

SISTEM REKOMENDASI LAPTOP MENGGUNAKAN COLLABORATIVE FILTERING DAN CONTENT-BASED FILTERING

*Anderias Eko Wijaya*¹, *Deni Alfian*²
Program Studi Teknik Informatika, STMIK Subang
E-mail: ekowjy09@yahoo.com¹, denialfian92@yahoo.co.id²

Abstract

Laptop is needed for students and for office workers because it is better than a desktop computer. In this era, laptops have a variety of brands and specifications that sometimes make people have difficulty in finding, choosing or buying the right laptop for their needs. Therefore there should be a recommendation system that can provide advice or recommendations, based on interest and needs in the search for references.

In commonly used algorithm recommendation system is collaborative filtering (CF) and content based filtering (CB). Collaborative filtering is a concept whereby the opinions of other users are used to predict items that a user might like / interest. For content based filtering using the availability of an item's content as a basis for recommendation.

In this research, the algorithm for collaborative filtering uses Adjusted-cossine similarity to calculate the similarity between user and weighted sum algorithm for prediction calculation, for content based filtering algorithm used is tf-idf to search availability of existing content.

This recommendation system combines collaborative filtering and content based filtering methods using mixed hybrid techniques, the system has also been tested using the blackbox method. The result of the required execution time is influenced by the number of items and content based filtering method has the fastest execution time compared to collaborative filtering and mixed hybrid methods.

Keywords : *recommender system, collaborative filtering, content based filtering, mixed hybrid, Adjusted-cossine similarity, weightes sum.*

SISTEM REKOMENDASI LAPTOP MENGGUNAKAN COLLABORATIVE FILTERING DAN CONTENT-BASED FILTERING

Abstrak

Sekarang ini laptop sangat diperlukan baik untuk pelajar maupun bagi para pekerja kantor karena kelebihanannya dibandingkan dengan komputer desktop. Dengan semakin berkembangnya zaman laptop sekarang ini mempunyai beragam merk dan spesifikasi yang terkadang membuat orang-orang mengalami kesusahan dan kesulitan dalam mencari, memilih ataupun membeli laptop yang tepat dan sesuai bagi kebutuhannya. Maka dari itu harus ada sebuah sistem rekomendasi (*recommendation system*) yang dapat memberikan saran ataupun rekomendasi laptop berdasarkan ketertarikan dan kebutuhan dalam pencarian referensi.

Dalam sistem rekomendasi algoritma yang umum digunakan adalah *collaborative filtering (CF)* dan *content based filtering (CB)*. *collaborative filtering* adalah suatu konsep dimana opini dari pengguna lain yang ada digunakan untuk memprediksi item yang mungkin disukai/diminati oleh seorang pengguna. Sedangkan *content based filtering* menggunakan ketersediaan konten sebuah item sebagai basis dalam pemberian rekomendasi.

Pada penelitian ini algoritma untuk *collaborative filtering* menggunakan *Adjusted-cosine similarity* untuk menghitung kemiripan antar user dan algoritma *weighted sum* untuk perhitungan prediksinya, sedangkan untuk *content based filtering* algoritma yang digunakan adalah *tf-idf* untuk pencarian ketersediaan konten yang ada.

Sistem rekomendasi ini menggabungkan antara metode *collaborative filtering* dan *content based filtering* dengan teknik *mixed hybrid*, sistem ini juga telah diuji menggunakan metode *blackbox*. Hasil waktu eksekusi yang dibutuhkan dipengaruhi oleh jumlah item dan metode *content based filtering* memiliki waktu eksekusi tercepat dibandingkan metode *collaborative filtering* dan *mixed hybrid*.

Kata kunci : sistem rekomendasi, *collaborative filtering*, *content based filtering*, *mixed hybrid*, *Adjusted-cosine similarity*, *weightes sum*.

PENDAHULUAN

Sekarang ini laptop sangat diperlukan baik untuk pelajar maupun bagi para pekerja kantor karena kelebihanannya dibandingkan dengan komputer desktop. Dengan semakin berkembangnya zaman laptop sekarang ini mempunyai beragam merk dan spesifikasi yang terkadang membuat orang-orang mengalami kesusahan dan kesulitan dalam mencari, memilih ataupun membeli laptop yang tepat dan sesuai bagi kebutuhannya. Dengan masalah tersebut tentunya orang-orang akan terbantu jika ada yang dapat

merekendasikan laptop yang sesuai bagi kebutuhannya. Maka dari itu harus ada sebuah sistem rekomendasi (*recommendation system*) yang dapat memberikan saran ataupun rekomendasi laptop berdasarkan ketertarikan dan kebutuhan dalam pencarian referensi.

Dalam sistem rekomendasi (*recommendation system*) algoritma yang umum digunakan adalah *collaborative filtering (CF)* dan *content based filtering (CB)*. *collaborative filtering (CF)* terbagi atas dua jenis yaitu *user-based CF* dan *item-based CF*. *User-based Collaborative Filtering*

berasumsi bahwa cara yang baik dalam menemukan konten yang dirasa akan disukai oleh konsumen adalah dengan menemukan orang lain dengan ketertarikan yang sama dengan konsumen tersebut, kemudian merekomendasikan hal yang disukai oleh orang lain tersebut kepada konsumen (handrico, 2012). *Item-based Collaborative Filtering* berasumsi bahwa jika mayoritas pengguna memberi penilaian beberapa item secara serupa, pengguna yang kita targetkan juga akan memberi penilaian terhadap item-item tersebut secara serupa dengan mayoritas pengguna lain (Sarwar, 2001).. Berbeda dengan *content-based filtering* yang mana tidak seperti *collaborative filtering* yang menggunakan parameter semacam rating untuk menghasilkan rekomendasi. Melainkan menggunakan deskripsi dari profil pengguna, atau dari deskripsi suatu item untuk menghasilkan suatu rekomendasi (Ricci, 2011).

