

**MODEL FRAMEWORK EVALUASI USABILITY TERINTEGRASI MENGGUNAKAN
USE, RMT, IPA (STUDI KASUS SIAKAD BERBASIS WEB)**

SKRIPSI

Oleh :
MUHAMMAD PRIMA SAIFULLAH
NIM. 210605110159



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**MODEL FRAMEWORK EVALUASI USABILITY TERINTEGRASI MENGGUNAKAN
USE, RMT, IPA (STUDI KASUS SIAKAD BERBASIS WEB)**

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana (S.Kom)

Oleh:
MUHAMMAD PRIMA SAIFULLAH
NIM. 210605110159

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**MODEL FRAMEWORK EVALUASI USABILITY TERINTEGRASI
MENGUNAKAN USE, RMT, IPA (STUDI KASUS SIAKAD BERBASIS WEB)**

Oleh :

MUHAMMAD PRIMA SAIFULLAH

NIM. 210605110159

Telah Diperiksa dan Disetujui untuk Diuji:

Tanggal: 13 Maret 2026

Pembimbing I



A'la Syaifuli, M. Kom

NIP. 19771201 200801 1 007

Pembimbing II



Supriyono, M. Kom

NIP. 19841010 201903 1 012

Mengetahui,

Ketua Program Studi Teknik Informatika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M. Kom

NIP: 19841010 201903 1 012

HALAMAN PENGESAHAN

MODEL FRAMEWORK EVALUASI USABILITY TERINTEGRASI
MENGUNAKAN USE, RMT, IPA (STUDI KASUS SIAKAD BERBASIS WEB)

SKRIPSI

Oleh:
MUHAMMAD PRIMA SAIFULLAH
NIM. 210605110159

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar (S.Kom)
Tanggal : 13 April 2026

Susunan Dewan Penguji

Ketua Penguji : Roro Inda Melani, M.T, M.Sc
NIP. 19780925 200501 2 008

Anggota Penguji I : Nurizal Dwi Priandani, M.Kom
NIP. 19920830 202203 1 001

Anggota Penguji II : A'la Syauqi, M.Kom
NIP. 19771201 200801 1 007

Anggota Penguji III : Supriyono, M.Kom
NIP. 19841010 201903 1 012

(.....)
(.....)
(.....)
(.....)

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M.Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Muhammad Prima Saifullah
NIM : 210605110159
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : MODEL FRAMEWORK EVALUASI
USABILITY TERINTEGRASI
MENGUNAKAN USE, RMT, IPA (STUDI
KASUS SIAKAD BERBASIS WEB)

Menyatakan dengan sebenarnya bahwa Skripsi yang Saya tulis ini benar-benar merupakan hasil karya Saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang Saya akui sebagai hasil tulisan atau pikiran Saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka Saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 13 April 2026
Yang membuat pernyataan,



Muhammad Prima Saifullah
NIM. 210605110159

MOTTO

”Measurement is the only way to know if we are making progress.”

-Benjamin D. Wright-

KATA PENGANTAR

Assalamu'alaikum wr. wb.

Alhamdulillahillāhirabbil‘ālamīn, puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “MODEL FRAMEWORK EVALUASI USABILITY TERINTEGRASI MENGGUNAKAN USE, RMT, IPA (STUDI KASUS SIAKAD BERBASIS WEB)” dengan baik. Shalawat serta salam semoga senantiasa tercurah kepada junjungan Nabi Besar Muhammad SAW., yang telah membimbing umat manusia dari masa kegelapan menuju era yang penuh dengan ilmu pengetahuan dan kebenaran.

Melalui kesempatan yang berharga ini, penulis dengan penuh rasa syukur ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan, bantuan, serta kontribusi, baik secara langsung maupun tidak langsung, sehingga penyusunan karya ilmiah ini dapat terselesaikan dengan baik, khususnya kepada:

- 1) Prof. Dr. Hj. Ilfi Nur Diana, M.Si selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang, atas kepemimpinan dan kebijakan yang mendukung terselenggaranya proses pendidikan dengan baik.
- 2) Dr. H. Agus Mulyono, S.Pd., M.Kes selaku Dekan Fakultas Sains dan Teknologi, atas dukungan serta fasilitas yang diberikan selama proses perkuliahan hingga penyusunan skripsi ini.

- 3) Supriyono, M.Kom selaku Ketua Program Studi Teknik Informatika sekaligus Pembimbing II, atas arahan, bimbingan, serta masukan yang sangat berarti dalam penyelesaian skripsi ini.
- 4) A'la Syauqi, M.Kom selaku Pembimbing I, atas bimbingan, kesabaran, serta saran konstruktif yang diberikan selama proses penelitian dan penulisan skripsi.
- 5) Roro Inda Melani, M.T., M.Sc selaku Ketua Penguji, atas kritik dan saran yang membangun dalam penyempurnaan skripsi ini.
- 6) Nurizal Dwi Priandani, M.Kom selaku Anggota Penguji, atas masukan dan evaluasi yang berharga demi peningkatan kualitas penelitian ini.
- 7) Seluruh dosen, admin, dan laboran serta civitas akademika Program Studi Teknik Informatika, atas ilmu, bantuan, dan pelayanan yang telah diberikan selama masa studi.
- 8) Kedua orang tua penulis, atas doa, dukungan moral maupun material, serta kasih sayang yang tiada henti kepada penulis.
- 9) Sahabat dan keluarga besar Teknik Informatika, khususnya angkatan 2021, atas kebersamaan, dukungan, dan semangat yang telah diberikan.
- 10) Serta seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan bantuan baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa penyusunan skripsi ini masih memiliki keterbatasan dan belum mencapai kesempurnaan. Oleh karena itu, penulis terbuka terhadap saran, kritik, dan masukan yang konstruktif sebagai upaya perbaikan dan

pengembangan di masa mendatang. Besar harapan penulis agar skripsi ini dapat memberikan kontribusi dan manfaat bagi perkembangan ilmu pengetahuan.

Wassalamu'alaikum wr. wb.

Malang, 13 April 2026

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
KATA PENGANTAR	vii
ABSTRAK	xiv
ABSTRACT	xv
مستخلص البحث	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Batasan Masalah	5
BAB II STUDI PUSTAKA	7
2.1 Penelitian Terdahulu	7
2.2 <i>USE Questionnaire</i>	16
2.3 <i>Rasch Measurement Theory (RMT)</i>	19
2.4 <i>Rating Scale Model</i>	22
2.5 <i>Maximum Marginal Likelihood Estimation (MMLE)</i>	25
2.6 <i>Gauss–Hermite Quadrature</i>	28
2.7 <i>Posterior Probability</i>	31
2.8 <i>Fit Statistic</i>	34
2.9 <i>Importance Performance Analysis</i>	37
BAB III DESAIN DAN IMPLEMENTASI	41
3.1 Desain Penelitian	41
3.1.1 Persiapan Kuesioner	42
3.1.2 Data Kuesioner	43
3.1.3 Analisa Kuesioner	47
3.1.4 Analisa Data	50
3.1.5 Hasil Evaluasi dan Rekomendasi	52
3.2 Implementasi Sistem	54
3.2.1 <i>Input</i>	55
3.2.2 <i>Processing Data</i>	58
3.2.3 Penentuan Parameter	60
3.2.4 <i>Gauss-Hermite Quadrature</i>	62
3.2.5 Estimasi MMLE	64
3.2.6 Hitung Probabilitas	66
3.2.7 <i>Posterior Distribution</i>	68
3.2.8 <i>Expected Score Varians</i>	70
3.2.9 <i>Fit Statistic</i>	72
3.2.10 <i>Reliability Estimation</i>	74
3.2.11 <i>Output</i>	76

BAB IV HASIL DAN PEMBAHASAN	79
4.1 Hasil Responden	79
4.2 Hasil Pengukuran Menggunakan <i>Rasch Measurement Theory</i>	80
4.2.1 <i>Item Fit</i>	81
4.2.2 <i>Person Fit</i>	83
4.3 Hasil <i>Importance Performance Analysis</i>	86
4.4 Integrasi Islam	89
BAB V KESIMPULAN DAN SARAN	95
5.1 Kesimpulan	95
5.2 Saran.....	96
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 2.1 Diagram <i>Importance Performance Analysis</i>	40
Gambar 3.1 Bagan Penelitian.....	41
Gambar 3.2 Tampilan Kuesioner	43
Gambar 3.3 Bagan Sistem.....	55
Gambar 4.1 Diagram Sebaran Jurusan.....	80
Gambar 4.2 <i>Item Fit Scatter Plot</i>	81
Gambar 4.3 <i>Person Fit Scatter Plot</i>	84
Gambar 4.4 Diagram IPA	87

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	10
Tabel 2.2 USE <i>Questionnaire</i>	17
Tabel 3.1 <i>Item</i> Kuesioner	45

ABSTRAK

Saifullah, Muhammad Prima. 2026. **MODEL FRAMEWORK EVALUASI USABILITY TERINTEGRASI MENGGUNAKAN USE, RMT, IPA (STUDI KASUS SIAKAD BERBASIS WEB)**. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) A'la Syauqi, M.Kom (II) Supriyono, M.Kom.

Kata kunci: *USE Questionnaire, Rasch Measurement Theory (RMT), Rating Scale Model (RSM), Importance–Performance Analysis (IPA), Sistem Informasi Akademik Berbasis Web, Evaluasi Pengalaman Pengguna*

Penelitian ini bertujuan untuk mengevaluasi tingkat *usability* Sistem Informasi Akademik (SIAKAD) berbasis web di Universitas Islam Negeri Maulana Malik Ibrahim Malang melalui pendekatan terintegrasi yang menggabungkan *USE Questionnaire, Rasch Measurement Theory (RMT)*, dan *Importance Performance Analysis (IPA)*. Pendekatan *USE* digunakan untuk mengukur empat dimensi utama, yaitu utilitas, kemudahan penggunaan, kepuasan pengguna, dan kemudahan pembelajaran, guna memperoleh gambaran komprehensif mengenai pengalaman pengguna. Selanjutnya, validitas dan reliabilitas instrumen diuji menggunakan *RMT* dengan model *Rating Scale Model (RSM)*, di mana *Maximum Marginal Likelihood Estimation (MMLE)* berperan dalam mengestimasi parameter *item* dan kemampuan responden, *Gauss-Hermite Quadrature* digunakan untuk mendekati integral dalam perhitungan *likelihood*, *Posterior Probability* dimanfaatkan untuk mengestimasi distribusi kemampuan laten responden, serta *Fit Statistics* digunakan untuk mengidentifikasi kesesuaian *item* dan responden terhadap model. Hasil pengukuran kemudian dianalisis menggunakan *IPA* untuk memetakan hubungan antara tingkat kepentingan dan kinerja atribut sistem. Berdasarkan hasil analisis, tingkat *usability* SIAKAD berada pada kategori cukup baik, namun masih memerlukan optimalisasi pada beberapa atribut prioritas, dengan distribusi 12 *item* pada Kuadran I, 2 *item* pada Kuadran II, 13 *item* pada Kuadran III, dan 3 *item* pada Kuadran IV, yang menunjukkan adanya kesenjangan antara kepentingan dan kinerja, khususnya pada aspek utilitas dan kemudahan penggunaan. Hasil validasi *RMT* menunjukkan bahwa seluruh item kuesioner memiliki tingkat kesesuaian yang baik tanpa adanya *item misfit*, meskipun terdapat 5 responden yang dikeluarkan karena tidak memenuhi kriteria *fit* masing-masing sebesar (1,76;1,75), (2,13;2,08), (1,60;1,60), (2,17;2,16), dan (1,91;1,90). Secara keseluruhan, integrasi *USE, RMT, dan IPA* terbukti mampu menghasilkan evaluasi *usability* yang valid, reliabel, dan berbasis data, serta memberikan dasar yang kuat dalam penyusunan rekomendasi perbaikan sistem.

ABSTRACT

Saifullah, Muhammad Prima. 2026. **INTEGRATED USABILITY EVALUATION FRAMEWORK MODEL USING USE, RMT, IPA (CASE STUDY OF A WEB-BASED ACADEMIC INFORMATION SYSTEM)**. Undergraduate Thesis. Informatics Engineering Study Program, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University of Malang. Supervisors: (I) A'la Syauqi, M.Kom (II) Supriyono, M.Kom.

This study aims to evaluate the usability level of the web-based Academic Information System (SIKAD) at Maulana Malik Ibrahim State Islamic University of Malang through an integrated approach combining the USE Questionnaire, Rasch Measurement Theory (RMT), and Importance Performance Analysis (IPA). The USE approach is employed to measure four main dimensions, namely usefulness, ease of use, user satisfaction, and ease of learning, in order to obtain a comprehensive understanding of user experience. Furthermore, the validity and reliability of the instrument are tested using RMT with the Rating Scale Model (RSM), in which Maximum Marginal Likelihood Estimation (MMLE) is used to estimate item parameters and respondent abilities, Gauss-Hermite Quadrature is applied to approximate integrals in likelihood calculations, Posterior Probability is utilized to estimate the distribution of respondents' latent abilities, and Fit Statistics are used to identify the conformity of items and respondents to the model. The measurement results are then analyzed using IPA to map the relationship between the importance and performance levels of system attributes. Based on the analysis, the usability level of SIKAD is categorized as fairly good; however, optimization is still required for several priority attributes, with a distribution of 12 items in Quadrant I, 2 items in Quadrant II, 13 items in Quadrant III, and 3 items in Quadrant IV, indicating a gap between importance and performance, particularly in the aspects of usefulness and ease of use. The RMT validation results show that all questionnaire items demonstrate good fit with no misfitting items, although 5 respondents were excluded for not meeting the fit criteria, with respective values of (1,76;1,75), (2,13;2,08), (1,60;1,60), (2,17;2,16), and (1,91;1,90). Overall, the integration of USE, RMT, and IPA is proven to produce a valid, reliable, and data-driven usability evaluation, while also providing a strong foundation for formulating system improvement recommendations.

Keywords: *USE Questionnaire, Rasch Measurement Theory (RMT), Rating Scale Model (RSM), Importance-Performance Analysis (IPA), Web-Based Academic Information Systems, User Experience Assessment*

مستخلص البحث

سيف الله، محمد بريما. 2026. نموذج إطار عمل متكامل لتقييم قابلية الاستخدام باستخدام USE و RMT و IPA (دراسة حالة على نظام معلومات أكاديمي قائم على الويب). رسالة جامعية (بكالوريوس). برنامج دراسة هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانغ. المشرفان: (الأول) علاء شوقي، ماجستير (الثاني) سوبريونو، ماجستير.

الكلمات المفتاحية: استبيان تجربة المستخدم، نظرية راش للقياس (RMT)، نموذج مقياس التقييم (RSM)، تحليل الأهمية والأداء (IPA)، أنظمة المعلومات الأكاديمية القائمة على الويب، تقييم تجربة المستخدم

تهدف هذه الدراسة إلى تقييم مستوى قابلية الاستخدام لنظام المعلومات الأكاديمي القائم على الويب (SIAKAD) في جامعة مولانا مالك إبراهيم الإسلامية الحكومية في مالانغ من خلال فحص متكامل يجمع بين استبيان USE ونظرية القياس لراش (RMT) وتحليل الأهمية والأداء (IPA). تم استخدام منهج USE لقياس أربعة أبعاد رئيسية، وهي الفائدة، وسهولة الاستخدام، ورضا المستخدم، وسهولة التعلم، بهدف الحصول على فهم شامل لتجربة المستخدم. علاوة على ذلك، تم اختبار صلاحية وموثوقية الأداة باستخدام نظرية راش من خلال نموذج مقياس التقييم (RSM)، حيث تم استخدام تقدير الاحتمالية العظمى الهامشية (MML) لتقدير معالم البنود وقدرات المستجيبين، كما تم تطبيق تكامل غاوس-هيرميت لتقريب التكاملات في حسابات الاحتمالية، واستخدام الاحتمال اللاحق لتقدير توزيع القدرات الكامنة للمستجيبين، واستخدام إحصاءات المطابقة لتحديد مدى توافق البنود والمستجيبين مع النموذج. بعد ذلك، تم تحليل نتائج القياس باستخدام تحليل الأهمية والأداء لرسم العلاقة بين مستوى الأهمية ومستوى الأداء لخصائص النظام. وبناءً على نتائج التحليل، تم تصنيف مستوى قابلية الاستخدام لنظام SIAKAD على أنه جيد إلى حد ما؛ إلا أنه لا يزال يتطلب تحسينًا في بعض الخصائص ذات الأولوية، حيث توزعت البنود إلى 12 بنودًا في الربع الأول، وبندين في الربع الثاني، و13 بنودًا في الربع الثالث، و3 بنود في الربع الرابع، مما يشير إلى وجود فجوة بين الأهمية والأداء، خاصة في جانبي الفائدة وسهولة الاستخدام. وأظهرت نتائج التحقق باستخدام نظرية راش أن جميع بنود الاستبيان تتمتع بدرجة توافق جيدة دون وجود بنود غير ملائمة، على الرغم من استبعاد 5 مستجيبين لعدم استيفائهم معايير المطابقة، بالقيم التالية: (1,75؛1,76)، (2,08؛2,13)، (1,60؛1,60)، (2,16؛2,17)، و(1,90؛1,91). وبشكل عام، أثبت دمج منهجيات USE و RMT و IPA قدرته على إنتاج تقييم لقابلية الاستخدام يتمتع بالصدق والموثوقية ويستند إلى البيانات، مع توفير أساس قوي لصياغة توصيات تحسين النظام.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Usability merupakan aspek krusial dalam pengembangan sistem informasi berbasis web, karena berhubungan langsung dengan pengalaman pengguna (*user experience*). Berbagai pendekatan telah dikembangkan untuk mengukur *usability*, salah satunya adalah *USE Questionnaire*, yang digunakan untuk menilai kepuasan pengguna terhadap suatu sistem berdasarkan empat dimensi utama: *usefulness*, *ease of use*, *satisfaction*, dan *ease of learning* (Hariyanto, Triyono, & Köhler, 2020). Pendekatan ini telah terbukti efektif dalam mengevaluasi berbagai sistem berbasis web, termasuk sistem *e-learning* (Hendra & Arifin, 2018) dan aplikasi kesehatan *mobile* (Melin et al., 2020). Dalam pengembangan sistem informasi yang bermanfaat, konsep kebermanfaatan bagi manusia juga sejalan dengan prinsip yang diajarkan dalam hadis Nabi Muhammad SAW:

خَيْرُ النَّاسِ أَنْفَعُهُمْ لِلنَّاسِ

"Sebaik-baik manusia adalah yang paling bermanfaat bagi orang lain." (HR. Ahmad)

Hadist ini menekankan pentingnya penciptaan sistem yang memberikan manfaat maksimal bagi penggunanya.

Dalam konteks ini, konsep-konsep yang terkait dengan pengetahuan, keteraturan, dan pengembangan sistem yang bermanfaat tidak hanya dapat ditemukan dalam kajian ilmiah modern, tetapi juga tercermin dalam ajaran agama

yang mengedepankan pentingnya ilmu pengetahuan sebagai sarana untuk mencapai kemajuan dan kebaikan umat manusia. Al-Qur'an, sebagai sumber utama ajaran Islam, berulang kali mengajak umatnya untuk memanfaatkan akal dan pengetahuan dalam setiap aspek kehidupan. Allah SWT berfirman dalam Surah Al-'Alaq (96:1-5):

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ١ خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ ٢ اقْرَأْ وَرَبُّكَ الْأَكْرَمُ ٣ الَّذِي عَلَّمَ بِالْقَلَمِ ٤ عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمْ ٥

“Bacalah dengan (menyebut) nama Tuhanmu yang menciptakan. Dia telah menciptakan manusia dari segumpal darah. Bacalah, dan Tuhanmulah yang Maha Pemurah, yang mengajar (manusia) dengan pena. Dia mengajarkan kepada manusia apa yang tidak diketahuinya.” (QS. Al-Alaq:1-5)

Ayat ini menegaskan bahwa pencarian ilmu pengetahuan merupakan kewajiban, dan dengan ilmu, umat manusia dapat mengembangkan teknologi dan sistem yang bermanfaat, termasuk dalam konteks pengembangan sistem informasi yang dapat mempermudah kehidupan.

Namun, untuk memberikan gambaran yang lebih komprehensif mengenai *usability*, evaluasi penting juga dilakukan melalui metode *Importance-Performance Analysis* (IPA). IPA menggabungkan analisis pentingnya atribut-atribut tertentu dalam sistem dengan kinerja aktualnya, sehingga memberikan wawasan yang lebih mendalam mengenai area yang perlu diperbaiki (Andry, Christianto, & Wilujeng, 2019). Dalam konteks sistem informasi akademik, metode ini dapat membantu mengidentifikasi elemen-elemen sistem yang dianggap penting oleh pengguna dan seberapa baik elemen-elemen tersebut

berfungsi. Konsep ini dapat dihubungkan dengan ajaran dalam Al-Qur'an, seperti dalam Surah At-Tawbah (9:71):

وَالْمُؤْمِنُونَ وَالْمُؤْمِنَاتُ بَعْضُهُمْ أَوْلِيَاءُ بَعْضٍ يَأْمُرُونَ بِالْمَعْرُوفِ وَيَنْهَوْنَ عَنِ الْمُنْكَرِ

“Dan orang-orang yang beriman, laki-laki dan perempuan, sebagian mereka (adalah) penolong bagi sebagian yang lain. Mereka menyuruh (mengerjakan) yang ma'ruf dan mencegah dari yang munkar” (QS. At-Taubah: 71)

Ayat ini mengajarkan tentang pentingnya kerja sama dalam memperbaiki sistem yang ada untuk kepentingan bersama.

