UNIVERSITAS GADJAH MADA
Pemodelan Topik Menggunakan Latent Dirichlet Allocation Pada Dokumen Berita Sustainable Development Goals (SDGs)
HIDAYATUL FITRI, Widyawan, S.T., M.Sc., Ph.D.; Dr. Indah Soesanti, S.T., M.T.
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

LEMBAR PENGESAHAN ..... **ii**

PERNYATAAN ..... **iii**

PRAKATA ..... **iv**

ARTI LAMBANG DAN SINGKATAN ..... **vi**

ABSTRACT ..... **vii**

INTISARI ..... **viii**

DAFTAR ISI ..... **ix**

DAFTAR GAMBAR ..... **xii**

DAFTAR TABEL ..... **xiii**

BAB I PENDAHULUAN **1**

1.1 Latar Belakang **1**

1.2 Perumusan masalah **6**

1.3 Keaslian penelitian **6**

1.4 Tujuan Penelitian **11**

1.5 Manfaat Penelitian **11**

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI **12**

2.1 Tinjauan Pustaka **12**

2.2 Landasan Teori **16**

2.2.1 *Text Mining* 16

2.2.2 *Topic Modeling* 17

2.2.3 *Latent Dirichlet Allocation* (LDA) 18

2.3 Hipotesis **22**

BAB III METODOLOGI **24**

3.1 Alat dan Bahan **24**

3.1.1 Alat 24

UNIVERSITAS GADJAH MADA
Pemodelan Topik Menggunakan Latent Dirichlet Allocation Pada Dokumen Berita Sustainable Development Goals (SDGs)
HIDAYATUL FITRI, Widyawan, S.T., M.Sc., Ph.D.; Dr. Indah Soesanti, S.T., M.T.
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

3.1.2 Bahan 25
3.2 Jalannya Penelitian 26
3.2.1 Tahap Pengumpulan Data 28
3.2.2 Tahap *Preprocessing* 28
3.2.3 Pembentukan N-Gram 29
3.2.4 Representasi *Dictionary* dan *Corpus* 29
3.2.5 Pembobotan Kata 30
3.2.6 Pemodelan LDA 31
3.2.7 *Coherence Score* 31
3.2.8 *Perplexity* 33
3.2.9 LDA – Mallet 34
3.3 Cara Analisis 35
BAB IV HASIL DAN PEMBAHASAN 37
4.1 Pengumpulan Data 37
4.2 *Text Preprocessing* 38
4.2.1. Implementasi *Case Folding* 39
4.2.2 Implementasi *Stopwords Removal* 40
4.2.3 Implementasi *Stemming* 41
4.2.4 Implementasi *Tokenization* 42
4.3 Pembentukan N-Gram 44
4.4 Pembentukan *Dictionary* dan *Corpora* 46
4.5 Pembobotan Kata TF-IDF 47
4.6 *Topic Modeling* dengan LDA 48
4.7 Membentuk Model Topik 49
4.8 Validasi Model Topik 52
4.8.1 *Coherence Score* 52
4.8.2 *Perplexity* 54
4.9 LDA – Mallet 55
4.10 Visualisasi LDA 56
4.11 Representasi Topik 61

UNIVERSITAS
GADJAH MADA
Pemodelan Topik Menggunakan Latent Dirichlet Allocation Pada Dokumen Berita Sustainable Development Goals (SDGs)
HIDAYATUL FITRI, Widyawan, S.T., M.Sc., Ph.D.; Dr. Indah Soesanti, S.T., M.T.
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

BAB V KESIMPULAN DAN SARAN 65

5.1 Kesimpulan 65

5.2 Saran 66

DAFTAR PUSTAKA 67

LAMPIRAN 72