Beberapa penelitian yang pernah dilakukan untuk menyelesaikan masalah sistem rekomendasi, diantaranya : sistem rekomendasi pencarian pekerjaan berbasis web menggunakan metode *hybrid-based recommendation*(nurul,2013) dimana menggabungkan *content based dan collaborative filtering* dengan menggunakan algoritma *decision tree* dan *nears neighbour* dalam menghasilkan rekomendasi, selanjutnya ada juga yang menggunakan *multi-criteria rating* yang di implementasikan menggunakan *hybrid (content based dan collaborative Filtering)* pada sistem rekomendasi software antivirus dengan *multi-criteria rating* (Arifin,2014), banyak metode kombinasi dalam penggabungan antara *content based dan collaborative Filtering* salah satunya menggunakan *switching hybrid* dimana sistem memanfaatkan dua metode sekaligus, jika salah satu metode gagal

menghasilkan prediksi atau memiliki nilai confident yang kecil maka akan digunakan metode yang lain, seperti yang digunakan pada sistem rekomendasi pengambilan mata kuliah pilihan menggunakan metode *hybrid* (Naufal, 2013).

Berdasarkan latar belakang di atas, maka didapat sebuah masalah yaitu:

- a) Sulitnya dalam memilih ataupun membeli laptop yang tepat dan sesuai keinginan.
- b) Dibutuhkannya suatu sistem yang dapat memberikan rekomendasi atau saran dalam memilih ataupun membeli laptop.

Adapun tujuan dalam penelitian yang dilakukan ini adalah sebagai berikut:

- a) Membuat sebuah sistem rekomendasi laptop menggunakan *collaborative filtering* dan *content based filtering*.
- b) Menjelaskan kelebihan dan kekurangan dari setiap metode.

Manfaat dari penelitian ini adalah sebagai berikut:

- a) Mempermudah dalam memilih ataupun membeli laptop yang sesuai.
- b) Sebagai salah satu alternatif untuk membantu memberikan rekomendasi dan saran laptop yang sesuai.

METODE PENELITIAN

Beberapa metodologi untuk pengerjaan Tugas Akhir ini adalah sebagai berikut :

1. Identifikasi Masalah

Pertama-tama masalah yang akan dipecahkan diidentifikasi terlebih dahulu supaya bisa mengetahui secara detail inti dari permasalahan yang akan diselesaikan dan juga bagaimana langkah-langkah dan metode yang akan diambil untuk menyelesaikan masalah tersebut

2. Perancangan sistem

Di tahap ini penulis mulai merancang kebutuhan-kebutuhan sistem seperti *dataset* yang dibutuhkan, metode-

metode yang akan dilakukan, dan juga hasil yang diharapkan.

3. Implementasi dan analisis sistem

Implementasi sistem dilakukan dalam dua tahap sebagai berikut :

- a. Mengimplementasikan dan menguji metode *item-based collaborative filtering* dengan langkah-langkah sebagai berikut :
 1. Melakukan pemrosesan data *rating* dari suatu *item* untuk mendapatkan data pengguna yang telah merating *item*.
 2. Menghitung *similarity* antara *item* satu dengan lainnya menggunakan metode *cosine similarity* berdasarkan *rating* yang diberikan pengguna.
 3. Melakukan perhitungan prediksi *rating* pengguna pada suatu *item* menggunakan metode *weighted sum*.
- b. Mengimplementasikan dan menguji metode *content-based filtering* dengan langkah-langkah sebagai berikut :
 1. Melakukan pemrosesan data profil pengguna yang digunakan sebagai *query*.
 2. Menghitung bobot *TF-IDF* (*term frequency-inverse document frequency*) pada tiap-tiap content laptop berdasarkan *query*.
 3. Menggunakan bobot *TF-IDF* untuk menghasilkan rekomendasi laptop.
 4. Menyimpulkan hasil dari penelitian Setelah hasil dianalisis dan dievaluasi, terdapat sebuah kesimpulan yang bisa diambil dan diharapkan berguna bagi kemajuan topik yang penulis teliti.

KAJIAN TEORI

Collaborative filtering

Collaborative filtering adalah suatu konsep dimana opini dari pengguna lain yang ada digunakan untuk memprediksi item yang mungkin disukai/diminati oleh seorang pengguna (Ricci , 2011).

Kualitas rekomendasi yang diberikan dengan menggunakan metode ini sangat bergantung dari opini pengguna lain (*neighbor*) terhadap suatu item. Belakangan diketahui bahwa melakukan reduksi *neighbor* (yaitu dengan memotong *neighbor* sehingga hanya beberapa pengguna yang memiliki kesamaan / *similarity* tertinggi sajalah yang akan digunakan dalam perhitungan) mampu meningkatkan kualitas rekomendasi yang diberikan (Adomavicious dan Kwon, 2007).

Collaborative filtering memberikan rekomendasi berdasarkan kumpulan dari pendapat, minat dan ketertarikan beberapa user yang biasanya diberikan dalam bentuk *rating* yang diberikan user kepada suatu item.

Untuk memperoleh data *rating* dari user yang digunakan dalam sistem rekomendasi, dibedakan menjadi dua cara (Wibowo,2010) yaitu:

- 1) Secara *Eksplisit*, yaitu proses pengumpulan data dimana user memberikan data secara sadar/sengaja.
- 2) Secara *Implisit*, yaitu proses pengumpulan data dimana user tidak menyadari bahwa ia telah memberikan masukan terhadap sistem.

Pendekatan *collaborative filtering* pada dasarnya dibagi menjadi dua kategori yaitu *user-based collaborative filtering* disebut juga *memory-based*, dan *item based collaborative filtering* yang disebut juga *model-based*, Pada pendekatan *user based collaborative filtering* sistem memberikan rekomendasi kepada user item-item

yang disukai atau dirating oleh user – user lain yang memiliki banyak kemiripan dengannya. Misalnya, user a menyukai atau merating item 1,2 dan 3, kemudian user b menyukai item 1,2 dan 4 maka sistem akan merekomendasikan item 3 kepada user b dan item 4 kepada user a. Kelebihan dari pendekatan user based collaborative filtering adalah dapat menghasilkan rekomendasi yang berkualitas baik. Sedangkan kekurangannya adalah kompleksitas perhitungan akan semakin bertambah seiring dengan bertambahnya pengguna sistem, semakin banyak pengguna (user) yang menggunakan sistem maka proses per Rekomendasi akan semakin lama (handrico, 2012).

