**Skema Fundamental UNIB tahun 2021**



**USULAN PENELITIAN**

**PEMBINAAN UNIB 2021**

**PENERAPAN *WEB SCRAPING* DAN *ITEM-BASED COLLABORATIVE* *FILTERING* PADA SISTEM REKOMENDASI ARTIKEL ILMIAH BERBASIS WEB BERDASARKAN KETERKAITANTOPIK SKRIPSI**

TIM PENELITI

**Ferzha Putra Utama, S.T., M.Eng. (NIDN 0023068905)**

**Nanang Sugianto, S.Si., M.Si. (NIDN 0013019004)**

**Widhia Oktoeberza KZ., S.T., M.Eng. (NIDN**

**FAKULTAS TEKNIK**

**UNIVERSITAS BENGKULU**

**2021**

# HALAMAN PENGESAHAN

|  |  |  |
| --- | --- | --- |
| Judul Penelitian | : | **Penerapan Web Scraping dan Item-Based Collaborative Filtering pada Sistem Rekomendasi Artikel Ilmiah Berbasis Web Berdasarkan Keterkaitan Topik Skripsi** |
| Ketua Peneliti |  |  |
| 1. Nama Lengkap | : | Ferzha Putra Utama, S.T., M.Eng. |
| 1. NIP | : | 19890623 2018031001 |
| 1. NIDN | : | 0023068905 |
| 1. Pangkat/Golongan | : | Penata Muda Tk.1/IIIb |
| 1. Jabatan/Fungsional | : | Asisten Ahli |
| 1. Fakultas/Jurusan | : | Teknik/Sistem Informasi |
| 1. Pusat Penelitian | : | Universitas Bengkulu |
| 1. Alamat Institusi | : | Jalan W.R. Supratman Kandang Limun 38371 |
| 1. Telepon/Fax/Email | : | (0736) 21170 |
| Dosen Anggota | : | 1 Orang |
| Nama Anggota | : | Julia Purnama Sari, S.T., M.Kom. |
| NIDN/NUP | : | 0009079004 |
| Bidang Keahlian | : | Sistem Informasi & Multimedia |
| Biaya yang diusulkan | : | Rp. 8.500.000,00 |

Bengkulu, 24 Agustus 2020

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Menyetujui, | |  | | |  | |
| Dekan Fakultas Teknik | |  | | | Ketua Peneliti | |
| Faisal Hadi, S.T., M.T.  NIP.197707132002121005 | | |  | Ferzha Putra Utama, S.T., M.Eng.  NIP. 198906232018031001 | | |
|  | Mengetahui:  Ketua LPPM  Dr. Ir. Hery Suharyanto, M.Sc.  NIP. 196306251987031002 | | | | |  |

# ABSTRAK

Skripsi merupakan syarat wajib untuk menyelesaikan studi dalam program sarjana (strata 1) di Universitas Bengkulu. Dalam proses menyusun skripsi masih banyak mahasiswa mengalami kebingunan untuk menentukan topik bahasan atau judul yang akan digunakan sebagai penelitian. Ketentuan pengajuan topik yang berdasarkan keterkaitan mata kuliah diambil memberikan dampak pada mahasiswa untuk membaca dan mengulas topik yang diajukan. Hal ini memungkinkan seorang mahasiswa mengalami kesulitan dalam mencari artikel yang sesuai dengan kebutuhannya. **Penelitian ini bertujuan untuk merancang dan membangun sistem rekomendasi artikel ilmiah menggunakan metode *web scraping* dan *item-based collaborative filtering* dalam bahasa pemrograman Python**. Penelitian ini berfokus dengan metode *web scraping* dan *item-based* *collaborative filtering* pada sistem rekomendasi yang dapat membantu pengguna dalam merekomendasikan artikel yang sesuai/relevan dengan bahasan yang dicari. Hasil Penelitian ini diharapkan dapat dimanfaatkan dengan baik untuk para pembuat tugas akhir dengan menggunakan sistem ini dapat membantu dalam mencari artikel yang berkaitan dengan topik skripsi. **Luaran** yang diharapkan penelitian ini adalah sistem rekomendasi artikel ilmiah berbasis web berdasarkan keterkaitan topik skripsi mahasiswa dan publikasi pada **jurnal nasional terkareditasi Sinta 5**.

Kata Kunci: Skripsi, Sistem Rekomendasi, *Web Scraping*, *Item-Based Collaborative filtering*.

# DAFTAR ISI

[HALAMAN PENGESAHAN i](#_Toc68619276)

[ABSTRAK ii](#_Toc68619277)

[DAFTAR ISI iii](#_Toc68619278)

[BAB I](#_Toc68619279) [PENDAHULUAN 1](#_Toc68619280)

[1.1 Latar Belakang 1](#_Toc68619281)

[1.2 Rumusan Masalah 3](#_Toc68619282)

[1.3 Batasan Masalah 3](#_Toc68619283)

[1.4 Tujuan Penelitian 4](#_Toc68619284)

[1.5 Manfaat Penelitian 4](#_Toc68619285)

[BAB II](#_Toc68619286) [TINJAUAN PUSTAKA 5](#_Toc68619287)

[2.1 Sistem Rekomendasi 5](#_Toc68619290)

[2.2 Skripsi 5](#_Toc68619291)

[2.3 *Google Scholar* (Google Cendekia) 5](#_Toc68619292)

[2.4 *Hyperteks Markup Language - Document Object Model* (HTML - DOM) 6](#_Toc68619293)

[2.5 *Web Scraping* 6](#_Toc68619294)

[2.6 Legalitas *Web Scraping* 8](#_Toc68619295)

[2.7 *Collaborative filtering* 9](#_Toc68619296)

[2.8 *Item-Based Collaborative filtering* 12](#_Toc68619297)

[2.9 Skala Penilaian 13](#_Toc68619298)

[2.10 Pengujian Sistem Rekomendasi 14](#_Toc68619299)

[2.11 *Mean Absolute Error* (MAE) 14](#_Toc68619300)

[2.12 Pengujian Presisi (Precision) 15](#_Toc68619301)

[2.13 *User Acceptance Test* (UAT) 15](#_Toc68619302)

[2.14 Flask 16](#_Toc68619303)

[2.15 Penelitian Terkait 16](#_Toc68619304)

[BAB III](#_Toc68619305) [PETA JALAN PENELIATAN DAN LUARAN PENELITIAN 19](#_Toc68619306)

[3.1 Peta Jalan Peneltian 19](#_Toc68619308)

[3.2 Luaran Penelitian 19](#_Toc68619309)

[BAB IV](#_Toc68619310) [METODE PENELITIAN 20](#_Toc68619311)

[4.1 Input Data 20](#_Toc68619313)

[4.2 Pengolahan Data 21](#_Toc68619314)

[4.3 Analisis Data 21](#_Toc68619315)

[4.4 Output 21](#_Toc68619316)

[4.5 Alat dan Bahan 22](#_Toc68619317)

[4.6 Lokasi Penelitian 22](#_Toc68619318)

[4.7 Target Luaran 22](#_Toc68619319)

[BAB V JADWAL PELAKSANAAN 24](#_Toc68619320)

[BAB VI](#_Toc68619321)  [REANCANAGAN ANGGARAN BIAYA DAN JUSTIFIKASI ANGGARAN 25](#_Toc68619322)

[DAFTAR PUSTAKA 26](#_Toc68619324)

# BAB I

# PENDAHULUAN

## Latar Belakang

Skripsi merupakan syarat wajib untuk menyelesaikan studi dalam program sarjana (strata 1) di Universitas Bengkulu. Skripsi bertujuan agar mahasiswa mampu mengembangkan sebuah penelitian dengan memadukan pengetahuan dan keterampilan dalam memahami, menganalisis, menggambarkan, dan menjelaskan masalah yang berhubungan dengan bidang keilmuan yang diambilnya.

