

**SISTEM REKOMENDASI KOMUNITAS AKADEMIK
MENGUNAKAN METODE *HYBRID FILTERING***

THESIS

**Oleh:
ABDURROZZAAQ ASHSHIDDIQI ZUHRI
NIM. 230605210021**



**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**SISTEM REKOMENDASI KOMUNITAS AKADEMIK MENGGUNAKAN
METODE *HYBRID FILTERING***

THESIS

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)

Oleh:
ABDURROZZAAQ ASHSHIDDIQI ZUHRI
NIM. 230605210021

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

SISTEM REKOMENDASI KOMUNITAS AKADEMIK MENGGUNAKAN
METODE *HYBRID FILTERING*

THESIS

Oleh:
ABDURROZZAAQ ASHSHIDDIQI ZUHRI
NIM. 230605210021

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 3 Juni 2025

Pembimbing I,



Dr. Ririen Kusumawati, S.Si, M.Kom
NIP. 19720309 200501 2 002

Pembimbing II,



Dr. Muhammad Ainul Yaqin, M.Kom
NIP. 19761013 200604 1 004

Mengetahui,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Carys Crysdian
19740424 200901 1 008

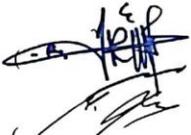
**SISTEM REKOMENDASI KOMUNITAS AKADEMIK MENGGUNAKAN
METODE *HYBRID FILTERING***

THESIS

Oleh:
ABDURROZZAAQ ASHSHIDDIQI ZUHRI
NIM. 230605210021

Telah Dipertahankan di Depan Dewan Penguji Thesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 3 Juni 2025

Susunan Dewan Penguji

Penguji I	: <u>Prof. Dr. Sri Harini, M.Si</u> NIP. 19731014 200112 2 002	()
Penguji II	: <u>Dr. Usman Pagalay, M.Si</u> NIP. 19650414 200312 1 001	()
Pembimbing I	: <u>Dr. Ririen Kusumawati, S.Si. M.Kom</u> NIP. 19720309 200501 2 002	()
Pembimbing II	: <u>Dr. Muhammad Ainul Yaqin, M.Kom</u> NIP. 19761013 200604 1 004	()

Mengetahui dan Mengesahkan,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang


Dr. Cahya Crysdiyan
NIP. 19740424 200901 1 008

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : ABDURROZZAAQ ASHSHIDDIQI ZUHRI
NIM : 230605210021
Fakultas / Prodi : Sains dan Teknologi / Magister Informatika
Judul Tesis : SISTEM REKOMENDASI KOMUNITAS AKADEMIK
MENGGUNAKAN *HYBRID FILTERING*

Menyatakan dengan sebenarnya bahwa Thesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan thesis ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang,
Yang membuat pernyataan,



Abdurrozzaaq Ashshiddiqi Zuhri
NIM. 230605210021

MOTTO

“Lakukan apapun yang ingin anda lakukan, jika gagal setidaknya kalian tidak menyesal karena tidak melakukannya”.

HALAMAN PERSEMBAHAN

Thesis ini saya persembahkan khusus kepada kedua orang tua, keluarga, dosen, sahabat, dan semua pihak yang telah membantu secara aktif memberikan support dalam menyelesaikan penelitian ini. Semoga Allah SWT selalu memberikan keberkahan dan hal-hal baik kepada mereka.

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh

Dengan penuh rasa syukur, penulis panjatkan puji syukur kehadiran Allah SWT atas limpahan rahmat dan karunia-Nya, serta shalawat dan salam semoga senantiasa tercurah kepada junjungan kita Nabi Muhammad SAW. Berkat berkah-Nya, penulis dapat menyelesaikan penulisan thesis ini dengan judul “Sistem Rekomendasi Komunitas Akademik Menggunakan *Hybrid Filtering*”. Thesis ini disusun sebagai salah satu syarat untuk mencapai gelar magister di Program Studi Magister Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Ucapan rasa syukur dan terima kasih penulis sampaikan kepada seluruh pihak yang telah membantu berupa kritik dan saran agar terselesaikannya thesis ini. Dengan rasa hormat penulis mengucapkan terima kasih kepada :

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Hariani, M.Si., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Cahyo Crysdiyan, selaku Ketua Program Studi Magister Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Ibu Dr. Ririen Kusumawati, M.Kom selaku dosen pembimbing I yang telah dengan sabar memberikan arahan, saran, kritik, serta motivasi yang baik dalam penulisan hingga program yang dibuat dalam menyelesaikan thesis ini.

5. Bapak Dr. Muhammad Ainul Yaqin, M.Kom selaku dosen pembimbing II yang telah membimbing serta memberikan bantuan dan arahan, sehingga penulis dapat menyelesaikan thesis ini.
6. Ibu Prof. Dr. Hj. Sri Harini, M.Si selaku dosen penguji I dan Bapak Dr. Usman Pagalay, M.Si selaku dosen penguji II yang telah memberikan kritik dan saran, sehingga dapat menyelesaikan thesis ini.
7. Selaku dosen dan jajaran staff Program Studi Magister Informatika yang telah memberikan dukungan selama pengerjaan thesis ini.
8. Bunda, Ayah, serta saudara saya yang selalu memberikan dukungan dan motivasi untuk terus berusaha, dan doa yang tak putus-putusnya selalu disampaikan agar dapat menuntaskan thesis ini dengan lancar dan baik.
9. Saudara Fast Track Magister Informatika 2020 yang selalu memberikan semangat untuk lulus sama-sama.
10. Seluruh pihak yang telah banyak membantu dalam penyusunan thesis ini.

Penulis menyadari bahwa thesis ini masih banyak kekurangan, dari keilmuan maupun penulisan. Maka dari itu, kritik dan saran yang membangun sangat diharapkan agar lebih baik lagi kedepannya. Semoga dengan penyusunan thesis ini bisa memberikan manfaat bagi banyak pihak.

Malang, 11 Juni 2025

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
MOTTO v	
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
ABSTRAK	xi
ABSTRACT	xii
البحث مستخلص.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Pernyataan Masalah.....	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	7
BAB II STUDI PUSTAKA	8
2.1 Sistem Rekomendasi	8
2.2 Komunitas Akademik.....	12
2.3 Content-based Filtering	14
2.4 Collaborative Filtering.....	16
2.5 Hybrid Filtering	18
2.6 Kerangka Teori.....	20
BAB III METODOLOGI PENELITIAN	22
3.1 Kerangka Konsep	23
3.2 Desain Penelitian	24
3.3 Desain Sistem	25
3.4 Pengumpulan Data.....	26
3.4.1 Pengolahan Data	27
3.4.1.1 Content Based Filtering	27
3.4.1.1.1 Preprocessing	28
3.4.1.1.2 Meta Data	30
3.4.1.1.3 TF-IDF	31
3.4.1.2 Collaborative Filtering.....	32
3.4.1. Cosine Similarity.....	32
3.4.2. Weighted Sum.....	33
3.4.3. Hybrid Filtering.....	33
3.5 Pengujian System	34
3.5.1 Mean Average Precission	34
BAB IV CONTENT-BASED FILTERING	36
4.1 Desain Metode Content-Based Filtering	36

4.1.1	Pengumpulan Data	37
4.1.2	Preprocessing Data	38
4.1.3	Perhitungan TF-IDF	39
4.1.4	Perhitungan Cosine Similarity	39
4.2	Implementasi Sistem	40
4.2.1	Preprocessing	41
4.2.2	Metadata	41
4.2.3	TF-IDF	41
4.2.4	Cosine Similarity.....	42
4.3	Uji Coba.....	44
4.3.1	Skenario 1	44
4.3.2	Skenario 2	45
4.3.3	Skenario 3	46
4.4	Pengujian Sistem MAP.....	47
4.4.1	Skenario 1	48
4.4.2	Skenario 2	49
4.4.3	Skenario 3	50
BAB V COLLABORATIVE FILTERING.....		52
5.1	Desain Metode Collaborative Filtering	52
5.1.1	Pengumpulan Data	53
5.1.2	Perhitungan Item Similarity	54
5.1.3	Perhitungan Weighted SUM	56
5.2	Implementasi Sistem	56
5.2.1	Perhitungan Item Similarity	57
5.2.2	Perhitungan Weighted SUM	58
5.3	Pengujian System MAP.....	59
BAB VI HYBRID FILTERING.....		61
6.1	Desain Metode Hybrid Filtering.....	61
6.2	Implementasi Hybrid Filtering	62
6.2.1	Normalisasi Hasil Collaborative Filtering	63
6.2.2	Perhitungan Skor Hybrid	65
6.3	Pengujian MAP Metode Hybrid Filtering	67
BAB VII PEMBAHASAN		69
BAB VIII KESIMPULAN DAN SARAN		78
8.1.	Kesimpulan.....	78
8.2.	Saran	78
DAFTAR PUSTAKA		80

DAFTAR GAMBAR

Gambar 3. 1 Kerangka Konsep	23
Gambar 3. 2 Desain Penelitian	24
Gambar 3. 3 Desain Sistem	25
Gambar 3. 4 Alur Pegolahan Data Mahasiswa.....	28
Gambar 3. 5 Alur Pegolahan Data Komunitas	28
Gambar 3. 6 Alur Pengolahan Data Collaborative Filtering	32
Gambar 4 1. Flowchart Metode Content-based Filtering	36
Gambar 4 2. Visualisasi Sampel data Mahasiswa sebelum dan setelah preprocessing.....	38
Gambar 4 3. Diagram Hasil Cosine Similarity Teratas	43
Gambar 4 4. Diragram Hasil MAP Skenario 1.....	44
Gambar 4 5. Bar Chart hasil cosine similarity dari skenario 2.....	46
Gambar 4 6. Bar Chart hasil cosine similarity dari skenario 2.....	47
Gambar 4 7. Diargram Hasil Average Precision	50
Gambar 4 8. Diagram Hasil Average Precision	51
Gambar 5. 1. Flowchart Metode Collaborative Filtering	52
Gambar 5. 2. Gambar Hitmap dari Similarity Item.....	55
Gambar 5. 3. Hitmap Similarity antar Item	57
Gambar 5. 1. Flowchart Metode Collaborative Filtering	52
Gambar 5. 2. Gambar Hitmap dari Similarity Item.....	55
Gambar 5. 3. Hitmap Similarity antar Item	57
Gambar 6. 1. Flowchart Metode Hybrid Filtering	61
Gambar 6. 2. Normalisasi Data Hasil CF	63
Gambar 6. 3. Grafik Skor Hybrid	67
Gambar 6. 4. Gambar Perbandingan 3 Metode (CBF-CF-HB).....	76
Gambar 7. 1. Hasil Cosine Similarity CBF	70
Gambar 7. 2. Hasil CBF dan CF.....	72
Gambar 7. 3 Gambar Perbandingan 3 Metode (CBF-CF-HB).....	76

DAFTAR TABEL

Tabel 2. 1. Penelitian Terkait Sistem Rekomendasi.....	10
Tabel 3. 1. Pertanyaan / Kriteria Survey	26
Tabel 3. 2 Case Folding pada Data.....	28
Tabel 3. 3 Tokenizing pada Data.....	29
Tabel 3. 4 Stopword pada Data	29
Tabel 3. 5 Stemming pada Data	30
Tabel 3. 6 Meta Data	30
Tabel 3. 7 Sampel Data	31
Tabel 4. 1 Perhitungan TF-IDF	39
Tabel 4. 2. Perhitungan Cosine Similarity	40
Tabel 4. 3. Contoh Input Minat Komuntas Akademik	40
Tabel 4. 4. Data Sebelum dan sesudah Preprocessing.....	41
Tabel 4. 5. Meta Data	41
Tabel 4. 6. Hasil TF-IDF Metdata	42
Tabel 4. 7. Hasil Cosine Similarity	43
Tabel 4. 8. Metadata Skenario 2.....	46
Tabel 4. 9. Metadata Skenario 3.....	47
Tabel 4. 10. Hasil Rekomendasi CBF dengan Data Aktual	48
Tabel 4. 11. Perhitugan MAP Content-based Filtering	48
Tabel 5. 1. Sampel 5 Data mahssiswa dengan rating komunitasnya.....	53
Tabel 5. 2. Tabel Hasil Prediksi Rating.....	56
Tabel 5. 3. Sampel Data Rating Mahasiswa.....	57
Tabel 5. 4. Hasil Prediksi Rating.....	58
Tabel 5. 5. Hasil Rekomendasi CF dengan Data Aktual.....	59
Tabel 5. 6. Perhitugan MAP Collaborative Filtering	59
Tabel 6. 1. Hasil Hybrid Filtering	66
Tabel 6. 2. Hasil Rekomendasi CF dengan Data Aktual.....	67
Tabel 6. 3. Perhitugan MAP Hybrid Filtering.....	68
Tabel 7. 1. Tabel Prediksi Rating CF	71

ABSTRAK

Zuhri, Abdurrozzaaq Ashshiddiqi. 2025. **Sistem Rekomendasi Komunitas Akademik Menggunakan Metode *Hybrid Filtering***. Thesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Ririen Kusumawati, S.Si, M.Kom (II) Dr. Muhammad Ainul Yaqin, M.Kom.

Kata kunci: *Sistem Rekomendasi, Hybrid Filtering, Content-Based Filtering, Collaborative Filtering, Komunitas Akademik, Mean Average Precision.*

Ketidaksesuaian antara jumlah dan kualitas talenta TIK dengan kebutuhan industri di Indonesia menimbulkan tantangan besar, khususnya bagi mahasiswa Teknik Informatika yang sering kesulitan menentukan bidang profesi yang tepat. Oleh karena itu, penelitian ini bertujuan untuk merancang dan membangun sebuah sistem rekomendasi menggunakan metode hybrid filtering untuk memberikan rekomendasi komunitas akademik yang lebih personal dan akurat kepada mahasiswa. Penelitian ini menerapkan metode hybrid filtering yang menggabungkan Content-Based Filtering (CBF) dan Collaborative Filtering (CF). Pendekatan CBF menganalisis profil deskriptif mahasiswa menggunakan TF-IDF dan Cosine Similarity, sementara CF memanfaatkan data rating untuk menemukan pola preferensi kolektif. Hasil dari kedua metode, yang diuji pada data dari 200 mahasiswa dan 10 komunitas, kemudian diintegrasikan menggunakan pendekatan simple average. Hasil evaluasi menggunakan metrik Mean Average Precision (MAP) menunjukkan bahwa metode hybrid filtering mencapai performa paling unggul dengan skor 0,866, secara signifikan melampaui metode tunggal Collaborative Filtering (0,816) dan Content-Based Filtering (0,733). Temuan ini menyimpulkan bahwa penggabungan kedua metode secara efektif menyeimbangkan relevansi konten personal dengan preferensi kolektif, sehingga terbukti menjadi solusi yang valid dan lebih adaptif untuk meningkatkan akurasi sistem rekomendasi di lingkungan akademik.

ABSTRACT

Zuhri, Abdurrozzaaq Ashshiddiqi. 2025. **Academic Community Recommendation System Using Hybrid Filtering Method**. Thesis. Thesis. Master of Informatics Study Program, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University, Malang. Supervisor: (I) Dr. Ririen Kusumawati, M.Kom (II) Dr. Muhammad Ainul Yaqin, M.Kom.

Keywords: *Recommendation System, Hybrid Filtering, Content-Based Filtering, Collaborative Filtering, Academic Community, Mean Average Precision.*

The mismatch between the number and quality of ICT talent and the needs of the industry in Indonesia poses a big challenge, especially for Informatics Engineering students who often have difficulty determining the right professional field. Therefore, this study aims to design and build a recommendation system using the hybrid filtering method to provide more personalized and accurate recommendations to the academic community to students. This study applies a hybrid filtering method that combines Content-Based Filtering (CBF) and Collaborative Filtering (CF). The CBF approach analyzes the descriptive profiles of students using TF-IDF and Cosine Similarity, while CF leverages rating data to find patterns of collective preferences. The results of the two methods, which were tested on data from 200 students and 10 communities, were then integrated using a simple average approach. The results of the evaluation using the Mean Average Precision (MAP) metric showed that the hybrid filtering method achieved the highest performance with a score of 0.866, significantly surpassing the single methods of Collaborative Filtering (0.816) and Content-Based Filtering (0.733). These findings conclude that the combination of the two methods effectively balances the relevance of personal content with collective preferences, thus proving to be a valid and more adaptive solution to improve the accuracy of recommendation systems in the academic environment.

البحث مستخلص

زهري، عبد الرزاق أششديقي. 2025. نظام توصيات المجتمع الأكاديمي باستخدام طريقة/التصفية الهجينة. أطروحة. أطروحة. برنامج دراسة ماجستير المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية، ملانج. المشرف: (أنا) د. ريرين كوسوماواتي، S.Si، م. كوم (الثاني) د. محمد عين اليقين، م. كوم.

الكلمات المفتاحية: نظام التوصية، التصفية الهجينة، التصفية القائمة على المحتوى، التصفية التعاونية، المجتمع الأكاديمي، متوسط الدقة.

يشكل عدم التوافق بين عدد وجودة مواهب تكنولوجيا المعلومات والاتصالات واحتياجات الصناعة في إندونيسيا تحدياً كبيراً، خاصة بالنسبة لطلاب هندسة المعلوماتية الذين غالباً ما يجدون صعوبة في تحديد المجال المهني المناسب. لذلك، تهدف هذه الدراسة إلى تصميم وبناء نظام توصية باستخدام طريقة التصفية الهجينة لتقديم توصيات أكثر تخصيصاً ودقة للمجتمع الأكاديمي للطلاب. تطبق هذه الدراسة طريقة تصفية هجينة تجمع بين التصفية المستندة إلى المحتوى (CBF) والتصفية التعاونية (CF). يحلل نهج CBF الملامح الوصفية للطلاب الذين يستخدمون TF-IDF و Cosine Similarity، بينما يستفيد CF من بيانات التصنيف للعثور على أنماط التفضيلات الجماعية. ثم تم دمج نتائج الطريقتين، اللتين تم اختبارهما على بيانات من 200 طالب و 10 مجتمعات، باستخدام نهج متوسط بسيط. أظهرت نتائج التقييم باستخدام مقياس متوسط الدقة (MAP) أن طريقة التصفية الهجينة حققت أعلى أداء بدرجة 0.866، متجاوزة بشكل كبير الطرق الفردية للتصفية التعاونية (0.816) والتصفية القائمة على المحتوى (0.733). وخلصت هذه النتائج إلى أن الجمع بين الطريقتين يوازن بشكل فعال بين أهمية المحتوى الشخصي والتفضيلات الجماعية، مما يثبت أنه حل صالح وأكثر تكيفاً لتحسين دقة أنظمة التوصية في البيئات الأكاديمية.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan laman www.dataindonesia.id ketidaksesuaian kuantitas dan kualitas talenta TIK menjadi hambatan serius dalam memenuhi kebutuhan industri, terutama dalam menghadapi lonjakan permintaan yang signifikan. Menurut proyeksi Kementerian Ketenagakerjaan, kebutuhan tenaga kerja di sektor Teknologi, Informasi, dan Komunikasi (TIK) di Indonesia diperkirakan mencapai 1,23 juta pada 2022, meningkat 21,4% menjadi 1,49 juta pada 2023, dan mencapai 1,74 juta pada 2024. Namun, laporan Badan Litbang SDM Kementerian Komunikasi dan Informatika (Kemenkominfo) pada 2020 menunjukkan bahwa hanya ada sekitar 430.000 lulusan TIK di Indonesia. Kesenjangan antara permintaan industri yang terus berkembang dan ketersediaan talenta yang terbatas menciptakan tantangan serius dalam menghadapi era digitalisasi yang membutuhkan keahlian TIK yang canggih.

Selain itu, menurut penelitian Arifitama (2015) hadirnya jenis pekerjaan yang sangat beragam pada industri informatika menjadikan mahasiswa khususnya jurusan Teknik Informatika cenderung memiliki kesulitan dalam menentukan bidang profesi mana yang ingin digeluti. Mengutip hal tersebut, perguruan tinggi dituntut agar mempunyai lulusan mahasiswa yang ahli dalam bidang tertentu agar tidak salah pilih profesi terutama di industri informatika. Dalam keadaan ini, sangat kurang apabila mahasiswa hanya mengandalkan kemampuan akademik pada

perkuliahan saja, terlebih mahasiswa program studi Teknik Informatika UIN Malang.

Menurut Siregar *et al.* (2020) banyak mahasiswa prodi Teknik Informatika tidak mengetahui minat atau bidang yang ada pada ilmu informatika itu sendiri. Hal ini menyebabkan tidak hanya ketidakpastian dalam memilih profesi, tetapi juga bisa berpengaruh pada nilai dan prestasi akademik mahasiswa di prodi Teknik Informatika. Salah satu hal yang bisa mendukung masalah diatas adalah adanya Komunitas Akademik. Menurut RIRIEN KUSUMAWATI (2020) Komunitas Akademik adalah sekumpulan mahasiswa yang mempunyai latar belakang berbeda tetapi mempunyai minat tertentu untuk mencapai tujuan yang sama. Penelitian tersebut menyebutkan bahwasannya peran Komunitas Akademik sangat penting bagi Mahasiswa khususnya dalam belajar di beberapa minat tertentu. Adapun berkenaan dengan mahasiswa yang mempunyai latar belakang yang berbeda-beda tetapi dari perbedaan itu maka bisa saling kenal mengenal. Seperti yang disebutkan pada Al-Quran surah Al-Hujurat ayat 13 yang berbunyi.

يَا أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِنْ ذَكَرٍ وَأُنْثَىٰ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارَفُوا ۗ إِنَّ أَكْرَمَكُمْ عِنْدَ اللَّهِ أَتْقَاكُمْ ۗ إِنَّ اللَّهَ عَلِيمٌ خَبِيرٌ

“Hai manusia, sesungguhnya Kami menciptakan kamu dari seorang laki-laki dan seorang perempuan dan menjadikan kamu berbangsa-bangsa dan bersuku-suku supaya kamu saling kenal-mengenal. Sesungguhnya orang yang paling mulia di antara kamu di sisi Allah ialah orang yang paling takwa di antara kamu. Sesungguhnya Allah Maha Mengetahui lagi Maha Mengenal” (QS Al-Hujurat:13).

Dalam surat ini Allah SWT mengajarkan bahwa umat islam diciptakan dari seorang laki-laki dan Perempuan agar saling mengenal. Maka Allah SWT menekankan bahwasannya manusia adalah makhluk sosial yang saling bergantung satu sama lain (Jannah *et al.*, 2021). Menurut tafsir Quraish Shihab, Allah SWT

telah menciptakan manusia berbangsa-bangsa dan bersuku-suku agar saling mengenal dan saling menolong.

Mengutip hal itu Program studi teknik informatika di UIN Malang telah menyediakan fasilitas bagi mahasiswanya untuk mengembangkan bakat melalui berbagai komunitas. Komunitas-komunitas ini memberikan tempat bagi mahasiswa untuk mengejar dan mengembangkan minat di bidang IT Suprihatin & Setiowati, (2021). Dalam program studi ini, terdapat 12 komunitas, termasuk 10 komunitas akademik dan 2 komunitas non-akademik seperti MOCAP (Android), WEBBOENDER (Web), UINUX (Desain interface), MAMUD (Multimedia), ETH0 (Jaringan), UINBUNTU (Sistem Operasi), FUN JAVA (Pemrograman Java), GDSC (Google), DSE (Data Science), ONTAKI (Robotic), ALFATAA (Sholawat), dan ISC (Olahraga). Komunitas-komunitas ini berfungsi sebagai platform untuk berbagi pengetahuan, ide, dan kolaborasi antara mahasiswa dengan tujuan menghasilkan proyek aplikasi teknologi informasi yang inovatif.

