

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi diciptakan untuk mempermudah pekerjaan manusia dan meningkatkan produktivitas [1]. Di era digital yang serba cepat ini, masyarakat semakin bergantung pada teknologi untuk mendapatkan informasi secara akurat dan instan. Kecerdasan buatan atau *artificial intelligent* adalah salah satu jenis teknologi yang semakin populer dan telah digunakan di banyak bidang, termasuk pendidikan. *Chatbot* berbasis AI menjadi bagian penting dalam sektor pendidikan karena dapat membantu siswa dan pendidik dengan memberikan jawaban cepat, mengurangi waktu tunggu, dan meningkatkan aksesibilitas informasi [2]. Kecerdasan buatan, atau AI, memungkinkan sistem komputer, perangkat lunak, atau bahkan robot untuk beroperasi dan berpikir layaknya manusia [1]. Salah satu contoh penerapan AI dalam bentuk *chatbot* adalah *Chatgpt* dan *Perplexity AI*. Kedua aplikasi tersebut telah menarik perhatian yang signifikan sebagai cara untuk memenuhi permintaan informasi pengguna dengan cepat dan efektif. Data menunjukkan bahwa *Chatgpt* mendominasi penggunaan *chatbot* AI di Indonesia dengan pangsa pasar 79%, sementara *Perplexity AI* hanya mencapai 5% [2].

Meskipun kedua aplikasi ini memiliki kemampuan yang terus berkembang dalam memahami dan merespons pertanyaan pengguna, masih sering muncul pertanyaan mengenai tingkat akurasi dan relevansi jawaban yang dihasilkan. Pengguna sering kali dihadapkan pada tantangan untuk menilai mana dari kedua aplikasi ini yang lebih tepat dan relevan dalam memberikan informasi yang dibutuhkan. Masalah ini semakin diperkuat oleh data yang menunjukkan bahwa hasil yang diberikan oleh aplikasi AI tidak selalu konsisten dalam hal akurasi dan relevansi. Misalnya, studi yang dilakukan pada berbagai *platform AI* mengungkapkan bahwa tingkat akurasi jawaban bisa bervariasi antara 60-80%, tergantung pada kompleksitas pertanyaan dan algoritma yang digunakan. Hal ini

menunjukkan bahwa meskipun AI telah berkembang pesat, masih terdapat celah dalam memberikan jawaban yang sepenuhnya akurat dan relevan.

Data ini menggarisbawahi pentingnya dilakukan evaluasi komparatif untuk mengetahui sejauh mana *Chatgpt* dan *Perplexity AI* dapat diandalkan dalam pencarian informasi dan dalam memberikan jawaban yang akurat serta relevan bagi pengguna. Oleh karena itu, studi ini berfokus pada menganalisis serta membandingkan dua aplikasi tersebut mengenai ketepatan dan kesesuaian jawaban yang mereka hasilkan. Dalam rangka melakukan perbandingan ini, studi ini akan memanfaatkan dua pendekatan utama, yaitu Proses Hierarki Analitik (AHP) dan *Blue Score*. Metode AHP akan diterapkan untuk memberikan nilai pada berbagai kriteria evaluasi berdasarkan hierarki kepentingan yang ditentukan sebelumnya, seperti akurasi, relevansi, kelengkapan, dan ketepatan waktu jawaban. Dengan melibatkan pakar dan pengguna dalam menentukan prioritas kriteria, AHP memungkinkan peneliti untuk mengevaluasi performa kedua aplikasi secara lebih terstruktur dan sistematis.

Sementara itu, *Blue Score* akan digunakan untuk mengukur kualitas jawaban dengan membandingkan kesamaan antara jawaban yang dihasilkan oleh *Chatgpt* dan *Perplexity AI* dengan jawaban referensi yang dianggap ideal. *Blue Score* sering digunakan dalam penilaian kualitas teks dalam pemrosesan bahasa alami, mengukur ketepatan dan kelengkapan jawaban berdasarkan *n-gram precision*, sehingga dapat memberikan pandangan yang lebih netral mengenai sejauh mana respons yang diberikan memenuhi kriteria kualitas yang diharapkan.

Pendekatan ini diharapkan mampu menyajikan pandangan yang lebih menyeluruh dan netral mengenai kelebihan serta kekurangan *Chatgpt* dan *Perplexity AI* dalam memberikan respons yang sesuai dengan keinginan pengguna. Studi ini diharapkan dapat memberikan pemahaman yang lebih baik tentang efektivitas kedua aplikasi AI dalam menciptakan teks yang tepat dan sesuai. Selain itu, hasil analisis ini juga diharapkan dapat memberikan rekomendasi untuk pengembangan AI di masa depan, terutama dalam hal peningkatan kualitas teks yang dihasilkan dan pemahaman konteks yang lebih baik oleh model-model AI tersebut.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah:

1. Bagaimana tingkat akurasi jawaban yang diberikan oleh aplikasi *Chatgpt* dan *Perplexity* dalam pencarian informasi?
2. Seberapa relevan jawaban yang diberikan oleh *Chatgpt* dibandingkan dengan *Perplexity* berdasarkan metode *AHP*?

1.3 Batasan Masalah

Batasan masalah pada penelitian kali ini adalah:

1. Penelitian ini hanya membandingkan dua aplikasi, yaitu *Chatgpt* dan *Perplexity*, dalam hal akurasi dan relevansi jawaban.
2. Pendekatan yang diterapkan dalam studi ini ialah *AHP* dan Blue Score untuk menilai ketepatan respon.
3. Studi ini memanfaatkan sumber informasi yang dapat diandalkan, yaitu data yang didapat dari situs resmi *Amazon*.
4. Penelitian ini tidak mencakup pembuatan sistem, baik dalam bentuk website maupun aplikasi mobile, dan hanya berfokus pada analisis dan evaluasi kinerja aplikasi *Chatgpt* dan *Perplexity AI* berdasarkan hasil jawaban yang dihasilkan.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Menganalisis tingkat akurasi jawaban yang diberikan oleh aplikasi *Chatgpt* dan *Perplexity* dalam pencarian informasi.
2. Mengevaluasi relevansi jawaban yang dihasilkan oleh *Chatgpt* dan *Perplexity* menggunakan metode *AHP*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini ialah:

1. Dengan menerapkan pendekatan dengan metode *AHP*, dapat meningkatkan keterampilan analisis data, terutama dalam mengevaluasi akurasi dan relevansi informasi yang dihasilkan oleh model AI melalui penentuan prioritas kriteria evaluasi secara sistematis.

2. Memanfaatkan temuan dari studi ini untuk menentukan aplikasi kecerdasan buatan, seperti Chatgpt atau Perplexity, yang paling cocok dengan kebutuhan pencarian data, agar pengguna dapat meningkatkan efisiensi dan hasil kerja dalam pengambilan keputusan yang didasarkan pada data.
3. Memperoleh pemahaman yang lebih mendalam mengenai teknologi AI, khususnya model bahasa seperti *Chatgpt* dan *Perplexity*, serta bagaimana teknologi ini berfungsi dalam konteks pencarian informasi.
4. Memahami kualitas informasi yang dihasilkan oleh aplikasi AI untuk digunakan dalam konteks akademik, profesional, dan public, dengan pendekatan evaluasi yang terstruktur melalui metode AHP.
5. Penggunaan *AHP* dan *Blue Score* sebagai alat evaluasi dapat memperkaya metode yang digunakan dalam menilai kinerja aplikasi AI lainnya. Hal ini membuka peluang untuk penelitian lebih lanjut dalam metode evaluasi performa model AI yang lebih mendalam dan komprehensif.