

Perancangan Sistem Rekomendasi Kursus Online di Bidang Teknologi Menggunakan Metode Content-Based Filtering

Dina Asritama Wulandari^{1*}, Yola Aulia Wanergi², Haffiant Resditya Akbar³, Andri Khoirul Manaf⁴, Satria Abdul Rofiq⁵, Vihi Atina⁶

¹Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

^{1*}230103056@mhs.udb.ac.id

²Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

²230103081@mhs.udb.ac.id

³Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

³230103062@mhs.udb.ac.id

⁴Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

⁴230103050@mhs.udb.ac.id

⁵Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

⁵230103077@mhs.udb.ac.id

⁶Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

⁶vihi_atina@udb.ac.id

Abstrak— Perkembangan platform kursus online telah menghadirkan beragam pilihan pembelajaran di bidang teknologi, namun banyaknya pilihan tersebut justru menyulitkan pengguna dalam menemukan kursus yang paling sesuai dengan kebutuhan mereka. Penelitian ini bertujuan untuk merancang sistem rekomendasi kursus online menggunakan metode Content-Based Filtering guna membantu pengguna memperoleh rekomendasi kursus yang relevan secara otomatis. Pengembangan sistem dilakukan menggunakan metode *Prototype* yang terdiri dari tahapan *Communication*, *Quick Plan*, dan *Modelling Quick Design*. Proses rekomendasi memanfaatkan pembobotan TF-IDF dan perhitungan *Cosine Similarity* untuk mengukur tingkat kemiripan antara preferensi pengguna dan konten kursus yang tersedia dalam dataset. Hasil pengujian dengan query “Python for Data Science” menunjukkan bahwa sistem berhasil menghasilkan tiga rekomendasi terbaik, yaitu Python for Data Science dengan nilai similarity 0,6907, Python for Beginners sebesar 0,4667, dan Data Analysis with Pandas sebesar 0,3783. Rancangan sistem mencakup pemodelan proses menggunakan flowchart dan use case diagram, serta antarmuka berupa halaman beranda, halaman hasil rekomendasi, dan halaman admin. Sistem yang dihasilkan terbukti mampu mengidentifikasi kursus dengan tingkat kemiripan konten tertinggi dan memberikan kontribusi nyata dalam meningkatkan efisiensi pencarian kursus online di bidang teknologi.

Kata kunci— Sistem Rekomendasi, Content-Based Filtering, TF-IDF, Cosine Similarity, Kursus Online

Abstract— The rapid growth of online course platforms in the technology sector has provided users with a wide range of learning options, yet the abundance of choices makes it difficult to find courses that best match their needs. This study aims to design an online course recommendation system using the Content-Based Filtering method to automatically provide users with relevant course recommendations. The system was developed using the Prototype method, consisting of Communication, Quick Plan, and Modelling Quick Design phases. The recommendation process employs TF-IDF weighting and Cosine Similarity calculation to measure the degree of similarity between user preferences and the course content available in the dataset. Testing with the query “Python for Data Science” demonstrated that the system successfully generated three top recommendations: Python for Data Science with a similarity score of 0.6907, Python for Beginners at 0.4667, and Data Analysis with Pandas at 0.3783. The system design includes process modeling through flowcharts and use case diagrams, as well as a user interface comprising a homepage, recommendation results page, and admin page. The results confirm that the proposed system effectively identifies courses with the highest content similarity and contributes to improving the efficiency of online course discovery in the technology domain.

Keywords— Recommendation System, Content-Based Filtering, TF-IDF, Cosine Similarity, Online Course

I. PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam dunia

pendidikan, khususnya melalui hadirnya berbagai platform kursus online yang memungkinkan masyarakat memperoleh

pengetahuan dan keterampilan secara fleksibel tanpa dibatasi oleh ruang dan waktu [1]. Platform kursus online menyediakan beragam materi pembelajaran dari berbagai bidang seperti pemrograman, analisis data, dan kecerdasan buatan, sehingga memberikan banyak pilihan bagi pengguna untuk meningkatkan kompetensi mereka [2]. Keberadaan layanan tersebut menjadi salah satu solusi dalam mendukung pembelajaran sepanjang hayat serta peningkatan kualitas sumber daya manusia di era digital [3].

Meningkatnya jumlah kursus yang tersedia pada berbagai platform pembelajaran digital memberikan keuntungan berupa banyaknya alternatif pembelajaran yang dapat dipilih pengguna [1]. Namun di sisi lain, kondisi tersebut juga menimbulkan tantangan berupa kesulitan dalam menemukan kursus yang paling sesuai dengan kebutuhan dan minat pengguna karena banyaknya pilihan yang tersedia [4]. Pengguna sering kali harus melakukan pencarian secara manual dengan membaca deskripsi kursus satu per satu sebelum menentukan kursus yang akan diikuti [2]. Proses tersebut tidak hanya membutuhkan waktu yang cukup lama tetapi juga berpotensi menghasilkan pemilihan kursus yang kurang sesuai dengan tujuan pembelajaran pengguna [5].

