

# **Pengembangan Sistem Tinjauan Makalah dengan Large Language Models**

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Perkembangan teknologi informasi telah menjadi pendorong utama transformasi digital yang telah mengubah pola kehidupan manusia secara fundamental. Teknologi informasi telah menyebar ke berbagai sektor serta memfasilitasi akses ke berbagai informasi, mempercepat pertukaran data, dan mengubah cara manusia berinteraksi. Saat ini teknologi bukan hanya sekedar kebutuhan sekunder namun dapat dikatakan sebagai kebutuhan primer. Perkembangan teknologi mempengaruhi berbagai aspek kehidupan seperti sosial, politik dan budaya. Salah satu bentuk dari perkembangan teknologi adalah kecerdasan buatan atau *artificial intelligence* (AI).

Kecerdasan buatan atau AI mengacu pada aplikasi algoritma dan teknik perangkat lunak yang memungkinkan komputer dan mesin untuk mensimulasikan persepsi manusia dan proses pengambilan keputusan untuk menyelesaikan tugas dengan sukses (Murphy, 2019). Teknologi kecerdasan buatan (AI) berkembang mengiringi perkembangan teknologi informasi, hal ini membawa perubahan diberbagai bidang seperti pendidikan, kesehatan, ekonomi, industry, dan transportasi. Teknologi AI menawarkan berbagai kemudahan, efisiensi, analisis data yang mendalam dan pengenalan pola. Perkembangan kecerdasan buatan (AI) telah memainkan peran krusial dalam menghadirkan kemajuan signifikan diberbagai aplikasi. Melalui teknik-teknik seperti deep learning dan machine learning, AI mampu menganalisis data yang kompleks dan mengidentifikasi pola-pola yang tersembunyi di dalamnya. Pada pengenalan gambar misalnya, AI telah mencapai tingkat ketepatan yang mengesankan dalam mengenali objek, wajah, atau bahkan pola-pola mikroskopis dalam citra medis. pada bidang teks, AI dapat mengenali pola-pola kompleks dalam teks seperti struktur gramatikal, entitas, dan makna kata. AI juga dapat mengidentifikasi pola-pola seperti opini atau sentimen dalam teks, topik pembicaraan, atau bahkan pemahaman konteks dari suatu kalimat.

Pertumbuhan pesat penelitian ilmiah dan produksi artikel ilmiah dalam berbagai disiplin ilmu menghasilkan tantangan baru dalam proses tinjauan artikel ilmiah. Volume besar publikasi membuat tugas penelaahan sejawat (*peer-review*) semakin kompleks dan memakan waktu yang lebih banyak. Jumlah naskah yang diajukan ke jurnal untuk proses *peer-review* mengalami pertumbuhan tahunan sebesar 6,1% (Checco et al., 2021). Sementara itu, kebutuhan untuk memastikan kualitas dan keakuratan penilaian tetap menjadi prioritas. Adanya pertumbuhan pengajuan naskah untuk di tinjau menimbulkan potensi bias dan konsistensi penelaahan yang kurang baik. Sehingga terdapat peluang untuk mengembangkan suatu model yang mampu melakukan tinjauan secara mandiri dengan bantuan kecerdasan buatan.

Pada saat yang sama, perkembangan kecerdasan buatan dalam bidang pemrosesan teks atau yang dikenal dengan *Natural Language Processing (NLP)* telah membuka peluang baru untuk otomatisasi beberapa aspek dari proses tinjauan artikel ilmiah. Kemampuan algoritma tersebut dalam memahami bahasa alami, analisis konten, dan penyajian informasi dapat dimanfaatkan untuk mempermudah dan mempercepat proses tinjauan. Pemrosesan bahasa alami (NLP) adalah subbidang kecerdasan buatan dan linguistik komputasi. Bidang ini berfokus pada kemampuan komputer untuk memahami, menafsirkan, dan menghasilkan bahasa manusia dengan cara yang bermakna dan berguna (Amaratunga, 2023). Salah satu pendekatan dalam pemrosesan bahasa alami yang memungkinkan untuk melakukan tugas mendalam dengan data besar adalah model bahasa besar atau *Large Language Models (LLM)*.