Dalam proses menyusun skripsi masih banyak mahasiswa mengalami kebingunan untuk menentukan topik bahasan atau judul yang akan digunakan sebagai penelitian. Ketentuan pengajuan topik yang berdasarkan keterkaitan mata kuliah diambil memberikan dampak pada mahasiswa untuk membaca dan mengulas topik yang diajukan. Hal ini memungkinkan seorang mahasiswa mengalami kesulitan dalam mencari artikel yang sesuai dengan kebutuhannya.

Saat ini banyak mahasiswa memanfaatkan internet untuk mencari artikel sebagai referensi penelitian. Namun, informasi yang tersebar luas di internet tersebut tidak sepenuhnya dapat dipertanggungjawabkan kebenarannya. Karena setiap orang bebas untuk menulis dan menyebarkan informasi melalui internet. Untuk menghindari ini telah disediakan beberapa layanan yang menyediakan artikel ilmiah salah satu layanan yang paling umum digunakan *google scholar*.

*Google scholar* atau yang disebut juga dengan google cendekia di Indonesia ialah layanan yang untuk melakukan pencarian publikasi ilmiah seperti jurnal, buku dan tesis atau skripsi. Namun, terdapat kekurangan dalam layanan tersebut, di antaranya adalah artikel yang baru ditambahkan tidak akan pernah direkomendasikan karena hanya artikel yang paling banyak dirujuk atau memiliki *rating* tinggi yang ditampilkan di halaman awal. Hal ini memungkinkan pengguna mengalami masalah ketika artikel yang ditampilkan terkadang tidak relevan dengan apa yang dicari. Oleh karena itu, dibutuhkan sistem rekomendasi yang dapat membantu mahasiswa dalam mencari dan merekomendasikan artikel yang sesuai dengan bahasan yang dicari.

Menurut Fadelillah (2017), sistem rekomendasi merupakan suatu aplikasi membuat suatu keputusan untuk merekomendasikan sekumpulan *item* yang diinginkan oleh pengguna. Sistem ini seperti penggambaran kebutuhan dan keinginan pengguna dengan menggunakan metode rekomendasi untuk mencari suatu *item* berupa artikel. Kemudian sistem rekomendasi melakukan proses pencarian dan merekomendasikan artikel berdasarkan kemiripan dari artikel yang dipilih.

Ketidakrelevanan artikel yang dicari mengakibatkan kebingungan pada mahasiswa yang sedang mendalami atau memahami topik skripsi. Berdasarkan data kuisioner dari 43 sampel mahasiswa Fakultas Teknik Universitas Bengkulu yang mengambil mata kuliah skripsi menunjukan 74,4% menyatakan kesulitan dalam mencari artikel yang berkaitan dengan topik skripsi dan 76,7% menyatakan artikel yang ditemukan terkadang tidak relevan. Hal ini dikarenakan belum ada sistem yang secara efektif memberikan informasi yang akurat tentang perekomendasian artikel ilmiah. Maka diperlukan suatu aplikasi yang dapat memberikan rekomendasi artikel ilmiah yang berkaitan dengan topik skripsi mahasiswa. Salah satu metode yang digunakan adalah *web scraping* dan *item-based* *collaborative filtering*.

*Web scraping* merupakan salah satu metode untuk mendapatkan artikel dari *google scholar* secara otomatis tanpa harus menyalinnya atau menginputkan secara manual. Tujuan dari *web scraping* untuk mencari suatu informasi dan menampilkan dalam web yang baru [1]. Namun, terdapat kekurangan dalam metode *web scraping*, artikel yang didapatkan tidak disaring terlebih dahulu. Sehingga rendahnya tingkat akurasi kemiripan atau informasi yang dinginkan oleh pengguna. Untuk mendapatkan artikel yang serupa berdasarkan tingkat kemiripan perlu salah satu metode rekomendasi seperti *item-based* *collaborative filtering.*

*Item-based collaborative filtering* merupakan salah satu bagian dari metode *collaborative filtering* untuk menghasilkan prediksi rekomendasi *item* dengan menghitung kemiripan antara suatu *item* dengan *item* yang lain. Algoritme *item-based collaborative filtering* memberikan kinerja dan kualitas prediksi yang lebih baik dari pada algoritme *user-based collaborative filtering* [2]*.*

Berdasarkan paparan di atas, peneliti dilakukan sebuah penelitian dengan topik “*Penerapan web scraping dan item-based* *collaborative filtering pada sistem rekomendasi artikel ilmiah berbasis web berdasarkan keterkaitan topik skripsi*”*.* Penelitian ini berfokus dengan metode *web scraping* dan *item-based* *collaborative filtering* pada sistem rekomendasi yang dapat membantu pengguna dalam merekomendasikan artikel yang sesuai/relevan dengan bahasan yang dicari. Sistem informasi ini diharapkan dapat dimanfaatkan dengan baik untuk para pembuat tugas akhir dengan menggunakan sistem ini dapat membantu dalam mencari artikel yang berkaitan dengan topik skripsi.

## Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, dapat dirumuskan beberapa permasalahan sebagai berikut:

1. Bagaimana memberikan rekomendasi artikel ilmiah sebagai sarana mahasiswa dalam mencari referensi penulisan skripsi.
2. Bagaimana merancang dan mengimplementasikan sebuah sistem rekomendasi dengan menggunakan metode *web scraping* dan *item-based* *collaborative filtering* dalam mencari artikel ilmiah berdasarkan keterkaitan topik skripsi.

## Batasan Masalah

Agar aplikasi ini tidak terlalu luas dalam proses pengembangannya, maka peneliti membuat batasan-batasan permasalahan sebagai berikut :

1. Sistem rekomendasi artikel ilmiah menggunakan metode *web scraping* dan *item-based* *collaborative filtering.*
2. Sistem rekomendasi artikel ilmiah dirancang dengan menggunakan bahasa pemrograman Python.
3. Data pada sistem ini bersumber dari *Google Scholar* (Google Cendekia).
4. Sistem akan memberikan rekomendasi pada *user* terdaftar yang telah me-*rating* minimal 1 artikel, jika *user* belum terdaftar atau belum me-*rating*, *user* akan diberikan rekomendasi artikel yang melimiki *rating* terbanyak.
5. Sistem ini diuji cobakan pada mahasiswa Fakultas Teknik Universitas Bengkulu yang mengambil mata kuliah skripsi.

## Tujuan Penelitian

Tujuan penelitian yang ingin dicapai dalam penelitian ini adalah merancang dan membangun sistem rekomendasi artikel ilmiah menggunakan metode *web scraping* dan *item-based* *collaborative filtering* dalam bahasa pemrograman Python.

## Manfaat Penelitian

Manfaat dari penerapan *web scraping* dan *item-based* *collaborative filtering* pada sistem rekomendasi artikel ilmiah berbasis web berdasarkan keterkaitan topik skripsi adalah sebagai berikut :

1. Membantu dalam merekomendasikan artikel ilmiah yang berkaitan dengan topik skripsi pengguna.
2. Membantu meningkatkan referensi paling berkaitan pada skripsi pengguna.

# BAB II

# TINJAUAN PUSTAKA



## Sistem Rekomendasi

Sistem rekomendasi merupakan suatu aplikasi untuk membantu *user* dengan merekomendasikan sekumpulan *item* yang sesuai dengan profil pengguna. Sistem rekomendasi menganalisis informasi yang diperoleh melalui profil pengguna, preferensi *item* dan aktivitas yang terjadi pada sistem untuk dilakukan perhitungan untuk memprediksi *item* yang akan direkomendasikan kepada pengguna. Profil pengguna dapat berisi tentang informasi pengguna, ketertarikan pengguna pada suatu *item* dan juga *history* interaksi antara pengguna dengan *item* [3]. Misalnya ketika pengguna mencari atau mengunduh suatu artikel pada situs, data aktivitas atau *history* yang dilakukan akan dicatat dan disimpan pada profil pengguna.