Mengutip hal tersebut, bahwasannya komunitas adalah sarana bagi mahasiswa untuk saling tolong menolong dalam kebaikan. Seperti potongan ayat firman Allah SWT pada surat Al-Maidah ayat 2:

وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ

“...Dan tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa...” (QS Al-Maidah: 2)

Dalam ayat tersebut bahwa Allah SWT menyeru manusia untuk berkerjasama dalam kebaikan dan melarang melakukan keburukan yang menambah dosa atau pun permusuhan (PUSPITASARI, 2022). Menurut Tafsir Quraish Shihab, ayat tersebut menganjurkan kita untuk saling tolong menolong kepada kebaikan. Komunitas akademik yang saling mendukung dan berkolaborasi dalam

mencari pengetahuan dan memajukan ilmu adalah bentuk konkret dari pelaksanaan ayat ini.

Meski demikian, di Teknik Informatika UIN Malang, salah satu masalah utamanya adalah kesulitan mahasiswa dalam memilih Komunitas Akademik yang sesuai dengan minat dan passion mahasiswa tersebut. Dalam proses memilih komunitas akademik TI UIN Malang, tidak sedikit mahasiswa yang masih kebingungan. Karena mahasiswa TI memiliki beragam minat, keterampilan, dan pengalaman yang berbeda-beda. Saat ini masih banyak mahasiswa yang memilih hanya karena popularitasnya tanpa mempertimbangkan sejauh mana komunitas tersebut memberikan kontribusi pada pengembangan keahlian atau mendukung jalur karir yang sesuai dengan minat mahasiswa tersebut. Pada akibatnya mahasiswa akan menjadi pasif dan kurang kontribusi terhadap komunitas yang dipilihnya.

Sebagai solusi untuk memaksimalkan potensi minat dari mahasiswa TI di UIN Malang, perlu dikembangkan sebuah sistem rekomendasi yang efektif untuk memudahkan mahasiswa dalam memilih komunitas yang sesuai dengan minat, keterampilan dan pengalaman mahasiswa tersebut. Sistem rekomendasi ini dapat membantu komunitas untuk mendapatkan anggota baru yang berpotensi dengan menganalisis dan memperhitungkan preferensi dan kesamaan antara minat, keterampilan dan pengalaman mahasiswa dengan isi fokus bidang keahlian dan kegiatan-kegiatan komunitas akademik.

Sistem rekomendasi adalah sistem yang merekomendasikan item yang relevan kepada pengguna berdasarkan preferensi atau kebiasaan pengguna tersebut

Yan *et al.*, (2023). Sistem rekomendasi pada penelitian ini bertujuan untuk memberikan rekomendasi beberapa komunitas akademik yang sesuai dengan minat, keterampilan, dan pengalaman mahasiswa TI UIN Malang. Sistem rekomendasi secara umum memiliki tiga pendekatan utama yaitu *content-based filtering*, *collaborative filtering*, dan *hybrid filtering*. *Content-based filtering* memberikan rekomendasi berdasarkan karakteristik pengguna, seperti minat, keterampilan, dan pengalaman.

Metode ini cocok untuk pengguna baru karena tidak membutuhkan data dari pengguna lain, namun memiliki kelemahan dalam hal *overspecialization*, yakni hanya merekomendasikan item yang terlalu mirip dengan preferensi sebelumnya dan mengurangi keberagaman rekomendasi (Rahman *et al.*, 2022). *Collaborative filtering*, di sisi lain, memanfaatkan data interaksi antar pengguna untuk merekomendasikan item berdasarkan kesamaan preferensi. Keunggulannya terletak pada kemampuan menemukan pola baru dari komunitas pengguna, tetapi memiliki kekurangan pada masalah *cold start*, yaitu kesulitan memberikan rekomendasi kepada pengguna baru yang belum memiliki riwayat aktivitas.

Sebagai solusi yang lebih komprehensif, *hybrid filtering* menggabungkan keunggulan dari *content-based* dan *collaborative filtering*. Pendekatan ini mampu mengurangi risiko *cold start* dan *overspecialization* secara bersamaan, serta meningkatkan akurasi dan relevansi rekomendasi. Namun, implementasi *hybrid filtering* cenderung lebih kompleks karena memerlukan pengolahan data yang lebih besar dan integrasi dua sistem. Dalam penelitian ini, pendekatan *hybrid filtering* dipilih karena dianggap paling tepat untuk memberikan rekomendasi komunitas

akademik yang sesuai bagi mahasiswa Teknik Informatika UIN Malang. Metode ini akan memanfaatkan *content-based filtering* dengan pendekatan TF-IDF dan *cosine similarity* berdasarkan minat, keterampilan, dan pengalaman mahasiswa yang diperoleh melalui survei. Selain itu, sistem juga akan mengadopsi *collaborative filtering* berdasarkan rating mahasiswa terhadap komunitas yang pernah mereka ikuti. Dengan menggabungkan kedua pendekatan tersebut, diharapkan sistem rekomendasi ini dapat memberikan hasil yang akurat dan mendukung mahasiswa dalam memilih komunitas secara lebih terarah dan efektif.

1.2 Pernyataan Masalah

Pernyataan masalah dalam penelitian ini sebagai berikut:

- a. Bagaimana membangun sistem rekomendasi Komunitas Akademik yang sesuai dengan minat mahasiswa dengan metode *hybrid filtering*.
- b. Bagaimana performa metode *hybrid filtering* dalam memberikan rekomendasi Komunitas Akademik terhadap Mahasiswa.

1.3 Tujuan Penelitian

Adapun yang menjadi tujuan dalam penelitian ini adalah sebagai berikut:

- a. Membangun sistem rekomendasi Komunitas Akademik yang sesuai dengan minat dari Mahasiswa dengan metode *hybrid filtering*.
- b. Mengetahui performa metode *hybrid filtering* dalam memberikan rekomendasi.

1.4 Manfaat Penelitian

Hasil penelitian yang diharapkan oleh penulis adalah agar memberikan manfaat sebagai berikut:

- a. Menambah pengetahuan tentang sistem rekomendasi dalam konteks pemilihan Komunitas Akademik menggunakan metode hybrid filtering dalam menghasilkan rekomendasi yang sesuai bagi Mahasiswa Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
- b. Memberikan kemudahan kepada Mahasiswa Program Studi Universitas Islam Negeri Maulana Malik Ibrahim Malang dalam memilih Komunitas Akademik yang sesuai dengan preferensi mahasiswa tersebut.
- c. Memberikan manfaat kepada Komunitas Akademik TI UIN Malang dalam menemukan anggota-anggota yang berpotensi membuat komunitas lebih hidup dan berjalan dengan baik.
- d. Terbangunnya budaya kolaborasi dan sharing knowledge dalam lingkungan komunitas akademik Program Studi Teknik Informatika UIN Malang.

1.5 Batasan Masalah

Berikut adalah beberapa batasan masalah dalam penelitian ini agar tercipta ruang lingkup penelitian yang terarah dan terkendala:

- a. Objek yang digunakan yaitu Mahasiswa Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
- b. Sistem rekomendasi ini hanya digunakan sebagai rekomendasi bagi Mahasiswa, keputusan dalam pemilihan Komunitas Akademik tetap pada Mahasiswa
- c. Data yang digunakan hanya dari komunitas akademik Program Studi Teknik Informatika UIN Malang

BAB II

STUDI PUSTAKA

2.1 Sistem Rekomendasi

Sistem rekomendasi adalah peralatan perangkat lunak yang merekomendasikan item yang relevan kepada pengguna berdasarkan preferensi atau kebiasaan pengguna tersebut. Sistem rekomendasi menggunakan berbagai algoritma dan Teknik untuk menganalisis data pengguna dan memberikan rekomendasi yang akurat (Mishra & Vijayalaxmi, 2022).

Dalam penerapannya rekomendasi akan berkaitan dengan keputusan pengguna (*decision making*), pengguna akan memilih sesuai dengan kriteria yang ada. Seperti dalam Al-Quran pada surat Al-Qalam ayat 38 yang berbunyi :

إِنَّ لَكُمْ فِيهِ لَمَا تَخَيَّرُونَ

“Bahwa di dalamnya kamu benar-benar boleh memilih apa yang kamu sukai untukmu”(QS Al-Qalam:38).

Ayat tersebut menurut tafsir jalalain, membahas tentang bagaimana bahwa seluruh keputusan ada pada diri orang-orang ketika dipertemukan dengan keputusan untuk memilih sebuah kitab untuk dipelajari, atau disebutkan juga pada tafsir jalalain bahwa kalian bebas memilih dengan sesuka hati (Tafsirweb, 2015). Sama halnya dengan penelitian ini, sistem rekomendasi hanya sekedar saran tetapi keputusan semuanya ada pada pengguna.

Dalam penerapannya sistem rekomendasi bisa memberikan manfaat dalam berbagai bidang, seperti membantu dalam pengambilan keputusan memilih barang,

film, musik yang sesuai dengan preferensi penggunanya. Dalam penelitian ini juga membantu Mahasiswa dalam bidang Pendidikan yaitu pemilihan Komunitas Akademik pada Program Studi Teknik Informatika UIN Maulana Malik Ibrahim Malang. Sistem rekomendasi akan merekomendasikan item yang sesuai dengan latar belakang pengguna yang berbeda bukan hanya lewat rating saja. Artinya, setiap pengguna pasti memiliki hasil rekomendasi yang berbeda tergantung preferensi profil pengguna dan data komunitas akademik yang ada.

Oleh karena itu, sistem rekomendasi membutuhkan preferensi pengguna untuk menghasilkan item atau komunitas akademik yang sesuai. Dari preferensi pengguna itu akan dilakukan analisa terhadap sumber data yang ada. Data yang ada akan dihitung nilai kemiripan antar preferensi pengguna dan item menggunakan algoritma pengolahan data dan teknologi machine learning.

Sistem rekomendasi memiliki 4 jenis metode, yaitu *content-based filtering* atau sistem berbasis konten yang memiliki ciri khas informasi dan dekripsi item , *collaborative filtering* atau sistem rekomendasi kolaboratif yang memiliki ciri khas dari penilaian pengguna terhadap item , *demographic filtering* atau sistem rekomendasi demografis yang didasarkan dengan hal-hal umum terhadap pengguna seperti gender, dan *hybrid filtering* atau sistem rekomendasi gabungan antara 2 atau lebih metode (Sharma *et al.*, 2021). Berikut beberapa penelitian yang membahas tentang sistem rekomendasi yang menggunakan beberapa metode atau algoritma yang berbeda.

Tabel 2. 1. Penelitian Terkait Sistem Rekomendasi

Peneliti	Topik Sistem rekomendasi	Metode			Variabel	Hasil
		Pra	Main	Post		
Nastiti, (2019)	Tanaman Pangan	TF-IDF	Content-Based Filtering	Cosine Similarity	3 (Kel. Tani, varietas, jumlah, lokasi)	Skor Presisi 78,40%
Larasati & Februariyanti, (2021)	Produk Emina	TF-IDF	Content-Based Filtering	Cosine Similarity	9(Face, Lip, Eye, Cheek, Body, Skincare, Nail, Other)	Nilai Cosine Similarity tertinggi 0,75
Esteban et al., (2020)	Mata Kuliah Pilihan	Genetic Algorithm	Collaborative filtering & Content-based filtering	Semantic similarity, Euclidean distance, Taxicab distance, dll	7(Ratings, Grades, Branch, Professor, competences, knowledge area, contents)	Skor RSME 0,971
Mondi & Wijayanto, (2019)	Menu makanan	-	Content-Based Filtering	Cosine Similarity	6 (restoran, menu, rasa, fasilitas, kategori, lokasi)	Skor Presisi 0,8915 dan akurasi 0,5118
Nilashi et al., (2018)	Hotel	Clustering, Fuzzy	Collaborative Filtering	-	Rating Hotel, review hotel	Skor presisi 0,908
Yang et al.,(2017)	Pekerjaan	-	Hybrid Filtering & SRL (Statistical Relational Learning)	-	Data pekerjaan, data user	Skor Akurasi 0,964

Pada penelitian yang dilakukan oleh Nastiti, (2019) berhasil melakukan rekomendasi lahan pertanian yang sesuai dengan pengguna. Penelitian tersebut menggunakan metode *content-based filtering* dan menggunakan algoritma vector space model. Penelitian tersebut berhasil merekomendasikan dengan hasil presisi 78,40%.

Pada penelitian yang dilakukan oleh Larasati & Februariyanti, (2021) berhasil melakukan rekomendasi produk kosmetik Emina. Dalam penelitian tersebut menggunakan *content-based filtering* yang mana mengambil data dari

produk yang dicari customer dengan deskripsi produk yang ada. Penelitian tersebut menggunakan *tfidf* sebagai pembobotan deskripsi produk dan *cosine similarity* untuk menghitung kesamaan antar kalimat. Dalam penelitian ini berhasil mendapatkan rekomendasi produk dengan similaritas tertinggi dengan nilai sebesar 0,7195.

Pada penelitian Esteban *et al.*, (2020) melakukan perbandingan metode *collaborative filtering* dan *content-based filtering* untuk merekomendasikan mata kuliah pilihan yang sesuai kepada mahasiswa berdasarkan kriteria tertentu. Penelitian ini menggunakan Genetic Algorithm (GA) untuk mengoptimalkan parameter konfigurasi dan bobot pada setiap kriteria. Studi kasus penelitian ini ditujukan untuk program studi Ilmu Komputer. Dalam penelitian ini didapatkan skor RMSE (Root Mean Squared Error) sebesar 0,971, nDCG (Normalized Discount Cumulative Gain) 0,682, Reach 100% dan Time 3.022 detik.

Pada penelitian Mondy & Wijayanto, (2019) membangun sistem rekomendasi restoran pada aplikasi MANGAN dengan menggunakan metode *content-based filtering*. Dalam penelitian ini penulis menggunakan perbandingan dari Item Profile restoran yang diambil dari data aplikasi MANGAN yang selanjutnya akan dilakukan kesamaan similaritas menggunakan algoritma *cosine similarity*. Hasil yang didapatkan dalam penelitian ini menunjukkan nilai rata-rata presisi sebesar 0,8915 dan akurasi sebesar 0,5118.

Pada penelitian Nilashi *et al.*, (2018) membangun sistem rekomendasi menggunakan metode *collaborative filtering* yang datanya diambil dari aplikasi TripAdvisor. Menggunakan algoritma CART dan Teknik berbasis aturan fuzzy

untuk prediksi perangkat secara keseluruhan. Menghasilkan hasil skor presisi 0,908, F1 0,897, dan MAE 0,788.

Pada penelitian Yang *et al.* (2017) membangun sebuah sistem rekomendasi pekerjaan menggunakan metode *hybrid* antara *Content-Based Filtering*, *Collaborative Filtering* dan *Statistical Relational Learning*. Paper ini menghasilkan hasil yang sangat bagus mempunyai skor akurasi sebesar 0,964.

2.2 Komunitas Akademik

Komunitas akademik adalah suatu perkumpulan akademisi yang mendukung atau meningkatkan pengembangan profesional dan kolaborasi akademisi Zou, (2019). Komunitas akademik ini memberikan ruang bagi akademisi untuk terlibat dalam pertukaran pengetahuan atau pembelajaran kolaboratif (Azcoitia *et al.*, 2017). Di Program Studi informatika sendiri sudah mempunyai komunitas akademik sebanyak 10, yaitu MOCAP (Android), WEBOENDER (Web), UINUX (Desain interface), MAMUD (Multimedia), ETH0 (Jaringan), UINBUNTU (Sistem Operasi), FUN JAVA (Pemrograman Java), GDSC (Google), DSE (Data Science), dan ONTAKI (Robotic).

Dalam islam perbedaan tidak bisa dijadikan sebagai alasan untuk gagal dalam mencapai suatu tujuan, karena dengan adanya perbedaan itu bisa saling kenal mengenal satu sama lain, seperti yang disebutkan pada Al-Quran surah Al-Hujurat ayat 31 yang berbunyi

يَا أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِنْ ذَكَرٍ وَأُنْثَىٰ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارَفُوا ۗ إِنَّ أَكْرَمَكُمْ عِنْدَ اللَّهِ أَتَقْوَاهُ ۗ إِنَّ اللَّهَ عَلِيمٌ خَبِيرٌ

“Hai manusia, sesungguhnya Kami menciptakan kamu dari seorang laki-laki dan seorang perempuan dan menjadikan kamu berbangsa-bangsa dan bersuku-suku supaya kamu saling kenal-mengenal. Sesungguhnya orang yang paling mulia di antara kamu di sisi Allah SWT ialah orang yang paling takwa di antara kamu. Sesungguhnya Allah SWT Maha Mengetahui lagi Maha Mengenal” (QS Al-Hujurat:13).

Menurut tafsir Al-Madinah Al-Munawwaroh Allah SWT menyampaikan bahwa kami menjadikan kalian berbagai bangsa melalui perkembangbiakan, dan dari bangsa-bangsa itu menjadi berbagai kabilah dan suku agar kalian saling kenal mengenal. Sungguh yang paling baik derajatnya di sisi Allah SWT. adalah orang yang paling bertakwa diantara kalian. Allah SWT Maha Mengetahui keadaan serta urusan hamba-hambanya.

Dalam studi yang dilakukan oleh Kusumawati *et al.*, (2021) dan Praherdhiono *et al.*, (2019) mengenai penerapan metode pendidikan cMOOC dalam mendukung berbagai aspek pengetahuan dan kolaborasi di lingkungan akademik, mahasiswa jurusan Teknik Informatika di UIN Malang menyimpulkan bahwa keterlibatan dalam komunitas akademik dapat memberikan keuntungan signifikan. Manfaat tersebut meliputi peningkatan keterampilan dan pengetahuan, perluasan jaringan, serta kemampuan untuk berbagi informasi, yang pada gilirannya menciptakan lingkungan belajar yang bersifat kolaboratif dan produktif. Pada komunitas-komunitas itulah diharapkan Mahasiswa bisa menemukan dan mengembangkan minat dan bakatnya sebelum memasuki dunia kerja pada bidang IT yang sebenarnya.

2.3 Content-based Filtering

Content-Based Filtering (CBF) atau pemfilteran berbasis konten merupakan salah satu pendekatan populer dalam sistem rekomendasi yang bertujuan untuk memberikan rekomendasi secara personal kepada pengguna dengan menganalisis karakteristik dari item yang tersedia. Teknik ini berfokus pada informasi atau deskripsi dari item, seperti fitur, atribut, atau metadata, dan mencocokkannya dengan profil atau preferensi pengguna. Menurut Shambour et al. (2023), CBF memungkinkan sistem untuk mengusulkan item yang serupa dengan apa yang sebelumnya disukai atau dikonsumsi oleh pengguna tersebut.

Profil pengguna dalam content-based filtering biasanya dibentuk berdasarkan riwayat interaksi pengguna terhadap item tertentu, seperti penilaian (rating), klik, pembelian, atau waktu yang dihabiskan pada konten tertentu. Berdasarkan perilaku tersebut, sistem akan mengidentifikasi fitur-fitur yang sering diasosiasikan dengan minat pengguna. Misalnya, jika seorang pengguna sering menonton film bergenre aksi dan bertema petualangan, maka sistem akan cenderung merekomendasikan film lain dengan genre dan tema serupa.

Memandingkan diri sendiri atau orang lain dengan orang lain dalam agama Islam dapat dilarang, tetapi juga boleh. Yang tidak boleh kita lakukan adalah mengambil sikap membandingkan ini, karena itu akan membuat kita kurang bersyukur atas nikmat yang telah diberikan Allah SWT kepada kita. Perbandingan bisa dianjurkan apabila membandingkan diri dengan orang lain ketika memiliki

tujuan mendapat motivasi dari orang lain, seperti yang diriwayatkan dalam hadits dibawah ini :

عن أبي هريرة رضي الله عنه قال: قال رسول الله صلى الله عليه وسلم: انظروا إلى من هو أسفل منكم ولا تنظروا إلى من هو فوقكم فإنه أجدر أن لا تزدروا نعمة الله عليكم

“Dari Abu Hurairah Radiyallahu ‘anhu, beliau berkata, Rasulullah SAW bersabda, Lihatlah siapa yang berada dibawah kalian, dan jangan melihat orang yang berada diatas kalian, sebab yang demikian lebih patut agar kalian tidak memandang remeh nikmat Allah SWT atas kalian.” (HR al-Bukhari).

Dalam hadits ini, Imam Ibnu Hajar mengatakan bahwa dalam hal ibadah, seseorang harus membandingkan dirinya dengan orang lain yang memiliki kualitas yang lebih baik darinya sehingga mereka termotivasi untuk meningkatkan kualitas ibadah mereka. Sebaliknya, dalam hal urusan duniawi, seseorang harus membandingkan dirinya dengan orang lain yang nasibnya lebih buruk darinya sehingga mereka dapat tetap bersyukur atas kelebihan yang diberikan kepada mereka (Abror, 2022). Hal ini yang mungkin menjadi acuan dalam penerapan metode sistem rekomendasi *content-based filtering*, yang didalamnya terdapat proses membandingkan antara data inputan dengan data yang sudah ada.

Seperti dijelaskan oleh Al-Ajlan & Alshareef (2023), metode ini memanfaatkan representasi vektor dari fitur item, dan menggunakan teknik seperti cosine similarity atau korelasi Pearson untuk mengukur tingkat kemiripan antara item dalam basis data dengan preferensi pengguna. Dengan pendekatan ini, sistem tidak perlu mengandalkan data dari pengguna lain, sehingga cocok untuk mengatasi

masalah cold-start bagi pengguna baru, asalkan terdapat cukup informasi terkait item yang sudah diakses oleh pengguna tersebut.

Content-based filtering juga memiliki fleksibilitas yang tinggi dan dapat diterapkan dalam berbagai domain aplikasi. Contohnya, dalam layanan kesehatan, sistem dapat merekomendasikan perawatan atau obat berdasarkan riwayat medis pasien dan karakteristik penyakit; dalam platform streaming, sistem dapat menyarankan film atau musik berdasarkan genre, aktor, atau artis yang sering dinikmati oleh pengguna; dan dalam e-commerce, sistem dapat menawarkan produk berdasarkan fitur seperti merek, warna, harga, dan kategori produk (Nasser et al., 2023).

2.4 Collaborative Filtering

Collaborative Filtering (CF) adalah algoritma dasar yang digunakan dalam sistem pemberi rekomendasi untuk membuat rekomendasi yang dipersonalisasi berdasarkan preferensi banyak pengguna. Algoritma CF menambang atau mempelajari pola korelatif dari data untuk mencocokkan pengguna dengan item yang diminati. Metode CF tradisional mencakup pendekatan berbasis memori seperti CF berbasis pengguna/item, serta metode berbasis pembelajaran seperti faktorisasi matriks dan model pembelajaran mendalam (Li et al., 2020).

Konsep pengambilan keputusan berdasarkan masukan atau preferensi dari banyak pihak ini sejalan dengan prinsip musyawarah (syura) dalam Islam. Musyawarah adalah proses saling bertukar pendapat untuk mencapai keputusan

terbaik. Allah SWT memuji orang-orang yang senantiasa menjadikan musyawarah sebagai landasan dalam urusan mereka, sebagaimana disebutkan dalam Al-Qur'an surah Asy-Syura ayat 38:

وَالَّذِينَ اسْتَجَابُوا لِرَبِّهِمْ وَأَقَامُوا الصَّلَاةَ وَأَمْرُهُمْ شُورَى بَيْنَهُمْ وَمِمَّا رَزَقْنَاهُمْ يُنفِقُونَ

Artinya: "Dan (bagi) orang-orang yang menerima (mematuhi) seruan Tuhan dan melaksanakan salat, sedang urusan mereka (diputuskan) dengan musyawarah antara mereka; dan mereka menginfakkan sebagian dari rezeki yang Kami berikan kepada mereka" (QS. Asy-Syura: 38).