Model Bahasa besar adalah model bahasa yang telah dilatih sebelumnya dengan ukuran parameter jauh lebih besar dibandingkan dengan pendekatan lainnya (Liu et al., 2024). Model bahasa besar adalah hasil dari kombinasi pemrosesan bahasa alami, konsep pembelajaran mendalam, dan model kecerdasan buatan generative. Akhir-akhir ini model bahasa besar telah memberikan terobosan yang signifikan terutama yang berkaitan dengan transformer. Hal ini mencakup peningkatan komputasi dan ketersediaan pelatihan data dengan skala yang besar.

Perkembangan tersebut telah menghasilkan transformasi revolusioner dengan memberikan kemungkinan pembuatan LLM yang dapat mendekati kinerja manusia pada berbagai tugas (Naveed et al., 2023).

Beberapa penelitian sebelumnya membahas mengenai penelaahan sejawat dengan kecerdasan buatan seperti yang ditulis oleh Alessandro, Lorenzo, Pierpaolo, Stephen & Giuseppe pada tahun 2021 menunjukkan bahwa kecerdasan buatan mampu melakukan penelaahan sejawat dan memprediksi sesuai dengan hasil ahli (Checco et al., 2021). Penelitian lainnya dilakukan oleh Iddo Drori & Dov Te'eni tahun 2024 menunjukkan kecerdasan buatan khususnya model Bahasa besar mampu melakukan tugas telaah sejawat dengan cukup baik (Drori & Te'eni, 2024).

Penelitian ini bertujuan untuk mengembangkan model kecerdasan buatan untuk tinjauan artikel berbasis large language models dengan fokus pada peningkatan efisiensi dan kecepatan proses telaah sejawat, tanpa mengorbankan kualitas dan keakuratan penilaian ilmiah. Model ini diharapkan dapat membantu dalam identifikasi aspek-aspek kritis, serta memberikan analisis otomatis yang dapat digunakan sebagai dasar untuk penilaian lebih lanjut. Pengembangan model-tools ini diharapkan dapat mempercepat alur kerja tinjauan artikel, meminimalkan potensi kesalahan manusia, dan pada akhirnya dapat meningkatkan efisiensi penelitian ilmiah secara keseluruhan.

## **1.2 Batasan dan Rumusan Masalah**

### **1.2.1 Batasan Masalah**

Untuk mencegah meluasnya permasalahan dalam domain yang diteliti penulis membuat Batasan agar fokus pada penyelesaian masalah dapat tercapai. Adapun Batasan masalah dari penelitian ini adalah sebagai berikut :

1. Penelitian berfokus pada pengembangan model platform tinjauan artikel ilmiah dengan menggunakan LLM
2. Model LLM yang digunakan adalah GPT-4
3. Model difokuskan pada artikel ilmiah dalam disiplin ilmu komputer

### **1.2.2 Rumusan Masalah**

Berdasarkan latar belakang masalah yang telah disampaikan diperoleh beberapa permasalahan yang harus di selesaikan. Permasalahan tersebut dirumuskan sebagai berikut :

1. Bagaimana membangun model platform tinjauan artikel ilmiah dengan LLM?
2. Bagaimana algoritma untuk klasifikasi artikel ilmiah untuk publikasi jurnal?

### **1.3 Tujuan Penelitian**

Secara umum penelitian ini bertujuan untuk menciptakan platform tinjauan artikel ilmiah berbasis kecerdasan buatan dengan menggunakan model bahasa besar. Tujuan penelitian secara khusus adalah sebagai berikut :

1. Membangun model platform tinjauan artikel ilmiah dengan LLM
2. Mengembangkan algoritma untuk klasifikasi artikel ilmiah untuk publikasi jurnal

Berdasarkan penelitian telaah pustaka yang telah dilakukan dapat disimpulkan bahwa proses telaah sejawat dapat dilakukan secara otomatis dengan bantuan teknologi kecerdasan buatan. beberapa penelitian sebelumnya telah membahas mengenai penggunaan kecerdasan buatan terutama menggunakan model bahasa besar. Berdasarkan pemaparan tersebut maka terdapat peluang untuk dilakukan pengembangan platform tinjauan artikel ilmiah dalam bidang ilmu computer dengan menggunakan model bahasa besar.