Permasalahan dalam pemilihan kursus online tersebut menunjukkan pentingnya suatu mekanisme yang mampu membantu pengguna memperoleh rekomendasi kursus secara otomatis berdasarkan karakteristik dan preferensi yang dimiliki [6]. Sistem rekomendasi merupakan salah satu pendekatan yang banyak digunakan untuk membantu pengguna menemukan item yang relevan dari sejumlah besar alternatif yang tersedia [7]. Implementasi sistem rekomendasi telah diterapkan pada berbagai bidang seperti rekomendasi buku [8], *e-commerce* [9], pekerjaan [10], program magang [11], pembelajaran digital [2], dan pelatihan berbasis teknologi [6]. Hasil dari berbagai penelitian tersebut menunjukkan bahwa sistem rekomendasi mampu meningkatkan efisiensi

pencarian informasi serta membantu pengguna dalam proses pengambilan keputusan [12].

Salah satu metode yang banyak digunakan dalam pengembangan sistem rekomendasi adalah *Content-Based Filtering* karena mampu menghasilkan rekomendasi berdasarkan kesamaan karakteristik antara item dan preferensi pengguna [8]. Metode ini memanfaatkan atribut atau deskripsi konten yang dimiliki suatu item untuk menentukan tingkat relevansi terhadap kebutuhan pengguna [12]. Penelitian sebelumnya menunjukkan bahwa metode *Content-Based Filtering* berhasil diterapkan pada sistem rekomendasi buku [8], sistem rekomendasi pembelajaran berbasis Android [2], sistem rekomendasi pekerjaan [10], sistem rekomendasi magang [11] serta sistem rekomendasi pelatihan digital [6]. Oleh karena itu, metode *Content-Based Filtering* dinilai sesuai untuk diterapkan pada sistem rekomendasi kursus online karena dapat memberikan rekomendasi berdasarkan kesesuaian materi kursus dengan minat pengguna [1].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk merancang sistem rekomendasi kursus online menggunakan metode *Content-Based Filtering* guna membantu pengguna memperoleh rekomendasi kursus yang lebih relevan sesuai kebutuhan pembelajaran mereka. Pengembangan sistem dilakukan menggunakan metode *prototype* sehingga kebutuhan pengguna dapat diidentifikasi secara lebih cepat dan rancangan sistem yang dihasilkan dapat lebih sesuai dengan kebutuhan pengguna [13], [14], [15]. Diharapkan penelitian ini dapat memberikan kontribusi dalam pengembangan sistem rekomendasi pada bidang pendidikan digital dan membantu meningkatkan efektivitas proses pencarian kursus online.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode pengembangan sistem *prototype*. Metode *Prototype* dipilih karena mampu menghasilkan rancangan sistem secara cepat melalui interaksi

yang berkelanjutan antara pengembang dan pengguna sehingga kebutuhan sistem dapat diidentifikasi dengan lebih baik [13]. Pendekatan ini memungkinkan proses pengembangan dilakukan secara bertahap dengan fokus pada penyempurnaan rancangan berdasarkan kebutuhan pengguna yang telah dikumpulkan sebelumnya [14]. Dalam penelitian ini, metode *Prototype* terdiri dari tiga tahapan utama yaitu *Communication*, *Quick Plan*, dan *Modelling Quick Design*. Ketiga tahapan tersebut digunakan untuk menghasilkan rancangan sistem rekomendasi kursus online yang sesuai dengan kebutuhan pengguna dan tujuan penelitian.

2.1 Communication

Tahap *Communication* merupakan tahap awal yang dilakukan untuk mengidentifikasi kebutuhan pengguna serta permasalahan yang dihadapi dalam proses pencarian kursus online [13]. Pada tahap ini dilakukan pengumpulan informasi terkait karakteristik pengguna, jenis kursus yang tersedia, serta faktor-faktor yang mempengaruhi pemilihan kursus [1]. Berdasarkan informasi yang telah diperoleh, dilakukan identifikasi terhadap kebutuhan sistem sebagai dasar dalam proses perancangan sistem.

2.2 Quick Plan

Tahap *Quick Plan* dilakukan untuk menyusun perencanaan awal pengembangan sistem berdasarkan kebutuhan yang telah diperoleh pada tahap *Communication* [14], [15]. Pada tahap ini dilakukan identifikasi sumber data, perancangan alur sistem, serta penentuan metode rekomendasi yang akan digunakan dalam penelitian. Metode yang dipilih adalah *Content-Based Filtering* karena mampu memberikan rekomendasi berdasarkan kesamaan karakteristik konten kursus dengan preferensi pengguna [8]. Data yang digunakan dalam penelitian berupa data kursus online yang terdiri dari informasi nama kursus, platform, kategori, harga, rating, durasi, dan tingkat kesulitan. Data tersebut digunakan sebagai dasar dalam proses pembentukan profil item yang akan diproses menggunakan metode *Content-Based Filtering* [1].

2.3 Modelling Quick Design

Tahap *Modelling Quick Design* dilakukan untuk menghasilkan rancangan awal sistem rekomendasi kursus online yang akan dibangun. Pada tahap ini terdapat dua pemodelan, yaitu pemodelan proses, dan pemodelan antarmuka sistem. Pemodelan proses dilakukan dengan merancang proses pengolahan data kursus berdasarkan kata kunci atau minat yang diinputkan pengguna. Proses rekomendasi menggunakan pembobotan TF-IDF dan perhitungan *Cosine Similarity* untuk menentukan tingkat kemiripan antara preferensi pengguna dan data kursus yang tersedia [12], [6]. Selanjutnya, pemodelan antarmuka dilakukan dengan menggambarkan alur kerja sistem menggunakan *use case* diagram untuk menunjukkan interaksi antara pengguna dan sistem, merancang tampilan input pencarian, halaman hasil rekomendasi dan melihat detail kursus guna memudahkan pengguna dalam menggunakan sistem [1].