Menurut Tafsir Al-Muyassar, ayat ini menjelaskan bahwa salah satu sifat orang-orang beriman adalah bahwa urusan-urusan penting yang mereka hadapi diputuskan atas dasar musyawarah dan saling tukar pandangan di antara mereka, dan mereka tidak memutuskannya sendiri tanpa berkonsultasi. Prinsip ini mencerminkan esensi dari Collaborative Filtering, di mana rekomendasi untuk seorang pengguna tidak ditentukan secara terisolasi, melainkan 'dimusyawarahkan' dengan menganalisis data preferensi dari komunitas pengguna yang lebih luas. Dengan demikian, sistem merekomendasikan item berdasarkan 'kebijaksanaan kolektif' dari para pengguna yang memiliki kesamaan minat, layaknya hasil dari sebuah musyawarah yang matang.

Namun, ada kebutuhan untuk beralih dari pembelajaran korelatif ke pembelajaran kausal di CF. *Casual Collaborative Filtering (CCF)* adalah kerangka umum yang memodelkan kausalitas dalam CF dan memungkinkan pemikiran di luar data observasi untuk pemodelan dan personalisasi pengguna (Liu et al., 2019). CCF memberikan pandangan kausal terpadu tentang CF dan menawarkan

pendekatan intervensi bersyarat untuk memperkirakan hubungan sebab akibat berdasarkan data observasi (L. Zou *et al.*, 2020). Hasil eksperimen menunjukkan bahwa CCF dapat meningkatkan kinerja rekomendasi dari banyak algoritma CF (Xu *et al.*, 2023).

2.5 Hybrid Filtering

Hybrid filtering merupakan metode dalam sistem rekomendasi yang menggabungkan beberapa teknik atau pendekatan untuk meningkatkan kualitas rekomendasi. Pendekatan ini menggabungkan dua atau lebih metode seperti *content-based filtering* dan *collaborative filtering* dalam satu sistem rekomendasi (Muflihah & Wardhani, 2024). Menurut Burke, Hybrid recommender system adalah metode yang mengkombinasikan dua atau lebih teknik rekomendasi untuk meningkatkan performansi rekomendasi, yang biasanya digunakan untuk memecahkan masalah yang ada di masing-masing metode yang digunakan (Muflihah & Wardhani 2024).

Prinsip penggabungan ini untuk mencapai hasil yang lebih unggul dan seimbang memiliki kesamaan dengan konsep wasatiyyah atau moderasi dalam Islam. Islam mendorong umatnya untuk menjadi umat yang pertengahan (Ummatan Wasatan), yang adil dan seimbang, tidak condong pada ekstremitas. Hal ini seperti yang difirmankan Allah SWT dalam Al-Qur'an surah Al-Baqarah ayat 143:

وَكَذَلِكَ جَعَلْنَاكُمْ أُمَّةً وَسَطًا لِتَكُونُوا شُهَدَاءَ عَلَى النَّاسِ وَيَكُونَ الرَّسُولُ عَلَيْكُمْ شَهِيدًا

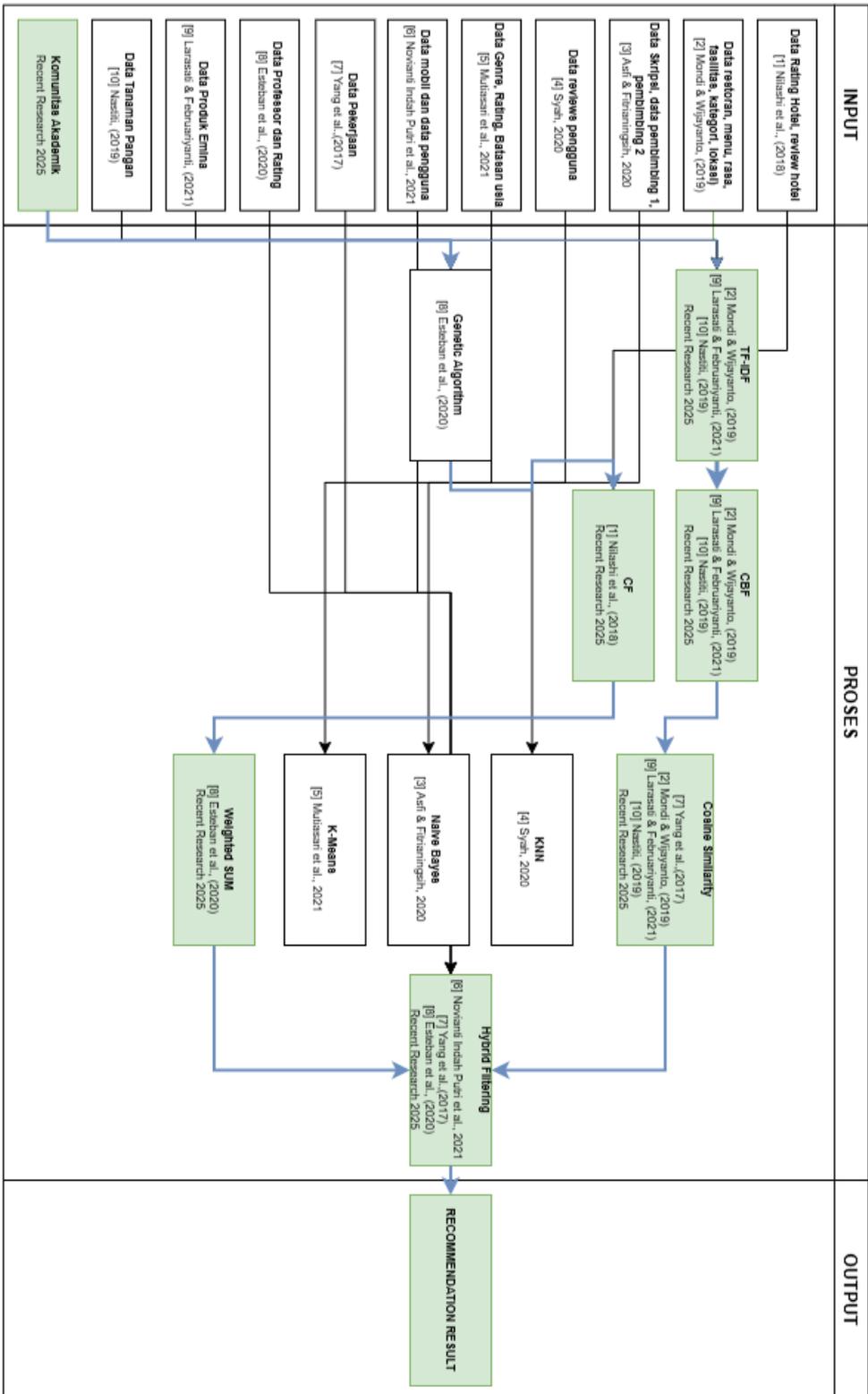
Artinya: “Dan demikian (pula) Kami telah menjadikan kamu (umat Islam), umat yang adil dan pilihan agar kamu menjadi saksi atas (perbuatan) manusia dan agar Rasul (Muhammad) menjadi saksi atas (perbuatan) kamu.” (QS. Al-Baqarah: 143).

Menurut Tafsir Al-Muyassar, makna Ummatan Wasatan adalah umat yang adil dan terbaik. Mereka dijadikan umat pertengahan agar menjadi saksi bagi umat-umat lain pada hari kiamat. Konsep 'pertengahan' ini dapat dianalogikan dengan metode hybrid filtering. Metode ini tidak bergantung sepenuhnya pada satu pendekatan saja—baik itu *content-based* yang berfokus pada profil pribadi pengguna, maupun collaborative yang berfokus pada preferensi kolektif. Sebaliknya, metode hybrid mengambil jalan tengah dengan mengintegrasikan kekuatan dari kedua pendekatan untuk menutupi kekurangan masing-masing. Dengan demikian, sistem ini menghasilkan rekomendasi yang lebih adil, seimbang, dan akurat, layaknya sebuah umat yang 'wasat' dalam mengambil keputusan.

Untuk membuat sistem hybrid terdapat 7 teknik yang bisa dipakai, yaitu *weighted, switching, mixed, feature combination, feature augmentation, cascade dan meta-level* (Muflihah & Wardhani, 2024). Dalam bukunya, Aggarwal mengelompokan teknik hybrid menjadi 3, yaitu *monolithic*, yang meliputi *combination dan meta-level*, kemudian *ensemble design*, yang meliputi *feature augmentation, cascade, weighted, dan switching*, dan yang terakhir *mixed system* yang berdiri sendiri dan tidak dikategorikan sebagai *monolithic* ataupun *ensemble* (Lucchese *et al.*, 2014).

2.6 Kerangka Teori

Dalam membuat penelitian tentang Sistem Rekomendasi Komunitas Akademik menggunakan metode *Hybrid Filtering* perlu mengacu pada jurnal-jurnal terdahulu yang telah melakukan kerangka teori seperti yang terlihat pada gambar 2.1.



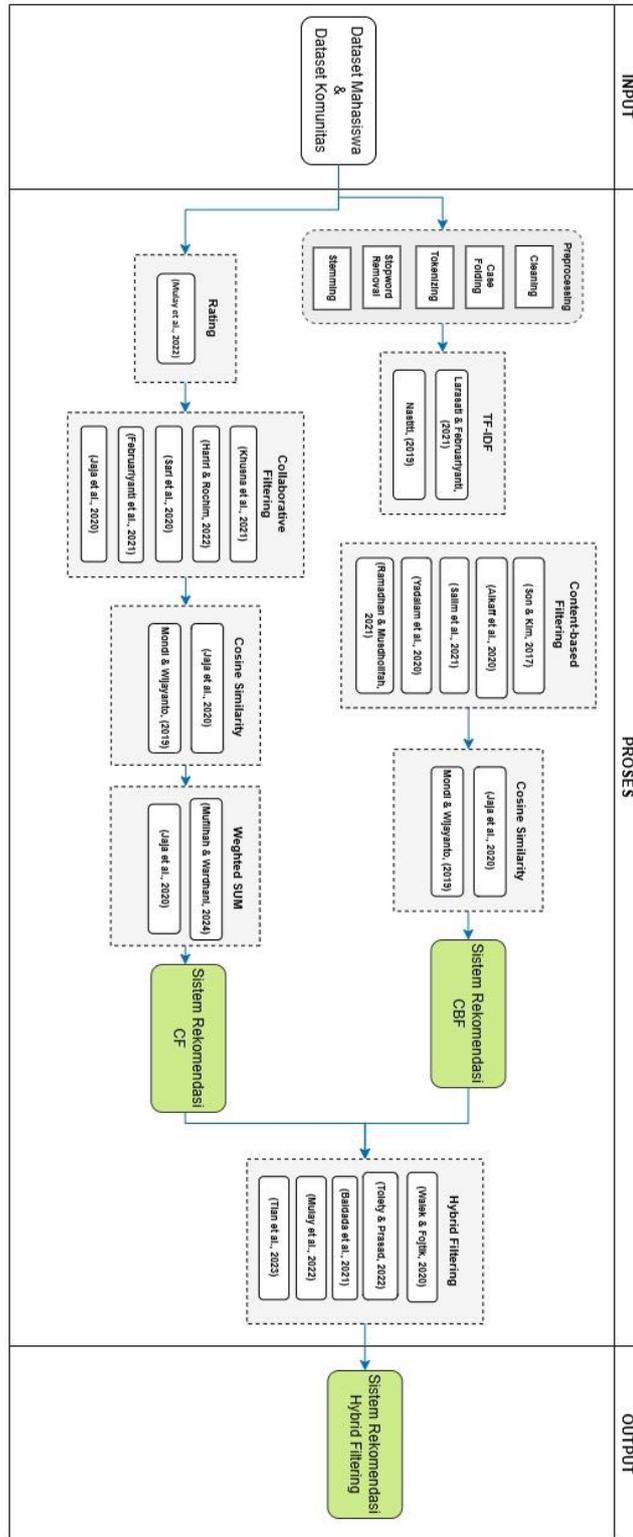
Gambar 2 1 Kerangka Teori

BAB III

METODOLOGI PENELITIAN

Pada bab ini akan membahas secara rinci tentang metode *Hybrid Filtering* yang digunakan untuk Sistem Rekomendasi Komunitas Akademik. Selain itu, bab ini juga akan membahas tentang kerangka konsep, desain dan implementasi sistem yang digunakan dalam penelitian ini.

3.1 Kerangka Konsep



Gambar 3. 1 Kerangka Konsep

Gambar 3.1 menunjukkan kerangka konsep sistem rekomendasi yang dikembangkan dalam penelitian ini. Input berupa Dataset Mahasiswa dan Dataset Komunitas diproses dengan dua pendekatan berbeda. Untuk Content-based Filtering (CBF), data konten proyek diproses terlebih dahulu melalui tahap *preprocessing* seperti *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*, lalu dilakukan pembobotan kata menggunakan TF-IDF. Selanjutnya dilakukan perhitungan kemiripan antar item menggunakan Cosine Similarity. Sementara itu, Collaborative Filtering (CF) menggunakan data interaksi berupa rating tanpa perlu *preprocessing* teks, dan menghitung kesamaan antar pengguna juga dengan Cosine Similarity. Hasil dari kedua metode ini kemudian digabungkan menggunakan pendekatan *Hybrid Filtering*, di mana skor rekomendasi dari CBF dan CF diberi bobot dan dijumlahkan untuk menghasilkan rekomendasi akhir. Keterbaruan dari penelitian ini terletak pada integrasi metode CBF dan CF menggunakan pendekatan *hybrid* yang diharapkan dapat meningkatkan akurasi sistem rekomendasi untuk komunitas akademik.

3.2 Desain Penelitian

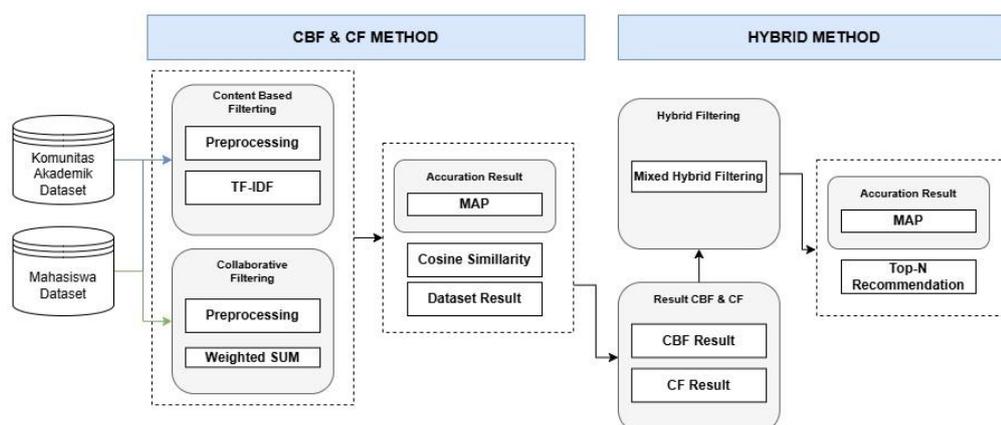


Gambar 3. 2 Desain Penelitian

Penelitian ini dimulai dengan mengidentifikasi permasalahan yang akan dijadikan fokus, dilanjutkan dengan pencarian sumber literatur yang relevan untuk mendukung penelitian yang direncanakan. Kemudian dilakukan pengumpulan data yang diperlukan. Pada penelitian ini melakukan survey melalui Google Form untuk mendapatkan data minat, keterampilan, dan pengalaman Mahasiswa yang akan dihitung kemiripan dengan dataset dari Komunitas Akademik Teknik Informatika UIN Malang. Data tersebut nantinya akan dianalisis agar menghasilkan informasi yang berguna dalam sistem rekomendasi yang lebih akurat (Esteban *et al.*, 2020). Dari hasil analisis tersebut data diproses melalui implementasi *recommender system* yang nantinya akan menghasilkan rekomendasi bagi user. Dalam tahap evaluasi, hasil rekomendasi tersebut membutuhkan evaluasi performa untuk mengukur keakuratan dari hasil perhitungan metode.

3.3 Desain Sistem

Berikut alur desain sistem rekomendasi yang digunakan dalam penelitian ini seperti pada Gambar 3.3 :



Gambar 3.3 Desain Sistem

Gambar di atas menunjukkan desain sistem rekomendasi yang menggabungkan metode *Content-Based Filtering* (CBF) dan *Item-Based Collaborative Filtering* (CF). Sistem dimulai dari dua sumber data, yaitu Dataset Mahasiswa dan Dataset Komunitas Akademik. Untuk CBF, data melalui proses *preprocessing* dan pembobotan menggunakan TF-IDF, lalu dihitung kemiripannya menggunakan *Cosine Similarity*. Sementara CF menggunakan pendekatan *item-based* dengan hasil dari interaksi pengguna tanpa perlu pembobotan teks. Kedua metode ini menghasilkan hasil rekomendasi yang diuji akurasinya menggunakan metrik MAP (*Mean Average Precision*). Pada tahap selanjutnya, hasil dari CBF dan CF digabungkan dalam metode *hybrid* untuk menghasilkan rekomendasi Top-N yang lebih akurat dan relevan.

3.4 Pengumpulan Data

Metode pengumpulan data yang digunakan penulis dalam penelitian ini adalah sebagai berikut:

1. Melakukan survey menggunakan kuisisioner Google Form kepada Mahasiswa Program Studi Teknik Informatika UIN Malang. Isi dari form tersebut yaitu untuk mendapatkan data minat, keterampilan dan pengalaman dari Mahasiswa tersebut dan rating dari Komunitas Akademik. Berikut adalah kriteria pertanyaan yang akan diajukan kepada Mahasiswa:

Tabel 3. 1. Pertanyaan / Kriteria Survey

Pertanyaan/Kriteria	Sumber
Apa Minat dan Passion anda dalam dunia TI?	Esteban et al., (2020)

Pertanyaan/Kriteria	Sumber
Sebutkan pengetahuan khusus sebelumnya dalam bidang teknologi informatika yang anda sudah pelajari	Talaghzi et al., (2023)
Lebih suka manakah anda, bekerja dengan tim besar atau tim kecil?	Pirker et al., (2014)
Apa Bahasa Pemograman/Teknologi yang ingin anda pelajari dari Komunitas Akademik?	Komunitas Akademik
Apa motivasi anda ingin mengikuti komunitas akademik di TI UIN?	Pirker et al., (2014)
Rating Komunitas	

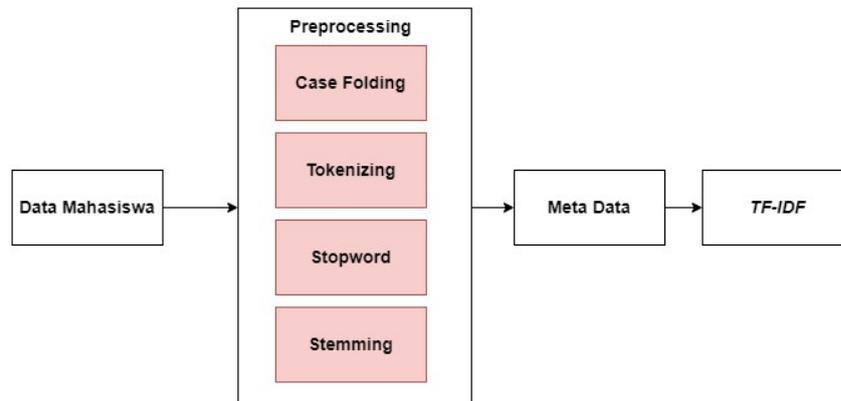
- Mengambil data dari website Program Studi Teknik Informatika UIN Malang secara manual untuk mendapatkan data Komunitas Akademik.

3.4.1 Pengolahan Data

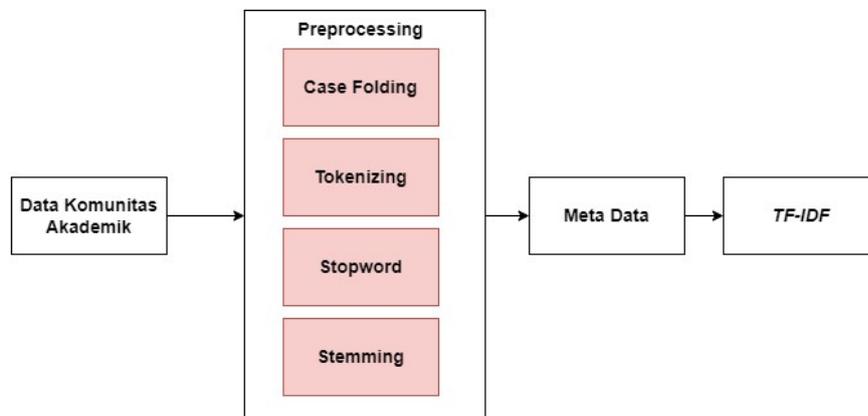
Selanjutnya setelah data terkumpul maka tahapan selanjutnya yaitu pengolahan data. Data yang sudah ada akan diproses menggunakan 2 metode yaitu *collaborative filtering* dan *content-based filtering*. Berikut alur pengolahan kedua metode tersebut.

3.4.1.1 Content Based Filtering

Data yang sudah ada akan di proses menggunakan pembobotan TF-IDF yang kemudian akan dihitung kemiripannya menggunakan algoritma cosine similarity. Berikut alur pengolahan data pada penelitian ini:



Gambar 3. 4 Alur Pegolahan Data Mahasiswa



Gambar 3. 5 Alur Pegolahan Data Komunitas

3.4.1.1.1 Preprocessing

1. Case Folding

Case Folding adalah proses penyamaan case dalam sebuah dokumen. Tidak semua dokumen teks konsisten dalam penggunaan huruf kapital. Oleh karena itu case folding dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar dalam hal ini huruf kecil atau lowercase.

Tabel 3. 2 Case Folding pada Data

Sebelum Case Folding	Sesudah Case Folding
Data Science	data science

Pemograman Web, Pemograman Mobile, Pemograman Java, Multimedia, Data Science, UI/UX	pemograman web, pemograman mobile, pemograman java, multimedia, data science, ui/ux
Tim Besar: membuat kegiatan seperti seminar, workshop	tim besar: membuat kegiatan seperti seminar, workshop
PHP, Python	php, python
Ingin menambah ilmu	ingin menambah ilmu

2. Tokenizing

Tokenizing adalah proses memecah dokumen menjadi Kumpulan kata. Tokenizing dapat dilakukan dengan menghilangkan tanda baca dan memisahkannya per-spasi. Tahapan ini juga menghilangkan karakter-karakter tertentu seperti tanda baca.

Tabel 3. 3 Tokenizing pada Data

Sebelum Tokenizing	Sesudah Tokenizing
Data Science	“data”, “science”
Pemograman Web, Pemograman Mobile, Pemograman Java, Multimedia, Data Science, UI/UX	“pemograman”, “web”, “pemograman”, “mobile”, “pemograman”, “java”, “multimedia”, “data”, “science”, “uiux”
Tim Besar: membuat kegiatan seperti seminar, workshop	“tim”, “besar”, “membuat”, “kegiatan”, “seperti”, “seminar”, “workshop”
PHP, Python	“php”, “python”
Ingin menambah ilmu	“ingin”, “menambah”, “ilmu”

3. Stopword Removal

Stopword Removal adalah proses penghilangan kata-kata yang tidak penting pada dokumen.