Selain itu, untuk memastikan validitas pengukuran *usability*, *Rasch Measurement Theory* telah diadopsi sebagai pendekatan statistik yang dapat mengukur kualitas dan reliabilitas instrumen evaluasi (Melin et al., 2020). Teori Rasch memberikan pemahaman yang lebih mendalam mengenai *item-item* yang digunakan dalam instrumen pengukuran dan memungkinkan pengukuran yang lebih tepat dan adil terhadap tingkat kepuasan pengguna. Prinsip ini selaras dengan ajaran dalam Surah Al-Baqarah (2:286):

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا

"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya" (QS. Al-Baqarah: 286)

Ayat ini mengajarkan bahwa setiap sistem harus dirancang dengan mengedepankan keadilan dan kecocokan antara tuntutan dan kapasitas pengguna.

Penelitian ini bertujuan untuk mengevaluasi *usability* dari SIAKAD berbasis web milik Universitas Islam Negeri Maulana Malik Ibrahim Malang menggunakan pendekatan terintegrasi yang melibatkan *USE Questionnaire*, IPA,

dan RMT. Dengan menggunakan ketiga pendekatan ini secara bersamaan, diharapkan dapat diperoleh gambaran yang lebih menyeluruh mengenai kualitas sistem dan dapat memberikan rekomendasi perbaikan yang lebih spesifik bagi pengembangan sistem informasi akademik di masa depan. Sejalan dengan prinsip ilmiah dan ajaran agama yang mengutamakan manfaat, keadilan, dan pengembangan ilmu, penelitian ini diharapkan dapat memberikan kontribusi positif dalam menciptakan sistem yang tidak hanya efisien, tetapi juga bermanfaat bagi pengguna dan masyarakat secara luas.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

- 1) Seberapa tingkat *usability* SIAKAD berbasis web Universitas Islam Negeri Maulana Malik Ibrahim Malang berdasarkan aspek utilitas, kemudahan penggunaan, kepuasan, dan kemudahan pembelajaran menggunakan USE *Questionnaire* dengan analisis *Importance Performance Analysis (IPA)*?
- 2) Seberapa valid pengukuran hubungan antara tingkat kepentingan dan kinerja atribut-atribut sistem informasi akademik berbasis web milik Universitas Islam Negeri Maulana Malik Ibrahim Malang menggunakan *Importance Performance Analysis* dan *Rasch Measurement Theory*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dikemukakan, tujuan penelitian ini adalah sebagai berikut:

- 1) Mengevaluasi nilai *usability* dari SIAKAD berbasis web milik Universitas Islam Negeri Maulana Malik Ibrahim Malang dengan menggunakan pendekatan *USE Questionnaire*, yang mencakup dimensi utilitas, kemudahan penggunaan, kepuasan pengguna, dan kemudahan pembelajaran, untuk memperoleh pemahaman yang komprehensif mengenai pengalaman pengguna terhadap sistem tersebut.
- 2) Menganalisis hubungan antara tingkat kepentingan dan kinerja atribut-atribut SIAKAD berbasis web dengan menggunakan metode *Importance Performance Analysis*, serta memastikan validitas dan reliabilitas pengukuran melalui penerapan *Rasch Measurement Theory*, guna memberikan rekomendasi perbaikan yang berbasis data yang valid dan objektif.

1.4 Batasan Masalah

Agar penelitian ini memiliki ruang lingkup yang jelas dan terarah, beberapa batasan masalah ditetapkan sebagai berikut:

- 1) Sistem yang menjadi objek penelitian adalah SIAKAD milik Universitas Islam Negeri Maulana Malik Ibrahim Malang. Fokus pengujian *usability* hanya dilakukan pada versi SIAKAD yang digunakan oleh mahasiswa dalam aktivitas akademik sehari-hari.
- 2) Responden yang digunakan dalam penelitian ini dibatasi pada mahasiswa Fakultas Sains dan Teknologi. Data responden berjumlah 133 orang mahasiswa yang didapat melalui Google Form. Google Form disebarluaskan

selama kurang lebih 1 bulan, dibuka pada tanggal 07 November 2025 dan ditutup pada tanggal 07 Desember 2025.

- 3) Aspek yang dievaluasi dalam penelitian ini mencakup fitur atau layanan SIAKAD yang berhubungan langsung dengan aktivitas akademik mahasiswa. Beberapa di antaranya meliputi pengisian Kartu Rencana Studi (KRS), akses Kartu Hasil Studi (KHS), pengelolaan jadwal kuliah, serta layanan akademik lain yang digunakan secara rutin oleh mahasiswa.

Dengan batasan-batasan tersebut, penelitian ini diharapkan dapat memberikan hasil evaluasi usability yang lebih fokus, terarah, dan relevan terhadap kebutuhan utama mahasiswa sebagai pengguna SIAKAD.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terdahulu

Sebelum penelitian dimulai, kajian terhadap penelitian terdahulu dilakukan untuk memahami perkembangan metodologi evaluasi *usability* pada berbagai jenis sistem informasi, khususnya yang menggunakan pendekatan kuesioner, analisis performa, dan model pengukuran psikometrik seperti *Rasch Measurement Theory* (RMT). Tinjauan ini kemudian menjadi dasar konseptual dalam merancang model evaluasi *usability* terintegrasi pada penelitian ini.

Hendra dan Arifin (2018) melakukan pengukuran *usability* pada sistem informasi penilaian mahasiswa berbasis web. Evaluasi dilakukan melalui instrumen kuesioner yang mengukur aspek kemudahan penggunaan, efisiensi, dan kepuasan pengguna. Studi ini menunjukkan bahwa pendekatan kuantitatif berbasis persepsi pengguna efektif dalam mengidentifikasi masalah antarmuka dan alur penggunaan. Namun, penelitian tersebut belum mengintegrasikan analisis performa maupun model validasi data pengukuran berbasis Rasch sehingga ruang untuk penguatan metodologi masih terbuka.

Harrati et al. (2016) mengeksplorasi tingkat kepuasan pengguna pada sistem *e-learning* dengan memadukan metrik penggunaan dan analisis *System Usability Scale* (SUS). Hasil penelitian menegaskan bahwa kombinasi data objektif (*usage metrics*) dan subjektif (SUS) memberikan gambaran lebih komprehensif mengenai *experience* pengguna. Meskipun demikian, model evaluasi yang digunakan tidak

menilai kesesuaian tingkat kepentingan terhadap performa sistem, sehingga belum mengakomodasi *gap analysis* seperti yang disediakan oleh *Importance Performance Analysis* (IPA).

Hariyanto et al. (2020) menilai *usability* sistem *e-learning* adaptif menggunakan *USE Questionnaire*. Penelitian ini menghasilkan bukti kuat bahwa empat dimensi utama dalam *USE Questionnaire*, yakni *usefulness*, *ease of use*, *ease of learning*, dan *satisfaction*, mampu memberikan representasi yang baik terhadap persepsi pengguna. Meski begitu, penelitian tersebut belum menerapkan teknik validasi psikometrik tingkat lanjut seperti RMT sehingga reliabilitas butir instrumen belum dianalisis secara mendalam.

AlSalem dan AlShamari (2023) mengembangkan pendekatan evaluasi *usability* berbasis perilaku untuk menilai sistem web interaktif. Studi ini menekankan pentingnya data empiris seperti pola klik, *error rate*, dan waktu penyelesaian tugas sebagai indikator kualitas *usability*. Penelitian ini memberi landasan mengenai pentingnya triangulasi data dalam evaluasi pengalaman pengguna, meskipun tidak menggunakan instrumen kuesioner maupun model validasi seperti Rasch.

Melin et al. (2020) menyusun sebuah instrumen kuesioner untuk menilai kepuasan pengguna aplikasi *mobile health* dengan pendekatan RMT. Penelitian ini menunjukkan bahwa Rasch mampu mengidentifikasi tingkat kesulitan item, mengukur kemampuan responden secara akurat, serta memastikan unidimensionalitas instrumen. Studi ini relevan karena menunjukkan bagaimana

RMT dapat digunakan sebagai alat validasi untuk meningkatkan kualitas instrumen *usability* yang digunakan pada penelitian ini.

Gomes et al. (2024) melakukan *usability* testing pada sebuah *web tool* untuk model-model diseminasi dan implementasi sains. Penelitian ini mengintegrasikan evaluasi heuristik, pengukuran kepuasan pengguna, serta uji coba berbasis skenario. Hasilnya memperlihatkan pentingnya pendekatan evaluasi *usability* yang komprehensif dan berlapis. Studi ini mendukung perlunya model integratif yang menggabungkan beberapa metode evaluasi seperti yang dilakukan pada penelitian ini.

Sukmasetya et al. (2020) mengevaluasi tingkat *usability* website universitas melalui pendekatan kuantitatif berbasis kuesioner. Hasil penelitian menunjukkan bahwa *usability* website dapat diukur secara efektif melalui indikator persepsi pengguna, meskipun tidak memanfaatkan pendekatan analisis performa dan validasi instrumen tingkat lanjut.

Medina et al. (2019) melakukan studi *usability* pada platform rehabilitasi motorik berbasis web dengan memadukan *performance testing* dan *feedback* pengguna. Studi ini menyoroti pentingnya desain antarmuka yang sederhana dan aksesibel, serta memberikan bukti bahwa evaluasi *usability* dapat meningkatkan kualitas sistem yang menyangkut kebutuhan kesehatan. Meskipun pendekatan yang digunakan komprehensif, penelitian ini tidak memanfaatkan metode analisis persepsi seperti IPA maupun validasi Rasch.

Thowaf (2021) menguji *usability m-commerce* Tokopedia menggunakan *USE Questionnaire* dan *performance test*. Penelitian ini menarik karena

menunjukkan bagaimana kombinasi instrumen persepsi dan pengukuran performa dapat mengungkap aspek-aspek *usability* yang tidak terlihat melalui kuesioner saja. Pendekatan ini relevan sebagai dasar integrasi metode dalam penelitian ini.

Pal dan Vanijja (2020) menilai *perceived usability* Microsoft Teams menggunakan SUS dan *Technology Acceptance Model* (TAM). Hasil penelitian menunjukkan faktor *usability* memiliki pengaruh signifikan terhadap penerimaan teknologi. Penelitian ini memperkuat posisi *usability* sebagai aspek kritis dalam keberhasilan sistem digital, tetapi belum menggunakan metode validasi data seperti Rasch maupun analisis gap performa seperti IPA.

Secara keseluruhan, penelitian terdahulu menunjukkan bahwa metode evaluasi *usability* berbasis kuesioner seperti USE dan SUS banyak digunakan, tetapi umumnya berdiri sendiri tanpa integrasi dengan teknik analisis performa atau validasi psikometrik lanjutan. Hanya sebagian penelitian yang memanfaatkan Rasch untuk validasi instrumen, dan sedikit yang melibatkan IPA dalam mengidentifikasi prioritas perbaikan sistem. Oleh karena itu, penelitian ini diharapkan dapat menawarkan kontribusi baru dengan mengintegrasikan USE *Questionnaire*, IPA, dan RMT ke dalam satu model *framework* evaluasi *usability* yang komprehensif untuk mengukur *usability* SIAKAD berbasis web.

Tabel 2.1 Penelitian Terdahulu

No.	Judul	Penulis	Metode	Hasil
1	<i>Web-based Usability Measurement for Student Grading</i>	Hendra, S.Kom., M.T.; Yulyani Arifin,	Peneliti menggunakan Kuesioner USE sebagai instrumen utama untuk mengukur kegunaan sistem. Data	Sistem Informasi Penilaian Mahasiswa memperoleh tingkat kelayakan kegunaan

No.	Judul	Penulis	Metode	Hasil
	<i>Information System</i>	S.Kom., M.M.	dikumpulkan dari pengguna sistem penilaian mahasiswa, kemudian dianalisis untuk menilai tiga aspek utama: <i>usefulness, ease of use,</i> dan <i>ease of learning.</i>	sebesar 75,23%, dikategorikan layak digunakan berdasarkan hasil pengukuran dengan <i>USE Questionnaire.</i>
2	<i>Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis</i>	Nouzha Harrati, Imed Bouchrika, Abdelkam el Tari, Ammar Ladjailia	Penelitian menerapkan kombinasi metrik penggunaan (<i>usage-based metrics</i>) yang diperoleh dari log sistem serta <i>System Usability Scale (SUS)</i> untuk mengevaluasi kepuasan pengguna. Analisis dilakukan lintas kelompok usia untuk melihat variasi persepsi <i>usability.</i>	Kelompok usia 25–29 tahun memiliki tingkat penyelesaian tugas dan kepuasan tertinggi, dengan skor SUS 74,67%. Kelompok usia 50–65 tahun juga memiliki skor SUS tinggi (69,57%) meskipun efisiensi lebih rendah. Secara umum, platform Moodle dinilai memuaskan oleh kedua kelompok.
3	<i>Usability Evaluation of a Personalized Adaptive E-</i>	Didik Hariyanto, Moch. Bruri Triyono,	Kuesioner USE diberikan kepada siswa dengan empat indikator utama. Data kuantitatif dari skala Likert	Sebanyak 66,13% siswa menyatakan puas. Variabel kepuasan memperoleh skor

No.	Judul	Penulis	Metode	Hasil
	<i>Learning System Using the USE Questionnaire</i>	Thomas Köhler	digabungkan dengan validasi silang menggunakan umpan balik terbuka (<i>open-ended feedback</i>) untuk memperkuat hasil evaluasi <i>usability</i> .	72,89, sesuai dengan rentang penerimaan. Hasil skala Likert dan umpan balik terbuka konsisten, sehingga aplikasi adaptif dinilai memuaskan.
4	<i>Assessing Interactive Web-Based Systems Using Behavioral Measurement Techniques</i>	Thanaa Saad AlSalem, Majed Aadi AlShamari	Penelitian menggabungkan tiga pendekatan, yaitu evaluasi heuristik oleh ahli, pengujian kegunaan (<i>usability testing</i>), dan pelacakan mata (<i>eye tracking</i>). Ketiga teknik digunakan untuk mengidentifikasi dan mengklasifikasikan masalah kegunaan dalam website <i>e-commerce</i> .	Ditemukan total masalah kegunaan: <ul style="list-style-type: none"> • 66,66% teridentifikasi oleh evaluasi heuristik • 13,9% oleh uji kegunaan • 19,44% oleh pelacakan mata Dari semua masalah, 38,89% merupakan masalah besar, sementara sisanya terdiri dari masalah kecil, sangat serius, dan kosmetik. Evaluasi heuristik menjadi metode paling efektif.
5	<i>A Questionnaire</i>	Jeanette Melin,	Peneliti mengembangkan	Skala mHealth dinilai optimal

No.	Judul	Penulis	Metode	Hasil
	<i>e for Assessing User Satisfaction with Mobile Health Apps: Development Using Rasch Measurement Theory</i>	PhD; Stephanie Erika Bonn, PhD; Leslie Pendrill, PhD; Ylva Trolle Lagerros, MPH, MD, PhD	kuesioner mHealth berbasis Likert 5 poin dan memvalidasi properti pengukurannya menggunakan <i>Rasch Measurement Theory</i> . Analisis mencakup pemeriksaan <i>fit item</i> , pemisahan kategori respon, deteksi DIF (<i>Differential Item Functioning</i>), dan evaluasi unidimensionalitas skala.	setelah kategori respons disederhanakan. Dua <i>item</i> dibuang karena <i>misfit</i> . <i>Person Separation Index</i> 0,79, cukup baik untuk perbandingan kelompok (bukan individu). Skala tidak sepenuhnya unidimensional dan menunjukkan tiga dimensi pengalaman pengguna.
6	<i>Usability Testing of a Web Tool for Dissemination and Implementation Science Models</i>	Rebekah Natalie Gomes, Bryan S. Ford, Rachel G. Tabak, Ross C. Brownson, Sara Malone, Maggie Padek, Russell E.	Studi menggunakan protokol pengujian kegunaan bertahap yang mengumpulkan data kuantitatif dan kualitatif. Peserta menggunakan berbagai model implementasi dalam <i>web tool</i> , dan pengalaman mereka dianalisis menggunakan kombinasi observasi, wawancara, serta log penggunaan.	Peserta menggunakan berbagai <i>framework</i> implementasi seperti RE-AIM, EPIS, PRISM, dan CFIR. Hasil uji menunjukkan variasi penggunaan dan preferensi model, serta memberikan wawasan terhadap pengalaman kegunaan masing-

No.	Judul	Penulis	Metode	Hasil
		Glasgow, Borsika Rabin		masing kerangka kerja dalam platform tersebut.
7	<i>Usability Evaluation of a University Website: A Case Study</i>	P. Sukmasetya, A. Setiawan, dan E. R. Arumi	Metode yang digunakan adalah observasi tradisional dengan melibatkan dosen dan mahasiswa. Peserta diminta mengevaluasi lima aspek <i>usability</i> (<i>learnability</i> , <i>memorability</i> , <i>efficiency</i> , <i>errors</i> , <i>satisfaction</i>) menggunakan instrumen penilaian berbasis skala.	Skor <i>usability</i> rata-rata keseluruhan adalah 2,77. Per indikator: • <i>Learnability</i> : 2,83 • <i>Efficiency</i> : 2,73 • <i>Memorability</i> : 2,82 • <i>Errors</i> : 2,65 • <i>Satisfaction</i> : 2,79 Kesimpulannya, kegunaan <i>website</i> tergolong cukup, namun masih membutuhkan banyak perbaikan.
8	<i>Usability Study of a Web-Based Platform for Home Motor Rehabilitation</i>	Jorge Luis Pérez Medina, Mario González, Henry Mauricio Pilco, Karina Beatriz Jimenes	Penelitian memanfaatkan IBM <i>Computer System Usability Questionnaire</i> (CSUQ) untuk memperoleh persepsi subjektif pengguna mengenai kegunaan platform. Pendekatan utama adalah survei kuantitatif yang	Platform ePHoRt dinilai berguna, efektif, efisien, mudah digunakan, dan memiliki antarmuka yang dapat diterima pengguna berdasarkan <i>Computer System Usability</i>

No.	Judul	Penulis	Metode	Hasil
		Vargas, Patricia Acosta- Vargas, Sandra Sanchez- Gordon, Tania Calle- Jimenez, Danilo Esparza, Yves Rybarczyk	menangkap efektivitas, efisiensi, dan kepuasan pengguna.	<i>Questionnaire.</i> Secara keseluruhan, platform diterima baik oleh peserta rehabilitasi.
9	Usability Testing pada M- Commerce Menggunakan Kuesioner USE	Thowaf Fuad Hasan, Tenia Wahyunin grum, Ariq Cahya Wardhana	Peneliti menggunakan kuesioner USE untuk mengumpulkan data dari pengguna aplikasi Tokopedia. Setelah data terkumpul, dilakukan analisis <i>usability</i> berdasarkan empat dimensi utama dari USE, disertai uji kinerja (<i>performance test</i>) sebagai pelengkap.	Aplikasi Tokopedia memperoleh skor usability 4,14, menunjukkan bahwa aplikasi dinilai baik, nyaman, dan memuaskan untuk digunakan oleh pengguna.
10	<i>Perceived usability evaluation of</i>	Debajyoti Pal, Vajirasak	Penelitian menerapkan dua pendekatan evaluasi usability, yaitu	Analisis ANOVA menunjukkan tidak ada efek signifikan

No.	Judul	Penulis	Metode	Hasil
	<i>Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India</i>	Vaniija	<i>System Usability Scale (SUS) dari ranah HCI dan Technology Acceptance Model (TAM) dari ranah Information Systems. Analisis dilanjutkan dengan ANOVA untuk menguji efek perbedaan platform terhadap persepsi pengguna.</i>	dari platform konsumsi terhadap skor SUS maupun TAM. Tidak ditemukan perbedaan berarti dalam tingkat <i>usability Microsoft Teams</i> berdasarkan platform yang digunakan.

2.2 USE Questionnaire

USE *Questionnaire* merupakan instrumen evaluasi yang dikembangkan untuk mengukur *usability* dari suatu sistem atau aplikasi. Instrumen ini pertama kali diperkenalkan oleh Lund (2001), yang menekankan bahwa *usability* tidak hanya berkaitan dengan kemampuan sistem untuk menyelesaikan tugas, tetapi juga melibatkan persepsi pengguna terhadap pengalaman mereka dalam menggunakan sistem tersebut. USE *Questionnaire* dirancang untuk mengukur empat dimensi utama: *Usefulness*, *Ease of Use*, *Ease of Learning*, dan *Satisfaction*. Dimensi-dimensi ini memberikan pandangan menyeluruh mengenai kualitas *usability* suatu sistem.