III. HASIL DAN PEMBAHASAN

3.1 Communication

Tahap *Communication* menghasilkan informasi mengenai kebutuhan pengguna dan data yang akan digunakan dalam pengembangan sistem rekomendasi kursus online. Berdasarkan hasil identifikasi kebutuhan, pengguna membutuhkan sistem yang mampu membantu proses pencarian kursus secara lebih cepat karena banyaknya pilihan kursus yang tersedia pada berbagai platform pembelajaran digital. Selain itu, pengguna menginginkan rekomendasi yang sesuai dengan topik pembelajaran yang diminati tanpa harus melakukan pencarian dan perbandingan kursus secara manual. Hasil pengumpulan data pada tahap ini menghasilkan dataset kursus online yang terdiri dari atribut nama kursus, kategori, *platform*, harga, rating, durasi dan level yang selanjutnya digunakan sebagai sumber data dalam proses rekomendasi. Data yang berhasil dikumpulkan kemudian disusun ke dalam bentuk dataset dan dipersiapkan untuk proses pengolahan menggunakan metode *Content-Based Filtering*. Berdasarkan hasil komunikasi tersebut, diperoleh

kebutuhan sistem berupa kemampuan untuk menerima input kata kunci dari pengguna, melakukan pencocokan terhadap data kursus yang tersedia, serta menghasilkan rekomendasi kursus yang relevan sesuai dengan kebutuhan pembelajaran pengguna.

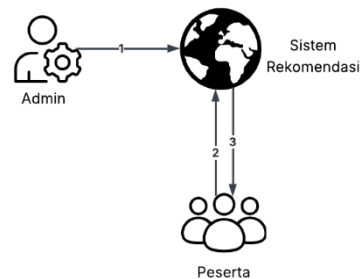
Tabel 1. Hasil Pengumpulan Data

No	Course_Name	Category	Platform	Price	Rating	Duration_Hours	Level
19	Agile Project Management	Business	Coursera	49	4.7	20	Intermediate
20	Artificial Intelligence	Data Science	edX	150	4.9	60	Advanced
1	Python for Beginners	Programming	Coursera	49	4.8	20	Beginner
2	Machine Learning A-Z	Data Science	Udemy	15	4.5	40	All Levels
3	Deep Learning Specialization	Data Science	Coursera	49	4.9	80	Advanced
4	Full Stack Web Development	Programming	Udemy	20	4.3	50	Intermediate
5	Digital Marketing Masterclass	Marketing	Udemy	15	4.6	15	All Levels
6	Data Analysis with Pandas	Data Science	edX	0	4.7	10	Beginner
7	Advanced Excel Formulas	Business	Udemy	15	4.8	12	Advanced
8	Python for Data Science	Data Science	Coursera	49	4.9	25	Intermediate
9	UI/UX Design Fundamentals	Design	Coursera	39	4.5	30	Beginner
10	Java Programming Basics	Programming	edX	0	4.2	15	Beginner
11	Graphic Design Masterclass	Design	Udemy	15	4.7	20	All Levels
12	SQL for Data Analysis	Data Science	Udemy	15	4.6	18	Intermediate
13	Financial Modeling	Business	Coursera	79	4.8	40	Advanced
14	Intro to Cloud Computing	IT	AWS	0	4.4	10	Beginner
15	Social Media Marketing	Marketing	Coursera	49	4.5	12	Beginner
16	React Web Developer	Programming	Udemy	25	4.8	35	Intermediate
17	SEO Optimization	Marketing	Udemy	15	4.3	8	Beginner
18	Tableau for Beginners	Data Science	edX	0	4.6	15	Beginner

No	Course_Name	Category	Platform	Price	Rating	Duration_Hours	Level
19	Agile Project Management	Business	Coursera	49	4.7	20	Intermediate
20	Artificial Intelligence	Data Science	edX	150	4.9	60	Advanced

3.2 Quick Plan

Tahap *Quick Plan* menghasilkan rancangan alur kerja sistem rekomendasi kursus online yang akan digunakan sebagai acuan pada proses pengembangan sistem. Berdasarkan hasil perencanaan yang telah dilakukan, proses rekomendasi diawali dengan pengguna memasukkan kata kunci atau preferensi kursus yang diinginkan ke dalam sistem. Selanjutnya sistem mengambil data kursus yang tersedia pada dataset untuk dilakukan proses pengolahan menggunakan metode *Content-Based Filtering*. Data kursus dan *query* pengguna kemudian diproses melalui tahapan *preprocessing*, pembobotan TF-IDF, dan perhitungan *Cosine Similarity* untuk memperoleh tingkat kemiripan antar data. Nilai kemiripan yang dihasilkan digunakan sebagai dasar dalam melakukan pemeringkatan kursus sehingga diperoleh rekomendasi dengan tingkat relevansi tertinggi. Alur kerja sistem dapat dilihat pada gambar 1.



