

BAB II

KAJIAN LITERATUR

2.1 Interaksi Manusia dan Komputer

Human-Computer Interaction (HCI) atau interaksi antara manusia dan komputer adalah proses dengan mengintegrasikan berbagai bidang ilmu terhadap pengalaman pada teknologi digital menjadi lebih optimal. HCI sendiri adalah gabungan dari keahlian informatika, psikologi, desain, dan ergonomis untuk menghasilkan teknologi yang dapat dipahami dan digunakan oleh pengguna. Fokus utama pada HCI adalah berfokus pada antarmuka yang dapat memungkinkan interaksi antara manusia dan komputer menjadi lebih efektif sehingga dapat menciptakan pengalaman digital yang lebih interaktif dan menyenangkan. HCI sendiri terdiri dari berbagai komponen-komponen antara muka yang digunakan sebagai jembatan utama antara manusia dan komputer yaitu *User Interface* (UI) dan *User Experience* (UX) [12], [13].

2.2 Kecerdasan Buatan (*Artificial Intelligence*)

Artificial Intelligence (AI) atau kecerdasan buatan merupakan cabang ilmu komputer yang berfokus pada pembuatan mesin dan sistem yang dapat melakukan tugas dengan kecerdasan yang mirip dengan manusia [14]. Tujuan dari kecerdasan ini adalah memberi komputer kemampuan untuk berpikir dan belajar melalui pemrograman dan algoritma yang kompleks, sehingga mereka dapat memecahkan masalah secara mandiri. Russell dan Norvig [15] mengatakan bahwa tujuan utama kecerdasan buatan adalah menciptakan agen cerdas yang dapat merasakan informasi dari lingkungan mereka dan melakukan yang terbaik untuk mencapai tujuan mereka.

Tujuan implementasi teknologi kecerdasan buatan difokuskan untuk membantu manusia melakukan tugas yang rumit termasuk membuat Keputusan (*Decision Making*), memecahkan masalah (*Problem-Solving*), memahami input sensorik, dan memahami pola manusia berkomunikasi satu sama lain. Banyak model AI menggunakan teknik Pembelajaran Mesin (*Machine Learning*) untuk mendapatkan informasi dari data [16]. Pembelajaran Mendalam (*Deep Learning*) yang menggunakan jaringan saraf tiruan berlapis untuk berinteraksi dengan data dan memproses informasi pada tingkat abstraksi yang lebih tinggi [17]. Kecerdasan buatan bisa saja mampu menciptakan sistem yang lebih akurat,

responsif, dan mampu bekerja dengan data dunia nyata dengan menerapkan kombinasi berbagai metode.

Kecerdasan buatan dikategorikan menjadi lima kategori utama berdasarkan tantangan yang dapat diatasi dan tugas yang dapat dilaksanakan [14]:

1. Kecerdasan buatan Analitis: Kategori kecerdasan buatan ini fokus pada mengidentifikasi, memahami, dan menyebarkan pola signifikan dalam data. Kecerdasan buatan analitis berusaha menemukan pola dan korelasi di antara variabel untuk memfasilitasi pengambilan keputusan berbasis data.
2. Kecerdasan buatan Fungsional: Bekerja dengan data yang serupa dengan kecerdasan buatan analitis. Namun, kategori kecerdasan buatan ini menekankan pada pelaksanaan tugas nyata secara otomatis, bukan hanya memberikan saran dan rekomendasi teknis.
3. Kecerdasan buatan Interaktif: Dirancang untuk memungkinkan komunikasi yang otomatis dan responsif antara manusia dan sistem. Teknologi ini sering digunakan untuk membuat asisten pribadi pintar dan *chatbot* yang dapat meniru percakapan dengan natural.
4. Kecerdasan buatan Teks: Sistem ini memiliki berbagai kemampuan pemrosesan bahasa alami (*Natural Language*), memungkinkan untuk mendeteksi teks, menyalin ucapan menjadi teks, menerjemahkan antar-bahasa, dan membuat konten teks secara mandiri.
5. Kecerdasan buatan Visual: Berfokus pada kemampuan komputer dalam mengenali, mengategorikan, dan memproses informasi visual dari gambar dan video. Pendekatan ini mengubah data optik menjadi informasi fungsional dan berharga yang meningkatkan pengembangan komputer visual.

2.3 Chatbot

Chatbot adalah kecerdasan buatan (AI) yang memiliki fungsi untuk berkomunikasi atau berinteraksi dengan manusia. Komunikasi *chatbot* pada dasarnya dilakukan dengan *text-based* atau *voice-based* melalui pemrosesan bahasa alami atau *Natural Language Processing* (NLP) yang dapat memungkinkan *chatbot* untuk mengenali, memahami, dan menghasilkan teks atau ucapan yang dapat dipahami oleh manusia [18], [19].

Chatbot pertama kali diperkenalkan pada tahun 1960 oleh Joseph Weizenbaum yang mengembangkan ELIZA sebagai salah satu *chatbot* yang menggunakan program pengenalan pola untuk mensimulasikan dialog seperti manusia. Pada tahun 2020, fenomena *chatbot* pintar pertama kali diperkenalkan oleh OpenAI yaitu GPT-3 sebagai salah satu model AI terkuat yang dapat mampu untuk menghasilkan rangkuman dari artikel yang diberikan,

namun model ini tidak dapat melakukan percakapan seperti manusia, dan hanya dapat menangani satu tugas saja. Pada tahun 2023, Open AI mengembangkan GPT-4 sebagai salah satu model AI yang dapat melakukan pencarian internet untuk meningkatkan akurasi dan jawaban yang akan diberikan kepada pengguna [20].

Seiring berkembangnya waktu, model AI pada *chatbot* sering berkembang yang dapat membuat *chatbot* mampu untuk menghasilkan jawaban yang lebih kompleks dan akurat dalam memahami instruksi manusia. Salah satu model ini disebut sebagai *Large Language Models* (LLMs). LLMs merupakan model AI yang telah dilatih oleh beberapa data dengan jumlah yang besar sehingga AI dapat memprediksi atau menghasilkan jawaban sesuai dengan relevansi terhadap pertanyaan dan pemasukan yang diberikan oleh pengguna [21]. Terdapat beberapa platform *chatbot* berbasis model LLMs yang sering dipakai oleh beberapa pengguna di dunia seperti ChatGPT, Gemini, Deepseek, Perplexity, Claude, Grok dan platform AI lainnya.

ChatGPT merupakan platform AI yang didirikan oleh perusahaan OpenAI sebagai *chatbot* yang digunakan untuk menghasilkan bahasa manusia seperti kalimat, paragraf, dan sebuah dokumen. ChatGPT pada awalnya dirilis pada tahun 2018 sebagai GPT-1 yang awalnya hanya digunakan sebagai mesin penerjemah. Setelah itu GPT-3 sebagai salah satu model AI yang terkuat yang dapat menghasilkan teks berbasis dari pelatihan sebelumnya seperti artikel, halaman website, buku, dan berbagai jenis sumber teks lainnya [22]. Salah satu model yang dipakai oleh ChatGPT untuk saat ini adalah GPT-5, yang merupakan model AI yang dapat menjawab semua pertanyaan secara mendalam dan menyelesaikan berbagai tugas sesuai dengan kebutuhan dan permintaan pengguna [23].

Gemini atau dikenal sebagai Google Bard sebelumnya merupakan *chatbot* AI yang dikembangkan oleh Google pada tahun 2023, pengembangan ini pada awalnya dilakukan secara mendesak terhadap ancaman dari popularitas ChatGPT yang dapat menggantikan kegunaan dan fungsi dari Google Search. Seiring perkembangan waktu, Gemini digunakan sebagai *chatbot* asisten utama bagi Google yang didesain untuk bersifat komunikatif, multimodal, dan bermanfaat bagi pengguna terutama pada pengguna Android [24]. Salah satu model AI terbaru dari Gemini adalah Gemini 3 yang dapat melakukan penelitian lebih dalam, pemahaman terhadap percakapan sebelumnya, dan dapat melakukan tugas kompleks yang diberikan oleh pengguna [25].

2.4 User Experience

Pengalaman Pengguna (UX) adalah studi tentang bagaimana perasaan orang ketika mereka menggunakan suatu produk, sistem, atau layanan [26]. Ketika membuat antarmuka aplikasi, UX merupakan cara untuk mengatasi masalah pengguna dengan mengutamakan kebahagiaan pengguna. Tujuan desain UX adalah untuk mengungkap masalah, mencari tahu apa yang dibutuhkan orang, dan mempelajari konteks pengguna sehingga aplikasi atau sistem tersebut benar-benar bermanfaat dan sesuai dengan orang-orang yang akan menggunakannya. Desain antarmuka yang mudah digunakan dan ramah merupakan fitur kunci dari pengalaman pengguna yang sukses. Desain ini akan mempermudah orang untuk menggunakan aplikasi dengan mempermudah mereka untuk bekerja sama. Hal ini akan membuat segalanya lebih baik bagi pelanggan dan membantu perusahaan mencapai tujuannya.

Tujuan dasar desain UX adalah untuk mengembangkan interaksi yang baik, bermanfaat, dan bermakna. Proyek ini khususnya tentang membuat orang bahagia dengan menawarkan antarmuka yang bagus yang memenuhi kebutuhan mereka dengan sempurna [26]. Tujuan lain dari UX adalah untuk memastikan bahwa sistem mudah digunakan sehingga pengguna tidak bingung dan dapat melakukan pekerjaan mereka dengan cepat dan mudah. Tujuan penting lainnya adalah membuat aplikasi mudah digunakan. Hal ini melibatkan penempatan informasi dan navigasi sedemikian rupa sehingga mengurangi kemungkinan terjadinya kesalahan saat menggunakan aplikasi. Tujuan dari pengalaman yang konsisten secara keseluruhan ini adalah untuk membangun citra merek yang positif, yang akan membuat orang merasa nyaman dan membuat mereka terus kembali untuk menggunakan aplikasi tersebut.

