

**IMPLEMENTASI SISTEM REKOMENDASI PADA PRODUK LAPTOP
MENGUNAKAN TEKNIK CONTENT BASED FILTERING**

SKRIPSI



disusun oleh

Mashudi Rohmat

20.22.2421

**PROGRAM SARJANA
PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**IMPLEMENTASI SISTEM REKOMENDASI PADA PRODUK LAPTOP
MENGUNAKAN TEKNIK CONTENT BASED FILTERING**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Sistem Informasi



disusun oleh

Mashudi Rohmat

20.22.2421

**PROGRAM SARJANA
PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

IMPLEMENTASI SISTEM REKOMENDASI PADA PRODUK LAPTOP MENGUNAKAN TEKNIK CONTENT BASED FILTERING

yang dipersiapkan dan disusun oleh

Mashudi Rohmat

20.22.2421

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal - Agustus 2021

Dosen Pembimbing,

Arif Dwi Laksito, M.Kom.

NIK. 190302150

PENGESAHAN

SKRIPSI

**IMPLEMENTASI SISTEM REKOMENDASI PADA PRODUK LAPTOP
MENGUNAKAN TEKNIK CONTENT BASED FILTERING**

yang dipersiapkan dan disusun oleh

Mashudi Rohmat

20.22.2421

telah dipertahankan di depan Dewan Penguji
pada tanggal - September 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Afrig Aminuddin, S.Kom., M.Eng
NIK. 190302351

