

DAFTAR ISI

HALAMAN JUDUL	
LEMBAR PERSETUJUAN	
LEMBAR PENGESAHAN	
SURAT PERNYATAAN KARYA SENDIRI	
ABSTRAK	i
ABSTRACT	ii
DAFTAR RIWAYAT HIDUP	iii
Kata Pengantar	iv
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Batasan Masalah	5
1.5 Kerangka Pemikiran	6
1.6 Metode Penelitian	7
1.7 Sistematika Penulisan	9
BAB II LANDASAN TEORI	11
2.1 The State of the Art	11
2.2 Landasan Teori	22
2.2.1 Definisi Al-quran	22
2.2.2 Data Mining	22
2.2.3 Text Mining	23
2.2.4 Clustering	23
2.2.5 K-means Clustering	24
2.2.6 K-means++ Clustering	25
2.2.7 Text Preprocessing	26
2.2.8 Term Frequency Inverse Document Frequency (Tf-Idf)	27
2.2.9 Python	28
2.2.10 Framework Django	28
2.2.11 UML	29

BAB III ANALISIS DAN PERANCANGAN	31
3.1 Analisis Sistem	31
3.1.1 Analisis Masalah	31
3.1.2 Analisis Sumber Data	32
3.2 Analisis Proses Perhitungan	32
3.2.1 Analisis Text Preprocessing	33
3.2.2 Analisis Algoritma TF-IDF	35
3.2.3 Analisis Algoritma	39
3.3 Analisis Fungsional	43
3.1 Analisis Kebutuhan Hardware	43
3.4.1 Analisis Kebutuhan <i>Hardware</i>	44
3.4.2 Analisis Kebutuhan <i>Software</i>	44
3.5 Arsitektur sistem	44
3.6 Perancangan Sistem	46
3.6.1 <i>Use Case Diagram</i>	46
3.6.2 <i>Definisi Use Case</i>	46
3.6.3 <i>Class Diagram</i>	47
3.6.4 <i>Sequence Diagram</i>	48
3.6.5 <i>Activity Diagram</i>	52
3.7 Flowchart Sistem	54
3.8 Perancangan Interface (Tampilan)	56
BAB IV IMPLEMENNTASI DAN PENGUJIAN	58
4.1 Implementasi	58
4.1.1 Persiapan Arsitektur	58
4.2 Implementasi Basis data	59
4.3 Tampilan Antarmuka	61
4.4 Source Code	63
4.5 Pengujian Cluster	65
4.6 Hasil Pengujian	68
4.1 Hasil Clustering	70
4.8 World Cloud	71
BAB V PENUTUP	75
5.1 Kesimpulan	75
5.2 Saran	75
DAFTAR PUSTAKA	77
LAMPIRAN	79

DAFTAR TABEL

Tabel 1. 1 Tabel kerangka pemikiran	6
Tabel 2. 1 The State of the Art	11
Tabel 3. 1 Sampel Data Terjemahan Ayat Al-quran	32
Tabel 3. 2 Proses Tokenizing.....	33
Tabel 3. 3 Proses Stopword.....	34
Tabel 3. 4 Proses Stemming	34
Tabel 3. 5 Analisis TF.....	35
Tabel 3. 6 Analisis DF	36
Tabel 3. 7 Analisis IDF.....	37
Tabel 3. 8 Analisis TF-IDF	38
Tabel 3. 9 Hasil TF-IDF.....	39
Tabel 3. 10 Perhitungan jarak ke pusat Centroid pertama (Data C).....	40
Tabel 3. 11 Pemilihan Centroid kedua.....	41
Tabel 3. 12 Perhitungan jarak ke pusat Centroid ke 2 (Data B).....	41
Tabel 3. 13 Hasil Perhitungan centroid Pertama (Yang baru)	42
Tabel 3. 14 Hasil Perhitungan centroid Kedua (Yang baru).....	42
Tabel 3. 15 Perhitungan jarak ke Centroid (Yang Baru)	42
Tabel 3. 16 Tabel Analisis Fungsional	43
Tabel 3. 17 Definisi Use Case	47
Tabel 4. 1 Pengamatan Indeks Silhouette coefficient.....	65
Tabel 4. 2 Tabel Hasil pengujian.....	68
Tabel 4. 3 Tabel hasil clustering.....	70

DAFTAR GAMBAR

Gambar 3. 1 Arsitektur sistem.....	45
Gambar 3. 2 Use Case Diagram	46
Gambar 3. 3 Class Diagram	48
Gambar 3. 4 Squence Diagram preprocessing	49
Gambar 3. 5 Sequence Diagram Clustering	50
Gambar 3. 6 Sequence Diagram hasil Clustering.....	51
Gambar 3. 7 Menampilkan surat	51
Gambar 3. 8 Activity melakukan Preprocessing	52
Gambar 3. 9 Activity dagram melakukan clustering	53
Gambar 3. 10 Activity Diagram hasil Clustering	54
Gambar 3. 11 Activity Diagram menampilkan surat.....	54
Gambar 3. 12 Flowchart sistem.....	55
Gambar 3. 13 Halaman melihat hasil cluster	57
Gambar 3. 14 Halaman jumlah cluster	57
Gambar 4. 1 Tabel keseluruhan data pada SQLite	59
Gambar 4. 2 Tabel data hasil clustering	60
Gambar 4. 3 Tabel data surat.....	60
Gambar 4. 4 Tabel detail surat	61
Gambar 4. 5 Tampilan halaman home	61
Gambar 4. 6 Halaman data hasil cluster	62
Gambar 4. 7 Halaman detail surat	62
Gambar 4. 8 inisialisasi data awal	63
Gambar 4. 9 sourcecode preprocessing.....	63
Gambar 4. 10 proses TFIDF.....	64
Gambar 4. 11 algoritma k-means++ clustering	64
Gambar 4. 12 Silhouette coefficient 1	65
Gambar 4. 13 Silhouette Coefficient II.....	65
Gambar 4. 14 Percobaan silhouette coefficient dengan 5 cluster.....	66
Gambar 4. 15 Percobaan silhouette coefficient dengan 7 cluster.....	67
Gambar 4. 16 Percobaan silhouette coefficient dengan 10 cluster.....	67
Gambar 4. 17 Percobaan silhouette coefficient dengan 15 cluster.....	68