Oleh karena itu, penerapan UX yang matang akan membantu perusahaan dalam berbagai hal nyata, termasuk mempertahankan lebih banyak pelanggan, mengurangi biaya layanan pelanggan, dan membuat produk mereka lebih berharga dan kompetitif di pasar yang ramai [26].

Dalam merancang pengalaman pengguna tersebut, proses desain yang berpusat pada pengguna (*user-centered design*) harus memenuhi beberapa elemen fundamental. Berdasarkan teori yang dikemukakan oleh Jesse James Garrett [27], terdapat lima elemen atau lapisan dalam membentuk user experience pada sebuah produk, yaitu:

1. Strategi (*Strategy*): Merupakan fondasi awal dan mengacu pada kebutuhan pengguna serta tujuan bisnis.

2. Batasan (*Scope*): Penetapan batasan fungsionalitas dan fitur aplikasi yang akan dikembangkan untuk menjawab kebutuhan dasar pengguna.
3. Struktur (*Structure*): Penjelasan konseptual tentang bagaimana interaksi pengguna akan berlangsung dan bagaimana informasi disusun agar efisien.
4. Kerangka (*Skeleton*): Perwujudan nyata dari konsep struktur menjadi desain antarmuka dasar (*wireframe*) tempat elemen interaktif mulai dirancang.
5. Wajah atau Tampilan (*Surface*): Lapisan terluar yang menyediakan pengalaman visual bagi pengguna yang berfokus pada tipografi, warna, gambar, dan tata letak akhir.

Untuk mengetahui apakah elemen-elemen tersebut telah memberikan pengalaman yang baik bagi produk, diperlukan sebuah alat ukur yang valid. Salah satu metode yang komprehensif untuk mengukur UX adalah *HEART Framework*, sebuah metode pengukuran berbasis *user-centered* yang awalnya dikembangkan oleh tim riset Google. Mengutip pemaparan Noer dkk. [28], *HEART Framework* memiliki lima indikator utama, yaitu:

1. *Happiness*: Mengukur aspek subjektif dari pengalaman pengguna seperti tingkat kepuasan, kemudahan penggunaan, dan kemungkinan pengguna merekomendasikan produk tersebut.
2. *Engagement*: Mengukur tingkat keterlibatan atau partisipasi pengguna dengan produk, seperti intensitas kunjungan dan durasi waktu interaksi.
3. *Adoption*: Mengukur seberapa banyak jumlah pengguna baru yang mulai menggunakan produk atau fitur tertentu dalam kurun waktu tertentu.
4. *Retention*: Mengukur tingkat loyalitas pengguna, yaitu seberapa banyak pengguna yang terus kembali dan bertahan menggunakan produk tersebut di periode waktu berikutnya.
5. *Task Success*: Mencakup metrik perilaku tradisional UX seperti tingkat efektivitas, efisiensi waktu penyelesaian tugas, dan tingkat kesalahan (*error rate*) yang dilakukan pengguna saat menggunakan aplikasi.

2.5 User Experience Questionnaire Plus

User Experience Questionnaire (UEQ) adalah skala penelitian yang menggunakan metode kuesioner untuk mengukur pengalaman pengguna melalui berbagai aspek seperti ketertarikan, kemudahan dalam menggunakan, efisien, dapat diandalkan, stimulasi, dan kebaruan terhadap suatu objek yang diteliti. Skala ini dikembangkan Bettina Laughwitz, Theo Held, dan Martin Schrepp pada tahun 2008 melalui penelitian yang berjudul “*Construction and Evaluation of a User Experience Questionnaire*” [29]. UEQ mulai mengalami berbagai perubahan dengan beberapa penambahan skala baru yang membuat

UEQ tidak hanya berfokus pada antarmuka sederhana, namun lebih terfokus pada peralatan yang spesifik seperti peralatan rumah tangga, peralatan kompleks, hingga peralatan yang memberikan status sosial terhadap seseorang. Untuk itu *User Experience Questionnaire Plus* (UEQ+) diciptakan sebagai salah satu metode pengukuran yang lebih fleksibel dan dapat disesuaikan dengan objek secara spesifik.

User Experience Questionnaire Plus (UEQ+) merupakan pendekatan modular yang memberikan fleksibilitas kepada peneliti untuk menentukan skala yang memiliki kesesuaian relevansi terhadap fokus objek yang ingin diteliti. UEQ+ pada dasarnya merupakan ekstensi modular dari *User Experience Questionnaire* (UEQ) yang paling banyak digunakan pada kuesioner untuk mengukur pengalaman pengguna atau user experience (UX) dengan menggunakan 6 skala, yaitu *attractiveness*, *efficiency*, *perspicuity*, *dependability*, *stimulation*, dan *novelty*. Sedangkan UEQ+ berisi daftar skala pengalaman pengguna yang lebih luas dari UEQ dan dapat memberikan peneliti kebebasan dalam menyesuaikan kuesioner yang paling relevan, sehingga objek yang akan diteliti akan menghasilkan hasil penelitian yang lebih terfokus dan konkret [11].

Tabel 2.1 Kategori Produk Pada UEQ+

Kategori Produk	Skala yang Relevan
<i>Word Processing</i>	<i>Dependability, Usefulness, Efficiency, Clarity, Perspicuity</i>
<i>Spreadsheet</i>	<i>Usefulness, Dependability, Efficiency, perspicuity, Clarity</i>
<i>Messenger</i>	<i>Trust, Intuitive Use, Dependability, Efficiency, Identity</i>
<i>Social Networks</i>	<i>Trust, Identity, Dependability, Intuitive Use, Stimulation, Quality of Content, Trustworthiness of Content</i>
<i>Video Conferencing</i>	<i>Trust, Dependability, Efficiency, Intuitive Use, Usefulness</i>
<i>Web Shops</i>	<i>Trust, Quality of Content, Trustworthiness of Content, Dependability, Clarity, Value, Intuitive Use, Visual Aesthetics</i>
<i>News Portals</i>	<i>Quality of Content, Content Reliability, Clarity</i>
<i>Booking Systems</i>	<i>Trust, Dependability, Quality of Content, Trustworthiness of Content, Efficiency, Clarity, Intuitive Use, Value, Usefulness</i>
<i>Info-Web-Sites</i>	<i>Content Quality, Trustworthiness of Content, Clarity</i>
<i>Learning Platforms</i>	<i>Quality of Content, Trustworthiness of Content, Usefulness, Clarity, Perspicuity, Efficiency, Trust, Dependability</i>
<i>Programming Tools</i>	<i>Dependability, Usefulness, Efficiency, Adaptability, Clarity, Perspicuity</i>
<i>Drawing Tools</i>	<i>Dependability, Usefulness, Efficiency, Adaptability, Clarity, Perspicuity</i>

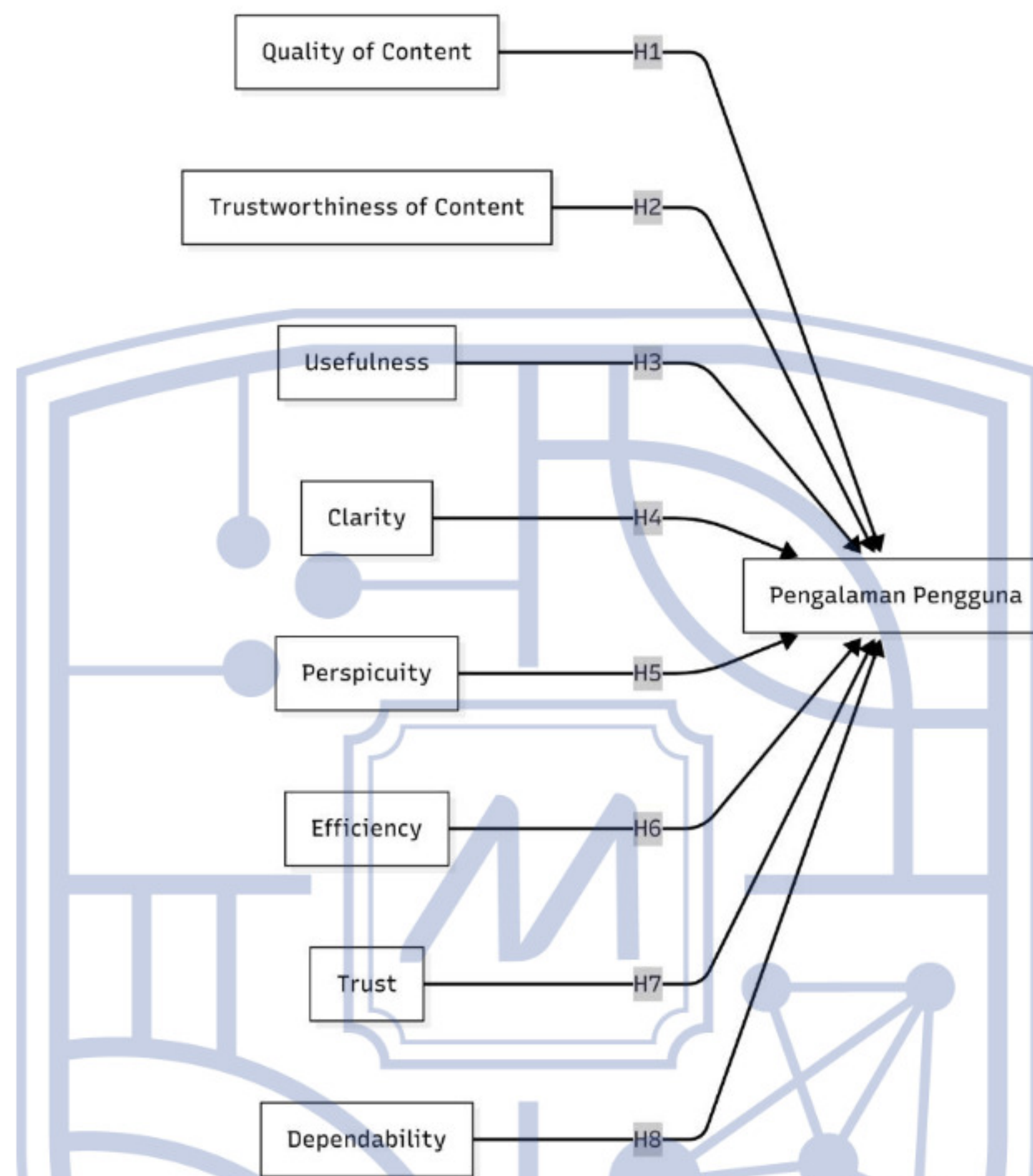
<i>Online-Banking</i>	<i>Trust, Dependability, Quality of Content, Trustworthiness of Content, Value, Clarity, Intuitive Use, Efficiency, Usefulness</i>
<i>Video Portals</i>	<i>Intuitive Use, Immersion, Clarity, Quality of Content, Trustworthiness of Content, Trust</i>
<i>Games</i>	<i>Immersion, Stimulation, Visual Aesthetics, Novelty, Dependability, Intuitive Use</i>
<i>Household Appliance</i>	<i>Usefulness, Intuitive Use, Efficiency, Haptics, Acoustics</i>
<i>Complex Medical Devices</i>	<i>Dependability, Efficiency, Usefulness, Clarity, Result Quality, Trust, Risk Handling, Hardware Security, Perspicuity, Trustworthiness of Content</i>