Tabel 3. 4 Stopword pada Data

Sebelum Stopword Removal	Sesudah Stopword Removal
data science	data science
pemograman web, pemograman mobile, pemograman java, multimedia, data science, uiux	pemograman web mobile java multimedia data science uiux

tim besar membuat kegiatan seperti seminar, workshop	tim besar kegiatan seminar workshop
php, python	php python
ingin menambah ilmu	menambah ilmu

4. Stemming

Stemming merupakan suatu proses yang terdapat dalam sistem IR (Information retrieval) yang mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (root word) dengan menggunakan aturan-aturan tertentu

Tabel 3. 5 Stemming pada Data

Sebelum Stemming	Sesudah Stemming
“data”, “science”	“data”, “science”
“pemograman”, “web”, “mobile”, “java”, “multimedia”, “data”, “science”, “uiux”	“web”, “mobile”, “java”, “multimedia”, “data”, “science”, “uiux”
“tim”, “besar”, “kegiatan”, “seminar”, “workshop”	“tim”, “besar”, “kegiatan”, “seminar”, “workshop”
“php”, “python”	“php”, “python”
“menambah”, “ilmu”	“tambah”, “ilmu”

3.2.1.1.2 Meta Data

Setelah dilakukan tahapan preprocessing maka seluruh data akan digabung menjadi sebuah meta data yang nantinya akan dilakukan proses berikutnya untuk diberi bobot menggunakan tf-idf.

Tabel 3. 6 Meta Data

Meta Data
data science web mobile java multimedia data science uiux tim besar kegiatan seminar workshop php python tambah ilmu

3.2.1.1.3 TF-IDF

Dalam tahapan ini, TF-IDF digunakan untuk menentukan bobot setiap term dari data Mahasiswa dan Komunitas Akademik. Bobot dalam dokumen dihitung dengan menggunakan tfidf. TF- IDF dikenal efisien, mudah dan memiliki hasil yang akurat. Metode ini menghitung nilai Term Frequency (TF) dan Inverse Document Frequency (IDF) pada setiap kata di setiap dokumen dalam data. Untuk menghitung pembobotan kata dengan melakukan perhitungan manual dengan menerapkan 3 data komunitas sampel yang disajikan pada tabel berikut:

Tabel 3. 7 Sampel Data

Dokumen	Meta Data Komunitas
Q	data science web mobile java multimedia uiux tim besar kegiatan seminar workshop php python tambah ilmu
D1	data science komunitas membahas pengolahan data. pemrograman python. workshop pengolahan data
D2	webonder komunitas bergerak dibidang pemrograman web, mempelajari html, css, php, javascript
D3	robotik komunitas bergerak dibidang robot teknologi digunakan iot bahasa pemograman c++. Kegiatan didalamnya produk robot

Tahapan pertama yang dilakukan dalam TF-IDF adalah menghitung nilai TF. Nilai TF digunakan untuk mengukur seberapa sering sebuah kata muncul dalam dokumen. Perhitungan TF dapat ditunjukkan pada persamaan 3.1.

$$TF = \frac{\text{Jumlah kemunculan kata dalam dokumen}}{\text{Jumlah kata dalam satu dokumen}} \quad (3.1)$$

Setelah melakukan perhitungan TF, kemudian perhitungan IDF yang ditunjukkan pada persamaan 3.2 dan contoh perhitungannya pada tabel 3.9.

$$IDF(t, D) = \text{Log}\left(\frac{N}{df(t)}\right) \quad (3.2)$$

Keterangan:

IDF(t,D) = nilai IDF untuk kata kunci (term) t dalam dokumen D.

N = Jumlah total dokumen

$df(t)$ = jumlah dokumen yang mengandung kata kunci (term) t

Setelah mendapatkan nilai TF dan IDF, langkah selanjutnya melakukan perhitungan bobot TF-IDF yang ditunjukkan pada persamaan 3.3 dan contoh perhitungannya ditunjukkan pada tabel 3.10.

$$W = TF \times (IDF + 1) \quad (3.3)$$

Keterangan:

TF: Jumlah kemunculan kata atau term dalam dokumen

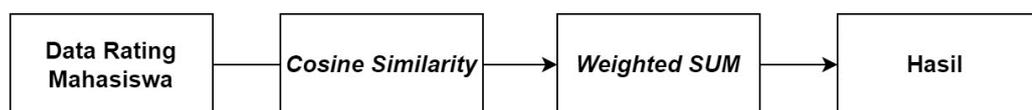
IDF: Inverse document frequency

D: Jumlah semua dokumen

W: bobot setiap dokumen

3.4.1.2 Collaborative Filtering

Data yang sudah ada akan dihitung kemiripan antar penilaian user menggunakan *cosine similarity* kemudian akan dihitung nilai prediksi rating pada komunitas yang belum dinilai menggunakan algoritma *weighted sum*. Berikut alur pengolahan data pada penelitian ini:



Gambar 3. 6 Alur Pengolahan Data Collaborative Filtering

3.4.1. Cosine Similarity

Setelah dilakukan pembobotan dengan TF-IDF maka dilakukan similaritas pada setiap dokumen menggunakan metode cosine similarity. Berikut adalah rumus dari cosine similarity:

$$\text{Cosine Similarity}(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|} \quad (3.4)$$

Keterangan:

A: Vektor A atau dokumen A

B: Vektor B atau dokumen B

$\|A\|$: Nilai mutlak A

$\|B\|$: Nilai mutlak B

3.4.2. Weighted Sum

Weighted Sum akan digunakan untuk memperkirakan rating yang akan diberikan seorang mahasiswa ke komunitas berdasarkan rating yang telah mahasiswa berikan ke komunitas-komunitas lain yang mirip. Berikut adalah rumus untuk algoritma *Weighted Sum*.

$$\hat{r}_{u,i} = \frac{\sum_{j \in N(i)} \text{sim}(i,j) \cdot r_{u,j}}{\sum_{j \in N(i)} |\text{sim}(i,j)|} \quad (3.5)$$

Keterangan:

$\hat{r}_{u,i}$: Rating yang diprediksi untuk user **u** terhadap item **i**

$N(i)$: Sekumpulan item yang mirip dengan item **i** dan telah dirating oleh user **u**

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

$r_{u,j}$: Rating yang diberikan user **u** terhadap item **j**

3.4.3. Hybrid Filtering

Pada penelitian ini, digunakan teknik *mixed hybrid* untuk menggabungkan kedua metode rekomendasi yang diterapkan, yaitu *Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF). Teknik *mixed hybrid* adalah pendekatan yang mengkombinasikan hasil rekomendasi dari berbagai metode secara bersamaan untuk menghasilkan rekomendasi yang lebih akurat dan relevan bagi pengguna. Pendekatan ini memungkinkan sistem untuk memanfaatkan kelebihan dari masing-

masing metode, sehingga dapat memberikan rekomendasi yang lebih baik, terutama pada kondisi di mana salah satu metode mungkin kurang efektif.

Proses penggabungan kedua metode ini melibatkan normalisasi hasil dari CF agar berada dalam rentang yang sama dengan hasil dari CBF, sehingga keduanya dapat dihitung secara seimbang dalam tahap selanjutnya. Selanjutnya, hybrid score dihitung sebagai rata-rata dari hasil CBF dan *Normalized CF*, dan item dengan *hybrid score* di atas ambang batas tertentu (misalnya 0,4) akan dipilih sebagai rekomendasi untuk pengguna (Tommy *et al.*, 2020).

3.5 Pengujian System

Pada tahap ini akan dilakukan perencanaan pengujian terhadap sistem rekomendasi. Adapun metode pengujian yang digunakan dalam penelitian yaitu *Mean Average Precision* (MAP).

3.5.1 Mean Average Precision

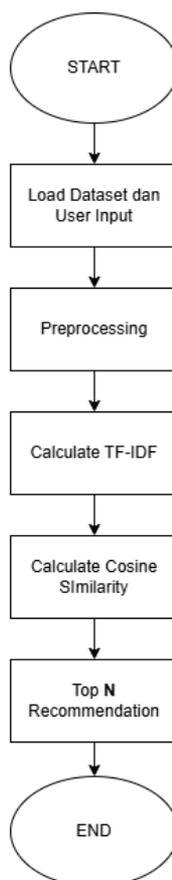
Mean Average Precision atau MAP merupakan salah satu metrik evaluasi yang digunakan untuk menilai performa sistem pencarian informasi atau dokumen. Metrik ini sangat tepat digunakan pada algoritma yang menghasilkan urutan item berdasarkan peringkat, di mana item tersebut bisa saja relevan atau tidak relevan bagi pengguna (Arfisko & Wibowo, 2022). Dalam penelitian ini, sistem menghasilkan rekomendasi berupa nilai kemiripan (*similarity*) serta nama mahasiswa. Rekomendasi tersebut kemudian dievaluasi menggunakan metode MAP.

MAP terdiri dari tiga tahap utama, yaitu menghitung *precision*, menghitung *mean precision*, dan menghitung nilai akhir MAP. *Precision* dihitung dengan membagi jumlah item yang relevan terhadap total item yang ditampilkan. Selanjutnya, *mean precision* didapat dengan menjumlahkan seluruh nilai *precision* lalu dibagi dengan jumlah item relevan. Terakhir, nilai MAP diperoleh dengan membagi total *mean precision* (MP) dengan jumlah percobaan yang dilakukan. Semakin mendekati angka 1, maka performa sistem dianggap semakin baik.

BAB IV

CONTENT-BASED FILTERING

4.1 Desain Metode Content-Based Filtering



Gambar 4 1. Flowchart Metode Content-based Filtering (Mathew *et al.*, 2016)

Desain metode Content-Based Filtering (CBF) yang ditampilkan dalam diagram ini menggambarkan alur sistem rekomendasi komunitas akademik berbasis konten, dimulai dari dua sumber data utama: Dataset Komunitas Akademik dan Dataset Mahasiswa. Kedua dataset ini menjadi input utama dalam proses sistem rekomendasi. Langkah pertama adalah tahapan preprocessing, di mana data dari kedua dataset dibersihkan dan diproses agar memiliki format teks yang seragam.

Proses ini meliputi normalisasi teks, penghapusan stopwords, stemming, dan proses lain yang umum dalam Natural Language Processing (NLP).

Setelah preprocessing, data yang telah dibersihkan kemudian diolah menggunakan metode TF-IDF (Term Frequency - Inverse Document Frequency). Metode ini digunakan untuk mengekstrak fitur dari teks dan mengubahnya menjadi representasi vektor numerik yang bisa dihitung kesamaannya. Dengan representasi TF-IDF ini, sistem kemudian melakukan perhitungan Cosine Similarity antara vektor mahasiswa dan komunitas akademik, untuk mengetahui seberapa relevan suatu komunitas dengan profil mahasiswa.

Hasil dari proses perhitungan similarity ini kemudian digunakan untuk menghasilkan Dataset Result, yaitu daftar komunitas yang direkomendasikan untuk masing-masing mahasiswa berdasarkan nilai kesamaan tertinggi. Terakhir, sistem mengevaluasi performa rekomendasi dengan menghitung nilai MAP (Mean Average Precision), yang merepresentasikan akurasi sistem dalam memberikan rekomendasi yang relevan. Dengan pendekatan ini, sistem dapat secara otomatis merekomendasikan komunitas yang paling sesuai dengan karakteristik masing-masing mahasiswa, berdasarkan konten deskriptif yang ada pada kedua dataset.

4.1.1 Pengumpulan Data

Data didapatkan dari survey secara online kepada Mahasiswa Teknik Informatika UIN Malang melalui Google Form dan mengambil dari website Teknik Informatika UIN Malang untuk komunitas akademik. Penyebaran survey dilakukan mulai tanggal 05 Februari hingga 13 Februari 2024. Dari rentan waktu tersebut

didapatkan sebanyak 200 Mahasiswa dengan minat dan data passion mereka pada komunitas akademik khususnya di Teknik Informatika UIN Malang. Pengambilan data komunitas diambil langsung dari website jurusan yang mana didapatkan data 10 komunitas akademik. Kedua data tersebut akan digunakan sebagai data utama penelitian ini. Data mahasiswa dan komunitas akan ditampilkan pada Lampiran 1 dan 2.

4.1.2 Preprocessing Data

Tahap preprocessing merupakan langkah awal yang sangat krusial dalam sistem rekomendasi berbasis konten. Pada tahap ini, data yang bersumber dari dataset mahasiswa dan komunitas akademik akan diproses agar siap untuk dianalisis. Proses preprocessing dilakukan terhadap data teks yang menjelaskan profil mahasiswa serta deskripsi komunitas. Beberapa tahapan dalam preprocessing meliputi *case folding* (mengubah semua huruf menjadi huruf kecil), tokenisasi (memecah teks menjadi kata-kata), *stopword removal* (menghapus kata-kata umum yang tidak memiliki makna penting), dan *stemming* (mengubah kata ke bentuk dasarnya). Hasil akhir dari preprocessing ini adalah data teks yang bersih dan terstandarisasi, siap untuk diekstraksi menggunakan metode TF-IDF. Visualisasi hasil *preprocessing* akan ditunjukkan pada gambar 4.2.

Before and After Processing

Original (1): Teknologi Google Pemograman Web, Pemograman Mobile, Pemograman Java, Jaringan, Linux, UI/UX Tim Besar: membuat kegiatan seperti seminar, workshop dll Java, Figma Memperluas networking

Processed (1): teknologi google pemograman web pemograman mobile pemograman java jaring linux ui ux tim giat seminar workshop dll java figma luas networking

Gambar 4 2. Visualisasi Sampel data Mahasiswa sebelum dan setelah *preprocessing*

4.1.3 Perhitungan TF-IDF

Setelah data dilakukan proses *preprocessing* dan digabung dalam kolom metadata, maka data akan dilakukan pembobotan kata ke dalam bentuk sebuah matrik dengan menggunakan algoritma TF-IDF. Proses perubahan atau transformasi data ke dalam bentuk matrik sudah disediakan fungsi-fungsinya dalam *library* 'TfidfVectorizer'.

$$W = TF \times (IDF + 1) \quad (4.1)$$

Keterangan:

TF: Jumlah kemunculan kata atau term dalam dokumen

IDF: Inverse document frequency

D: Jumlah semua dokumen

Log yang digunakan *library* TF-IDF di *sklearn* adalah logaritma natural.

Setelah melalui perhitungan tersebut didapatkan hasil seperti pada tabel 4.1.

Tabel 4. 1 Perhitungan TF-IDF

No	Term	Tf	Idf	W
1	3d	0.1249	2.9459	0.4932
2	acara	0.0494	2.9459	0.1952
3	acaraacara	0.0426	2.5404	0.1511
...
16	work	0.0641	1.3863	0.2530
17	workshop	0.1733	1.3863	0.6136
18	wujud	0.1249	0.2877	0.4932

4.1.4 Perhitungan Cosine Similarity

Tahap selanjutnya setelah didapatkan pembobotan kata yaitu perhitungan kemiripan antar konten atau data menggunakan algoritma *Cosine Similarity*. Perhitungan Tingkat kemiripan dihitung berdasarkan nilai *cosine* yang didapatkan dari perbandingan perkalian antar matriks.

Untuk menghitung nilai *cosine* dari seluruh matrik yang sebelumnya di proses pada tfidf. Nilai *cosine* dari perhitungan tersebut akan dijadikan acuan untuk

menentukan tingkat kemiripan antar dokumen yang diinputkan dengan dokumen lainnya. Beberapa dokumen yang memiliki tingkat kemiripan yang paling tinggi akan dijadikan dokumen yang direkomendasikan kepada user. Hasil perhitungan skor *cosine similarity* yang dihasilkan dari perbandingan dokumen x dan dokumen y ditunjukkan pada tabel 4.2.

Tabel 4. 2. Perhitungan Cosine Similarity

Dokumen	0	1	2	3	4	5
0	0.0678	0.075	0	0.0415	0.0786	0.0678
1	0.0316	0.0224	0.1644	0.0296	0.0091	0.0687
2	0	0.0182	0.0306	0.0196	0.1144	0
3	0.0573	0.1136	0.1011	0.0925	0.0748	0.0903
4	0.055	0.0639	0.1335	0.0611	0.0334	0.2003

4.2 Implementasi Sistem

Setelah semua data dihitung bobot dan nilai kemiripannya dengan menggunakan TF-IDF dan *cosine similarity*, untuk mendapatkan hasil rekomendasi user harus melakukan input beberapa kriteria/minat komunitas akademik yang sudah disediakan di sistem berupa form. Sebagai sampel penulis menginputkan data seperti tabel 4.3.

Tabel 4. 3. Contoh Input Minat Komuntas Akademik

Kriteria/Minat	Hasil input
Passion	Data Science
Pengetahuan Sebelumnya	Data Science
Preferensi Tim	Tim Besar
Minat Skill	Python
Motivasi	Ingin mempelajari lebih dalam data science

Dari inputan yang dilakukan user tersebut, data input akan melalui beberapa langkah sebelum mendapatkan rekomendasi yaitu *preprocessing*, *tf-idf* dan *cosine similarity*.

4.2.1 Preprocessing

Pada tahap ini data yang ada akan dilakukan pra-pemrosesan agar data yang diproses oleh sistem sudah bersih dan siap digunakan. Tabel akan memberikan hasil sebelum dan sesudah data di preprocessing.

Tabel 4. 4. Data Sebelum dan sesudah Preprocessing

Sebelum Preprocessing Data	Sesudah Preprocessing Data
data science data science tim besar pyhton Ingin mempelajari lebih dalam data science	data science data science tim pyhton ajar data science

4.2.2 Metadata

Setelah dilakukan tahapan preprocessing maka seluruh data akan digabung menjadi sebuah meta data yang nantinya akan dilakukan proses berikutnya untuk diberi bobot menggunakan *tf-idf*.

Tabel 4. 5. Meta Data

Meta Data
data science data science tim pyhton ajar data science

4.2.3 TF-IDF

Setelah data dilakukan proses *preprocessing* dan digabung dalam kolom metadata, maka data akan dilakukan pembobotan kata ke dalam bentuk sebuah matrik dengan menggunakan algoritma TF-IDF. Proses pengubahan atau transformasi data ke dalam bentuk matrik sudah disediakan fungsi-fungsinya dalam

library 'TfidfVecorizer'. Tabel akan menunjukkan hasil *tf-idf* dari metadata yang sudah ada.

Tabel 4. 6. Hasil TF-IDF Metdata

No	Term	Tf	Idf	W
1	science	0.68927	2.386294	2.33407
2	data	0.606174	2.098612	1.878298
3	tim	0.268795	2.791759	1.019208
4	pyhton	0.268795	2.791759	1.019208
5	ajar	0.113836	1.182322	0.248427

4.2.4 Cosine Similarity

Setelah dilakukan tahap pembobotan fitur menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF), langkah selanjutnya adalah menghitung tingkat kemiripan antara mahasiswa dan komunitas akademik menggunakan algoritma Cosine Similarity. Algoritma ini digunakan untuk mengukur kesamaan antar dua vektor berdasarkan sudut kosinus di antara keduanya, di mana semakin kecil sudutnya maka semakin besar tingkat kesamaan yang dimiliki. Nilai cosine similarity berkisar antara 0 hingga 1, dengan nilai yang lebih tinggi menunjukkan kemiripan yang lebih besar antara vektor fitur mahasiswa dan komunitas.

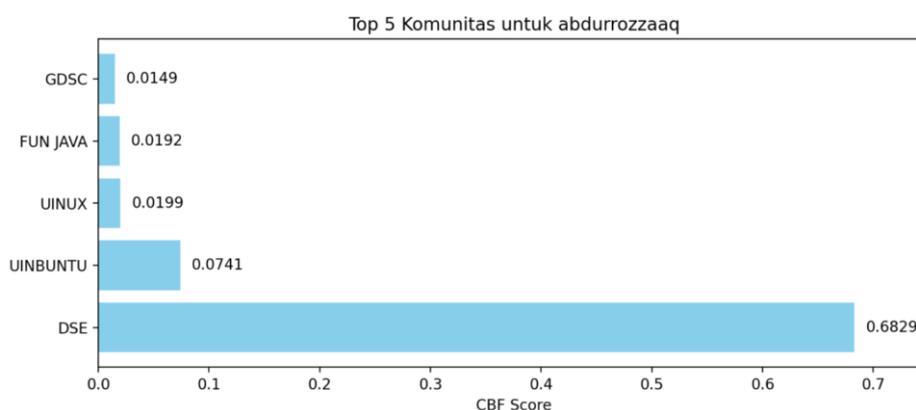
Tabel di bawah ini menyajikan hasil perhitungan cosine similarity antara seorang mahasiswa (dalam hal ini abdurrozaaq) terhadap sepuluh komunitas akademik yang tersedia. Komunitas dengan nilai kemiripan tertinggi adalah DSE dengan skor 0.6828, yang berarti komunitas tersebut memiliki karakteristik paling relevan dengan preferensi atau profil mahasiswa yang bersangkutan. Diikuti oleh komunitas UINBUNTU (0.0741), UINUX (0.0199), FUN JAVA (0.0192), dan

GDSC (0.0148). Adapun komunitas seperti MOCAP dan ETH0 memperoleh nilai 0, yang mengindikasikan tidak adanya kemiripan fitur antara komunitas tersebut dengan mahasiswa dalam kasus ini.

Tabel 4. 7. Hasil Cosine Similarity

No	Komunitas	Nilai Cosine Similarity
1	DSE	0.6828
2	UINBUNTU	0.0741
3	UINUX	0.0199
4	FUN JAVA	0.0192
5	GDSC	0.0148
6	MAMUD	0.0115
7	ONTAKI	0.0063
8	WEBONDER	0.002
9	MOCAP	0
10	ETH0	0

Visualisasi hasil rekomendasi dalam bentuk diagram batang ditunjukkan pada gambar memperkuat interpretasi dari hasil tabel, di mana terlihat bahwa komunitas DSE memiliki skor tertinggi secara signifikan dibandingkan komunitas lainnya. Hasil ini menjadi dasar dalam sistem rekomendasi berbasis Content-Based Filtering untuk menyarankan komunitas yang paling relevan kepada mahasiswa berdasarkan kesesuaian fitur konten yang dimiliki masing-masing entitas.



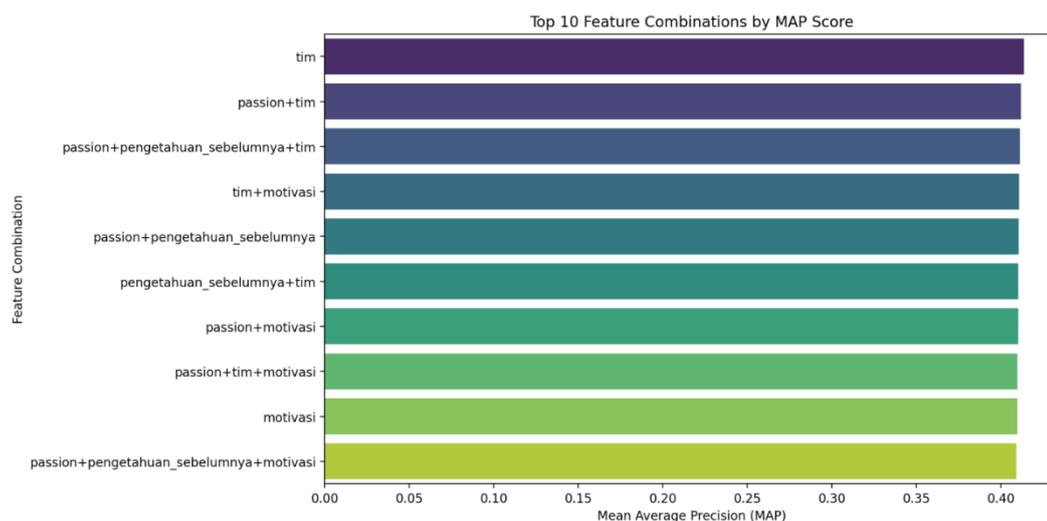
Gambar 4 3. Diagram Hasil Cosine Similarity Teratas

4.3 Uji Coba

Uji coba dilakukan untuk mencoba semua skenario yang sudah dirancang. Pada proses selanjutnya akan di terangkan lebih jelas.

