

BAB I

PENDAHULUAN

1.1 Latar Belakang

Informasi adalah salah satu elemen fundamental bagi kehidupan manusia, berfungsi sebagai fondasi untuk pengambilan keputusan, pembelajaran, dan komunikasi. Sejak dulu manusia telah mengandalkan informasi untuk bertahan hidup dan berkembang. Revolusi percetakan pada abad ke-15 membawa lompatan besar dalam penyebaran informasi, membuat buku dan surat kabar menjadi lebih mudah diakses oleh masyarakat luas (Rahayu Ningrum et al., 2024).

Berita merupakan salah satu sumber informasi utama yang sering diakses oleh pembaca (Ansharulhaq Aminsyah et al., 2025). Di era digital saat ini, akses terhadap berita semakin mudah dan cepat melalui berbagai portal berita online, memungkinkan masyarakat untuk mendapatkan informasi terkini dalam hitungan detik (Adek et al., 2022). Salah satu portal berita yang berperan penting dalam ekosistem ini adalah *detiknews*, sebuah sub-portal dari *Detik.com* bersama dengan sub-portal lainnya seperti Detik Sport, Detik *Finance*, dan sebagainya.

detiknews menyajikan berita terbaru dari berbagai kategori dengan puluhan hingga ratusan artikel yang diterbitkan setiap hari, menjadikannya salah satu sumber informasi terpercaya di Indonesia. Kecepatan dan akurasi informasi yang disampaikan oleh *detiknews* membuatnya menjadi pilihan utama bagi masyarakat yang ingin selalu terhubung dengan berita terkini.

Setiap hari, jumlah berita yang disediakan sangat banyak sehingga pembaca dapat menerima berbagai informasi dari berbagai artikel. Namun, banyaknya berita yang tersedia terutama dengan topik yang sama sering kali membuat pembaca kesulitan untuk menemukan inti dari informasi yang mereka cari, pembaca harus menelusuri dan membaca sejumlah besar berita secara lengkap dan satu per satu untuk menemukan poin penting atau informasi inti dari berita tersebut. Proses ini tidak hanya memakan waktu tetapi juga dapat menyebabkan kelelahan terutama jika pembaca harus menyaring berita-berita yang serupa atau berulang demi mendapatkan informasi yang relevan dan penting.

Sistem *summarise* berita adalah sistem yang meringkas berita berdasarkan topik yang dimasukkan, sehingga memuat inti penting dari berita yang ditemukan tanpa perlu membaca seluruh berita secara lengkap dan satu per satu untuk menemukan informasi yang relevan. Dengan banyaknya berita yang disediakan setiap hari, terutama pada topik yang sama, sistem ini membantu pembaca mengatasi masalah kelebihan informasi. Pembaca tidak lagi harus merasa kewalahan atau lelah akibat membaca sejumlah besar berita, karena sistem ini mampu menyaring dan menampilkan poin-poin penting secara ringkas dan efisien. Dengan demikian, pembaca dapat mengakses informasi yang diperlukan dengan lebih cepat dan mudah tanpa kehilangan esensi dari berita yang disajikan.

Proses *summarise* akan menghasilkan ringkasan berita berdasarkan topik yang diinput dengan menggunakan metode *Large Language Models (LLM)* dan bantu diperhalus dengan *relevance-augmented generation*. Sistem ini melakukan peringkasan berita melalui *API* dari *Groq*, yang diintegrasikan dengan mekanisme scraping pada *API* portal berita *detiknews* untuk mengumpulkan berita sesuai topik yang dimasukkan oleh pengguna. Dengan pendekatan ini, sistem dapat menyaring dan meringkas berita secara efisien, sehingga membantu pengguna menemukan inti informasi yang relevan tanpa perlu membaca keseluruhan berita secara manual.

Metode *Large Language Models (LLM)* memiliki kemampuan untuk memahami konteks secara mendalam, sehingga dapat menghasilkan ringkasan yang akurat dan relevan. Selain itu, teknik *Relevance-Augmented Generation* menambahkan lapisan filter yang memastikan hanya informasi yang relevan dengan topik yang dimasukkan oleh pengguna yang disertakan dalam ringkasan.

Penelitian sebelumnya yang dilakukan oleh Anggun Tri Utami Br. Lubis, Nazruddin Safaat Harahap, Surya Agustian, Muhammad Irsyad dan Iis Afrianty pada tahun 2024, dalam penelitian “*Question Answering System* pada *Chatbot Telegram Menggunakan Large Language Models (LLM) dan Langchain (Studi Kasus UU Kesehatan)*”, mereka berhasil memanfaatkan teknologi *Large Language Models (LLM)* dan *Langchain* untuk mengembangkan sebuah sistem *Question Answering* yang terintegrasi dengan *chatbot*.

Berdasarkan penjelasan di atas, maka judul skripsi yang peneliti ambil adalah “Penerapan *Relevance-Augmented Generation Large Language Models* Untuk Summarise Berita Pada Portal Berita *Detiknews*”

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diuraikan diatas, maka permasalahan yang penulis rumuskan adalah:

1. Bagaimana melakukan *summarise* berita berdasarkan sebuah topik?
2. Bagaimana mendapatkan berita berdasarkan topik yang di input?
3. Bagaimana mengekstrak kalimat-kalimat dari berita tersebut?
4. Bagaimana skor kualitas hasil *summarise* berita ditentukan berdasarkan analisis sistem yang ditampilkan oleh aplikasi?

1.3 Tujuan Penelitian

Tujuan dari penelitian yang akan penulis laksanakan antara lain sebagai berikut:

1. Menyediakan hasil *summary* dari berita.
2. Memberikan inti dari berita dalam waktu singkat, memungkinkan pembaca untuk cepat mendapatkan pokok-pokok penting tanpa membaca semua berita.
3. Mempermudah menyaring dan meringkas berita secara efisien.
4. Memahami cara aplikasi menilai kualitas ringkasan berita agar hasilnya relevan dan informatif.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini ditetapkan adalah sebagai berikut:

1. *Summarise* ini berdarkan berita yang discraping pada *API* yang tersedia di *detiknews*.
2. Pembuatan *system summarise* menggunakan bahasa pemrograman *python*.
3. Membuat *interface* sederhana untuk menampilkan hasil *summary*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Memberikan solusi untuk mempermudah dalam membaca inti berita tanpa harus membaca satu-persatu.
2. Menghemat waktu pembaca terutama saat ada banyak berita yang perlu diakses.