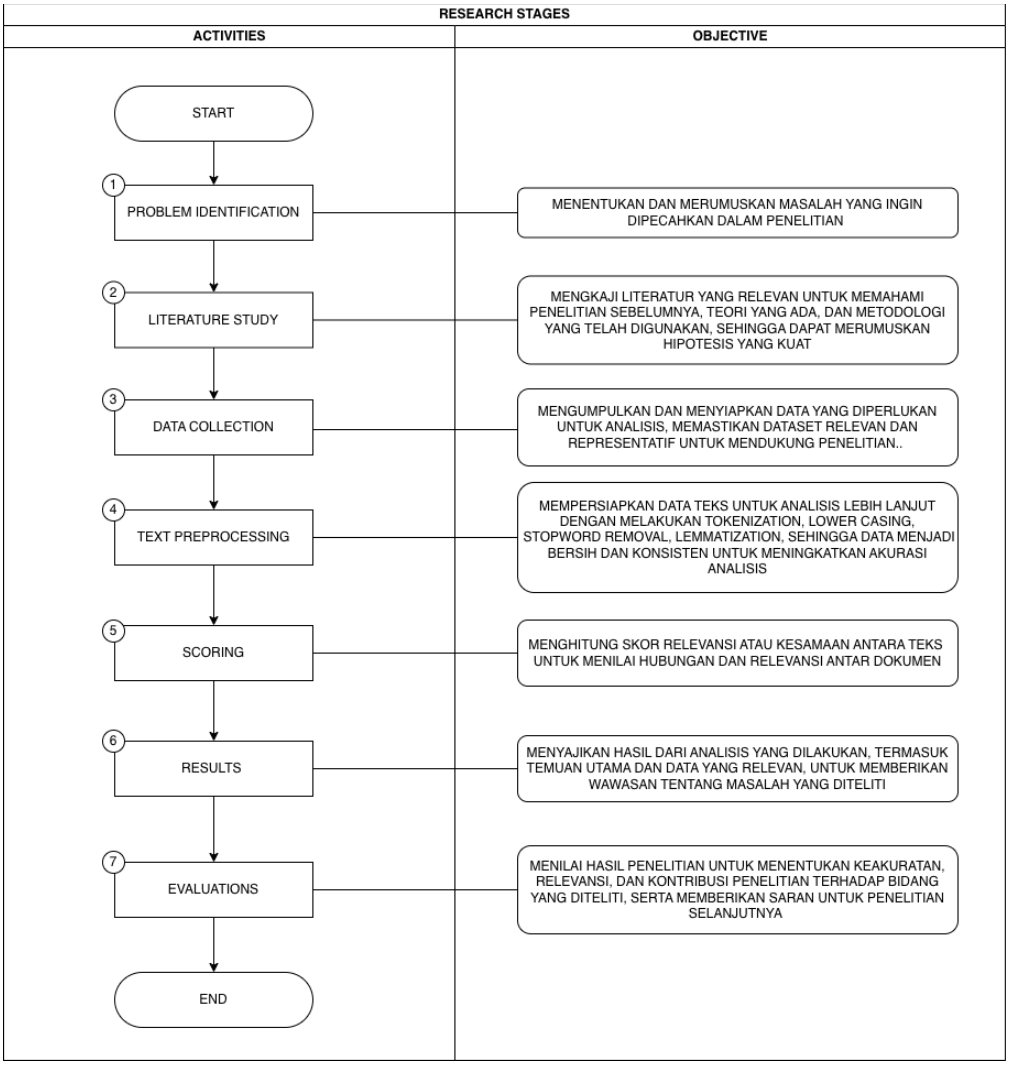
BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian *(Research Stages)*

Tahapan penelitian dijelaskan pada Gambar 3.1, yang menggambarkan alur proses dari identifikasi masalah hingga evaluasi. Gambar ini diadaptasi dari penelitian Sihombing (2022) dan dimodifikasi sesuai kebutuhan penelitian ini. Setiap langkah memiliki tujuan jelas agar penelitian dapat memberikan wawasan dan kontribusi signifikan terhadap bidang yang dikaji.