4.3.1 Skenario 1

Pada skenario pertama, fokus pengujian adalah untuk menemukan kombinasi fitur terbaik dari data mahasiswa yang digunakan dalam sistem rekomendasi. Tabel akan menunjukkan hasil kombinasi fitur terbaik yang bisa digunakan dalam proses metode *content-based filtering*.



Gambar 4 4. Diragram Hasil MAP Skenario 1

Gambar menyajikan sepuluh kombinasi fitur teratas berdasarkan nilai Mean Average Precision (MAP) dalam evaluasi sistem rekomendasi komunitas akademik berbasis content-based filtering. Nilai MAP digunakan sebagai indikator utama untuk mengukur keakuratan sistem dalam merekomendasikan komunitas yang relevan sesuai dengan preferensi pengguna. Hasil menunjukkan bahwa fitur tunggal "tim" menghasilkan skor MAP tertinggi sebesar 0.4137, mengindikasikan bahwa

aspek keterlibatan dalam kerja tim memiliki korelasi yang sangat kuat terhadap relevansi komunitas yang direkomendasikan.

Kombinasi fitur "passion+tim" dan "passion+pengetahuan_sebelumnya+tim" berada pada peringkat kedua dan ketiga, dengan skor MAP masing-masing sebesar 0.4119 dan 0.4113. Meskipun melibatkan lebih banyak dimensi informasi, peningkatan akurasi yang diperoleh dibandingkan fitur "tim" tunggal relatif kecil, yang menunjukkan bahwa keberadaan fitur "tim" berkontribusi paling signifikan dalam model rekomendasi ini.

Rentang nilai MAP dari keseluruhan kombinasi fitur yang ditampilkan berkisar antara 0.4093 hingga 0.4137, dengan deviasi yang sangat kecil, menandakan kestabilan performa sistem terhadap variasi fitur masukan. Meskipun fitur-fitur seperti "motivasi", "passion", dan "pengetahuan_sebelumnya" juga memberikan kontribusi terhadap peningkatan akurasi, pengaruhnya tidak melebihi kontribusi dominan dari fitur "tim".

Secara keseluruhan, temuan ini mengindikasikan bahwa preferensi terhadap kerja tim merupakan prediktor paling informatif dalam memodelkan relevansi komunitas akademik yang direkomendasikan, dan perlu mendapatkan perhatian utama dalam desain serta pengembangan sistem rekomendasi berbasis konten di ranah pendidikan tinggi.

4.3.2 Skenario 2

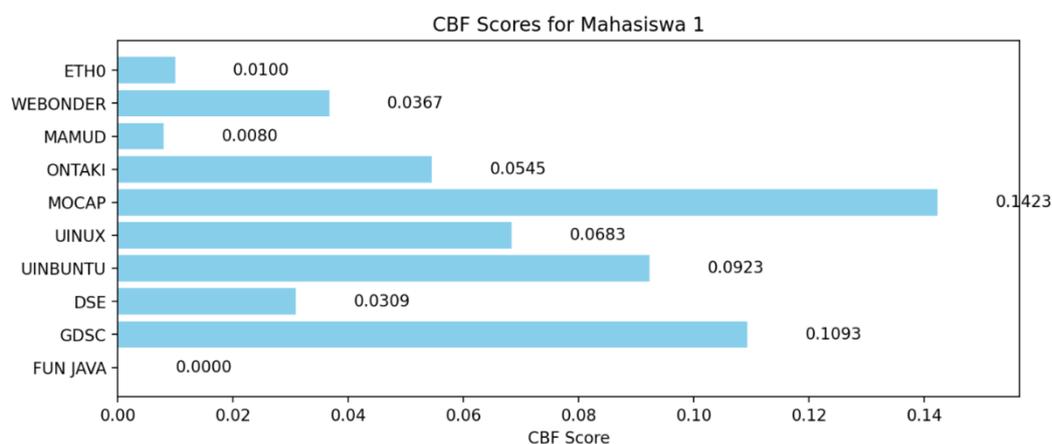
Skenario kedua bertujuan untuk menguji performa sistem dalam kondisi ideal, di mana data input sangat sesuai dengan konteks bidang Teknik Informatika.

Langkah pertama dilakukan dengan mengubah data mahasiswa menjadi bentuk esai naratif yang menggambarkan secara menyeluruh profil mahasiswa, dengan isi yang memiliki kesesuaian 100% terhadap atribut yang umum dimiliki oleh mahasiswa Teknik Informatika. Tabel akan menunjukkan metadata yang digunakan sampel untuk skenario ini.

Tabel 4. 8. Metadata Skenario 2

Meta Data
<p> milik minat kembang perangkat lunak tantang logika kreativitas butuh alam cakup kembang aplikasi python javascript framework react flask biasa tim fokus kembang aplikasi mobile bas flutter milik terampil tools docker postman figma desain ui kembang terampil teknis kontribusi proyek proyek teknologi dampak </p>

Visualisasi akan menunjukkan hasil dari 100% kesesuaian dalam atribut data mahasiswa terhadap Teknik informatika.



Gambar 4 5. Bar Chart hasil cosine similarity dari skenario 2

4.3.3 Skenario 3

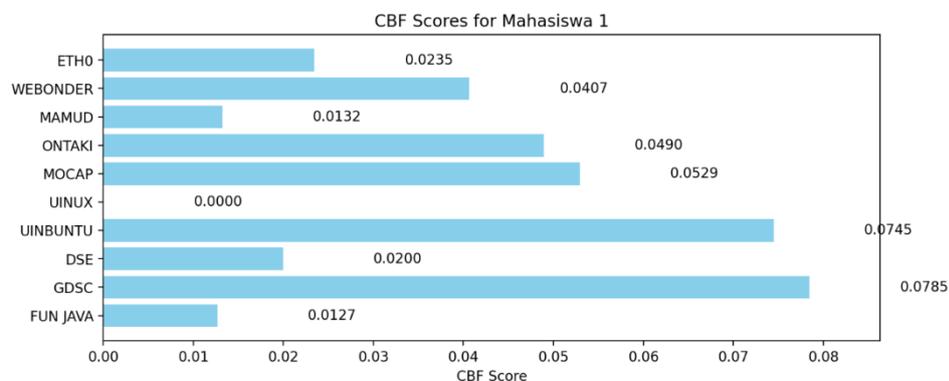
Pada skenario ketiga, dilakukan pengujian dengan pendekatan yang sama seperti skenario kedua, namun dengan kondisi input yang tidak ideal. Data mahasiswa juga diubah menjadi bentuk esai, tetapi hanya mencerminkan 20%

kesesuaian terhadap profil mahasiswa Teknik Informatika. Tabel akan menunjukkan metadata yang digunakan sampel untuk skenario ini.

Tabel 4. 9. Metadata Skenario 3

Meta Data
ikut kembang teknologi cloud computing fokus kolaborasi tim kembang software tarik paham basis data bantu bangun aplikasi teknik informatika pikir logis struktur kreativitas aspek hidup kolaborasi tim kembang software tarik percaya empati kunci bangun relasi keingintahuan dorong coba tarik dalam kait kembang kolaborasi tim kembang software tarik suka ikut seminar latih tambah terampil nikmat giat tuntutan kerja komunikasi kolaborasi tim kembang software tarik luang senang baca buku motivasi kolaborasi tim kembang software tarik hadap tantang semangat tarik dalam kait kembang adaptasi ubah salah kuat adaptasi ubah salah kuat tulis jurnal pribadi bantu refleksi alam suka tim dinamis kreatif atur prioritas utama selesai tugas suka tantang debugging proses coding lingkungan kerja positif produktivitas tulis jurnal pribadi bantu refleksi alam

Visualisasi akan menunjukkan hasil dari performa 20% kesesuaian dalam atribut data mahasiswa terhadap Teknik informatika.



Gambar 4 6. Bar Chart hasil cosine similarity dari skenario 2

4.4 Pengujian Sistem MAP

Pengujian sistem dilakukan untuk mengevaluasi seberapa baik performa sistem rekomendasi dalam memberikan hasil yang relevan. Pengujian dilakukan berdasarkan skenario-skenario yang telah dirancang sebelumnya, dengan menggunakan metode evaluasi Mean Average Precision (MAP) sebagai metrik

utama. Tujuannya adalah untuk mengetahui kombinasi fitur terbaik serta seberapa akurat sistem saat diberikan input data yang ideal maupun yang tidak relevan. Pengujian ini tidak hanya melihat hasil akhir, tetapi juga mencermati stabilitas sistem dalam berbagai kondisi input.

4.4.1 Skenario 1

Pada scenario 1 dilakukan evaluasi terhadap performa sistem rekomendasi dengan pendekatan content-based filtering menggunakan metrik Mean Average Precision (MAP). Pengujian ini bertujuan untuk mengukur seberapa akurat sistem dalam merekomendasikan komunitas yang relevan bagi pengguna. Nilai MAP Content-based filtering ini di dapat dari 5 data mahasiswa aktual sebagai sampel.

Tabel 4. 10. Hasil Rekomendasi CBF dengan Data Aktual

Data	Nilai Similarity				
	<i>Recomm 1</i>	<i>Recomm 2</i>	<i>Recomm 3</i>	<i>Recomm 4</i>	<i>Recomm 5</i>
Dokumen 1	0.2654	0.1363	0.0762	0.0666	0.0508
Dokumen 2	0.0815	0.0744	0.054	0.0466	0.0457
Dokumen 3	0.1325	0.1272	0.0766	0.0707	0.0655
Dokumen 4	0.0903	0.0814	0.0771	0.067	0.0639
Dokumen 5	0.0778	0.0564	0.0441	0.026	0.014

Tabel 4. 11. Perhitungan MAP Content-based Filtering

Data	Nilai Similarity					AP
	<i>Recomm 1</i>	<i>Recomm 2</i>	<i>Recomm 3</i>	<i>Recomm 4</i>	<i>Recomm 5</i>	
Dokumen 1	1	0	0	0	0	1
Dokumen 2	0	0	0	0,25	0	0,25
Dokumen 3	1	1	1	0	0	1
Dokumen 4	1	0	0	0	0	1
Dokumen 5	0	0	0,333333333	0,5	0	0,415
Total						3,665
$MAP = \frac{\text{total nilai AP}}{\text{total dokumen}} \times 100\%$						73,3 %

Tabel 4.11 menunjukkan hasil perhitungan Mean Average Precision (MAP) dari sistem rekomendasi berbasis *Content-based Filtering* (CBF). Pada eksperimen ini, setiap dokumen atau pengguna diberikan lima rekomendasi komunitas, dan

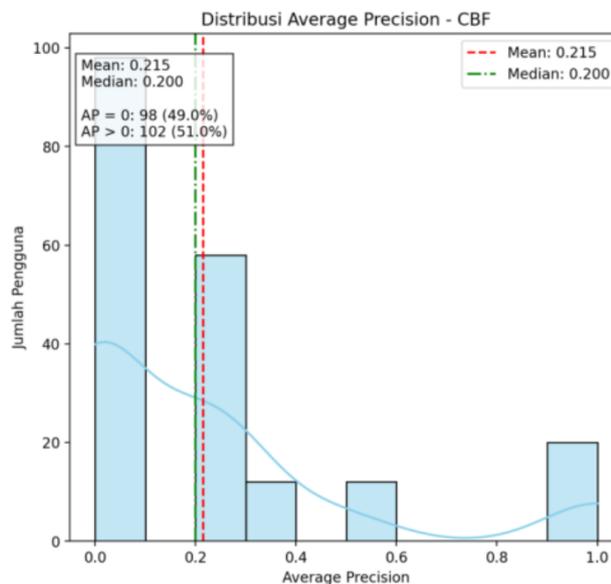
nilai Average Precision (AP) dihitung berdasarkan relevansi hasil rekomendasi. Perhitungan AP dilakukan dengan mengamati posisi rekomendasi yang relevan dan menghitung rasio akurasi kumulatif hingga posisi tersebut. Sebagai contoh, pada Dokumen 1, rekomendasi yang relevan berada pada posisi pertama, sehingga memperoleh nilai AP sebesar 1. Sedangkan pada Dokumen 2, rekomendasi yang relevan muncul pada posisi keempat, sehingga nilai AP-nya hanya sebesar 0,25.

Hasil perhitungan menunjukkan bahwa total nilai AP dari lima dokumen adalah sebesar 3,665. Nilai ini kemudian dirata-ratakan dengan membaginya terhadap jumlah dokumen, sehingga diperoleh nilai MAP sebesar 0,733 atau 73,3%. Nilai ini mengindikasikan bahwa metode *Content-based Filtering* memiliki tingkat akurasi yang cukup baik dalam menghasilkan rekomendasi yang relevan.

4.4.2 Skenario 2

Distribusi Average Precision untuk metode Content-Based Filtering (CBF) menunjukkan polarisasi kinerja yang signifikan. Dengan nilai Mean Average Precision 0,215 dan median 0,200, data terdistribusi bimodal dengan konsentrasi pada nilai ekstrem. Sebanyak 49,0% mahasiswa (98 pengguna) mendapatkan AP = 0, mengindikasikan tidak adanya rekomendasi relevan, sementara 51,0% lainnya (102 mahasiswa) menerima rekomendasi dengan tingkat relevansi bervariasi. Adanya peningkatan frekuensi pada nilai AP = 1,0 (sekitar 20 mahasiswa) menunjukkan bahwa CBF sangat efektif untuk sebagian kecil pengguna yang memiliki keselarasan sempurna antara profil dan item yang direkomendasikan, namun kurang efektif untuk hampir separuh pengguna lainnya, menggarisbawahi

sensitivitas metode ini terhadap kualitas representasi profil pengguna dan karakteristik item. Nilai-nilai ini dipengaruhi oleh ketidaksesuaian antara data preferensi pengguna yang diinput dengan ground truth yang digunakan, mengingat sistem hanya melakukan pencocokan berbasis kesesuaian konten antara pengguna dan komunitas. Oleh karena itu, keputusan akhir dalam mengikuti komunitas tetap menjadi hak prerogatif pengguna sepenuhnya.

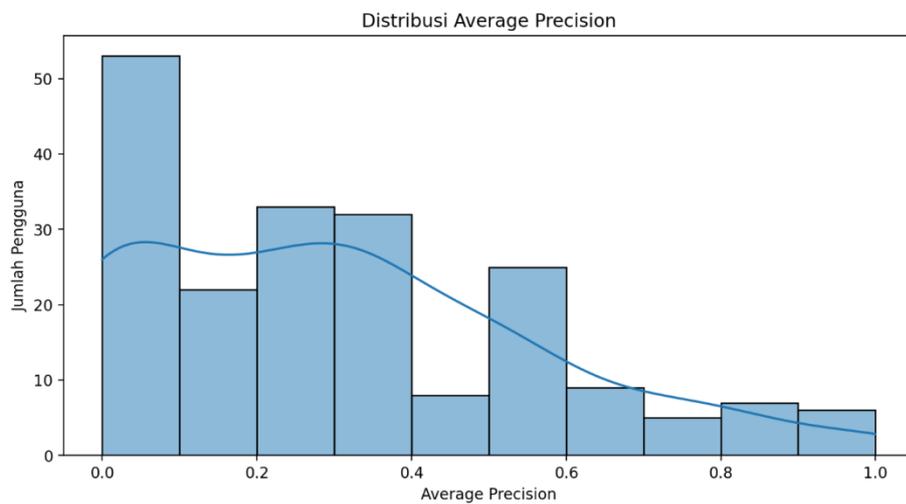


Gambar 4.7. Diagram Hasil Average Precision

4.4.3 Skenario 3

Gambar menunjukkan distribusi nilai Average Precision dari hasil rekomendasi sistem terhadap seluruh pengguna. Dari visualisasi histogram, terlihat bahwa sebagian besar pengguna memiliki nilai Average Precision yang rendah, dengan puncak distribusi berada pada rentang 0.0 hingga 0.1. Hal ini menunjukkan bahwa rekomendasi yang diberikan belum sepenuhnya relevan bagi mayoritas pengguna. Nilai Mean Average Precision (MAP) sebesar 0.3031 mengindikasikan

bahwa sistem memiliki tingkat akurasi yang sedang dalam merekomendasikan komunitas yang relevan, di mana semakin mendekati nilai 1 maka semakin baik akurasi. Salah satu faktor penyebab rendahnya nilai ini adalah ketidaksesuaian antara data yang diinput oleh pengguna dan ground truth, yaitu komunitas yang benar-benar diikuti, sehingga menurunkan relevansi hasil rekomendasi yang dihasilkan oleh sistem.

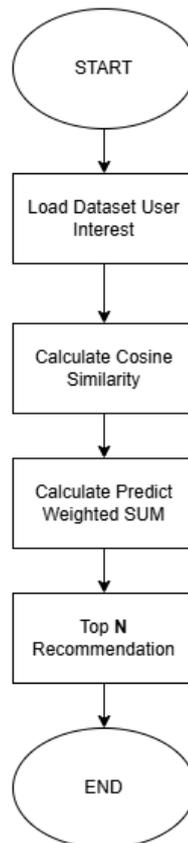


Gambar 4 8. Diagram Hasil Average Precision

BAB V

COLLABORATIVE FILTERING

5.1 Desain Metode Collaborative Filtering



Gambar 5. 1. Flowchart Metode Collaborative Filtering (Valentino & Setiawan, 2024)

Flowchart ini menggambarkan alur proses dalam sistem rekomendasi berbasis Content-Based Filtering (CBF). Proses diawali dengan memuat data minat pengguna yang berasal dari atribut mahasiswa, seperti jurusan, minat, atau keikutsertaan dalam kegiatan tertentu. Selanjutnya dilakukan perhitungan cosine similarity untuk mengukur tingkat kemiripan antara profil mahasiswa dengan masing-masing komunitas. Nilai kemiripan ini menjadi dasar dalam menentukan seberapa relevan suatu komunitas bagi mahasiswa tersebut.

Tahap berikutnya adalah perhitungan skor prediksi menggunakan metode weighted sum, di mana setiap komunitas diberi bobot berdasarkan nilai kemiripannya. Dari hasil skor tersebut, sistem akan memilih beberapa komunitas terbaik dengan nilai tertinggi yang kemudian ditampilkan sebagai Top-N Recommendation. Proses ini menghasilkan rekomendasi yang bersifat personal dan relevan dengan karakteristik pengguna. Setelah itu, proses berakhir dan sistem siap memberikan hasil rekomendasi kepada pengguna.

5.1.1 Pengumpulan Data

Data didapatkan dari survey secara online kepada Mahasiswa Teknik Informatika UIN Malang melalui Google Form dan mengambil dari website Teknik Informatika UIN Malang untuk komunitas akademik. Penyebaran survey dilakukan mulai tanggal 05 Februari hingga 13 Februari 2024. Dari rentan waktu tersebut didapatkan sebanyak 58 Mahasiswa dengan ketertarikan mereka pada komunitas akademik dalam bentuk rating khususnya di Teknik Informatika UIN Malang. Data tersebut akan digunakan sebagai data utama penelitian ini. Contoh data yang akan digunakan pada metode collaborative filtering akan ditunjukkan pada tabel.

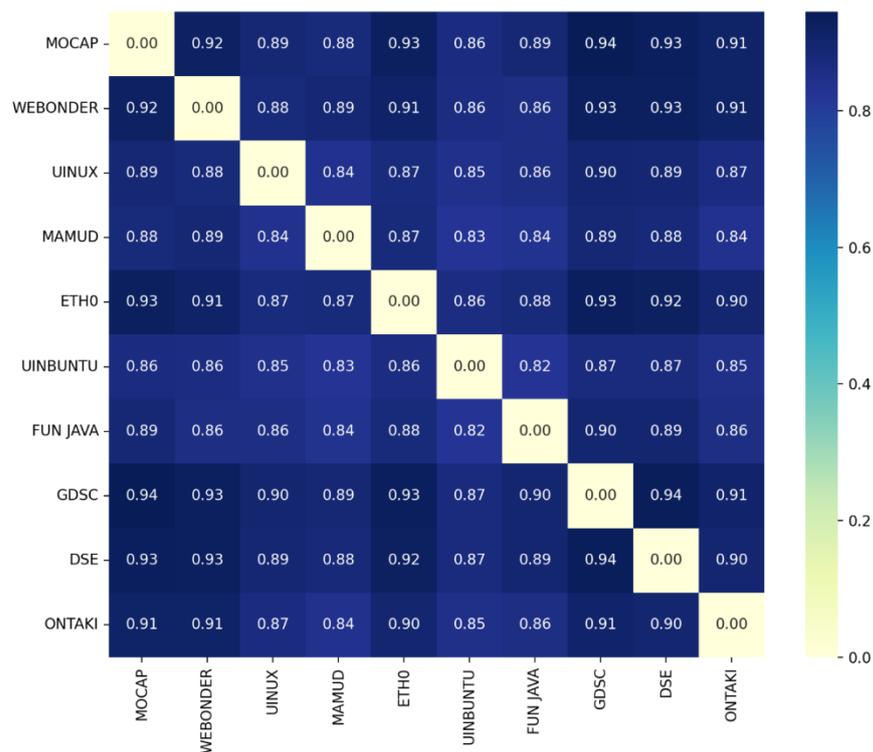
Tabel 5. 1. Sampel 5 Data mahasiswa dengan rating komunitasnya

MHS	Rating Komunitas									
	MCP	WB	UIX	MM	ETH	UBT	FJ	GDSC	DSE	ONT
1	3	5	2	5	2	3	1	5	4	4
2	2	3	4	1	4	3	2	5	4	4
3	3	4	2	3	3	5	1	4	3	3

MHS	Rating Komunitas									
	MCP	WB	UIX	MM	ETH	UBT	FJ	GDSC	DSE	ONT
4	3	2	2	5	3	5	4	4	4	4
5	4	4	2	3	5	5	2	4	4	2

5.1.2 Perhitungan Item Similarity

Pada tahap ini dijelaskan proses perhitungan kemiripan antar komunitas akademik yang digunakan dalam sistem rekomendasi. *Item similarity* menjadi dasar untuk mengetahui hubungan antar komunitas berdasarkan data interaksi mahasiswa, seperti rating mahasiswa terhadap komunitas. Sistem menghitung kemiripan komunitas dengan menggunakan metode *cosine similarity*, sehingga komunitas yang sering dipilih atau diminati oleh mahasiswa yang sama akan memiliki nilai kemiripan yang tinggi. Hasil perhitungan ini akan menjadi acuan dalam merekomendasikan komunitas yang relevan kepada mahasiswa.



