# DAFTAR ISI

HALAMAN JUDUL ..... i
HALAMAN PENGESAHAN DOSEN PEMBIMBING ..... ii
HALAMAN PENGESAHAN DOSEN PENGUJI ..... iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR ..... iv
HALAMAN PERSEMBAHAN ..... v
HALAMAN MOTO ..... vi
KATA PENGANTAR ..... vii
SARI ..... viii
GLOSARIUM ..... ix
DAFTAR ISI ..... x
DAFTAR TABEL ..... xiii
DAFTAR GAMBAR ..... xiv
BAB I PENDAHULUAN ..... 1
1.1 Latar Belakang ..... 1
1.2 Rumusan Masalah ..... 3
1.3 Batasan Masalah ..... 3
1.4 Tujuan Penelitian ..... 3
1.5 Manfaat Penelitian ..... 3
1.6 Metodologi Penelitian ..... 3
1.7 Sistematika Penulisan ..... 4
BAB II LANDASAN TEORI ..... 5
2.1 REST API ..... 5
2.2 Twitter ..... 5
2.3 Text Mining ..... 6
2.4 Elasticsearch ..... 6
2.5 Logstash ..... 7
2.6 Kibana ..... 7
2.7 *Latent Dirichlet Allocation* (LDA) ..... 7
2.8 *Coefficient of Variation* (CV) ..... 8
2.9 Studi Literatur ..... 9

BAB III METODOLOGI PENELITIAN ..... 13
3.1 Metode Penelitian ..... 13
3.1.1 Studi Literatur ..... 13
3.1.2 Mempersiapkan Aplikasi dan Bahan ..... 14
3.1.3 Pengambilan Data ..... 14
3.1.4 Perancangan Sistem ..... 15
3.1.5 *Preprocessing* ..... 15
3.1.6 Identifikasi Frasa ..... 15
3.1.7 *Topic Modelling* Menggunakan LDA ..... 15
3.1.8 Menghitung *Coefficient of Variation* ..... 16
3.2 Analisis Kebutuhan Aplikasi ..... 16
3.2.1 Analisis Kebutuhan Perangkat Keras ..... 16
3.2.2 Analisis Kebutuhan Perangkat Lunak ..... 16
3.3 Perancangan Sistem ..... 17
3.3.1 Perancangan *Activity Diagram* ..... 17
3.3.2 Perancangan *Use Case Diagram* ..... 18
3.3.3 Desain Antar Muka ..... 19
BAB IV HASIL DAN PEMBAHASAN ..... 21
4.1 Pengambilan data dan Membuat Akun Twitter Developer ..... 21
4.1.1 Membuat Akun Twitter Developer ..... 21
4.1.2 Menentukan *Keyword* dan Membuat *Script* ..... 23
4.1.3 Pengambilan Data ..... 24
4.1.4 Membuat *Index Pattern* di *Kibana* ..... 24
4.1.5 Mengeluarkan Hasil Pengambilan Data ..... 26
4.2 Implementasi ..... 26
4.2.1 *Preprocessing* ..... 26
4.2.2 Identifikasi Frasa ..... 34
4.2.3 Pemodelan LDA ..... 35
4.2.4 Menghitung Coefficient of Variation ..... 36
4.2.5 Analisis Hasil Pemodelan LDA ..... 37
4.2.6 Analisis Hasil Perhitungan *Coefficient of Variation* ..... 43
BAB V KESIMPULAN DAN SARAN ..... 44
5.1 Kesimpulan ..... 44
5.2 Saran ..... 44

DAFTAR PUSTAKA .....45
LAMPIRAN.....47

# DAFTAR TABEL

Tabel 2.1 Penelitian sebelumnya ..... 9
Tabel 4.1 Contoh hasil proses menghapus *non-ascii* ..... 27
Tabel 4.2 Contoh hasil proses menghapus URL ..... 28
Tabel 4.3 Contoh hasil proses menghapus simbol ..... 30
Tabel 4.4 Contoh hasil proses menghapus angka atau nomor ..... 30
Tabel 4.5 Contoh hasil proses *case folding* ..... 31
Tabel 4.6 Contoh hasil proses menghapus kelebihan spasi ..... 32
Tabel 4.7 Contoh hasil proses menghapus *stopword* ..... 33

# DAFTAR GAMBAR

Gambar 2.2 Proses LDA ..... 8
Gambar 3.1 Tahapan penelitian ..... 13
Gambar 3.2 Daftar calon pilpres 2019 di situs KPU (KPU, n.d.) ..... 14
Gambar 3.3 Tahapan proses *preprocessing* ..... 15
Gambar 3.4 *Activity diagram* sistem ..... 18
Gambar 3.5 *Use case diagram* sistem ..... 19
Gambar 3.6 Desain halaman utama ..... 19
Gambar 3.7 Desain halaman upload file ..... 20
Gambar 3.8 Desain halaman hasil ..... 20
Gambar 4.1 Form membuat Twitter Developer ..... 22
Gambar 4.2 *Access token* dan *API key* ..... 23
Gambar 4.3 *Script* untuk mengambil data menggunakan API ..... 24
Gambar 4.4 Menjalankan ketiga aplikasi ..... 24
Gambar 4.5 Halaman utama *Kibana* ..... 25
Gambar 4.6 Halaman management ..... 25
Gambar 4.7 Membuat *index pattern* ..... 26
Gambar 4.8 Kode program untuk menghapus kalimat ..... 27
Gambar 4.9 Kode program untuk menghapus *non-ascii* ..... 27
Gambar 4.10 Kode program untuk menghapus URL ..... 28
Gambar 4.11 Kode program untuk menghapus simbol ..... 29
Gambar 4.12 Kode program untuk menghapus angka atau nomor ..... 30
Gambar 4.13 Kode program untuk *case folding* ..... 31
Gambar 4.14 Kode program untuk menghapus kelebihan spasi ..... 32
Gambar 4.15 Kode program untuk menghapus *stopword* ..... 33
Gambar 4.16 Kode program untuk semua tahap *preprocessing* ..... 34
Gambar 4.17 Kode program untuk identifikasi frasa ..... 34
Gambar 4.18 Kode program untuk *dictionary* ..... 35
Gambar 4.19 Kode program untuk *bag of words* ..... 35
Gambar 4.20 Kode program untuk *tf-idf* ..... 35
Gambar 4.21 Kode program untuk menyimpan hasil pemodelan LDA ..... 35
Gambar 4.22 Kode program untuk menghitung rata-rata *coefficient of variation* ..... 36
Gambar 4.23 Kode program untuk menghitung *coefficient of variation* ..... 37

Gambar 4.24 Kode program untuk menyimpan hasil *coefficient of variation* .....37
Gambar 4.25 Kode program untuk membuat *wordcloud* .....37
Gambar 4.26 Hasil Visualisasi *Intertopic Distance Map* .....38
Gambar 4.27 Hasil Visualisasi *wordcloud*.....39
Gambar 4.28 Hasil Visualisasi *Top-30 Most Salient Terms*.....40
Gambar 4.29 Pencarian Google terkait kata “online_online” (Google, n.d.-a) .....41
Gambar 4.30 Pencarian Google terkait kata “ton_produksi” dan “impor_jagung” (Google, n.d.-b).....42
Gambar 4.31 Hasil perhitungan CV.....43