## Skripsi

Skripsi merupakan syarat wajib untuk menyelesaikan studi dalam program sarjana (strata 1) di Program Studi Informatika Universitas Bengkulu. Skripsi juga memiliki syarat minimal 120 sks dengan nilai minimal C. Sudah lulus kerja praktik dan sudah lulus matakuliah metodologi penelitian. Skripsi normal-nya dilaksanakan pada saat mahasiswa duduk di bangku semester 8, namun tidak menutup kemungkinan untuk diambil lebih awal asalkan persyaratan untuk mengambil tugas akhir dapat dipenuhi [4].

## *Google Scholar* (Google Cendekia)

*Google Scholar* merupakan situs *web* yang menyediakan layanan pencarian publikasi ilmiah seperti jurnal, buku dan tesis atau skripsi. Situs ini dibangun oleh Alex Verstak pada 20 November 2004 [5]. *Google Scholar* menyediakan berbagai fitur layanan untuk memudahkan pencarian artikel ilmiah, sebagai berikut :

1. Pencarian meliputi pencarian artikel ilmiah (jurnal, buku dan tesis atau skripsi), penulis, sumber pustaka, tahun dan penerbit.
2. Unduhan sitasi yang memudahkan pengguna mengambil artikel (sitasi) berbentuk *BibTeX, EndNote, RefMan* dan *RefWorks*.
3. Unduhan artikel ilmiah secara utuh (*fulltext)* dalam format PDF.
4. Artikel dirujuk dan artikel terkait untuk memudahkan pengguna untuk mencari artikel yang berhubungan.

*Google Scholar* dapat digunakan hanya dengan memasukkan kata kunci ke dalam kotak pencarian untuk menemukan artikel ilmiah yang dibutuhkan. *Google Scholar* juga menyediakan fasilitas pencarian lanjut dengan memilih opsi kata kunci yang dipilih, sesuai dengan judul, abstrak, pengarang, dan penerbitnya. *Google Scholar* juga dapat memfilter hasil pencarian dengan menyesuaikan tahun terbit, mengurutkan berdasarkan relevansi dan tanggal.

## *Hyperteks Markup Language - Document Object Model* (HTML - DOM)

HTML-DOM adalah model standar penulisan sebuah dokumen HTML. HTML-DOM berfungsi untuk mengatur elemen-elemen html disusun memenuhi dan bagaimana untuk mendapatkan, mengubah, menambah, atau menghapus elemen html.

Elemen html adalah kode html yang dimulai dari sebuah *tag* pembuka dan diakhiri dengan *tag* penutup, *tag* itu sendiri adalah kode-kode tertentu yang menjadi pengenal bahwa kode tersebut adalah kode html dan bisa diterjemahkan oleh browser. Ada tiga *tag* utama yang membangun sebuah halaman html, yaitu *tag* html (<html> </html>), *tag head* (<head> </head>) dan *tag body* (<body> </body>) [6].

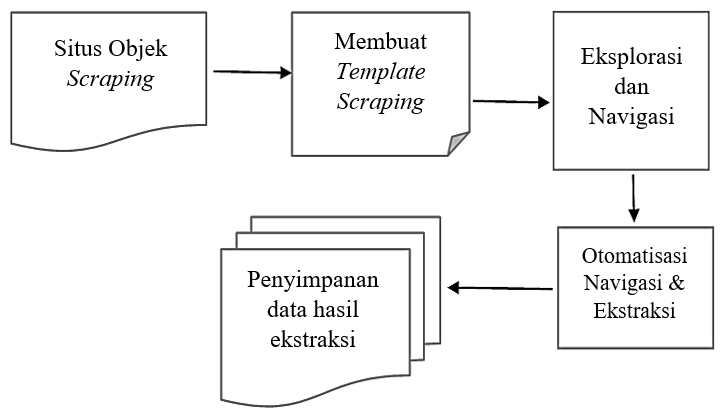
*Tag body* adalah *tag* yang mengapit semua informasi yang ingin ditampilkan pada sebuah halaman *website*. Situs pada umumnya menempatkan isi informasi di antara *tag* yang ada di dalam *tag body*. Untuk mengambil informasi tersebut, maka harus dilakukan proses ekstraksi untuk memisahkan konten dari *tag html* yang mengapitnya, proses ekstraksi ini dikenal dengan istilah *Web Scraping*.

## *Web Scraping*

*Web scraping* adalah sebuah metode untuk mendapatkan informasi dari sebuah halaman *website*.Metode *Scraping* merupakan salah satu metode untuk melakukan pengumpulan informasi dari internet. *Web scraping* dirancang untuk mengenali konten yang berbeda didalam sebuah *website* dan untuk mendapatkan serta menyimpan kategori informasi yang diinginkan oleh pengguna. Sebagai contoh, *web scraping* memungkinkan pengguna untuk mencari sebuah artikel pada suatu *website* dan menyimpan data judul, tahun,abstrak, sitasi dan data pendukung lainnya.

*Web scraping* memiliki sejumlah tahapan, menurut penelitian Maria Rosario B, dkk pada tahun 2017 menjelaskan secara umum *web scraping* memiliki beberapa langkah yaitu membuat *scraping template,* melakukan eksplorasi navigasi situs, mengotomatisasi navigasi dan mengekstrak informasi serta menyimpannya [7].

1. Membuat *template scraping,* proses ini melakukan observasi terhadap dokumen HTML situs dijadikan objek *scraping*. Cara yang digunakan adalah dengan menganalisis *tag* HTML yang mengapit informasi yang akan di ambil.
2. Eksplorasi navigasi situs, proses ini dilakukan dengan menelusuri navigasi pada situs objek *scraping* untuk ditirukan oleh *web scraper* yang akan dibangun.
3. Mengotomatisasi navigasi dan mengekstraksi informasi berdasarkan hasil dari langkah satu dan dua, aplikasi dibuat untuk mengotomatisasi pengambilan informasi pada situs objek *scraping*.
4. Ekstraksi data dan menyimpan histori, pada proses ini informasi yang diekstrak dari langkah tiga disimpan pada tabel-tabel di dalam *basis data*. Cara kerja teknik *scraping* dapat dilihat pada gambar 2.1.



*Gambar 3.1 Ilustrasi Cara Kerja* Web Scraping *(Maria Rosario B, dkk, 2017)*

Pada gambar 1 merupakan langkah-langkah pada *web scraping.* Langkah pertama dengan menentukan situs yang akan dijadikan objek *scraping*, misalkan situs yang digunakan *google scholar.* Kemudian membuat *template scraping* dengan cara mempelajari HTML dari *google scholar* untuk diambil informasi pada *tag* HTML. Proses berikutnya adalah dengan mengeksplorasi navigasi *google scholar* untuk mempelajari teknik navigasi pada *google scholar* untuk ditirukan pada aplikasi yang dibuat. Kemudian dilakukan otomatisasi informasi yang didapat dari *google scholar* disebut juga sebagai proses ekstraksi informasi. Setelah informasi berhasil di ektraksi maka proses berikutnya adalah melakukan penyimpanan informasi ke dalam basisdata.