Untuk penelitian yang mengevaluasi pengalaman pengguna pada platform web Gemini dan ChatGPT dalam konteks mahasiswa, UEQ+ digunakan karena memiliki pendekatan modular yang memungkinkan pemilihan skala berdasarkan kebutuhan dan konteks penelitian. Salah satu template yang tersedia dalam UEQ+ adalah *Learning Platforms*, yang terdiri atas delapan skala, yaitu *Quality of Content*, *Trustworthiness of Content*, *Usefulness*, *Clarity*, *Perspicuity*, *Efficiency*, *Trust*, dan *Dependability*.

Pemilihan template *Learning platforms* dalam penelitian ini bukan didasarkan pada karakteristik responden sebagai mahasiswa, tetapi juga sebagai pertimbangan terhadap kesesuaian aspek yang diukur dengan karakteristik dari penggunaan Gemini dan ChatGPT pada konteks penelitian. Kedua platform digunakan untuk memperoleh dan mengolah informasi dalam membantu penyelesaian tugas dan aktivitas akademik pengguna. Oleh karena itu, aspek *Quality of Content* dan *Trustworthiness of Content* menjadi salah satu aspek terpenting dalam mengevaluasi informasi yang diberikan oleh sistem, sedangkan aspek *Usefulness*, *Clarity*, *Perspicuity*, dan *Efficiency* digunakan untuk menilai manfaat serta kemudahan pengguna dalam berinteraksi dengan platform. Sementara *Trust* dan *Dependability* digunakan untuk menilai kepercayaan layanan serta keadaan sistem dari pandangan pengguna.

Dengan pertimbangan tersebut, template *Learning Platforms* digunakan sebagai ladsan terhadap pemilihan kumpulan skala UEQ+ yang dinilai sesuai dengan aspek pengalaman pengguna yang ingin dievaluasi dalam penelitian ini. Penggunaan template tersebut tidak dimaksudkan untuk mengklasifikasi Gemini dan ChatGPT sebagai platform pembelajaran secara keseluruhan, melainkan sebagai pemilihan skala pengukuran yang sesuai dengan konteks penelitian. Delapan skala tersebut akan digunakan sebagai variabel

dalam kerangka konseptual penelitian untuk memperoleh pengukuran pengalaman pengguna yang lebih terarah dan terstruktur [30].



Gambar 2.1 Model UEQ+ Pada Kategori *Learning Platforms*

Berikut adalah penjelasan dari masing-masing skala yang ditampilkan pada Gambar 2.1 adalah sebagai berikut:

1. *Quality of Content* (Kualitas Konten), skala ini menentukan penyediaan informasi terhadap suatu produk sesuai dengan fakta dan tersusun secara baik.
2. *Trustworthiness of Content* (Tingkat Kepercayaan Terhadap Konten), skala ini menentukan informasi yang disediakan oleh produk sangat berkualitas dan dapat diandalkan.
3. *Usefulness* (Kegunaan), skala ini menentukan manfaat terhadap kegunaan suatu produk.
4. *Clarity* (Kejelasan Tampilan), skala ini menentukan struktur dan kejelasan terhadap antarmuka suatu produk.
5. *Perspicuity* (Kejelasan), skala ini menentukan kemudahan dalam mempelajari cara kerja suatu produk.

6. *Efficiency* (Efisiensi), skala ini menentukan tugas yang dapat dilakukan oleh produk tanpa adanya usaha yang tidak diperlukan.
7. *Trust* (Kepercayaan), skala ini menentukan keamanan terhadap sensitivitas data pengguna dari suatu produk.
8. *Dependability* (Tingkat Keandalan), skala ini menentukan keandalan terhadap kebutuhan pada suatu produk.

Kuesioner akan dirancang sesuai dengan 8 skala pengalaman pengguna pada kategori produk *Learning Platforms* yang disediakan oleh UEQ+. Setiap skala dikelompokkan menjadi 4 *item* pertanyaan dan 1 pertanyaan tambahan, jika dijumlahkan maka terdapat 40 pertanyaan per objek yang ingin diteliti. Akan tetapi penelitian berfokus pada dua objek, yaitu Gemini dan ChatGPT, maka terdapat 80 pertanyaan yang akan diajukan kepada responden dalam memberikan jawaban berdasarkan 1-7 penilaian skala *semantic differential*.

Tabel 2.2 Skala Tingkat Kualitas Konten

1.Kualitas konten (<i>Quality of Content</i>)								
Menurut saya, informasi dan data pada produk ini								
usang	☐	☐	☐	☐	☐	☐	☐	terbaru
tidak menarik	☐	☐	☐	☐	☐	☐	☐	menarik
tidak dipersiapkan dengan baik	☐	☐	☐	☐	☐	☐	☐	dipersiapkan dengan baik
tidak dapat dipahami	☐	☐	☐	☐	☐	☐	☐	dapat dipahami
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

Tabel 2.3 Skala Tingkat Kepercayaan Terhadap Konten

2.Tingkat Kepercayaan Terhadap Konten (<i>Trustworthiness of Content</i>)								
Menurut saya, informasi dan data pada produk ini								
tidak berguna	☐	☐	☐	☐	☐	☐	☐	berguna
tidak masuk akal	☐	☐	☐	☐	☐	☐	☐	masuk akal
tidak dapat dipercaya	☐	☐	☐	☐	☐	☐	☐	terpercaya
tidak akurat	☐	☐	☐	☐	☐	☐	☐	akurat
Saya pikir ciri produk yang dijelaskan di atas sebagai								

Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting
---------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	----------------

Tabel 2.4 Skala Tingkat Kegunaan

3.Kegunaan (<i>Usefulness</i>)								
Saya pikir produk ini								
tidak berguna	☐	☐	☐	☐	☐	☐	☐	berguna
tidak membantu	☐	☐	☐	☐	☐	☐	☐	membantu
tidak menguntungkan	☐	☐	☐	☐	☐	☐	☐	menguntungkan
tidak bermanfaat	☐	☐	☐	☐	☐	☐	☐	bermanfaat
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

Tabel 2.5 Skala Tingkat Kejelasan Tampilan

4.Kejelasan Tampilan (<i>Clarity</i>)								
Menurut pendapat saya, tampilan antarmuka produk terlihat								
Dikelompokkan dengan buruk	☐	☐	☐	☐	☐	☐	☐	Dikelompokkan dengan baik
Tidak terstruktur	☐	☐	☐	☐	☐	☐	☐	Terstruktur
Tidak teratur	☐	☐	☐	☐	☐	☐	☐	Terurut
Tidak teratur	☐	☐	☐	☐	☐	☐	☐	Teratur
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

Tabel 2.6 Skala Tingkat Kejelasan

5.Kejelasan (<i>Perspiciuity</i>)								
Menurut saya cara menggunakan produk ini								
tidak dapat dipahami	☐	☐	☐	☐	☐	☐	☐	dapat dipahami
sulit dipelajari	☐	☐	☐	☐	☐	☐	☐	mudah dipelajari
rumit	☐	☐	☐	☐	☐	☐	☐	mudah
membingungkan	☐	☐	☐	☐	☐	☐	☐	jelas
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

Tabel 2.7 Skala Tingkat Efisiensi

6.Efisiensi (<i>Efficiency</i>)								
Untuk mencapai tujuan saya, menurut saya produk ini								
lambat	☐	☐	☐	☐	☐	☐	☐	cepat
tidak efisien	☐	☐	☐	☐	☐	☐	☐	efisien
tidak praktis	☐	☐	☐	☐	☐	☐	☐	praktis
tidak terstruktur	☐	☐	☐	☐	☐	☐	☐	terstruktur
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

Tabel 2.8 Skala Tingkat Kepercayaan

7.Tingkat Kepercayaan (<i>Trust</i>)								
Berkaitan dengan penggunaan informasi dan data pribadi saya, produk ini								
tidak aman	☐	☐	☐	☐	☐	☐	☐	aman
tidak dapat dipercaya	☐	☐	☐	☐	☐	☐	☐	dapat dipercaya
tidak dapat diandalkan	☐	☐	☐	☐	☐	☐	☐	dapat diandalkan
tidak transparan	☐	☐	☐	☐	☐	☐	☐	transparan
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

Tabel 2.9 Skala Tingkat Keandalan

8.Tingkat Keandalan (<i>Dependability</i>)								
Menurut saya, respons dari produk setelah saya berikan masukan dan perintah								
tidak dapat diprediksi	☐	☐	☐	☐	☐	☐	☐	dapat diprediksi
menghambat	☐	☐	☐	☐	☐	☐	☐	mendukung
tidak terjamin	☐	☐	☐	☐	☐	☐	☐	terjamin
tidak sesuai harapan	☐	☐	☐	☐	☐	☐	☐	sesuai harapan
Saya pikir ciri produk yang dijelaskan di atas sebagai								
Sama sekali tidak relevan	☐	☐	☐	☐	☐	☐	☐	Sangat penting

2.6 Pengujian Kualitas Data

Proses pengujian kualitas data pada penelitian menggunakan alat ukur (instrumen) yang dapat diterima sesuai dengan standar yang telah diuji melalui validitas dan reliabilitas

data, uji pengukuran ini bertujuan untuk mengukur seberapa tepat suatu data dan sejauh mana instrumen yang digunakan dapat dipercaya.