Gambar 1. Workflow Sistem Rekomendasi

Keterangan *Workflow*:

1. Admin mengelola data yang digunakan oleh sistem rekomendasi.
2. Peserta mencari rekomendasi dengan memasukkan kata kunci atau kriteria pencarian.
3. Peserta menerima output rekomendasi berupa daftar kursus yang paling relevan berdasarkan hasil perhitungan sistem.

3.3 Modelling Quick Design

3.3.1 Pemodelan Proses Perhitungan

Pada tahap ini dilakukan proses pengolahan data kursus yang telah diperoleh pada tahap sebelumnya untuk menghasilkan rekomendasi yang sesuai dengan kebutuhan pengguna. Pengolahan data dilakukan menggunakan beberapa *library* pendukung seperti Pandas untuk mengelola dataset, TfidfVectorizer untuk melakukan pembobotan kata menggunakan metode TF-IDF, serta *Cosine Similarity* untuk menghitung tingkat kemiripan antara *query* pengguna dan data kursus.

Tabel 2. Hasil Preprocessing

Course_Name	Hasil_Preprocessing
Python for Beginners	python, beginners, programming, coursera, beginner, price49, rating4, duration20h
Machine Learning A-Z	machine, learning, data, science, udemy, levels, price15, rating4, duration40h
Deep Learning Specialization	deep, learning, specialization, data, science, coursera, advanced, price49, rating4, duration80h
Full Stack Web Development	stack, web, development, programming, udemy, intermediate, price20, rating4, duration50h
Digital Marketing Masterclass	digital, marketing, masterclass, marketing, udemy, levels, price15, rating4, duration15h
Data Analysis with Pandas	data, analysis, pandas, data, science, edx, beginner, price0, rating4, duration10h
Advanced Excel Formulas	advanced, excel, formulas, business, udemy, advanced, price15, rating4, duration12h
Python for Data Science	python, data, science, data, science, coursera, intermediate, price49, rating4, duration25h
UI/UX Design Fundamentals	ui, ux, design, fundamentals, design, coursera, beginner, price39, rating4, duration30h
Java Programming Basics	java, programming, basics, programming, edx, beginner, price0, rating4, duration15h
Graphic Design Masterclass	graphic, design, masterclass, design, udemy, levels, price15, rating4, duration20h
SQL for Data Analysis	sql, data, analysis, data, science, udemy, intermediate, price15, rating4, duration18h
Financial Modeling	financial, modeling, business, coursera, advanced, price79, rating4, duration40h
Intro to Cloud Computing	intro, cloud, computing, aws, beginner, price0, rating4, duration10h
Social Media Marketing	social, media, marketing, marketing, coursera, beginner, price49, rating4, duration12h
React Web Developer	react, web, developer, programming, udemy, intermediate, price25, rating4, duration35h
SEO Optimization	seo, optimization, marketing, udemy, beginner, price15, rating4, duration8h
Tableau for Beginners	tableau, beginners, data, science, edx, beginner, price0, rating4, duration15h
Agile Project Management	agile, project, management, business, coursera, intermediate, price49, rating4, duration20h

Course_Name	Hasil_Preprocessing
Artificial Intelligence	artificial, intelligence, data, science, edx, advanced, price150, rating4, duration60h

3.3.1.1 Perhitungan TF

Setelah melalui tahapan *preprocessing* yang meliputi *case folding*, *tokenizing*, *stopword removal*, dan *stemming*, diperoleh kumpulan *term* yang digunakan dalam proses perhitungan *Term Frequency* (TF). Pada tahap ini, setiap *term* dihitung berdasarkan jumlah kemunculannya pada masing-masing dokumen kursus. Nilai 0 menunjukkan bahwa *term* tidak ditemukan pada dokumen, sedangkan nilai 1 dan 2 menunjukkan bahwa *term* muncul sebanyak satu kali dan dua kali. Dalam penelitian ini, *query* yang digunakan adalah "Python for Data Science", sehingga frekuensi *term* pada *query* tersebut dihitung terhadap seluruh dokumen kursus.

Tabel 3. Hasil Perhitungan TF

Course_Name	advanced	agile	analyses	...	web
Python for Beginners	0	0	0	...	0
Machine Learning A-Z	0	0	0	...	0
Deep Learning Specialization	1	0	0	...	0
Full Stack Web Development	0	0	0	...	1
Digital Marketing Masterclass	0	0	0	...	0
Data Analysis with Pandas	0	0	1	...	0
Advanced Excel Formulas	2	0	0	...	0
Python for Data Science	0	0	0	...	0
UI/UX Design Fundamentals	0	0	0	...	0
Java Programming Basics	0	0	0	...	0
Graphic Design Masterclass	0	0	0	...	0
SQL for Data Analysis	0	0	1	...	0
Financial Modeling	1	0	0	...	0
Intro to Cloud Computing	0	0	0	...	0
Social Media Marketing	0	0	0	...	0
React Web Developer	0	0	0	...	1
SEO Optimization	0	0	0	...	0
Tableau for Beginners	0	0	0	...	0

Course_Name	advance d	agile	analysi s	...	web
Agile Project Management	0	1	0	...	0
Artificial Intelligence	1	0	0	...	0