## **2.6 Peer Review**

Peer Review Process adalah proses di mana jurnal menilai kualitas naskah sebelum diterbitkan, ditinjau oleh para ahli yang relevan di bidangnya untuk mereview dan mengomentari naskah yang diterima. Proses ini bertujuan untuk membantu editor menentukan apakah naskah harus diterbitkan dalam jurnal.

Poin penting dalam Peer Review Process :

1. Naskah yang dikirimkan ke jurnal terlebih dahulu melalui penyaringan awal oleh tim editorial.
2. Naskah yang lolos pemeriksaan akan dikirim pada minimal dua peer reviewer untuk ditinjau.
3. Peer reviewer secara independen membuat rekomendasi kepada editor jurnal, apakah naskah harus ditolak atau diterima (dengan atau tanpa revisi).
4. Editor jurnal mempertimbangkan semua umpan balik dari peer reviewer dan membuat keputusan untuk menerima atau menolak naskah.
5. Peer Review Process untuk publikasi jurnal pada dasarnya adalah mekanisme kendali mutu, dimana para ahli mengevaluasi naskah yang bertujuan untuk memastikan kualitas dari naskah yang diterbitkan.

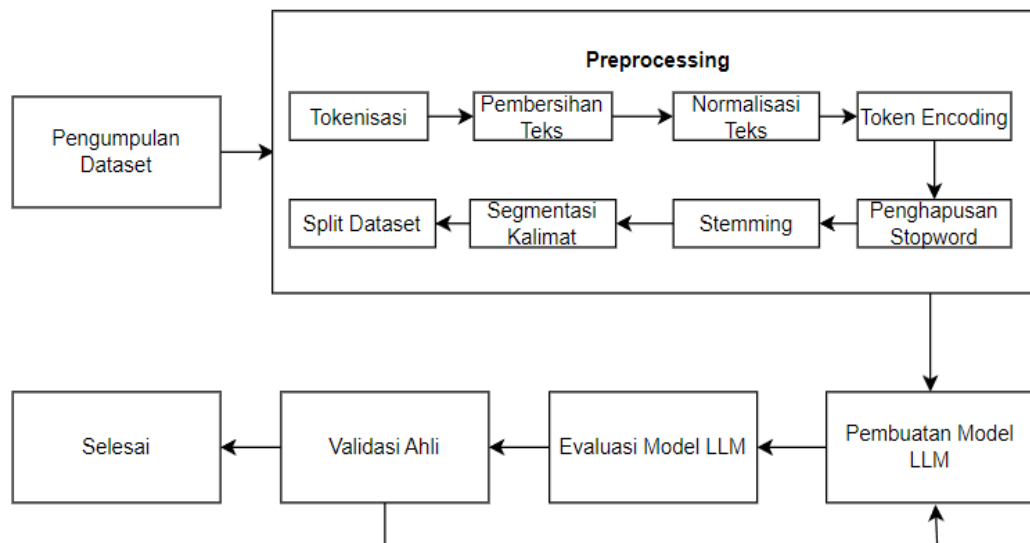
## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Gambaran Umum

Penelitian ini bertujuan untuk mengembangkan aplikasi berbasis Large Language Model (LLM) dengan arsitektur GPT-4 yang mampu melakukan telaah sejawat(peer review) secara otomatis pada artikel ilmiah dari jurnal komputer. Data utama yang digunakan adalah artikel ilmiah berbahasa Indonesia dalam bidang ilmu komputer dari berbagai jurnal akademik. Sebelum digunakan, data akan diperiksa untuk menghilangkan informasi pribadi yang dapat mengidentifikasi penulis atau reviewer. Aplikasi ini diharapkan dapat membantu para peneliti dan editor jurnal dalam menganalisis dan memperoleh wawasan dari artikel yang seringkali bersifat kompleks dan teknis.

Untuk melakukan penelitian ini perlu dilakukan beberapa tahapan hingga penelitian selesai, tahapan yang dilakukan mulai dari pengumpulan data, preprocessing data, melakukan pemodelan untuk telaah sejawat, mengevaluasi model dan validasi ahli. Untuk tahapan penelitian dapat dilihat pada gambar 3.1.