## Legalitas *Web Scraping*

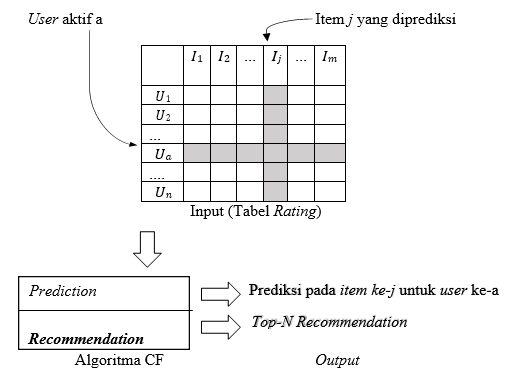
Sampai saat ini masih belum ada peraturan khusus yang mengatur tentang *scraping*. Pada akhir 2019, Pengadilan Amerika Serikat menolak banding dari permintaan LinkedIn kepada hiQ. hiQ adalah perusahaan analitik data yang menggunakan bot otomatis untuk mengumpulkan informasi dari profil LinkedIn. Keputusan itu adalah momen bersejarah dalam era privasi dan regulasi data yang menunjukkan bahwa data apapun yang tersedia untuk umum dan tidak dilindungi hak cipta. Namun, keputusan tersebut tidak memberikan kebebasan bagi *web scraping* untuk menggunakan data yang diperoleh untuk tujuan komersial. Misalnya, *web scraping* akan diizinkan untuk mencari judul video di Youtube, tetapi tidak dapat memposting ulang video Youtube di situsnya sendiri, karena video tersebut memiliki hak cipta. Keputusan ini juga tidak memberikan kebebasan kepada *web scraping* untuk mendapatkan data dari situs yang membutuhkan otentikasi. Misalnya, *web scraping* masuk ke Facebook dan mengunduh data pengguna adalah ilegal. Ada beberapa pedoman yang harus dipatuhi oleh para *web scrapper*,sebagai berikut [8] :

1. Data tidak digunakan untuk penipuan dan penyalahgunaan komputer.
2. Data yang diambil tidak dilindungi hak cipta.
3. Jika data dilindungi hak cipta, konten yang diambil harus mematuhi standar dengan tidak memodifikasi dan mencantumkan pemilik hak cipta.
4. Tindakan *scraping* tidak membebani dan menghambat kinerja layanan situs *web* yang diambil.
5. *Scraping* tidak boleh melanggar ketentuan penggunaan situs yang diambil.
6. *Scraping* tidak boleh mengumpulkan informasi pengguna yang bersifat sensitif.

## *Collaborative filtering*

*Collaborative Filtering* memberikan rekomendasi berdasarkan kumpulan dari pendapat, minat dan ketertarikan beberapa pengguna yang biasanya diberikan dalam bentuk *rating* yang diberikan pengguna kepada suatu *item*. Kualitas rekomendasi yang diberikan dengan menggunakan metode ini sangat bergantung dari peniliaian pengguna lain terhadap suatu *item*. Algoritme *collaborative filtering* memiliki dua hal untuk melakukan sistem rekomendasi sebagai berikut [9] :

1. *Prediction*, merupakan nilai prediksi pada *rating item* yang mungkin disukai oleh *user*.
2. *Recommendation*, memberikan rekomendasi berupa daftar *item* dengan nilai prediksi tertinggi yang mungkin akan disukai pengguna. Dengan catatan bahwa *item* yang direkomendasikan belum pernah dilihat atau di-*rating* oleh user tersebut. Hal ini sering juga disebut *top-N recommendation*.



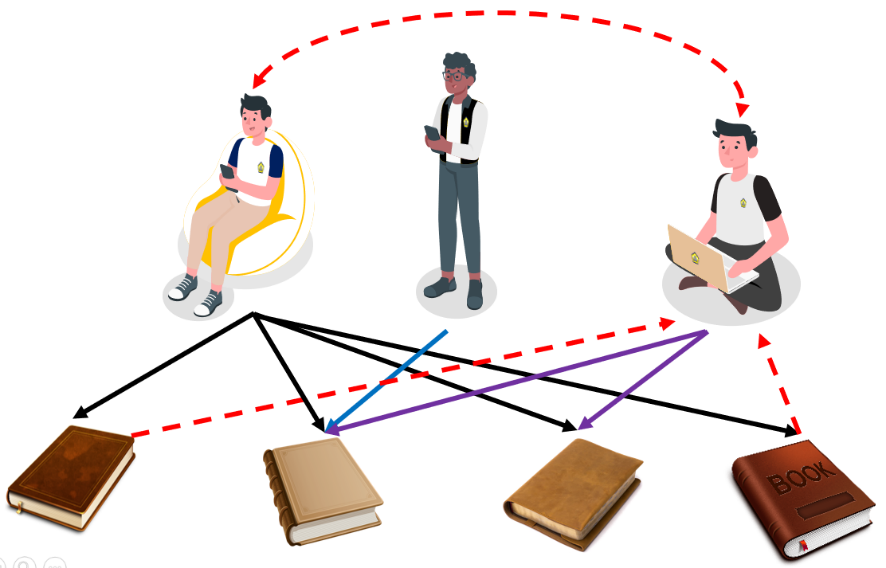
*Gambar 3.2 Proses* Collaborative Filtering[9]

Gambar 2 menunjukkan skema kerja dari sistem rekomendasi yang menggunakan *collaborative filtering* berbentuk *user-item rating* matriks berukuran m x n, m adalah jumlah *user* U = {u1,u2,u3,...,um}, dan n adalah jumlah *item* I = {i1,i2,i3,...in}. Setiap *user (*ui) mempunyai daftar *item* Iui yang merupakan *rating* dari *user* pada setiap *item*.

Untuk memperoleh data *rating* dari *user* yang digunakan dalam sistem rekomendasi, dibedakan menjadi dua cara yaitu [10] :

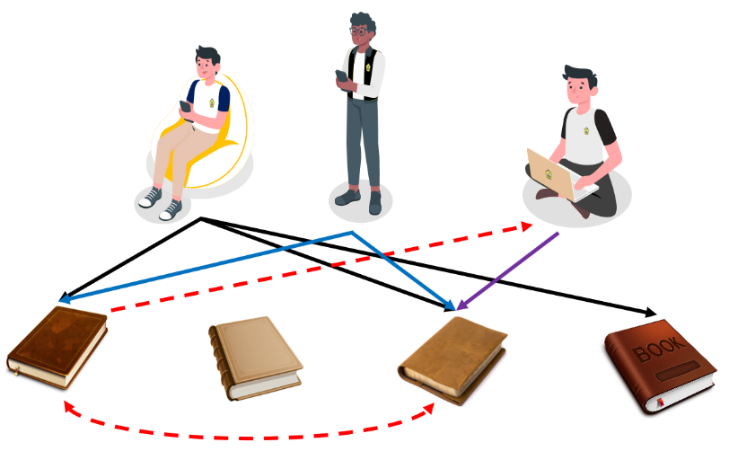
1. Secara *Eksplisit*, yaitu proses pengumpulan data saat user memberikan data secara sadar atau sengaja. Contoh seperti pemberian *rating*, pengisian form tentang data *user*, dan komentar *user* pada suatu *item*.
2. Secara *Implisit*, yaitu proses pengumpulan data saat user tidak menyadari bahwa ia telah memberikan masukan terhadap sistem. Contoh seperti catatan *item* yang diunduh, *item* yang diklik, dan *item* yang dicari *user*.

*Collaborative filtering* pada dasarnya dibagi menjadi dua metode yaitu *user-based collaborative filtering* disebut juga *memory-based*, dan *item-based collaborative filtering* yang disebut juga *model-based*. Pada metode *user based* sistem memberikan rekomendasi kepada *user, item-item* yang disukai atau di-*rating* oleh *user* lain. Misalnya, *user* a menyukai atau me-*rating item* 1 2 dan 3, kemudian *user* b menyukai *item* 1 2 dan 4 maka sistem akan merekomedasikan *item* 3 kepada *user* b dan *item* 4 kepada *user* a. Sedangkan metode *item based* memberikan rekomendasi berdasarkan kemiripan antar *item.* Metode *item based* merupakan metode rekomendasi yang dengan adanya kesamaan antara pemberian *rating* terhadap suatu *item* dengan *item* yang pernah di-*rating user* [10].



Gambar 2. 1 Ilustrasi Metode *User Based Collaborative Filtering*

*User based collaborative filtering* bekerja berdasarkan asumsi bahwa pengguna (pengguna aktif) merupakan bagian dari kelompok yang memiliki kesamaan dengan pengguna lainnya. Dengan kata lain, pengguna yang memiliki kesamaan hubungan (atribut) akan tertarik terhadap *item* yang sama.