Pendekatan *item based collaborative filtering* memberikan rekomendasi berdasarkan kemiripan antar item. Metode ini merupakan metode rekomendasi yang didasari atas adanya kesamaan antara pemberian rating terhadap suatu item dengan item yang pernah dirating user lain (handrico,2009). Item yang telah di rating oleh user akan menjadi patokan untuk mencari sejumlah item lainnya yang berkorelasi dengan item yang telah dirating user. Motivasi kunci dibalik metode ini adalah user akan cenderung menyukai item yang sejenis atau mempunyai korelasi dengan item yang telah disukainya (Sarwar, 2001).

Secara umum proses pemberian rekomendasi pada collaborative filtering terdiri atas 2 langkah (Sarwar, 2001), yaitu: Penemuan *similar* item dan Penghitungan prediksi.

Terdapat beberapa algoritma untuk menemukan similar item, yaitu: (Sarwar,2001)

1. Algoritma *Cosine-based Similarity* Pada kasus ini dua item dianggap sebagai 2 vektor. Kesamaan antara 2 item ini diukur dengan menghitung kosinus dari

sudut antara 2 vektor item. Item dibandingkan misalnya u dan v, dianggap sebagai sebuah vektor baris dengan anggotanya adalah nilai rating yang diberikan terhadap kedua item tersebut. Dua vektor dikatakan sama jika membentuk sudut 0^0 atau nilai

kosinusnya 1. Dengan kata lain dua item dikatakan mirip jika nilai kosinus dari perhitungan mendekati 1 (Wiranto, 2010).

Persamaan *cosine-based similarity*:

$$sim(i, j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\|_2 \|\vec{j}\|_2}$$

Dimana \vec{i} dan \vec{j} merupakan vektor vektor baris dengan anggota nilai rating pada item i dan item j. $\cos(\vec{i}, \vec{j})$ merupakan nilai kosinus sudut yang dibentuk vektor baris ratingitemi dan j.

2. Algoritma *Correlation-based Similarity* Pada algoritma ini kemiripan antara dua item i dan j diukur dengan menghitung korelasi Pearson-r correlation. Agar perhitungan korelasi yang diperoleh akurat, terlebih dahulu dilakukan pemisahan terhadap co-rateditem (item-item yang kedua item i dan j nya di-rating oleh user).

Persamaan *correlation-based similarity*:

$$sim(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_i)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_i)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_j)^2}}$$

3. *Adjusted-cosine similarity* Persamaan *adjusted cosine similarity* digunakan untuk menghitung nilai kemiripan antar item. Perhitungan kemiripan ini merupakan modifikasi dari perhitungan kemiripan berbasis vektor dimana dengan melihat fakta bahwa setiap user memiliki skema rating yang berbeda-beda. Terkadang user memberi rating yang tinggi terhadap item a disisi lain user memberi rating yang sangat rendah pada item b. Maka
-

dari itu untuk setiap rating dikurangi dengan rata-rata rating yang diberikan user.

$$\text{sim}(i,j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u)(R_{u,j} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}}$$

Keterangan :

$\text{sim}(i,j)$ = Nilai kemiripan antara item i dan item j .

$u \in U$ = Himpunan user u yang merating item i dan item j .

$R_{u,i}$ = Rating user u pada item i .

$R_{u,j}$ = Rating user u pada item j .

\bar{R}_u = Nilai rata-rata rating user u

Untuk menghitung nilai kemiripan (similarity) antar 2 item, diperlukan himpunan user yang me-rating item tersebut. Nilai yang dihasilkan pada persamaan *adjusted-cosine similarity* adalah berkisar antara +1.0 dengan -1.0. Item dianggap saling berkolerasi jika nilai similarity antara kedua item tersebut mendekati +1, begitu juga sebaliknya item dianggap tidak berkolerasi apabila nilai similarity-nya mendekati -1.

Content based filtering

Sistem rekomendasi berbasis konten (*Content-based Recommendation System*) menggunakan ketersediaan konten (sering juga disebut dengan fitur, atribut atau karakteristik) sebuah item sebagai basis dalam pemberian rekomendasi (Ricci, 2011). Sebagai contoh, sebuah film mempunyai konten seperti genre, author, tahun rilis, dan lain-lain, atau sebuah file dokumen memiliki konten berupa tulisan yang ada di dalamnya.

Metode content-based filtering biasa digunakan untuk merekomendasikan berita, artikel maupun situs web. Metode tersebut akan mengekstrak informasi yang terdapat pada item kemudian membandingkannya dengan informasi

item yang pernah dilihat atau disukai oleh user.

Teknik – teknik yang digunakan dalam content-based seperti TF-IDF, Bayesian Classifiers, Cluster analysis, decision trees dan artificial neural networks.

Sistem rekomendasi berbasis konten memiliki beberapa kelebihan, yaitu (Afifi, 2014):

- a. Sistem rekomendasi berbasis konten dapat menjelaskan bagaimana hasil rekomendasi didapatkan.
- b. Sistem rekomendasi berbasis konten dapat merekomendasikan item-item yang bahkan belum pernah di-rate oleh siapapun. Namun, sistem rekomendasi berbasis konten juga memiliki beberapa kelemahan, yaitu (Arifin, 2014):
- c. Sistem rekomendasi berbasis konten memerlukan sebuah profil user yang berisikan ketertarikan dan minat pengguna. Bagi pengguna baru yang belum pernah melakukan aktivitas apapun dan tidak memiliki profil user yang cukup, sistem rekomendasi tidak dapat memberikan rekomendasi yang handal kepadanya (*Cold Start Problem*).

Hybrid recommendation

Masing-masing teknik sistem rekomendasi memiliki kelebihan dan kelemahannya tersendiri. Karenanya, sistem rekomendasi hybrid yang menggabungkan kekuatan dari model dan algoritma yang berbeda untuk mengatasi masalah-masalah yang telah disebutkan di atas menjadi target

penelitian yang sedang marak dikembangkan (Jannach, 2011).

