

BAB 3

METODOLOGI DAN PERANCANGAN PENELITIAN

3.1 Metodologi Penelitian

Dalam penelitian ini, terdapat beberapa metodologi dan perancangan yang dilakukan untuk memenuhi metodologi dan perancangan sistem. Berikut merupakan tahap - tahap yang dilakukan:

1. Identifikasi Masalah, Kurangnya personalisasi dalam pembelajaran, terutama dalam penyediaan soal quiz yang relevan dengan kelemahan siswa pada mata kuliah pertanian. Banyak siswa menghadapi kesulitan menguasai topik tertentu karena soal latihan yang diberikan tidak selalu sesuai dengan kebutuhan belajar mereka. Penelitian ini bertujuan mengembangkan sistem rekomendasi soal quiz berbasis *content-based filtering* yang dapat mengidentifikasi kelemahan siswa berdasarkan hasil kuis sebelumnya dan merekomendasikan soal yang relevan untuk meningkatkan pemahaman konseptual dan kesiapan akademik.
2. Studi Literatur, dilakukan dengan mempelajari dan memahami beberapa sumber yang berkaitan dengan topik. Untuk meningkatkan pengetahuan penulis terhadap *Intelligent tutoring Systems* dan *Content-Based Filtering*.
3. Analisa Kebutuhan, tahap ini dilakukan untuk menentukan komponen yang diperlukan untuk modul rekomendasi. sistem memerlukan dataset berupa bank soal ata hasil kuis mahasiswa. Algoritma utama yang digunakan adalah *Cosine Similarity* untuk mengukur kesamaan antara vektor kelemahan mahasiswa dan vektor bab. sistem harus mampu mengidentifikasi bab terlemah siswa, merekomendasikan soal yang sesuai, dan menyimpan hasil rekomendasi dalam format CSV. Untuk evaluasi, menggunakan *confusion matrix* untuk menguji performa sistem.
4. Perancangan, sistem dirancang untuk merekomendasikan soal quiz yang relevan dengan kelemahan siswa. Input sistem meliputi data hasil kuis (vektor jawaban benar/salah) dan bank soal. Pemrosesan dilakukan dengan mengonversi jawaban siswa ke vektor kelemahan (1 untuk salah, 0 untuk benar) per bab, membuat vektor referensi bab, menghitung *Cosine Similarity* untuk mengidentifikasi bab terlemah. Alur kerja

sistem mencakup pengambilan data, pembuatan vektor, perhitungan Cosine Similarity, pemilihan soal, dan penyimpanan hasil dalam CSV.

5. Pengumpulan Dataset, Dataset dikumpulkan berupa bank soal dalam format CSV. Dataset yang digunakan dalam penelitian ini berasal dari buku berjudul *Teknik Pemuliaan Tanaman* karya Dr. Muhamad Syukur, SP., M.Si. Data hasil kuis mahasiswa diperoleh dari simulasi kuis, mencakup jawaban siswa, jawaban benar, skor, dan timestamp. Jawaban mahasiswa dikonversi ke vektor kelemahan per bab (misalnya, [0, 0, 6, 1, 0, ...] untuk jumlah jawaban salah).
6. Implementasi, sistem diimplementasikan menggunakan Python. Data dimuat dan dibersihkan dengan pandas, fitur soal dibobotkan menggunakan TF-IDF dari scikit-learn untuk menonjolkan karakteristik unik, dan *Cosine Similarity* dihitung untuk mengidentifikasi bab terlemah. Fungsi `recommended_questions_by_chapter` memilih soal dari bank soal berdasarkan bab terlemah, dengan jumlah soal sesuai input pengguna. Rekomendasi ditampilkan dalam format terstruktur (nomor, teks soal, opsi, jawaban benar) melalui fungsi `display_recommendations`. Hasil rekomendasi dan kuis disimpan dalam file CSV berdasarkan NIM siswa dan timestamp. Visualisasi *confusion matrix* dibuat dengan seaborn dan matplotlib untuk analisis performa.
7. Pengujian Sistem, sistem diuji menggunakan dataset simulasi kuis dengan skala terbatas. Prosedur pengujian meliputi memasukkan data kuis, menghasilkan vektor kelemahan, menghitung *Cosine Similarity*, merekomendasikan soal, dan menyimpan hasil. Pengujian dilakukan pada Jupyter Notebook atau PyCharm, dengan kasus uji meliputi validitas data, keakuratan rekomendasi, dan performa waktu pemrosesan. Tujuan pengujian adalah memastikan bahwa soal yang direkomendasikan sesuai dengan bab terlemah dan sistem berjalan tanpa error.
8. Evaluasi Sistem, sistem dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Fungsi `evaluate_recommendation` menghitung True Positive (soal relevan yang direkomendasikan), False Positive (soal tidak relevan yang direkomendasikan), True Negative (soal tidak relevan yang tidak direkomendasikan), dan False Negative (soal relevan yang tidak

direkomendasikan). *Confusion matrix* divisualisasikan sebagai heatmap untuk analisis performa.

