

DAFTAR ISI

HALAMAN JUDUL

LEMBAR PERSETUJUAN

LEMBAR PENGESAHAN

PERNYATAAN KEASLIAN SKRIPSI

ABSTRAK

ABSTRACT

KATA PENGANTAR	i
DAFTAR ISI	iii
DAFTAR GAMBAR	viii
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	6
1.3. Batasan Masalah	6
1.4. Tujuan Penelitian	7
1.5. Metode Penelitian	8
1.6. Sistematika Penulisan	8
BAB II LANDASAN TEORI	10
2.1 Penggalian Data (<i>Data Mining</i>)	10
2.2 Penggalian Teks (<i>Text Mining</i>)	11
2.3 Pengelompokkan Teks (<i>Text Clustering</i>)	12

2.4	Algoritma <i>Clustering</i>	13
2.5	Metode <i>Kmeans</i>	14
2.6	Metode <i>Kmedoids</i>	15
2.7	Pra-Pemrosesan Teks (<i>Text Preprocessing</i>)	16
2.7.1	<i>Case Folding</i>	17
2.7.2	<i>Tokenizing</i>	17
2.7.3	<i>Stopword Removal</i>	17
2.7.4	<i>Stemming</i>	17
2.8	Pembobotan Kata (<i>Term Weighting</i>)	18
2.8.1	Metode <i>TF-IDF</i>	18
2.9	Reduksi Dimensi Fitur	20
2.9.1	<i>Feature Extraction</i>	20
2.10	<i>Proximity Measure</i>	23
2.10.1	<i>Euclidean Distance</i>	24
2.10.2	<i>Cosine Similarity</i>	24
2.11	<i>Silhouette Coefficient</i>	25
2.12	<i>Python</i>	28
BAB III ALGORITMA CLUSTERING K-MEANS DAN K-MEDOIDS		
DENGAN KOMBINASI PROXIMITY MEASURE DAN REDUKSI		
DIMENSI PRINCIPAL COMPONENT ANALYSIS		29
3.1	Kitab Hadis	31
3.2	Pengumpulan Data	31
3.3	Analisis <i>Text Preprocessing</i>	34
3.3.1	<i>Case Folding</i>	36

3.3.2	<i>Tokenizing</i>	36
3.3.3	<i>Stopword Removal</i>	37
3.3.4	<i>Stemming</i>	39
3.4	<i>Analisis Term Weighting</i>	40
3.4.1	<i>Analisis Metode TF-ID</i>	41
3.5	<i>Reduksi Dimensi</i>	47
3.6	<i>Algoritma Clustering K-Means</i>	49
3.6.1	Skenario ke-1	49
3.6.2	Skenario ke-2	50
3.7	<i>Algoritma Clustering K-Medoids</i>	51
3.7.1	Skenario ke-3	51
3.7.2	Skenario ke-4	52
BAB IV ANALISIS HASIL PENENTUAN JUMLAH CLUSTER TERBAIK PADA ALGORITMA K-MEANS DAN K-MEDOIDS		53
4.1	<i>Dataset</i>	53
4.2	<i>Analisis Hasil Clustering Skenario ke-1</i>	54
4.2.1	<i>Analisis Hasil Clustering K-Means Menggunakan Euclidean Distance Tanpa Reduksi Dimensi</i>	54
4.2.2	<i>Analisis Hasil Clustering K-Means Menggunakan Cosine Similarity Tanpa Reduksi Dimensi</i>	56
4.2.3	<i>Perbandingan Hasil Algoritma Clustering K-Means Dengan Kombinasi Proximity Measure Pada Skenario Pertama</i>	58
4.3	<i>Analisis Hasil Clustering Skenario ke-2</i>	59
4.3.1	<i>Analisis Hasil Clustering K-Means Menggunakan Euclidean Distance setelah Reduksi Dimensi PCA</i>	60

4.3.2 Analisis Hasil <i>Clustering</i> K-Means Menggunakan <i>Cosine Similarity</i> setelah Reduksi Dimensi PCA	63
4.3.3 Perbandingan Hasil Algoritma <i>Clustering K-Means</i> Dengan Kombinasi <i>Proximity Measure</i> Pada Skenario Kedua	65
4.4 Analisis Perbandingan Hasil <i>Clustering</i> dengan Algoritma <i>K-Means</i>	66
4.5 Analisis Hasil <i>Clustering</i> Skenario ke-3	68
4.5.1 Analisis Hasil <i>Clustering K-Medoids</i> Menggunakan <i>Euclidean Distance</i> Tanpa Reduksi Dimensi	68
4.5.2 Analisis Hasil <i>Clustering K-Medoids</i> Menggunakan <i>Cosine Similarity</i> Tanpa Reduksi Dimensi	71
4.5.3 Perbandingan Hasil Algoritma <i>Clustering K-Means</i> Dengan Kombinasi <i>Proximity Measure</i> Pada Skenario Ketiga	73
4.6 Analisis Hasil <i>Clustering</i> Skenario ke-4	74
4.6.1 Analisis Hasil <i>Clustering</i> K-Medoids Menggunakan <i>Euclidean Distance</i> setelah Reduksi Dimensi PCA	74
4.6.2 Analisis Hasil <i>Clustering</i> K-Medoids Menggunakan <i>Cosine Similarity</i> setelah Reduksi Dimensi PCA	77
4.6.3 Perbandingan Hasil Algoritma <i>Clustering K-Medoids</i> Dengan Kombinasi <i>Proximity Measure</i> Pada Skenario Keempat	80
4.7 Analisis Perbandingan Hasil <i>Clustering</i> Algoritma <i>K-Medoids</i>	81
4.8 Analisis Perbandingan Algoritma <i>K-Means</i> dan <i>K-Medoids</i> Tanpa Reduksi Dimensi	82
4.8.1 Analisis Hasil <i>Clustering</i> Algoritma K-Means dan K-Medoids dengan Perhitungan Jarak <i>Euclidean Distance</i>	82
4.8.2 Analisis Hasil <i>Clustering</i> Algoritma K-Means dan K-Medoids dengan Perhitungan Jarak <i>Cosine Similarity</i>	83

4.9 Analisis Perbandingan Algoritma <i>K-Means</i> dan <i>K-Medoids</i> dengan Reduksi Dimensi PCA	84
4.9.1 Analisis Hasil <i>Clustering</i> Algoritma <i>K-Means</i> dan <i>K-Medoids</i> dengan Perhitungan Jarak <i>Euclidean Distance</i>	85
4.9.2 Analisis Penentuan Jumlah <i>Cluster</i> Terbaik pada Algoritma <i>K-Means</i> dan <i>K-Medoids</i> dengan Perhitungan Jarak <i>Cosine Similarity</i>	86
4.10 Analisis Penentuan Jumlah <i>Cluster</i> Terbaik pada Algoritma <i>K-Means</i> dan <i>K-Medoids</i>	87
BAB V PENUTUP	89
5.1 Kesimpulan	89
5.2 Saran	89
DAFTAR PUSTAKA	91
LAMPIRAN	94
RIWAYAT HIDUP	