Gambar 2. 2 Ilustrasi Metode *Item-Based Collaborative Filtering*

*Item-Based collaborative filtering* adalah algoritme yang bekerja untuk mencari hubungan antar *item* berdasarkan tabel *rating* untuk membentuk sebuah rekomendasi terhadap suatu *item* kepada *user*. *Item-based collaborative filtering* berasumsi bahwa jika mayoritas pengguna memberi penilaian beberapa *item* secara serupa, pengguna yang ditargetkan (pengguna aktif) juga akan memberi penilaian terhadap *item-item* tersebut secara serupa dengan mayoritas pengguna lain.

Menurut penelitian yang dilakukan oleh Christanti (2013) dalam melakukan perbandingan metode *user-based collaborative filtering* dan *item-based collaborative filtering* pada studi kasus sistem rekomendasi tempat wisata untuk wilayah Solo Dan Yogyakarta. Menyimpukan bahwasanya metode *item-based* dapat menghasilkan rekomendasi yang lebih baik dari *user-based.* Hal ini dikarenakan pada perhitungan prediksi dari metode *user-based* banyak ditemukan nilai prediksi di luar *range,* selain itu semakin bertambah *item* semakin meningkat nilai MAE dan NMAE yang dihasilkan dan proses perekomendasian akan semakin lama. Sedangkan pada metode *item-based* tidak ditemukan nilai prediksi di luar *range* dan memiliki tingkat akurasi yang lebih baik dari metode *user-based* [11].

Penelitian ini mengguanakan metode *item-based collaborative filtering*, karena *item-based* membutuhkan perhitungan yang lebih sedikit tetapi mempunyai kualitas yang sama dibandingkan dengan metode *user-based*. Pada pengambilan *rating* dilakukan secara *implisit* untuk mengumpulkan data *rating* tanpa disadari oleh *user*. *Rating* diperoleh secara langsung pada saat pengguna melakukan pencarian, melakukan klik dan mengunduh suatu artikel*.* Dengan banyaknya *rating* yang terkumpul melalui aktifitas *user* akan sangat berguna pada *item-based* *collaborative filtering,* karena metode ini membutuhkan banyak *rating* agar rekomendasi yang dihasilkan semakin berkualitas.

## *Item-Based Collaborative filtering*

Metode *item-based collaborative filtering* merupakan metode rekomendasi dengan menghitung kemiripan *item* yang sudah di *rating* dengan *item* lain, kemudian dipilih sekelompok *item* yang mempunyai nilai kemiripan dengan *item* yang telah di *rating*. *Item* yang telah di *rating* oleh *user* akan menjadi patokan untuk mencari sejumlah *item* lainnya yang berkorelasi dengan *item* yang telah di *rating* *user* [12].

Untuk membuat sistem rekomendasi menggunakan metode *item-based collaborative filtering* ada dua langkah yang harus dilakukan, sebagai berikut [13] :

1. Menghitung *Similarity*

Menghitung kemiripan (*Similarity)* antara satu *item* dengan *item* lain. *Cosine similarity* merupakan metode yang sering digunakan untuk menghitung kesamaan *item*, tetapi metode ini memiliki satu kekurangan. Perbedaan skala *rating* antara berbagai pengguna akan menghasilkan *simarity* yang sangat berbeda. *Adjusted cosine similarity* mengatasi kelemahan dari *cosine similarity*. Metode *adjusted cosine similarity* dapat ditunjukkan oleh persamaan [14].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | (1) |
| Keterangan : | | | |
| *sim(i,j)* | = | Nilai kemiripan antara artikel *i* dan artikel *j*, | |
| *ΣuϵU* | = | Himpunan *user* u yang memberikan *rate* pada artikel *i* dan artikel *j*, | |
|  | = | *Rating user u* pada artikel *i*, | |
|  | = | *Rating user u* pada artikel *j*, | |
|  | = | Rata-rata *rating user u* | |

Dari perhitungan kemiripan tersebut akan mengurutkan *item*-*item* berdasakan nilai *similarity*, *item*-*item* yang mempunyai *similarity* besar akan berada pada urutan atas dan sebalikanya. Hasil dari pada persamaan *cosine similarity* adalah berkisar antara-1 s/d 1. Jika nilai *similarity* antara kedua *item* mendekati +1, maka kedua *item* dianggap dianggap semakin berkolerasi. Sebaliknya, jika nilai *similarty* mendekati -1, maka kedua *item* itu akan semakin tidak berkolerasi.

1. Menghitung prediksi *rating*

Menghitung prediksi *rating* dari *item-item* tersebut dengan membandikan *rating* yang pernah diberikan pengguna pada suatu *item* dengan kemiripan antara *item* tersebut dengan *item* lainnya. Metode yang digunakan merupakan metode *Weighted Sum* dapat ditunjukkan oleh persamaan*.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | (2) |
| Keterangan : | | | |
| *P(u,j)* | = | Prediksi *rating* pada artikel *j* oleh *user u*, | |
| *iϵI* | = | Himpunan artikel yang mirip dengan artikel *j*, | |
|  | = | *Rating user u* pada artikel i*,* | |
|  | = | Nilai kemiripan antara artikel *i* dan artikel *j* | |

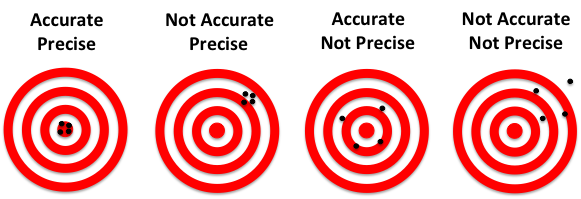
## Skala Penilaian

Skala penilaian adalah salah satu pengukuruan penilaian yang sering digunakan untuk mengetahui pendapat pengguna terhadap suatu penelitian. Peneliti meminta pengguna untuk memberikan penilaian terhadap sekumpulan *item*. Hal ini dilakukan untuk mengetahui seberapa besar ketertarikan responden terhadap *item* tersebut. Dalam penelitian ini menggunakan skala *Rating* (*Rating Scale*).

Menurut [16] *Rating scale* merupakan skala penilaian yang lebih fleksibel, skala penilaian ini tidak hanya untuk mengukur sikap tetapi dapat juga digunakan untuk mengukur persepsi responden terhadap fenomena lingkungan, seperti skala untuk mengukur status sosial, ekonomi, pengetahuan, kemampuan dan lain-lain. penilaian *rating scale* merupakan keputusan peneliti yang menentukan skala mana yang digunakan untuk penilaian yang digunakan, apakah 1 sampai 3, 1 sampai 5, 1 sampai 7, 1 sampai 10 atau 1 sampai 100.

## Pengujian Sistem Rekomendasi

Sebuah sistem pengukuran dapat bernilai akurat dan tepat, atau akurat tetapi tidak tepat, atau tepat tetapi tidak akurat atau tidak tepat dan tidak akurat. Keakuratan metode estimasi kesalahan pengukuran diindikasikan dengan adanya *error* yang kecil. Metode estimasi yang mempunyai *error* lebih kecil dikatakan lebih akurat daripada metode estimasi yang mempunyai *error* lebih besar. Presisi menunjukkan seberapa dekat perbedaan nilai pada saat dilakukan pengulangan pengukuran [15].



* Presisi
  + Akurasi
    - * Presisi
    - Akurasi
      * Presisi
  + Akurasi
* Presisi
* Akurasi

Gambar 2. 3 Ilustrasi Perbedaan antara Akurasi dan Presisi [15]

Gambar 2.3 ilustrasi perbedaan antara akurasi dan presisi dalam pengukuran berulang, diibaratkan dengan anak panah yang menembak target beberapa kali. Akurasi menggambarkan kedekatan panah panah dengan pusat sasaran. Panah yang menancap lebih dekat dengan pusat sasaran dianggap lebih akurat. Semakin dekat sistem pengukuran terhadap nilai yang diterima, sistem dianggap lebih akurat. Sedangkan presisi adalah ukuran kedekatan dari masing-masing anak panah dalam kumpulan tersebut. Semakin menyempit kumpulan anak panah tersebut, sistem dianggap semakin presisi.

