

**PENERAPAN TEKNIK *MODIFIED FUZZY K-PARTITION*,
FUZZY C-MEANS DAN *POPULARITAS ITEM* UNTUK MENGATASI
MASALAH *COLD-START* PADA SISTEM REKOMENDASI**

Moh. Iqbal Okthapian
(160411100108)

Pembimbing I : Dr. Noor Ifada S.T., M.ISD.

Pembimbing II : Moch. Kautsar Sophan, S.Kom., M.MT.

ABSTRAK

Perkembangan teknologi yang semakin pesat menjadikan sistem rekomendasi sangat dibutuhkan. Masalah yang sering terjadi pada sistem rekomendasi adalah *cold-start*. *Cold-start* merupakan keadaan pada *user* baru yang belum memiliki riwayat atau catatan apapun di dalam sistem. Hal ini akan membuat sistem rekomendasi akan sulit mendapatkan hasil rekomendasi yang tepat. Tujuan dari penelitian ini adalah untuk mengatasi masalah *cold-start* dengan teknik *clustering* serta *popularitas item*. Teknik *clustering* yang digunakan adalah *Modified Fuzzy K-Partition* dan *Fuzzy C-Means*. *Modified Fuzzy K-Partition* digunakan untuk mengelompokkan *user* berdasarkan atribut umur, jenis kelamin serta pekerjaan *user*. Untuk mengelompokkan *item* menggunakan *Fuzzy C-Means* berdasarkan *rating* dari *item* tersebut, sedangkan untuk menghasilkan rekomendasi *Top-N* digunakan teknik *Popularitas Item*. Dengan metode tersebut sistem dapat memberikan rekomendasi *item* untuk *user* meskipun dalam keadaan *cold-start*. Dari hasil percobaan yang telah dilakukan, metode yang diusulkan mampu mendapatkan akurasi lebih baik dari metode dengan pendekatan *memory based (Item-based)*. Prosentase kenaikan rata-rata, maksimal dan minimal DCG dari metode usulan terhadap metode *item-based* adalah sebesar 115,3863%; 141,1612%; dan 34,2446%.

Kata Kunci: *Cold-start* , *Fuzzy C-Means*, *Modified Fuzzy K-Partition*, *Popularitas Item*, *Sistem Rekomendasi*.

DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN ORISINALITAS.....	iv
ABSTRAK	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR KODE PROGRAM.....	xiii
DAFTAR ALGORITMA.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.2.1 Permasalahan.....	3
1.2.2 Metode Usulan	3
1.2.3 Pertanyaan Peneliti	3
1.3 Tujuan Dan Manfaat.....	3
1.3.1 Tujuan	3
1.3.2 Manfaat	4
1.4 Batasan Masalah.....	4
1.5 Sistematika Proposal	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Landasan Teori	6
2.1.1 Sistem Rekomendasi	6
2.1.2 Content-Based	6

2.1.3	<i>Collaborative Filtering</i>	7
2.1.4	<i>Cold-Start</i>	8
2.1.5	<i>Modified Fuzzy K-Partition</i>	9
2.1.6	<i>Fuzzy C-Means</i>	12
2.1.7	Teknik <i>Popularitas Item</i>	15
2.2	Penelitian Terkait	15
BAB III METODOLOGI PENELITIAN		17
3.1	Langkah-langkah Penelitian	17
3.1.1	Survei Pendahuluan	17
3.1.2	Studi Pustaka	17
3.1.3	Pengambilan Data	17
3.2	Arsitektur Sistem	19
3.3	Perhitungan Manual	21
3.3.1	Data Matrik Demografi <i>User</i>	21
3.3.2	<i>Modified Fuzzy K-Partition</i>	22
3.3.3	Data matrik <i>Rating</i>	28
3.3.4	<i>Fuzzy C-Means</i>	29
3.3.5	Teknik <i>Popularitas Item</i>	35
3.3.6	Hasil Rekomendasi	37
3.4	Metode Evaluasi	37
3.5	Skenario Pengujian	38
BAB IV HASIL DAN PEMBAHASAN		40
4.1	Lingkungan Uji Coba	40
4.2	Pembuatan Program	40
4.2.1	Metode	40
4.2.1.1	Proses Pembuatan Data Demografi <i>User</i> Menjadi Matrik	40

4.2.1.2	<i>Modified Fuzzy K-Partition</i>	41
4.2.1.3	Proses Pembuatan Data <i>Rating</i> Menjadi Matrik	43
4.2.1.4	<i>Fuzzy C-Means</i>	44
4.2.1.5	Teknik <i>Popularitas Item</i>	46
4.2.1.6	Menentukan Hasil Rekomendasi	48
4.2.2	Evaluasi.....	48
4.3	Hasil Pengujian Akurasi	49
4.4.1	SKENARIO 1.....	49
4.4.2	SKENARIO 2.....	51
4.4.3	SKENARIO 3.....	52
4.4	Contoh Evaluasi per- <i>user</i> dan Hasil Pengujian Manual.....	55
BAB V KESIMPULAN DAN SARAN.....		57
5.1	Kesimpulan.....	57
5.2	Saran.....	57
DAFTAR PUSTAKA		58

