

Daftar Isi

LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI	iv
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI.....	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR.....	viii
Daftar Isi.....	x
Daftar Gambar	xiii
Daftar Tabel	xiv
Daftar Simbol.....	xv
BAB I.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Metode Penelitian	4
1.6 Sistematika Penulisan.....	5
BAB II.....	7
2.1 Penambangan Data (<i>Data Mining</i>).....	7
2.2. Penambangan Teks (<i>Text Mining</i>)	10
2.2.1. <i>Pre-processing</i>	11
2.2.2. Transformasi teks.....	15
2.2.3. <i>Feature selection</i>	16
2.2.4. Penemuan pola	16
2.2.5. Interpretasi dan evaluasi	16
2.3. <i>Natural language processing</i>	17
2.4. <i>Words Embedding</i>	18
2.4.1. <i>Word2Vec</i>	20
2.4.2. <i>Global Vectors for Word Representation (GloVe)</i>	26
2.4.3. <i>Fast-Text Model</i>	29
2.5. <i>Clustering</i>	30

2.5.1. <i>Hierarchical Clustering Based</i>	33
2.5.2. <i>Partitioning Clustering Method</i>	33
2.5.3. <i>Density Based Clustering Method</i>	34
2.6. <i>Evaluasi Clustering</i>	38
2.6.1. <i>Silhouette Score</i>	39
2.6.2. <i>Analisis Similaritas kata</i>	39
2.6.3. <i>Calinski Harabsz Index</i>	39
2.6.4. <i>Density-Based Clustering Validation</i>	39
BAB III	40
3.1. <i>Persiapan Dataset Penelitian</i>	41
3.2. <i>Pra-pemrosesan data</i>	41
3.2.1. <i>Cleaning Dataset</i>	42
3.2.2. <i>Case-folding</i>	42
3.2.3. <i>Tokenisasi</i>	42
3.2.4. <i>Stop-words Removal</i>	43
3.2.5. <i>Penggabungan token menjadi teks</i>	43
3.3. <i>Pelatihan Data Menggunakan Model Word2Vec</i>	43
3.4. <i>Pengelompokan Kata Menggunakan OPTICS clustering</i>	47
3.4.1. <i>Visualisasi Cluster</i>	52
3.4.2. <i>Menghitung Silhouette Score</i>	53
3.4.3. <i>Menghitung Sebaran Cluster Menggunakan Calinski-Harabasz</i>	54
3.4.4. <i>Menghitung similarity kata</i>	55
3.4.5. <i>Validasi Hasil Clustering</i>	55
BAB IV	58
4.1. <i>Dataset Penelitian</i>	58
4.2. <i>Training Model</i>	59
4.3. <i>OPTICS CLUSTERING</i>	71
4.3.1. <i>Analisis Hasil Clustering</i>	72
4.3.2. <i>Pemilihan Cluster Terbaik</i>	79
4.3.3. <i>Analisis Relasi Data Points Cluster pada Ayat Al-Qur'an</i>	85
BAB V	88
5.1. <i>Kesimpulan</i>	88
5.2. <i>Saran</i>	89
DAFTAR PUSTAKA	91

RIWAYAT HIDUP	94
LAMPIRAN	95

