

## DAFTAR ISI

<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>LEMBAR PERSETUJUAN</b> .....	<b>ii</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>iii</b>
<b>LEMBAR PERNYATAAN</b> .....	<b>iv</b>
<b>ABSTRAK</b> .....	<b>v</b>
<b>ABSTRACT</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	<b>vii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>xii</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiii</b>
<b>DAFTAR KODE SUMBER</b> .....	<b>xv</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	4
1.5 Manfaat Penelitian .....	4
1.6 Sistematika Penulisan .....	4
<b>BAB II KAJIAN PUSTAKA DAN DASAR TEORI</b> .....	<b>6</b>
2.1 Kajian Pustaka .....	6
2.2 <i>Information Retrieval System</i> .....	7
2.3 <i>Question Answering System</i> .....	9
2.4 <i>Text Mining</i> .....	9
2.4.1 <i>Pengertian Text Mining</i> .....	9
2.4.2 <i>Ruang Lingkup Text Mining</i> .....	11
2.4.3 <i>Teks Preprocessing</i> .....	11
2.4.3.1 <i>Case Folding dan Tokenizing</i> .....	12
2.4.3.2 <i>Filtering</i> .....	13
2.4.3.3 <i>Stemming</i> .....	14
2.5 <i>Algoritma Stemming Arifin</i> .....	15

2.6 Metode <i>Hashing</i> .....	18
2.7 Algoritma Rabin-Karp .....	19
2.8 Pembobotan Heuristik <i>Question Answering System</i> .....	23
2.9 Pembobotan <i>Rule Based Question Answering System</i> .....	25
<b>BAB III METODE PENELITIAN DAN PERANCANGAN .....</b>	<b>27</b>
3.1 Studi Literatur .....	28
3.2 Analisa Sistem .....	28
3.2.1 Kebutuhan <i>User Interface</i> .....	31
3.2.1.1 Perancangan Input pada Sistem .....	31
3.2.1.2 Perancangan <i>Prototype</i> Sistem .....	31
3.2.2 Kebutuhan Data .....	33
3.2.3 Kebutuhan Fungsional .....	33
3.2.4 Diagram UML ( <i>Unified Model Language</i> ) .....	34
3.3 Perancangan Sistem .....	38
3.4 Pemrosesan <i>Query</i> .....	40
3.4.1 <i>Preprocessing</i> .....	41
3.4.1.1 <i>Case Folding</i> dan <i>Tokenizing</i> .....	42
3.4.1.2 <i>Filtering</i> .....	44
3.4.1.3 Cek Kamus .....	46
3.4.1.4 <i>Stemming</i> .....	47
3.5 Pencarian dan <i>Retrieval Engine</i> .....	48
3.6 Ekstraksi Jawaban dengan Pembobotan Rabin-Karp .....	49
3.6.1 <i>Parsing</i> K-Gram .....	51
3.6.2 <i>Hashing</i> .....	52
3.7 Ekstraksi Jawaban dengan Pembobotan <i>Rule-Based</i> .....	53
3.8 Implementasi .....	54
3.8.1 Pengumpulan Data .....	54
3.8.1 Pemrosesan <i>Query</i> .....	55
3.8.2 Modul Sistem Pencarian .....	56
3.8.3 Modul Ekstraksi Jawaban .....	56
3.9 Perhitungan manual .....	59
3.10 Pengujian .....	70