Tabel 2.2 USE Questionnaire

No.	Factor	Item
1	Usefulness	<i>It helps me be more effective.</i>
2		<i>It help me be more productive.</i>
3		<i>It is useful.</i>
4		<i>It gives me more control over the activities in my life.</i>
5		<i>It makes the things I want to accomplish easier to get done.</i>
6		<i>It saves me time when I use it.</i>
7		<i>It meets my needs.</i>
8		<i>It does everything I would expect it to do.</i>
9	Ease of Use	<i>It is easy to use.</i>
10		<i>It is simple to use.</i>
11		<i>It is user friendly.</i>
12		<i>It requires the fewer steps possible to accomplish what I want to do with it.</i>
13		<i>It is flexible.</i>
14		<i>Using it is effortless.</i>
15		<i>I can use it without written instructions.</i>
16		<i>I don't notice any inconsistencies as I use it.</i>
17		<i>Both occasional and regular users would like it.</i>
18		<i>I can recover from mistakes quickly and easily.</i>
19		<i>I can use it successfully every time.</i>
20	Ease of Learning	<i>I learned to use it quickly.</i>
21		<i>I easily remember how to use it.</i>
22		<i>It is easy to learn to use it.</i>
23		<i>I quickly become skillful with it</i>
24	Satisfaction	<i>I am satisfied with it.</i>
25		<i>I would recommend it to a friend.</i>
26		<i>It is fun to use.</i>
27		<i>It works the way I want it to work.</i>
28		<i>It is wonderful.</i>

No.	Factor	Item
29		<i>I feel I need to have it.</i>
30		<i>It is pleasant to use.</i>

Lund (2001) menyatakan bahwa USE *Questionnaire* memiliki keunggulan dalam kesederhanaan dan kemudahan penerapannya, sehingga dapat digunakan untuk mengevaluasi berbagai jenis sistem, termasuk perangkat lunak, situs web, dan aplikasi *mobile*. Dalam konteks SIAKAD, instrumen ini telah terbukti relevan dan efektif dalam mengevaluasi pengalaman pengguna, seperti yang ditunjukkan oleh penelitian Fadhilah dan Meuthia (2022), yang mengevaluasi *usability* SIAKAD di Politeknik Negeri Padang. Hasil penelitian tersebut menunjukkan bahwa USE *Questionnaire* mampu mengidentifikasi aspek-aspek yang memengaruhi kepuasan pengguna serta memberikan rekomendasi perbaikan yang konkret.

Hariyanto, Triyono, dan Köhler (2020) juga menggunakan USE *Questionnaire* untuk mengevaluasi sistem *e-learning* adaptif. Penelitian mereka menunjukkan bahwa instrumen ini dapat mengungkapkan kekuatan dan kelemahan sistem berdasarkan persepsi pengguna. Dimensi *Ease of Use* dan *Ease of Learning* dianggap sangat penting dalam meningkatkan adopsi dan efektivitas sistem. Studi ini menggarisbawahi pentingnya desain yang intuitif dan kemampuan sistem untuk memfasilitasi pembelajaran mandiri.

Selain itu, Kilis et al. (2022) menggunakan USE *Questionnaire* untuk mengevaluasi sistem operasi Android. Hasil penelitian mereka mengungkapkan bahwa dimensi *Usefulness* dan *Satisfaction* memiliki korelasi yang kuat dengan

tingkat loyalitas pengguna terhadap sistem. Penelitian ini menegaskan bahwa pengalaman pengguna yang positif dapat mendorong penggunaan berulang dan meningkatkan kepuasan secara keseluruhan.

Dalam tinjauan yang lebih luas, Moulaei, Moulaei, dan Bahaadinbeigy (2024) mengidentifikasi USE *Questionnaire* sebagai salah satu instrumen yang paling sering digunakan dalam evaluasi *usability* perangkat pintar dan robot. Penelitian ini menunjukkan bahwa USE *Questionnaire* telah berkembang menjadi alat yang universal dalam berbagai konteks teknologi, membuktikan fleksibilitas dan validitasnya dalam mengukur *usability*.

Keseluruhan studi ini menunjukkan bahwa USE *Questionnaire* merupakan alat evaluasi yang komprehensif dan serbaguna. Penggunaannya dalam berbagai konteks, termasuk SIAKAD, *e-learning*, dan perangkat pintar, membuktikan relevansinya dalam mengukur pengalaman pengguna dan memberikan wawasan yang berharga untuk pengembangan sistem lebih lanjut.

2.3 Rasch Measurement Theory (RMT)

Rasch Measurement Theory (RMT) merupakan metode analisis statistik yang dikembangkan oleh Georg Rasch untuk mengukur kemampuan, sikap, atau persepsi dengan menggunakan model probabilistik. Pendekatan ini bertujuan untuk mengubah data ordinal menjadi data interval, memungkinkan interpretasi yang lebih objektif dan konsisten dalam evaluasi instrumen pengukuran (Rasch, 1980). Dalam konteks pengujian *software usability*, teori ini membantu mengevaluasi kualitas dan keandalan alat ukur, seperti kuesioner, dengan

mengidentifikasi seberapa baik *item-item* dalam instrumen tersebut sesuai dengan model Rasch.

Rasch Measurement Theory (RMT) melakukan pendekatan matematis yang digunakan untuk mengevaluasi validitas dan reliabilitas suatu instrumen pengukuran berbasis data respon individu terhadap serangkaian *item* atau pertanyaan. Model ini pertama kali diperkenalkan oleh Georg Rasch (1960) dalam karyanya *Probabilistic Models for Some Intelligence and Attainment Tests*.

Model Rasch berbasis pada asumsi bahwa probabilitas seseorang untuk memberikan jawaban tertentu terhadap suatu *item* bergantung pada perbedaan antara kemampuan individu (*person ability*, θ) dan tingkat kesulitan *item* (*item difficulty*, β). Secara matematis, probabilitas seseorang dalam menjawab suatu *item* dengan benar dinyatakan sebagai:

$$\log\left(\frac{P_{nik}}{P_{ni(k-1)}}\right) = \theta_n - \beta_i - \tau_k \quad (2.1)$$

dimana P_{nik} adalah probabilitas responden n memilih kategori ke- k pada item i , θ_n adalah kemampuan individu, β_i adalah tingkat kesulitan *item*, τ_k adalah *threshold* ke- k (jarak antar kategori), dan k adalah jumlah kategori – 1

Model ini memiliki beberapa karakteristik, antara lain:

- 1) Independensi Invarian – Estimasi tingkat kesulitan suatu *item* tidak tergantung pada karakteristik sampel responden yang digunakan dalam

pengukuran, dan sebaliknya, estimasi kemampuan individu tidak bergantung pada kumpulan *item* tertentu.

2) *Unidimensionality* – Model Rasch mengasumsikan bahwa hanya satu faktor (kemampuan tertentu) yang mempengaruhi respons individu terhadap suatu set *item*.

3) *Local Independence* – Respons individu terhadap suatu *item* tidak dipengaruhi oleh respons terhadap *item* lain, kecuali oleh faktor yang diukur.

Aspek probabilistik dari model ini memungkinkan evaluasi yang lebih objektif terhadap performa suatu instrumen pengukuran dan memberikan interpretasi yang lebih akurat dibandingkan pendekatan klasik, seperti teori tes klasik (*Classical Test Theory*, CTT).

Aryadoust, Tan, dan Ng (2019) mengidentifikasi perkembangan RMT sebagai metode yang esensial dalam pengukuran psikometrik. Studi mereka menunjukkan bahwa RMT tidak hanya berguna dalam pendidikan tetapi juga diadopsi dalam berbagai bidang seperti pengujian bahasa dan evaluasi aplikasi kesehatan digital. Penggunaan RMT dalam konteks ini memastikan bahwa instrumen pengukuran memiliki validitas dan reliabilitas yang tinggi.

Dalam penerapannya, *Rasch Measurement Theory* memiliki beberapa jenis model yang disesuaikan dengan bentuk data dan jenis respons yang digunakan dalam instrumen penelitian. Model yang paling dasar adalah *Dichotomous Rasch Model*, yang digunakan untuk data dengan dua kategori respons, misalnya benar dan salah pada tes kemampuan. Model ini kemudian dikembangkan untuk mengakomodasi data dengan lebih dari dua kategori respons (*polytomous data*).

Salah satu pengembangannya adalah *Rating Scale Model* (RSM) yang digunakan ketika seluruh *item* memiliki struktur kategori respons yang sama, seperti skala Likert 1 sampai 5. Model lainnya adalah *Partial Credit Model* (PCM) yang digunakan ketika setiap *item* memiliki struktur kategori yang berbeda. Selain itu terdapat pula model lain seperti *Many-Facet Rasch Model* (MFRM) yang digunakan ketika pengukuran melibatkan lebih dari dua komponen, misalnya responden, *item*, dan penilai.

2.4 *Rating Scale Model*

Model Rasch klasik pada awalnya dirancang untuk data dikotomis, yaitu respons yang hanya memiliki dua kemungkinan kategori seperti benar–salah atau setuju–tidak setuju. Namun, dalam praktik penelitian sosial dan sistem informasi, instrumen pengukuran sering kali menggunakan skala penilaian bertingkat seperti skala Likert yang memiliki lebih dari dua kategori respons. Oleh karena itu, diperlukan perluasan model Rasch untuk menangani data *polytomous*.

Pengembangan model Rasch untuk data *polytomous* kemudian dilakukan oleh David Andrich melalui model yang dikenal sebagai *Rating Scale Model* (RSM). Model ini merupakan ekstensi dari model Rasch yang dirancang khusus untuk instrumen dengan kategori respons yang sama pada setiap item, seperti skala Likert lima tingkat yang umum digunakan dalam penelitian usability. RSM mengasumsikan bahwa seluruh *item* dalam instrumen memiliki struktur kategori yang identik sehingga jarak antar kategori respon dapat dimodelkan secara konsisten di seluruh *item*.

Secara konseptual, RSM memodelkan probabilitas seorang responden memilih kategori tertentu pada suatu item sebagai fungsi dari tiga komponen utama, yaitu kemampuan responden (*person ability*), tingkat kesulitan item (*item difficulty*), dan ambang kategori (*category threshold*). Kemampuan responden merepresentasikan kecenderungan individu dalam memberikan penilaian yang lebih tinggi atau lebih rendah terhadap suatu konstruk laten, sedangkan kesulitan *item* menggambarkan tingkat kemudahan atau kesulitan suatu pernyataan untuk disetujui oleh responden. Sementara itu, ambang kategori menunjukkan titik transisi probabilistik antara dua kategori respons yang berdekatan pada skala penilaian.

Model matematis dari RSM dapat dinyatakan sebagai berikut:

$$P_{nik} = \frac{e^{\sum_{m=0}^k (\beta_n - \delta_i - \tau_m)}}{\sum_{j=0}^m e^{\sum_{m=0}^j (\beta_n - \delta_i - \tau_m)}} \quad (2.2)$$

di mana β_n menyatakan kemampuan responden, δ_i menyatakan parameter kesulitan item, dan τ_m merupakan parameter ambang kategori pada skala respons. Melalui model ini, data ordinal dari kuesioner dapat ditransformasikan menjadi ukuran interval logit yang memungkinkan analisis statistik yang lebih akurat dan bermakna.

Penggunaan RSM dalam penelitian ini didasarkan pada karakteristik instrumen yang digunakan, yaitu *USE Questionnaire* yang memiliki struktur skala Likert yang konsisten pada setiap *item*. Instrumen *USE Questionnaire* mengukur empat dimensi utama *usability*, yaitu *usefulness*, *ease of use*, *ease of learning*, dan

satisfaction. Karena seluruh *item* dalam instrumen ini menggunakan kategori respons yang sama, maka RSM menjadi model Rasch yang paling sesuai untuk mengestimasi parameter responden dan item secara simultan. Selain itu, penerapan RSM memungkinkan peneliti untuk mengevaluasi kualitas instrumen melalui indikator kesesuaian model (*fit statistics*), reliabilitas *item* dan responden, serta struktur kategori skala penilaian.

Dalam konteks penelitian ini, penerapan RSM dilakukan setelah proses pengumpulan data kuesioner dari responden pengguna SIAKAD. Data respons yang diperoleh terlebih dahulu disusun dalam bentuk matriks responden-item, di mana setiap sel matriks merepresentasikan skor respons pada skala Likert. Selanjutnya, data tersebut dianalisis menggunakan pendekatan estimasi parameter Rasch untuk memperoleh nilai kemampuan responden, tingkat kesulitan item, serta ambang kategori skala penilaian. Hasil estimasi kemudian digunakan untuk mengevaluasi reliabilitas instrumen, mengidentifikasi item yang tidak sesuai dengan model, serta memastikan bahwa skala pengukuran yang digunakan memenuhi asumsi pengukuran Rasch.

Melalui proses ini, data kuesioner yang pada awalnya bersifat ordinal dapat dikonversi menjadi ukuran interval dalam satuan logit. Transformasi ini sangat penting dalam penelitian usability karena memungkinkan interpretasi yang lebih objektif terhadap persepsi pengguna terhadap sistem. Selain itu, hasil pengukuran berbasis Rasch juga memberikan dasar yang lebih kuat untuk analisis lanjutan, termasuk pemetaan tingkat kepentingan dan kinerja pada metode IPA yang digunakan dalam penelitian ini. Dengan demikian, integrasi antara USE

Questionnaire, RMT melalui RSM, dan IPA membentuk suatu kerangka evaluasi *usability* yang komprehensif dan terukur secara ilmiah.

2.5 Maximum Marginal Likelihood Estimation (MMLE)

Dalam analisis berbasis RMT, estimasi parameter merupakan tahap penting untuk memperoleh ukuran kemampuan responden dan karakteristik *item* secara akurat. Salah satu pendekatan estimasi yang banyak digunakan dalam model Rasch *polytomous* adalah *Maximum Marginal Likelihood Estimation* (MMLE). Metode ini dikembangkan dalam kerangka teori respons butir modern untuk mengatasi keterbatasan pendekatan estimasi tradisional yang tidak mampu menangani distribusi kemampuan responden secara efisien pada model dengan banyak parameter.

Secara konseptual, MMLE merupakan metode estimasi berbasis *likelihood* yang menghitung parameter *item* dengan cara memarginalkan parameter kemampuan responden dari fungsi *likelihood*. Dalam pendekatan ini, kemampuan responden diperlakukan sebagai variabel laten yang mengikuti distribusi probabilitas tertentu, umumnya distribusi normal standar. Dengan memarginalkan parameter kemampuan tersebut, proses estimasi dapat difokuskan pada parameter *item* sehingga perhitungan menjadi lebih stabil dan efisien, khususnya pada model *polytomous* seperti RSM.

Penggunaan MMLE dalam RSM menjadi penting karena model ini melibatkan beberapa parameter sekaligus, yaitu parameter kemampuan responden, parameter kesulitan *item*, serta parameter ambang kategori. Jika seluruh parameter diestimasi secara simultan menggunakan metode *likelihood* biasa, proses

komputasi akan menjadi sangat kompleks. Oleh karena itu, MMLE memisahkan estimasi parameter item dari distribusi kemampuan responden dengan melakukan integrasi terhadap parameter kemampuan. Pendekatan ini memungkinkan estimasi parameter item dilakukan secara lebih efisien tanpa harus mengestimasi setiap parameter responden secara langsung pada tahap awal.

Secara matematis, fungsi *likelihood* marginal dalam MMLE dapat dituliskan sebagai berikut:

$$L(\delta, \tau) = \prod_{n=1}^N \int \left[\prod_{i=1}^I P_{ni}(X_{ni} | \theta_n, \delta_i, \tau) \right] f(\theta_n) d\theta_n \quad (2.3)$$

di mana P_{ni} merupakan probabilitas respon responden n pada item i , θ_n menyatakan kemampuan laten responden, δ_i adalah parameter kesulitan item, dan τ adalah parameter ambang kategori pada skala respons. Fungsi $f(\theta_n)$ merepresentasikan distribusi probabilitas kemampuan responden yang biasanya diasumsikan mengikuti distribusi normal. Proses integrasi pada persamaan tersebut menghasilkan *likelihood* marginal yang hanya bergantung pada parameter item, sehingga parameter tersebut dapat diestimasi melalui proses optimisasi maksimum *likelihood*.

Dalam implementasi praktis, proses integrasi dalam MMLE biasanya tidak dapat diselesaikan secara analitik sehingga diperlukan pendekatan numerik. Salah satu teknik yang umum digunakan adalah integrasi numerik menggunakan metode kuadratur, seperti *Gauss-Hermite quadrature*, yang memungkinkan pendekatan

integral terhadap distribusi kemampuan responden secara efisien. Setelah nilai *likelihood* marginal diperoleh, parameter *item* diestimasi melalui prosedur iteratif hingga mencapai kondisi konvergensi.

Pada penelitian ini, pendekatan MMLE digunakan untuk mengestimasi parameter dalam RSM yang diterapkan pada data kuesioner *usability*. Data respons yang diperoleh dari pengguna Sistem Informasi Akademik terlebih dahulu disusun dalam bentuk matriks responden-*item*. Selanjutnya, probabilitas respons pada setiap kategori dihitung menggunakan formulasi RSM. Berdasarkan probabilitas tersebut, fungsi *likelihood* marginal kemudian dihitung dengan mengintegrasikan distribusi kemampuan responden. Proses optimisasi dilakukan secara iteratif untuk memperoleh estimasi parameter kesulitan *item* dan ambang kategori yang memaksimalkan fungsi *likelihood*.

Setelah parameter *item* diperoleh, langkah berikutnya adalah mengestimasi kemampuan responden berdasarkan parameter *item* yang telah diketahui. Nilai kemampuan responden kemudian dinyatakan dalam satuan logit, sehingga memungkinkan interpretasi yang lebih objektif terhadap persepsi pengguna terhadap sistem. Hasil estimasi tersebut selanjutnya digunakan untuk mengevaluasi kualitas instrumen, termasuk analisis reliabilitas responden, reliabilitas *item*, serta kesesuaian model terhadap data empiris.

Dengan menggunakan pendekatan MMLE, proses estimasi parameter dalam RSM dapat dilakukan secara lebih stabil dan efisien, terutama ketika jumlah responden dan *item* relatif besar. Oleh karena itu, metode ini menjadi salah satu teknik estimasi yang umum digunakan dalam analisis Rasch modern, khususnya

pada penelitian yang melibatkan instrumen pengukuran dengan skala respons bertingkat seperti kuesioner usability dalam evaluasi sistem informasi.

2.6 *Gauss–Hermite Quadrature*

Dalam proses estimasi parameter pada RMT, khususnya ketika menggunakan metode MMLE, diperlukan proses integrasi terhadap distribusi kemampuan laten responden. Integrasi ini muncul karena parameter kemampuan responden diperlakukan sebagai variabel acak yang mengikuti distribusi probabilitas tertentu, umumnya distribusi normal. Namun, pada banyak kasus, integral yang muncul dalam fungsi *likelihood* marginal tidak dapat diselesaikan secara analitik. Oleh karena itu, digunakan pendekatan integrasi numerik untuk mendekati nilai integral tersebut. Salah satu metode integrasi numerik yang paling umum digunakan dalam model Rasch *polytomous* adalah *Gauss–Hermite Quadrature*.

Gauss–Hermite Quadrature merupakan teknik integrasi numerik yang dirancang khusus untuk menghitung integral dari fungsi yang dikalikan dengan fungsi bobot eksponensial berbentuk distribusi normal. Metode ini berasal dari teori kuadratur Gaussian yang memanfaatkan titik evaluasi tertentu (*quadrature points*) dan bobot tertentu (*quadrature weights*) untuk memperoleh pendekatan integral yang akurat dengan jumlah perhitungan yang relatif kecil. Dalam konteks model Rasch, teknik ini sangat sesuai karena distribusi kemampuan responden umumnya diasumsikan mengikuti distribusi normal standar.

Pada model *polytomous* seperti RSM, proses estimasi parameter dengan pendekatan MMLE melibatkan integral terhadap distribusi kemampuan responden.

Integral tersebut muncul pada fungsi *likelihood* marginal yang mengintegrasikan probabilitas respons terhadap seluruh kemungkinan nilai kemampuan responden. Karena integral ini tidak memiliki solusi tertutup dalam bentuk analitik, maka *Gauss–Hermite Quadrature* digunakan untuk menghitung pendekatan numeriknya secara efisien. Metode ini memungkinkan integral kontinu terhadap kemampuan responden diaproksimasi sebagai penjumlahan terbobot dari sejumlah titik diskret.

Secara matematis, integral yang muncul dalam proses estimasi MMLE dapat dituliskan sebagai berikut:

$$\int_{-\infty}^{\infty} g(\theta) e^{-\theta^2} d\theta \quad (2.4)$$

Integral tersebut kemudian didekati menggunakan *Gauss–Hermite Quadrature* dengan formulasi:

$$\int_{-\infty}^{\infty} g(\theta) e^{-\theta^2} d\theta \approx \sum_{q=1}^Q w_q g(x_q) \quad (2.5)$$

di mana x_q merupakan titik kuadratur (*quadrature nodes*), w_q merupakan bobot kuadratur (*quadrature weights*), dan Q adalah jumlah titik kuadratur yang digunakan dalam proses integrasi. Fungsi $g(\theta)$ merepresentasikan bagian dari fungsi *likelihood* yang bergantung pada parameter kemampuan responden.

Dalam implementasi pada RSM, fungsi $g(\theta)$ umumnya merepresentasikan hasil perkalian probabilitas respons terhadap seluruh *item* bagi seorang responden. Dengan demikian, integral pada *likelihood* marginal dapat dihitung sebagai penjumlahan terbobot dari nilai fungsi probabilitas pada sejumlah titik kemampuan yang telah ditentukan sebelumnya. Semakin banyak titik kuadratur yang digunakan, maka pendekatan terhadap integral yang sebenarnya akan semakin akurat, meskipun dengan konsekuensi peningkatan kompleksitas komputasi.