Gambar 5. 2. Gambar Hitmap dari Similarity Item

Gambar 5.2 menunjukkan matriks item similarity yang merepresentasikan tingkat kemiripan antar komunitas akademik berdasarkan interaksi mahasiswa. Setiap sel dalam matriks ini berisi nilai antara 0 hingga 1, yang menunjukkan seberapa mirip dua komunitas satu sama lain—semakin mendekati 1, semakin tinggi kemiripannya. Contohnya, komunitas WEBONDER memiliki nilai kemiripan sebesar 0.92 dengan MOCAP, yang berarti kedua komunitas ini sering memiliki anggota mahasiswa yang sama atau memiliki pola ketertarikan yang serupa. Matriks ini akan digunakan dalam metode Collaborative Filtering (CF) sebagai dasar dalam proses prediksi rating, di mana sistem akan memanfaatkan nilai-nilai kemiripan ini untuk memperkirakan seberapa besar kemungkinan seorang mahasiswa akan tertarik pada komunitas yang belum pernah ia ikuti,

berdasarkan komunitas lain yang sudah ia minati dan kemiripan antar komunitas tersebut.

5.1.3 Perhitungan Weighted SUM

Tahap ini memaparkan bagaimana sistem menggunakan hasil kemiripan antar komunitas untuk memprediksi minat mahasiswa terhadap komunitas yang belum mereka ikuti. Dengan metode *weighted sum*, sistem mengalikan nilai kemiripan komunitas dengan rating mahasiswa terhadap komunitas serupa, kemudian menjumlahkan nilai tersebut untuk mendapatkan skor rekomendasi. Proses ini memungkinkan sistem memberikan rekomendasi komunitas yang paling sesuai dan berpotensi diminati mahasiswa berdasarkan interaksi sebelumnya. Contoh Prediksi rating 4 mahasiswa terhadap 10 komunitas akademik.

Tabel 5. 2. Tabel Hasil Prediksi Rating

Komunitas	Predict Ratings Mhs 1	Predict Ratings Mhs 2	Predict Ratings Mhs 3	Predict Ratings Mhs 4
FUN JAVA	1	2	1	4
GDSC	5	5	4	4
DSE	4	4	3	4
UINBUNTU	3	3	5	5
UNIX	2	4	2	2
MOCAP	3	2	3	3
ONTAKI	4	4	3	4
MAMUD	5	1	3	5
WEBONDER	5	3	4	2
ETH0	2	4	3	3

5.2 Implementasi Sistem

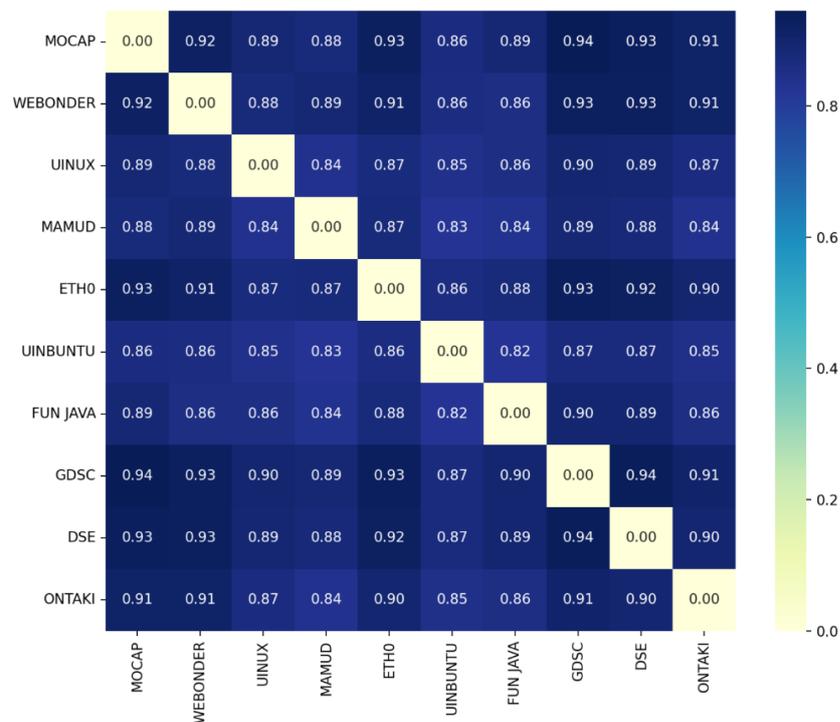
Dalam tahapan ini akan dilakukan percobaan menggunakan input manual untuk data mahasiswa dan rating mahasiswa terhadap komunitas. Untuk menjalankannya data mahasiswa diambil dari tabel. Sedangkan untuk ratingnya akan ditunjukkan oleh tabel.

Tabel 5. 3. Sampel Data Rating Mahasiswa

Komunitas	Rating
FUN JAVA	1
GDSC	2
DSE	4
UINBUNTU	0
UINUX	0
MOCAP	0
ONTAKI	0
MAMUD	0
WEBONDER	1
ETH0	0

5.2.1 Perhitungan Item Similarity

Sistem menghitung kemiripan komunitas dengan menggunakan metode *cosine similarity*, sehingga komunitas yang sering dipilih atau diminati oleh mahasiswa yang sama akan memiliki nilai kemiripan yang tinggi. Hasil perhitungan ini akan menjadi acuan dalam merekomendasikan komunitas yang relevan kepada mahasiswa.

**Gambar 5. 3.** Hitmap Similarity antar Item

5.2.2 Perhitungan Weighted SUM

Tabel 5. 4. Hasil Prediksi Rating

Komunitas	Predict Ratings
FUN JAVA	1
GDSC	2
DSE	4
UINBUNTU	2
UINUX	2
MOCAP	2
ONTAKI	2
MAMUD	2
WEBONDER	1
ETHO	2

Tabel menampilkan hasil prediksi rating dari sistem rekomendasi berbasis *Collaborative Filtering* (CF) untuk mahasiswa bernama abdurrozzaaq terhadap sejumlah komunitas akademik. Dalam tabel ini, rating dengan nilai bulat (seperti 1.0000, 2.0000, atau 4.0000) merupakan rating yang sudah diberikan langsung oleh mahasiswa, sedangkan nilai desimal (seperti 2.0131, 2.0121, dan seterusnya) merupakan hasil prediksi dari sistem CF.

Prediksi ini dihitung berdasarkan kemiripan antar komunitas (*item similarity*) dan riwayat interaksi mahasiswa dengan komunitas lainnya. Misalnya, mahasiswa abdurrozzaaq belum pernah memberi rating pada komunitas UINUBUNTU, namun sistem memprediksi bahwa ia kemungkinan akan memberi rating sebesar 2.0131 berdasarkan preferensinya terhadap komunitas lain yang mirip dengan UINUBUNTU. Prediksi tertinggi di antara komunitas yang belum dinilai mahasiswa adalah ETHO (2.0159) dan MOCAP (2.0135), sehingga komunitas-komunitas ini dapat direkomendasikan sebagai komunitas yang kemungkinan besar

akan diminati oleh mahasiswa tersebut. Dengan pendekatan ini, sistem dapat memberikan rekomendasi yang lebih personal dan relevan bagi setiap mahasiswa.

5.3 Pengujian System MAP

Pengujian sistem dilakukan untuk mengevaluasi seberapa baik performa sistem rekomendasi dalam memberikan hasil yang relevan. Pengujian dilakukan dengan menggunakan metode evaluasi Mean Average Precision (MAP) sebagai metrik utama. Tujuannya adalah untuk mengetahui kombinasi fitur terbaik serta seberapa akurat sistem saat diberikan input data yang ideal maupun yang tidak relevan. Pengujian ini tidak hanya melihat hasil akhir, tetapi juga mencermati stabilitas sistem dalam berbagai kondisi input.

Tabel 5. 5. Hasil Rekomendasi CF dengan Data Aktual

Data	Nilai Similarity				
	<i>Recomm 1</i>	<i>Recomm 2</i>	<i>Recomm 3</i>	<i>Recomm 4</i>	<i>Recomm 5</i>
Dokumen 1	0.5077	0.4432	0.4131	0.4083	0.4004
Dokumen 2	0.3970	0.3464	0.3403	0.3332	0.3085
Dokumen 3	0.5000	0.5000	0.3769	0.3764	0.3780
Dokumen 4	0.75	0.75	0.5839	0.5831	0.5825
Dokumen 5	0.7500	0.5019	0.5014	0.5047	0.5034

Tabel 5. 6. Perhitungan MAP Collaborative Filtering

Data	Nilai Similarity					AP
	<i>Recomm 1</i>	<i>Recomm 2</i>	<i>Recomm 3</i>	<i>Recomm 4</i>	<i>Recomm 5</i>	
Dokumen 1	1	0	0	0	0	1
Dokumen 2	0	0	0,33	0	0	0,33
Dokumen 3	1	1	0	0,75	0	0,75
Dokumen 4	0	1	1	0	0	1
Dokumen 5	1	0	0	0	0	1
Total						4,08
$MAP = \frac{\text{total nilai AP}}{\text{total dokumen}} \times 100\%$						81,6%

Tabel 5.6 menyajikan hasil perhitungan Mean Average Precision (MAP) dari sistem rekomendasi menggunakan pendekatan Collaborative Filtering (CF). Dalam

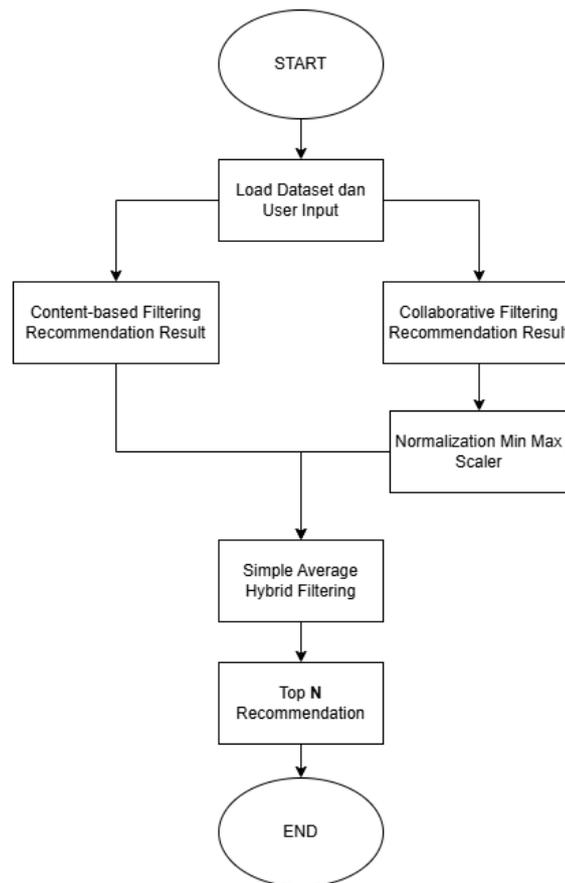
eksperimen ini, setiap dokumen memperoleh lima rekomendasi, dan relevansi masing-masing rekomendasi dinilai berdasarkan kemunculan nilai 1 (relevan) pada posisi tertentu. Nilai Average Precision (AP) dihitung untuk masing-masing dokumen dengan mempertimbangkan urutan kemunculan rekomendasi yang relevan. Sebagai contoh, Dokumen 1 dan Dokumen 4 memperoleh nilai AP sempurna, yaitu 1, karena seluruh rekomendasi relevan muncul di posisi awal. Sementara itu, Dokumen 2 dan Dokumen 3 memperoleh nilai AP masing-masing 0,33 dan 0,75 karena rekomendasi relevan muncul di posisi yang lebih rendah.

Secara keseluruhan, total nilai AP dari kelima dokumen adalah 4,08. Nilai ini kemudian dirata-ratakan dengan membaginya terhadap jumlah dokumen, sehingga diperoleh nilai MAP sebesar 0,816 atau 81,6%. Nilai ini menunjukkan bahwa metode Collaborative Filtering memiliki performa yang tinggi dalam menghasilkan rekomendasi yang relevan, dan lebih unggul dibandingkan metode Content-based Filtering yang hanya memperoleh MAP sebesar 0,733. Temuan ini menunjukkan bahwa kolaborasi antar preferensi pengguna memberikan kontribusi yang signifikan dalam meningkatkan akurasi sistem rekomendasi.

BAB VI

HYBRID FILTERING

6.1 Desain Metode Hybrid Filtering



Gambar 6. 1. Flowchart Metode Hybrid Filtering (Nasser *et al.*, 2023)

Flowchart 6.1 menjelaskan alur kerja sistem rekomendasi berbasis *hybrid filtering* yang menggabungkan pendekatan *content-based filtering* dan *collaborative filtering*. Proses diawali dengan memuat dataset serta input pengguna sebagai dasar perhitungan. Data ini kemudian diproses secara terpisah oleh dua metode: *content-based filtering* yang menghitung kesesuaian berdasarkan karakteristik pengguna, serta *collaborative filtering* yang mempertimbangkan preferensi pengguna lain yang serupa. Hasil dari *collaborative filtering* kemudian

dinormalisasi menggunakan teknik *Min-Max Scaler* agar berada dalam skala yang sebanding dengan hasil *content-based filtering*.

Langkah selanjutnya adalah menggabungkan kedua hasil rekomendasi tersebut menggunakan pendekatan *simple average* dalam metode *hybrid filtering*. Tujuannya adalah memperoleh rekomendasi yang lebih seimbang dengan mempertimbangkan kekuatan dari kedua pendekatan. Setelah skor gabungan dihitung, sistem akan menentukan sejumlah rekomendasi terbaik yang disebut sebagai Top-N Recommendation. Proses ini berakhir ketika sistem menghasilkan daftar rekomendasi yang paling relevan bagi pengguna berdasarkan hasil penggabungan kedua metode tersebut.

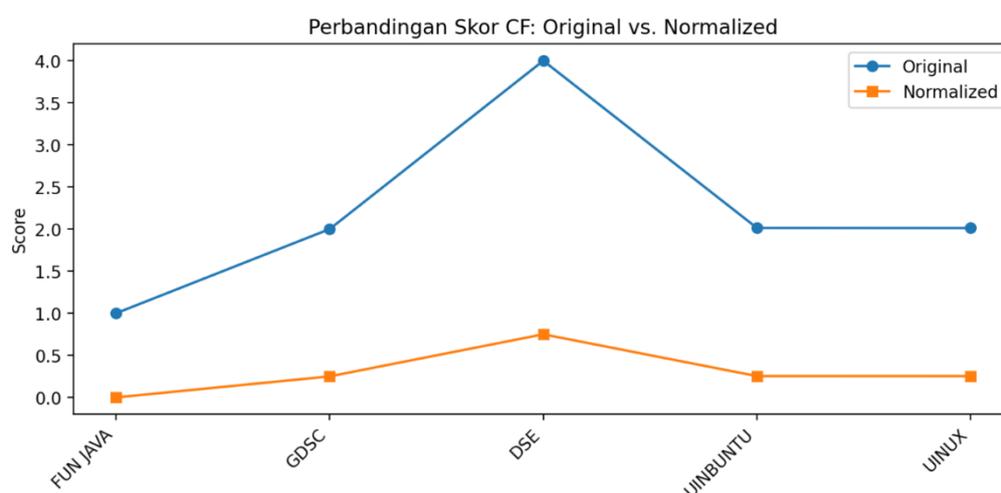
6.2 Implementasi Hybrid Filtering

Implementasi hybrid filtering dengan pendekatan simple average merupakan metode integrasi yang menggabungkan hasil rekomendasi dari dua pendekatan berbeda, yaitu content-based filtering dan collaborative filtering, dengan cara menghitung rata-rata dari skor relevansi yang dihasilkan masing-masing metode. Pendekatan ini bertujuan untuk menyeimbangkan kelebihan dan kekurangan dari kedua algoritma, sehingga mampu menghasilkan rekomendasi yang lebih akurat dan personal. Dalam prosesnya, skor dari collaborative filtering biasanya dinormalisasi terlebih dahulu agar memiliki skala yang setara dengan content-based filtering, sehingga proses penggabungan skor menjadi valid dan proporsional. Dengan metode simple average, sistem dapat menghasilkan rekomendasi yang tidak hanya mempertimbangkan kesesuaian konten, tetapi juga mempertimbangkan

preferensi pengguna lain yang serupa, sehingga memperkuat kualitas dan relevansi rekomendasi akhir.

6.2.1 Normalisasi Hasil Collaborative Filtering

Pada tahap ini membahas proses normalisasi terhadap hasil perhitungan Collaborative Filtering sebelum digabungkan dengan metode lainnya dalam pendekatan *Hybrid Filtering*. Normalisasi diperlukan karena rentang nilai yang dihasilkan oleh metode *Collaborative Filtering* dapat berbeda dengan metode *Content-Based Filtering*, sehingga diperlukan penyesuaian skala agar proses penggabungan kedua metode dapat dilakukan secara proporsional. Dalam penelitian ini, proses normalisasi dilakukan dengan menggunakan pendekatan *Min-Max Scaler*, yang mengubah nilai prediksi ke dalam rentang $[0, 1]$. Tahapan ini bertujuan untuk memastikan bahwa kontribusi dari masing-masing metode terhadap hasil akhir rekomendasi bersifat seimbang dan tidak saling mendominasi.



Gambar 6. 2. Normalisasi Data Hasil CF

Gambar 6.2 memperlihatkan hasil perbandingan skor *Collaborative Filtering* (CF) antara data asli (*Original*) dan hasil normalisasi (*Normalized*) terhadap lima entitas komunitas: FUN JAVA, GDSC, DSE, UINBUNTU, dan UINUX. Dalam konteks sistem rekomendasi, khususnya yang akan digunakan pada pendekatan *Simple Average Hybrid Filtering*, proses normalisasi ini sangat penting untuk menyetarakan skala skor dari berbagai sumber agar dapat digabungkan secara adil.

Secara ilmiah, normalisasi dilakukan untuk mengatasi perbedaan skala antar nilai skor, yang dapat menyebabkan bias terhadap nilai-nilai yang memiliki rentang lebih besar. Pada grafik tersebut, skor asli memiliki rentang dari 1 hingga 5, sedangkan skor yang telah dinormalisasi memiliki rentang lebih sempit antara 0 dan 0.75. Hal ini menunjukkan bahwa data telah ditransformasikan ke dalam skala yang seragam, menggunakan metode *min-max normalization*. Dengan melakukan normalisasi, setiap entitas komunitas memiliki kontribusi yang sebanding terhadap hasil akhir ketika digabungkan dengan teknik rekomendasi lainnya, seperti *Content-Based Filtering*, dalam kerangka *hybrid system*.

Penggunaan skor yang telah dinormalisasi dalam *Simple Average Hybrid Filtering* memungkinkan integrasi berbagai metode rekomendasi dengan cara yang objektif dan konsisten. Ini penting karena tanpa normalisasi, metode dengan skala skor yang lebih besar akan mendominasi hasil akhir, sehingga mengurangi akurasi dan keadilan sistem rekomendasi. Oleh karena itu, normalisasi menjadi langkah krusial dalam meningkatkan performa dan keandalan sistem rekomendasi berbasis *hybrid filtering*.

6.2.2 Perhitungan Skor Hybrid

Pada tahap ini menjelaskan proses penggabungan hasil dari Content-Based Filtering dan Collaborative Filtering menggunakan metode perhitungan rata-rata sederhana atau simple average. Setelah kedua hasil rekomendasi dinormalisasi ke dalam skala yang sebanding, dilakukan proses agregasi dengan cara menghitung rata-rata dari masing-masing skor rekomendasi yang dihasilkan oleh kedua metode. Pendekatan ini dipilih karena bersifat mudah diimplementasikan dan efektif dalam menggabungkan keunggulan masing-masing metode. Dengan metode rata-rata ini, sistem dapat menghasilkan skor akhir yang merepresentasikan kombinasi preferensi berdasarkan konten dan perilaku pengguna lain secara seimbang. Rumus matematika dari rata-rata skor hybrid ditunjukkan pada persamaan.

$$\text{Hybrid Score} = \frac{\text{CBF Score} + \text{CF Score}}{2} \quad (6.1)$$

Keterangan:

Hybrid Score: Skor akhir hasil gabungan dari kedua metode.

CBF Score : Skor dari metode *Content-Based Filtering*.

CF Score : Skor dari metode *Collaborative Filtering*.

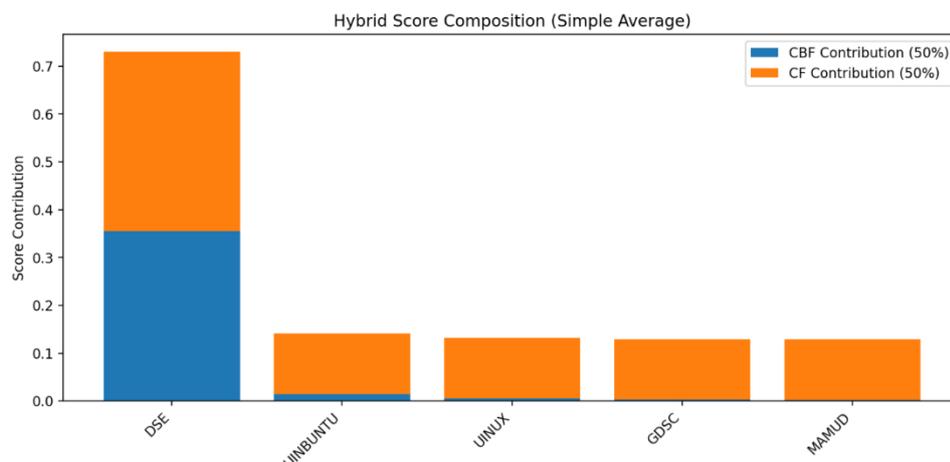
Hasil skor hybrid ditunjukkan pada tabel 6.1 menggunakan metode *Simple Average Hybrid Filtering* menunjukkan bahwa komunitas DSE memiliki skor tertinggi (0.7297), menandakan bahwa komunitas ini sangat relevan baik dari sisi kesesuaian konten (CBF Score 0.7094) maupun preferensi pengguna lain (CF Score 0.75), sehingga menjadi rekomendasi utama yang paling kuat. Sementara itu, komunitas seperti UINBUNTU, UINUX, GDSC, dan MAMUD memiliki skor hybrid yang jauh lebih rendah, berkisar antara 0.1275 hingga 0.1408, yang

menunjukkan bahwa meskipun mereka memiliki kontribusi dari skor CF (0.25), nilai CBF yang sangat rendah mengindikasikan kurangnya kesesuaian konten dengan preferensi pengguna. Dengan demikian, metode hybrid ini memberikan hasil yang seimbang dengan menggabungkan dua pendekatan, dan secara efektif mengutamakan komunitas yang unggul di kedua sisi.

Tabel 6. 1. Hasil Hybrid Filtering

Komunitas	CBF Score	CF Score	Hybrid Score
DSE	0.7094	0.75	0.7297
UINBUNTU	0.0283	0.25	0.1408
UINUX	0.0101	0.25	0.1315
GDSC	0.0072	0.25	0.1286
MAMUD	0.0050	0.25	0.1286

Grafik akan menunjukkan komposisi skor hybrid dari lima komunitas berdasarkan metode *Simple Average Hybrid Filtering*, dengan kontribusi masing-masing dari metode Content-Based Filtering (CBF) dan Collaborative Filtering (CF) yang dibagi secara merata (50%). Terlihat jelas bahwa komunitas DSE memiliki kontribusi skor tertinggi baik dari sisi CBF maupun CF, sehingga menghasilkan skor hybrid yang dominan dibandingkan komunitas lainnya. Sebaliknya, komunitas seperti UINBUNTU, UINUX, GDSC, dan MAMUD memiliki skor hybrid yang rendah, dengan kontribusi utama berasal dari CF, sedangkan CBF hampir tidak memberikan pengaruh signifikan. Hal ini mengindikasikan bahwa rekomendasi terhadap komunitas-komunitas tersebut lebih banyak didorong oleh kesamaan preferensi dengan pengguna lain, bukan karena kesesuaian konten, dan hanya DSE yang menunjukkan kekuatan dari kedua sisi.