2.6.1 Uji Validitas

Uji validitas adalah langkah pengujian yang dilakukan terhadap isi suatu konten penelitian dari suatu instrumen dengan tujuan untuk mengukur ketepatan instrumen yang akan digunakan pada suatu penelitian. Uji validitas juga digunakan untuk menunjukkan sejauh mana alat ukur yang digunakan dalam mengukur suatu objek yang diukur. Untuk itu kuesioner pada penelitian dapat dikatakan valid jika pertanyaan suatu kuesioner dapat mengungkapkan sesuatu yang dapat diukur oleh kuesioner.

Uji validitas pada penelitian ini menggunakan program SPSS (*Statistical Package for the Social Sciences*) yang merupakan perangkat lunak yang digunakan untuk membandingkan hasil probabilitas koefisien korelasi r dengan taraf signifikan 5%.

Semakin tinggi validitas instrumen yang ditunjukkan maka semakin akurat alat pengukur dalam mengukur suatu data. Jika $r_{hitung} \geq$ dari nilai r_{tabel} maka hasilnya dapat dinyatakan valid, akan tetapi jika $r_{hitung} <$ dari nilai r_{tabel} maka hasil dinyatakan tidak valid [31], [32].

Berikut adalah rumus yang digunakan untuk melakukan uji validitas:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \dots\dots\dots (1)$$

Keterangan pada setiap nilai:

r = koefisien korelasi

n = jumlah responden

x = skor pertanyaan tiap nomor

y = jumlah skor total pertanyaan

2.6.2 Uji Reliabilitas

Uji reliabilitas adalah metode pengujian untuk menunjukkan sejauh mana suatu instrumen atau alat ukur dapat dipercaya dalam mengukur data. Teknik yang digunakan pada pengukuran reliabilitas menggunakan teknik *Cronbach Alpha* (α) yang merupakan uji

koefisien terhadap nilai jawaban responden yang dihasilkan dari penggunaan instrumen penelitian. Suatu variabel dapat dikatakan reliabel jika hasil $\alpha > 0,60$ maka data dapat dikatakan reliabel, akan tetapi jika hasil $\alpha < 0,60$ maka data dinyatakan tidak reliabel [33], [34].

Berikut rumus yang digunakan untuk melakukan uji reliabilitas dengan metode *Cronbach Alpha*:

$$r_i = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum s_i^2}{s_t^2} \right] \dots \dots \dots (2)$$

Keterangan pada setiap nilai:

r_i = koefisien reliabilitas instrumen

k = jumlah pertanyaan yang valid

s_t^2 = varian skor total

$\sum s_i^2$ = jumlah varian skor setiap item

2.7 Uji Normalitas

Uji normalitas adalah jenis pengujian yang digunakan oleh peneliti untuk mengetahui apakah variabel terhadap suatu data bersifat normal atau tidak. Jika data normal maka peneliti dapat menggunakan statistik parametrik, jika data tidak normal maka peneliti hanya dapat menggunakan statistik non parametrik [30].

Salah satu metode statistik yang dapat digunakan oleh peneliti untuk mencari distribusi normalitas data dengan sampel di atas 300 dengan menggunakan metode *Skewness & Kurtosis*. *Skewness* adalah pengukuran terhadap simetri suatu distribusi data, jika *Skewness* berada pada 0 maka dapat dikatakan bahwa distribusi data bersifat asimetri, jika ekor memanjang ke kanan dari distribusi data maka dapat dikatakan bahwa *Skewness* bersifat *positive skew*, begitu juga sebaliknya jika ekor memanjang ke kiri dari distribusi data maka dapat dikatakan bahwa *Skewness* bersifat *negative skew*. *Kurtosis* adalah pengukuran terhadap puncak suatu distribusi data, jika puncak *excess kurtosis* berada pada 0 maka distribusi data bersifat normal sempurna, jika puncak *excess kurtosis* bersifat positif dapat dikatakan bahwa distribusi puncak suatu data bersifat tinggi (*leptokurtic*), begitu juga jika puncak *excess kurtosis* bersifat negatif dapat dikatakan bahwa distribusi puncak suatu data bersifat rendah (*platykurtic*) [35].

Berikut adalah pengujian normalitas data dengan menggunakan *Skewness & Kurtosis* melalui *z-test*.

$$Z = \frac{\text{Skew value}}{SE_{\text{skewness}}}, Z = \frac{\text{Excess kurtosis}}{SE_{\text{excess kurtosis}}} \dots\dots\dots (3)$$

Normalitas data pada pengujian *Skewness & Kurtosis* berdasarkan jumlah sampel yang digunakan. Berikut adalah jumlah sampel dan syarat untuk memenuhi normalitas suatu distribusi data.

Tabel 2.10 Syarat Normalitas Data *Skewness & Kurtosis*

Ukuran Sampel	Syarat Normalitas Data
Sampel kecil ($n < 50$)	Jika <i>z-score</i> pada <i>Skewness</i> dan <i>kurtosis</i> di bawah 1,96, maka distribusi data bersifat normal.
Sampel menengah ($50 < n < 300$)	Jika <i>z-score</i> pada <i>Skewness</i> dan <i>kurtosis</i> di bawah 3,29, maka distribusi data bersifat normal.
Sampel besar (> 300)	Jika nilai pada <i>skewness</i> di bawah 2 dan <i>kurtosis</i> di bawah 7, maka distribusi data bersifat normal.

2.8 Uji Hipotesis

Hipotesis adalah dugaan sementara atau kesimpulan sementara terhadap pertanyaan-pertanyaan yang diajukan dalam rumusan masalah suatu penelitian. Oleh karena itu, pengujian hipotesis pada dasarnya adalah proses evaluasi statistik yang penting untuk membuktikan kebenaran dugaan tersebut. Melalui pengujian ini, peneliti dapat secara objektif menentukan apakah hipotesis awal dapat diterima atau harus ditolak berdasarkan bukti empiris dari data yang dikumpulkan. Berdasarkan penelitian kuantitatif, uji hipotesis, data yang digunakan harus dianalisis dengan menggunakan uji normalitas. Jika data menunjukkan distribusi normal maka dilakukan analisis parametrik. Apabila data tidak berdistribusi normal maka dilakukan analisis non parametrik [36], [37].

2.8.1 Uji Parsial (Uji T)

Uji t (Uji Parsial) digunakan untuk mengetahui apakah masing-masing variabel independen berpengaruh secara signifikan terhadap variabel dependen. Uji ini menunjukkan seberapa besar pengaruh variabel independen dalam menjelaskan variasi variabel dependen

dengan asumsi variabel dependen lainnya konstan. Kriteria pengujiannya adalah sebagai berikut [38], [39]:

1. Jika nilai $t_{hitung} > t_{tabel}$, maka H_0 ditolak dan H_1 diterima. Hal ini berarti variabel independen secara parsial memiliki pengaruh yang positif dan signifikan terhadap variabel dependen.
2. Jika nilai $t_{hitung} < t_{tabel}$, maka H_0 diterima dan H_1 ditolak. Hal ini berarti variabel independen secara parsial tidak memiliki pengaruh yang signifikan terhadap variabel dependen.

2.8.2 Uji Simultan (Uji F)

Uji F (uji simultan) digunakan untuk menentukan apakah, secara kolektif, variabel independen memiliki pengaruh signifikan terhadap variabel dependen. Uji ini juga bertujuan untuk mengukur pengaruh variabel independen dengan variabel dependen [37], [40]. Dalam pengujiannya nilai F pada tabel *ANOVA*, dengan tingkat nilai signifikansi tersebut dibandingkan dengan batas toleransi kesalahan yang telah ditetapkan, yaitu sebesar 0,05. Adapun kriteria pengambilan keputusannya diuraikan sebagai berikut [41]:

1. Apabila nilai signifikansi $F \leq 0,05$, maka H_0 ditolak dan H_1 diterima. Artinya, semua variabel independen (bebas) secara simultan terbukti memiliki pengaruh yang signifikan terhadap variabel dependen (terikat).
2. Apabila nilai signifikansi $F > 0,05$, maka H_0 diterima dan H_1 ditolak. Artinya, semua variabel independen (bebas) secara simultan tidak memiliki pengaruh yang signifikan terhadap variabel dependen (terikat).

2.8.3 Uji Koefisien Determinasi (R^2)

Uji koefisien determinasi (R^2) digunakan untuk mengukur persentase seberapa besar variabel independen mempengaruhi penjelasan variabel dependen [38], [41].

Nilai Koefisien selalu antara 0 dan 1. Kedekatan angka 1 menunjukkan model tersebut semakin baik, karena variabel independen mampu menjelaskan fluktuasi dari variabel dependennya dengan sangat kuat. Di sisi lain, jika nilainya mendekati 0, kemampuan variabel independen dalam menjelaskan variabel dependen sangat rendah [41].