Adapun tujuh teknik kombinasi untuk hybridization (burke, n.d):

1. *Weighted* : bobot dari dua atau lebih metode rekomendasi digabungkan secara numerik
2. *Switching* : sistem memilih salah satu atau lebih metode rekomendasi dan menerapkan salah satu metode rekomendasi yang dipilih.
3. *Mixed* : teknik rekomendasi dari berbagai metode ditampilkan menampilkannya secara bersamaan.
4. *Feature Combination* : output dari salah satu teknik rekomendasi digunakan sebagai input yang lain.
5. *Feature Augmentatio* : fitur-fitur dari sumber data rekomendasi yang berbeda digabung bersama-sama ke dalam algoritma rekomendasi tunggal.
6. *Cascade* : satu rekomen der mengolah rekomendasi yang diberikan oleh yang lainnya.
7. *Meta-level* : model dipelajari oleh satu rekomen der yang digunakan sebagai inputan yang lainnya.

ANALISA

Collaborative filtering

Untuk dapat memberikan rekomendasi laptop menggunakan metode *collaborative filtering* dibutuhkan data informasi rating dari setiap *user* yang telah memberi rating terhadap suatu *item* atau laptop. Dari data rating tersebut nantinya akan dihitung nilai kemiripannya menggunakan algoritma *adjusted cosine similarity* dan melakukan bobot perhitung prediksi menggunakan algoritma *weighted sum*. langkah-langkah atau proses perhitungan menggunakan metode *collaborative filtering* adalah sebagai berikut :

1. pemberian rating
Dalam hal ini kita membutuhkan data rating yang berisi data user yang telah merating suatu item. Untuk

skala pemberian ratingnya adalah antara 1-5, sebagai contoh kita memiliki data nilai rating user terhadap item sebagai berikut :

Tabel 1 data rating

	lapto p a	lapto p b	lapto p c	lapto p d	lapto p e	lapto p f	rata- rata rating
use r 1		5	4	3			4
use r 2			3	2	4	1	2,5
use r 3		3				3	3
use r 4	4			1			2,5
use r 5		2	2	4		5	3,25
use r 6		5		4			4,5

2. menghitung kemiripan antar item
Setelah data rating terbentuk langkah selanjutnya adalah menghitung nilai kemiripan antar item menggunakan algoritma *adjusted cosine similarity*, berikut adalah langkah-langkah dalam menghitung kemiripan dengan algoritma *adjusted cosine similarity* :

Rumus yang digunakan dalam *adjusted cosine similarity* adalah :

$$\text{sim}(i,j) = \frac{\sum_{u \in U} (R_{ui} - \bar{R}_u)(R_{uj} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{ui} - \bar{R}_u)^2} + \sqrt{\sum_{u \in U} (R_{uj} - \bar{R}_u)^2}}$$

Keterangan :

$\text{sim}(i,j)$ = Nilai kemiripan antara item i dan item j.

$u \in U$ = Himpunan user u yang merating item i dan item j.

R_{ui} = Rating user u pada item i.

R_{uj} = Rating user u pada item j.

\bar{R}_u = Nilai rata-rata rating user u

Berikut adalah perhitungan algoritma *adjusted cosine similarity* :

Perhitungan nilai kemiripan dilakukan jika terdapat 2 atau lebih rating dari *user*

lain terhadap kedua *item* tersebut, sebagai contoh kita akan menghitung nilai kemiripan antara laptop a dan laptop d.

$$\text{sim}(a, d) = \frac{(4 - 2,5)(1 - 2,5)}{\sqrt{(4 - 2,5)^2} + \sqrt{(1 - 2,5)^2}} = -1$$

dari perhitungan tersebut diperoleh nilai kemiripan antara laptop a dan laptop d adalah -1. Lakukan perhitungan pada semua item hingga

hasilnya menjadi seperti tabel di bawah ini :

Tabel 2 nilai kemiripan antar item

	lapto p a	lapto p b	lapto p c	lapto p d	lapto p e	lapt op f
lapto p a	1	0	0	-1	0	0
lapto p b	0	1	0,78	-0,96	0	-1
lapto p c	0	0,78	1	-0,65	1	-0,94
lapto p d	-1	-0,96	-0,65	1	-1	0,99
lapto p e	0	0	1	-1	1	-1
lapto p f	0	-1	-0,94	0,99	-1	1

3. menghitung bobot prediksi

Setelah kita mendapatkan nilai kemiripan antar *item* langkah terakhir dalam menghasilkan rekomendasi adalah dengan menghitung nilai bobot prediksi dengan menggunakan algoritma *weighted sum*. Berikut adalah rumus dari algoritma *weighted sum*:

$$P(u, j) = \frac{\sum_{i \in j} (R_{u,i} * S_{i,j})}{\sum_{i \in j} |S_{i,j}|}$$

Keterangan :

$P(u, j)$ = Prediksi untuk user u pada item j.

$\sum_{i \in j}$ = Himpunan item yang mirip dengan item j.

$R_{u,i}$ = Rating user u pada item i.

$S_{i,j}$ = Nilai kemiripan antara item i dan item j.

Misal kita akan menghitung bobot prediksi user 1 terhadap laptop a sebagai berikut :

$$p(\text{user1, laptop a}) = \frac{(5 * 0) + (4 * 0) + (3 * -1) + (0 * 0) + (0 * 0)}{|0| + |0| + |-1| + |0| + |0|} = -3$$

Dari perhitungan di atas kita bisa lihat bahwa bobot prediksi user1 terhadap laptop a adalah -3. Lakukan perhitungan prediksi pada semua user terhadap setiap laptop. Hasilnya dapat dilihat pada tabel 3 dibawah ini:

Tabel 3 bobot prediksi

	laptop a	laptop b	laptop c	laptop d	laptop e	laptop f
user 1	-3					-1,47
user 2	-2					
user 3	0		-0,14		-1	
user 4		-0,35			-0,33	
user 5	-4					
user 6	-4				-1,33	-0,26

Dari tabel 3 dapat dilihat bahwa yang memiliki bobot nilai yang mendekati 1 adalah item yang akan direkomendasikan, dengan begitu member 1 mendapatkan rekomendasi laptop e, member 2 mendapatkan rekomendasi laptop b dan seterusnya.

