

“PENGGUNAAN ATRIBUT *GENRE* DALAM MODEL NAIVE BAYES UNTUK *PROBABILISTIC RANKING* DALAM SISTEM REKOMENDASI”

SELGY ZAHRANIDA SUGIHARTO

14.04.111.00.013

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

Pembimbing II : Dr. Arif Muntasa, S.Si, M.T.,

ABSTRAK

Sistem rekomendasi merupakan sebuah sistem yang dibuat dengan tujuan memberikan sebuah rekomendasi *item* kepada *user*. Salah satu model pendekatan sistem rekomendasi yang sering digunakan adalah *Collaborative Filtering* (CF), dengan pendekatan *item-based*. Model ini memberikan hasil rekomendasi yang cukup baik, namun *item-based* CF tidak memanfaatkan *atribut item* dalam perhitungan rekomendasinya. Penerapan *atribut item* dapat dilakukan dengan dua cara, yaitu dengan satu tahap permodelan dan dua tahap permodelan. Satu tahap permodelan menghasilkan kualitas rekomendasi tidak terlalu signifikan (kurang baik). Cara yang kedua adalah dua tahap permodelan, *candidate* rekomendasi dan tahap *re-ordering*. Misal, tahap *candidate* rekomendasi menggunakan pendekatan *collaborative filtering*, dan *re ordering* dilakukan *content-based* atau *naïve bayes*. Pada penelitian sebelumnya sudah dilakukan penerapan *candidate* rekomendasi menggunakan *collaborative filtering*, dan *re ordering* dengan *naive bayes* pada data tagging menghasilkan rekomendasi yang baik. Oleh sebab itu pada penelitian ini penulis mencoba menerapkan dua tahap permodelan *candidate ranking* dengan *collaborative filtering*, dan *re-ordering* menggunakan *nave bayes*. dengan *study* kasus *movie* atribut *genre* pada data *explicit* (2 dimensi) . Dari hasil percobaan mendapatkan nilai yang cukup baik, dengan kenaikan nilai akurasi yang dihasilkan *metode evaluasi NDCG* sebesar 40.94486%.

Kata kunci: *Sistem Rekomendasi, Item-Based Collaborative Filtering, Naive Bayes, NDCG, Probabilistic Ranking.*

DAFTAR ISI

HALAMAN PENGESAHAN.....	v
HALAMAN PERNYATAAN ORISINALITAS.....	vii
HALAMAN PERSEMBAHAN.....	ix
ABSTRAK.....	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xv
DAFTAR GAMBAR.....	xviii
DAFTAR TABEL.....	xx
BAB.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	3
1.6 Model Penelitian.....	4
1.7 Sistematika Penulisan.....	5
BAB II.....	6
2.1 Penelitian Sebelumnya.....	7
2.2 Sistem Rekomendasi.....	9
2.3 Dataset MovieLens.....	10
2.4 Collaborative Filtering.....	11
2.5 Memory Based.....	11
2.6 Item-based Collaborative Filtering.....	12
2.7 Naive Bayes at probabilistic Ranking.....	13
2.7.1 Inverse User Frequency Matrik S (User-user).....	16
2.7.2 Genre Preference Generation Matrik UG (User-genre).....	17
2.7.3 Candidate Genre Set Matrik CGS.....	17
2.8 Matrik Evaluasi.....	18
BAB III.....	21
3.1 Analisa Sistem.....	21
3.2 Analisa Kebutuhan.....	21
3.2.1 Analisa Perangkat Keras.....	21
3.2.2 Analisa Perangkat Lunak.....	22
3.2.3 Analisa Data.....	22
3.3 Rancangan Sistem.....	22

3.4 Arsitektur Sistem	24
3.4.1 Input	24
3.4.2 Process	26
3.4.3 Output	26
3.5 Evaluasi	30
BAB IV	33
4.1. Lingkungan Uji Coba	33
4.2. <i>Dataset</i>	33
4.3. Pembuatan Program	34
4.3.1 Proses Matrik UI (User Item)	34
4.3.2 Proses Mean Centered Rating	34
4.3.3 Proses Adjusted Cosine Similarity	36
4.3.4 Prediksi Rating Candidate Item	37
4.3.5 Proses Matrik GI (Genre Item)	38
4.3.6 Proses Matrik UG (User Genre)	39
4.3.7 Proses Inverse User Frequency (Matrik S)	40
4.3.8 Proses Genre Preference Generation (Matrk UG)	41
4.3.9 Proses Candidate Genre Set (Matrik CGS)	42
4.3.10 Proses Naive Bayes at Probabilistic Ranking	42
4.3.11 Proses Evaluasi NDCG	43
4.4 Skenario Uji Coba	45
4.4.1 Skenario Uji Coba Nilai k	46
4.4.2 Skenario Uji Coba Nilai D	46
4.4.3 Skenario Uji Coba Nilai CGS	47
4.4.4 Skenario Uji Coba Nilai TOP-N	48
BAB V	52
PENUTUP	52
5.1. Kesimpulan	52
5.2. Saran	52
DAFTAR PUSTAKA	54

~Halaman Ini Sengaja Dikosongkan~

DAFTAR GAMBAR

Gambar 2. 1 Model Sistem Rekomendasi	8
Gambar 2.2 <i>u.data</i>	9
Gambar 2.3 <i>u.genre</i>	10
Gambar 2.4 <i>Genre</i> dalam <i>user-item</i>	12
Gambar 3.1 Diagram IPO	22
Gambar 3. 2 <i>Input</i>	23
Gambar 3. 3 Matrik <i>User-item</i>	23
Gambar 3. 4 Matrik <i>Item genre</i>	24
Gambar 3. 5 <i>Process</i>	24
Gambar 3. 6 Matrik <i>Mean Centered Rating</i>	25
Gambar 3. 7 Matrik <i>Adjusted Cosine Similarity</i>	25
Gambar 3. 8 <i>Output</i>	28
Gambar 4. 1 Skenario Uji Coba Nilai <i>top-N</i>	47



~Halaman Ini Sengaja Dikosongkan~

DAFTAR TABEL

Tabel 2.1 Penelitian sebelumnya	7
Tabel 2.2 Matrik <i>UI (User-item)</i>	13
Tabel 2.3 Matrik <i>IG (item-genre)</i>	13
Tabel 2.4 Matrik <i>GI (Genre-item)</i>	14
Tabel 2.5 Matrik <i>UG (user-genre)</i>	14
Tabel 2.6 Matrik <i>S (user-user)</i>	15
Tabel 4. 1 Spesifikasi perangkat	30
Tabel 4. 2 Skenario Uji Coba Nilai <i>k</i>	43
Tabel 4. 3 Skenario Uji Coba Nilai <i>D</i>	44
Tabel 4. 4 Skenario Uji Coba Nilai <i>CGS</i>	45
Tabel 4. 5 Skenario Uji Coba Nilai <i>top-N</i>	46