2.8.4 Paired Sample t-Test

Paired Sample t-Test adalah uji statistik parametrik yang digunakan untuk membandingkan dua kelompok data berpasangan untuk melihat apakah ada perbedaan

signifikan dalam rata-rata antara keduanya. Data dari subjek atau responden yang sama diukur di bawah dua kondisi berbeda seperti sebelum dan setelah evaluasi intervensi atau perbandingan antara sistem A dan B. Uji ini mensyaratkan perbedaan nilai antara kedua kondisi harus terdistribusi normal. Kriteria pengambilan keputusan uji *Paired Sample t-Test* berdasarkan nilai signifikansi dengan $\alpha = 0,05$ dan ditentukan dari nilai signifikansi hasil adalah sebagai berikut [42]:

1. Jika nilai Sig. $\leq 0,05$, maka H_0 ditolak dan H_1 diterima, yang berarti terdapat perbedaan rata-rata yang signifikan antara dua kondisi yang diuji.
2. Jika nilai Sig. $> 0,05$, maka H_0 diterima dan H_1 ditolak, yang berarti tidak terdapat perbedaan rata-rata yang signifikan antara dua kondisi yang diuji.

2.9 Populasi dan Sampel

2.9.1 Populasi

Populasi dalam konteks penelitian adalah suatu hal yang merujuk pada suatu kelompok, yang memiliki karakteristik tertentu sehingga menjadi suatu objek yang ingin diteliti. Populasi sendiri dapat terdiri dari individu, objek, peristiwa, dan hal lain memiliki keterkaitan dengan penelitian yang sedang dilakukan untuk mendapatkan sebuah kesimpulan dan hasil akhir dari suatu penelitian [43].

2.9.2 Sampel

Sampel dalam konteks penelitian adalah bagian dari individu atau entitas dari suatu populasi yang dipilih untuk mewakili karakteristik dari setiap populasi. Tujuan sampel digunakan oleh peneliti untuk memahami suatu karakteristik dari populasi tanpa perlu mempelajari setiap individu yang terdapat pada populasi, dengan kata lain populasi menjadi salah satu teknik yang digunakan oleh peneliti untuk mengatasi keterbatasan dana, tenaga, dan waktu terhadap penelitian yang sedang dilakukan [43].

2.10 Kuesioner

Kuesioner adalah sebuah teknik yang sering digunakan pada penelitian kuantitatif untuk mengumpulkan data melalui serangkaian pertanyaan yang telah dirancang dengan tujuan untuk mengukur suatu variabel yang telah direncanakan dari suatu penelitian. Kuesioner pada penelitian kuantitatif umumnya menggunakan pengumpulan data numerik yang akan dianalisis lebih lanjut untuk mendapatkan hasil akhir dari suatu penelitian [44].

2.11 Skala *Semantic Differential*

Skala *Semantic Differential* adalah skala pengukuran yang digunakan untuk menilai suatu masalah atau objek berdasarkan sifat dari setiap individu. Penilaian pada skala ini terdiri dari 1-7 penilaian (*rating*) dari kata yang saling berlawanan (*bipolar*) seperti "baik/buruk", "kuat/lemah", "cepat/lambat", dan sebagainya. Skala ini pertama kali dikembangkan oleh Charles Osgood pada tahun 1952 sebagai salah satu teknik pengukuran yang menggunakan kata sifat sebagai cara untuk menilai suatu objek berdasarkan alasan dari kata sifat yang saling berlawanan [45].

2.12 Teknik Sampel

Teknik pengambilan sampel (*sampling*) merupakan teknik yang digunakan oleh peneliti untuk menentukan sampel atau bagian dari jumlah karakteristik yang terdapat pada populasi untuk digunakan pada penelitian [30].

Tujuan dari teknik pengambilan sampel digunakan untuk mempelajari hubungan dan korelasi terhadap variabel dan sasaran populasi untuk menghasilkan kesimpulan terhadap keterkaitan teori yang digunakan [46].

Jenis-jenis teknik *sampling* yang ada, yaitu:

1. *Probability Sampling*

Probability Sampling (Pengambilan Sampel Probabilitas) adalah teknik di mana setiap anggota yang terdapat pada populasi memiliki kesempatan yang untuk dijadikan sebagai sampel. Beberapa jenis *Probability Sampling*, yaitu:

a. *Simple Random Sampling*

Simple Random Sampling (Pengambilan Sampel Acak Sederhana) merupakan metode pengambilan yang memerlukan kerangka lengkap dari seluruh unit pada populasi, teknik ini dapat disebar secara geografis namun diperlukan biaya yang sangat tinggi untuk mendapatkan sampel.

b. *Systematic Sampling*

Systematic Sampling (Pengambilan Sampel Sistematis) merupakan metode pengambilan yang dilakukan acak secara awal atau disebut dengan nilai n yang akan terus mewakili pengambilan sampel selanjutnya tanpa melakukan proses pengacakan.

c. *Stratified Random Sampling*

Stratified Random Sampling (Pengambilan Sampel Acak Bertingkat) merupakan metode pengambilan yang membagi populasi menjadi strata atau sub kelompok untuk dijadikan sebagai sampel acak.

d. Cluster Sampling

Cluster Sampling (Pengambilan Sampel Klaster) merupakan metode pembagian kelompok terhadap suatu populasi yang akan digunakan untuk proses pengambilan sampel acak sampai sampel akhir.

e. Multi-stage Sampling

Multi-stage Sampling (Pengambilan Sampel Multi-Tahap) merupakan metode perpindahan dari sampel yang luas ke sampel yang sempit dengan menggunakan proses bertahap untuk mendapatkan sampel yang lebih terkonsentrasi pada beberapa wilayah geografis.

2. Non-probability Sampling

Non-probability Sampling (Pengambilan Sampel Tidak Acak) adalah metode pengambilan sampel yang berfokus pada penelitian kualitatif yang memiliki fokus pada sampel kecil yang akan digunakan untuk meneliti suatu fenomena yang terdapat pada kehidupan nyata. Beberapa jenis *Non-probability Sampling*, yaitu:

a. Quota Sampling

Quota Sampling (Pengambilan Sampel Kuota) merupakan metode pengambilan sampel di mana partisipan dipilih berdasarkan kesesuaian terhadap karakteristik yang telah ditentukan pada fokus penelitian.

b. Snowball Sampling

Snowball Sampling (Pengambilan Sampel Bola Salju) merupakan metode pengambilan sampel yang menggunakan beberapa kasus untuk mendorong kasus lain menjadi bagian dari penelitian untuk meningkatkan ukuran sampel yang diperlukan.

c. Convenience Sampling

Convenience Sampling (Pengambilan Sampel Keinginan) merupakan metode pengambilan sampel yang pada umumnya partisipan mudah tersedia pada setiap waktu.

d. Purposive or Judgment Sampling

Purposive or Judgment Sampling (Pengambilan Sampel yang Bertujuan atau Pertimbangan) merupakan metode pengambilan sampel dengan menggunakan strategi melalui peristiwa tertentu untuk memberikan informasi penting yang tidak dapat diperoleh dari pilihan lain.

2.13 Slovin

Rumus *Slovin* adalah rumus yang digunakan oleh peneliti untuk menghitung ukuran sampel yang sesuai dari suatu populasi. Rumus ini digunakan pada penelitian ketika peneliti tidak mengetahui bagaimana atribut yang diukur akan terdistribusi dalam populasi suatu penelitian [47].

Berikut adalah rumus slovin untuk mendapatkan jumlah sampel penelitian:

$$n = \frac{N}{1+N(e)^2} \dots \dots \dots (4)$$

Keterangan:

n adalah ukuran sampel

N adalah total populasi

e adalah tingkat presisi pada penelitian (*margin of error*)

2.14 Penelitian Terdahulu

Berikut ini adalah penelitian terdahulu yang menggunakan metode UEQ dan UEQ+ oleh peneliti untuk menjadi sumber pendukung:

Tabel 2.11 Penelitian Terdahulu

No	Nama & Tahun penelitian	Judul Penelitian	Variabel Penelitian	Hasil Penelitian
1.	Zulpa Salsabila, Fandi Halim, Ariel Febryan First Putra Lumban Gaol, Anggi Alfredo Hutaeruk (2023)	Penggunaan User <i>Experience Questionnaire Plus</i> (UEQ+) untuk Mengevaluasi Pengalaman Pengguna Aplikasi MyPertamina [48]	Menggunakan 8 variabel skala UEQ+: 1. <i>Quality of Content</i> 2. <i>Trustworthiness of Content</i> 3. <i>Clarity</i> 4. <i>Value</i> 5. <i>Visual Aesthetics</i> 6. <i>Intuitive Use</i> 7. <i>Trust</i> 8. <i>Dependability</i>	Penelitian ini melakukan evaluasi terhadap pengalaman pengguna pada aplikasi MyPertamina menggunakan metode UEQ+. Dari hasil penelitian menunjukkan bahwa aplikasi ini mendapatkan <i>positive evaluation</i> (nilai <i>mean</i> > 0,8) terhadap setiap skala pengukuran yang membuat aplikasi ini bernilai positif terhadap pengalaman pengguna.