3.3.1.2 Perhitungan IDF

Berdasarkan hasil pengolahan data kursus menggunakan metode *Content-Based Filtering*, diperoleh sebanyak 70 *term* unik yang digunakan dalam proses pembobotan *Inverse Document Frequency* (IDF). Hasil perhitungan menunjukkan bahwa beberapa *term* seperti *agile*, *artificial*, *aws*, *cloud*, *computing*, dan *term* lain yang hanya muncul pada satu dokumen memiliki nilai IDF tertinggi sebesar 3,397895. Sebaliknya, *term* seperti *rating4*, *beginner*, *coursera*, *data*, dan *science* memiliki nilai IDF yang lebih rendah karena muncul pada lebih banyak dokumen dalam dataset. Perbedaan nilai IDF menunjukkan bahwa semakin jarang suatu *term* muncul pada dokumen, maka semakin besar bobot yang diberikan sehingga *term* tersebut memiliki kemampuan yang lebih baik dalam membedakan karakteristik setiap kursus. Rumus perhitungan IDF:

$$IDF = \log \frac{D}{DF} \quad (1)$$

Tabel 4. Hasil Perhitungan IDF

Kata	IDF
advanced	2,4816
agile	3,3979
analysis	2,9924
artificial	3,3979
aws	3,3979
...	...
ux	3,3979
web	2,9924

3.3.1.3 Perhitungan TF-IDF Dokumen

Setelah diperoleh nilai *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF), langkah selanjutnya adalah menghitung bobot TF-IDF untuk setiap *term* pada masing-masing dokumen kursus. Hasil pembobotan menunjukkan bahwa setiap *term* memiliki nilai yang berbeda sesuai dengan frekuensi kemunculannya dan tingkat

kepentingannya dalam membedakan dokumen. Nilai TF-IDF yang semakin besar menunjukkan bahwa *term* tersebut memiliki kontribusi yang lebih tinggi terhadap karakteristik suatu dokumen, sedangkan nilai 0 menunjukkan bahwa *term* tidak muncul pada dokumen tersebut. Rumus menghitung W:

$$W = TF \times IDF \quad (2)$$

Tabel 5. Hasil Perhitungan TF-IDF Dokumen

Course_Name	advance d	agile	analysi s	...	web
Python for Beginners	0	0	0	...	0
Machine Learning A-Z	0	0	0	...	0
Deep Learning Specialization	0,3034	0	0	...	0
Full Stack Web Development	0	0	0	...	0,3545
Digital Marketing Masterclass	0	0	0	...	0
Data Analysis with Pandas	0	0	0,3741	...	0
Advanced Excel Formulas	0,5798	0	0	...	0
Python for Data Science	0	0	0	...	0
UI/UX Design Fundamentals	0	0	0	...	0
Java Programming Basics	0	0	0	...	0
Graphic Design Masterclass	0	0	0	...	0
SQL for Data Analysis	0	0	0,3725	...	0
Financial Modeling	0,3158	0	0	...	0
Intro to Cloud Computing	0	0	0	...	0
Social Media Marketing	0	0	0	...	0
React Web Developer	0	0	0	...	0,3545
SEO Optimization	0	0	0	...	0
Tableau for Beginners	0	0	0	...	0
Agile Project Management	0	0,4230	0	...	0
Artificial Intelligence	0,3037	0	0	...	0

3.3.1.4 Perhitungan TF-IDF Query

Setelah dilakukan pembobotan TF-IDF pada *query*, diperoleh bobot untuk setiap *term* yang digunakan sebagai representasi vektor *query*. Nilai bobot yang lebih tinggi menunjukkan bahwa *term* tersebut memiliki tingkat

kepentingan yang lebih besar dalam *query*. Berdasarkan hasil perhitungan, *term* python memiliki bobot TF-IDF tertinggi sebesar 0,5864, diikuti oleh *term* data, dan *science* yang masing-masing memiliki bobot 0,4106. Bobot TF-IDF pada *query* ini selanjutnya digunakan dalam proses perhitungan *Cosine Similarity* untuk menentukan tingkat kemiripan antara *query* dan setiap dokumen kursus.

Tabel 6. Hasil TF-IDF Query

Kata	TF-IDF
python	0,5864
data	0,4106
science	0,4106

3.3.1.5 Perhitungan Cosine Similarity

Hasil perhitungan *Cosine Similarity* menunjukkan tingkat kemiripan antara *query* dengan setiap dokumen kursus berdasarkan bobot TF-IDF. Semakin besar nilai similarity, maka semakin tinggi tingkat kesesuaian dokumen terhadap *query* yang diberikan. Berdasarkan hasil perhitungan, kursus Python for Data Science memperoleh nilai similarity tertinggi sebesar 0,6907, diikuti oleh Python for Beginners sebesar 0,4667 dan Data Analysis with Pandas sebesar 0,3783. Sementara itu, beberapa kursus seperti Full Stack Web Development, Graphic Design Masterclass, Advanced Excel Formulas, Digital Marketing Masterclass, dan React Web Developer memiliki nilai similarity sebesar 0, yang menunjukkan bahwa dokumen tersebut tidak memiliki kemiripan dengan *query*. Hasil perhitungan ini kemudian digunakan untuk menentukan urutan rekomendasi kursus, di mana kursus dengan nilai similarity tertinggi ditempatkan pada peringkat teratas. Hasil perhitungan *Cosine Similarity* secara lengkap dapat dilihat pada tabel 7. Rumus perhitungan *cosine similarity*:

