

Daftar Isi

ABSTRAK	I
ABSTRACT	II
LEMBAR PERSEMBAHAN.....	III
KATA PENGANTAR.....	IV
DAFTAR ISI.....	V
DAFTAR GAMBAR.....	VII
DAFTAR TABEL	VIII
DAFTAR ISTILAH	IX
1. PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 RUMUSAN MASALAH	2
1.3 BATASAN MASALAH	3
1.4 TUJUAN	3
1.5 METODOLOGI PENYELESAIAN MASALAH	3
2. LANDASAN TEORI.....	5
2.1 INTERNET DAN WEB	5
2.2 KNOWLEDGE DISCOVERY IN DATABASE	5
2.3 TEXT PREPROCESSING	6
2.4 PEMBOBOTAN	6
2.4.1 TF (Term Frequency)	6
2.4.2 IDF (Inverse Document Frequency)	7
2.4.3 TF-IDF	7
2.5 INFORMATION RETRIEVAL	8
2.6 WEB CRAWLER	8
2.6.1 Best First Search	10
2.6.2 Cosine Similarity	11
2.6.3 Classifier.....	12
2.6.3.1 Teorema Bayes.....	12
2.6.3.2 Naïve Bayes Classifier.....	13
2.7 PARAMETER PENGUKURAN PERFORMANSI	15
2.7.1 Akurasi	15
2.7.2 Recall	15
2.7.3 Precision.....	16
2.7.4 F-Measure.....	16
2.8 HALAMAN WEB OLAHRAGA	17
3. PERANCANGAN SISTEM.....	18
3.1 DESKRIPSI UMUM SISTEM.....	18
3.2 KONFIGURASI SISTEM.....	19
3.2.1 Kebutuhan perangkat lunak:	19
3.2.2 Kebutuhan perangkat keras:.....	19
3.3 DIAGRAM UML (UNIFIED MODELLING LANGUAGE)	19
3.3.1 Diagram Use Case.....	19
3.3.2 Skenario Use Case	20
3.3.3 Diagram Aktivitas	22

3.3.4 Diagram Sequence	24
3.3.4.1 Diagram Sequence untuk SetDataTrain	24
3.3.4.2 Diagram Sequence untuk ViewDataTrain	24
3.3.4.3 Diagram Sequence untuk SetSeed	25
3.3.4.4 Diagram Sequence untuk Crawling	25
3.3.4.5 Diagram Sequence untuk ViewTree	26
3.3.5 Class Diagram.....	27
3.3.6 Component Diagram	28
4. PENGUJIAN DAN ANALISIS.....	29
4.1 PENGUJIAN	29
4.1.1 Link Pengujian.....	29
4.1.2 Parameter Pengujian.....	29
4.1.3 Skenario Pengujian.....	29
4.1.4 Data Pengujian.....	30
4.1.4.1 Pengaruh stopword pada pengujian	31
4.1.5 Hasil Pengujian dan Analisis	33
5. KESIMPULAN DAN SARAN.....	37
5.1 KESIMPULAN	37
5.2 SARAN	37
REFERENSI.....	39
LAMPIRAN A	40
MENGHILANGKAN TERM STOPWORD.....	40
TANPA MENGHILANGKAN TERM STOPWORD	42
DENGAN MENGGNAKAN LINK PENGUJIAN NETRAL.....	43
LAMPIRAN B	44
CONTOH SPIDER TRAP	44