



DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN	iv
PRAKATA.....	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metodologi penelitian	3
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
BAB III LANDASAN TEORI.....	9
3.1 <i>Information Retrieval</i>	9
3.2 Praproses	10
3.2.1 Tokenisasi.....	10
3.2.2 Normalisasi.....	11
3.2.3 Stop word.....	11
3.2.4 Lemmatisasi.....	12
3.3 <i>Term Frequency dan Weighting</i>	12
3.3.1 <i>Inverse document frequency</i>	12
3.3.2 <i>Tf-idf weighting</i>	13
3.4 <i>Vector Space Model (VSM)</i>	13
3.4.1 <i>Dot product</i>	14
3.4.2 Kueri sebagai vektor.....	15



3.4.3	Evaluasi	16
3.5	<i>Clustering</i> Dokumen	17
3.5.1	<i>Hierarchical agglomerative clustering</i>	18
3.5.2	<i>K-means</i>	19
3.5.3	<i>Bisecting k-means</i>	21
3.5.4	Pengukuran jarak	22
3.5.5	Evaluasi <i>cluster</i>	23
BAB IV ANALISIS DAN RANCANGAN		27
4.1	Data	30
4.2	Praproses	31
4.2.1	Tokenisasi	31
4.2.2	Normalisasi	31
4.2.3	Penghapusan <i>stopword</i>	32
4.2.4	Penghapusan kata dengan panjang kurang dari tiga huruf	32
4.2.5	Lemmatisasi	32
4.3	Rancangan pembobotan	33
4.4	Rancangan Implementasi <i>Clustering</i> pada Dokumen	34
4.4.1	<i>Hierarchical agglomerative centroid clustering</i>	34
4.4.2	<i>K-means</i>	36
4.4.3	<i>Bisecting k-means</i>	37
4.5	Rancangan Pencarian Hadits	38
4.5.1	Rancangan pencarian data hadits tanpa <i>clustering</i>	38
4.5.2	Rancangan pencarian terhadap gugusan hasil <i>clustering</i>	40
4.6	Rancangan Pengujian	40
4.6.1	Pengujian kualitas hasil gugusan	41
4.6.2	Pengujian sistem pencarian hadits	42
BAB V IMPLEMENTASI		44
5.1	Praproses	44
5.2	Pembobotan	45
5.3	<i>Clustering</i> Hadits	46
5.3.1	<i>Hierarchical agglomerative centroid clustering</i>	46
5.3.2	<i>K-means</i>	49
5.3.3	<i>Bisecting k-means</i>	50
5.4	Sistem Pencarian Hadits	52



5.4.1	Pembentukan kamus kata	53
5.4.2	Pembentukan vektor kueri	53
5.4.3	Perhitungan similaritas	54
5.4.4	Pencarian data hadits tanpa <i>clustering</i>	55
5.4.5	Pencarian terhadap gugusan hasil <i>clustering</i>	57
5.5	Pengujian	59
5.5.1	Pengujian kualitas hasil gugusan	59
5.5.2	Pengujian sistem pencarian hadits.....	60
BAB VI HASIL PENGUJIAN DAN PEMBAHASAN		61
6.1	Pengujian Kualitas Hasil Gugusan	61
6.1.1	<i>Purity score</i>	61
6.1.2	<i>Silhouette value</i>	62
6.2	Pengujian Sistem Pencarian Hadits.....	63
6.2.1	Pencarian dengan kueri <i>top term</i> dari masing-masing bab.....	63
6.2.2	Pencarian dengan kueri responden	66
BAB VII KESIMPULAN DAN SARAN		68
7.1	Kesimpulan.....	68
7.2	Saran.....	68
DAFTAR PUSTAKA		69
LAMPIRAN A		71
LAMPIRAN B		72
LAMPIRAN C		75
LAMPIRAN D.....		77