

## DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iii
PRAKATA.....	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusah Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Keaslian Penelitian.....	4
BAB II TINJUAN PUSTAKA.....	5
BAB III LANDASAN TEORI.....	10
3.1 Sistem Rekomendasi.....	10
3.2 Metode <i>Content Based</i> .....	11
3.2.1 <i>Term frequency – inversed document frequency</i> .....	11
3.2.2 <i>Cosine similarity</i> .....	12
3.3 Metode <i>Item Based Collaborative Filtering</i> .....	13
3.4 Metode Penggabungan <i>Linear (Linear Combination)</i> .....	16
3.5 Transformasi Data pada Pemrosesan Paralel Apache Spark.....	17
3.5.1 Transformasi fungsi <i>map</i> .....	18
3.5.2 Transformasi fungsi <i>reduceByKey</i> .....	18
3.5.3 Transformasi fungsi <i>join</i> .....	19
3.5.4 Transformasi fungsi cartesian.....	19
3.5.5 Transformasi fungsi <i>reduce</i> .....	20
3.5.6 Transformasi fungsi <i>takeOrdered</i> .....	21
3.6 Apache Spark.....	22
3.6.1 Komponen apache spark.....	22
3.6.2 <i>Cluster</i> apache spark.....	24
3.6.3 Apache spark API.....	26
3.7 Skalabilitas Sistem.....	28
BAB IV ANALISIS DAN PERANCANGAN.....	29
4.1 Data Penelitian.....	29
4.2 Diagram Alir Program.....	30
4.2.1 Gambaran umum rancangan program.....	30
4.2.2 Konfigurasi dan parameter.....	32

4.2.3	Membaca dataset.....	33
4.2.4	Metode <i>content based</i> .....	34
4.2.5	Metode collaborative filtering.....	44
4.2.6	Penggabungan metode .....	51
4.3	Contoh Perhitungan .....	53
4.3.1	Metode <i>content based</i> .....	53
4.3.2	Metode <i>collaborative filtering</i> .....	56
4.3.3	Penggabungan metode .....	60
4.4	Rancangan Pengujian .....	61
4.4.1	Kriteria evaluasi .....	61
4.4.2	Skenario pengujian.....	62
BAB V	IMPLEMENTASI.....	63
5.1	Tahapan Awal Program.....	63
5.2	Implementasi Membaca Dataset.....	64
5.3	Implementasi Perhitungan Metode <i>Content Based</i> .....	67
5.3.1	Mengelompokkan frasa.....	67
5.3.2	Menghitung nilai TF .....	70
5.3.3	Menghitung nilai IDF.....	71
5.3.4	Menghitung nilai TF-IDF.....	72
5.3.5	Menghitung kemiripan dengan <i>cosine similarity</i> .....	73
5.4	Implementasi Perhitungan Metode <i>Collaborative Filtering</i> .....	76
5.4.1	Menghitung nilai rata - rata <i>rating</i> .....	76
5.4.2	Menghitung selisih.....	77
5.4.3	Menghitung <i>similarity</i> .....	78
5.4.4	Prediksi <i>rating</i> .....	81
5.5	Implementasi Penggabungan Metode.....	82
5.6	Menjalankan Program di <i>Standalone Cluster</i> .....	83
5.6.1	Membangun <i>cluster</i> .....	83
5.6.2	Menjalankan program rekomendasi.....	84
5.7	Menjalankan Program di <i>Cluster Google Cloud Dataproc</i> .....	87
5.7.1	Membuat <i>cluster</i> .....	87
5.7.2	Penyimpanan data .....	89
5.7.3	Submit program ke <i>cluster</i> .....	89
BAB VI	HASIL DAN PEMBAHASAN .....	92
6.1	Hasil perhitungan.....	92
6.2	Hasil uji coba pada <i>cluster</i> .....	93
BAB VII	PENUTUP .....	96
7.1	Kesimpulan.....	96
7.2	Saran .....	96
DAFTAR PUSTAKA	.....	98
LAMPIRAN SOURCE CODE	.....	102
LAMPIRAN UJI COBA	.....	105