Metode *collaborative filtering* memiliki beberapa kelebihan yaitu rekomendasi tetap akan berkerja dalam keadaan dimana konten sulit dianalisis sekalipun, namun metode ini juga memiliki kekurangan yaitu membutuhkan parameter rating, sehingga jika ada item baru sistem tidak akan merekomendasikan item tersebut.

Content based filtering

Dalam *content based filtering* kita tidak membutuhkan parameter rating dalam menghasilkan rekomendasi melainkan dengan deskripsi dari suatu item atau deskripsi dari profil pengguna sesuai

dengan konten/fitur yang diinginkannya untuk menghasilkan rekomendasi. Untuk mendapatkan data konten dari suatu item atau pun profil pengguna kita bisa menggunakan algoritma TF-IDF dalam pemrosesan datanya dalam mendapatkan nilai bobot pada tiap-tiap laptop. Tahapannya sebagai berikut :

1. Menentukan *query term*

Sebagai contoh misal user1 ingin rekomendasi laptop dengan spesifikasi yang disukainya sebagai berikut :

Tabel 4 spekulasi user

Brand	apple, asus
Series	Aspire,MacBook-Air,MacBook-Pro
OS	Linux Mac-OS-X
Tahun rilis	2016
Ukuran layar	14 inch
Resolusi layar	1366-x-768 pixel
Prosesor	Intel-Core-i5
kecepatan	2.2 Ghz
Ram	4 GB
Storage	512 GB
Harga	6149000

Dari tabel 4 di atas data dari spekulasi laptop yang user sukai digabungkan untuk menjadi sebuah *query*, untuk pembuatan dokumennya hanya spesifikasi laptopnya saja yang diambil dan digabungkan untuk menjadi dokumen.

Tabel 5 dokument data laptop

Doc	Isi dokument
Q	apple, asus Aspire,MacBook-Air,MacBook-Pro Linux Mac-OS-X 2016 14 inch 1366-x-768 Intel-Core-i5 2.2 Ghz 4 gb 512 gb 6149000
1	Nama : Aspire E5-475G Core i5-7200U Spesifikasi: acer aspire DOS 2015 14 inch 1366 x 768 pixel Intel Core i5 2.5 Ghz 4 gb 1000 gb 6149000
2	Nama : Acer Aspire ES1-432-

	C56Y / C5GA / C52R Spesifikasi: acer aspire linux 2016 14 inch 1366 x 768 pixel Intel Celeron 1.1 Ghz 2 gb 500 gb 3150000
3	Nama : Apple Macbook Pro MF839 Retina Spesifikasi: apple MacBook Pro Mac OS X 2014 13 inch 2560 x 1600 pixel Intel Core i5 2.7 Ghz 8 gb 128 gb 14454000
4	Nama : Apple MacBook Air MMGG2 Spesifikasi: apple MacBook Air Mac OS X 2016 13 inch 1440 x 900 pixel Intel Core i7 2.2 Ghz 8 gb 512 gb 13954600
5	Nama : Asus X441SA-BX001D / BX002D / BX003D / BX004D Spesifikasi: asus x series DOS 2016 14 inch 1366 x 768 pixel Intel Celeron 2.48 Ghz 2 gb 500 gb 3185000
6	Nama : Asus A455LA-WX667D / WX668D / WX669D Spesifikasi: asus x series DOS 2016 14 inch 1366 x 768 pixel Intel Core i3 2 Ghz 4 gb 500 gb 4725000

2. Perhitungan bobot TF-IDF

Dalam menghitung bobot TF-IDF dipakai rumus sebagai berikut:

$$IDF = (D/DF)$$

$$W = TF * (IDF + 1)$$

Keterangan :

W : bobot setiap dokumen

TF : jumlah kemunculan kata atau term dalam dokumen

D = jumlah semua dokumen

DF = jumlah dokumen yang mengandung kata (term)

IDF = inverse document frequency

Wijaya,
Sistem Rekomendasi Laptop Menggunakan *Collaborative Filtering* Dan *Content-Based Filtering*

Berikut adalah perhitungan untuk TF, DF dan IDF :

Tabel 6 perhitungan TF,DF dan IDF

Query	TF						D F	D / D F	ID F	ID F+ 1
	d 1	d 2	d 3	d 4	d 5	d 6				
Apple			1	1			2	3	0,4 771 21	1,4 771 21
Asus					1	1	2	3	0,4 771 21	1,4 771 21
Aspire	1	1					2	3	0,4 771 21	1,4 771 21
Mac Book air				1			1	6	0,7 781 51	1,7 781 51
Mac Book pro			1				1	6	0,7 781 51	1,7 781 51
Linux		1					1	6	0,7 781 51	1,7 781 51
2016		1		1	1	1	4	5	0,1 760 91	1,1 760 91
Mac OS X			1	1			2	3	0,4 771 21	1,4 771 21
Intel-Core i5	1		1				2	3	0,4 771 21	1,4 771 21
14 inch	1	1			1	1	4	5	0,1 760 91	1,1 760 91
1366-x-768	1	1			1	1	4	5	0,1 760 91	1,1 760 91
2.2 Ghz				1			1	6	0,7 781 51	1,7 781 51
4 GB	1					1	2	3	0,4 771 21	1,4 771 21
512 GB				1			1	6	0,7 781 51	1,7 781 51
6149 000	1						1	6	0,7 781 51	1,7 781 51

Selanjutnya menghitung nilai bobot TF-IDF sebagai berikut :

Tabel 7 Bobot TF-IDF

W = TF * IDF + 1					
d1	d2	d3	d4	d5	d6
0	0	1,47712 1255	1,477 121	0	0
0	0	0	0	1,477 121	1,477 121

1,47712 1255	1,477 121	0	0	0	0
0	0	0	1,778 151	0	0
0	0	1,77815 125	0	0	0
0	1,778 151	0	0	0	0
0	1,176 091	0	1,176 091	1,176 091	1,176 091
0	0	1,47712 1255	1,477 121	0	0
1,47712 1255	0	1,47712 1255	0	0	0
1,17609 1259	1,176 091	0	0	1,176 091	1,176 091
1,17609 1259	1,176 091	0	0	1,176 091	1,176 091
0	0	0	1,778 151	0	0
1,47712 1255	0	0	0	0	1,477 121
0	0	0	1,778 151	0	0
1,77815 125	0	0	0	0	0
Jumlah bobot setiap dokumen					
8,56169 7533	6,783 546	6,20951 5015	9,464 788	5,005 395	6,482 516

Hasil dari perhitungan bobot diatas diketahui bahwa nilai bobot tertinggi terdapat pada document 4 bernilai 9,464788, sehingga rekomendasi laptop untuk user1 adalah d4.