Pada penelitian ini, *Gauss–Hermite Quadrature* digunakan untuk menghitung integral pada fungsi *likelihood* marginal dalam proses estimasi parameter menggunakan metode MMLE. Proses ini dimulai dengan menentukan sejumlah titik kuadratur dan bobot yang sesuai dengan distribusi normal standar. Selanjutnya, untuk setiap titik kemampuan tersebut dihitung probabilitas respons responden terhadap seluruh *item* menggunakan formulasi RSM. Nilai probabilitas yang diperoleh kemudian dikalikan dengan bobot kuadratur yang sesuai dan dijumlahkan untuk memperoleh pendekatan nilai integral pada fungsi *likelihood* marginal.

Setelah nilai integral diperoleh, langkah selanjutnya adalah melakukan optimisasi fungsi *likelihood* untuk memperoleh estimasi parameter *item* dan parameter ambang kategori yang memaksimalkan *likelihood* tersebut. Proses ini dilakukan secara iteratif hingga nilai parameter mencapai kondisi konvergensi. Hasil estimasi yang diperoleh kemudian digunakan untuk menghitung ukuran

kemampuan responden dalam satuan logit serta mengevaluasi kualitas instrumen pengukuran.

Dengan demikian, penggunaan *Gauss-Hermite Quadrature* dalam penelitian ini berperan sebagai teknik integrasi numerik yang memungkinkan proses estimasi parameter pada RSM dilakukan secara efisien dan akurat. Integrasi metode ini dengan pendekatan MMLE memungkinkan analisis Rasch terhadap data kuesioner *usability* dilakukan secara komputasional, sehingga menghasilkan estimasi parameter yang stabil dan dapat digunakan sebagai dasar evaluasi *usability* SIAKAD yang diteliti.

2.7 Posterior Probability

Dalam kerangka RMT, analisis data respons tidak hanya bertujuan untuk mengestimasi parameter *item*, tetapi juga untuk memperoleh estimasi kemampuan laten setiap responden berdasarkan pola respons yang diberikan. Salah satu konsep penting yang digunakan dalam proses ini adalah *Posterior Probability*. *Posterior probability* merupakan probabilitas suatu parameter laten setelah mempertimbangkan informasi dari data yang diamati. Dalam konteks model Rasch, probabilitas ini digunakan untuk memperkirakan tingkat kemampuan responden dengan memanfaatkan hubungan antara parameter *item* dan pola respons yang dihasilkan oleh responden.

Konsep *posterior probability* berasal dari pendekatan probabilitas Bayesian yang memanfaatkan hubungan antara probabilitas awal (*prior probability*), fungsi *likelihood*, dan probabilitas setelah observasi data (*posterior probability*). Dalam analisis Rasch, *prior probability* biasanya merepresentasikan asumsi distribusi

kemampuan responden, yang umumnya diasumsikan mengikuti distribusi normal standar. Setelah data respons diperoleh, informasi tersebut dikombinasikan dengan *prior distribution* untuk menghasilkan distribusi *posterior* yang menggambarkan probabilitas kemampuan responden berdasarkan pola respons yang diamati.

Penggunaan *posterior probability* menjadi penting dalam model polytomous seperti RSM karena estimasi parameter responden tidak dapat diperoleh secara langsung dari fungsi likelihood marginal yang digunakan dalam MMLE. Pada tahap estimasi MMLE, parameter kemampuan responden telah dimarginalkan untuk memudahkan estimasi parameter *item*. Oleh karena itu, setelah parameter *item* berhasil diestimasi, diperlukan langkah tambahan untuk memperkirakan kemampuan responden secara individual. *Posterior probability* digunakan pada tahap ini untuk menentukan distribusi probabilitas kemampuan responden yang paling konsisten dengan pola respons yang diberikan.

Secara matematis, *posterior probability* dapat dinyatakan menggunakan formulasi berikut:

$$P(\theta_n|X_n) = \frac{P(X_n|\theta_n)f(\theta_n)}{\int P(X_n|\theta_n)f(\theta_n)d\theta_n} \quad (2.6)$$

di mana $P(\theta_n|X_n)$ merupakan probabilitas posterior dari kemampuan responden n setelah mempertimbangkan data respons X_n . Simbol $P(X_n|\theta_n)$ menyatakan fungsi *likelihood* dari pola respons responden terhadap seluruh *item* berdasarkan nilai kemampuan tertentu, sedangkan $f(\theta_n)$ merupakan distribusi *prior* dari

kemampuan responden. Penyebut pada persamaan tersebut merupakan faktor normalisasi yang memastikan bahwa total probabilitas posterior bernilai satu.

Dalam implementasi pada RSM, fungsi *likelihood* $P(X_n|\theta_n)$ dihitung sebagai hasil perkalian probabilitas respons terhadap seluruh *item* berdasarkan parameter *item* yang telah diestimasi sebelumnya. Probabilitas respons pada setiap *item* diperoleh dari formulasi probabilitas dalam RSM yang mempertimbangkan parameter kesulitan *item* dan ambang kategori skala respons. Dengan demikian, *posterior probability* menggambarkan seberapa besar kemungkinan seorang responden memiliki tingkat kemampuan tertentu berdasarkan pola respons yang diberikan terhadap seluruh *item* dalam kuesioner.

Pada penelitian ini, *posterior probability* digunakan untuk mengestimasi kemampuan responden setelah parameter *item* dan parameter ambang kategori diperoleh melalui proses MMLE. Proses perhitungannya dilakukan dengan mengevaluasi fungsi *likelihood* pada sejumlah titik kemampuan yang telah ditentukan sebelumnya, yang biasanya diperoleh dari titik kuadratur dalam metode *Gauss–Hermite Quadrature*. Nilai *likelihood* pada setiap titik kemampuan kemudian dikalikan dengan nilai *prior distribution* dan dinormalisasi untuk menghasilkan distribusi posterior bagi setiap responden.

Setelah distribusi posterior diperoleh, nilai estimasi kemampuan responden dapat dihitung menggunakan pendekatan statistik tertentu, seperti nilai ekspektasi posterior atau nilai maksimum dari distribusi posterior tersebut. Nilai kemampuan yang dihasilkan kemudian dinyatakan dalam satuan logit, sehingga

memungkinkan interpretasi yang lebih objektif terhadap tingkat persepsi atau pengalaman pengguna terhadap sistem yang dievaluasi.

Dengan demikian, penggunaan *posterior probability* dalam penelitian ini memungkinkan estimasi kemampuan responden dilakukan secara konsisten dengan kerangka Rasch setelah parameter *item* diperoleh melalui MMLE. Pendekatan ini memberikan estimasi kemampuan yang lebih stabil karena memanfaatkan seluruh informasi yang tersedia, baik dari distribusi prior maupun dari pola respons yang diamati. Hasil estimasi kemampuan tersebut selanjutnya digunakan sebagai dasar dalam analisis lanjutan untuk mengevaluasi tingkat *usability* SIAKAD yang menjadi objek penelitian.

2.8 *Fit Statistic*

Dalam analisis berbasis RMT, salah satu aspek penting yang harus diperhatikan adalah kesesuaian antara data empiris dengan model pengukuran yang digunakan. Untuk mengevaluasi tingkat kesesuaian tersebut digunakan ukuran yang dikenal sebagai *Fit Statistic*. *Fit statistic* merupakan indikator statistik yang digunakan untuk menilai sejauh mana pola respons yang dihasilkan oleh responden sesuai dengan ekspektasi probabilistik yang ditentukan oleh model Rasch. Dengan kata lain, ukuran ini berfungsi untuk mengidentifikasi apakah data yang diperoleh konsisten dengan asumsi pengukuran yang mendasari model Rasch.

Dalam model *polytomous* seperti RSM, evaluasi kesesuaian model menjadi sangat penting karena model ini melibatkan beberapa parameter sekaligus, seperti kemampuan responden, kesulitan *item*, serta ambang kategori respons. Apabila

terdapat *item* atau responden yang menunjukkan pola respons yang sangat berbeda dari yang diprediksi oleh model, maka estimasi parameter yang dihasilkan dapat menjadi kurang akurat. Oleh karena itu, analisis *fit statistic* digunakan untuk mendeteksi adanya ketidaksesuaian (*misfit*) pada data sehingga peneliti dapat mengevaluasi kualitas instrumen maupun perilaku respons responden.

Dalam praktik analisis Rasch, dua jenis *fit statistic* yang paling umum digunakan adalah *Infit Mean Square (Information-Weighted Mean Square)* dan *Outfit Mean Square (Outlier-Sensitive Mean Square)*. Kedua indikator ini mengukur besarnya deviasi antara respons yang diamati dengan respons yang diprediksi oleh model Rasch. Perbedaan utama keduanya terletak pada sensitivitas terhadap respons ekstrem. *Infit* lebih sensitif terhadap respons yang berada dekat dengan tingkat kemampuan responden, sedangkan *outfit* lebih sensitif terhadap respons yang sangat tidak terduga atau bersifat ekstrem.

Secara matematis, statistik kesesuaian model dapat dinyatakan melalui formulasi berikut:

$$Infit\ MNSQ = \frac{\sum w_{ni}(X_{ni} - E_{ni})^2}{\sum w_{ni}} \quad (2.7)$$

$$Outfit\ MNSQ = \frac{\sum (X_{ni} - E_{ni})^2}{\sum V_{ni}} \quad (2.8)$$

di mana X_{ni} merupakan skor respons aktual responden n pada *item* i , E_{ni} adalah nilai ekspektasi respons yang diprediksi oleh model, V_{ni} merupakan varians dari

respons yang diprediksi, dan w_{ni} adalah bobot informasi yang digunakan dalam perhitungan *infit statistic*. Nilai residual ($X_{ni} - E_{ni}$) menunjukkan selisih antara respons yang diamati dan respons yang diperkirakan oleh model.

Interpretasi nilai *fit statistic* umumnya didasarkan pada nilai *mean square* (MNSQ). Nilai MNSQ yang mendekati 1 menunjukkan bahwa data sesuai dengan ekspektasi model Rasch. Nilai yang lebih besar dari 1 menunjukkan adanya variasi respons yang lebih besar dari yang diprediksi oleh model, yang mengindikasikan potensi *misfit* atau respons yang tidak konsisten. Sebaliknya, nilai yang jauh lebih kecil dari 1 menunjukkan bahwa variasi respons terlalu kecil, yang dapat mengindikasikan adanya redundansi pada *item* atau respons yang terlalu deterministik.

Dalam konteks penelitian ini, analisis *fit statistic* digunakan untuk mengevaluasi kualitas instrumen *usability* yang diukur menggunakan kuesioner. Setelah parameter *item* dan kemampuan responden diestimasi menggunakan pendekatan Rasch, nilai residual antara respons yang diamati dan yang diprediksi oleh model dihitung untuk setiap pasangan responden dan *item*. Berdasarkan nilai residual tersebut, dihitung nilai *infit* dan *outfit* MNSQ untuk masing-masing *item* dan responden.

Hasil analisis ini digunakan untuk mengidentifikasi item yang tidak sesuai dengan model pengukuran. *Item* dengan nilai *fit statistic* yang berada di luar rentang yang dapat diterima dapat dipertimbangkan untuk ditinjau kembali, diperbaiki, atau bahkan dieliminasi dari instrumen pengukuran. Selain itu, analisis ini juga dapat membantu mendeteksi responden yang memberikan respons secara

tidak konsisten, misalnya karena kurang memahami pertanyaan atau menjawab secara acak.

Dengan demikian, penggunaan *fit statistic* dalam penelitian ini berperan penting dalam memastikan bahwa data yang dianalisis benar-benar memenuhi asumsi pengukuran dalam model Rasch. Evaluasi kesesuaian model ini memberikan dasar yang kuat untuk memastikan bahwa hasil pengukuran *usability* SIAKAD yang diperoleh dari kuesioner dapat diinterpretasikan secara valid dan reliabel.

2.9 Importance Performance Analysis

Importance Performance Analysis (IPA) merupakan sebuah metode evaluasi yang pertama kali diperkenalkan oleh Martilla dan James (1977) dalam konteks pemasaran jasa, yang dirancang untuk membantu manajer dalam mengidentifikasi prioritas peningkatan layanan berdasarkan persepsi pengguna. Dalam perkembangannya, metode ini banyak diadaptasi dalam berbagai studi, termasuk dalam studi *usability* untuk mengevaluasi atribut *usability* yang dianggap penting oleh pengguna serta sejauh mana atribut tersebut telah terpenuhi oleh sistem yang diuji.

Secara konseptual, IPA berangkat dari logika sederhana yang menggabungkan dua dimensi kunci: *importance* (tingkat kepentingan) dan *performance* (tingkat kinerja). Dimensi *importance* digunakan untuk menangkap sejauh mana suatu atribut *usability* dianggap penting oleh pengguna dalam mendukung tujuan penggunaan sistem. Sedangkan dimensi *performance* mengukur sejauh mana sistem telah memenuhi ekspektasi pengguna terkait atribut

usability tersebut. Dengan demikian, IPA secara runtut memperhitungkan kedua dimensi tersebut sehingga dapat menghasilkan pemetaan atribut *usability* yang lebih objektif dan relevan dengan pengalaman pengguna.

Menurut Martilla dan James (1977), metode ini membantu dalam mengalokasikan sumber daya secara efektif dengan memetakan atribut *usability* ke dalam empat kuadran yang merepresentasikan tindakan strategis yang berbeda. Atribut yang berada pada *importance* tinggi dan *performance* rendah menunjukkan area yang memerlukan perhatian khusus, sedangkan atribut dengan *importance* tinggi dan *performance* tinggi mencerminkan area yang harus dipertahankan. Adapun atribut dengan *importance* rendah memiliki prioritas yang lebih rendah dalam upaya perbaikan. Pendekatan ini menekankan keseimbangan antara persepsi pengguna terhadap kebutuhan sistem dan realisasi performa aktual, sehingga sangat berguna dalam konteks evaluasi *usability* SIAKAD.

Langkah awal dalam penerapan IPA adalah pengumpulan data melalui kuesioner yang dirancang untuk mengukur kedua dimensi tersebut. Responden diminta untuk memberikan skor pada skala Likert (misalnya 1–5) terhadap setiap atribut *usability*, baik dari sisi *importance* maupun *performance*. Selanjutnya, nilai rata-rata *importance* (\bar{I}) dan *performance* (\bar{P}) untuk setiap atribut *usability* dihitung menggunakan rumus masing-masing sebagai berikut:

$$\bar{I}_j = \frac{\sum_{i=1}^n I_{ij}}{n} \quad (2.9)$$

$$\bar{P}_j = \frac{\sum_{i=1}^n P_{ij}}{n} \quad (2.10)$$

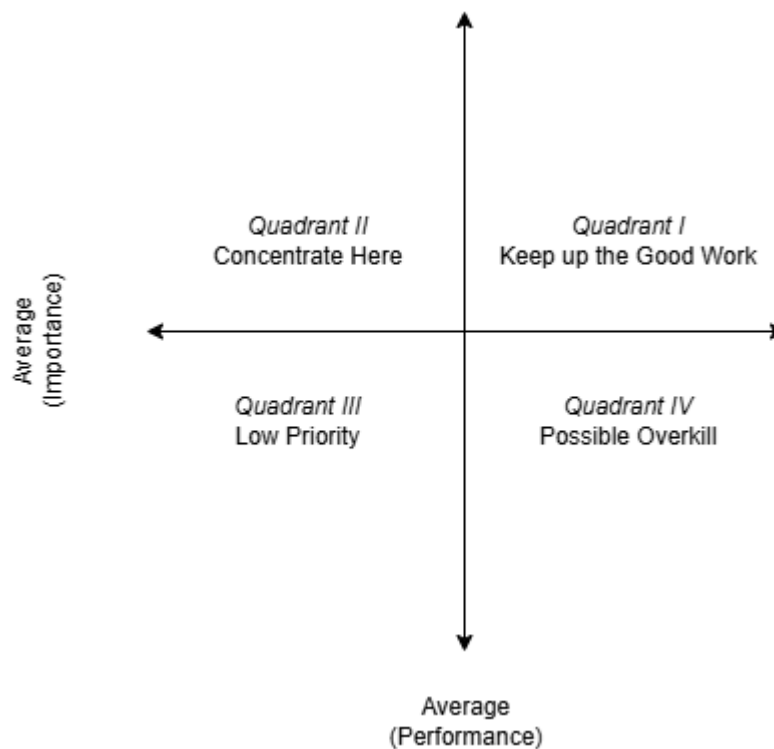
dimana \bar{I}_j adalah rata-rata *importance* atribut ke-j, \bar{P}_j adalah rata-rata *performance* atribut ke-j, I_{ij} adalah skor *importance* dari responden ke-i pada atribut ke-j, P_{ij} adalah skor *performance* dari responden ke-i pada atribut ke-j, dan n = jumlah responden.

Hasil perhitungan ini kemudian digunakan untuk memetakan atribut *usability* pada diagram IPA dua dimensi, di mana sumbu horizontal (X) merepresentasikan *importance* dan sumbu vertikal (Y) merepresentasikan *performance*. Dengan menarik garis pembagi (*cut-off*) pada masing-masing sumbu berdasarkan nilai rata-rata keseluruhan *importance* dan *performance*, diagram tersebut dibagi menjadi empat kuadran:

Dengan kerangka tersebut, IPA memberikan dasar analisis yang sistematis untuk mengidentifikasi atribut *usability* yang memerlukan perhatian prioritas, atribut yang perlu dipertahankan, serta atribut yang cenderung memiliki prioritas lebih rendah dalam upaya perbaikan. Pendekatan ini menjadikan IPA sebagai alat bantu yang efektif dalam pengambilan keputusan strategis, khususnya dalam pengembangan SIAKAD, dengan tetap mempertimbangkan perspektif pengguna secara integral.

- 1) Kuadran I (*Keep Up the Good Work*) — *importance* tinggi dan *performance* tinggi;
- 2) Kuadran II (*Concentrate Here*) — *importance* tinggi tetapi *performance* rendah;

- 3) Kuadran III (*Low Priority*) — *importance* rendah dan *performance* rendah;
- 4) Kuadran IV (*Possible Overkill*) — *importance* rendah tetapi *performance* tinggi.



Gambar 2.1 Diagram *Importance Performance Analysis*

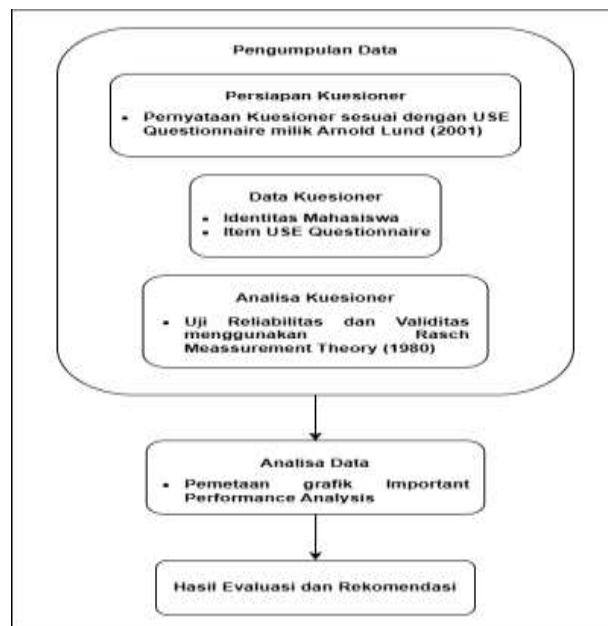
Dengan kerangka tersebut, IPA memberikan dasar analisis yang sistematis untuk mengidentifikasi atribut *usability* yang memerlukan perhatian prioritas, atribut yang perlu dipertahankan, serta atribut yang cenderung memiliki prioritas lebih rendah dalam upaya perbaikan. Pendekatan ini menjadikan IPA sebagai alat bantu yang efektif dalam pengambilan keputusan strategis, khususnya dalam pengembangan SIAKAD, dengan tetap mempertimbangkan perspektif pengguna secara integral.

BAB III

DESAIN DAN IMPLEMENTASI

3.1 Desain Penelitian

Desain penelitian disusun secara sistematis untuk menggambarkan alur pelaksanaan penelitian mulai dari tahap perencanaan hingga tahap analisis dan penarikan kesimpulan. Rangkaian tahapan tersebut menunjukkan hubungan logis antarproses yang dilakukan, mencakup penentuan objek penelitian, pengumpulan data, pengolahan data, serta metode analisis serta evaluasi yang digunakan untuk menjawab tujuan penelitian. Alur pelaksanaan penelitian secara keseluruhan disajikan dalam bentuk bagan untuk memberikan gambaran yang jelas dan terstruktur mengenai proses penelitian yang dilakukan, sebagaimana dapat dilihat pada bagan berikut ini.



Gambar 3.1 Bagan Penelitian

3.1.1 Persiapan Kuesioner

Kuesioner yang digunakan dalam penelitian ini disusun berdasarkan instrumen *USE Questionnaire* yang dikembangkan oleh Lund (2001). Instrumen ini terdiri dari empat dimensi utama, yaitu *usefulness*, *ease of use*, *ease of learning*, dan *satisfaction*. Setiap dimensi diuraikan menjadi beberapa butir pernyataan yang dirancang untuk menggambarkan persepsi pengguna terhadap kualitas *usability* sistem.

