

DAFTAR ISI

HALAMAN PENGESAHAN	ii
PERNYATAAN	iii
HALAMAN MOTTO DAN PERSEMBAHAN	iv
PRAKATA	v
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
INTISARI	xiii
ABSTRACT	xiv
BAB I	
PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Metodologi Penelitian	3
1.6 Sistematika penulisan	4
BAB II	
TINJAUAN PUSTAKA	6
BAB III	
LANDASAN TEORI	9
3.1 HTML	9
3.1.1 Struktur HTML	10
3.2 Web Scraping	11
3.2.1 HTML Scraping	12
3.3 Text Processing	14
3.3.1 Stopword	14
3.4 Topic Modelling	15
3.4.1 Latent Dirichlet Allocation	15

3.5 Tabular Data.....	18
3.5.1 Tabular Data Model	19
3.5.2 CSV	20
BAB IV	
ANALISIS DAN PERANCANGAN SISTEM.....	21
4.1 Deskripsi Sistem	21
4.2 Analisis Kebutuhan Sistem	21
4.3 Perancangan Sistem	22
4.3.1 Perancangan <i>Web Scraping</i>	23
4.3.2 Perancangan <i>Topic Modelling</i>	24
4.3.2 Flowchart	25
4.4 Activity Diagram.....	27
4.5 Use Case Diagram.....	27
4.6 Diagram Alir Data.....	28
4.6.1 Diagram Alir Data Level Konteks	28
4.6.2 Diagram Alir Data Level 1.....	29
4.6.2 Diagram Alir Data Level 2.....	30
4.7 Perancangan Antar Muka.....	31
4.7.1 Halaman Eksekusi <i>Scraping</i> dan <i>Modelling</i>	31
BAB V	
IMPLEMENTASI	33
5.1 Pengembangan Sistem	33
5.2 Implementasi <i>Web Scraping</i>	33
5.2.1 Identifikasi Kelas	33
5.2.2 <i>Scraping Template</i>	35
5.3 Implementasi <i>Topic Modelling</i>	37
BAB VI	
PENGUJIAN DAN PEMBAHASAN	40
6.1 Pengujian Sistem.....	40
6.2 Hasil Pemrosesan Pada <i>Web Scraping</i>	40
6.3 Hasil Pemrosesan Pada <i>Topic Modelling</i>	43

BAB VII

KESIMPULAN DAN SARAN	48
7.1 Kesimpulan	48
7.2 Saran.....	48
DAFTAR PUSTAKA	49
LAMPIRAN A	
<i>Source Code Web Scraping dan Topic Modelling</i>	<i>51</i>
LAMPIRAN B	
<i>Source Code Gabungan.....</i>	<i>54</i>
LAMPIRAN C	
Hasil Total Artikel Dari Proses <i>Scraping</i>	58
LAMPIRAN D	
Hasil Pemodelan Dalam 7 Hari.....	61

DAFTAR TABEL

Tabel 2.1 Perbandingan dengan penelitian sebelumnya.....	8
Tabel 3.1 Tabel contoh <i>stopword</i>	15
Tabel 4.1 Daftar link proses <i>scraping</i>	24
Tabel 4.2 Contoh <i>dictionary</i>	25
Tabel 6.1 Total artikel yang diambil selama 7 hari.....	42
Tabel 6.2 Bobot topik untuk gambar pertama.....	44
Tabel 6.3 Bobot topik untuk gambar kedua.....	45
Tabel 6.4 Bobot topik untuk gambar ketiga.....	46
Tabel 6.5 Bobot topik untuk gambar keempat.....	46
Tabel 6.6 Bobot topik untuk gambar kelima.....	47

DAFTAR GAMBAR

Gambar 3.1 Struktur HTML.....	11
Gambar 3.2 Skema web scraping.....	12
Gambar 3.3 Modul import request scraping	13
Gambar 3.4 Request link scraping	13
Gambar 3.5 Dokumen HTML dengan pemodelan tree	13
Gambar 3.6 Tag div HTML dengan pemodelan tree	14
Gambar 3.7 Intuisi dasar LDA.....	16
Gambar 3.8 Plate notation LDA.....	17
Gambar 3.9 Bentuk dokumen CSV.....	20
Gambar 4.1 Pemodelan tree tag div HTML	23
Gambar 4.2 Pemodelan tree tag div HTML artikel	24
Gambar 4.3 Flowchart sistem	26
Gambar 4.4 Activity Diagram sistem.....	27
Gambar 4.5 Use Case Diagram	28
Gambar 4.6 DAD level konteks	28
Gambar 4.7 DAD level 1.....	29
Gambar 4.8 DAD level 2 web scraping	30
Gambar 4.9 DAD level 2 topic modelling	31
Gambar 4.10 Antar muka proses scraping	32
Gambar 4.11 Antar muka proses topic modelling	32
Gambar 5.1 Identifikasi kelas yang mengapit judul artikel	34
Gambar 5.2 Identifikasi kelas yang mengapit artikel berita	34
Gambar 5.3 Identifikasi kelas yang mengapit tanggal artikel	35
Gambar 5.4 Impor modul (library) python scraping	35
Gambar 5.5 Template Scraping berisi tag kelas HTML	36
Gambar 5.6 Memasukkan indeks halaman portal	36
Gambar 5.7 Menkonversi hasil scraping ke CSV	37
Gambar 5.8 Impor modul (library) python modeling	37
Gambar 5.9 Mengolah dokumen CSV	37
Gambar 5.10 Pembuatan dictionary dan corpus Indonesia	38

Gambar 5.11 Penghilangan stopwords	38
Gambar 5.12 Proses LDA	39
Gambar 5.13 Pemodelan dan keluaran topik	39
Gambar 6.1 Proses Eksekusi Scraping.....	41
Gambar 6.2 Hasil artikel web scraping 1 Mei.....	41
Gambar 6.3 Total artikel web scraping 1 Mei	42
Gambar 6.4 Proses topic modelling	43
Gambar 6.5 Hasil topic modelling	43
Gambar 6.6 Keluaran topik dengan wordcloud0.....	44
Gambar 6.7 Keluaran topik dengan wordcloud1.....	45
Gambar 6.8 Keluaran topik dengan wordcloud2.....	45
Gambar 6.9 Keluaran topik dengan wordcloud3.....	46
Gambar 6.10 Keluaran topik dengan wordcloud4.....	47