Metode content-based filtering memiliki beberapa kelebihan yaitu : dapat memberikan rekomendasi item yang belum pernah dirating sekalipun, kekurangan metode ini adalah tidak dapat merekomendasikan item bagi user baru yang belum pernah melakukan aktivitas apapun.

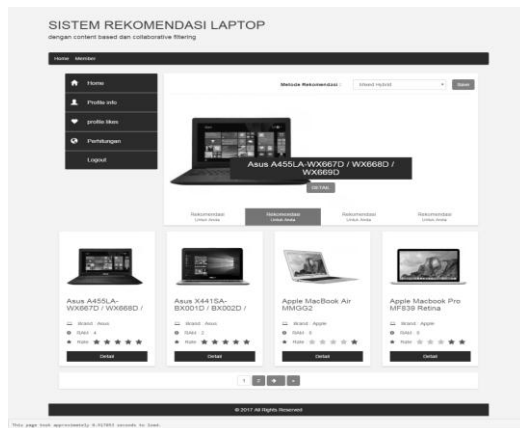
HASIL DAN PENGUJIAN

Implementasi antar muka

1. Member (halaman awal)

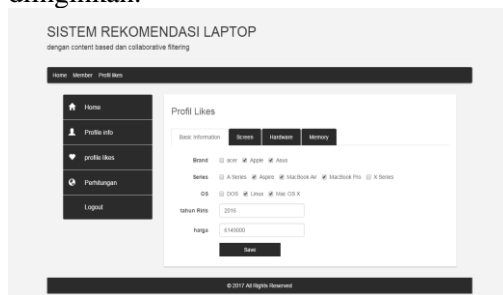
Merupakan tampilan awal bagi member saat berhasil login dengan benar.

Wijaya,
Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based Filtering



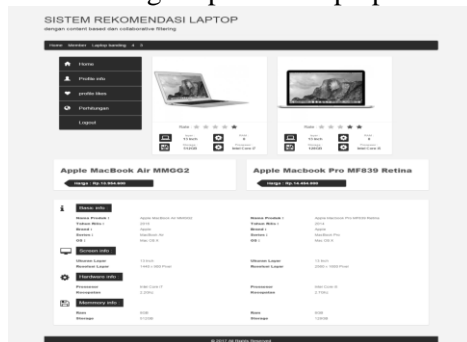
Gambar 2 halaman awal bagi member/user

2. Member (profil likes)
Merupakan tampilan bagi member untuk mengisi ketertarikan laptop yang diinginkan.



Gambar 3 halaman profil likes bagi member/user

3. Member (banding laptop)
Tampilan yang digunakan untuk membandingkan spesifikasi laptop.



Gambar 4 halaman laptop banding bagi member/user

Pengujian metode

Pada pengujian metode dilakukan dengan cara membandingkan perhitungan manual dengan

perhitungan yang dihasilkan oleh sistem untuk menguji kebenaran sistem rekomendasi laptop menggunakan collaborative filtering dan content based filtering.

Hasil pengujian metode

1. Collaborative filtering

Berikut adalah perhitungan collaborative filtering untuk merekomendasikan laptop untuk user1 dibawah ini merupakan data laptop yang dipakai.

Tabel 8 laptop data

item	Nama
1	Aspire E5-475G Core i5-7200U
2	Acer Aspire ES1-432-C56Y / C5GA / C52R
3	Apple Macbook Pro MF839 Retina
4	Apple MacBook Air MMGG2
5	Asus X441SA-BX001D / BX002D / BX003D / BX004D / BX005D
6	Asus A455LA-WX667D / WX668D / WX669D
7	Sony Vaio SVD13213SG
8	Sony Vaio SVF14216SG
9	Dell Inspiron 14-3458
10	Dell XPS 13 Core i5-5200U

Berikut adalah matrix user x item antara rating item user1 dengan pengguna lain :

Tabel 9 rating user1 dan pengguna lain

	i 1	i 2	i 3	i 4	i 5	i 6	i 7	i 8	i 9	i 10	Rata-rata
u 1		3	4	2	5	4		3			3,5
u 2			5		4			3			4
u 3			2		2					5	3
u 4		3							4		3,5
u 5	5	3					5	2			3,75

Kemudian, hitung kemiripan antar item dengan rumus adjusted cosine similarity :

$$sim(i, j) = \frac{\sum_{u \in U} (R_{ui} - \bar{R}_u)(R_{uj} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{ui} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{uj} - \bar{R}_u)^2}}$$

$$\begin{aligned} \text{sim}(i1, i2) &= \\ \frac{(5-3,75)(3-3,75)}{\sqrt{(5-3,75)^2} \sqrt{(3-3,75)^2}} &= \\ \frac{-0,9375}{0,9375} &= -1 \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i7) &= \\ \frac{(3-3,75)(5-3,75)}{\sqrt{(3-3,75)^2} \sqrt{(5-3,75)^2}} &= \\ \frac{-0,9375}{0,9375} &= -1 \end{aligned}$$

$$\begin{aligned} \text{sim}(i1, i7) &= \\ \frac{(5-3,75)(5-3,75)}{\sqrt{(5-3,75)^2} \sqrt{(5-3,75)^2}} &= \\ \frac{1,5625}{1,5625} &= 1 \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i8) &= \\ \frac{(3-3,5)(3-3,5) + (3-3,75)(2-3,75)}{\sqrt{(3-3,5)^2 + (3-3,75)^2} \sqrt{(3-3,5)^2 + (2-3,75)^2}} &= \\ \frac{1,5625}{1,6405} &= 0,9524 \end{aligned}$$