Selain itu, penggunaan kuesioner dalam evaluasi *usability* juga didukung oleh sejumlah penelitian terdahulu, salah satunya penelitian oleh Zhou et al. (2019) yang mengembangkan dan memvalidasi instrumen *MHealth App Usability Questionnaire* (MAUQ). Studi tersebut menegaskan pentingnya penyusunan instrumen yang terstruktur, valid, dan sesuai konteks sistem yang diuji. Pada ranah aplikasi komersial, Thowaf (2021) juga menunjukkan bahwa *USE Questionnaire* efektif dalam mengukur *usability* aplikasi *M-Commerce* melalui analisis persepsi dan *performance test*, sehingga memperkuat dasar metodologis penyusunan instrumen dalam penelitian ini.

Pada tahap persiapan, seluruh butir kuesioner disusun ulang dalam format yang lebih sistematis dan mudah dipahami oleh responden. Setiap pernyataan diberikan dalam bentuk skala Likert, dan untuk kepentingan analisis *Importance-Performance Analysis* (IPA), setiap butir disertai dua jenis pertanyaan, yaitu pertanyaan mengenai tingkat kepentingan pernyataan bagi pengguna serta pertanyaan mengenai tingkat kinerja atau relevansinya terhadap kondisi SIAKAD saat ini.

Membantu saya menjadi lebih efektif dalam mengikuti kegiatan perkuliahan.

Description (optional)

Seberapa penting pernyataan di atas menurut Anda? (U1) *

1 2 3 4 5

Sangat Tidak Setuju Sangat Setuju

Menurut Anda, apakah pernyataan di atas relevan dengan kondisi SIAKAD saat ini? (U1) *

1 2 3 4 5

Sangat Tidak Setuju Sangat Setuju

Gambar 3.2 Tampilan Kuesioner

Instrumen kuesioner ini terdiri dari total 60 pertanyaan yang berasal dari 30 pernyataan utama pada *USE Questionnaire*. Setiap pernyataan dipecah menjadi dua bentuk pertanyaan untuk memenuhi kebutuhan analisis *Importance–Performance Analysis*. Bentuk pertama digunakan untuk menilai tingkat kepentingan suatu aspek menurut pengguna, sedangkan bentuk kedua digunakan untuk menilai tingkat kinerja atau relevansi aspek tersebut terhadap kondisi SIAKAD saat ini. Dengan pengandaan ini, peneliti dapat memperoleh gambaran yang lebih mendalam mengenai kesenjangan antara harapan dan pengalaman pengguna.

3.1.2 Data Kuesioner

Proses evaluasi *usability* pada penelitian ini diawali dengan pengisian kuesioner oleh responden yang sebelumnya telah menggunakan Sistem Informasi Akademik (SIAKAD). Pengisian kuesioner dilakukan setelah responden berinteraksi langsung dengan sistem agar penilaian yang diberikan mencerminkan

pengalaman nyata, baik dari segi tampilan, kemudahan penggunaan, alur navigasi, maupun fungsi-fungsi utama yang tersedia.

Subjek penelitian dalam studi ini adalah mahasiswa Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Berdasarkan data yang tersedia, jumlah mahasiswa pada fakultas tersebut adalah sebanyak 3.830 orang. Untuk menentukan jumlah minimal responden dengan tingkat kesalahan (*margin of error*) sebesar 10 persen, digunakan rumus Slovin dengan mempertimbangkan jumlah populasi yang ada.

$$\frac{3830}{1 + 3830 \cdot (10\%)^2} \approx 97.45$$

Sesuai dengan perhitungan di atas, maka jumlah minimal responden yang diperlukan agar jumlah sampel dapat merepresentasikan populasi adalah sebanyak 98 responden.

Kuesioner disusun berdasarkan *USE Questionnaire* dan disajikan melalui Google Form agar mudah diakses oleh responden. Instrumen ini berfungsi untuk memperoleh umpan balik yang akurat terkait persepsi pengguna terhadap kualitas *usability* SIAKAD, mencakup aspek kegunaan, kemudahan penggunaan, kemudahan mempelajari sistem, serta tingkat kepuasan pengguna.

Setiap responden diminta untuk memberikan penilaian terhadap butir-butir pertanyaan yang telah disesuaikan dengan konteks penggunaan SIAKAD. Dalam konteks IPA, setiap pernyataan disajikan dalam dua bentuk, yaitu penilaian tingkat kepentingan serta penilaian tingkat kinerja. Dengan demikian, data yang

terkumpul dapat menggambarkan area sistem yang dianggap penting oleh pengguna sekaligus menunjukkan sejauh mana sistem memenuhi ekspektasi tersebut.

Kuesioner ini terdiri dari dua bagian utama, yaitu:

- 1) Data responden, meliputi informasi dasar seperti program studi, dan tingkat semester
- 2) *Item* pertanyaan berdasarkan *USE Questionnaire* yang telah diadaptasi sesuai kebutuhan penelitian.

Data yang terkumpul diekspor ke dalam format *spreadsheet* untuk diproses lebih lanjut pada tahap analisis menggunakan RMT dan IPA. Proses ini memastikan bahwa data yang digunakan valid, reliabel, dan siap diolah untuk menghasilkan kesimpulan yang akurat mengenai *usability* SIAKAD.

Berikut adalah tabel *item* dari *questionnaire* yang telah disesuaikan dengan keperluan penilaian SIAKAD dan *Importance Performance Analysis*.

Tabel 3.1 *Item* Kuesioner

No.	Dimensi	Pernyataan
1	<i>Usefulness</i>	Membantu Saya menjadi lebih efektif dalam mengikuti kegiatan perkuliahan.
2		SIAKAD membantu Saya menjadi lebih produktif dalam bidang akademik.
3		Selama ini SIAKAD bermanfaat bagi Saya
4		SIAKAD membantu Saya merencanakan dan memonitor aktivitas perkuliahan Saya.
5		Menggunakan SIAKAD membuat proses pencapaian akademik Saya menjadi lebih mudah.
6		Menggunakan SIAKAD membuat proses administrasi

No.	Dimensi	Pernyataan
		akademik lebih efisien.
7		SIAKAD menyediakan fitur dan fungsi yang sesuai dengan kebutuahn akademik Saya.
8		Kinerja dan fitur SIAKAD sudah sesuai dengan ekspektasi saya.
9		Saya dapat menggunakan SIAKAD dengan mudah dalam berbagai aktivitas akademik.
10		SIAKAD memiliki tampilan dan langkah penggunaan yang sederhana.
11		SIAKAD memberikan pengalaman penggunaan yang nyaman dan bersahabat.
12		SIAKAD meminimalkan jumlah langkah yang diperlukan untuk menyelesaikan pekerjaan saya.
13		Fitur-fitur SIAKAD dapat digunakan dalam berbagai situasi sesuai kebutuhan.
14	<i>Ease of Use</i>	Penggunaan SIAKAD terasa alami dan tidak melelahkan.
15		Saya dapat menggunakan SIAKAD tanpa perlu membaca petunjuk tertulis terlebih dahulu.
16		SIAKAD konsisten dalam tampilan dan cara kerjanya saat digunakan.
17		Aplikasi SIAKAD mudah digunakan oleh pengguna baru maupun yang sudah berpengalaman.
18		Saya dapat dengan mudah memperbaiki kesalahan yang saya buat saat menggunakan aplikasi SIAKAD.
19		Saya tidak pernah merasa kesulitan atau bingung saat menggunakan aplikasi SIAKAD.
20	<i>Ease of Learning</i>	Saya tidak membutuhkan waktu lama untuk memahami cara menggunakan SIAKAD.
21		Setelah mempelajari cara menggunakan SIAKAD, Saya

No.	Dimensi	Pernyataan
		bisa langsung ingat cara menggunakannya lagi.
22		Saya dapat dengan cepat memahami cara menggunakan SIAKAD pada saat pertama kali mencobanya.
23		Saya tidak membutuhkan waktu lama untuk menguasai penggunaan SIAKAD untuk kegiatan akademik Saya.
24		Saya merasa penggunaan SIAKAD ini memenuhi harapan saya.
25		Saya akan mendorong teman-teman saya untuk menggunakan SIAKAD ini.
26		Menggunakan SIAKAD memberikan pengalaman yang menyenangkan bagi Saya.
27	<i>Satisfaction</i>	Fungsi-fungsi yang ada di SIAKAD berjalan sesuai dengan yang saya butuhkan.
28		Saya sangat terkesan dengan SIAKAD ini.
29		Saya merasa SIAKAD ini sangat penting untuk kegiatan akademik saya.
30		Saya merasa senang dan nyaman ketika menggunakan SIAKAD.

3.1.3 Analisa Kuesioner

Analisis menggunakan *Rating Scale Model* (RSM) berlangsung melalui serangkaian proses estimasi yang saling berkaitan, dimulai dari pembentukan struktur parameter model hingga evaluasi kesesuaian respons melalui indeks *Infit* dan *Outfit*. Pada tahap awal, respons Likert 1 sampai 5 diolah menjadi kategori 0 sampai 4 untuk memenuhi struktur dasar RSM yang memandang kategori pertama sebagai titik awal.

Kemampuan responden direpresentasikan melalui parameter θ , sedangkan tingkat kesulitan *item* dilambangkan dengan β . Keduanya berinteraksi dengan *threshold* kategori τ , yaitu batas logit yang memisahkan satu kategori dengan kategori berikutnya. *Threshold* dianggap konsisten untuk seluruh *item* dalam RSM dan dibentuk sebagai langkah-langkah kumulatif yang meningkat secara bertahap dari kategori rendah menuju kategori tinggi. Dengan demikian, struktur model menghubungkan θ , β , dan τ dalam bentuk ekspresi logit:

Nilai logit tersebut kemudian diubah menjadi probabilitas kategori untuk setiap *item* menggunakan transformasi eksponensial sehingga seluruh kategori memiliki probabilitas yang terstandarisasi. Probabilitas inilah yang menjadi dasar untuk menghitung skor ekspektasi. Skor ekspektasi pada tiap *item* diperoleh melalui penjumlahan kategori yang diberi bobot oleh probabilitasnya.

Setelah probabilitas kategori dihitung untuk seluruh nilai θ , langkah selanjutnya adalah menentukan distribusi posterior kemampuan responden. Distribusi posterior merepresentasikan estimasi kemampuan laten responden berdasarkan pola respons yang diberikan. Perumusan posterior diperoleh dari penggabungan antara probabilitas respons dan distribusi prior θ , sebagaimana ditunjukkan pada rumus (2.6).

Distribusi posterior menghasilkan dua parameter penting, yaitu rata-rata posterior (*mean ability*) dan varians posterior. Berdasarkan nilai ekspektasi yang telah diperoleh, residual dihitung sebagai selisih antara respons aktual dan nilai prediksi. Residual tersebut menjadi dasar dalam penghitungan indeks kesesuaian model (*fit statistics*), yaitu *Infit* dan *Outfit Mean Square*, yang dirumuskan pada

persamaan (2.7) dan (2.8). Secara khusus, *Outfit Mean Square* dihitung sebagai rata-rata kuadrat residual tanpa pembobotan. Nilai-nilai tersebut dihitung untuk setiap responden dan setiap *item*, sehingga dapat memberikan gambaran komprehensif mengenai sejauh mana pola respons sesuai dengan prediksi model Rasch.

Tahap akhir analisis melibatkan pengukuran reliabilitas responden dan reliabilitas *item*. Reliabilitas responden mencerminkan stabilitas estimasi kemampuan peserta, sedangkan reliabilitas *item* menunjukkan kestabilan estimasi kesulitan *item*. Reliabilitas dihitung melalui perbandingan antara varians sejati dan total varians.

Varians sejati dihitung sebagai varians dari estimasi parameter (baik kemampuan maupun *item*), sedangkan galat diperoleh dari varians posterior untuk responden atau dari invers Hessian untuk *item*. Nilai reliabilitas yang tinggi menunjukkan bahwa parameter dapat diestimasi secara konsisten dan akurat.

Setelah nilai reliabilitas serta indeks kesesuaian *Infit* dan *Outfit* diperoleh, tahap analisis selanjutnya adalah melakukan peninjauan terhadap kualitas data responden dan *item* untuk memastikan bahwa instrumen benar-benar mengukur konstruk yang dimaksud. Nilai reliabilitas memberikan gambaran mengenai konsistensi estimasi kemampuan responden dan kesulitan *item*, sedangkan nilai *Infit* dan *m* memberikan informasi mengenai validitas internal, yaitu seberapa baik pola respons aktual selaras dengan prediksi model Rasch. Oleh karena itu, evaluasi kualitas data dilakukan dengan mengacu pada kedua indikator tersebut.

Dalam konteks responden, nilai *Infit* dan *Outfit* digunakan untuk menilai apakah pola respons yang diberikan konsisten. Responden dengan nilai *Outfit*

yang sangat tinggi, umumnya di atas 2.0, mengindikasikan adanya respons ekstrem yang tidak sesuai dengan pola prediksi model, misalnya karena pengisian respons secara acak, ketidaktelitian, atau kesalahan dalam memahami instrumen. Responden dengan *Infit* di atas 1.5 juga menunjukkan ketidakteraturan pada pola respons inti sehingga perlu dipertimbangkan untuk dikeluarkan dari analisis lanjutan. Selain itu, pola respons yang tidak bervariasi, seperti mengisi seluruh item dengan kategori yang sama, atau jumlah respons yang hilang terlalu banyak, dapat mengganggu estimasi kemampuan responden dan juga perlu dieksklusi. Sementara itu, responden dengan nilai *Infit* atau *Outfit* berada pada kisaran 1.5 hingga 2.0 tidak langsung harus dibuang; respons mereka perlu dikaji ulang untuk memastikan apakah ketidaksesuaian terjadi hanya pada satu atau dua item tertentu atau merupakan pola yang lebih luas.

3.1.4 Analisa Data

Setelah data dari *USE Questionnaire* diuji reliabilitas serta validitasnya, setiap dimensi *usability* (*usefulness*, *ease of use*, *ease of learning*, dan *satisfaction*) dianalisis berdasarkan dua aspek utama dalam IPA, yaitu tingkat kepentingan (*importance*) dan tingkat kinerja (*performance*). Dalam penelitian ini, tingkat kepentingan mencerminkan sejauh mana pengguna menganggap suatu aspek *usability* penting, sementara tingkat kinerja menunjukkan bagaimana mereka menilai performa aktual sistem dalam aspek tersebut.

Setiap responden diminta untuk memberikan penilaian terhadap setiap pernyataan dalam *USE Questionnaire* berdasarkan dua perspektif:

- 1) Kepentingan (*Importance*): Seberapa penting aspek *usability* tertentu bagi pengguna.
- 2) Kinerja (*Performance*): Seberapa baik sistem memenuhi harapan pengguna terkait aspek tersebut.

Kedua aspek tersebut diukur menggunakan skala Likert lima tingkat, di mana nilai rata-rata pada masing-masing dimensi dihitung menggunakan rumus sebagaimana tercantum pada persamaan (2.9) dan (2.10).

Setelah skor kepentingan dan kinerja dihitung, hasilnya dipetakan ke dalam diagram IPA, yang terdiri dari empat kuadran:

- 1) Kuadran I (*Keep Up the Good Work*) – Aspek dengan kepentingan dan kinerja tinggi, yang berarti sistem telah memenuhi harapan pengguna.
- 2) Kuadran II (*Concentrate Here*) – Aspek dengan kepentingan tinggi tetapi kinerja rendah, yang menunjukkan area kritis yang memerlukan perbaikan segera.
- 3) Kuadran III (*Low Priority*) – Aspek dengan kepentingan dan kinerja rendah, yang berarti aspek ini tidak terlalu berpengaruh terhadap kepuasan pengguna.
- 4) Kuadran IV (*Possible Overkill*) – Aspek dengan kepentingan rendah tetapi kinerja tinggi, yang mengindikasikan bahwa sumber daya yang dialokasikan mungkin berlebihan.

Pemetaan ini dilakukan dengan menghitung nilai rata-rata keseluruhan untuk kepentingan (\bar{X}) dan kinerja (\bar{Y}) guna menentukan titik tengah kuadran.

Setiap dimensi *usability* kemudian dipetakan berdasarkan nilai kepentingan (\bar{X}) dan kinerja (\bar{Y}) ke dalam diagram.

Setelah pemetaan dilakukan, interpretasi hasil analisis IPA bertujuan untuk menentukan rekomendasi perbaikan pada SIAKAD. Jika suatu aspek *usability* berada dalam Kuadran II (*Concentrate Here*), maka aspek tersebut menjadi prioritas utama untuk perbaikan guna meningkatkan kepuasan pengguna. Sebaliknya, aspek yang berada dalam Kuadran IV (*Possible Overkill*) dapat dipertimbangkan untuk pengurangan sumber daya jika diperlukan.

3.1.5 Hasil Evaluasi dan Rekomendasi

Data hasil analisis dapat dimanfaatkan sebagai dasar untuk menyusun evaluasi *usability* menggunakan metode *USE Questionnaire* dan *Importance Performance Analysis*, yang kemudian divalidasi melalui pendekatan *Rasch Measurement Theory*. Penyampaian hasil ini tidak hanya ditujukan untuk memberikan gambaran umum mengenai tingkat kegunaan sistem, tetapi juga mengidentifikasi aspek-aspek spesifik yang memerlukan perbaikan serta rekomendasi yang dapat digunakan oleh pengembang sistem dan pemangku kepentingan dalam meningkatkan kualitas layanan akademik berbasis web.

Hasil evaluasi pertama yang dapat dilaporkan adalah evaluasi tingkat *usability* berdasarkan *USE Questionnaire*, yang terdiri dari empat dimensi utama, yaitu *Usefulness*, *Ease of Use*, *Ease of Learning*, dan *Satisfaction*. Metrik yang diperoleh dari angket ini akan diolah untuk menunjukkan sejauh mana pengguna merasakan manfaat sistem, kemudahan dalam penggunaannya, serta tingkat kepuasan mereka. Selain itu, dilakukan analisis lebih lanjut untuk melihat

bagaimana persepsi pengguna terhadap sistem dalam konteks penggunaannya dalam kehidupan akademik sehari-hari. Dengan demikian, penelitian ini diharapkan dapat menyajikan gambaran kuantitatif mengenai *usability* sistem yang dapat dijadikan acuan dalam pengambilan keputusan strategis oleh pihak universitas.

Hasil evaluasi IPA memungkinkan pemetaan hubungan antara tingkat kepentingan dan tingkat kinerja yang dirasakan oleh pengguna terhadap berbagai fitur dalam SIAKAD. Hasil evaluasi ini dapat disajikan dalam bentuk kuadran yang membagi aspek *usability* menjadi empat kategori utama, yaitu *Keep Up the Good Work*, *Concentrate Here*, *Low Priority*, dan *Possible Overkill*. Penyajian ini bertujuan untuk mengidentifikasi area yang harus segera diperbaiki karena memiliki kepentingan tinggi tetapi kinerjanya masih rendah, serta area yang sudah memiliki kinerja baik dan perlu dipertahankan. Dengan adanya laporan ini, pihak pengembang sistem dapat memahami prioritas dalam peningkatan fitur dan pengalaman pengguna secara lebih terarah.

Selain kedua evaluasi tersebut, hasil validasi RMT dapat digunakan untuk mengukur konsistensi dan reliabilitas data yang diperoleh dari kuesioner, serta memastikan bahwa instrumen penelitian yang digunakan memiliki validitas yang kuat. Melalui pendekatan ini, penelitian dapat mengidentifikasi apakah terdapat bias dalam respons pengguna, sejauh mana pertanyaan dalam kuesioner mampu membedakan tingkat persepsi *usability* antar pengguna, serta bagaimana distribusi tingkat kemampuan pengguna dalam menilai sistem. Hasil evaluasi ini akan

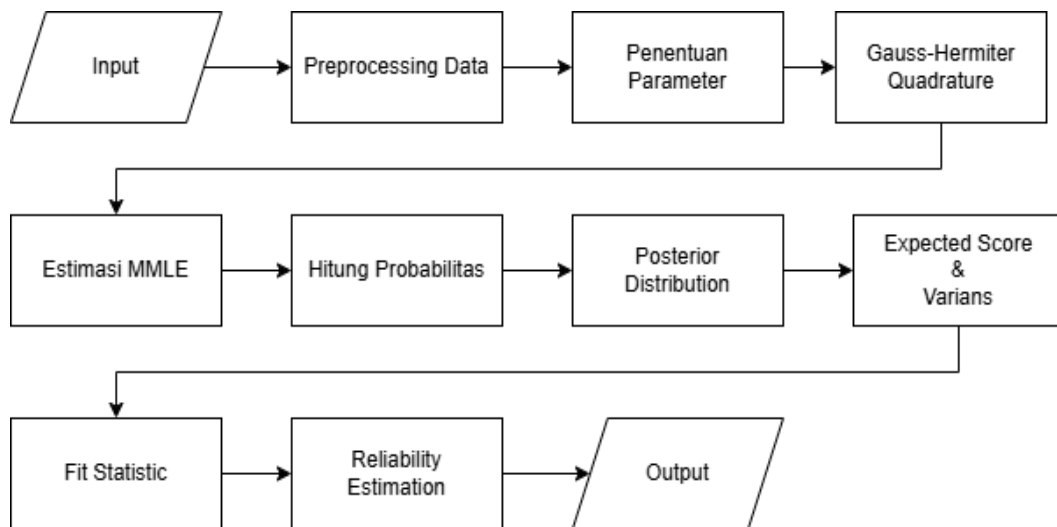
memberikan wawasan lebih dalam terkait validitas instrumen evaluasi *usability* serta memperkuat argumentasi ilmiah mengenai kesimpulan yang diperoleh.