3.10.1 Perhitungan Presisi dan <i>Recall</i> .....	71
3.10.2 Rancangan Uji Coba .....	72
3.11 Pengambilan Kesimpulan dan saran .....	73
<b>BAB IV IMPLEMENTASI .....</b>	<b>74</b>
4.1 Lingkungan Implementasi .....	74
4.1.1 Lingkungan Implementasi Perangkat Keras .....	74
4.1.2 Lingkungan Implementasi Perangkat Lunak .....	74
4.2 Implementasi Program .....	75
4.2.1 Kelas dan Fungsi.....	75
4.2.2 Tahapan Pemrosesan.....	79
4.3 Implementasi <i>User Interface</i> .....	88
<b>BAB V PENGUJIAN DAN ANALISIS.....</b>	<b>93</b>
5.1 Skenario Pengujian .....	93
5.2 Hasil Pengujian .....	94
5.2.1 Hasil Pengujian Presisi dan Akurasi .....	94
5.2.1.1 Perbandingan Hasil Pengujian Tanpa <i>Stemming</i> .....	112
5.2.1.2 Perbandingan Hasil Pengujian Dengan <i>Stemming</i> .....	115
5.2.2 Hasil Pengujian K-Gram.....	117
5.2.3 Hasil Pengujian <i>Modulo</i> .....	125
5.3 Analisis Hasil Pengujian.....	127
<b>BAB VI PENUTUP .....</b>	<b>135</b>
6.1 Kesimpulan .....	135
6.2 Saran .....	135
<b>DAFTAR PUSTAKA .....</b>	<b>137</b>
<b>LAMPIRAN.....</b>	<b>139</b>

**DAFTAR TABEL**

Tabel 3.1 Skenario *Use Case Diagram*..... 37

Tabel 3.2 Daftar pasangan kata tanya dan *named entity*..... 55

Tabel 3.3 Klasifikasi Kelas pada Perhitungan Akurasi, Presisi, dan *Recall*.... 71

Tabel 3.4 Perancangan uji coba ..... 72

Tabel 5.1 Tabel Hasil Uji Coba..... 95

Tabel 5.2 Tabel Hasil Uji Coba K-Gram = 1 dengan *modulo* 101 ..... 117

Tabel 5.3 Tabel Hasil Uji Coba K-Gram = 2 dengan *modulo* 101 ..... 119

Tabel 5.4 Tabel Hasil Uji Coba K-Gram = 3 dengan *modulo* 101 ..... 120

Tabel 5.5 Tabel Hasil Uji Coba K-Gram = 4 dengan *modulo* 101 ..... 122

Tabel 5.6 Tabel Hasil Uji Coba K-Gram = 5 dengan *modulo* 101 ..... 123

Tabel 5.7 Tabel uji *modulo* dengan kgram = 1 ..... 125

Tabel 5.8 Tabel uji *modulo* dengan kgram = 2 ..... 126

Tabel 5.9 Tabel uji *modulo* dengan kgram = 3 ..... 126

Tabel 5.10 Tabel uji *modulo* dengan kgram = 4 ..... 126

Tabel 5.11 Tabel uji *modulo* dengan kgram = 5 ..... 127



## DAFTAR GAMBAR

Gambar 2.1 Bagian – bagian <i>Information Retrieval (IR) System</i> .....	8
Gambar 2.2 Tahap <i>preprocessing</i> .....	12
Gambar 2.3 Tahap <i>case folding</i> dan <i>tokenizing</i> .....	12
Gambar 2.4 Tahap <i>filtering</i> .....	13
Gambar 2.5 Tahap <i>Stemming</i> .....	15
Gambar 2.6 Cara kerja algoritma <i>Rabin-Karp</i> .....	20
Gambar 2.7 Pengecekan tiga karakter pertama.....	21
Gambar 2.8 Pengecekan terhadap <i>substring</i> berikutnya.....	21
Gambar 2.9 Pengecekan <i>pattern</i> “c a b” dengan <i>substring</i> “a b b” .....	22
Gambar 2.10 Perbandingan <i>pattern</i> dengan <i>substring</i> berikutnya.....	22
Gambar 2.11 Perbandingan <i>pattern</i> dengan <i>hash</i> sama dengan <i>substring</i> .....	22
Gambar 2.12 Hasil pencarian <i>pattern</i> ditemukan .....	23
Gambar 2.13 Aturan metode <i>Rule-Based</i> .....	26
Gambar 3.1 Diagram alir metode penelitian .....	27
Gambar 3.2 Arsitektur <i>Question Answering System</i> .....	29
Gambar 3.3 Rancangan <i>prototype system</i> .....	32
Gambar 3.4 Diagram <i>Use Case</i> Sistem.....	35
Gambar 3.5 Diagram alir sistem .....	39
Gambar 3.6 Diagram alir pemrosesan <i>query</i> .....	40
Gambar 3.7 Diagram alir <i>preprocessing</i> .....	41
Gambar 3.8 Diagram alir <i>case folding</i> dan <i>tokenizing</i> .....	43
Gambar 3.9 Diagram alir <i>filtering</i> .....	45
Gambar 3.10 Diagram alir cek kamus.....	46
Gambar 3.11 Diagram alir <i>stemming</i> .....	48
Gambar 3.12 Diagram alir pencarian .....	49
Gambar 3.13 Diagram alir pembobotan <i>Rabin-Karp</i> .....	50
Gambar 3.14 Diagram alir parsing K-Gram .....	51
Gambar 3.15 Diagram alir <i>hashing</i> .....	52
Gambar 3.16 Diagram alir pembobotan <i>Rule-Based</i> .....	53
Gambar 4.1 Halaman utama.....	89