$$\text{Cos}(\theta) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (3)$$

Tabel 7. Hasil Perhitungan Cosine Similarity

Rank	Course_Name	Similarity
1	Python for Data Science	0,6907
2	Python for Beginners	0,4667
3	Data Analysis with Pandas	0,3783
4	Tableau for Beginners	0,3130
5	SQL for Data Analysis	0,2904
6	Deep Learning Specialization	0,2852
7	Machine Learning A-Z	0,2132
8	Artificial Intelligence	0,1903
9	Social Media Marketing	0,1705
10	UI/UX Design Fundamentals	0,1461
11	Financial Modeling	0,0990
12	Agile Project Management	0,0968
13	SEO Optimization	0,0940
14	Intro to Cloud Computing	0,0856
15	Java Programming Basics	0,0819
16	Full Stack Web Development	0
17	Graphic Design Masterclass	0
18	Advanced Excel Formulas	0
19	Digital Marketing Masterclass	0
20	React Web Developer	0

3.3.1.6 Hasil Rekomendasi Kursus

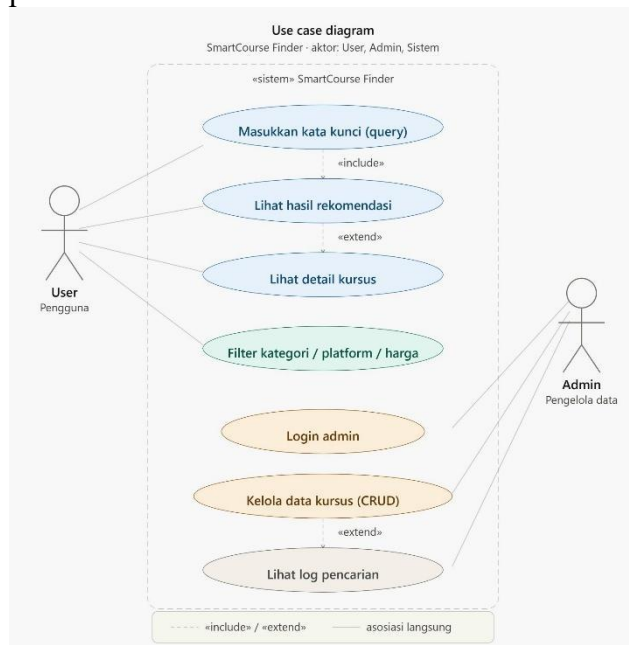
Berdasarkan proses pemeringkatan nilai *Cosine Similarity*, sistem berhasil menghasilkan tiga rekomendasi kursus terbaik yang paling sesuai dengan *query* pengguna. Kursus Python for Data Science menempati peringkat pertama dengan nilai similarity 0,6907, diikuti oleh Python for Beginners dengan nilai 0,4667, dan Data Analysis with Pandas dengan nilai 0,3783. Ketiga kursus tersebut berasal dari kategori yang relevan dengan kata kunci Python dan Data Science yang digunakan pada *query* pencarian. Hasil ini menunjukkan bahwa metode *Content-Based Filtering* mampu mengidentifikasi kursus yang memiliki tingkat kemiripan konten paling tinggi berdasarkan hasil pembobotan TF-IDF dan perhitungan *Cosine Similarity*. Hasil rekomendasi yang dihasilkan sistem dapat dilihat pada Tabel 8.

Tabel 8. Top 3 Hasil Rekomendasi

Rank	Course_Name	Similarity
1	Python for Data Science	0,6907
2	Python for Beginners	0,4667
3	Data Analysis with Pandas	0,3783

3.3.2 Use Case

Use Case Diagram digunakan untuk menggambarkan interaksi antara aktor dengan sistem rekomendasi kursus yang dikembangkan. Sistem memiliki dua aktor utama, yaitu User dan Admin. User dapat memasukkan kata kunci (*query*), melihat hasil rekomendasi, melihat detail kursus, serta melakukan penyaringan hasil berdasarkan kategori, platform, atau harga. Sementara itu, Admin memiliki hak akses untuk melakukan login, mengelola data kursus (*Create, Read, Update, dan Delete*), serta melihat log pencarian pengguna. Diagram ini menunjukkan bahwa setiap fungsi yang tersedia telah disesuaikan dengan hak akses masing-masing aktor sehingga proses pencarian dan pengelolaan data kursus dapat berjalan secara terstruktur. *Use Case Diagram* sistem rekomendasi dapat dilihat pada Gambar 2.