Rizqi Sukma Kharisma, M.Kom
NIK. 190302215

Arif Dwi Laksito, M.Kom
NIK. 190302150

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 April 2021

**DEKAN FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA**

Hanif Al Fatta, M. Kom
NIK. 190302096

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 25 Agustus 2021

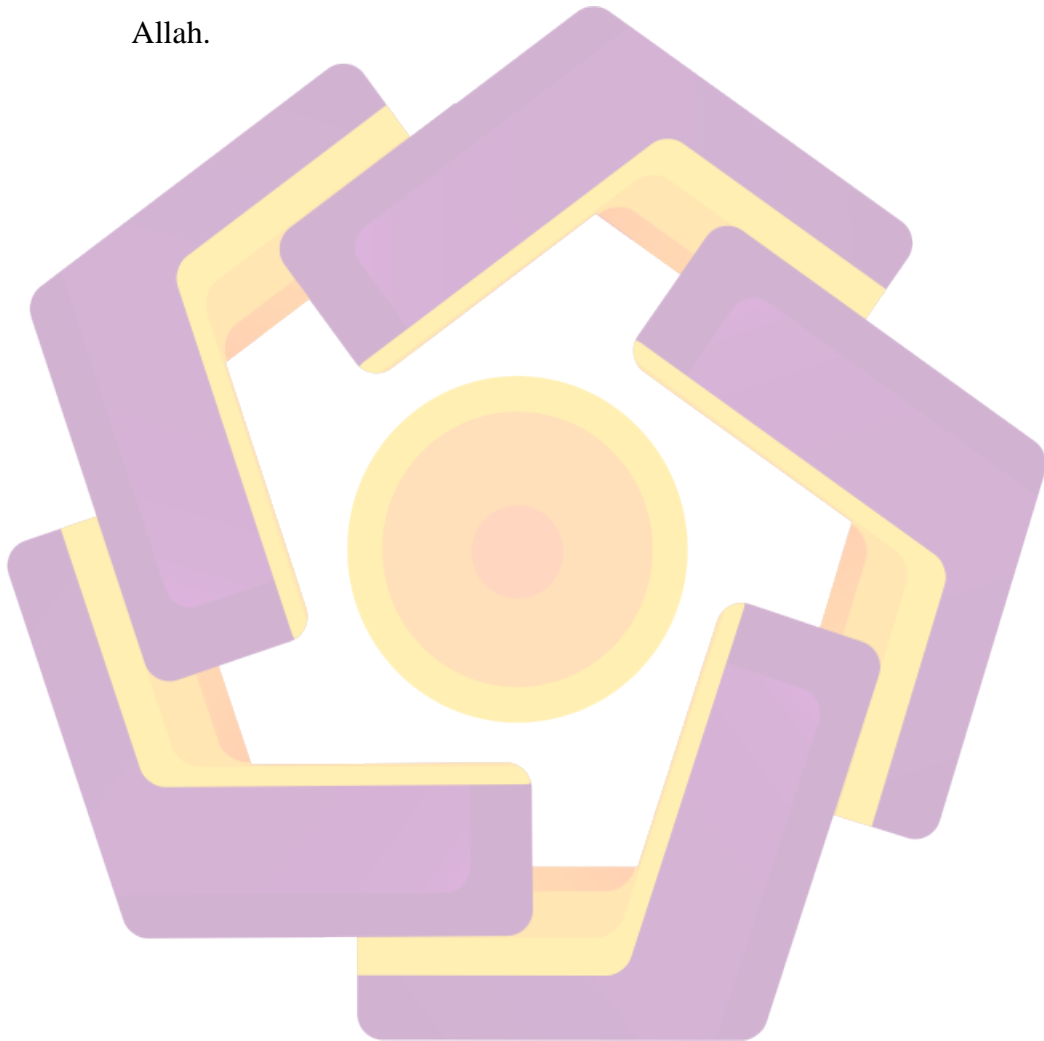


Mashudi Rohmat

NIM. 20.22.2421

MOTTO

- Sebaik – baiknya orang adalah orang yang bermanfaat bagi orang lain.
- Gagal setelah mencoba lebih terhormat dari pada gagal sebelum mencoba sama sekali.
- Tidak diciptakan Jin dan Manusia kecuali hanya untuk beribadah kepada Allah.



PERSEMBAHAN

Skripsi ini saya persembahkan untuk orang – orang di belakang layar:

- Kedua orang tua tercinta.
- Keluarga Bapak Sularto (Pakde).
- Keluarga besar saya.
- Bapak Arif Dwi Laksito M.Kom selaku Dosen Pembimbing.
- Teman – teman yang memberikan support yang tidak bisa saya sebutkan satu per satu.



KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan Skripsi dengan judul “IMPLEMENTASI SISTEM REKOMENDASI PADA PRODUK LAPTOP MENGGUNAKAN TEKNIK CONTENT BASED FILTERING” ini dengan baik dan lancar. Penyelesaian laporan ini merupakan bagian dari Skripsi sebagai syarat wajib kelulusan untuk menyelesaikan program Strata 1.

Dalam penyusunan laporan ini penulis banyak mendapatkan bantuan dari berbagai pihak, oleh sebab itu pada kesempatan ini penulis ingin menyampaikan rasa terima kasih kepada: Bp. Arif Dwi Laksito selaku Dosen Pembimbing, yang telah memberikan dukungan, arahan, bimbingan, dan motivasi selama proses penyelesaian Skripsi.

Penulis menyadari bahwa laporan ini masih jauh dari kesempurnaan, baik dari materi maupun teknik penyajiannya. Untuk itu penulis mengharapkan kritik dan saran yang membangun dari pembaca demi penyempurnaan dalam menyelesaikan laporan ini. Semoga laporan ini dapat memberikan banyak manfaat bagi penulis dan pembaca.