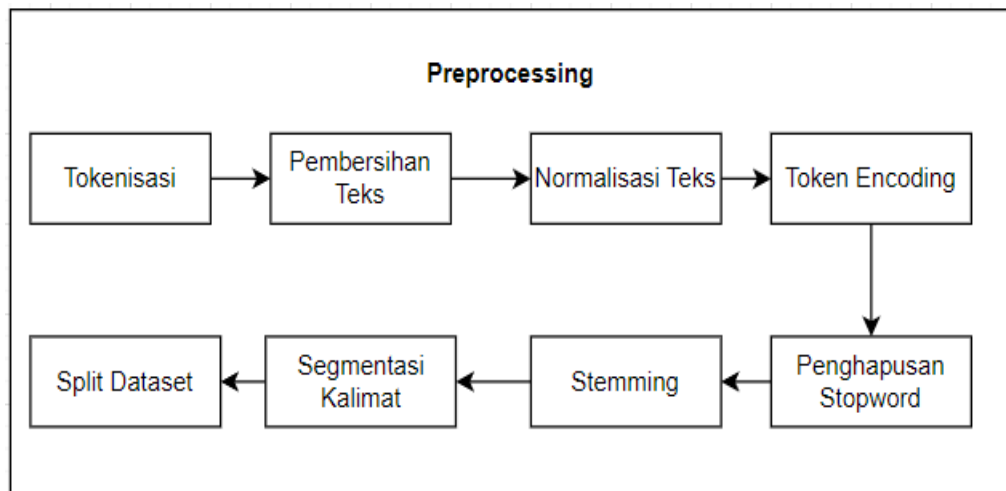
**Gambar 3. 1 Tahapan Penelitian**

### 3.1.1 Pengumpulan data

Proses pengumpulan data dilakukan dengan cara mengumpulkan artikel ilmiah dari berbagai sumber terbuka dengan topik artikel ilmu computer. Pengumpulan data menggunakan teknik webscraping, artikel yang telah dikumpulkan akan diproses melalui tahap preprocessing.

### 3.1.2 Preprocessing data

Proses preprocessing data merupakan langkah yang sangat penting dalam persiapan data untuk pemodelan LLM. Proses ini melibatkan beberapa tahap penting yang bertujuan untuk membersihkan dan menyiapkan data teks agar sesuai dengan kebutuhan model serta meningkatkan kualitas dan konsistensi representasi teks. Proses preprocessing dilakukan melalui beberapa tahap seperti tokenisasi, pembersihan teks, normalisasi, token encoding, penghapusan stopwords, stemming, segmentasi kalimat dan pemisahan dataset.



**Gambar 3. 2 Tahapan Preprocessing**

Tahap pertama adalah tokenisasi, di mana teks dipecah menjadi unit-unit yang lebih kecil yang dikenal sebagai token, memungkinkan model untuk menganalisis teks pada tingkat yang lebih granular. Selanjutnya, dilakukan pembersihan teks untuk menghilangkan karakter atau simbol yang tidak diinginkan seperti tanda baca, angka, dan karakter khusus lainnya, serta penghapusan spasi berlebih dan karakter yang tidak relevan. Normalisasi juga dilakukan untuk mengubah teks



menjadi bentuk standar, termasuk mengubah semua huruf menjadi huruf kecil, menghapus aksentuasi dari huruf, dan menangani variasi penulisan yang berbeda untuk kata yang sama. Setelah itu, token yang dihasilkan dari tokenisasi perlu diubah menjadi representasi numerik melalui token encoding, menggunakan teknik embeddings dari model transformer. Penghapusan stopwords, yaitu kata-kata umum yang sering muncul dalam teks tetapi tidak memiliki makna khusus yang penting untuk analisis, juga dilakukan untuk mengurangi dimensi data dan fokus pada kata-kata yang lebih bermakna. Proses selanjutnya adalah stemming dan lemmatisasi yang bertujuan untuk mengurangi kata-kata ke bentuk dasar atau akar katanya, dengan stemming memotong akhiran kata dan lemmatisasi menggunakan kamus bahasa untuk mengembalikan kata ke bentuk dasar yang benar secara gramatikal. Selanjutnya segmentasi kalimat dilakukan untuk memisahkan teks menjadi kalimat-kalimat individu yang bisa dianalisis lebih lanjut secara terpisah. Tahap terakhir dalam preprocessing adalah pemisahan dataset menjadi bagian-bagian yang berbeda, seperti data latih, data validasi, dan data uji, yang penting untuk mengevaluasi kinerja model secara adil dan menghindari overfitting. Melalui proses preprocessing yang cermat dan terstruktur, data teks menjadi lebih bersih, terorganisir, dan siap digunakan dalam pemodelan, sehingga tidak hanya meningkatkan efisiensi pemrosesan data tetapi juga memungkinkan model untuk belajar dan melakukan prediksi dengan lebih akurat.