Gambar 4.2 Halaman jawaban pembobotan <i>Rabin Karp</i> .....	90
Gambar 4.3 Halaman jawaban <i>pembobotan Rule Based</i> .....	90
Gambar 4.4 Halaman proses <i>preprocessing</i> .....	91
Gambar 4.5 Halaman login pengelolaan dokumen .....	92
Gambar 4.6 Halaman pengelolaan <i>named entity tagging</i> .....	92
Gambar 5.1 Hasil pengujian tanpa <i>stemming</i> .....	114
Gambar 5.2 Hasil pengujian dengan <i>stemming</i> .....	117
Gambar 5.3 Grafik perbandingan waktu dengan kgram 1-5 pada soal 1-10 tanpa <i>stemming</i> .....	129
Gambar 5.4 Grafik perbandingan waktu dengan kgram 1-5 pada soal 1-10 dengan <i>stemming</i> .....	129
Gambar 5.5 Grafik perbandingan <i>heuristic</i> dengan kgram 1-5 pada soal 1-10 tanpa <i>stemming</i> .....	130
Gambar 5.6 Grafik perbandingan <i>heuristic</i> dengan kgram 1-5 pada soal 1-10 dengan <i>stemming</i> .....	131
Gambar 5.7 Grafik perbandingan waktu dengan kgram 1-5 tanpa <i>stemming</i> ..	132
Gambar 5.8 Grafik perbandingan waktu dengan kgram 1-5 dengan <i>stem</i> .....	132
Gambar 5.9 Grafik perbandingan <i>heuristic score</i> dengan kgram 1-5 tanpa <i>stemming</i> .....	133
Gambar 5.10 Grafik perbandingan <i>heuristic score</i> dengan kgram 1-5 dengan <i>stemming</i> .....	134

## DAFTAR KODE SUMBER

Kode Sumber 4.1 <i>Source code</i> proses cari dokumen.....	79
Kode Sumber 4.2 <i>Source code</i> proses <i>case folding</i> dan <i>tokenizing</i> .....	80
Kode Sumber 4.3 <i>Source code</i> proses <i>filtering</i> .....	81
Kode Sumber 4.4 <i>Source code</i> proses pencarian kandidat berdasarkan <i>rule</i> ... ..	81
Kode Sumber 4.5 <i>Source code</i> proses <i>parsing</i> K-Gram .....	82
Kode Sumber 4.6 <i>Source code</i> proses <i>hashing</i> .....	83
Kode Sumber 4.7 <i>Source code</i> proses perhitungan <i>threshold</i> .....	84
Kode Sumber 4.8 <i>Source code</i> proses perhitungan <i>count match</i> .....	84
Kode Sumber 4.9 <i>Source code</i> proses perhitungan nilai SM.....	85
Kode Sumber 4.10 <i>Source code</i> proses perhitungan nilai ORD .....	85
Kode Sumber 4.11 <i>Source code</i> proses menyimpan kata kerja .....	86
Kode Sumber 4.12 <i>Source code</i> proses perhitungan <i>wordmatch</i> .....	86
Kode Sumber 4.13 <i>Source code</i> proses perhitungan <i>rule</i> apa.....	87
Kode Sumber 4.14 <i>Source code</i> proses perhitungan <i>rule</i> dimana .....	87