Gambar 3. 1 Tahapan Penelitian (Sihombing, 2022)

3.1.1 Identifikasi Masalah *(Problem Identification)*

Identifikasi Masalah merupakan langkah fundamental dalam penelitian ini, berfokus pada masalah penilaian esai manual yang tidak objektif dan konsisten. Dalam penilaian esai manual, ada beberapa masalah yang harus dipertimbangkan. Ini termasuk perbedaan dalam subjektivitas penilai, waktu yang dihabiskan untuk melakukan penilaian, dan hasil yang tidak konsisten. Pada tahap ini, kebutuhan akan dirumuskan untuk sistem penilaian esai otomatis yang menggabungkan *Cosine Similarity* dan *Model* (LLM). Oleh karena itu, diperlukan sistem otomatis untuk memberikan penilaian yang lebih akurat, konsisten, dan objektif.

3.1.2 Studi Literatur *(Literature Study)*

Studi literatur dilakukan dengan meninjau publikasi ilmiah dari berbagai platform akademik, seperti *Research Gate, Google Scholar, Directory of Open Access Journals (DOAJ), Springer Open, Academia.edu, arXiv,* dan *OALib*. Model penilaian esai otomatis, *Cosine Similarity, Text Mining*, *Large Language Model*s, dan Teknologi Pengolahan Bahasa Alam adalah topik artikel yang menjadi fokus penelusuran. Tinjauan literatur bertujuan mengidentifikasi *state of the art*, metode yang digunakan, keberhasilan, dan area penelitian yang masih perlu dikembangkan. Hasil dari studi literatur ini akan menjadi fondasi konseptual untuk merancang metodologi penelitian yang inovatif.

3.1.3 Pengumpulan Data *(Data Collection)*

1. Sumber *Dataset*

Dataset yang digunakan dalam penelitian ini bersumber dari penelitian sebelumnya oleh (Sihombing, 2022). Penelitian tersebut telah mengumpulkan data relevan yang mencakup tiga soal esai mengenai konsep dasar struktur data, kunci jawaban ideal, serta berbagai variasi jawaban yang diberikan oleh mahasiswa. Hal ini memastikan bahwa *dataset* yang digunakan memiliki landasan yang kuat dan relevan dengan topik yang diteliti. *Dataset* ini terdiri dari tiga soal esai yang dirancang untuk menguji pemahaman mahasiswa tentang struktur data. Setiap soal direspon oleh lima mahasiswa, sehingga menghasilkan beragam perspektif dan pendekatan dalam menjawab. Variasi jawaban ini penting untuk menciptakan *dataset* yang komprehensif dan mencerminkan beragam tingkat pemahaman mahasiswa.

2. Pelabelan Manual oleh Dosen

Proses pelabelan dilakukan secara manual oleh dua dosen yang berpengalaman di bidangnya. Ditugaskan untuk menilai dan memberikan skor pada jawaban mahasiswa, sehingga dapat diketahui seberapa akurat dan relevan model yang dikembangkan. Penilaian ini juga bertujuan untuk memastikan bahwa jawaban yang diberikan sesuai dengan kriteria yang telah ditetapkan.

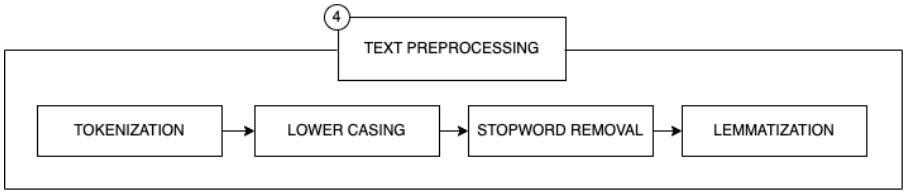
3. Validasi dan Normalisasi *Dataset*

Setelah semua jawaban terkumpul, tahap selanjutnya adalah validasi dan normalisasi *dataset*. Proses ini melibatkan analisis dan perbandingan jawaban dari

berbagai sumber untuk memastikan konsistensi dan kualitas data. Pembersihan data, penghapusan duplikasi, dan standarisasi format jawaban dilakukan untuk menghasilkan *dataset* yang tidak hanya akurat tetapi juga siap digunakan untuk penelitian lebih lanjut atau aplikasi pendidikan. Dengan demikian, *dataset* ini memberikan nilai tambah yang signifikan bagi pengembangan kurikulum dan proses pembelajaran.