Yogyakarta ,November 2021

Penulis

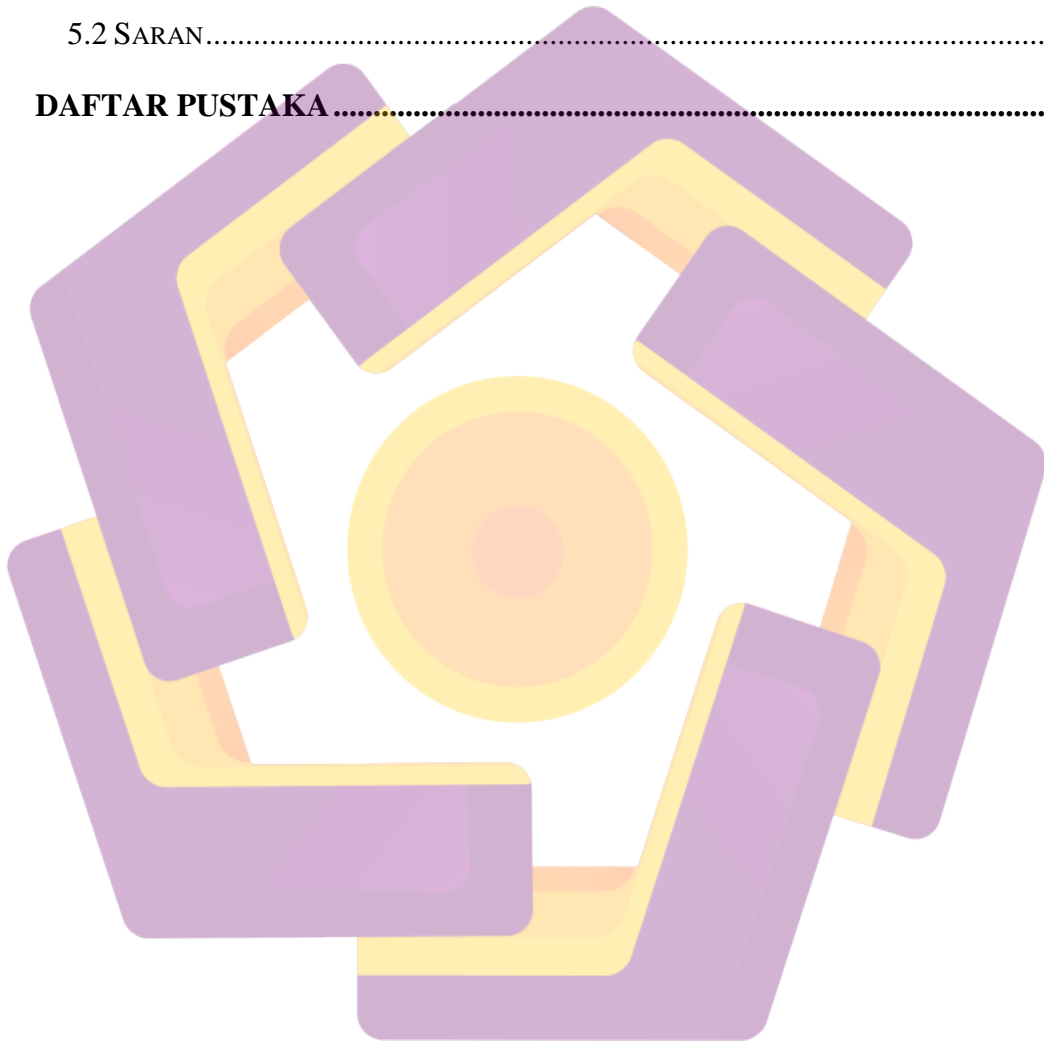
DAFTAR ISI

JUDUL	II
PERNYATAAN.....	IV
MOTTO	VI
PERSEMBAHAN.....	VI
KATA PENGANTAR.....	VIII
DAFTAR ISI.....	IX
DAFTAR TABEL	XIII
DAFTAR GAMBAR.....	XV
INTISARI	XVI
ABSTRACT.....	XVII
BAB I PENDAHULUAN.....	1
1.1 LATAR BELAKANG	1
1.2 RUMUSAN MASALAH	3
1.3 BATASAN MASALAH	3
1.4 MAKSUD DAN TUJUAN PENELITIAN	3
1.5 MANFAAT PENELITIAN.....	4
1.6 METODE PENELITIAN	4
1.6.1 Metode Pengumpulan Data.....	4
1.6.2 Metode Analisis.....	4
1.6.3 Metode Perancangan	5
1.6.4 Metode Pengujian	5
1.7 SISTEMATIKA PENULISAN.....	5
BAB II LANDASAN TEORI	7
2.1 KAJIAN PUSTAKA	7
2.2 KAGGLE	12

2.3 SISTEM REKOMENDASI.....	12
2.4 METODE SISTEM REKOMENDASI.....	13
2.5 CONTENT BASED FILTERING.....	14
2.6 TF-IDF.....	15
2.7 COSINE SIMILARITY.....	16
2.8 FLASK.....	17
2.9 PYTHON.....	17
2.10 SDLC (SOFTWARE DEVELOPMENT LIVE CYCLE).....	18
2.11 WATERFALL SDLC MODEL.....	18
2.12 DATABASE.....	21
2.13 ERD (ENTITY RELATIONAL DIAGRAM).....	22
2.14 FLOWCHART.....	23
2.15 BLACK BOX TESTING.....	24
2.16 NDCG (NORMALIZED DISCOUNTED CUMULATIVE GAIN).....	24
BAB III METODE PENELITIAN.....	26
3.1 TAHAPAN PENELITIAN.....	26
3.1.1 <i>Persiapan Data</i>	27
3.1.1.1 Pengumpulan Data.....	27
3.1.1.2 Pemilihan Data.....	27
3.1.2 <i>Pengolahan Data</i>	27
3.1.3 <i>Perhitungan Manual Pengolahan Data</i>	28
3.1.3.1 Penentuan Dokumen.....	28
3.1.3.2 Perhitungan TF-IDF.....	30
3.1.3.2 Perhitungan Cosine Similarity.....	38
3.2 EVALUASI NDCG (NORMALIZED DOCUMENT CUMULATIVE GAIN).....	44
3.3 PERANCANGAN HALAMAN <i>USER INTERFACE</i>	46
3.3.1 <i>Halaman Utama</i>	46
3.3.2 <i>Halaman Rekomendasi</i>	46
3.3.2 <i>Halaman Info</i>	47