Lebih lanjut, visualisasi data dalam bentuk grafik, tabel, serta diagram kuadran IPA dapat membantu dalam memudahkan pemahaman terhadap data yang telah diolah. Selain itu, laporan juga akan mencakup interpretasi temuan yang mengaitkan hasil analisis dengan studi-studi terdahulu dalam bidang *Software Usability Testing*, sehingga memberikan konteks yang lebih luas mengenai implikasi dari hasil penelitian ini.

Pada akhirnya, hasil evaluasi ini tidak hanya ditujukan untuk memberikan evaluasi deskriptif terhadap *usability* SIAKAD yang diteliti, tetapi juga menyajikan rekomendasi strategis bagi pengembang sistem, administrator universitas, dan pemangku kepentingan lainnya. Dengan adanya laporan ini, diharapkan langkah-langkah perbaikan yang berbasis data dapat diimplementasikan untuk meningkatkan pengalaman pengguna dan efisiensi dalam pengelolaan informasi akademik secara digital.

3.2 Implementasi Sistem

Agar penelitian dapat berjalan secara sistematis, terstruktur, dan sesuai dengan tujuan yang telah ditetapkan, maka disusun suatu desain sistem yang digunakan sebagai acuan dalam proses pengolahan dan analisis data. Desain sistem ini bertujuan untuk menggambarkan alur kerja penelitian mulai dari tahap pengumpulan data kuesioner, proses pengolahan data menggunakan pendekatan *Rasch Measurement Theory* (RMT), hingga analisis hasil menggunakan metode *Importance Performance Analysis* (IPA).



Gambar 3.3 Bagan Sistem

Desain penelitian yang ditunjukkan pada Gambar 3.3 merepresentasikan keseluruhan tahapan proses yang dilakukan dalam penelitian ini. Setiap tahapan dalam desain tersebut dirancang untuk memastikan bahwa data yang diperoleh dalam desain tersebut dirancang untuk memastikan bahwa data yang diperoleh dapat diolah secara optimal serta menghasilkan informasi yang valid dan reliabel dalam mengevaluasi usability sistem.

Selanjutnya, desain penelitian pada Gambar 3.3 diimplementasikan ke dalam bahasa pemrograman Python sebagai alat bantu dalam proses komputasi dan analisis data. Implementasi ini mencakup proses input data kuesioner, estimasi parameter menggunakan model *Rating Scale Model* (RSM), perhitungan statistik kelayakan model, serta visualisasi hasil analisis dalam bentuk grafik dan peta IPA.

3.2.1 Input

Tahap input data merupakan proses awal dalam sistem estimasi RSM, yang bertujuan untuk membaca, memvalidasi, dan menyiapkan data kuesioner sebelum

dilakukan proses estimasi parameter. Data yang digunakan berbentuk *file* CSV yang berisi respons responden terhadap *item* dengan skala Likert 1–5.

Proses dimulai dengan pemanggilan antarmuka pemilihan *file* menggunakan pustaka *tkinter*, sehingga pengguna dapat memilih *file* CSV secara interaktif. Baris kode berikut digunakan untuk membuka dialog pemilihan *file*:

Pseudocode 3.1 Input File CSV Menggunakan tkinter

```

root = tk.Tk()
root.withdraw()
csv_path = filedialog.askopenfilename(
    title="Pilih file CSV (data kuesioner)",
    filetypes=[("CSV", "*.csv")]
)

```

Kode tersebut membuat jendela tersembunyi (`root.withdraw()`) dan memunculkan dialog pemilihan *file*. Apabila pengguna tidak memilih *file*, proses dihentikan dengan:

Pseudocode 3.2 Validasi File CSV

```

if not csv_path:
    print("No file selected. Exiting.")
    return

```

Setelah *file* dipilih, data dibaca menggunakan pustaka *pandas*. Data dibaca menggunakan fungsi:

Pseudocode 3.3 Input Data File

```

df_raw = pd.read_csv(csv_path)

```

Pada tahap ini seluruh isi *file* CSV dimuat ke dalam objek *DataFrame*. Data biasanya terdiri atas kolom metadata (misalnya nama responden, NIM, jurusan) serta kolom *item* kuesioner. Oleh karena itu, dilakukan pemisahan antara kolom metadata dan kolom *item* agar analisis Rasch hanya menggunakan respons *item*.

Pseudocode 3.4 Seleksi Kolom Item

```

meta_cols = ["Timestamp", "Nama", "NIM", "Jurusan",
"Angkatan"]
item_cols = [c for c in df_raw.columns if c not in meta_cols]
df_items = df_raw[item_cols].copy()

```

Baris kode tersebut mengidentifikasi kolom selain metadata sebagai item penelitian, kemudian menyimpannya dalam variabel `df_items`. Selanjutnya, seluruh data *item* dikonversi menjadi numerik untuk memastikan dapat diproses secara matematis dalam model Rasch:

Pseudocode 3.5 Konversi Data Numerik

```

df_items = df_items.apply(pd.to_numeric, errors='coerce')

```

Parameter `errors='coerce'` berfungsi mengubah nilai yang tidak dapat dikonversi menjadi angka menjadi `NaN`. Nilai `NaN` ini nantinya diperlakukan sebagai *missing value* dan tidak diikuti dalam perhitungan *likelihood*.

Sebagai langkah validasi awal, sistem mendeteksi kategori maksimum dalam data untuk memastikan skala yang digunakan sesuai dengan skala Likert 1–

5. Validasi dilakukan melalui fungsi:

Pseudocode 3.6 Deteksi Skala Likert

```

max_cat = int(np.nanmax(df_items.values))
print("Detected max category (should be 5 for Likert 1-5):",
max_cat)

```

Jika kategori maksimum tidak sama dengan 5, sistem akan menampilkan pesan konfirmasi kepada pengguna. Proses ini dilakukan melalui fungsi:

Pseudocode 3.7 Konfirmasi Skala Likert

```

if max_cat != 5:
    msg = ("Terlihat data tidak memiliki kategori 1..5.\n"
          f"Detected max category = {max_cat}.\n"
          "Jika Anda yakin data 1..5, periksa CSV. Lanjutkan? (OK
= lanjut, Cancel = stop)")
    if not messagebox.askokcancel("Konfirmasi kategori", msg):
        return

```

3.2.2 Processing Data

Tahap *preprocessing* data dilakukan untuk menyiapkan struktur data agar sesuai dengan kebutuhan komputasi Model Rasch *Rating Scale Model* (RSM). Pada tahap ini, data yang telah dibaca dalam bentuk DataFrame dikonversi ke dalam format numerik matriks, ditentukan dimensi dasarnya, serta disesuaikan dengan format kategori yang digunakan dalam formulasi matematis model.

Langkah awal dalam *preprocessing* adalah mengubah objek DataFrame menjadi array numerik bertipe *float* agar dapat diproses menggunakan operasi matriks NumPy. Proses ini dilakukan melalui baris kode berikut:

Pseudocode 3.8 Konversi Data Ke Array

```
data_arr = data_df.values.astype(float)
```

Variabel *data_arr* merepresentasikan matriks respons dengan ukuran $n \times I$, di mana n adalah jumlah responden dan I adalah jumlah *item*. Selanjutnya, sistem mengidentifikasi jumlah responden dan jumlah *item* melalui:

Pseudocode 3.9 Dimensi Data

```
n person, n item = data_arr.shape
```

Baris kode tersebut mengekstrak dimensi matriks untuk digunakan dalam proses estimasi parameter. Setelah itu, dilakukan tahap berikutnya yakni mendeteksi kategori respons maksimum dalam data. Karena instrumen menggunakan skala Likert 1–5, sistem memastikan bahwa kategori tertinggi adalah 5. Deteksi dilakukan dengan:

Pseudocode 3.10 Deteksi Kategori Respon

```
max_cat = int(np.nanmax(data_arr))
```

Fungsi `np.nanmax()` digunakan agar nilai missing (`NaN`) tidak memengaruhi proses identifikasi kategori maksimum. Dalam formulasi matematis Model Rasch RSM, kategori respons harus dimulai dari 0 hingga m , dengan $m = C - 1$, dimana C adalah jumlah kategori. Oleh karena itu dilakukan transformasi skor dari 1–5 menjadi 0–4 melalui:

Pseudocode 3.11 Transformasi Skor

	<pre>m = max_cat - 1 data_arr0 = data_arr - 1</pre>
--	---

Transformasi ini penting karena model probabilitas kategori dalam RSM menggunakan indeks kategori mulai dari nol. Dengan konversi tersebut, respons 1 menjadi 0, respons 2 menjadi 1, dan seterusnya hingga respons 5 menjadi 4. Nilai missing tetap dipertahankan sebagai `NaN` sehingga tidak ikut dihitung dalam proses *likelihood*.

Untuk efisiensi perhitungan pada tahap estimasi, sistem kemudian membangun struktur data yang hanya memuat respons yang terobservasi (*non-missing*). Hal ini dilakukan melalui fungsi `build_observed()` yang dijalankan dengan:

Pseudocode 3.12 Data Respons Terobservasi

	<pre>data_obs = build_observed(data_arr0)</pre>
--	---

Fungsi ini menghasilkan tiga vektor indeks, yaitu indeks responden (`p_idx`), indeks *item* (`i_idx`), dan nilai respons (`r_list`). Struktur ini mempercepat proses komputasi *likelihood* karena sistem tidak perlu melakukan iterasi pada seluruh sel matriks, melainkan hanya pada respons yang benar-benar tersedia.

3.2.3 Penentuan Parameter

Tahap penentuan parameter merupakan proses inialisasi nilai awal parameter yang akan diestimasi dalam Model Rasch *Rating Scale Model* (RSM). Pada tahap ini ditentukan parameter kesulitan *item* (β), parameter *threshold* kategori (τ), serta parameter bantu (γ) yang digunakan untuk menjamin sifat monotonik *threshold*. Inialisasi parameter awal diperlukan karena proses estimasi dilakukan menggunakan algoritma optimasi numerik yang membutuhkan titik awal (*starting values*). Setelah tahap preprocessing menghasilkan matriks respons `data_arr0`, sistem telah mengetahui jumlah responden (`n_person`), jumlah *item* (`n_item`), serta jumlah kategori langkah (`step`) yang dinyatakan sebagai:

Pseudocode 3.13 Jumlah Step Rating Scale

	<code>m = max_cat - 1</code>
--	------------------------------

Nilai m merepresentasikan jumlah *threshold* dalam *Rating Scale Model*, dengan kategori respons berada pada rentang 0 sampai m . Parameter pertama yang ditentukan adalah parameter kesulitan *item* (β). Pada tahap awal, seluruh nilai β diinisialisasi dengan nol menggunakan:

Pseudocode 3.14 Inialisasi Awal Nilai β

	<code>beta0 = np.zeros(n_item)</code>
--	---------------------------------------

Inialisasi nol berarti pada iterasi pertama seluruh *item* diasumsikan memiliki tingkat kesulitan yang sama. Nilai ini kemudian akan diperbarui selama proses optimasi hingga mencapai nilai yang memaksimalkan *likelihood*.

Parameter berikutnya adalah parameter γ , yang merupakan parameter bantu untuk membentuk *threshold* kategori (τ). Dalam model ini, *threshold* tidak diestimasi secara langsung, melainkan dibangun dari parameter γ melalui fungsi eksponensial agar bersifat positif dan meningkat secara monoton. Inisialisasi γ dilakukan dengan:

Pseudocode 3.15 Inisialisasi Awal Nilai γ

```
gamma0 = np.log(np.linspace(1.0, 1.0, m))
```

Karena `np.linspace(1.0, 1.0, m)` menghasilkan vektor berisi angka 1 sebanyak m elemen, maka `np.log(1)` menghasilkan nilai nol. Dengan demikian, seluruh γ pada awal estimasi bernilai nol yang selanjutnya digabungkan seluruh parameter awal ke dalam satu vektor parameter yang akan dioptimasi.

Pseudocode 3.16 Vektor Parameter

```
x0 = np.concatenate([beta0, gamma0])
```

Vektor x_0 berisi seluruh parameter yang akan diestimasi, dengan susunan:

- 1) Elemen pertama hingga ke- N Item adalah parameter β (*difficulty item*)
- 2) Elemen berikutnya hingga akhir adalah parameter γ (parameter pembentuk *threshold*).

Selama proses optimasi, parameter γ akan diubah menjadi *threshold* τ melalui transformasi:

Pseudocode 3.17 Transformasi Parameter γ ke *Threshold* τ

```
step = np.exp(gamma)
tau_raw = np.cumsum(step)
tau = tau_raw - np.mean(tau_raw)
```

Transformasi ini memiliki dua tujuan utama. Pertama, penggunaan fungsi eksponensial menjamin bahwa setiap langkah *threshold* bernilai positif sehingga

urutan kategori meningkat secara logis. Kedua, proses pengurangan rata-rata (`np.mean(tau_raw)`) dilakukan untuk memenuhi syarat identifikasi model, yaitu agar rata-rata *threshold* bernilai nol.

3.2.4 Gauss-Hermite Quadrature

Tahap *Gauss-Hermite Quadrature* merupakan bagian penting dalam proses estimasi parameter menggunakan pendekatan *Marginal Maximum Likelihood Estimation* (MMLE). Pada pendekatan ini, parameter kemampuan responden (θ) tidak diestimasi secara langsung pada tahap awal, melainkan diperlakukan sebagai variabel acak yang mengikuti distribusi normal standar. Oleh karena itu, diperlukan proses integrasi terhadap distribusi θ untuk menghitung marginal *likelihood*. Karena integral tersebut tidak dapat diselesaikan secara analitik, maka digunakan pendekatan numerik melalui metode *Gauss-Hermite Quadrature*.

Secara umum, *Gauss-Hermite Quadrature* digunakan untuk mendekati integral berbobot fungsi eksponensial berbentuk distribusi normal. Metode ini menghasilkan sejumlah titik integrasi (*nodes*) dan bobot (*weights*) yang digunakan untuk mendekati nilai integral secara diskrit.

Pseudocode 3.18 Perhitungan Nodes dan Weights

<code>nodes_raw, weights_raw = hermgauss(quadature_points)</code>

Parameter `quadature_points` menentukan jumlah titik integrasi yang digunakan, misalnya 41 titik. Semakin banyak titik yang digunakan, semakin tinggi tingkat akurasi pendekatan integral, meskipun waktu komputasi juga meningkat.

Fungsi `hermgauss()` secara default menghasilkan titik dan bobot untuk integral dengan bentuk fungsi berbobot e^{-x^2} . Namun, dalam model Rasch, distribusi kemampuan diasumsikan mengikuti distribusi normal standar. Oleh karena itu, dilakukan transformasi agar sesuai dengan distribusi normal standar. Transformasi dilakukan dengan:

Pseudocode 3.19 Transformasi Nodes dan Weights ke Normal Standar

```
theta_nodes = np.sqrt(2) * nodes_raw
theta_weights = weights_raw / np.sqrt(np.pi)
```

Perkalian dengan $\sqrt{2}$ mengubah titik integrasi agar sesuai dengan variansi distribusi normal standar. Pembagian bobot dengan $\sqrt{\pi}$ menyesuaikan konstanta normalisasi distribusi. Setelah itu, bobot dinormalisasi agar jumlah totalnya sama dengan satu:

Pseudocode 3.20 Normalisasi Bobot

```
theta_weights = theta_weights / np.sum(theta_weights)
```

Normalisasi ini penting agar bobot dapat diinterpretasikan sebagai proporsi kontribusi masing-masing titik θ dalam proses integrasi. Hasil dari tahap ini adalah:

- 1) `theta_nodes` \rightarrow sekumpulan titik kemampuan hipotetik $(\theta_1, \theta_2, \dots, \theta_q)$,
- 2) `theta_weights` \rightarrow bobot masing-masing titik integrasi (w_1, w_2, \dots, w_q) .

Dalam perhitungan *marginal likelihood* pada metode MMLE, diperlukan integrasi terhadap distribusi kemampuan laten sebagaimana dinyatakan pada persamaan (2.3). Karena integral tersebut sulit diselesaikan secara analitik, maka digunakan pendekatan numerik berupa penjumlahan diskrit dengan metode

kuadratur Gauss, sebagaimana ditunjukkan pada persamaan (2.4), untuk mempermudah proses estimasi parameter model.

Dengan demikian, tahap *Gauss-Hermite Quadrature* memungkinkan proses integrasi distribusi kemampuan dilakukan secara numerik sehingga estimasi parameter item dapat dilakukan melalui MMLE. Tahap ini berperan krusial dalam menjembatani asumsi distribusi kemampuan normal dengan proses optimasi *likelihood* pada model Rasch RSM.

3.2.5 Estimasi MMLE

Tahap estimasi parameter menggunakan metode *Marginal Maximum Likelihood Estimation* (MMLE) merupakan inti dari proses kalibrasi model Rasch RSM. Pada tahap ini dilakukan pencarian nilai parameter kesulitan *item* (β) dan *threshold* kategori (τ) yang memaksimalkan fungsi *likelihood* marginal. Berbeda dengan pendekatan *Joint Maximum Likelihood* (JMLE), metode MMLE tidak mengestimasi kemampuan responden (θ) secara langsung pada tahap ini, melainkan mengintegrasikan distribusi θ menggunakan *Gauss-Hermite Quadrature*. Dalam program, fungsi *likelihood* negatif didefinisikan dalam fungsi:

Pseudocode 3.21 Perhitungan Likelihood Negatif

```
def neg_marginal_loglik(params, data_obs, n_person, n_item, m,
theta_nodes, theta_weights):
```

Parameter *params* berisi gabungan parameter β dan γ . Di dalam fungsi tersebut, parameter dipisahkan melalui fungsi:

Pseudocode 3.22 Ekstraksi Parameter β dan γ

```
beta = params[:n_item]
gamma = params[n_item:]
```

Parameter γ kemudian ditransformasikan menjadi *threshold* (τ) agar bersifat meningkat dan teridentifikasi dengan baik. Proses ini dilakukan melalui fungsi:

Pseudocode 3.23 Transformasi Parameter γ Menjadi Threshold τ

```
step = np.exp(gamma)
tau_raw = np.cumsum(step)
tau = tau_raw - np.mean(tau_raw)
```

Transformasi eksponensial memastikan setiap langkah *threshold* bernilai positif, sedangkan pengurangan rata-rata menjaga identifikasi model (rata-rata *threshold* sama dengan nol). Selanjutnya, sistem menghitung probabilitas kategori untuk setiap kombinasi titik kemampuan dan *item*. Perhitungan ini dilakukan melalui:

Pseudocode 3.24 Perhitungan Probabilitas Kategori

```
delta_mat = theta_nodes[:, None] - beta[None, :]
probs_mat[q] = rsm_category_probs(delta_mat[q], tau)
```

Nilai `delta_mat` merepresentasikan selisih antara kemampuan hipotetik (θ_q) dan kesulitan *item* (β_i). Fungsi `rsm_category_probs()` kemudian menghitung probabilitas setiap kategori respons berdasarkan formulasi RSM. Untuk setiap responden, sistem menghitung probabilitas pola respons dengan mengalikan probabilitas tiap *item* pada setiap titik θ_q , kemudian menjumlahkannya dengan bobot *quadrature*:

Pseudocode 3.25 Perhitungan Likelihood Marginal

```
prod_q *= probs_mat[:, item, r]
marg = np.dot(theta_weights, prod_q)
total_loglik += np.log(marg)
```

Nilai *log-likelihood* dijumlahkan untuk seluruh responden. Karena proses optimasi umumnya meminimalkan fungsi, maka fungsi yang digunakan adalah negatif *log-likelihood*:

Pseudocode 3.26 Transformasi *Log-Likelihood* menjadi Negatif

	<code>return -total_loglik</code>
--	-----------------------------------

Setelah *Log-Likelihood* ditransformasikan menjadi negatif, dilakukan optimasi parameter. Optimasi parameter dilakukan menggunakan algoritma L-BFGS-B dari pustaka SciPy:

Pseudocode 3.27 Optimasi Parameter dengan L-BFGS-B

	<pre> res = minimize(neg_marginal_loglik, x0, args=(data_obs, n_person, n_item, m, theta_nodes, theta_weights), method='L-BFGS-B', options={'maxiter': maxiter, 'disp': False}) </pre>
--	--

Algoritma ini bekerja secara iteratif dengan memperbarui parameter β dan γ hingga mencapai nilai yang meminimalkan negatif *log-likelihood*. Proses berhenti ketika konvergensi tercapai atau jumlah iterasi maksimum terpenuhi. Setelah optimasi selesai, parameter optimal diekstrak kembali.