2.	Desi Arisandy, Riche, Julie Eva Shinta, Cindy Kalyana (2022)	Pengukuran dan Evaluasi Pengalaman Pengguna Aplikasi WhatsApp dan Telegram dengan Metode UEQ+ pada Mahasiswa Universitas Mikroskil [49]	Menggunakan variabel skala UEQ+: 1. <i>Trust</i> 2. <i>Intuitive Use</i> 3. <i>Dependability</i> 4. <i>Efficiency</i>	4	Penelitian ini melakukan pengukuran dan evaluasi terhadap pengalaman pengguna pada aplikasi WhatsApp dan Telegram yang berfokus pada mahasiswa Universitas Mikroskil. Dari hasil penelitian ini kedua aplikasi mendapatkan <i>positive evaluation</i> (nilai <i>mean</i> > 0.8) terhadap setiap skala pengukuran sehingga kedua aplikasi bernilai positif terhadap pengalaman pengguna. Selain itu terdapat hasil uji perbandingan menunjukkan bahwa terdapat perbedaan signifikan terhadap <i>Intuitive Use</i>, <i>Dependability</i>, dan <i>Efficiency</i> yang membuat WhatsApp lebih unggul terhadap pengalaman pengguna.
3.	Arif Setia Sandi Ariyanto, Deny Nugroho Triwibowo, Indah Trivilia (2025)	Analisis Pengalaman Pengguna terhadap Platform Konferensi Video untuk Mendukung Perkuliahan: Studi Evaluasi Menggunakan UEQ+ [50]	Menggunakan variabel skala UEQ+: 1. <i>Attractiveness</i> 2. <i>Efficiency</i> 3. <i>Intuitive Use</i> 4. <i>Visual Aesthetics</i> 5. <i>Quality of Content</i> 6. <i>Trust</i>	6	Penelitian ini melakukan analisis terhadap pengalaman pengguna terhadap platform konferensi video seperti Zoom, Google Meet, dan Microsoft Teams sebagai alat untuk mendukung perkuliahan dengan metode UEQ+. Dari hasil penelitian menunjukkan bahwa ketiga aplikasi mendapatkan <i>positive evaluation</i> (nilai <i>mean</i> > 0.8) terhadap setiap skala pengukuran yang membuat ketiga aplikasi

				bernilai positif terhadap pengalaman pengguna.
4.	Galih Aditama, Bambang Soedijono, & Alva Hendi Muhammad (2023)	Pengukuran <i>User Experience</i> Penggunaan Krapyak-U Menggunakan <i>Framework User Experience Questionnaire Plus</i> [51]	Menggunakan 8 variabel skala UEQ+: 1. <i>Efficiency</i> 2. <i>Perspiciuity</i> 3. <i>Dependability</i> 4. <i>Trust</i> 5. <i>Usefulness</i> 6. <i>Intuitive Use</i> 7. <i>Quality of Content</i> 8. <i>Clarity</i>	Penelitian ini melakukan analisis terhadap pengalaman pengguna pada aplikasi Krapyak-U dengan menggunakan metode UEQ+. Dari hasil penelitian menunjukkan bahwa aplikasi ini memiliki nilai <i>mean</i> > 1 pada setiap skala yang membuat aplikasi ini bernilai sangat positif pada pengalaman pengguna.
5.	Feby Rahma & Miftahul Huda (2024)	Analisis <i>User Experience</i> (UX) Menggunakan Metode <i>User Experience Questionnaire Plus</i> (UEQ+) Pada Aplikasi Perpustakaan Digital Nasional Republik (iPusnas) [52]	Menggunakan 13 variabel skala UEQ+: 1. <i>Efficiency</i> 2. <i>Usefulness</i> 3. <i>Perspiciuity</i> 4. <i>Adaptability</i> 5. <i>Dependability</i> 6. <i>Intuitive Use</i> 7. <i>Stimulation</i> 8. <i>Novelty</i> 9. <i>Trust</i> 10. <i>Attractiveness</i> 11. <i>Visual Aesthetics</i> 12. <i>Clarity</i> 13. <i>Value</i>	Penelitian ini melakukan analisis terhadap pengalaman pengguna pada aplikasi iPusnas dengan menggunakan metode UEQ+. Dari hasil penelitian aplikasi ini mendapatkan berbagai kelebihan dan kekurangan pada setiap skala yang membuat aplikasi ini bernilai sangat <i>positive evaluation</i> (nilai <i>mean</i> > 0.8) pada aspek <i>usefulness</i> , <i>stimulation</i> , <i>perspiciuity</i> , <i>intuitive use</i> , <i>efficiency</i> , dan <i>clarity</i> .
6.	Widya Ariyanti, & RG Guntur Alam (2025)	Analisis Kepuasan Pengguna ChatGPT Gemini dan di Kalangan Mahasiswa Menggunakan Metode UEQ [7]	Menggunakan 6 variabel skala UEQ 1. <i>Attractiveness</i> 2. <i>Perspiciuity</i> 3. <i>Efficiency</i> 4. <i>Dependability</i> 5. <i>Stimulation</i> 6. <i>Novelty</i>	Penelitian ini melakukan analisis terhadap kepuasan pengguna pada aplikasi ChatGPT dan Gemini di kalangan mahasiswa dengan menggunakan metode UEQ. Hasil penelitian menunjukkan bahwa

					ChatGPT mendapat <i>positive evaluation</i> (nilai <i>mean</i> > 0.8) pada skala <i>attractiveness</i> , <i>perspicuity</i> , <i>efficiency</i> , <i>dependability</i> , <i>stimulation</i> yang membuat ChatGPT lebih unggul terhadap pengalaman pengguna dibandingkan Gemini yang memiliki nilai <i>mean</i> < 0,8 pada setiap skala pengukuran.
7.	Ilham Nasrul, Muhammad Luthfi Hamzah, & Eki Saputra (2024)	<i>Analysis Of User Experience Of ChatGPT And Gemini Users Using The User Experience Questionnaire (UEQ) For Education</i> [9]	Menggunakan variabel skala UEQ 1. <i>Attractiveness</i> 2. <i>Perspicuity</i> 3. <i>Efficiency</i> 4. <i>Dependability</i> 5. <i>Stimulation</i> 6. <i>Novelty</i>	6	Penelitian ini melakukan analisis terhadap pengalaman pengguna pada ChatGPT dan Gemini untuk kebutuhan edukasi dengan metode UEQ+. Dari hasil penelitian menunjukkan bahwa ChatGPT dan Gemini memiliki <i>mean</i> < 0.8, yang membuat kedua platform AI tidak unggul terhadap pengalaman pengguna.
8.	Ulya Nur Chamidah, Hario Jati Setyadi, & Muhammad Rivani Ibrahim (2025)	Evaluasi pengalaman penggunaan <i>Artificial Intelligence</i> pada ChatGPT dan Gemini dalam proses pembelajaran di Samarinda dengan metode <i>User Experience Questionnaire</i> [10]	Menggunakan variabel skala UEQ 1. <i>Attractiveness</i> 2. <i>Perspicuity</i> 3. <i>Efficiency</i> 4. <i>Dependability</i> 5. <i>Stimulation</i> 6. <i>Novelty</i>	6	Penelitian ini melakukan analisis terhadap pengalaman pengguna AI seperti ChatGPT dan Gemini dalam mendukung proses pembelajaran di Samarinda dengan metode UEQ. Hasil penelitian menunjukkan bahwa kedua AI mendapatkan <i>positive evaluation</i> (nilai <i>mean</i> > 0.8) pada setiap skala pengukuran yang membuat kedua aplikasi AI bernilai positif terhadap pengalaman pengguna.

				Hasil perbandingan terhadap penelitian ini menunjukkan bahwa ChatGPT sangat unggul terhadap pengalaman terutama pada aspek <i>attractiveness</i> , <i>perspicuity</i> , <i>efficiency</i> , dan <i>stimulation</i> jika dibandingkan dengan Gemini.
9.	Dila Okta Dwi Putri, & Dinda Lestari (2024)	<i>Assessing User Experience of ChatGPT Website Employing the User Experience Questionnaire (UEQ) [53]</i>	Menggunakan variabel skala UEQ 1. <i>Attractiveness</i> 2. <i>Perspicuity</i> 3. <i>Efficiency</i> 4. <i>Dependability</i> 5. <i>Stimulation</i> 6. <i>Novelty</i>	6 Penelitian ini melakukan evaluasi terhadap pengalaman pengguna pada platform web ChatGPT melalui metode UEQ. Dari hasil penelitian menunjukkan bahwa aplikasi ChatGPT telah mendapat <i>average mean</i> (nilai <i>mean</i> > 0.8) pada setiap skala pengukuran yang membuat aplikasi ini bernilai cukup positif pada pengalaman pengguna.
10.	Paolo Brie, Nicolas Burny, Arthur Sluyters, & Jean Vanderdonckt. (2023)	<i>Evaluating a Large Language Model on Searching for GUI Layouts [54]</i>	Menggunakan variabel skala UEQ+: 1. <i>Efficiency</i> 2. <i>Usefulness</i> 3. <i>Perspicuity</i> 4. <i>Adaptability</i> 5. <i>Dependability</i> 6. <i>Intuitive Use</i> 7. <i>Stimulation</i> 8. <i>Novelty</i> 9. <i>Trustworthiness of content</i> 10. <i>Attractiveness</i> 11. <i>Visual Aesthetics</i> 12. <i>Quality of Content</i> 13. <i>Value</i>	13 Penelitian ini melakukan evaluasi terhadap model berbasis <i>Large Language Model</i> (LLM) dalam mencari <i>Graphical User Interface</i> (GUI) melalui metode UEQ+. Dari hasil penelitian menunjukkan bahwa LLM mendapatkan nilai positif (<i>mean</i> > 1) pada beberapa aspek skala seperti <i>attractiveness</i> , <i>efficiency</i> , <i>perspicuity</i> , <i>dependability</i> , <i>stimulation</i> , <i>novelty</i> , <i>usefulness</i> , <i>value</i> , <i>intuitive use</i> , <i>trustworthiness of content</i> , dan <i>quality of content</i> .