$$\begin{aligned} \text{sim}(i1, i8) &= \\ \frac{(5-3,75)(2-3,75)}{\sqrt{(5-3,75)^2} \sqrt{(2-3,75)^2}} &= \\ \frac{-2,1875}{2,1875} &= -1 \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i9) &= \\ \frac{(3-3,5)(4-3,5)}{\sqrt{(3-3,5)^2} \sqrt{(4-3,5)^2}} &= \frac{-0,25}{0,25} = \\ -1 & \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i3) &= \\ \frac{(3-3,5)(4-3,5)}{\sqrt{(3-3,5)^2} \sqrt{(4-3,5)^2}} &= \frac{-0,25}{0,25} = \\ -1 & \end{aligned}$$

$$\begin{aligned} \text{sim}(i3, i4) &= \\ \frac{(4-3,5)(2-3,5)}{\sqrt{(4-3,5)^2} \sqrt{(2-3,5)^2}} &= \frac{-0,75}{0,75} = \\ -1 & \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i4) &= \\ \frac{(3-3,5)(2-3,5)}{\sqrt{(3-3,5)^2} \sqrt{(2-3,5)^2}} &= \frac{0,75}{0,75} = 1 \end{aligned}$$

$$\begin{aligned} \text{sim}(i3, i5) &= \\ \frac{(4-3,5)(5-3,5) + (5-4)(4-4) + (2-3)(2-3)}{\sqrt{(4-3,5)^2 + (5-4)^2 + (2-3)^2} \sqrt{(5-3,5)^2 + (4-4)^2 + (2-3)^2}} &= \\ \frac{1,75}{2,7041} &= 0,6471 \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i5) &= \\ \frac{(3-3,5)(5-3,5)}{\sqrt{(3-3,5)^2} \sqrt{(5-3,5)^2}} &= \frac{-0,75}{0,75} = \\ -1 & \end{aligned}$$

$$\begin{aligned} \text{sim}(i3, i6) &= \\ \frac{(4-3,5)(4-3,5)}{\sqrt{(4-3,5)^2} \sqrt{(4-3,5)^2}} &= \frac{0,25}{0,25} = 1 \end{aligned}$$

$$\begin{aligned} \text{sim}(i2, i6) &= \\ \frac{(3-3,5)(4-3,5)}{\sqrt{(3-3,5)^2} \sqrt{(4-3,5)^2}} &= \frac{-0,25}{0,25} = \\ -1 & \end{aligned}$$

$$\begin{aligned} \text{sim}(i3, i8) &= \\ \frac{(4-3,5)(3-3,5) + (5-4)(3-4)}{\sqrt{(4-3,5)^2 + (5-4)^2} \sqrt{(3-3,5)^2 + (3-4)^2}} &= \\ \frac{-1,25}{1,25} &= -1 \end{aligned}$$

$$\text{sim}(i3, i10) = \frac{(2-3)(5-3)}{\sqrt{(2-3)^2} \sqrt{(5-3)^2}} = \frac{-2}{2} = -1$$

$$\text{sim}(i4, i5) = \frac{(2-3,5)(5-3,5)}{\sqrt{(2-3,5)^2} \sqrt{(5-3,5)^2}} = \frac{-2,5}{2,5} =$$

-1

$$\text{sim}(i4, i6) = \frac{(2-3,5)(4-3,5)}{\sqrt{(2-3,5)^2} \sqrt{(4-3,5)^2}} = \frac{-0,75}{0,75} =$$

-1

$$\text{sim}(i4, i8) = \frac{(2-3,5)(3-3,5)}{\sqrt{(2-3,5)^2} \sqrt{(3-3,5)^2}} = \frac{0,75}{0,75} = 1$$

$$\text{sim}(i5, i6) = \frac{(5-3,5)(4-3,5)}{\sqrt{(5-3,5)^2} \sqrt{(4-3,5)^2}} = \frac{0,75}{0,75} = 1$$

$$\text{sim}(i5, i8) = \frac{(5-3,5)(3-3,5)+(4-4)(3-4)}{\sqrt{(5-3,5)^2+(4-4)^2} \sqrt{(3-3,5)^2+(3-4)^2}} = \frac{-0,75}{1,6770} = -0,447$$

$$\text{sim}(i5, i10) = \frac{(2-3)(5-3)}{\sqrt{(2-3)^2} \sqrt{(5-3)^2}} = \frac{-2}{2} = -1$$

$$\text{sim}(i6, i8) = \frac{(4-3,5)(3-3,5)}{\sqrt{(4-3,5)^2} \sqrt{(3-3,5)^2}} = \frac{-0,25}{0,25} =$$

-1

$$\text{sim}(i7, i8) = \frac{(5-3,75)(2-3,75)}{\sqrt{(5-3,75)^2} \sqrt{(2-3,75)^2}} = \frac{-2,1875}{-2,1875} = -1$$

berikut adalah matrik kemiripan antar item dari hasil perhitungan diatas :

tabel 10 kemiripan antar item

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10
i1		-1	0	0	0	0	1	-1	0	0
i2	-1		-1	1	-1	-1	-1	0,95 24	-1	0
i3	0	-1		-1	0,64 71	1	0	-1	0	-1
i4	0	1	-1		-1	-1	0	1	0	0
i5	0	-1	0,64 71	-1		1	0	-0,44 7	0	-1
i6	0	-1	1	-1	1		0	-1	0	0
i7	1	-1	0	0	0	0		-1	0	0
i8	-1	0,95 24	-1	1	-0,44 7	-1	-1		0	0
i9	0	-1	0	0	0	0	0	0		0
i10	0	0	-1	0	-1	0	0	0	0	

hasil perhitungan kemiripan yang dilakukan oleh sistem, terlihat bahwa hasil dari perhitungan manual dan hasil dari perhitungan oleh sistem rekomendasi memiliki hasil yang sama. Selanjutnya adalah perhitungan prediksi menggunakan algoritma *weighted sum* :

$$P(u, j) = \frac{\sum_{i \in J} (R_{u,i} * S_{i,j})}{\sum_{i \in J} |S_{i,j}|}$$

$$p(\text{user1}, i1) = \frac{(3*-1)+(4*0)+(2*0)+(5*0)+(4*0)+(0*1)+(3*-1)+(0*0)+(0*0)}{|-1|+|0|+|0|+|0|+|0|+|1|+|-1|+|0|+|0|} = \frac{-6}{3} = -2$$

