UNIVERSITAS
GADJAH MADA
EKSTRAKSI LOKASI DAN TOPIK PADA BERITA ONLINE BERBAHASA INDONESIA
AZIZ MUSTIKA AJI, Anny Kartika Sari, S.Si., M.Sc., Ph.D
Universitas Gadjah Mada, 2018 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

PRAKATA.....iv
INTISARI.....x
ABSTRACT.....xi
BAB I PENDAHULUAN.....12
1.1 Latar Belakang.....12
1.2 Rumusan Masalah.....14
1.3 Batasan Masalah.....14
1.4 Tujuan Penelitian.....14
1.5 Manfaat Penelitian.....14
1.6 Metode Penelitian.....14
1.7 Sistematika Penulisan.....15
BAB II TINJAUAN PUSTAKA.....17
BAB III LANDASAN TEORI.....21
3.1 Berita.....21
3.2 Ekstraksi Informasi.....21
3.3 Named Entity Recognition (NER).....21
3.4 Ekstraksi Topik.....23
3.5 Evaluasi.....25
BAB IV ANALISIS DAN PERANCANGAN SISTEM.....28
4.1 Ekstraksi Informasi.....28
4.2 Perancangan Sistem.....30
4.3 *Preprocessing*.....33
4.4 Ekstraksi Lokasi.....36
4.5 Ektraksi Topik.....37
4.6 Rancangan Percobaan.....38
BAB V IMPLEMENTASI.....40
5.1 Deskripsi Implementasi.....40
5.2 Implementasi Pengumpulan Data.....40
5.3 Implementasi Pelatihan Data.....44
5.4 Implementasi Preprocessing.....47
5.5 Implementasi Ekstraksi Lokasi.....48
5.6 Implementasi Ekstraksi Topik.....51
BAB VI PEMBAHASAN DAN PENGUJIAN.....54
6.1 Pembahasan.....54
6.2 Pengujian Ekstraksi Lokasi.....54
6.3 Pengujian Ekstraksi Topik.....55
6.4 Hasil Pembahasan Ekstraksi Topik.....65
BAB VII KESIMPULAN DAN SARAN.....66

UNIVERSITAS
GADJAH MADA
EKSTRAKSI LOKASI DAN TOPIK PADA BERITA ONLINE BERBAHASA INDONESIA
AZIZ MUSTIKA AJI, Anny Kartika Sari, S.Si., M.Sc., Ph.D
Universitas Gadjah Mada, 2018 | Diunduh dari http://etd.repository.ugm.ac.id/

7.1 Kesimpulan.....66
7.2 Saran.....67
DAFTAR PUSTAKA.....68
LAMPIRAN A KAMUS *STOP WORD*.....71

UNIVERSITAS
GADJAH MADA
EKSTRAKSI LOKASI DAN TOPIK PADA BERITA ONLINE BERBAHASA INDONESIA
AZIZ MUSTIKA AJI, Anny Kartika Sari, S.Si., M.Sc., Ph.D
Universitas Gadjah Mada, 2018 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR TABEL

Tabel 2.1 Perbedaan penelitian ini dengan penelitian-penelitian sebelumnya.....19
Tabel 3.1 Tabel kontingensi.....26
Tabel 4.1 Contoh korpus sebagai data latih.....29
Tabel 4.2 Contoh korpus sebagai data latih yang telah diklasifikasi.....30
Tabel 4.3 Contoh data hasil *crawling* dari *website* Kompas.....33
Tabel 4.4 Hasil proses tokenisasi.....35
Tabel 4.5 Hasil proses penghapusan *stop word*.....35
Tabel 4.6 Hasil proses identifikasi lokasi menggunakan NER.....36
Tabel 4.7 Hasil proses validasi lokasi.....37
Tabel 5.1 Variabel di dalam URL indeks berita Detik.....42
Tabel 6.1 Pengujian *10-fold cross-validation* pada data latih.....55
Tabel 6.2 Daftar 30 istilah berbobot tertinggi rentang waktu 1 hari.....56
Tabel 6.3 Daftar 25 kalimat berbobot tertinggi rentang waktu 1 hari.....56
Tabel 6.4 Daftar vektor unit rentang waktu 1 hari.....57
Tabel 6.5 Daftar 30 istilah berbobot tertinggi rentang waktu 2 hari.....59
Tabel 6.6 Daftar 25 kalimat berbobot tertinggi rentang waktu 2 hari.....59
Tabel 6.7 Daftar vektor unit rentang waktu 2 hari.....61
Tabel 6.8 Daftar 30 istilah berbobot tertinggi rentang waktu 7 hari.....62
Tabel 6.9 Daftar 25 kalimat berbobot tertinggi rentang waktu 7 hari.....63
Tabel 6.10 Daftar vektor unit rentang waktu 7 hari.....64
Tabel 9.1 Daftar *stop word* (Tala, 2003).....71

UNIVERSITAS
GADJAH MADA
EKSTRAKSI LOKASI DAN TOPIK PADA BERITA ONLINE BERBAHASA INDONESIA
AZIZ MUSTIKA AJI, Anny Kartika Sari, S.Si., M.Sc., Ph.D
Universitas Gadjah Mada, 2018 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR GAMBAR

Gambar 3.1 Hidden Markov Model.....23
Gambar 3.2 Ilustrasi proses *k-fold cross-validation* (Eckroth, 2014).....25
Gambar 4.1 Diagram alir data sistem.....31
Gambar 4.2 Diagram alir data *preprocessing*.....34
Gambar 5.1 Cuplikan kode untuk mengambil konten dari halaman indeks berita Detik.....41
Gambar 5.2 Cuplikan kode untuk memotong teks berita dari Detik.....43
Gambar 5.3 Cuplikan tabel Detik pada basis data.....44
Gambar 5.4 Cuplikan corpus yang diklasifikasikan secara manual.....45
Gambar 5.5 Cuplikan kode prosedur pemanggilan *Class* Train.....45
Gambar 5.6 Cuplikan kode *function* wordSeqWithTag.....46
Gambar 5.7 Cuplikan kode *function* extractSentence.....47
Gambar 5.8 Cuplikan kode *function* stopwordcheck.....48
Gambar 5.9 Cuplikan kode prosedur klasifikasi entitas.....49
Gambar 5.10 Cuplikan kode prosedur pencocokan nama lokasi.....50
Gambar 5.11 Cuplikan kode validasi token sebagai nama lokasi.....50
Gambar 5.12 Cuplikan kode prosedur pembobotan istilah pada *Class* TFPDF.....52
Gambar 5.13 Cuplikan kode prosedur pembuatan vektor kalimat.....53