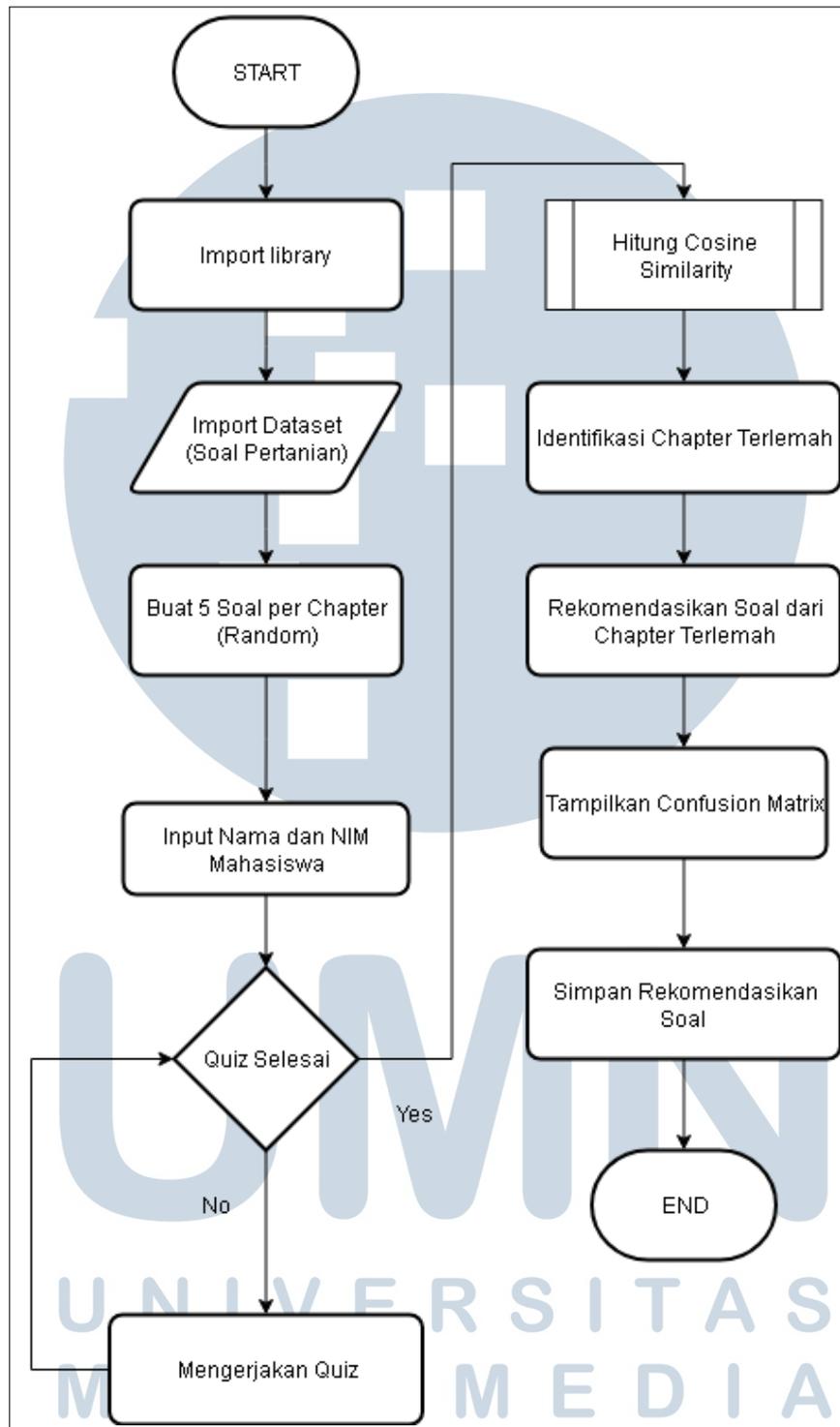
3.2 Perancangan

Perancangan sistem implementasi algoritma *Content-Based Filtering* untuk merekomendasikan soal sesuai dengan kelemahan mahasiswa yang akan dijabarkan dalam *flowchart* yang terdiri dari *flowchart* utama dan *flowchart Cosine Similarity*.