Gambar 6. 3. Grafik Skor Hybrid

6.3 Pengujian MAP Metode Hybrid Filtering

Pada subbab ini dilakukan evaluasi terhadap performa sistem rekomendasi dengan pendekatan hybrid filtering menggunakan metrik Mean Average Precision (MAP). Pengujian ini bertujuan untuk mengukur seberapa akurat sistem dalam merekomendasikan komunitas yang relevan bagi pengguna. Metode hybrid filtering yang digunakan merupakan gabungan dari pendekatan content-based filtering (CBF) dan collaborative filtering (CF) dengan pembobotan sederhana (simple average). Nilai MAP Hybrid filtering ini di dapat dari 5 data mahasiswa aktual sebagai sampel.

Tabel 6. 2. Hasil Rekomendasi CF dengan Data Aktual

Data	Nilai Similarity				
	<i>Recomm 1</i>	<i>Recomm 2</i>	<i>Recomm 3</i>	<i>Recomm 4</i>	<i>Recomm 5</i>
Dokumen 1	0.5077	0.4432	0.4131	0.4083	0.4004
Dokumen 2	0.3970	0.3464	0.3403	0.3332	0.3085
Dokumen 3	0.3136	0.2883	0.2547	0.2235	0.2218
Dokumen 4	0.4202	0.4135	0.3319	0.3243	0.3228
Dokumen 5	0.3820	0.2898	0.2789	0.2653	0.2584

Tabel 6.3. Perhitungan MAP Hybrid Filtering

Data	Nilai Similarity					AP
	<i>Recomm 1</i>	<i>Recomm 2</i>	<i>Recomm 3</i>	<i>Recomm 4</i>	<i>Recomm 5</i>	
Dokumen 1	1	0	0	0	0	1
Dokumen 2	0	0	0,33	0	0	0,33
Dokumen 3	1	1	0	0	0	1
Dokumen 4	1	0	0	0	0	1
Dokumen 5	1	0	0	0	0	1
Total						4,33
$MAP = \frac{\text{total nilai AP}}{\text{total dokumen}} \times 100\%$						86%

Berdasarkan Tabel 6.3, dapat dilihat bahwa sistem rekomendasi yang dibangun menggunakan metode Hybrid Filtering menghasilkan nilai Mean Average Precision (MAP) sebesar 0,866 atau setara dengan 86,6%. Nilai ini diperoleh dari rata-rata nilai Average Precision (AP) pada lima dokumen uji yang mewakili lima mahasiswa berbeda. Pada masing-masing dokumen, sistem memberikan lima rekomendasi komunitas akademik. AP dihitung dengan mengidentifikasi posisi rekomendasi yang relevan, lalu menghitung rasio jumlah relevansi yang ditemukan terhadap urutan kemunculannya. Sebagai contoh, pada Dokumen 1, rekomendasi yang relevan muncul tepat di posisi pertama, sehingga mendapatkan nilai AP sempurna sebesar 1. Sementara itu, pada Dokumen 2, rekomendasi yang relevan berada di posisi ketiga, sehingga hanya memperoleh nilai AP sebesar 0,33. Adapun dokumen lainnya menunjukkan variasi posisi relevansi, yang secara keseluruhan memberikan total nilai AP sebesar 4,33. Dengan membagi total nilai AP tersebut terhadap jumlah dokumen, diperoleh nilai MAP sebesar 0,866. Nilai ini mencerminkan bahwa sistem memiliki performa yang cukup baik dalam merekomendasikan komunitas akademik yang relevan kepada mahasiswa berdasarkan profil dan preferensi mereka.

BAB VII

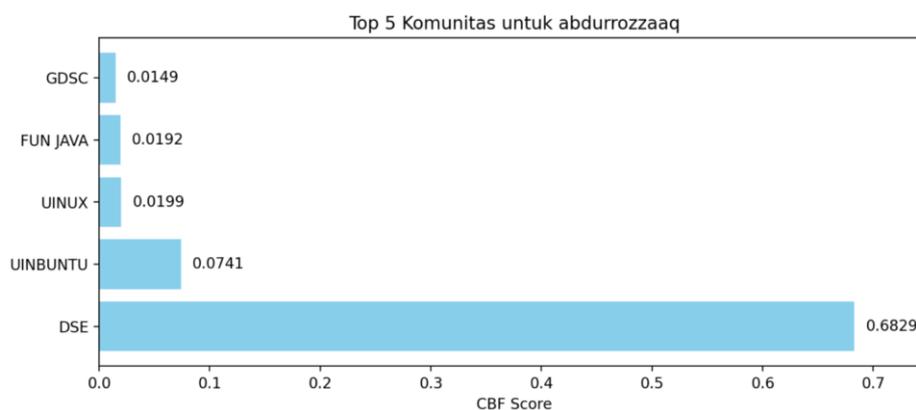
PEMBAHASAN

Berdasarkan seluruh pengujian yang telah dilakukan, bab ini menjelaskan temuan utama dan diskusi mendalam mengenai implementasi serta evaluasi sistem rekomendasi komunitas akademik yang dikembangkan. Penelitian ini menguji tiga pendekatan, yaitu Content-Based Filtering (CBF), Collaborative Filtering (CF), dan Hybrid Filtering. Tujuan utama sistem ini adalah untuk membantu mahasiswa, baik baru maupun lama, dalam menemukan komunitas akademik yang cocok dengan kemampuan dan minat mereka, sehingga dapat mendukung proses belajar serta meningkatkan produktivitas di bidang IT. Proses pembuatan sistem ini meliputi beberapa tahap penting: preprocessing data, pembobotan kata menggunakan TF-IDF, serta perhitungan kemiripan data memakai Cosine Similarity. Data yang digunakan dalam penelitian ini berasal dari 200 mahasiswa dan 10 komunitas akademik.

Fitur atau kriteria utama yang diambil dari data mahasiswa adalah minat, pengetahuan IT sebelumnya, preferensi tim (lebih suka tim kecil atau besar), skill yang telah dimiliki, motivasi mengikuti komunitas, serta rating atau nilai ketertarikan mahasiswa terhadap komunitas akademik. Sementara itu, data komunitas mencakup deskripsi, teknologi yang digunakan, aktivitas, serta visi dan misi komunitas. Data-data ini pada dasarnya menggambarkan preferensi dari mahasiswa dan ciri khas komunitas itu sendiri, yang mana cocok untuk digunakan dalam sistem hybrid filtering.

Setelah sistem berhasil dibangun, dilakukan pengujian untuk mengevaluasi performa masing-masing pendekatan, yaitu Content-Based Filtering, Collaborative Filtering, dan Hybrid Filtering. Evaluasi ini difokuskan pada analisis daftar rekomendasi yang dihasilkan oleh setiap metode, penentuan skenario data yang akan digunakan pada Content-Based Filtering, serta perbandingan performa metode menggunakan Mean Average Precision (MAP). Hasil evaluasi ini menjadi dasar untuk menentukan seberapa efektif gabungan metode ini dalam menemukan rekomendasi komunitas akademik yang paling relevan dan akurat.

Dalam implementasinya, sebelum dilakukan perhitungan menggunakan hybrid filtering, terlebih dahulu dilakukan pengujian terhadap dua metode individu, yaitu *content-based filtering* dan *collaborative filtering*. Pengujian ini menggunakan satu data dummy sebagai input pengguna untuk memastikan bahwa kedua metode berjalan sebagaimana mestinya serta menghasilkan daftar rekomendasi yang dapat dianalisis lebih lanjut. Hasil dari kedua pendekatan ini kemudian dibandingkan untuk mengetahui karakteristik dan perbedaan hasil rekomendasinya.



Gambar 7. 1. Hasil Cosine Similarity CBF

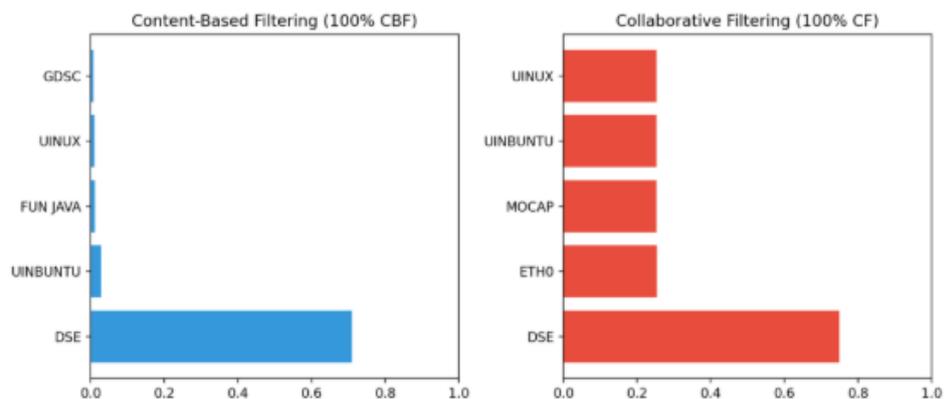
Pada implementasi Content-Based Filtering (CBF), data input dummy pengguna disimulasikan dengan memasukkan beberapa kriteria minat, seperti Passion 'Data Science', Minat Skill 'Python', dan preferensi tim 'Tim Besar'. Data input ini kemudian melalui tahap preprocessing, yang mencakup case folding, tokenizing, stopwords removal, dan stemming, untuk membersihkan dan menyeragamkan teks. Setelah bersih, data digabungkan menjadi metadata dan dihitung bobot fiturnya menggunakan algoritma TF-IDF. Langkah terakhir adalah menghitung tingkat kemiripan antara vektor fitur input pengguna dengan vektor fitur setiap komunitas menggunakan algoritma Cosine Similarity. Hasil perhitungan untuk input dummy ini menunjukkan bahwa komunitas DSE memperoleh skor kemiripan tertinggi, yaitu 0.6828, diikuti oleh UINBUNTU dengan skor 0.0741, dan UINUX dengan skor 0.0199. Hasil ini kemudian divisualisasikan dalam bentuk diagram batang untuk menampilkan 5 komunitas teratas yang direkomendasikan.

Tabel 7. 1. Tabel Prediksi Rating CF

Komunitas	Predict Ratings
FUN JAVA	1
GDSC	2
DSE	4
UINBUNTU	2
UINUX	2
MOCAP	2.0135
ONTAKI	2
MAMUD	2
WEBONDER	1
ETH0	2.0159

Selanjutnya, metode CF, yang berfokus pada rating mahasiswa, diimplementasikan menggunakan *item-based similarity* dan *Weighted Sum*. Tabel menampilkan hasil prediksi rating dari sistem rekomendasi berbasis *Collaborative*

Filtering (CF) untuk mahasiswa bernama abdurrozzaaq terhadap sejumlah komunitas akademik. Prediksi ini dihitung berdasarkan kemiripan antar komunitas (*item similarity*) dan riwayat interaksi mahasiswa dengan komunitas lainnya. Misalnya, mahasiswa abdurrozzaaq belum pernah memberi rating pada komunitas UINBUNTU, namun sistem memprediksi bahwa ia kemungkinan akan memberi rating sebesar 2.0131 berdasarkan preferensinya terhadap komunitas lain yang mirip dengan UINBUNTU. Prediksi tertinggi di antara komunitas yang belum dinilai mahasiswa adalah ETHO (2.0159) dan MOCAP (2.0135), sehingga komunitas-komunitas ini dapat direkomendasikan sebagai komunitas yang kemungkinan besar akan diminati oleh mahasiswa tersebut. Dengan pendekatan ini, sistem dapat memberikan rekomendasi yang lebih personal dan relevan bagi setiap mahasiswa.



Gambar 7. 2. Hasil CBF dan CF

Perbedaan signifikan antara hasil rekomendasi yang disajikan oleh metode Content-Based Filtering (CBF) dan Collaborative Filtering (CF) dapat diamati, yang merefleksikan karakteristik fundamental dari masing-masing pendekatan.

CBF cenderung menghasilkan rekomendasi yang bersifat lebih personal dan kontekstual, karena bekerja dengan menganalisis dan mencocokkan atribut atau metadata komunitas dengan profil minat pengguna, sehingga komunitas yang direkomendasikan memiliki kemiripan fitur yang tinggi. Sementara itu, CF menyajikan rekomendasi berdasarkan pola interaksi kolektif atau preferensi dari pengguna lain yang serupa; akibatnya, komunitas akademik yang disarankan mungkin tidak selalu memiliki kesamaan konten yang eksplisit, namun dianggap relevan karena diminati oleh kelompok pengguna dengan pola ketertarikan yang sejalan.

Keterbatasan pada masing-masing metode inilah yang mendorong perlunya sebuah pendekatan gabungan. Prinsip untuk menyatukan dua kekuatan yang berbeda untuk mencapai tujuan yang lebih baik dan saling menutupi kekurangan sejalan dengan ajaran Islam untuk saling tolong-menolong dalam kebaikan (ta'awun). Sebagaimana disebutkan dalam firman Allah SWT pada surat Al-Maidah ayat 2:

وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ ۖ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ

Artinya: "...Dan tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan pelanggaran..." (QS. Al-Maidah: 2)

Menurut Tafsir Quraish Shihab, ayat ini menganjurkan untuk saling bahu-membahu dalam segala bentuk kebaikan. Dalam konteks penelitian ini, prinsip ta'awun dapat dianalogikan pada cara kerja metode hybrid filtering. Metode CBF dan CF seakan "bekerja sama" untuk menghasilkan rekomendasi terbaik. CBF

memberikan "pertolongan" dengan memahami detail preferensi personal pengguna, sementara CF memberikan "pertolongan" dengan menyumbangkan kearifan kolektif dari pengguna lain. Hasil dari "kerja sama" ini adalah sistem rekomendasi yang lebih seimbang dan akurat, yang merupakan bentuk "kebaikan" bagi pengguna dalam menemukan komunitas yang tepat.

Pada tahap implementasi, setiap metode menghasilkan daftar rekomendasi yang berbeda sesuai dengan prinsip kerja masing-masing algoritma. Metode CBF menghasilkan rekomendasi berdasarkan kesamaan konten antar item dan preferensi pengguna, sedangkan CF menghasilkan rekomendasi dengan memanfaatkan pola interaksi dan kesamaan antar pengguna. Sementara itu, metode hybrid menggabungkan keunggulan dari kedua pendekatan tersebut dengan menerapkan bobot tertentu untuk hasil dari CBF dan CF.

Untuk mengukur efektivitas dari ketiga metode tersebut, dilakukan evaluasi menggunakan metrik Mean Average Precision (MAP). Metrik ini dipilih karena mampu mengukur tingkat ketepatan rekomendasi dengan memperhitungkan relevansi dan posisi item yang direkomendasikan. Hasil evaluasi menunjukkan adanya perbedaan nilai MAP dari masing-masing metode, yang menjadi dasar dalam menilai efektivitas dan performa sistem dalam memberikan rekomendasi yang relevan.

Untuk menganalisis performa secara mendalam, dilakukan perbandingan langsung antara hasil rekomendasi dari metode *Content-Based Filtering*, *Collaborative Filtering*, dan *Hybrid Filtering* menggunakan data uji dari lima

mahasiswa sebagai sampel. Rincian hasil perbandingan ini, termasuk skor similaritas (*Sim*) dan nilai *Average Precision* (AP) untuk setiap dokumen (mahasiswa), disajikan secara lengkap pada lampiran 3.

Metode *Simple Average* menunjukkan hasil yang lebih seimbang, dengan DSE tetap dominan namun komunitas seperti GDSC, UINUX, dan MAMUD juga memperoleh skor tinggi. Ini menunjukkan bahwa hybrid filtering mampu menggabungkan keunggulan kedua pendekatan, menghasilkan rekomendasi yang lebih akurat dan adaptif terhadap variasi karakteristik pengguna dan komunitas.

Analisis perbandingan performa secara mendalam menegaskan bahwa metode *Hybrid Filtering* memberikan kinerja paling unggul di antara ketiga pendekatan yang diuji. Berdasarkan rekapitulasi pengujian yang rinciannya tersaji pada lampiran 3 metode *Hybrid Filtering* berhasil mencapai nilai *Mean Average Precision* (MAP) sebesar 0,866 (86,6%). Nilai ini secara signifikan melampaui performa *Collaborative Filtering* (CF) yang memperoleh MAP 0,816 (81,6%) dan *Content-Based Filtering* (CBF) dengan MAP 0,733 (73,3%). Keunggulan metode hybrid terletak pada kemampuannya untuk beradaptasi dan menyeimbangkan kekuatan dari kedua metode pembentuknya sesuai dengan karakteristik data setiap pengguna.

Sebagai contoh, pada Dokumen 3 di lampiran 3 di mana profil personal pengguna sangat relevan, metode Hybrid (AP=1) mampu menyamai performa CBF (AP=1) yang lebih unggul dari CF (AP=0,75). Sebaliknya, pada Dokumen 5, di mana preferensi kolektif lebih akurat, metode Hybrid (AP=1) kembali menyamai

performa CF ($AP=1$) yang jauh melampaui CBF ($AP=0,415$). Kemampuan untuk secara dinamis memanfaatkan sinyal terkuat—baik dari kesamaan konten maupun dari pola rating—menjadikan pendekatan hybrid lebih stabil dan relevan, yang pada akhirnya menghasilkan akurasi keseluruhan yang tertinggi.



Gambar 7.3 Gambar Perbandingan 3 Metode (CBF-CF-HB)

Gambar 6.1 menyajikan visualisasi perbandingan hasil rekomendasi berdasarkan tiga pendekatan berbeda, yaitu *Content-Based Filtering* (CBF), *Collaborative Filtering* (CF), dan kombinasi keduanya melalui metode *Simple Average* (50% CBF + 50% CF). Tujuan dari perbandingan ini adalah untuk mengevaluasi performa tiap pendekatan dalam merekomendasikan komunitas yang paling relevan untuk pengguna (dalam hal ini mahasiswa). Pada CBF, komunitas DSE mendapat skor tertinggi karena kesesuaiannya dengan profil pengguna berdasarkan fitur konten. Sementara itu, CF juga menempatkan DSE sebagai rekomendasi utama, namun menyertakan komunitas lain seperti ETHO, MOCAP, dan UINUX berdasarkan pola preferensi pengguna serupa.

Dari visualisasi ini dapat disimpulkan bahwa metode *hybrid filtering* mampu menghasilkan hasil rekomendasi yang lebih stabil, di mana rekomendasi tidak hanya bergantung pada kesamaan konten, tetapi juga mempertimbangkan pola perilaku pengguna lain. Dengan demikian, pendekatan gabungan menjadi solusi yang lebih adaptif dalam menangani perbedaan karakteristik pengguna dan komunitas.

BAB VIII

KESIMPULAN DAN SARAN

8.1. Kesimpulan

Penelitian ini berhasil membangun sistem rekomendasi komunitas akademik berbasis *hybrid filtering* yang menggabungkan pendekatan *content-based filtering* (CBF) dan *collaborative filtering* (CF). Hasil pengujian menunjukkan bahwa metode hybrid memberikan kinerja yang lebih baik dibandingkan pendekatan tunggal, dengan nilai Mean Average Precision (MAP) sebesar 0,866, lebih tinggi dari CBF (0,733) maupun CF (0,816). Hal ini menunjukkan bahwa penggabungan kedua metode mampu menyeimbangkan antara kesesuaian konten dan pola perilaku pengguna, sehingga sistem rekomendasi menjadi lebih relevan dan adaptif terhadap variasi preferensi mahasiswa. Dengan demikian, metode hybrid filtering terbukti efektif dalam memberikan rekomendasi komunitas akademik yang sesuai dengan minat mahasiswa.

8.2. Saran

Berdasarkan hasil yang telah dicapai, disarankan agar pengembangan sistem rekomendasi komunitas akademik ke depannya terus memanfaatkan pendekatan *hybrid filtering* karena terbukti menghasilkan performa yang lebih baik. Selain itu, untuk peningkatan akurasi, sistem dapat dikembangkan lebih lanjut dengan mempertimbangkan bobot dinamis antara CBF dan CF, serta melibatkan fitur tambahan seperti aktivitas mahasiswa, riwayat keikutsertaan, dan umpan balik pengguna. Penelitian selanjutnya juga disarankan untuk menguji sistem ini dengan

data dalam skala lebih besar dan lingkungan nyata agar validitas dan skalabilitas sistem dapat lebih terukur.