Pengujian digunakan untuk mengetahui seberapa akurat hasil rekomendasi yang diberikan sistem kepada *user*. Dapat dilakukan dengan beberapa cara untuk mengevaluasi sistem rekomendasi yaitu dengan menghitung tingkat *error* pada hasil rekomendasi dan juga menghitung nilai presisi dari hasil UAT (*User Acceptance Test*) yang diberikan *user.*

## *Mean Absolute Error* (MAE)

Untuk melihat keefektifan sistem rekomendasi dengan mengukur tingkat akurasi hasil rekomendasi dapat dilakukan dengan melihat nilai *error* pada hasil rekomendasi yang diberikan. *Mean Absolute Error* (MAE) merupakan metode perhitungan nilai *error* yang paling sering digunakan untuk menghitung nilai *error* dari hasil rekomendasi. Nilai MAE merepresentasikan rata – rata kesalahan (*error*) absolut antara nilai prediksi dengan nilai yang sebenarnya [15]. Metode MAE dapat ditunjukkan oleh Persamaan.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | (3) |
| Keterangan : | | | |
| MAE | = | Nilai rata-rata kesalahan hitungan | |
| *n* | = | Jumlah itemyang dihitung, | |
|  | = | Nilai prediksi artikelke i*,* | |
|  | = | Nilai *rating* sebenarnya artikelke i | |

## Pengujian Presisi (Precision)

Presisi juga dapat digunakan untuk evaluasi kualitas hasil dari sekumpulan *item* sebagai hasil rekomendasi untuk *user*. Dari beberapa *item-item* tersebut tentu tidak semua *item* yang relevan atau yang sesuai dengan kebutuhan *user.* Untuk mengetahui kualitas hasil rekomendasi, dapat menggunakan persamaan relevansi presisi yang membadingkan antara *item* yang relevan dengan total *item* yang dihasilkan atau yang direkomendasikan kepada *user* [15]*.* Metode relevansi presisi dapat ditunjukkan oleh Persamaan.

|  |  |
| --- | --- |
|  | (3) |

*Relevant items retrieved* adalah jumlah *item* relevan yang direkomendasikan sedangkan *retrieved items* adalah jumlah total *item* yang direkomendasikan

## *User Acceptance Test* (UAT)

*User Acceptance Test* (UAT) adalah pengujian hasil rekomendasi yang dilakukan oleh *user* bahwa seberapa besar hasil rekomendasi dapat diterima pengguna. Pada pengujian ini dilakukan dengan menyebarkan kuesioner yang berisikan pertanyaan mengenai kesesuaian antara hasil rekomendasi dengan keinginan pengguna. Untuk menguji hasil rekomendasi yaitu jumlah *item* yang sesuai dengan selera pengguna berbanding dengan jumlah dokumen yang ditampilkan. Pengujian ini juga dapat dilakukan dengan persamaan persisi. Hasil akhir dari kuesioner dapat memperkuat kualitas dari sistem rekomendasi [15].

## Flask

Flask adalah sebuah *web framework* yang ditulis dengan bahasa Python dan tergolong sebagai jenis *microframework*. Flask berfungsi sebagai kerangka kerja aplikasi dan tampilan dari suatu web. Dengan menggunakan Flask dan bahasa Python, pengembang dapat membuat sebuah web yang terstruktur dan dapat mengatur *behaviour* suatu web dengan lebih mudah [17].

## Penelitian Terkait

Penelitian-penelitian sebelumnya yang menjadi acuan/referensi dalam penelitian ini, antara lain :

* + - * 1. *A Survey on Collaborative Filtering Based Recommendation System,* 2016 (G. Suganeshwari and S.P. Syed Ibrahim).

Penelitian ini memuat analisis dan survei mengenai metode pada sistem rekomendasi yang membandingkan 3 jenis metode yaitu *content based, collaborative filtering* dan *hybrid method*. Ketiga metode memiliki kelebihan dan batasan tersendiri. Pada artikel ini diusulkan untuk menggabungkan LARS dan KARS. LARS yang merupakan *item based collaborative filtering* yang mengurangi kompleksitas dengan mengurangi jumlah *user* dan *item* dengan mempertimbangkan sifat spasial mereka. KARS adalah *user based collaborative filtering* yang menangani data (*review*) yang tidak terstruktur yang diberikan oleh *user*. Hasil dari menggabungkan kedua metode telah terbukti bahwa rekomendasi yang dihasikan lebih akurat daripada teknik aslinya [18].

* + - * 1. *Scientific Paper Recommendation: A Survey,* 2019 (Xiaomei Bai, Mengyang Wang, Ivan Lee, Zhuo Yang, Xiangjie Kong dan Feng Xia)

Penelitian ini memuat analisis teknik rekomendasi dapat dibagi menjadi empat kategori utama: *content-based filtering* (CBF), *collaborative filtering* (CF), *graph-based method* (GB) dan *hybrid method*. Dalam paper ini, membahas mengenai metode rekomendasi dan metode evaluasi sistem rekomendasi artikel ilmiah. Serta beberapa masalah dan tantangan dari sistem rekomendasi, sepeti *cold start, sparsity, scalability, privacy, serendipity* dan standar data.

* + - * 1. *Sentiment analysis of tokopedia application review to service product recommender system using neural collaborative filtering for marketplace in Indonesia*, 2020 (Restu Meifitrah, Irfan Darmawan, dan Oktariani Nurul Pratiwi)

Penelitian ini melakukan pengklasifikasian pada *review* atau ulasan aplikasi Tokopedia di *Google Play* dengan menggunakan metode *Web scraping* dan *Naive Bayes.* Dalam paper ini, membahas mengenai komentar yang diberikan pengguna terhadap aplikasi berupa komentar positif, negatif dan netral. Terdapat perbedaan yang cukup besar antara data hasil perhitungan manual dan hasil prediksi sistem. Perbedaan terbesar pada data dengan sentimen netral. Ini terjadi karena salah tafsir kalimat, misalnya dalam kalimat netral ada kata-kata positif, menyebabkan sistem cenderung membaca data positif.

* + - * 1. *Sistem Rekomendasi Artikel Ilmiah Berbasis Web Menggunakan Content-based Learning dan Collaborative Filtering*, 2018 (Fitrianti, Betharia Sri Fachurrozi, Muhammad Yusliani, Novi)

Penelitian ini melakukan perhitungan nilai kesalahan rata-rata absolut (*Mean Absolute Error*) terhadap perbandingan antara *rating* yang diprediksi dan *rating* sebenarnya dari 100 sampel dengan menggunakan metode *Content-based Learning* dan *Collaborative Filtering.* Hasil pengujian menggunakan *Recall* pada metode *Content-based Learning* adalah 91,6%. Hasil pengujian menggunakan *Mean Absolute Error* *Hybrid Recommendation* (*Content-based Learning* dan *Collaborative Filtering*) adalah 0.85

* + - * 1. *Keyword-Based Collaborative Filter Recommendation System Using Scraping*, 2018 (Young Jun Kim, Jeong Min Park, Sung Taek Chung dan Jeong Joan Kim)

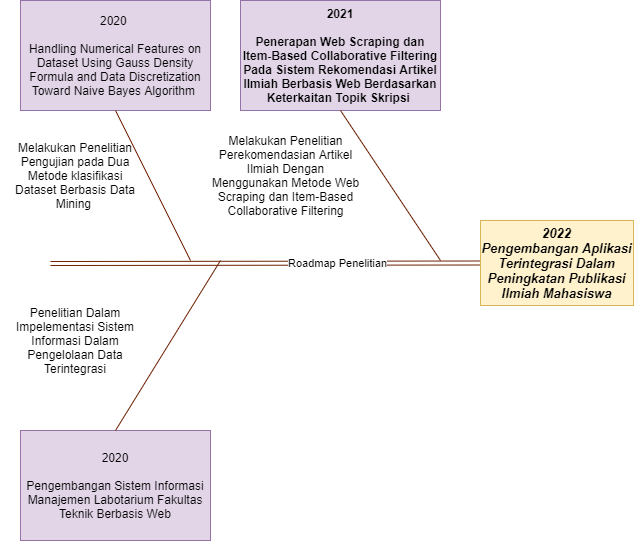
Penelitian ini melakukan melakukan ekstraksi data berdasarkan kata kunci pada protal film korea dengan metode *Scraping*, film yang didapat akan di hitung berdasarkan kesamaan, tetangga terdekat dari setiap film dengan metode *Collaborative Filter*. Penelitian ini melakukan pengkelompokan kata berdasarkan kata kunci pencarian, dari kelompok kata ini, akan dibentuk grup kata untuk merekomendasikan film serupa. Kata kunci yang diekstrak memiliki peran penting dalam menilai kesamaan film.