3.4	RANCANGAN PERANGKAT LUNAK	48
3.4.1	<i>Flowchart</i>	48
3.4.2	<i>Rancangan Database</i>	49
3.5	ALAT DAN BAHAN PENELITIAN.....	50
3.5.1	<i>Alat</i>	50
3.5.2	<i>Kebutuhan Software</i>	51
BAB IV IMPLEMENTASI DAN PEMBAHASAN		52
4.1	PERSIAPAN DATA.....	52
4.1.1	<i>Pengumpulan Data</i>	52
4.1.2	<i>Pemilihan Data</i>	55
4.2	PENGOLAHAN DATA	55
4.2.1	<i>Implementasi Pengolahan Data</i>	55
4.2.1.1	Penentuan Query atau Kata Kunci	55
4.2.1.2	Membaca Data Laptop	56
4.2.1.3	Membersihkan Data	56
4.2.1.4	Menggabungkan Data Parameter.....	57
4.2.1.5	Proses TF-IDF.....	57
4.2.1.6	Proses Cosine Similarity	57
4.2.1.7	Perangkingan.....	58
4.2.1.8	Menyimpan Data Rekomendasi ke Database.....	58
4.2.1.9	Memanggil Rekomendasi dari Database.....	59
4.2.1.10	Menampilkan Data Rekomendasi ke User Interface.....	59
4.3	IMPLEMENTASI USER INTERFACE.....	60
4.3.1	<i>Halaman Utama</i>	61
4.3.2	<i>Halaman Rekomendasi</i>	61
4.3.3	<i>Halaman Info</i>	62
4.4	UJI COBA SISTEM.....	63
4.4.1	<i>Black Box Testing</i>	63
4.4.2	<i>Evaluasi NDCG (Normalized Discounted Cumulative Gain)</i>	64
4.4.2.1	Data Uji	64

4.4.2.2	Menghitung NDCG.....	73
4.4.2.3	Menghitung Rata -rata NDCG	80
4.4.3	<i>Pengujian Waktu Pengolahan Data</i>	81
BAB V	PENUTUP	83
5.1	KESIMPULAN	83
5.2	SARAN.....	84
DAFTAR PUSTAKA	86



DAFTAR TABEL

Tabel 2.1 Tabel Rangkuman Kajian Pustaka.....	9
Tabel 2.2 History versi Python.....	18
Tabel 2.3 Simbol – simbol Entity Relational Diagram.....	22
Tabel 2.4 Simbol – simbol <i>Flowchart</i>	24
Tabel 3.1 Data Rekomendasi.....	29
Tabel 3.2 Perhitungan jumlah dokumen dan TF.....	30
Tabel 3.3 Perhitungan nilai IDF.....	33
Tabel 3.4 Perhitungan TF-IDF.....	36
Tabel 3.5 Perhitungan panjang vector dokumen.....	38
Tabel 3.6 Perhitungan TF-IDF Query.....	40
Tabel 3.7 Perkalian TF-IDF Kata kunci dengan TF-IDF Dokumen.....	42
Tabel 3.8 Peringkat rekomendasi.....	44
Tabel 3.9 Nilai relevansi.....	45
Tabel 3.10 Alat Penelitian.....	50
Tabel 3.11 Kebutuhan Software.....	51
Tabel 4.1 Data laptop.csv.....	53
Tabel 4.2 Black Box Testing.....	63
Tabel 4.2 Skenario 1.....	64
Tabel 4.3 Skenario 2.....	65
Tabel 4.4 Skenario 3.....	66
Tabel 4.5 Skenario 4.....	67

Tabel 4.6 Skenario 5.....	68
Tabel 4.7 Skenario 6.....	69
Tabel 4.8 Skenario 7.....	69
Tabel 4.9 Skenario 8.....	70
Tabel 4.10 Skenario 9.....	71
Tabel 4.11 Skenario 10.....	72
Tabel 4.12 Nilai Relevansi Skenario 1.....	73
Tabel 4.13 Nilai Relevansi Skenario 2.....	74
Tabel 4.14 Nilai Relevansi Skenario 3.....	75
Tabel 4.15 Nilai Relevansi Skenario 4.....	75
Tabel 4.16 Nilai Relevansi Skenario 5.....	76
Tabel 4.17 Nilai Relevansi Skenario 6.....	77
Tabel 4.18 Nilai Relevansi Skenario 7.....	77
Tabel 4.19 Nilai Relevansi Skenario 8.....	78
Tabel 4.20 Nilai Relevansi Skenario 9.....	79
Tabel 4.21 Nilai Relevansi Skenario 10.....	79
Tabel 4.22 Nilai Rata -rata NDCG.....	80
Tabel 4.23 Pengujian Waktu Pengolahan Data.....	81