3.2.1 *Flowchart* Utama

Gambar 3.1 merupakan alur utama dari pengimplementasian sistem *Content-Based Filtering* untuk merekomendasikan soal sesuai dengan kelemahan mahasiswa. Proses dimulai dengan mengimpor berbagai pustaka Python yang dibutuhkan untuk pengolahan data dan perhitungan algoritma, seperti *pandas*, *scikit-learn*, dan *seaborn*. Selanjutnya, sistem memuat dataset berupa bank soal pertanian dari file CSV, yang terdiri dari kumpulan soal lengkap dengan bab, opsi jawaban, dan kunci jawaban. Setelah itu, sistem akan secara acak memilih 5 soal dari setiap bab untuk membentuk satu set kuis. Mahasiswa kemudian diminta menginput nama dan NIM untuk memulai pengerjaan kuis.

Selama kuis berlangsung, jawaban mahasiswa akan dicatat dan disimpan. Setelah mahasiswa menyelesaikan seluruh soal, sistem akan menghitung cosine similarity antara vektor kelemahan mahasiswa (berdasarkan jawaban salah per bab) dan vektor konten soal. Hasil perhitungan ini digunakan untuk mengidentifikasi bab (chapter) terlemah yang menunjukkan pemahaman paling rendah. Berdasarkan bab tersebut, sistem akan merekomendasikan soal-soal tambahan yang relevan untuk latihan. Sistem kemudian mengevaluasi performa rekomendasi menggunakan metrik akurasi, presisi, recall, dan F1-score, dan menampilkannya dalam bentuk *confusion matrix*. Terakhir, soal-soal yang direkomendasikan disimpan dalam file dan proses berakhir. *Flowchart* ini menunjukkan alur kerja sistem secara terstruktur dan mendukung personalisasi pembelajaran berbasis data.