11.	Joy Ciliani (2025)	<i>University Orientation through Conversational Technologies: Evaluating the User Experience of an Embodied Conversational Agent</i> [55]	Menggunakan 14 variabel skala UEQ+: 1. <i>Attractiveness</i> 2. <i>Efficiency</i> 3. <i>Perspicuity</i> 4. <i>Dependability</i> 5. <i>Stimulation</i> 6. <i>Novelty</i> 7. <i>Trust</i> 8. <i>Acoustics</i> 9. <i>Usefulness</i> 10. <i>Value</i> 11. <i>Visual Aesthetics</i> 12. <i>Intuitive Use</i> 13. <i>Trustworthiness of Content</i> 14. <i>Quality of Content</i>	Penelitian ini melakukan evaluasi pengalaman pengguna pada orientasi universitas dengan teknologi komunikasi melalui <i>conversational agent</i> dengan metode UEQ+. Dari hasil penelitian menunjukkan bahwa setiap skala bernilai <i>mean</i> > 1 yang membuat teknologi ini dinilai positif pada pengalaman pengguna.
12.	Angela, Fandi Halim, Chatrine Sylvia (2022)	Pengukuran Pengalaman Pengguna Aplikasi Platform Pembelajaran dan Konferensi Video Menggunakan <i>Framework</i> UEQ+ [56]	Menggunakan 9 variabel skala UEQ+: 1. <i>Efficiency</i> 2. <i>Perspicuity</i> 3. <i>Dependability</i> 4. <i>Trust</i> 5. <i>Usefulness</i> 6. <i>Intuitive Use</i> 7. <i>Trustworthiness of Content</i> 8. <i>Quality of Content</i> 9. <i>Clarity</i>	Penelitian ini melakukan evaluasi pengalaman pengguna pada aplikasi Microsoft Teams melalui <i>framework</i> UEQ+. Dari hasil penelitian menunjukkan setiap skala bernilai <i>mean</i> > 2 yang membuat Microsoft Teams dinilai sangat baik dan mendapatkan <i>positive evaluation</i> pada pengalaman pengguna.

Berdasarkan 11 penelitian terdahulu yang mayoritas menerapkan kerangka *User Experience Questionnaire* (UEQ) dan *User Experience Questionnaire Plus* (UEQ+), Metode ini telah terbukti mampu mengevaluasi secara komprehensif tingkat kepuasan pengguna pada berbagai objek digital dengan hasil yang konsisten yang mengarah pada evaluasi positif, meskipun tingkat signifikansi dari setiap skala sangat bervariasi tergantung pada karakteristik layanan yang diuji. Jika diklasifikasikan berdasarkan objek perangkat lunak yang diteliti, kelompok pertama didominasi oleh teknologi kecerdasan buatan dan agen percakapan virtual (seperti ChatGPT, Gemini, *Large Language Models*, dan *Conversational*

Agent), di mana skala *Efficiency* (efisiensi), *Dependability* (keandalan), dan *Perspicuity* (kemudahan belajar) Kelompok kedua dari objek penelitian melibatkan aplikasi komunikasi dan kolaborasi (misalnya WhatsApp, Telegram, dan sistem konferensi video) yang menunjukkan bahwa *Efficiency* (efisiensi), *Trust* (kepercayaan), dan *Usefulness* (kegunaan) adalah skala prioritas yang paling umum diukur untuk memfasilitasi pertukaran pesan secara *real-time* yang lancar, dan privasi data antara pengguna tetap terjaga. Sementara, kelompok objek ketiga adalah layanan operasional, transaksional dan portal informasi (seperti MyPertamina, Krapyak-U, dan iPusnas). *Quality of Content* (kualitas konten) dan *Clarity* (kejelasan tampilan) terbukti menjadi kriteria yang paling banyak dievaluasi sebagai kunci keberhasilan dalam menyajikan literatur atau data transaksi yang akurat, lengkap, dan didukung visual antarmuka dengan tidak membingungkan mata. Perbedaan dalam kategori objek aplikasi memperkuat adanya variasi dalam prioritas penggunaan skala, yang menunjukkan bahwa harapan pengguna selalu disesuaikan dengan fungsi dan tujuan spesifik dari sistem yang digunakan. Penelitian ini sangat relevan untuk menguji dan membandingkan profil pengalaman pengguna secara komprehensif pada objek situs web kecerdasan buatan (Gemini dan ChatGPT) untuk menentukan platform mana yang paling kuat dan memuaskan untuk kebutuhan akademis mahasiswa di lingkungan Universitas Mikroskil.

2.15 Pengembangan Hipotesis

Berdasarkan rumusan masalah yang telah dikemukakan sebelumnya, hipotesis pada penelitian ini disusun sebagai jawaban sementara yang akan diuji kebenarannya melalui proses penelitian terdahulu dan tinjauan literatur. Adapun hipotesis yang diajukan dalam penelitian ini:

2.15.1 Tingkat Pengalaman Pengguna Website Gemini dengan Metode UEQ+ pada Mahasiswa Mikroskil

1. Pengaruh *Quality of Content* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Quality of Content* berfokus dengan kedalaman pemahaman, ketepatan, dan kualitas keluaran bahasa oleh model kecerdasan buatan saat memberikan respons. Skala *Quality of Content* mengukur sejauh mana pengguna puas dengan konten informasi yang disampaikan oleh sistem [48]. Pengujian pada model bahasa *Large Language Model* menunjukkan bahwa dimensi *Quality of Content* berhasil meraih skor yang sangat positif

dalam mengevaluasi pemenuhan kebutuhan informasi pengguna [55]. Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Quality of Content* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Quality of Content* signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

2. Pengaruh *Trustworthiness of Content* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Trustworthiness of Content* berfokus membangun kepercayaan pengguna bahwa informasi yang diberikan oleh sistem ini adalah berdasarkan fakta objektif dan tidak bias. Skala *Trustworthiness of Content* mengukur sejauh mana pengguna memiliki kepercayaan pada sumber dan akurasi informasi yang diberikan oleh sistem kepada mereka dan bahwa sistem tidak akan memberikan informasi yang menyesatkan atau salah [48]. Berdasarkan pengujian pada LLM dan aplikasi MyPertamina, aspek kepercayaan terhadap validitas informasi (*Trustworthiness of Content*) mencatat hasil evaluasi yang positif, membuktikan dominasinya dalam membentuk keyakinan pengguna [48], [55]. Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Trustworthiness of Content* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Trustworthiness of Content* berpengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

3. Pengaruh *Usefulness* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Usefulness* berfokus menekankan fitur-fitur pada sebuah platform yang juga dapat meningkatkan nilai nyata dalam rutinitas sehari-hari pengguna. Skala *Usefulness* mengukur pandangan pengguna tentang sejauh mana sistem membantu mereka mencapai tujuan mereka dengan lebih cepat, mudah, atau efisien [48]. Studi evaluasi terhadap aplikasi iPusnas membuktikan bahwa nilai *Usefulness* (kegunaan) memiliki kontribusi yang sangat kuat dan meraih evaluasi paling positif terhadap peningkatan produktivitas penggunanya [52]. Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Usefulness* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Usefulness* berpengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

4. Pengaruh *Clarity* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Clarity* adalah tampilan fungsional dari fungsionalitas antarmuka yang indah yang tidak dimaksudkan untuk membebani kapasitas visual penggunanya. Skala *Clarity* mengukur kesederhanaan dalam memahami struktur, aspek visual, dan navigasi antarmuka pengguna [48]. Temuan pada evaluasi sistem Krapyak-U menunjukkan bahwa aspek kejelasan desain (*Clarity*) sangat efektif dalam memandu kelancaran interaksi dengan skor pencapaian yang sangat tinggi [51]. Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Clarity* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Clarity* berpengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

5. Pengaruh *Perspiciuity* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Perspiciuity* berfokus pada tingkat transparansi dan kenyamanan bagi orang awam untuk menggunakan sistem pada penggunaan awalnya. Skala *Perspiciuity* mengukur kemudahan di mana seorang pengguna dapat mempelajari dan memahami sistem, terutama pada penggunaan pertama [48]. Penelitian pada ChatGPT dan Gemini mengonfirmasi bahwa aspek kemudahan belajar (*Perspiciuity*) merupakan faktor keunggulan kompetitif yang menentukan seberapa cepat pengguna mampu beradaptasi dengan sistem AI baru [7], [10]. Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Perspiciuity* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Perspiciuity* berpengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

6. Pengaruh *Efficiency* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Efficiency* berfokus pada keunggulan komputasi dari sistem untuk menerjemahkan pertanyaan dan menghasilkan balasan dalam siklus waktu terkecil yang

memungkinkan. Skala *Efficiency* mengukur kecepatan dan akurasi respons sistem terhadap masukan pengguna, serta usaha terkecil yang diperlukan untuk menyelesaikan sebuah pekerjaan [48]. bukti empiris pada pengujian aplikasi kecerdasan buatan menunjukkan bahwa dimensi efisiensi (*Efficiency*) menjadi skala yang menonjol secara signifikan, yang membuktikan bahwa respons instan tanpa hambatan memegang peranan krusial bagi perangkat digital [7]. Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Efficiency* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Efficiency* berpengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

7. Pengaruh *Trust* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil.

Skala *Trust* berfokus pada keandalan sistem dalam menciptakan lingkungan digital yang aman dalam hal perlindungan ruang obrolan, pengelolaan data pribadi, pemeliharaan privasi. Skala *Trust* mengukur pada persepsi tentang keamanan dan privasi pengguna selama berinteraksi dengan sistem dan keyakinan bahwa data dan informasi pribadi mereka akan dikelola dengan baik dan tidak disalahgunakan [41]. Penelitian mengenai agen percakapan virtual menegaskan bahwa elemen *Trust* adalah prasyarat fundamental yang meraih evaluasi sangat tinggi untuk menjembatani interaksi yang andal [55].