# BAB III

# PETA JALAN PENELIATAN DAN LUARAN PENELITIAN



## Peta Jalan Peneltian

**

Gambar 3.1 Diagram *fish bone* jalannya penelitian

Peta jalannya penelitian ini telah dimulai sejak tahun 2020 dilakukan penelitian dalam pengujian dua metode klasifikasi dataset berbasis data mining. Lebih lanjut penelitian ini berkembang menjadi suatu sistem informasi dalam pengelolaan data terintegrasi, diharapkan dapat memberikan informasi lebih akurat dalam pengelolaan data. Penelitian ini akan terus berlanjut untuk dalam melakukan perekomendasian data informasi. Diharapkan dapat membantu meningkatkan publikasi ilmiah mahasiswa. Tujuan yang ingin dicapai dalam penelitian berjangka ini adalah dengan adanya perekomendasian artikel ilmiah dapat membantu mahasiswa dalam pencarian referensi penelitian, memungkinkan mahasiswa untuk dapat mengefisiensikan waktu dan meningkatkan kualitas dari penelitian yang dilakukan.

## Luaran Penelitian

Luran pada penelitian ini adalah sistem rekomendasi artikel ilmiah berbasis *web* berdasarkan keterkaitan topik skripsi mahasiswa Universitas Bengkulu dengan menggunakan metode *web scraping* dan *item-based collaborative filtering.*

# BAB IV

# METODE PENELITIAN



## Input Data

Pada tahapan ini akan dilaksanakan perencanaan, penentuan tujuan, pengumpulan data, dan memasukkannya ke dalam sistem yang akan dibuat, data yang dibutuhkan dapat dijabarkan secara umum sebagai berikut:

1. Mahasiswa untuk mengakses sistem dan menyimpan tingkah laku atau aktivitas mahasiswa terhadap sistem. Aktivitas mahasiswa digunakan sebagai *rating* yang menjadi pedoman dalam perekomendasian artikel. Data ini mencakup nama, Nomor Pokok Mahasiswa (NPM), email, kata sandi.
2. Artikel didapatkan dari hasil pencarian mahasiswa kemudian di *scraping* dari *google scholar*. Informasi yang didapatkan judul, abstrak, penulis, penerbit, tahun terbit, kutipan, link, kode dan dokumen artikel dengan format pdf.
3. Pencarian berisi catatan artikel yang pernah dicari atau diklik oleh *user* ketika berinteraksi dengan sistem beserta jumlah klik yang telah dilakukannya.
4. Unduhan berisikan catatan unduhan *user* terhadap suatu artikel.
5. *Rating* berisi nilai *rate* yang diberi *user* terhadap artikel. Nilai ini diperoleh dari aktivitas *user* baik berupa pencarian atau klik dan juga unduhan dokumen*.*
6. *Similarity* berisi data kemiripan yang diperoleh dari perhitungan rumus *adjusted cosine similarty.* Data ini dipergunakan untuk tahap penentuan prediksi dimana data tersebut diurutkan dari yang besar ke yang kecil, kemudian diambil sebanyak *n* data untuk dilakukan tahap perhitungan prediksi.
7. Prediksi berisi data artikel yang belum pernah di-*rating user* beserta nilai prediksi *rating-*nya. Data ini diurutkan dari besar ke kecil. Kemudian sebanyak *n* data diambil untuk direkomendasikan ke *user.*

## Pengolahan Data

Pada tahap ini akan dilakukan klasifikasi, kompilasi, dan geoprocessing guna membangun peta yang menghasilkan informasi dengan lebih baik. Pada bagian ini akan mendapatkan peta dasar yang akan dianalisa dan diproses pada tahap selanjutnya.

## Analisis Data

Pada tahapan ini dilakukan perhitungan kemiripan dan perhitungan prediksi pada nilai *rating* pengguna terhadap *item*. Dengan *adjusted cosine similarity* untuk menghitung nilai *similarty* dan *Weighted Sum* menghitung nilai prediksi *rating*.

Pada tahapan juga dilakukan analisa keruangan, dalam metode *item-based collaborative filtering*, terdapat 2 permasalahan utama, yaitu *cold start problem* dan *data sparsity. Cold start problem* adalah keadaan di mana *item* atau *user* baru saja dimasukkan ke dalam sistem, sehingga sistem perekomendasi tidak memiliki pengetahuan atau informasi untuk menghasilkan rekomendasi yang akurat. Dalam sistem ini, *cold start problem* diatasi dengan penambahan fitur lain agar *user* baru tetap dapat me-*rating* atau *item* baru tetap dapat di-*rating* meskipun sistem belum dapat merekomendasikannya ke *user,* salah satunya dengan fitur pencarian. Sedangkan *data sparsity* adalah keadaan di mana banyak data yang kosong untuk setiap kolom dalam table matriks *rating user-item*. Untuk teknik *item-based*, hanya kolom matriks penilaian atau *rating* yang terisi yang akan diproses dalam perhitungan prediksi.

## Output

Pada tahap ini merupakan fase akhir di mana ini akan berkaitan dengan penyajian hasil analisa yang telah dilakukan. Pada bagian ini akan diperoleh hasil rekomendasi artikel berdasarkan perdasarkan perhitungan *similarity* dan perhitungan prediksi yang didapatkan dari metode *item-based* *collaborative filtering*. Kemudian dilakukan perbandingan sitasi atau kutipan (pada artikel yang telah *scraping* di *google scholar*) dengan nilai-nilai prediksi yang dihasilkan. Proses terakhir ini tidak termasuk dalam perhitungan *item-based* *collaborative filtering,* namun proses tersebut dilakukan dengan tujuan untuk mendapatkan hasil rekomendasi yang lebih akurat dengan menggabungkan tingkat kemiripan artikel berdasarkan *raintg* pengguna pada *google scholar* dan pengguna sistem yang akan dibangun.

## Alat dan Bahan

Dalam melaksanakan penelitian ini, dibutuhkan beberapa sarana pendukung yang merupakan perangkat-perangkat yang di butuhkan untuk penerapan *web scraping* dan *item-based* *collaborative filtering* pada sistem rekomendasi artikel ilmiah berbasis web berdasarkan keterkaitan topik skripsi. Untuk melaksanakan penelitian ini, peneliti akan menggunakan alat dan bahan sebagai berikut:

|  |  |
| --- | --- |
| Alat: | Bahan: |
| * Komputer * Processor AMD Ryzen™ 5 3500U * RAM 8 GB | * Beautiful soup 4 * Framewrok Flask * Browser * Sistem Operasi Windows 10 * Python 3.9.1 * Visual Studio Code |

## Lokasi Penelitian

Lokasi penelitian akan dilaksanakan di Fakultas Teknik Universitas Bengkulu.