DAFTAR GAMBAR

Gambar 2.1 Cara kerja Content Based Filtering pada sistem rekomendasi.....	14
Gambar 2.2 Waterfall SDLC Model.....	19
Gambar 3.1 Diagram Alir Tahapan Penelitian.....	26
Gambar 3.2 Diagram Alir Tahapan Pengolahan Data.....	28
Gambar 3.3 Rancangan Halaman Utama.....	46
Gambar 3.4 Rancangan Halaman Detail.....	47
Gambar 3.5 Rancangan Halaman Info.....	48
Gambar 3.6 Flowchart sistem.....	49
Gambar 3.7 Rancangan ERD.....	50
Gambar 4.1 Kode Kata Kunci Rekomendasi.....	55
Gambar 4.2 Kode Pemanggilan Dokumen.....	56
Gambar 4.3 Kode Pembersihan Data.....	56
Gambar 4.4 Kode Menggabungkan Data.....	57
Gambar 4.5 Kode Mencari TF-IDF.....	57
Gambar 4.6 Kode Mencari Cosine Similarity.....	58
Gambar 4.7 Kode Perangkingan.....	58
Gambar 4.8 Kode Menyimpan Rekomendasi ke Database.....	59
Gambar 4.9 Kode Memanggil Data Rekomendasi dari Database.....	59
Gambar 4.10 Kode Menampilkan Data Rekomendasi.....	60
Gambar 4.11 Implementasi Halaman Utama.....	61
Gambar 4.12 Implementasi Halaman Rekomendasi.....	62
Gambar 4.13 Implementasi Halaman Info.....	63

INTISARI

Variasi merk dan jenis laptop dari berbagai pabrikan mengalami perkembangan yang sangat pesat, banyak pilihan jenis laptop berdasarkan dari segmen pasar yang dibidik, mulai laptop untuk keperluan kantor, notebook, design grafis hingga *gaming*. Banyaknya pilihan ini membuat calon pembeli merasa kesulitan untuk menemukan jenis laptop yang sesuai dengan kebutuhannya.

Menjawab permasalahan di atas maka dibuatlah sistem rekomendasi yang mampu memberikan rekomendasi laptop kepada calon pembeli. Secara garis besar prinsip kerjanya memberikan rekomendasi laptop yang relevan dengan kata kunci yang dipilih oleh *user*, pemberian rekomendasi berdasarkan kemiripan atribut yang melekat pada item laptop seperti type, ukuran layar, jenis processor, RAM, memory, kartu grafis dan lain – lain. Prosesnya dengan melakukan pembobotan setiap *term* yang terdapat pada dokumen menggunakan *TF-IDF*, kemudian mencari kemiripan antara dokumen dengan kata kunci menggunakan *Cosine Similarity*.

Hasil dari penelitian ini berupa peringkat top 6 rekomendasi laptop yang paling relevan. Untuk mengevaluasi hasil rekomendasi menggunakan teknik NDCG (*Normalized Discounted Cumulative Gain*), evaluasi dilakukan dengan 10 skenario pengujian, hasilnya memberikan nilai rata – rata NDCG sebesar 0,9704 yang artinya sistem memberikan performa rekomendasi yang baik.

Kata Kunci: sistem rekomendasi, *content-based filtering*, *tf-idf*, *cosine similarity*, *ndcg*

ABSTRACT

The variety of brands and types of laptops from various manufacturers has growing very rapidly, there are many choices of types of laptops based on the targeted market segment, ranging from laptops for office purposes, notebooks, graphic design to gaming. This number of choices makes it difficult for potential buyers to find the type of laptop that suits their needs.

Answering the problems above, a recommendation system is made that able to provide laptop recommendations to prospective buyers. Broadly speaking, the working principle is to provide laptop recommendations that are relevant to the keywords chosen by the user, giving recommendations based on the similarity of attributes attached to laptop items such as type, screen size, type of processor, RAM, memory, graphics card and others. The process is by weighting each term contained in the document using TF-IDF, then looking for similarities between documents and keywords using Cosine Similarity.

The results of this study are in the form of a ranking of the top 6 most relevant laptop recommendations. To evaluate the results of recommendations using the NDCG (Normalized Discounted Cumulative Gain) technique, the evaluation was carried out with 10 test scenarios, the results gave an average NDCG value of 0.9704 which means the system provides a good recommendation performance.

Keyword: recommender system, content based filtering, tf-idf, cosine similarity, ndcg