.4. | <i>Parameter Pengujian</i> | 28 |
| 4.3 | ANALISIS HASIL PENGUJIAN | 28 |
| 4.3.1. | <i>Analisis Performansi Algoritma Wrapper Induction dari Pola yang Terbentuk Berdasarkan Jumlah Halaman yang Dibutuhkan Untuk Proses Learning</i> | 28 |
| 4.3.1.1. | <i>www.antaraneews.com</i> | 29 |
| 4.3.1.2. | <i>www.bbc.co.id</i> | 30 |
| 4.3.1.3. | <i>www.detik.com</i> | 31 |
| 4.3.1.4. | <i>www.kompas.com</i> | 32 |
| 4.3.1.5. | <i>www.liputan6.com</i> | 33 |
| 4.3.1.6. | <i>Penarikan Kesimpulan Analisis</i> | 34 |
| 4.3.2. | <i>Analisis Pengukuran Akurasi Ekstraksi Halaman Web Berdasarkan Data Uji</i> | 35 |
| 4.3.2.1. | <i>Data Uji dari Web yang Sama dengan Data Latih</i> | 36 |
| 4.3.2.2. | <i>Data Uji dari Web yang Berbeda dengan Data Latih</i> | 37 |
| 5. | KESIMPULAN DAN SARAN | 41 |
| 5.1. | KESIMPULAN | 41 |
| 5.2. | SARAN | 41 |
| | DAFTAR PUSTAKA | 42 |
| | LAMPIRAN A : STUDY KASUS | 43 |
| A.1. | PROSES LEARNING | 43 |
| A.1.1. | <i>Penentuan Calon Delimiters</i> | 43 |
| A.1.2. | <i>Pencocokan kandidat</i> | 45 |
| A.2. | PROSES EKSTRAKSI | 45 |
| A.2.1. | <i>Input Dokumen Ekstraksi</i> | 45 |
| A.2.2. | <i>Perhitungan Recall dan Precision</i> | 46 |
| A.2.3. | <i>Perhitungan F-Measure</i> | 46 |