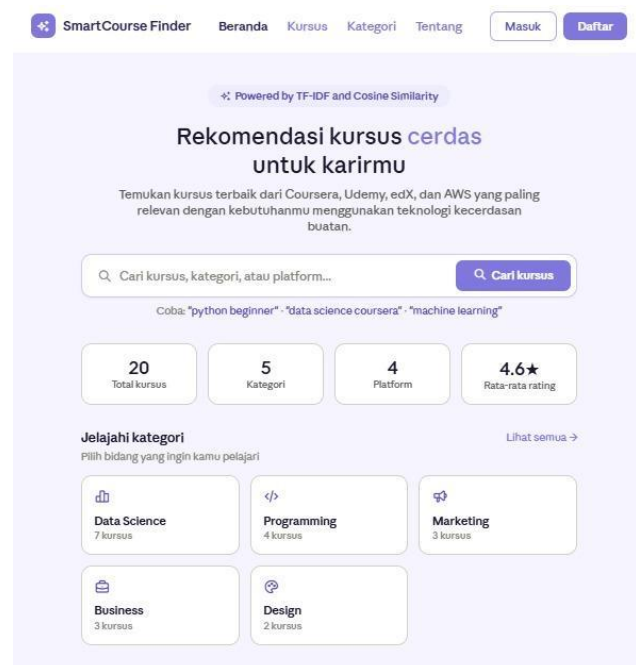
Gambar 2. Use Case Sistem Rekomendasi

3.3.3 Pemodelan Interface

3.3.3.1 Pemodelan Interface Halaman Beranda

Halaman beranda merupakan tampilan awal yang diakses oleh pengguna ketika membuka

sistem. Pada halaman ini tersedia kolom pencarian yang digunakan untuk memasukkan kata kunci sesuai kebutuhan atau minat pengguna terhadap kursus tertentu. Selain itu, halaman ini juga menampilkan informasi ringkas mengenai jumlah kursus, kategori, platform, atribut data, serta rata-rata rating yang tersedia dalam sistem. Pengguna juga dapat melihat daftar kategori kursus sebagai referensi sebelum melakukan pencarian. Halaman ini dirancang sebagai titik awal interaksi pengguna dalam memperoleh rekomendasi kursus menggunakan metode *Content-Based Filtering* berbasis TF-IDF dan *Cosine Similarity*.

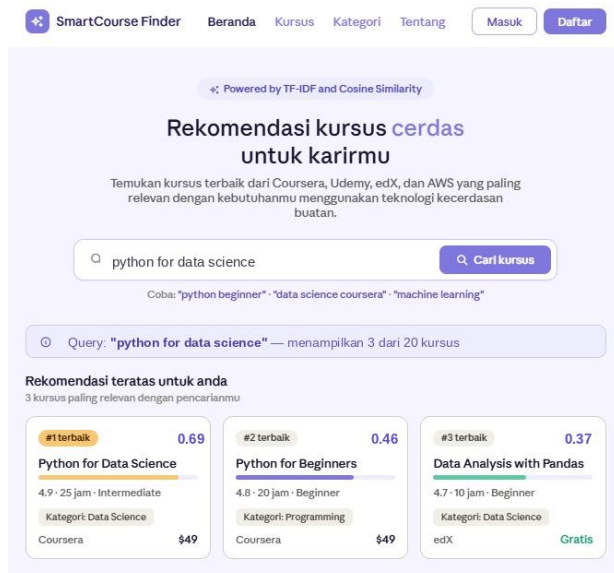


Gambar 3. Halaman Beranda User

3.3.3.2 Pemodelan Interface Hasil Rekomendasi

Halaman hasil rekomendasi ditampilkan setelah pengguna memasukkan kata kunci pencarian dan menekan tombol Cari. Sistem akan menghitung tingkat kemiripan antara *query* dengan seluruh dokumen kursus menggunakan metode TF-IDF dan *Cosine Similarity*, kemudian menampilkan tiga kursus dengan nilai similarity tertinggi. Setiap hasil rekomendasi dilengkapi dengan informasi nama kursus, kategori, platform, level, harga, rating, durasi, serta nilai similarity sebagai indikator tingkat kesesuaian terhadap *query* pengguna. Selain itu, pengguna

dapat menerapkan filter berdasarkan kategori, platform, harga, dan level untuk memperoleh rekomendasi yang lebih sesuai dengan kebutuhannya.



Gambar 4. Halaman Hasil Rekomendasi

IV. KESIMPULAN

Penelitian ini berhasil merancang sistem rekomendasi kursus online di bidang teknologi menggunakan metode *Content-Based Filtering* dengan pendekatan pengembangan *Prototype*. Sistem memanfaatkan pembobotan TF-IDF dan perhitungan *Cosine Similarity* untuk mengukur tingkat kemiripan antara preferensi pengguna dan konten kursus yang tersedia. Hasil pengujian dengan query "Python for Data Science" membuktikan bahwa sistem mampu menghasilkan rekomendasi yang tepat sasaran, dengan kursus Python for Data Science memperoleh nilai similarity tertinggi sebesar 0,6907, diikuti Python for Beginners sebesar 0,4667 dan Data Analysis with Pandas sebesar 0,3783. Rancangan antarmuka yang dihasilkan mencakup halaman beranda, halaman hasil rekomendasi, serta halaman admin yang terstruktur, sehingga tujuan penelitian untuk membantu pengguna memperoleh rekomendasi kursus yang relevan secara otomatis telah tercapai.

Penerapan metode *Content-Based Filtering* pada sistem ini memberikan kontribusi nyata

dalam bidang pendidikan digital, khususnya dalam membantu pengguna menemukan kursus yang sesuai secara efisien tanpa perlu melakukan pencarian manual. Sistem yang dirancang juga dilengkapi dengan fitur filter berdasarkan kategori, platform, harga, dan level sehingga semakin mempermudah pengguna dalam menyesuaikan hasil rekomendasi dengan kebutuhannya. Pengembangan ke depan dapat diarahkan pada integrasi metode *Collaborative Filtering* untuk menghasilkan rekomendasi yang lebih personal, perluasan dataset kursus, serta penambahan metrik evaluasi seperti *precision* dan *recall* guna mengukur kinerja sistem secara lebih komprehensif.