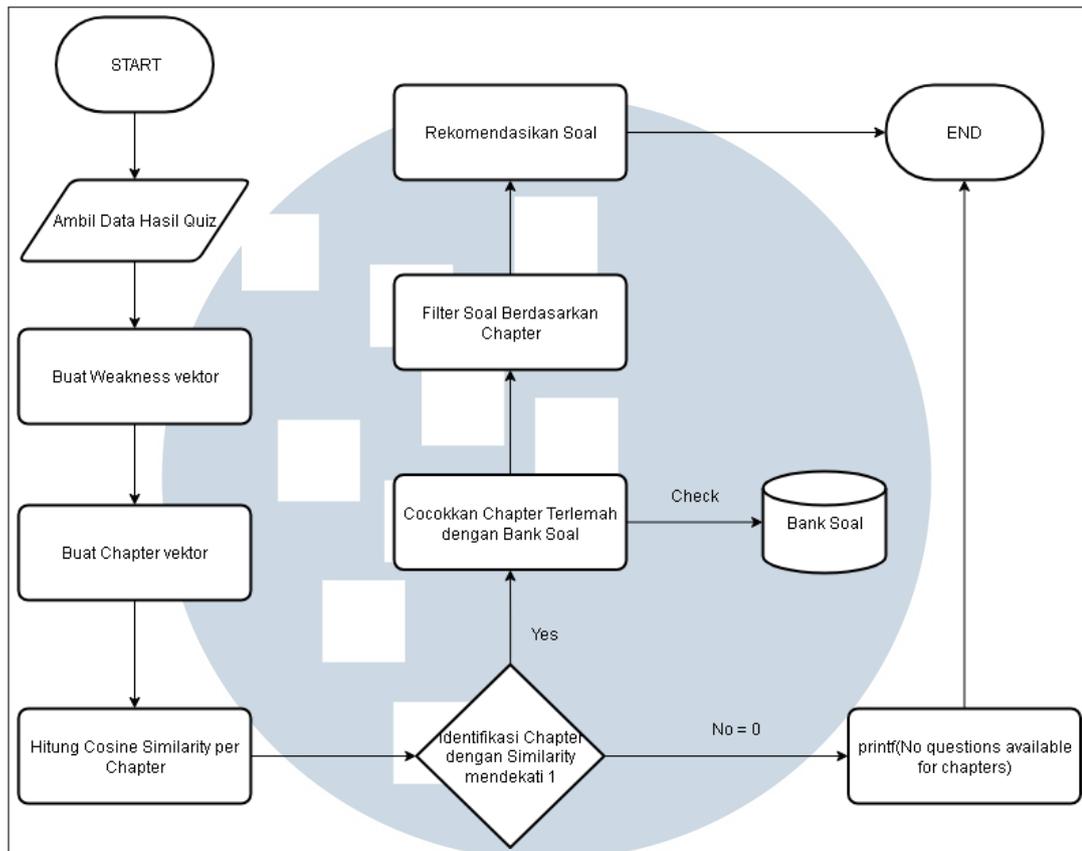
Gambar 3.1. Flowchart Utama

3.2.2 *Flowchart Cosine Similarity*

Gambar 3.2 menggambarkan proses sistem Content-Based Filtering (CBF) dalam merekomendasikan soal kepada mahasiswa berdasarkan kelemahan mereka setelah mengerjakan kuis. Proses dimulai dengan pengambilan data hasil kuis yang mencakup jawaban mahasiswa terhadap setiap soal. Berdasarkan data tersebut, sistem membuat weakness vector, yaitu representasi numerik dari kelemahan mahasiswa per bab atau chapter. Kemudian sistem juga membentuk chapter vector, yaitu representasi dari setiap bab dalam bentuk vektor biner. Setelah kedua vektor dibuat, sistem menghitung nilai cosine similarity antara weakness vector mahasiswa dan setiap chapter vector untuk mengetahui seberapa besar kemiripan kelemahan mahasiswa terhadap masing-masing bab.

Langkah selanjutnya adalah mengidentifikasi chapter yang memiliki nilai kemiripan (similarity) tertinggi atau mendekati angka 1, yang menunjukkan bahwa bab tersebut merupakan kelemahan utama mahasiswa. Setelah chapter terlemah diidentifikasi, sistem mencocokkan bab tersebut dengan soal yang tersedia di dalam bank soal. Jika soal untuk chapter tersebut tersedia, sistem akan memfilternya dan hanya menampilkan soal yang sesuai untuk direkomendasikan. Namun, jika tidak ada soal yang cocok ditemukan, sistem akan menampilkan pesan “no questions available for chapters” sebagai penanda bahwa tidak ada soal yang bisa diberikan. Jika soal berhasil difilter, sistem menampilkan dan merekomendasikan soal tersebut kepada mahasiswa, lalu proses berakhir. Flowchart ini menjelaskan secara rinci bagaimana sistem secara otomatis menyesuaikan soal latihan berdasarkan kelemahan individu mahasiswa dengan pendekatan berbasis konten.

U M W N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A



Gambar 3.2. Flowchart Cosine Similarity

3.2.3 Sumber Dataset: Buku Teknik Pemuliaan Tanaman

Buku Teknik Pemuliaan Tanaman [19] merupakan salah satu referensi utama dalam bidang ilmu pertanian yang membahas secara komprehensif prinsip, metode, dan aplikasi pemuliaan tanaman. Fokus utama dari buku ini adalah bagaimana menghasilkan varietas tanaman unggul melalui pendekatan ilmiah, seperti seleksi massa, seleksi galur murni, hibridisasi, serta penerapan teknologi modern seperti kultur jaringan dan penanda molekuler (*molecular markers*). Materi-materi tersebut sangat relevan untuk dikuasai oleh mahasiswa pertanian, karena menjadi dasar dalam pengembangan teknologi budidaya tanaman yang berkelanjutan dan adaptif terhadap perubahan lingkungan.

Dalam penelitian ini, buku tersebut digunakan sebagai sumber utama dalam penyusunan dataset soal yang mencakup berbagai topik penting dalam pemuliaan tanaman. Penyusunan soal dilakukan dengan bantuan teknologi *Artificial Intelligence* (AI), yaitu menggunakan platform Studio AI. Pendekatan ini dipilih untuk menghasilkan soal-soal pilihan ganda yang sistematis, terstruktur, dan sesuai

dengan cakupan materi yang terdapat dalam buku referensi.

Dataset yang dihasilkan dari proses ini digunakan sebagai basis pengujian sistem rekomendasi soal yang dikembangkan menggunakan pendekatan *Content-Based Filtering* (CBF) dalam kerangka kerja *Intelligent Tutoring System* (ITS). Dengan merujuk pada materi dari buku ini, sistem rekomendasi diharapkan mampu menyajikan soal-soal yang tidak hanya sesuai dengan kelemahan mahasiswa berdasarkan hasil kuis, tetapi juga relevan secara akademik terhadap konteks bidang pertanian.

Secara teknis, dataset disusun dalam format yang memuat atribut sebagai berikut: `question_id`, `question_text`, `option_a`, `option_b`, `option_c`, `option_d`, `correct_answer`, `chapter_number`, `chapter_title`, `topic`, `sub_topic`, `keywords`, dan `chapter_quiz`. Dataset ini terdiri dari 26 bab, masing-masing memuat 20 soal, sehingga total terdapat 520 soal. Seluruh bab kemudian dikelompokkan ke dalam empat kuis, dengan pembagian sebagai berikut: kuis 1 dan kuis 4 masing-masing mencakup 7 bab, sementara kuis 2 dan kuis 3 mencakup masing-masing 6 bab. Untuk keperluan implementasi dan pengujian sistem, data yang digunakan difokuskan pada soal-soal dari kuis 1.