### **3.1.3 Pembuatan Model LLM**

Setelah dataset yang di kumpulkan dan melalui proses preprocessing maka dilanjutkan tahap pemodelan dengan menggunakan LLM. Pada tahap ini dilakukan pemodelan dengan arsitektur GPT-4 untuk platform tinjauan artikel ilmiah. Proses pemodelan dimulai dengan fine-tuning GPT-4 menggunakan dataset yang telah dipreprocessing sebelumnya. Fine-tuning dilakukan untuk menyesuaikan model dengan gaya penulisan dan terminologi spesifik yang digunakan dalam artikel ilmiah. Selama fase pelatihan, model dievaluasi secara berkala untuk memastikan kinerjanya sesuai dengan harapan, dan parameter model dioptimalkan untuk meningkatkan kualitas output. Penggunaan GPT-4 untuk platform tinjauan artikel ilmiah dapat menyediakan analisis yang mendalam dan komprehensif, membantu

reviewer untuk lebih cepat dan efisien dalam menilai kualitas dan kontribusi sebuah artikel. Hal ini tidak hanya meningkatkan produktivitas tetapi juga memastikan bahwa artikel yang dipublikasikan memenuhi standar ilmiah yang tinggi.

#### **3.1.4 Evaluasi Model LLM**

Evaluasi model merupakan langkah yang penting dalam pengembangan sistem kecerdasan buatan, karena memungkinkan untuk menilai kinerja dan efektivitas model dalam menyelesaikan tugas tertentu. Proses evaluasi membantu mengidentifikasi kelemahan dan kekuatan model, serta memberikan wawasan tentang seberapa baik model dapat digunakan. Tanpa evaluasi yang tepat, model yang dikembangkan dapat menghasilkan prediksi yang tidak akurat atau tidak dapat diandalkan, yang berpotensi menyebabkan kinerja sistem yang buruk secara keseluruhan. Pada penelitian ini dilakukan evaluasi model dengan melihat nilai akurasi, presisi, recall dan F1-Score.

1. Akurasi memberikan gambaran umum tentang seberapa baik model klasifikasi melakukan prediksi secara keseluruhan.
2. Presisi memberikan informasi tentang seberapa banyak prediksi positif yang sebenarnya benar dari semua prediksi positif yang dilakukan oleh model.
3. Recall memberikan informasi tentang seberapa banyak instance positif yang berhasil diidentifikasi oleh model dari semua instance positif yang
4. sebenarnya dalam dataset.
5. F1-Score berguna ketika kelas target tidak seimbang dalam dataset, karena mencakup baik presisi maupun recall dalam perhitungannya.

#### **3.1.5 Validasi Ahli**

Proses validasi ahli ini memastikan bahwa model GPT-4 yang digunakan untuk telaah sejawat mampu memberikan evaluasi yang akurat, relevan, dan sesuai dengan standar akademik, dengan masukan berharga dari para ahli di bidangnya.

### **3.2 Jadwal Penelitian**

Jadwal penelitian bertujuan untuk mengatasi target waktu penelitian, memastikan bahwa penelitian ini dapat diselesaikan sesuai dengan batas waktu

yang telah ditetapkan. Adanya jadwal penelitian, diharapkan penelitian dapat berjalan secara efisien dan sesuai rencana, sehingga memberikan kepastian bahwa semua tahapan penelitian dapat diselesaikan tepat pada waktunya. Table jadwal penelitian dapat dilihat pada table 3.1

**Table 3. 1 Jadwal Penelitian**

[illegible]