Hipotesis yang dapat ditarik sebagai berikut:

H₀: *Trust* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a: *Trust* berpengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

8. Pengaruh *Dependability* terhadap Pengalaman Pengguna di Website Gemini pada mahasiswa Mikroskil

Skala *Dependability* terkait dengan keandalan dan jaminan stabilitas operasional sistem dalam setiap sesi penggunaan. Skala *Dependability* mengukur persepsi pengguna terhadap keandalan dan stabilitas operasi sistem, dengan gangguan teknis, kesalahan, atau gangguan layanan yang tidak terduga sekecil mungkin [48]. Kajian empiris pada komparasi layanan platform AI mencatat bahwa stabilitas aplikasi yang dapat diandalkan

(*Dependability*) secara konsisten menjadi skala pembeda yang penting untuk meminimalkan kendala pada sisi pengguna akhir [7]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Dependability* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

H_a : *Dependability* berpengaruh positif dan signifikan terhadap pengalaman pengguna website Gemini pada Mahasiswa Mikroskil.

2.15.2 Tingkat Pengalaman Pengguna Website ChatGPT dengan Metode UEQ+ pada Mahasiswa Mikroskil

1. Pengaruh *Quality of Content* terhadap Pengalaman Pengguna di Situs Web *Chatbot* pada mahasiswa Mikroskil.

Skala *Quality of Content* (kualitas konten berfokus seberapa jauh informasi yang diberikan oleh sistem mempunyai kualitas, kelengkapan, dan relevansi yang sesuai dengan kebutuhan penggunanya. Skala *Quality of Content* mengukur sejauh mana pengguna puas dengan informasi yang diberikan oleh sistem, mencakup akurasi, kedalaman, dan keandalan data yang dihasilkan [48]. Pada penelitian sebelumnya yang mengevaluasi model *Large Language Model* (LLM) berbasis teks, dimensi kualitas konten (*Quality of Content*) berhasil memperoleh evaluasi yang sangat positif yang berdampak langsung pada pemenuhan ekspektasi pengguna [54]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Quality of Content* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Quality of Content* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

2. Pengaruh *Trustworthiness of Content* terhadap Pengalaman Pengguna di Situs Web *Chatbot* pada mahasiswa Mikroskil.

Skala *Trustworthiness of Content* adalah tingkat kepercayaan pengguna terhadap ketepatan, objektif, dan tidak adanya manipulasi fakta pada informasi yang dihasilkan oleh sistem. Skala *Trustworthiness of Content* mengukur seberapa besar pengguna mempercayai sumber dan kualitas informasi yang disediakan oleh sistem serta kepercayaan mereka bahwa sistem tidak akan menyampaikan informasi yang menyesatkan atau salah [48]. Pengukuran pengalaman pengguna pada aplikasi iPusnas menunjukkan bahwa dimensi tingkat kegunaan (*Usefulness*) menjadi salah satu aspek yang meraih evaluasi sangat positif, membuktikan

korelasi positifnya yang kuat terhadap kelancaran aktivitas fungsional pengguna [55]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Trustworthiness of Content* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Trustworthiness of Content* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

3. Pengaruh *Usefulness* terhadap Pengalaman Pengguna di Website ChatGPT pada mahasiswa Mikroskil.

Skala *Usefulness* menghasilkan keuntungan pragmatis yang langsung dirasakan oleh pengguna ketika sistem dapat memberikan jawaban atau mempermudah pencapaian tujuan dengan lebih efektif. Skala *Usefulness* adalah ukuran dari kesan pengguna tentang sejauh mana sistem membantu mereka mencapai tujuan lebih cepat, lebih mudah, atau lebih efektif. Jika inovasi teknologi dianggap mampu mempercepat pekerjaan dan menghemat tenaga kerja, maka pengguna secara otomatis akan memiliki dorongan kuat untuk menggunakannya sebagai bantuan utama secara teratur [48]. Pengukuran pengalaman pengguna pada aplikasi iPusnas menunjukkan bahwa dimensi tingkat kegunaan (*Usefulness*) menjadi salah satu aspek yang meraih evaluasi sangat positif, membuktikan korelasi positifnya yang kuat terhadap kelancaran aktivitas fungsional pengguna [52]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Usefulness* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Usefulness* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

4. Pengaruh *Clarity* terhadap Pengalaman Pengguna di Website ChatGPT pada mahasiswa Mikroskil.

Skala *Clarity* berkaitan dengan aspek visual dan estetika dari antarmuka sistem, yang mencakup keterbacaan teks, proporsi warna, dan pengaturan elemen layar yang tersusun. *Clarity* mengukur seberapa baik pengguna dapat membaca tata letak, elemen visual, dan navigasi antarmuka pengguna serta seberapa baik informasi disajikan tanpa kebingungan [48]. Sebuah studi evaluasi pada sistem Krapyak-U membuktikan bahwa kejelasan antarmuka (*Clarity*) mencatat evaluasi positif yang sangat tinggi dalam memfasilitasi kelancaran interaksi penggunaannya [51]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Clarity* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Clarity* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

5. Pengaruh *Perspiciuity* terhadap Pengalaman Pengguna di Website ChatGPT pada mahasiswa Mikroskil.

Skala *Perspiciuity* merupakan tingkat kejelasan navigasi dan tingkat beban kognitif yang ditawarkan kepada pengguna baru untuk memahami cara kerja sistem. Skala *Perspiciuity* mengukur seberapa mudah pengguna dapat mempelajari dan memahami suatu sistem [48]. Berdasarkan evaluasi yang dilakukan terhadap ChatGPT di kalangan mahasiswa, dimensi *Perspiciuity* (kemudahan belajar) secara empiris terbukti menjadi salah satu skala keunggulan utama yang paling signifikan dalam mendorong penerimaan sistem [7], [10]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Perspiciuity* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Perspiciuity* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

6. Pengaruh *Efficiency* terhadap Pengalaman Pengguna di Website ChatGPT pada mahasiswa Mikroskil.

Skala *Efficiency* berfokus pada seberapa baik sistem merespons dan mengeksekusi perintah pengguna dengan meminimalkan waktu tunggu dan interaksi yang tidak perlu. Efisiensi mengukur kecepatan dan akurasi dengan mana sistem dapat merespons input pengguna dan meminimalkan usaha yang diperlukan untuk menyelesaikan tugas [13]. Studi evaluasi komparatif pada aplikasi komunikasi serta platform AI ChatGPT menunjukkan bahwa tingkat efisiensi interaksi (*Efficiency*) merupakan salah satu skala pembeda yang secara signifikan membuat sebuah sistem dinilai lebih unggul dan dapat diandalkan terus-menerus [11], [15]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Efficiency* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Efficiency* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

7. Pengaruh *Trust* terhadap Pengalaman Pengguna di Website ChatGPT pada mahasiswa Mikroskil.

Skala *Trust* sebagian besar berkaitan dengan jaminan rasa aman pengguna terkait integritas platform dan komitmen dari penyedia, bukan pada konten teks yang dihasilkan. Skala *Trust* mengukur kesan yang dimiliki pengguna tentang keamanan dan privasi suatu sistem serta keyakinan bahwa data dan informasi pribadi mereka akan diperlakukan dengan tepat dan tidak akan disalahgunakan [48]. Evaluasi pada teknologi agen percakapan virtual (*conversational agent*) membuktikan bahwa kepercayaan pengguna terhadap sistem keamanan (*Trust*) memiliki peran yang sangat krusial, di mana dimensi ini memperoleh evaluasi sangat positif dalam membangun pengalaman interaksi [55]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Trust* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Trust* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

8. Pengaruh *Dependability* terhadap Pengalaman Pengguna di Website ChatGPT pada mahasiswa Mikroskil.

Skala *Dependability* berfokus pada tingkat keandalan, terkait dengan derajat stabilitas, konsistensi, dan prediktabilitas kinerja perangkat lunak saat berfungsi di bawah berbagai kondisi. Skala *Dependability* mengukur kesan pengguna terhadap keandalan dan stabilitas suatu sistem yang meminimalkan gangguan teknologi, kesalahan, atau kehilangan layanan yang tidak terduga [48]. Literatur dari studi evaluasi ChatGPT menetapkan bahwa keandalan operasional (*Dependability*) perangkat lunak meraih evaluasi positif yang kuat dan menjadi pilar fundamental yang mengunggulkan sistem tersebut di mata pengguna [7]. Hipotesis yang dapat ditarik sebagai berikut:

H_0 : *Dependability* Tidak terdapat pengaruh signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

H_a : *Dependability* signifikan terhadap pengalaman pengguna website ChatGPT pada Mahasiswa Mikroskil.

2.15.3 Perbandingan Tingkat Pengalaman Pengguna Website Gemini dan ChatGPT dengan Metode UEQ+ pada Mahasiswa Mikroskil

Mengikuti perumusan masalah yang telah dijabarkan, penelitian ini juga bertujuan untuk menguji perbandingan tingkat pengalaman pengguna antara website Gemini dan ChatGPT menggunakan metode *User Experience Questionnaire Plus* (UEQ+). Berdasarkan penelitian terdahulu, keseluruhan skala pada instrumen pengukuran dengan metode *User Experience Questionnaire Plus* (UEQ+) mampu mempresentasikan kualitas interaksi dan memenuhi parameter terhadap pengalaman pengguna. Pengujian komparatif ini diperlukan untuk mengetahui sejauh mana perbedaan atau disparitas profil pengalaman yang ditawarkan oleh kedua platform kecerdasan buatan tersebut. Dengan demikian, dapat ditarik hipotesis sebagai berikut:

H₀: tidak terdapat perbedaan yang signifikan antara pengalaman pengguna website Gemini dan ChatGPT berdasarkan metode *User Experience Questionnaire Plus* (UEQ+) pada mahasiswa Mikroskil.

H_a: terdapat perbedaan yang signifikan antara pengalaman pengguna website Gemini dan ChatGPT berdasarkan metode *User Experience Questionnaire Plus* (UEQ+) pada mahasiswa Mikroskil.