Wijaya,
Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based Filtering

window 8						1	1		
2016		1		1	1	1			
14 inch	1	1			1	1		1	1
1366-x-768 pixel	1	1			1	1	1	1	1
Intel-Core-i5	1		1				1		1
2.2 Ghz				1					1
4 GB	1					1	1		1
1000 GB	1								
14454000			1						

Menghitung nilai IDF sebagai berikut:

Tabel 14 perhitungan nilai IDF

DF	d/df	Idf	idf + 1
2	5	0,698970004	1,698970004
2	5	0,698970004	1,698970004
1	10	1	2
1	10	1	2
2	5	0,698970004	1,698970004
1	10	1	2
2	5	0,698970004	1,698970004
4	2,5	0,397940009	1,397940009
6	1,666666667	0,22184875	1,22184875
8	1,25	0,096910013	1,096910013
4	2,5	0,397940009	1,397940009
2	5	0,698970004	1,698970004
4	2,5	0,397940009	1,397940009
1	10	1	2
1	10	1	2

Perhitungan manual untuk bobot tf-idf sebagai berikut :

Tabel 15 perhitungan bobot TF-IDF

W=TF * (idf+1)									
d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
0	0	1,697	1,697	0	0	0	0	0	0
0	0	0	0	0	0	1,6997	1,69897	0	0
0	0	0	2	0	0	0	0	0	0
0	0	2	0	0	0	0	0	0	0
0	0	0	0	0	0	1,6997	1,69897	0	0
0	0	0	0	0	0	0	0	0	2
0	0	0	0	0	0	1,6997	1,69897	0	0
0	1,39794	0	1,39794	1,39794	1,39794	0	0	0	0
1,1849	1,1849	0	0	1,1849	1,1849	0	1,1849	1,1849	0
1,09691	1,09691	0	0	1,09691	1,09691	1	1	1	1
1,39794	0	1,39794	0	0	0	1,39794	0	0	1,39794
0	0	0	1,69897	0	0	0	0	1,69897	0
1,39794	0	0	0	0	1,39794	1,39794	0	1,39794	0

						4			
d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
2	0	0	0	0	0	0	0	0	0
0	0	2	0	0	0	0	0	0	0
d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
7, 11, 46, 39	3, 71, 66, 99	7, 9, 6, 1	6, 9, 5, 8	3, 71, 66, 99	5, 11, 46, 39	8, 9, 8, 7	7, 41, 56, 69	5, 41, 56, 69	4, 9, 4, 5

KESIMPULAN

Setelah melewati tahap pengimplementasian dan pengujian terhadap sistem, maka terdapat beberapa kesimpulan yang ditemukan pada penelitian ini, yaitu :

1. Dengan teknik penggabungan secara *mixed hybrid* antara metode *collaborative filtering* dan *content-based filtering* dapat menghasilkan sistem rekomendasi laptop yang mampu menutupi kekurangan dari setiap metode yang digunakan.
2. Waktu eksekusi yang dibutuhkan dipengaruhi oleh jumlah data, dapat disimpulkan bahwa semakin besar jumlah data yang digunakan maka semakin besar juga waktu eksekusi. Pada hasil waktu eksekusi tersebut dapat disimpulkan bahwa metode *content-based filtering* memiliki waktu eksekusi lebih cepat dari metode *collaborative filtering*.

SARAN

Saran untuk penelitian di bidang sistem rekomendasi selanjutnya, yaitu:

1. Untuk penelitian selanjutnya diharapkan untuk menggunakan algoritma yang berbeda dalam merekomendasikan suatu item untuk mendapatkan hasil yang lebih baik.

2. Untuk penelitian selanjutnya dapat menambahkan beberapa parameter dalam memberikan rekomendasi seperti histori, komentar, ataupun *likes/dislikes*.
3. Untuk penelitian selanjutnya diharapkan dalam pembuatan user profil dilakukan secara implisit yaitu dimana user tidak menyadari ia telah memberikan masukan terhadap sistem.

DAFTAR PUSTAKA

- Adomavicius, G., & Kwon, Y. (2015). Multi-criteria recommender systems. In *Recommender systems handbook* (pp. 847-880). Springer, Boston, MA.
- Arifin, W. 2014. Implementasi hybrid (*content based* dan *collaborative filtering*) pada sistem rekomendasi software antivirus dengan multi-criteria rating. Fakultas ilmu komputer dan teknologi informasi, universitas sumatera utara. Medan.
- Burke, R. (2007). Hybrid web recommender systems. In *The adaptive web* (pp. 377-408). Springer, Berlin, Heidelberg.
- Handrico, A. (2012). Sistem rekomendasi buku perpustakaan fakultas sains dan teknologi dengan metode *collaborative filtering*. Jurusan teknik informatika, Fakultas sains dan Teknologi universitas Islam negeri Sultan Syarif Kasim Riau. Pekanbaru
- Jannach, D., Karakaya, Z., & Gedikli, F. (2012, June). Accuracy improvements for multi-criteria recommender systems. In *Proceedings of the 13th ACM conference on electronic commerce* (pp. 674-689). ACM.
- Ricci, F., Rokach, L., & Saphira, B. (2010). Introduction to recommender systems handbook. In F. Ricci, L. Rokach, B. Saphira, &

- P. B. Kantor (Eds.), *Recommender systems handbook* (pp. 1–29). New York: Springer.
- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001, April). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web* (pp. 285-295). ACM.
- Wibowo, A. (2010). Recommender System di Perpustakaan Universitas Kristen Petra menggunakan Rocchio Relevance Feedback dan Cosine Similarity. In *Industrial Electronic Seminar*.
- Wiranto, E. (2010). Konsep Multicriteria Collaborative Filtering Untuk Perbaikan Rekomendasi. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*.
-