## Target Luaran

Rencana capaianluaran penelitian ini adalah sebagai berikut:

1. Sistem rekomendasi artikel ilmiah berbasis *web* berdasarkan keterkaitan topik skripsi mahasiswa Universitas Bengkulu dengan menggunakan metode *web scraping* dan *item-based collaborative filtering.* Adapun fitur-fiturnya sebagai berikut :
   1. Pencarian meliputi pencarian artikel ilmiah (jurnal, buku dan tesis atau skripsi), penulis, sumber pustaka, tahun dan penerbit.
   2. Perekomendasian artikel imiah.
   3. Pencarian artikel dapat dilakukan dalam Bahasa Indonesia dan Bahasa Inggris.
   4. Unduhan sitasi yang memudahkan pengguna mengambil artikel (sitasi) berbentuk *BibTeX, EndNote, RefMan* dan *RefWorks*.
   5. Unduhan artikel ilmiah secara utuh (*fulltext)* dalam format PDF.
   6. Penelusuran artikel berdasarkan dirujuk dan artikel terkait pada *google scholar* untuk memudahkan pengguna untuk mencari artikel yang berhubungan.
   7. Memuat informasi atau detail artikel yang di *scraping* pada *google scholar.*
2. Publikasi jurnal ilmiah nasional terakreditasi sinta 5 di Journal Scientific and Applied Informatics (JSAI), Jurusan Informatika, Universitas Muhammadiyah Bengkulu ISSN 2614-3062.

Laman website: <http://jurnal.umb.ac.id/index.php/JSAI/index>



# BAB V

# JADWAL PELAKSANAAN

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kegiatan** | **Bulan ke- (selama 7 bulan)** | | | | | | | **Indikator Capaian** |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** |  |
|  | Penyusunan Proposal |  |  |  |  |  |  |  | Proposal |
|  | Pengumpulan Data |  |  |  |  |  |  |  | Data Alih Fungsi |
|  | Pengumpulan dan Analisis *Data* |  |  |  |  |  |  |  | Analisis data |
|  | Pembuatan Sistem/Program |  |  |  |  |  |  |  | Sistem rekomendasi artikel ilmiah |
|  | Penyusunan Laporan Penelitian |  |  |  |  |  |  |  | Laporan |
|  | Publikasi |  |  |  |  |  |  |  | Karya Ilmiah |

# BAB VI

# REANCANAGAN ANGGARAN BIAYA DAN JUSTIFIKASI ANGGARAN

|  |  |  |
| --- | --- | --- |
| **No.** | **Kegiatan dan Bahan** | **Jumlah** |
|  | Honorarium Tenaga Survey 5 orang x 10 hari x 80.000/hari | Rp 4.000.000 |
|  | Honorarium Pengolah Data 5 orang x 5 hari x 80.000/hari | Rp 2.000.000 |
|  | Konsumsi survey lapangan 5 x 10 x 20.000 | Rp 1000.000 |
|  | Konsumsi pengolah data 5 x 5 x 20.000 | Rp 500.000 |
|  | SSD 512GB Midas Force Solid State Drive 512 Gb Midasforce | Rp 820.000 |
|  | RAM Laptop Kingston DDR4 8GB 2400 MHz 19200 | Rp 575.000 |
|  | Bahan Habis Pakai   1. Kertas   4 rim x Rp. 47.500 = Rp. 190.000   1. Alat tulis pena, pensil = Rp. 100.000 2. Materai   10 x Rp. 12000 = Rp. 120.000   1. Tinta   4 warna x Rp. 25.000 = Rp 100.000   1. Cartridge 4 xRp150.000 = 600.000 | Rp. 1.110.000 |
|  | Biaya cetak dan jilid laporan 15 rangkap x Rp 50.000 | Rp 750.000 |
|  | Biaya Publikasi | Rp 700.000 |
|  | Biaya Berlangganan Grammarly Premium dan turnitin | Rp 500.000 |
|  |  |  |
| **Total Pembiayaan** | | **Rp 11.955.000** |



# DAFTAR PUSTAKA

[1] D. D. A. Yani, H. S. Pratiwi, and H. Muhardi, “Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace,” *J. Sist. dan Teknol. Inf.*, vol. 7, no. 4, p. 257, 2019, doi: 10.26418/justin.v7i4.30930.

[2] I. W. Jepriana and S. Hanief, “Analisis dan Implementasi Metode Item-based Collaborative Filtering untuk Sistem Rekomendasi Konsentrasi di STMIK Stikom Bali,” *J. Nas. Pendidik. Tek. Inform. JANAPATI*, vol. 9, no. 2, pp. 171–180, 2020.

[3] Muhammad Fadelillah, “SISTEM REKOMENDASI PENCARIAN ARTIKEL JURNAL INDONESIA MENGGUNAKAN METODE JACCARD’S COEFFICIENT,” UNISSULA, 2017.

[4] B. Panduan and P. Skripsi, *Buku panduan penyusunan skripsi*. FAKULTAS EKONOMI DAN BISNIS UNIVERSITAS BENGKULU, 2018.

[5] N. K. Susrini, *Google: Mesin Pencari yang Ditakuti Raksasa Microsoft*. Bentang-B First, 2009.

[6] V. Mitra, H. Sujaini, and A. Negara, “Rancang Bangun Aplikasi Web Scraping untuk Korpus Paralel Indonesia - Inggris dengan Metode HTML DOM,” *J. Sist. dan Teknol. Inf.*, vol. 5, no. 1, pp. 1–6, 2017.

[7] F. Maria Rosario B, Yovi Pratama, “Penerapan Web Scraping Pada Website Company Profile,” *Kntia*, vol. 4, no. 4, pp. 37–43, 2017.

[8] V. A. Flores, P. A. Permatasari, and L. Jasa, “Penerapan Web Scraping Sebagai Media Pencarian dan Menyimpan Artikel Ilmiah Secara Otomatis Berdasarkan Keyword,” *Maj. Ilm. Teknol. Elektro*, vol. 19, no. 2, p. 157, 2020, doi: 10.24843/mite.2020.v19i02.p06.

[9] L. Sabani, “PERPUSTAKAAN MENGGUNAKAN METODE ITEM-BASED COLLABORATIVE FILTERING,” Universitas Islam Negeri Maulana Malik Ibrahim Malang, 2020.

[10] A. E. Wijaya and D. Alfian, “Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based Filtering,” *J. Comput. Bisnis*, vol. 12, no. 1, pp. 11–27, 2018.

[11] Y. E. Christanti, “Perbandingan metode user-item based dan item-based collaborative filtering pada studi kasus sistem rekomendasi tempat wisata untuk wilayah Solo Dan Yogyakarta,” UNS, 2013.

[12] Y. V. L. Jaja, B. Susanto, and L. R. Sasongko, “Penerapan Meode Item-Based Collaborative Filtering Umtuk Sistem Rekomendasi Data MovieLens,” *J. Mat. dan Apl.*, 2020.

[13] Y. Setiawan, A. Nurwanto, and A. Erlansari, “Implementasi Item Based Collaborative Filtering Dalam Pemberian Rekomendasi Agenda Wisata Berbasis Android,” *Pseudocode*, vol. 6, no. 1, pp. 13–20, 2019, doi: 10.33369/pseudocode.6.1.13-20.

[14] A. Pamuji, “Sistem Rekomendasi Kredit Perumahan Rakyat Dengan Menggunakan Metode Collaborative Filtering,” *Fakt. Exacta*, vol. 10, no. 1, pp. 1–9, 2017.

[15] I. S. Wahyudi, “Mesin Rekomendasi Film Menggunakan Metode Kemiripan Genre Berbasis Collaborative Filtering,” Institut Teknologi Sepuluh Nopember Surabaya, 2017.

[16] J. R. Wimmer, Roger D and Dominick, *Mass media research*. Cengage learning, 2013.

[17] R. Irsyad, “Penggunaan Python Web Framework Flask Untuk Pemula,” 2018, doi: 10.31219/osf.io/t7u5r.

[18] G. S. and S. P. S. Ibrahim, “A Survey on Collaborative Filtering Based Recommendation System,” *Smart Innov. Syst. Technol.*, vol. 49, pp. v–vii, 2016, doi: 10.1007/978-3-319-30348-2.