3.1.4 Pra-pemrosesan Teks *(Text Preprocessing)*

Text preprocessing merupakan tahap penting untuk mempersiapkan data esai sebelum diproses. Tahapan *text preprocessing* pada Gambar 3.2 akan mencakup serangkaian teknik normalisasi dan transformasi teks dalam bahasa Indonesia.



Gambar 3. 2 Tahapan *Text Preprocessing* (Sihombing, 2022)

1. *Tokenization*

Memecah teks menjadi unit-unit yang lebih kecil (token), seperti kata atau frasa, untuk memudahkan analisis.

2. *Lower Casing*

Mengubah seluruh teks menjadi huruf kecil untuk menghilangkan variasi kapitalisasi.

3. *Stopword Removal*

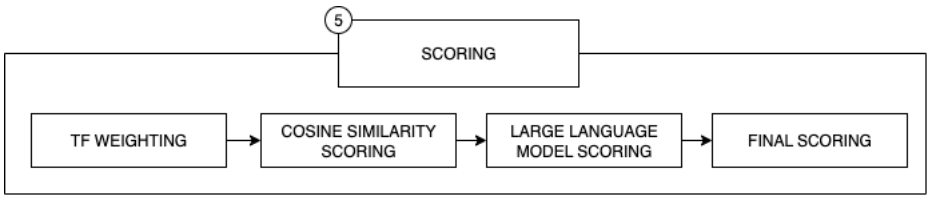
Menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis, seperti "yang", "dan", "di". Dalam proses ini, kita dapat menggunakan *package* Sastrawi untuk mengakses daftar *stopword* bahasa Indonesia yang telah ditentukan, sehingga proses penghapusan menjadi lebih efisien dan akurat. Dengan menghilangkan *stopword*, dapat fokus pada kata-kata yang lebih bermakna dan relevan dalam konteks analisis.

4. *Lemmatization*

Mengubah kata ke bentuk dasar yang valid dengan mempertimbangkan konteks dan struktur bahasa, misalnya "berlari" menjadi "lari". Proses *preprocessing* ini bertujuan untuk menghasilkan representasi teks yang bersih, terstandarisasi, dan siap untuk dianalisis oleh algoritma *Cosine Similarity* dan model bahasa besar.

3.1.5 Perhitungan skor *(Scoring)*

Pada Gambar 3.3, dilakukan *scoring* yaitu menghitung skor relevansi atau kesamaan antara teks untuk menilai hubungan dan relevansi antar dokumen.



Gambar 3. 3 Tahapan *Scoring* (Sihombing, 2022)

1. TF *Weighting*

Proses yang di lakukan pada tahap ini, bobot kata dihitung berdasarkan frekuensi kemunculannya dalam dokumen dan koleksi dokumen secara keseluruhan. Proses ini bertujuan untuk menilai pentingnya kata dalam konteks tertentu, sehingga dapat mengidentifikasi kata-kata kunci yang signifikan.

2. *Cosine Similarity Scoring*

Penilaian pada tahap ini menentukan kesamaan antara dua vektor teks diukur untuk menentukan sejauh mana dokumen satu mirip dengan dokumen lainnya. Hasil perhitungan cosine similarity diberikan bobot sebesar 40%, yang memberikan penilaian kuantitatif berbasis kesamaan teks.

3. *Large Language Model* (LLM) *Scoring*

Model LLM digunakan untuk memberikan penilaian kualitatif yang mencakup aspek-aspek seperti struktur argumen, kedalaman analisis, kohesi, dan ketepatan bahasa. Proses ini memiliki bobot sebesar 60% dalam evaluasi keseluruhan.