REFERENSI

- [1] A. D. Pangestu, G. P. Dini, F. A. Sariasih, S. N. Rakhmah, and I. Sutoyo, "PENGEMBANGAN APLIKASI WEB REKOMENDASI KURSUS ONLINE BERBASIS MINAT DENGAN PENCOCOKAN SEMANTIK," *Jurnal Sistem Informasi dan Teknologi Informasi*, 2026.
- [2] J. Teknologi Informasi, A. Arya Anggara, A. Ridho, J. Alue Peunyareng, U. Tanog Darat, and K. Aceh Barar, "Sistem Rekomendasi Pembelajaran dengan Menggunakan Metode Content-Based Filtering Berbasis Aplikasi Android," vol. 3, no. 1, pp. 11–19, 2024, [Online]. Available: <http://jurnal.utu.ac.id/JTI>
- [3] H. A. Firdaos, A. Sujjada, and S. Somantri, "Melangkah ke Masa Depan Literasi Digital: Rancang Bangun Sistem Genusian Course Academy dengan Pendekatan Hybrid Collaborative Filtering dan Content-Based Filtering," *Briliant: Jurnal Riset dan Konseptual*, vol. 10, no. 2, pp. 432–441, May 2025, doi: 10.28926/briliant.v10i2.1957.
- [4] D. B. Elnursa, V. Nofriana, A. Syamsuri, L. Cahyani, D. Budi Elnursa, "Sistem Rekomendasi Pemilihan Program MSIB Bagi Mahasiswa Pendidikan Informatika Kata Kunci: Recommendation System MSIB Program Content Based Filtering Cosine Similarity," *JOURNAL SHIFT VOL*, vol. 3, 2023, [Online]. Available: <https://kampusmerdeka.kemdikbud.go.id/>
- [5] S. Wahyu and A. S. Purnomo, "PENERAPAN CONTENT BASED FILTERING DALAM SISTEM REKOMENDASI PEMILIHAN JURUSAN SEKOLAH MENENGAH KEJURUAN BERDASARKAN FAKTOR MINAT BAKAT DAN NILAI AKADEMIK SISWA (STUDI KASUS: SLEMAN, YOGYAKARTA)," 2025.
- [6] A. L. Mutmainnah, A. Yusuf, and B. A. Nugroho, "Sistem Rekomendasi Pelatihan Digital Talent Scholarship Menggunakan Content Based Filtering dan Cosine Similarity," 2026.
- [7] M. P. Putri, "IMPLEMENTASI CONTENT-BASED FILTERING DAN COLLABORATIVE FILTERING PADA SISTEM INFORMASI," 2025.
- [8] P. R. Effendi, "PENGEMBANGAN SISTEM REKOMENDASI BUKU MENGGUNAKAN CONTENT-BASED FILTERING PADA INTEGRATED LIBRARY SYSTEM LITE (INLISLITE)," 2026.
- [9] R. Y. Ardhitian, "Pembuatan E-Commerce Berbasis Website Dengan Sistem Rekomendasi Content-Based Filtering," 2024.
- [10] A. Ro'uf, Y. M. Pranoto, and D. E. Setyati, "Sistem Rekomendasi Pekerjaan Menggunakan Content Based Similarity," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, 2023, [Online]. Available: <https://www.bps.go.id/>
- [11] R. U. Ichlas Amal Sukma Putra and J. Sasongko Wibowo, "Sistem Rekomendasi Magang Berbayar Menggunakan Metode Content-

- based Filtering,” *INNOVATIVE: Journal Of Social Science Research*, 2023.
- [12] F. Akhdan and L. Fitriani, “Aplikasi Mentor Pembelajaran Berbasis Sistem Rekomendasi Content-Based Filtering dengan Metode TF-IDF dan Cosine Similarity,” *Jurnal Algoritma*, vol. 22, no. 2, Nov. 2025, doi: 10.33364/algoritma/v.22-2.2493.
- [13] A. Pranata and F. Sulianta, “Implementasi dan Analisis Algoritma Content-Based Filtering Pada Sistem Rekomendasi Produk Tas pada Basis Data MySQL,” *Jurnal Ilmiah Global Education*, vol. 6, no. 3, pp. 1419–1444, Aug. 2025, doi: 10.55681/jige.v6i3.4017.
- [14] A. Sulami and V. Atina, “STRING (Satuan Tulisan Riset dan Inovasi Teknologi) PENERAPAN METODE CONTENT BASED FILTERING DALAM SISTEM REKOMENDASI PEMILIHAN PRODUK SKINCARE,” 2024.
- [15] K. S. Yanisa Putri, I Made Agus Dwi Suarjaya, and Wayan Oger Vihikan, “Sistem Rekomendasi Skincare Menggunakan Metode Content Based Filtering dan Collaborative Filtering,” *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 3, pp. 764–774, Aug. 2024, doi: 10.51454/decode.v4i3.601.