Pseudocode 3.28 Ekstraksi Parameter Optimal

	<pre> xopt = res.x beta_opt = xopt[:n_item] gamma_opt = xopt[n_item:] </pre>
--	--

3.2.6 Hitung Probabilitas

Tahap perhitungan probabilitas merupakan proses untuk menentukan peluang setiap kategori respons berdasarkan model Rasch RSM. Probabilitas ini menjadi dasar dalam menghitung *likelihood*, *expected score*, residual, serta

statistik kesesuaian (*fit statistics*). Pada model RSM, probabilitas kategori tidak hanya bergantung pada kemampuan responden (θ) dan kesulitan *item* (β), tetapi juga pada struktur *threshold* kategori (τ) yang sama untuk seluruh *item*. Dalam implementasi program, perhitungan probabilitas kategori didefinisikan dalam fungsi:

Pseudocode 3.29 Definisi Probabilitas Kategori

```
def rsm_category_probs(delta, tau):
```

Sedangkan parameter `delta` merepresentasikan selisih antara kemampuan dan kesulitan *item*. Parameter ini didapat melalui fungsi:

Pseudocode 3.30 Perhitungan Selisih Kemampuan dan Kesulitan Item

```
delta_mat = theta_nodes[:, None] - beta[None, :]
```

Nilai `delta` tersebut dihitung untuk setiap kombinasi titik kemampuan (θ_q) dan *item* (β_i). Langkah selanjutnya dalam fungsi adalah membentuk akumulasi *threshold* kategori:

Pseudocode 3.31 Pembentukan Akumulasi Threshold

```
s = np.concatenate(([0.0], np.cumsum(tau)))
```

Selanjutnya dihitung nilai logit untuk setiap kategori. Nilai logit didapat melalui fungsi:

Pseudocode 3.32 Perhitungan Nilai Logit Kategori

```
logits = np.stack([k * delta - s[k] for k in range(m+1)], axis=-1)
```

Ekspresi $k * \text{delta} - s[k]$ merepresentasikan bagian eksponen dalam rumus probabilitas RSM. Untuk menjaga stabilitas numerik dan mencegah *overflow* pada fungsi eksponensial, dilakukan normalisasi dengan mengurangi nilai maksimum logit:

Pseudocode 3.33 Normalisasi Nilai Logit

<pre>maxlog = np.max(logits, axis=-1, keepdims=True) ex = np.exp(logits - maxlog)</pre>

Langkah ini memastikan nilai eksponensial tetap dalam rentang komputasi yang aman tanpa mengubah hasil akhir probabilitas. Tahap terakhir adalah menghitung probabilitas dengan membagi setiap nilai eksponensial dengan jumlah totalnya:

Pseudocode 3.34 Perhitungan Probabilitas Kategori per θ dan Item

<pre>probs = ex / np.sum(ex, axis=-1, keepdims=True)</pre>
--

Hasil akhir `probs` adalah matriks probabilitas untuk setiap kategori (0 sampai m) pada setiap kombinasi θ dan *item*. Hasil akhir ini nantinya akan digunakan untuk menghitung *Posterior Distribution*.

3.2.7 Posterior Distribution

Tahap posterior distribution dilakukan setelah parameter *item* (β) dan *threshold* kategori (τ) berhasil diestimasi melalui MMLE. Pada tahap ini, kemampuan responden (θ) tidak lagi diperlakukan sebagai variabel laten yang diintegrasikan, melainkan dihitung distribusi posteriornya berdasarkan pola respons aktual yang diberikan responden. Estimasi kemampuan dilakukan menggunakan pendekatan *Expected A Posteriori* (EAP), yaitu rata-rata tertimbang dari distribusi posterior.

Secara matematis, distribusi posterior kemampuan responden ke- n dirumuskan sebagaimana dinyatakan pada persamaan (2.6). Dalam program, tahap ini diimplementasikan dalam fungsi:

Pseudocode 3.35 Perhitungan Posterior Kemampuan

```
def compute_person_posteriors_and_stats(beta, tau, data_arr,
theta_nodes, theta_weights):
```

Setelah posterior kemampuan didapat, selanjutnya adalah menghitung kembali probabilitas kategori untuk setiap kombinasi θ dan *item*. Hal ini dapat dilakukan menggunakan parameter hasil estimasi dengan fungsi:

Pseudocode 3.36 Perhitungan Ulang Probabilitas Kategori

```
delta_mat = theta_nodes[:, None] - beta[None, :]
probs_mat[q] = rsm_category_probs(delta_mat[q], tau)
```

Kemudian, untuk setiap responden, dihitung *likelihood* pola respons pada setiap titik kemampuan. Nilai *Likelihood* dapat dihitung melalui fungsi:

Pseudocode 3.37 Perhitungan Likelihood Pola Respons per θ

```
prod_q = np.ones(Q)
for i in np.where(mask)[0]:
    r = int(data_arr[p, i])
    prod_q *= probs_mat[:, i, r]
```

Variabel `prod_q` merepresentasikan hasil perkalian probabilitas respons seluruh *item* pada masing-masing titik θ_q . Selanjutnya dihitung distribusi posterior tak ternormalisasi dengan mengalikan bobot *quadrature*:

Pseudocode 3.38 Perhitungan Distribusi Posterior

```
unnorm = theta_weights * prod_q
denom = np.sum(unnorm)
posterior = unnorm / denom
```

Variabel `posterior` berisi probabilitas posterior untuk setiap titik kemampuan θ_q . Kemampuan responden kemudian dihitung sebagai nilai harapan (*mean posterior*):

Pseudocode 3.39 Estimasi Kemampuan Responden

```
mean_p = np.sum(posterior * theta_nodes)
```

Sementara itu, varians posterior dikalkulasi guna menentukan nilai varians skor harapan (*expected score varians*). Perhitungan varians tersebut dilakukan dengan menerapkan fungsi berikut:

Pseudocode 3.40 Perhitungan Varians Posterior

```
var_p = np.sum(posterior * (theta_nodes - mean_p)**2)
```

Nilai *mean posterior* dan varians ini kemudian disimpan kedalam memori untuk digunakan pada tahap selanjutnya. Proses ini dapat dilakukan dengan menggunakan fungsi:

Pseudocode 3.41 Penyimpanan Mean dan Varians Posterior

```
person_mean[p] = mean_p
person_var[p] = var_p
```

Selain menghitung kemampuan, pada tahap ini juga dihitung nilai *expected score* dan varians respons untuk setiap *item* berdasarkan distribusi posterior. Proses ini dapat dilakukan dengan fungsi:

Pseudocode 3.42 Perhitungan Expected Score Item

```
exp_q = np.sum(probs_mat[:, i, :] * cats[None, :], axis=1)
expected[p, i] = np.sum(posterior * exp_q)
```

3.2.8 *Expected Score Varians*

Tahap perhitungan *Expected Score* dan varians dilakukan setelah distribusi posterior kemampuan responden diperoleh. Tujuan dari tahap ini adalah menghitung nilai harapan skor respons dan varians respons pada setiap kombinasi responden–*item* berdasarkan Model Rasch RSM. Nilai ini menjadi dasar dalam perhitungan residual serta statistik kesesuaian model seperti *Infit* dan *Outfit*.

Karena kemampuan responden diestimasi menggunakan distribusi posterior, maka nilai harapan tersebut diintegrasikan terhadap distribusi posterior kemampuan. Dalam program, perhitungan ini dilakukan di dalam fungsi:

Pseudocode 3.43 Proses Perhitungan Posterior dan Statistik Responden

```
compute_person_posteriors_and_stats(...)
```

Hal pertama yang harus dilakukan adalah menentukan daftar kategori respons. Kategori respons ditentukan dengan fungsi:

Pseudocode 3.44 Penentuan Kategori Respons

```
cats = np.arange(len(tau) + 1)
```

Variabel `cats` berisi nilai kategori 0 sampai m . Untuk setiap titik kemampuan θ_q , dihitung nilai harapan kategori dengan fungsi:

Pseudocode 3.45 Perhitungan Nilai Harapan Kategori per θ

```
exp_q = np.sum(probs_mat[:, i, :] * cats[None, :], axis=1)
```

Variabel `exp_q` merepresentasikan nilai harapan skor pada *item* ke- i untuk setiap titik θ_q . Nilai *expected score* akhir dihitung dengan merata-ratakan `exp_q` menggunakan distribusi posterior sebagai bobot:

Pseudocode 3.46 Perhitungan *Expected Score* Berbobot Posterior

```
expected[p, i] = np.sum(posterior * exp_q)
```

Setelah *expected score* dihitung, dicari varians kategori respons. Varians respons dihitung melalui fungsi:

Pseudocode 3.47 Perhitungan Varians Kategori per θ

```
var_q = np.sum(probs_mat[:, i, :] * (cats[None, :] - exp_q[:, None])**2, axis=1)
```

Karena kemampuan responden tidak berupa satu nilai tetap melainkan distribusi posterior, maka varians total dihitung dengan menggabungkan varians dalam-kemampuan (*within-theta variance*) dan varians antar-kemampuan (*between-theta variance*):

Pseudocode 3.48 Perhitungan Varians Total Respons

```

variance[p, i] = (
    np.sum(posterior * var_q)
    + np.sum(posterior * (exp_q - np.sum(posterior * exp_q))**2)
)

```

Bagian pertama menghitung rata-rata varians kategori pada setiap θ_q , sedangkan bagian kedua mengakomodasi variasi akibat distribusi posterior kemampuan. Hasil akhir dari tahap ini adalah dua matriks:

`expected[p, i]` → nilai harapan skor respons,

`variance[p, i]` → varians respons.

Nilai *expected score* menggambarkan skor yang diprediksi oleh model untuk setiap responden pada setiap *item*, sedangkan varians menunjukkan tingkat ketidakpastian prediksi tersebut. Kedua komponen ini digunakan dalam tahap berikutnya untuk menghitung residual dan statistik *Infit* serta *Outfit*, yang menjadi indikator kesesuaian data dengan Model Rasch. Dengan demikian, tahap *Expected Score* dan Varians berperan penting dalam evaluasi validitas internal model pengukuran.

3.2.9 *Fit Statistic*

Tahap perhitungan statistik kesesuaian dilakukan untuk mengevaluasi sejauh mana data empiris sesuai dengan prediksi model Rasch RSM. Evaluasi ini dilakukan dengan membandingkan skor respons aktual dengan skor yang

diprediksi oleh model melalui nilai residual. Residual menggambarkan selisih antara respons yang diberikan responden dan nilai harapan yang dihitung berdasarkan parameter model.

Dalam program, residual dihitung setelah matriks *expected score* dan varians diperoleh. Untuk setiap responden, langkah pertama adalah memilih *item* yang memiliki respons (tidak *missing*):

Pseudocode 3.49 Seleksi Data Respons Valid

```
mask = ~np.isnan(data_arr0[p])
```

Setelah respons valid didapat, barulah nilai residual bisa dihitung. Nilai residual kemudian dihitung menggunakan:

Pseudocode 3.50 Perhitungan Residual

```
resid = (data_arr0[p, mask] - expected[p, mask])
```

Nilai residual kuadrat digunakan untuk mengukur besarnya penyimpangan terhadap model. Statistik pertama dihitung sebagai rata-rata sederhana dari residual kuadrat:

Pseudocode 3.51 Perhitungan *Outfit* Responden

```
outfit_person[p] = np.mean(resid**2)
```

Perhitungan ini memberikan ukuran sensitivitas terhadap respons ekstrem, karena tidak mempertimbangkan bobot varians. Nilai yang jauh lebih besar dari 1 menunjukkan adanya respons yang tidak terduga oleh model, sedangkan nilai jauh lebih kecil dari 1 menunjukkan respons yang terlalu konsisten.

Statistik kedua mempertimbangkan bobot varians prediksi model. Varians respons diperoleh dari tahap sebelumnya. Nilai statistik dihitung dengan:

Pseudocode 3.52 Perhitungan Infit Responden

```

denom = np.sum(var_p)
infit_person[p] = np.sum((resid**2) * var_p) / denom if denom > 0
else np.nan

```

Perhitungan ini memberikan bobot lebih besar pada *item* yang memiliki informasi tinggi (varians lebih kecil), sehingga lebih sensitif terhadap pola respons yang tidak konsisten pada *item* dengan tingkat kesulitan mendekati kemampuan responden. Prosedur yang sama diterapkan pada level *item*, dengan mengiterasi setiap *item* dan menghitung residual seluruh responden yang menjawab item tersebut:

Pseudocode 3.53 Perhitungan Residual Level Item

```

mask = ~np.isnan(data_arr0[:, i])
resid = (data_arr0[mask, i] - expected[mask, i])

```

Setelah residual level *item* ditemukan, maka selanjutnya adalah menemukan nilai statistik pada level *item*. Nilai statistik pada level *item* dihitung menggunakan formula yang sama:

Pseudocode 3.54 Perhitungan Outfit dan Infit Item

```

outfit_item[i] = np.mean(resid**2)
denom = np.sum(var_i)
infit_item[i] = np.sum((resid**2) * var_i) / denom if denom > 0
else np.nan

```

3.2.10 Reliability Estimation

Tahap estimasi reliabilitas dilakukan untuk menilai tingkat konsistensi dan ketepatan hasil pengukuran yang dihasilkan oleh model Rasch RSM. Dalam konteks Rasch, reliabilitas tidak dihitung menggunakan koefisien klasik seperti Cronbach's Alpha, melainkan berdasarkan rasio antara varians sejati parameter

terhadap total varians (varians sejati ditambah varians *error* pengukuran).

Estimasi dilakukan baik pada level responden maupun level *item*.

Pseudocode 3.55 Perhitungan Reliabilitas

```
def reliability(vals, se_vals):
    if np.all(np.isnan(se_vals)):
        return np.nan
    var_true = np.nanvar(vals, ddof=1)
    err_var = np.nanmean(se_vals**2)
    return var_true / (var_true + err_var) if (var_true +
err var) != 0 else np.nan
```

Parameter `vals` merepresentasikan nilai estimasi parameter (misalnya kemampuan responden atau kesulitan *item*), sedangkan `se_vals` merupakan standar *error* dari parameter tersebut. Pada level responden, standar *error* kemampuan diperoleh dari akar varians posterior:

Pseudocode 3.56 Perhitungan Standar Error Kemampuan

```
person se = np.sqrt(theta_post_var)
```

Setelah standar *error* dihitung, selanjutnya adalah menentukan reliabilitas responden. Reliabilitas responden dihitung dengan fungsi:

Pseudocode 3.57 Perhitungan Reliabilitas Responden

```
person_reliability = reliability(theta_post_mean, person_se)
```

Di sini, `theta_post_mean` adalah estimasi kemampuan responden (θ) sedangkan `person_se` adalah standar *error* kemampuan masing-masing responden. Pada level *item*, reliabilitas dihitung menggunakan parameter kesulitan *item* (β) dan standar *error*-nya:

Pseudocode 3.58 Perhitungan Reliabilitas Item

```
item_reliability = reliability(beta_opt, se_beta)
```

Secara umum, nilai reliabilitas mendekati 1 menunjukkan bahwa varians parameter lebih banyak disebabkan oleh perbedaan sejati antar responden atau antar *item*, bukan oleh kesalahan pengukuran. Sebaliknya, nilai yang rendah menunjukkan bahwa *error* pengukuran relatif besar dibandingkan varians sejati, sehingga ketepatan pengukuran kurang optimal.

Secara umum, nilai reliabilitas mendekati 1 menunjukkan bahwa varians parameter lebih banyak disebabkan oleh perbedaan sejati antar responden atau antar *item*, bukan oleh kesalahan pengukuran. Sebaliknya, nilai yang rendah menunjukkan bahwa error pengukuran relatif besar dibandingkan varians sejati, sehingga ketepatan pengukuran kurang optimal.

3.2.11 Output

Tahap keluaran merupakan proses akhir dalam sistem estimasi model Rasch RSM, yang bertujuan menyajikan hasil analisis dalam bentuk *file* terstruktur sehingga dapat digunakan untuk interpretasi, pelaporan, dan dokumentasi penelitian. Hasil estimasi yang telah diperoleh pada tahap sebelumnya—meliputi kemampuan responden, kesulitan *item*, *threshold* kategori, statistik kesesuaian, serta reliabilitas—diformat ulang ke dalam bentuk tabel dan diekspor ke *file* Microsoft Excel.

Proses dimulai dengan penyusunan data pada level responden. Sistem membentuk sebuah DataFrame yang memuat indeks responden, nama (jika tersedia), estimasi kemampuan, standar *error* kemampuan, serta statistik kesesuaian. Penyusunan tabel dilakukan dengan kode berikut:

Pseudocode 3.59 Penyusunan Data Responden

```

person_df = pd.DataFrame({
    'RespondentIndex': np.arange(1, len(results['theta_mean']) +
1),
    'Nama': df_raw['Nama'] if 'Nama' in df_raw.columns else [''] *
len(results['theta_mean']),
    'Ability': results['theta_mean'],
    'SE_Ability': results['theta_se'],
    'Infit_MNSQ': results['infit_person'],
    'Outfit_MNSQ': results['outfit_person']
})

```

Setelah tabel terbentuk, pengguna diminta menentukan lokasi penyimpanan *file* melalui dialog penyimpanan. Proses ini dilakukan melalui fungsi:

Pseudocode 3.60 Pemilihan Lokasi Penyimpanan File

```

save_person = filedialog.asksaveasfilename(
    defaultextension=".xlsx",
    filetypes=[("Excel", "*.xlsx")],
    title="Simpan Person.xlsx"
)

```

Apabila lokasi penyimpanan telah ditentukan, maka *file* akan diekspor.

Ekspor *file* dilakukan menggunakan fungsi:

Pseudocode 3.61 Ekspor Data ke Excel

```

person_df.to_excel(save_person, index=False)

```

Langkah serupa dilakukan untuk hasil pada level item. Tabel *item* memuat indeks *item*, nama *item*, estimasi kesulitan (β), standar *error*, serta statistik kesesuaian. Penyusunan dilakukan dengan fungsi:

Pseudocode 3.62 Penyusunan Data Item

```

item_df = pd.DataFrame({
    'ItemIndex': np.arange(1, len(results['beta']) + 1),
    'ItemName': item_cols,
    'Difficulty': results['beta'],
    'SE_Difficulty': results['se_beta'],
    'Infit_MNSQ': results['infit_item'],
    'Outfit_MNSQ': results['outfit_item']
})

```

Karena *threshold* kategori (τ) merupakan parameter global dalam RSM, nilai tersebut disimpan pada lembar terpisah dalam *file* yang sama menggunakan.

Hal ini dapat dilakukan menggunakan ExcelWriter dengan fungsi:

Pseudocode 3.63 Penyimpanan Data Item dan Threshold ke Excel

```
with pd.ExcelWriter(save_item, engine='openpyxl') as writer:
    item_df.to_excel(writer, sheet_name='Items', index=False)
    tau_df = pd.DataFrame({
        'Threshold': [f'tau_{k+1}' for k in
range(len(results['tau']))],
        'Value': results['tau'],
        'SE': results['se_tau']
    })
    tau_df.to_excel(writer, sheet_name='Thresholds', index=False)
```

Selain itu, sistem juga menghasilkan ringkasan hasil analisis yang mencakup reliabilitas responden, reliabilitas *item*, nilai negatif *log-likelihood*, serta pesan dari proses optimasi. Hal ini dapat diterapkan dengan fungsi:

Pseudocode 3.64 Penyusunan Ringkasan Hasil Analisis

```
summary_df = pd.DataFrame({
    'Statistic': ['Person reliability', 'Item reliability',
'NegLogLikelihood', 'OptimizerMessage'],
    'Value': [results['person_reliability'],
results['item_reliability'],
results['negloglik'],
results['optim_message']]
})
```

BAB IV

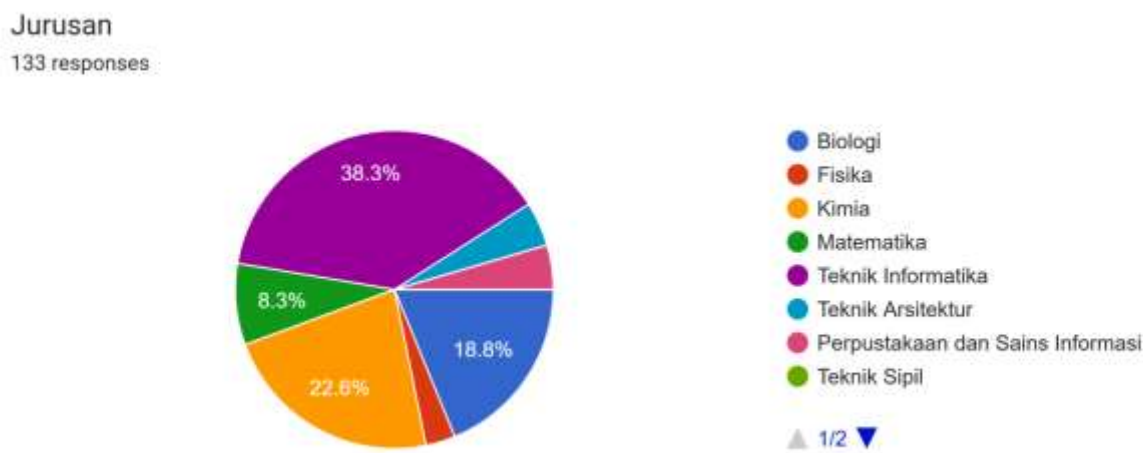
PEMBAHASAN

4.1 Hasil Responden

Pada penelitian ini, proses pengumpulan data dilakukan melalui penyebaran kuesioner daring menggunakan platform Google Form. Kuesioner ditujukan kepada mahasiswa Fakultas Sains dan Teknologi (Saintek) sebagai pengguna aktif Sistem Informasi Akademik (SIKAD) berbasis web. Pemilihan mahasiswa Saintek sebagai responden didasarkan pada tingkat penggunaan SIKAD yang tinggi dalam aktivitas perkuliahan sehari-hari, sehingga diharapkan mampu memberikan umpan balik yang representatif terkait tingkat *usability* sistem.