4. Final Scoring

Tahap akhir ini melibatkan akumulasi skor dari ketiga proses perhitungan yang telah dilakukan, berdasarkan bobot yang diberikan untuk proses *Cosine Similarity Scoring* dan *Model* (LLM) yang bisa di atur perbandingan bobotnya oleh tenaga penilai. Skor akhir mencerminkan evaluasi yang komprehensif, mengintegrasikan aspek kuantitatif dan kualitatif dari dokumen yang dianalisis.

3.1.6 Hasil *(Results)*

Hasil dari analisis disajikan secara komprehensif, mencakup beberapa aspek penting, yaitu:

1. Skor Penilaian Kuantitatif

Mencakup nilai yang diperoleh dari penerapan metode TF dan *Cosine Similarity*, yang digunakan untuk mengukur kesamaan dan relevansi antara teks jawaban mahasiswa dengan jawaban referensi. Skor ini memberikan indikasi yang jelas tentang seberapa baik jawaban mahasiswa sesuai dengan kriteria yang telah ditetapkan.

2. Penilaian Kualitatif

Penilaian ini dihasilkan oleh *Model* (LLM), yang mengevaluasi aspek-aspek seperti struktur argumen, kedalaman analisis, kohesi, dan ketepatan bahasa. Penilaian kualitatif ini memberikan perspektif yang lebih mendalam mengenai kualitas jawaban yang diberikan oleh mahasiswa.

3. Akurasi Sistem

Pada bagian ini, peneliti membandingkan hasil penilaian otomatis dengan penilaian manual yang dilakukan oleh pengajar. Ini bertujuan untuk menilai seberapa akurat sistem dalam mereplikasi penilaian manusia, serta untuk mengidentifikasi potensi area perbaikan.

3.1.7 Evaluasi *(Evaluations)*

Evaluasi dalam penelitian ini bertujuan mengukur efektivitas masing-masing pendekatan model menggunakan metode kuantitatif *Mean Absolute*

Percentage Error (MAPE), karena mampu merepresentasikan tingkat kesalahan prediksi secara relatif terhadap nilai aktual:

Evaluasi menggunakan metrik MAPE untuk menghitung rata-rata persentase kesalahan. Analisis dilakukan melalui tiga pengujian: (1) *Cosine Similarity*, (2) *Model* (LLM), dan (3) model hibrid yang menggabungkan keduanya.

1. Pengujian Model Berbasis *Cosine Similarity*

Pada tahap pertama, model diuji dengan hanya mengandalkan pendekatan *Cosine Similarity* untuk mengukur kesamaan antar dokumen atau teks. Pengujian ini bertujuan untuk mengevaluasi performa pendekatan berbasis statistik dan vektor dalam memodelkan kedekatan semantik antar teks.

2. Pengujian Model Berbasis *Model* (LLM)

Tahapan kedua mengevaluasi model yang hanya menggunakan *Model* (LLM). Evaluasi ini bertujuan untuk melihat seberapa baik LLM mampu memahami konteks dan makna secara semantik tanpa bantuan teknik pembobotan vektor.

3. Pengujian Model Hibrid (*Cosine Similarity* + LLM)

Tahapan ketiga merupakan evaluasi terhadap model hibrid yang menggabungkan pendekatan *Cosine Similarity* dengan LLM. Model ini diharapkan mampu mengombinasikan kekuatan statistik numerik dengan pemahaman semantik, sehingga memberikan hasil yang lebih akurat dan relevan. Evaluasi pada tahap ini menjadi penting untuk mengetahui apakah pendekatan gabungan dapat mengatasi keterbatasan dari masing-masing metode individual.

4. Interpretasi Hasil Evaluasi

Setelah ketiga model diuji, nilai MAPE yang dihasilkan akan dibandingkan untuk mengetahui pendekatan mana yang memberikan tingkat kesalahan prediksi paling rendah. Nilai MAPE yang lebih kecil menunjukkan tingkat akurasi yang lebih tinggi dan menjadi indikator bahwa model tersebut lebih optimal dalam konteks tugas yang diuji.