DAFTAR GAMBAR

Gambar 2. 1 Sistem Rekomendasi <i>Collaborativ-Filtering</i>	7
Gambar 3. 1 Contoh Data <i>Rating</i>	18
Gambar 3. 2 Contoh Data Demografi <i>User</i>	18
Gambar 3. 3 Arsitektur Sistem	19
Gambar 3. 4 <i>5-Fold Cross Validation</i>	37
Gambar 3. 5 Skenario Pengujian	38
Gambar 4. 1 Perbandingan Evaluasi DCG Metode Usulan dan <i>Item-Besed</i>	55
Gambar 4. 2 Perhitungan Manual DCG	55
Gambar 4. 3 <i>Graphical User interface</i>	56

DAFTAR TABEL

Tabel 2. 1 Rangkuman Penelitian Sebelumnya.....	16
Tabel 3. 1 <i>Genre Film</i>	18
Tabel 3. 2 Data Demografi <i>User</i>	21
Tabel 3. 3 Kelompok Umur	22
Tabel 3. 4 Data Demografi <i>User</i> Setelah Diproses	22
Tabel 3. 5 Inisialisasi <i>Metode Modified Fuzzy K-Partition</i>	22
Tabel 3. 6 Nilai <i>Random</i> Derajat Keanggotaan	23
Tabel 3. 7 Anggota Atribut	23
Tabel 3. 8 Nilai <i>Lamda</i>	24
Tabel 3. 9 Hasil <i>Ln Cluster 1</i>	25
Tabel 3. 10 Hasil <i>Ln Cluster 2</i>	25
Tabel 3. 11 Hasil <i>Ln Cluster 3</i>	26
Tabel 3. 12 Derajat keanggotaan Baru dan Hasil <i>Cluster</i>	27
Tabel 3. 13 Hasil <i>Cluster 3 Kali Iterasi</i>	27
Tabel 3. 14 Data <i>Rating (X)</i>	28
Tabel 3. 15 Inisialisasi Nilai Variabel Metode <i>Fuzzy C-Means</i>	29
Tabel 3. 16 Nilai <i>Random</i> Derajat Keanggotaan	29
Tabel 3. 17 Pusat <i>Cluster</i>	30
Tabel 3. 18 Hasil $(x_{ij} - v_{kj})^2$ Untuk Data Ke-1	31
Tabel 3. 19 Hasil Seluruh <i>Cluster</i>	31
Tabel 3. 20 Jumlah Tiap <i>Cluster</i>	32
Tabel 3. 21 Hasil $((x_{ij} - v_{kj})^2)^{\frac{-1}{w-1}}$	33
Tabel 3. 22 Derajat Keanggotaan Baru	34
Tabel 3. 23 Anggota <i>Cluster</i>	34
Tabel 3. 24 Anggota <i>Cluster</i> Setelah 3 kali <i>iterasi</i>	35
Tabel 3. 25 Hasil <i>Cluster User</i>	35
Tabel 3. 26 Hasil <i>Cluster Item</i>	36
Tabel 3. 27 Inisialisasi Target & <i>Cluster User</i>	36
Tabel 3. 28 <i>Cluster Item</i> Terpilih.....	36
Tabel 3. 29 Hasil <i>Top-N</i>	36

Tabel 3. 30 Urutan <i>Top-N</i>	37
Tabel 3. 31 Skenario Pengujian	39
Tabel 4. 1 Perangkat yang Digunakan	40
Tabel 4. 2 Nilai Parameter Skenario 1	50
Tabel 4. 3 Hasil Evaluasi Skenario <i>Modified Fuzzy K-Partition</i>	50
Tabel 4. 4 Nilai Parameter Skenario 2	51
Tabel 4. 5 Hasil Evaluasi <i>Fuzzy C-Means</i>	52
Tabel 4. 6 Nilai Parameter Skenario 3	53
Tabel 4. 7 Perbandingan Evaluasi DCG Metode Usulan dan <i>Item-Besed</i>	54

DAFTAR KODE PROGRAM

Kode Program 4. 1 Pengolah Data Demografi <i>User</i> Menjadi Matrik.....	41
Kode Program 4. 2 <i>Modified Fuzzy K-Partition</i>	42
Kode Program 4. 3 Pengolah Data <i>Rating</i> Menjadi Matrik	43
Kode Program 4. 4 <i>Fuzzy C-Means</i>	45
Kode Program 4. 5 Menentukan <i>Cluster Target</i>	47
Kode Program 4. 6 Menentukan <i>Cluster Item Target</i>	47
Kode Program 4. 7 Jumlah <i>Rating Item</i> Terpilih	48
Kode Program 4. 8 Hasil Rekomendasi <i>Top-N</i>	48
Kode Program 4. 9 DCG	49

DAFTAR ALGORITMA

Algoritma 2. 1 <i>Modified Fuzzy K-Partition</i>	12
Algoritma 2. 2 <i>Fuzzy C-Means</i>	14
Algoritma 3. 1 Teknik <i>Popularitas Item</i>	21