

DAFTAR ISI

LEMBAR PERSETUJUAN.....	
LEMBAR PENGESAHAN.....	
PERNYATAAN KARYA SENDIRI.....	
HALAMAN PERSEMBAHAN.....	
ABSTRAK.....	I
ABSTRACT.....	II
KATA PENGANTAR.....	III
DAFTAR ISI.....	IV
DAFTAR GAMBAR.....	VI
DAFTAR TABEL.....	VII
DAFTAR ISTILAH.....	VIII
DAFTAR LAMPIRAN.....	X
BAB I PENDAHULUAN.....	1
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH.....	3
1.3 TUJUAN PENELITIAN.....	3
1.4 BATASAN MASALAH.....	3
1.5 KERANGKA PEMIKIRAN.....	4
1.6 SISTEMATIKA PENULISAN.....	5
BAB II TINJAUAN LITERATUR.....	7
2.1 TINJAUAN PUSTAKA.....	7
2.2 LANDASAN TEORI.....	9
2.2.1 <i>SIMILARITY</i> TEKS.....	9
2.2.2 ANALISIS SEMANTIK.....	10
2.2.3 <i>COSINE SIMILARITY</i>	10
2.2.4 <i>NATURAL LANGUAGE PROCESSING</i> (NLP).....	11
2.2.5 <i>WORD EMBEDDINGS</i>	13
2.2.6 <i>TRANSFORMERS</i>	15
2.2.7 <i>BERT</i>	16
2.2.8 <i>HUGGING FACE</i>	17
2.2.9 MEAN ABSOLUTE ERROR (MAE).....	17
2.2.10 <i>AGILE</i>	18

2.2.11	<i>CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM)</i>	19
BAB III METODOLOGI PENELITIAN		23
3.1	<i>PLANNING</i>	24
3.1.1	IDENTIFIKASI TUJUAN PENELITIAN	24
3.1.2	RENCANA PENELITIAN	24
3.2	DESIGN	25
3.2.1	DESIGN ARSITEKTUR SISTEM	25
3.3	DEVELOPMENT	26
3.3.1	<i>DATA UNDERSTANDING</i>	26
3.3.2	<i>DATA PREPARATION</i>	27
3.3.3	<i>MODELLING</i>	31
3.4	TESTING.....	34
3.4.1	EVALUASI HASIL PEMODELAN	34
3.4.2	EVALUASI APLIKASI DETEKSI SIMILARITY	34
3.5	REVIEW	35
BAB IV HASIL DAN PEMBAHASAN.....		36
4.1	HASIL PLANNING.....	36
4.1.1	HASIL TUJUAN BISNIS	36
4.1.2	<i>TIMELINE</i> PENELITIAN	36
4.2	<i>HASIL DESIGN APLIKASI</i>	37
4.3	<i>DEVELOPMENT</i>	38
4.3.1	HASIL <i>DATA UNDERSTANDING</i>	38
4.3.2	HASIL <i>DATA PREPARATION</i>	40
4.3.3	HASIL <i>MODELLING</i>	46
4.4	HASIL TESTING PERANGKAT LUNAK	51
4.5	HASIL REVIEW	53
BAB V SIMPULAN DAN SARAN		55
5.1	SIMPULAN	55
5.2	SARAN	56
DAFTAR PUSTAKA		XI
DAFTAR RIWAYAT HIDUP		XIV
LAMPIRAN.....		XV