DAFTAR PUSTAKA

- Abror, M. (2022). *Ojo Dibandingke atau Jangan Membandingkan dengan Orang Lain Menurut Islam*. Nuonline.
- Al-Ajlan, A., & Alshareef, N. (2023). Recommender System for Arabic Content Using Sentiment Analysis of User Reviews. *Electronics (Switzerland)*, 12(13). <https://doi.org/10.3390/electronics12132785>
- Arfisko, H. H., & Wibowo, A. T. (2022). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering dan Content-Based Filtering. *E-Proceeding of Engineering*, 9(3), 2149–2159.
- Arifitama, B. (2015). Perancangan Sistem Pakar Minat Mahasiswa Berdasarkan Kurikulum Program Studi Teknik Informatika Universitas XYZ. *Sisfo*, 05(03), 193–199. <https://doi.org/10.24089/j.sisfo.2015.03.005>
- Azcoitia, C. M., Practice, D., & Ed, D. (2017). *The Development of Academic Identity in Community Schools*. 7881(May), 175–181.
- Esteban, A., Zafra, A., & Romero, C. (2020). Helping university students to choose elective courses by using a hybrid multi-criteria recommendation system with genetic optimization. *Knowledge-Based Systems*, 194, 105385. <https://doi.org/10.1016/j.knosys.2019.105385>
- Jannah, M., Wulandari, N., & Hasibuan, S. R. (2021). *Tarbiyah Islamiyah : Jurnal Ilmiah Pendidikan Agama Islam NILAI-NILAI PENDIDIKAN AKHLAK DALAM AL-QURAN : KAJIAN TAFSIR SURAT AL-HUJURAT AYAT 9-13*. 11(2), 113–124. <https://doi.org/10.18592/jtipai.v11i2.4910>
- Kusumawati, R., Setyosari, P., Degeng, I. N. S., & Ulfa, S. (2021). Designing a cmooc for academic communities to support awareness of scaffolding procedure on sharing knowledge. *Journal of Engineering Science and Technology*, 16(4), 3576–3587.
- Larasati, F. B. A., & Februariyanti, H. (2021). Sistem Rekomendasi Product Emina Cosmetics Dengan Menggunakan Metode Content - Based Filtering. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 45. <https://doi.org/10.36595/misi.v4i1.250>
- Li, X., Zhang, M., Wu, S., Liu, Z., Wang, L., & Yu, P. S. (2020). Dynamic graph collaborative filtering. *Proceedings - IEEE International Conference on Data Mining, ICDM, 2020-Novem*, 322–331. <https://doi.org/10.1109/ICDM50108.2020.00041>
- Liu, Y., Xu, Z., Hui, C., Xuan, Y., Chen, J., & Shan, Y. (2019). *Heterogeneous Collaborative Filtering*. <http://arxiv.org/abs/1909.01727>
- Lucchese, C., Muntean, C. I., Perego, R., Silvestri, F., Vahabi, H., & Venturini, R. (2014). Recommender systems. In *Mining User Generated Content*. <https://doi.org/10.4018/ijeei.2013100103>
- Mathew, P., Kuriakose, B., & Hegde, V. (2016). Book Recommendation System through content based and collaborative filtering method. *Proceedings of 2016 International Conference on Data Mining and Advanced Computing, SAPIENCE 2016*, 47–52. <https://doi.org/10.1109/SAPIENCE.2016.7684166>

- Mondi, R. H., & Wijayanto, A. (2019). Recommendation System With Content-Based Filtering Method for Culinary Tourism in Mangan Application. *ITSMART: Jurnal Ilmiah Teknologi Dan Informasi*, 8(2), 65–72.
- Muflihah, Y., & Wardhani, M. W. (2024). Implementasi Hybrid Filtering dalam Sistem Rekomendasi Berbasis Web untuk Persewaan Tenda Pesta. *Jurnal Rekayasa Sistem Informasi Dan Teknologi*, 1(4), 361–372. <https://doi.org/10.59407/jrsit.v1i4.814>
- Nasser, A. M., Bhagat, J., Agrawal, A., & Devadas, T. J. (2023). Mean-reversion based hybrid movie recommender system using collaborative and content-based filtering methods. *International Journal of Statistics and Applied Mathematics*, 8(3S), 121–137. <https://doi.org/10.22271/math.2023.v8.i3sb.1012>
- Nastiti, P. (2019). Penerapan Metode Content Based Filtering Dalam Implementasi Sistem Rekomendasi Tanaman Pangan. *Teknika*, 8(1), 1–10. <https://doi.org/10.34148/teknika.v8i1.139>
- Nilashi, M., Ibrahim, O., Yadegaridehkordi, E., Samad, S., Akbari, E., & Alizadeh, A. (2018). Travelers decision making using online review in social network sites: A case on TripAdvisor. *Journal of Computational Science*, 28, 168–179. <https://doi.org/10.1016/j.jocs.2018.09.006>
- Pirker, J., Riffnaller-Schiefer, M., & Gütl, C. (2014). Motivational active learning - Engaging university students in computer science education. *ITICSE 2014 - Proceedings of the 2014 Innovation and Technology in Computer Science Education Conference*, 297–302. <https://doi.org/10.1145/2591708.2591750>
- Praherdhiono, H., Kusumawati, R., Pramono, E. A., & Atmoko, A. (2019). Strengthening MOOCs of academic community through scaffolding electronic automation. *International Journal of Innovation, Creativity and Change*, 5(4), 431–446.
- PUSPITASARI, M. (2022). Kerjasama Dalam Lembaga Pendidikan Berdasarkan Tafsir Al-Qur'an Surat Al-Maidah Ayat 2. *LEARNING : Jurnal Inovasi Penelitian Pendidikan Dan Pembelajaran*, 2(3), 209–221. <https://doi.org/10.51878/learning.v2i3.1521>
- Rahman, M., Shama, I. A., Rahman, S., & Nabil, R. (2022). Hybrid Recommendation System To Solve Cold Start Problem. *Journal of Theoretical and Applied Information Technology*, 100(11), 3562–3580.
- RIRIEN KUSUMAWATI. (2020). *PENGEMBANGAN SISTEM PENGELOLAAN PENGETAHUAN MENGGUNAKAN PENDEKATAN HEUTAGOGI DAN cMOOCs UNTUK KOMUNITAS AKADEMIK PERGURUAN TINGGI* (Issue July).
- Shambour, Q. Y., Al-Zyoud, M. M., Hussein, A. H., & Kharma, Q. M. (2023). A doctor recommender system based on collaborative and content filtering. *International Journal of Electrical and Computer Engineering*, 13(1), 884–893. <https://doi.org/10.11591/ijece.v13i1.pp884-893>
- Siregar, O. S., Fauseh, F., & Rosalina Gustari, D. P. (2020). Sistem Pendukung Keputusan Dalam Pemilihan Minat Mahasiswa Jurusan Teknik Informatika Menggunakan Metode AHP (Analytical Hierarchy Process) di Universitas Muhammadiyah Pontianak. *Digital Intelligence*, 1(1), 11. <https://doi.org/10.29406/diligent.v1i1.2330>
- Suprihatin, T., & Setiowati, E. A. (2021). Dukungan Komunitas Belajar dan Ketekunan Akademik Pada Mahasiswa. *Psychopolytan : Jurnal Psikologi*, 5(1), 20–25.

<https://doi.org/10.36341/psi.v5i1.1713>

- Tafsirweb. (2015). *Surat Al-Anfal Ayat 53 Arab, Latin, Terjemah dan Tafsir | Baca di TafsirWeb*.
- Talaghzi, J., Bellafkih, M., Bennane, A., Himmi, M. M., & Amraouy, M. (2023). A Combined E-Learning Course Recommender System. *International Journal of Emerging Technologies in Learning*, 18(6), 53–70. <https://doi.org/10.3991/ijet.v18i06.36987>
- Tommy, L., Novianto, D., & Japriadi, Y. S. (2020). Sistem Rekomendasi Hybrid untuk Pemesanan Hidangan Berdasarkan Karakteristik dan Rating Hidangan. *Journal of Applied Informatics and Computing*, 4(2), 137–145. <https://doi.org/10.30871/jaic.v4i2.2687>
- Valentino, N., & Setiawan, E. B. (2024). Movie Recommender System on Twitter Using Weighted Hybrid Filtering and GRU. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(2), 159–172. <https://doi.org/10.22219/kinetik.v9i2.1941>
- Xu, S., Ge, Y., Li, Y., Fu, Z., Chen, X., & Zhang, Y. (2023). Causal Collaborative Filtering. *ICTIR 2023 - Proceedings of the 2023 ACM SIGIR International Conference on the Theory of Information Retrieval*, 235–245. <https://doi.org/10.1145/3578337.3605122>
- Yan, X., Qi, S., & Chen, C. (2023). Recommender Systems: Collaborative Filtering and Content-based Recommender System. *Applied and Computational Engineering*, 2(1), 346–351. <https://doi.org/10.54254/2755-2721/2/20220658>
- Yang, S., Korayem, M., AlJadda, K., Grainger, T., & Natarajan, S. (2017). Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive Statistical Relational Learning approach. *Knowledge-Based Systems*, 136, 37–45. <https://doi.org/10.1016/j.knosys.2017.08.017>
- Zou, L., Xia, L., Gu, Y., Zhao, X., Liu, W., Huang, J. X., & Yin, D. (2020). Neural Interactive Collaborative Filtering. *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 749–758. <https://doi.org/10.1145/3397271.3401181>
- Zou, T. X. P. (2019). Community-based professional development for academics: a phenomenographic study. *Studies in Higher Education*, 44(11), 1975–1989. <https://doi.org/10.1080/03075079.2018.1477129>

LAMPIRAN

Lampiran 1. Sampel Data Mahasiswa Aktual

Mahasiswa	Passion	pengetahuan_sebelumnya	tim	skill	motivasi	komunitas_diikuti
Mahasiswa 1	Data Science	Pemograman Mobile, Pemograman Java, Data Science	Tim Kecil: Lebih ke proyek-proyek tertentu	Java, C#, Dart, Python	Ingin mempelajari lebih dalam tentang ilmu ilmunya	DSE
Mahasiswa 2	Robotik	Robotik	Tim Kecil: Lebih ke proyek-proyek tertentu	PHP, C++, C#, Figma, Blender, Cloud	Menambah relasi & pengetahuan di bidang terkait	ONTAKI
Mahasiswa 3	Pemograman Web	Pemograman Web	Tim Kecil: Lebih ke proyek-proyek tertentu	PHP, HTML, CSS, Javascript	Untuk mendapatkan pengalaman baru serta belajar tanggung jawab	MOCAP, WEBONDER, DSE
Mahasiswa 4	Pemograman Web	Pemograman Web, Data Science	Tim Kecil: Lebih ke proyek-proyek tertentu	HTML, CSS, Figma, Python	Mengembangkan keahlian dan mendapatkan pengalaman dan ilmu	MOCAP, WEBONDER, MAMUD, ETH0
Mahasiswa 5	Robotik	Pemograman Web, Multimedia, Robotik, UI/UX	Tim Kecil: Lebih ke proyek-proyek tertentu	C++, C#, Figma, Ubuntu, Cloud, Linux, Golang	Menambah insight baru	MOCAP, UINBUNTU, ONTAKI

Lampiran 2. Sampel Data Rating Mahasiswa Aktual

Mahasiswa	RATING KOMUNITAS									
	MOCAP	WEBONDER	UINUX	MAMUD	ETH0	UINBUNTU	FUN JAVA	GDSC	DSE	ONTAKI
Mahasiswa 1	4	0	0	0	0	0	0	0	0	0
Mahasiswa 2	3	0	0	0	0	0	0	0	0	0
Mahasiswa 3	2	4	0	0	0	0	0	0	4	0
Mahasiswa 4	3	3	0	2	2	0	0	0	0	0
Mahasiswa 5	3	0	0	0	0	2	0	0	0	4

Lampiran 3. Hasil Perbandingan CBF-CF-HB

	Content Based Filtering			Collaborative Filtering			Hybrid Filtering		
	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP
Dokumen 1	DSE	0.2654	1	DSE	0.5077	1	DSE	0.5077	1
	FUN JAVA	0.1363	0	FUN JAVA	0.4432	0	FUN JAVA	0.4432	0
	UINUX	0.0762	0	UINUX	0.4131	0	UINUX	0.4131	0
	UINBUNTU	0.0666	0	UINBUNTU	0.4083	0	UINBUNTU	0.4083	0
	GDSC	0.0508	0	GDSC	0.4004	0	GDSC	0.4004	0
	Jumlah AP		1	Jumlah AP		1	Jumlah AP		1
Dokumen 2	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP
	ETH0	0.0815	0	GDSC	0.3970	0	GDSC	0.3970	0
	ONTAKI	0.0744	0	WEBONDER	0.3464	0	WEBONDER	0.3464	0
	UINBUNTU	0.054	0	ONTAKI	0.3403	0,33	ONTAKI	0.3403	0,33
	WEBONDER	0.0466	0,25	ETH0	0.3332	0	ETH0	0.3332	0
	MAMUD	0.0457	0	MOCAP	0.3085	0	MOCAP	0.3085	0
	Jumlah AP		0,25	Jumlah AP		0,33	Jumlah AP		0,33
Dokumen 3	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP
	DSE	0.1325	1	MOCAP	0.5000	1	WEBONDER	0.3136	1
	WEBONDER	0.1272	1	WEBONDER	0.5000	1	MOCAP	0.2883	1
	MOCAP	0.0766	1	ONTAKI	0.3769	0	DSE	0.2547	0
	UINBUNTU	0.0707	0	DSE	0.3764	0,75	UINBUNTU	0.2235	0
	ONTAKI	0.0655	0	UINUX	0.3780	0	ONTAKI	0.2218	0
Jumlah AP		1	Jumlah AP		0,75	Jumlah AP		1	
Dokumen 4	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP
	WEBONDER	0.0903	1	DSE	0.75	0	WEBONDER	0.4202	1
	UINUX	0.0814	0	WEBONDER	0.75	1	DSE	0.4135	0

	Content Based Filtering			Collaborative Filtering			Hybrid Filtering		
	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP
	DSE	0.0771	0	MAMUD	0.5839	1	UINUX	0.3319	0
	FUN JAVA	0.067	0	UINBUNT	0.5831	0	FUN JAVA	0.3243	0
	GDSC	0.0639	0	UINUX	0.5825	0	GDSC	0.3228	0
	Jumlah AP		1	Jumlah AP		1	Jumlah AP		1
Dokumen 5	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP	Nama Komunitas	Sim	AP
	UINUX	0.0778	0	ONTAKI	0.7500	1	ONTAKI	0.3820	1
	MAMUD	0.0564	0	WEBONDER	0.5019	0	UINUX	0.2898	0
	UINBUNTU	0.0441	0,33	ETH0	0.5014	0	MAMUD	0.2789	0
	WEBONDER	0.026	0	GDSC	0.5047	0	WEBONDER	0.2653	0
	ONTAKI	0.014	0,5	FUN JAVA	0.5034	0	GDSC	0.2584	0
	Jumlah AP		0,415	Jumlah AP		1	Jumlah AP		1
Total AP			3,665	Total AP		4,08	Total AP		4,33
MAP			73%	MAP		82%	MAP		86%

Lampiran 4. Data Komunitas Akademik

NO	Nama Komunitas	Deskripsi	Aktivitas	Teknologi	Visi Misi
	FUN JAVA	Komunitas fun java adalah komunitas yang menaungi mahasiswa teknik informatika untuk mendukung mahasiswa dalam memperdalam skill pemrograman dengan bahasa java	It incubation, study club, membuat project dan belajar bersama mentor atau mentoring.	Java	menjadikan wadah utama bagi mahasiswa untuk mempelajari bahasa java dan bisa belajar bahasa java dalam menunjang perkuliahan
	GDSC	Google Developer Student Clubs (GDSC) adalah program dari Google yang ditujukan untuk mahasiswa di perguruan tinggi di seluruh dunia. Tujuan dari program ini adalah untuk memberdayakan mahasiswa dalam pengembangan perangkat lunak dan teknologi terkini, serta membangun komunitas yang bersemangat dalam teknologi.	Mempelajari berbagai topik teknis dan dapatkan keterampilan baru melalui, acara, diskusi, dan aktivitas pembangunan proyek secara online dan tatap muka. pembelajaran baru untuk membangun solusi hebat bagi permasalahan lokal. Tingkatkan keterampilan, karier, dan jaringan. Feedback kepada komunitas Anda dengan membantu orang lain belajar.	php, javascript, flutter, figma, css, laravel, node.js, express.js	Visi : Menjadi wadah utama bagi mahasiswa di seluruh dunia untuk mengembangkan keterampilan teknis, berinovasi, dan menjadi pemimpin dalam industri teknologi. Misi : Memberdayakan Mahasiswa: GDSC bertujuan untuk memberdayakan mahasiswa dengan sumber daya, pelatihan, dan kesempatan untuk mengembangkan keterampilan teknis yang diperlukan dalam dunia industri teknologi yang terus berkembang, Mendorong Inovasi: GDSC mendorong mahasiswa untuk berinovasi dan menciptakan solusi teknologi yang berdampak positif bagi masyarakat dan lingkungan sekitar. Ini dilakukan melalui proyek kolaboratif, hackathon, dan kompetisi yang diadakan oleh klub, Membangun Komunitas: GDSC menciptakan lingkungan yang mendukung dan inklusif di mana mahasiswa dapat belajar, berbagi ide, dan berkolaborasi dalam pengembangan perangkat lunak dan teknologi. Ini membantu dalam membangun jaringan profesional yang kuat di kalangan mahasiswa dan profesional teknologi, Menyediakan Akses ke Sumber Daya: GDSC menyediakan akses ke sumber daya pembelajaran, pelatihan, dan dukungan teknis dari Google dan mitra-mitra teknologi lainnya. Hal ini membantu mahasiswa untuk terus belajar dan mengembangkan keterampilan mereka di bidang teknologi, Mendorong Keterlibatan dengan Komunitas Teknologi: GDSC mendorong keterlibatan aktif mahasiswa dalam komunitas teknologi di tingkat lokal dan global. Ini melibatkan partisipasi dalam

NO	Nama Komunitas	Deskripsi	Aktivitas	Teknologi	Visi Misi
					acara-acara industri, konferensi, dan kegiatan-kegiatan kolaboratif dengan pihak-pihak terkait.
	DSE	DSE (Data Science Enthusiast) merupakan sebuah komunitas yang berfokus pada bidang Data Science	Di komunitas DSE kami mempelajari mengenai Data Science Life Cycle dan belajar menerapkan setiap dari cycle tersebut dari awal hingga akhir	Python	Visi: Menciptakan lingkungan mahasiswa yang memiliki concern terhadap bidang keilmuan data science (terkhusus pada bagian machine learning) Misi: Membahas dan mempelajari setiap bagian dari data science life cycle agar semua anggota komunitas memiliki pengalaman dan pemahaman yang konkret tentang data science
	UINBUNTU	Uinbuntu adalah komunitas Open-Source dan Cloud Computing prodi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang. Komunitas ini melakukan pengembangan perangkat lunak berbasis sistem Operasi Linux UBUNTU. Komunitas ini sangat menunjang komunitas ETH0 dalam hal penyediaan platform jaringan komputer dan sistem operasi.	Open source dapat diartikan sebagai perangkat lunak yang kode sumber atau kode dasarnya dapat digunakan oleh banyak orang. Selain itu, dengan menggunakan perangkat lunak open source ini kamu dapat mengembangkan aplikasi dan membuat perangkat lunak versi terbaru. Mempelajari open source memiliki beragam manfaat, di antaranya adalah meningkatkan pemahaman tentang teknologi, memperluas kemampuan dalam pengembangan perangkat lunak, serta memungkinkan kolaborasi dengan komunitas global untuk memecahkan masalah dan menciptakan inovasi. Dengan memahami kode sumber terbuka, individu dapat belajar dari praktik terbaik yang telah ada, mengembangkan keterampilan dalam pemrograman, dan berkontribusi pada proyek-proyek yang dapat memberikan dampak positif bagi masyarakat. Selain itu, mempelajari open source juga dapat membantu dalam mengurangi ketergantungan pada vendor tertentu,	linux, vscode, php, javascript, python, apache http server	Visi kami adalah menjadi komunitas pembelajar open source dan cloud computing yang ulul albab. Kami berkomitmen untuk menjadi pusat pembelajaran dan pengembangan teknologi terbuka yang berkualitas dan berdaya saing tinggi. Misi kami adalah Menganalisis System Open Source, Mengembangkan System Open Source, Melakukan Kerjasama Terkait dengan Pengembangan Open Source

NO	Nama Komunitas	Deskripsi	Aktivitas	Teknologi	Visi Misi
			memberikan fleksibilitas dalam penggunaan dan pengembangan teknologi, serta meningkatkan keamanan dan keandalan sistem. Cloud computing adalah penyediaan sumber daya komputasi seperti server, penyimpanan data, jaringan, dan perangkat lunak melalui internet. Mempelajari cloud computing memberikan beragam manfaat, seperti memungkinkan akses mudah dan cepat terhadap sumber daya komputasi, penyimpanan data, dan aplikasi melalui internet. Dengan memahami konsep dan teknologi cloud computing, individu dapat mengoptimalkan efisiensi dan skalabilitas dalam pengelolaan infrastruktur IT, mengurangi biaya investasi awal, dan meningkatkan fleksibilitas dalam penyediaan layanan dan aplikasi. Selain itu, mempelajari cloud computing juga membuka peluang untuk eksplorasi dan adopsi teknologi baru, seperti machine learning, big data analytics, dan Internet of Things (IoT), yang dapat memberikan nilai tambah dalam berbagai bidang industri.		
	UINUX	Komunitas bagi mahasiswa TI yang ingin belajar tentang UI/UX Designer	Belajar desain bareng, meeting rutin pengurus, design challenge, webinar UI/UX	Figma	Sebagai wadah para mahasiswa dalam mempelajari dunia UI/UX Design dan sebagai jembatan untuk menjadi seorang UI/UX Designer
	MOCAP	Komunitas Aplikasi Mobile Teknik Informatika UIN Maulana Malik Ibrahim Malang	Course online, Workshop, Study Jam, IT INCU	Flutter, Java, Kotlin, Figma	Membantu para mahasiswa untuk mengembangkan minat dan skill dalam bidang pemrograman mobile, Update terhadap perkembangan teknologi mobile
	ONTAKI	Komunitas yang bergerak dibidang otomasi dan robotika	Dalam ontaki biasanya para pengurusnya difokuskan untuk mengikuti lomba robotika, seperti sumo robot, soccer, drone, dan lainnya. Selain mengikuti lomba, ontaki juga mengadakan workshop mengenai robotika, dan juga mengadakan	Arduino uno, python, c++, solid work	Visi: Menciptakan komunitas yang menjadi pusat inovasi dalam bidang otomasi dan robotika, di mana anggotanya dapat mengembangkan kreativitas dan keahlian mereka untuk merancang solusi teknologi masa depan. Misi: Mendorong partisipasi aktif anggota dalam berbagai lomba robotika seperti sumo

NO	Nama Komunitas	Deskripsi	Aktivitas	Teknologi	Visi Misi
			belajar bareng karena di beberapa semester masih ada mata kuliah yang berkaitan dengan ontaki seperti elektronika digital dan sistem komputer.		robot, soccer, drone, dan kompetisi sejenis untuk menguji dan memperbaiki keterampilan teknis mereka, Menyelenggarakan workshop rutin mengenai robotika untuk memperluas pengetahuan dan keterampilan anggota dalam pengembangan teknologi otomasi, Mengadakan kegiatan belajar bareng untuk membahas topik-topik yang berkaitan dengan mata kuliah seperti elektronika digital dan sistem komputer, sehingga anggota dapat menerapkan konsep-konsep tersebut dalam proyek-proyek praktis, Membangun jaringan kolaborasi dengan industri dan lembaga pendidikan untuk mendukung pengembangan karier anggota dan mendorong adopsi teknologi otomasi yang berkelanjutan, Menjadi wadah bagi inovasi dan eksperimen dalam otomasi dan robotika, dengan tujuan utama meningkatkan efisiensi, produktivitas, dan kualitas hidup melalui penerapan teknologi canggih.
	MAMUD	Mamud kepanjangan dari manusia multimedia , yang pastinya bergerak di bidang multimedia	Sharing session untuk memperdalam dunia ke multimediaan khususnya anggota mamud sendiri, study camp guna untuk belajar bersama untuk menyelesaikan permasalahan yang ada di mata kuliah yang bersangkutan dengan multimedia	Figma, unity , 3d blender, c#	Visi : mewujudkan Mamud sebagai komunitas yang kreatif dan inovatif Misi : 1. menyelenggarakan pembelajaran dalam mendalami dunia multimedia 2. Melaksanakan kegiatan dalam dunia kreatif
	WEBONDER	Weboender merupakan komunitas yang ada di prodi Teknik Informatika UIN Malang yang bergerak pada bidang web development. Komunitas ini menaungi mahasiswa UIN Malang khususnya prodi Teknik Informatika yang mempunyai minat atau ketertarikan di bidang web development. Komunitas ini juga akan mengajak para mahasiswa untuk menggali lebih	Ada beberapa aktivitas yang ada di Weboender yaitu Study Club, Bootcamp, dan Sharing Session. Study club merupakan event dari Weboender untuk berkumpul secara berkala untuk mendalami topik-topik tertentu dalam pengembangan web. Bootcamp merupakan event tahunan yang diselenggarakan oleh Weboender, Bootcamp ini juga merupakan event terbesar dalam komunitas ini. Untuk program yang sudah terlaksana yaitu Bootcamp Laravel. Bootcamp Laravel adalah program intensif yang dirancang	PHP, Javascript, Laravel, CSS, Figma.	Visi : Menjadikan Weboender sebagai wadah bagi mahasiswa dalam mengembangkan keahlian khususnya di bidang web development. Misi : Mengajak mahasiswa untuk menggali atau memahami lebih dalam tentang dunia web development. Memberikan edukasi kepada mahasiswa tentang web development melalui acara-acara yang ada pada Weboender. Menciptakan forum terbuka untuk pegiat web developer khususnya di Teknik Informatika UIN Malang.

NO	Nama Komunitas	Deskripsi	Aktivitas	Teknologi	Visi Misi
		dalam tentang pemahaman web development.	untuk memberikan pelatihan mendalam tentang pengembangan web menggunakan framework Laravel. Sharing session adalah forum terbuka dimana anggota komunitas dapat berbagi pengalaman, ide, atau pengetahuan tentang pengembangan web. Tujuan utama dari sharing session adalah untuk memfasilitasi pertukaran informasi dan pengalaman antar anggota sehingga mereka dapat belajar satu sama lain dan memperluas relasi atau wawasan mereka.		