Total responden yang berhasil dikumpulkan sebanyak 133 mahasiswa. Jumlah tersebut telah memenuhi batas minimum responden yang diperlukan agar jumlah sampel dapat merepresentasikan populasi dengan *margin of error* sebesar 10%. Variasi jurusan di dalam Fakultas Saintek juga memungkinkan analisis yang lebih komprehensif terhadap persepsi *usability* berdasarkan latar belakang studi yang beragam.

Untuk memberikan gambaran mengenai distribusi responden, berikut adalah grafik sebaran jurusan dari masing-masing responden, yang memberikan gambaran mengenai komposisi populasi pada penelitian ini.



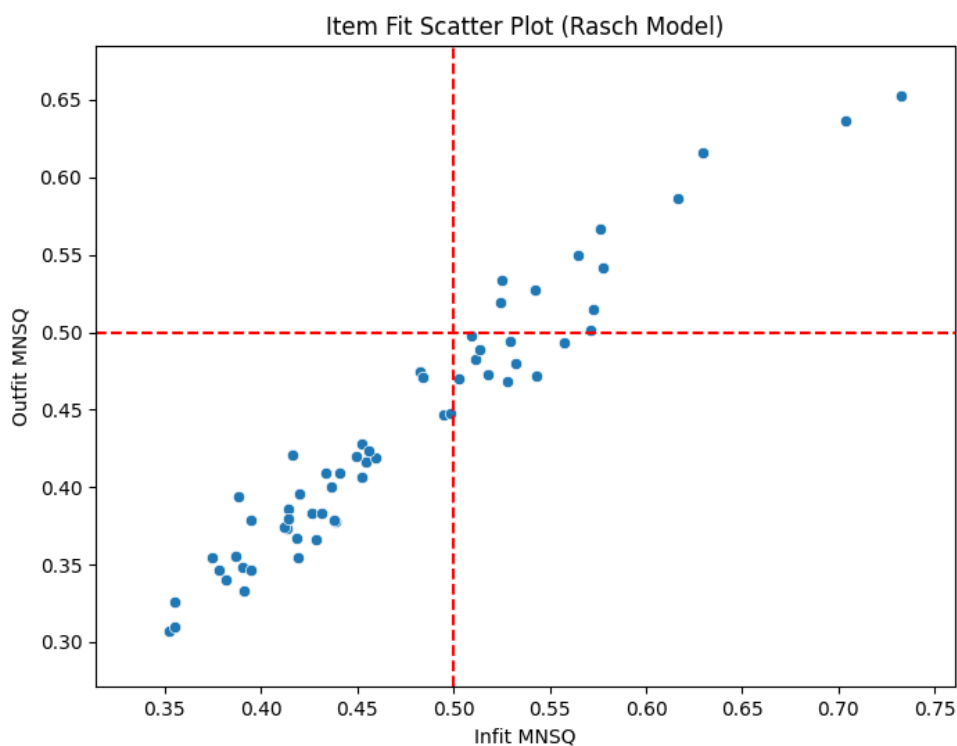
Gambar 4.1 Diagram Sebaran Jurusan

4.2 Hasil Pengukuran Menggunakan *Rasch Measurement Theory*

Pengukuran menggunakan *Rasch Measurement Theory* (RMT) dilakukan sebagai tahap awal dalam proses analisis data sebelum hasil penelitian dianalisis lebih lanjut menggunakan metode *Importance Performance Analysis* (IPA). Tahap ini diperlukan untuk memastikan bahwa data yang diperoleh dari kuesioner telah memenuhi kualitas pengukuran yang baik sehingga layak digunakan dalam analisis berikutnya. Data kuesioner pada umumnya bersifat ordinal, khususnya pada instrumen yang menggunakan skala Likert, sehingga diperlukan proses transformasi agar data tersebut dapat dianalisis secara lebih akurat dalam bentuk skala interval. Melalui RMT, respons responden dikonversi ke dalam satuan logit yang menggambarkan hubungan antara kemampuan responden dan tingkat kesulitan *item* dalam satu skala pengukuran yang sama.

4.2.1 Item Fit

Evaluasi kecocokan *item* (*item fit*) dilakukan untuk memastikan bahwa setiap butir pertanyaan pada instrumen USE *Questionnaire* berperilaku sesuai dengan asumsi model Rasch *Rating Scale Model* (RSM) yang digunakan dalam penelitian ini. Penilaian *item fit* difokuskan pada dua indeks utama, yaitu *Infit Mean Square* (*Infit MNSQ*) dan *Outfit Mean Square* (*Outfit MNSQ*). Kedua indeks ini memberikan gambaran mengenai sejauh mana pola respons responden terhadap masing-masing item konsisten dengan prediksi model Rasch.



Gambar 4.2 *Item Fit Scatter Plot*

Dalam model Rasch, nilai MNSQ yang dianggap ideal berada pada kisaran 0,5 hingga 1,5. Nilai di bawah 0,5 menunjukkan bahwa suatu *item* terlalu

'*predictable*' sehingga mungkin kurang memberikan informasi tambahan bagi model, sedangkan nilai di atas 1,5 mengindikasikan adanya ketidaksesuaian (*misfit*) akibat pola respons yang tidak stabil atau tidak konsisten. Oleh karena itu, rentang 0,5 hingga 1,5 digunakan sebagai dasar untuk menilai apakah suatu *item* dapat dinyatakan *fit* atau tidak dalam model.

Berdasarkan hasil analisis *item* dengan memakai skala Likert 1–5 serta pendekatan RSM, dapat dilihat pada Gambar 4.2 bahwa seluruh butir pada instrumen menunjukkan nilai *Infit* dan *Outfit* MNSQ yang berada di bawah 1,0, dengan rentang keseluruhan antara 0,30 hingga 0,73. Semua nilai tersebut berada jauh di dalam batas toleransi model Rasch, sehingga dapat dinyatakan bahwa tidak ada *item* yang mengalami *overfit* maupun *underfit*. Kondisi ini menunjukkan bahwa setiap *item* memberikan kontribusi informasi yang baik, bekerja secara stabil, dan tidak menunjukkan pola respons anomali.

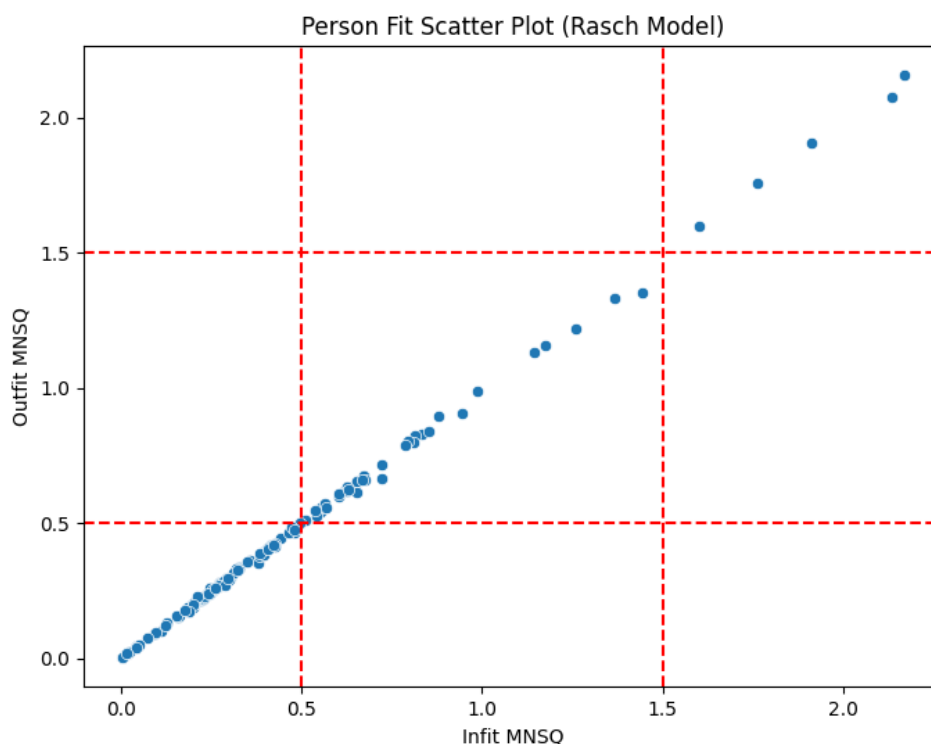
Secara umum, *item-item* yang memiliki nilai MNSQ mendekati 1 dianggap paling ideal karena menunjukkan keseimbangan antara konsistensi dan variasi respons. Sementara itu, beberapa *item* dalam instrumen ini memiliki nilai MNSQ di kisaran 0,30 hingga 0,40, yang menandakan tingkat prediktabilitas tinggi. Meskipun demikian, kondisi ini tidak menimbulkan masalah signifikan dalam konteks analisis Rasch, sebab nilai MNSQ yang terlalu rendah masih dapat diterima sepanjang tidak mempengaruhi fungsi diskriminasi instrumen. Hal ini menunjukkan bahwa responden relatif konsisten dalam menjawab pernyataan yang disajikan, dan struktur *item* telah dipahami dengan baik.

Seluruh *item* pada setiap dimensi, baik *Usefulness* (U), *Ease of Use* (EoU), *Ease of Learning* (EoL), maupun *Satisfaction* (S), menunjukkan karakteristik *fit* yang kuat. Tidak ditemukan butir dengan nilai MNSQ melebihi 1,0, apalagi melampaui batas 1,5. Hasil ini menegaskan bahwa instrumen bekerja secara harmonis dengan asumsi model Rasch dan setiap item mampu mengukur konstruksi yang sama secara seragam tanpa mengarah pada *misfit* ataupun *noise* yang berlebihan.

Dengan demikian, hasil analisis *item fit* ini dapat disimpulkan bahwa seluruh 60 *item* yang dianalisis memenuhi kriteria kecocokan model Rasch. Instrumen dapat dinyatakan memiliki kualitas psikometrik yang baik, valid dari perspektif model Rasch, serta layak digunakan dalam pengukuran usability Sistem Informasi Akademik berbasis web yang menjadi objek penelitian.

4.2.2 Person Fit

Analisis *Person Fit* dilakukan untuk menilai sejauh mana pola respons setiap responden sesuai dengan ekspektasi model Rasch. Model Rasch mengasumsikan bahwa responden dengan tingkat kemampuan atau sikap (*ability*) tertentu akan memberikan respons yang konsisten terhadap butir-butir dengan tingkat kesulitan yang berbeda. Parameter yang digunakan untuk mengevaluasi kesesuaian tersebut adalah *Infit* MNSQ dan *Outfit* MNSQ, sama seperti yang digunakan pada analisis *Item fit*. Dalam penelitian ini, kriteria kesesuaian yang digunakan adalah nilai MNSQ berada pada rentang 0.5 hingga 1.5. Nilai di luar rentang tersebut menunjukkan adanya ketidaksesuaian pola respons.



Gambar 4.3 *Person Fit Scatter Plot*

Seperti halnya pada analisis *item fit*, responden dikategorikan menjadi 3 dengan kriteria sebagai berikut:

- 1) Responden dengan Indikasi *Misfit* ($Infit/Outfit > 1.5$)

Misfit terjadi ketika pola respons responden bersifat terlalu acak, tidak stabil, atau tidak dapat diprediksi oleh model. Pada konteks skala Likert, kondisi ini dapat mencerminkan ketidakkonsistenan dalam memberikan *rating*, misalnya memberikan nilai tinggi pada pernyataan yang relatif negatif tetapi nilai rendah pada pernyataan positif, atau jawaban yang cenderung random. Pada Gambar 4.3, dapat dilihat terdapat 5 titik yang mengindikasikan data *misfit*.

Nilai-nilai tersebut berada jauh di atas ambang batas 1.5 sehingga dapat mengindikasikan adanya pola respons yang tidak valid, seperti jawaban yang tidak dipikirkan (*careless responding*) atau ketidaktelitian ketika mengisi instrumen. Responden dengan tingkat *misfit* ekstrem seperti ini perlu diperiksa lebih lanjut, termasuk kemungkinan dikeluarkan dari analisis apabila ditemukan pola respons yang tidak rasional.

2) Responden dengan Indikasi *Overfit* ($Infit/Outfit < 0.5$)

Overfit terjadi ketika pola respons terlalu sempurna atau terlalu dapat diprediksi oleh model. Hal ini sering muncul ketika responden memberikan jawaban yang sama pada hampir semua butir, misalnya selalu memberikan nilai 5 atau 1 tanpa mempertimbangkan variasi isi pernyataan.

Nilai MNSQ yang sangat rendah (mendekati 0) mengindikasikan adanya pola respons yang terlalu seragam. Responden dengan *ability* yang sangat tinggi ini kemungkinan memberikan nilai setuju maksimum pada hampir semua pernyataan, sehingga kehilangan variabilitas yang dibutuhkan untuk estimasi Rasch. Meskipun *Overfit* tidak selalu merusak model, nilai ekstrem seperti ini biasanya menandakan kurangnya diferensiasi respons antar butir.

3) Responden dengan Pola Respons *Fit* ($0.5 \leq Infit/Outfit \leq 1.5$)

Sebagian besar responden berada dalam kategori *fit*, menunjukkan bahwa respons yang diberikan konsisten dengan karakteristik butir dalam model. Responden dengan nilai MNSQ dalam batas toleransi menunjukkan pola respons yang stabil, rasional, serta sesuai dengan prediksi Rasch.

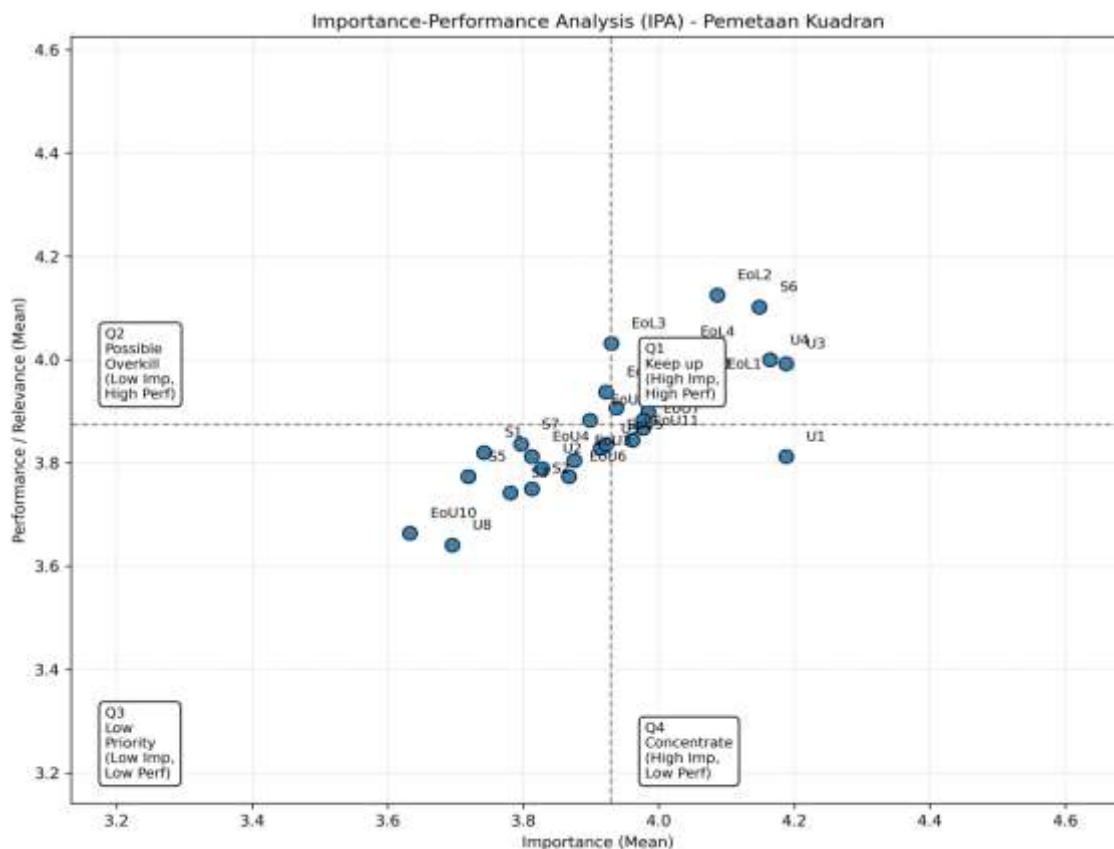
Hal ini menggambarkan bahwa mayoritas responden memahami instrumen dengan baik dan memberikan penilaian sesuai dengan tingkat kesepakatan mereka terhadap setiap pernyataan.

Berdasarkan hasil analisis terhadap 133 responden seperti yang dapat dilihat pada Gambar 4.3, dapat disimpulkan bahwa sebagian besar data respons berada dalam kategori *fit*, sehingga layak digunakan dalam analisis lebih lanjut. Hanya sebagian kecil responden yang teridentifikasi sebagai *outlier* dalam bentuk *Misfit*. Responden dengan *misfit* ekstrem diputuskan untuk tidak diikuti ke dalam analisa data lebih lanjut.

4.3 Hasil *Importance Performance Analysis*

Setelah 5 responden yang sebelumnya ditetapkan sebagai *Misfit* menurut perhitungan rasch, analisis *Importance Performance Analysis* (IPA) dilakukan untuk memetakan posisi setiap *item* pertanyaan berdasarkan tingkat kepentingan (*importance*) dan tingkat kinerja aktual sistem (*performance*) menurut responden. IPA memberikan gambaran mengenai prioritas perbaikan yang harus dilakukan pada Sistem Informasi Akademik (SIKAD) berdasarkan persepsi mahasiswa sebagai pengguna langsung.

Data yang digunakan dalam analisis ini berasal dari 30 item pernyataan yang mencakup empat dimensi usability, yaitu *Usefulness* (U), *Ease of Use* (EoU), *Ease of Learning* (EoL), dan *Satisfaction* (S). Setiap *item* dihitung nilai rata-rata *importance* dan *performancenya*, kemudian selisih keduanya (*Mean Difference*) digunakan untuk mengidentifikasi kesenjangan persepsi antara harapan dan kondisi saat ini.



Gambar 4.4 Diagram IPA

Berdasarkan grafik yang disediakan pada Gambar 4.4, secara umum, nilai rata-rata *importance* pada seluruh *item* berada dalam rentang 3,63 hingga 4,18, yang menunjukkan bahwa seluruh atribut SIAKAD dianggap penting oleh mahasiswa. Sementara itu, nilai rata-rata *performance* berada dalam rentang 3,64 hingga 4,12, yang menandakan bahwa sistem telah memberikan kinerja yang cukup baik, meskipun masih terdapat beberapa aspek yang belum memenuhi harapan pengguna.

Item dengan nilai *importance* tertinggi antara lain: U1 (4,1875), U3 (4,1875), EoL2 (4,0859), dan EoL1 (4,0703)

Hal ini mengindikasikan bahwa aspek manfaat dan kemudahan pembelajaran sistem masih memegang peran penting dalam pengalaman pengguna terhadap SIAKAD.

Perbandingan antara nilai *Importance* dan *Performance* menghasilkan *Mean Difference* yang menunjukkan sejauh mana kinerja sistem memenuhi tingkat kepentingan pengguna. Selisih positif menunjukkan bahwa harapan pengguna lebih tinggi dibandingkan kinerja aktual, sementara selisih negatif menunjukkan kondisi yang sudah sesuai atau sedikit melampaui harapan.

Beberapa temuan penting adalah sebagai berikut:

1) *Item* dengan Kesenjangan Tertinggi (Perlu Prioritas Perbaikan)

Beberapa *item* menunjukkan selisih positif yang cukup besar sehingga berada dalam kategori perlu ditingkatkan, antara lain: EoU11 (+0,1172), EoL1 (+0,1172), EoU7 (+0,1093), EoU9 (+0,09375), U3 (+0,1953)

Nilai ini menunjukkan bahwa mahasiswa menilai beberapa aspek kemudahan penggunaan (*Ease of Use*) serta kemudahan mempelajari sistem (*Ease of Learning*) masih belum sepenuhnya terpenuhi. Pada sisi *Usefulness*, *item* U3 memiliki gap yang cukup mencolok, menandakan adanya kebutuhan peningkatan terkait kemanfaatan fitur tertentu.

2) *Item* dengan Kinerja Relatif Memuaskan

Beberapa *item* memiliki selisih negatif, yang menunjukkan bahwa kinerja sistem sudah berada pada level yang sebanding atau bahkan sedikit lebih baik dari harapan pengguna. Di antaranya adalah EoU2 (-0,0156), EoU10 (-0,03125), EoL2 (-0,03906), S1 (-0,0781), EoL3 (-0,1015)

Item-item ini dapat dikategorikan sebagai “pertahankan kualitas” karena kinerjanya sudah selaras dengan tingkat kepentingan yang dirasakan pengguna.

Berdasarkan distribusi *Mean Difference* dan mempertimbangkan pola umum *importance* dan *performance*, dapat disimpulkan beberapa rekomendasi:

- 1) Fokus prioritas perbaikan perlu diberikan pada item dengan gap tinggi, terutama pada dimensi *Ease of Use* dan *Ease of Learning*. Hal ini menunjukkan bahwa meskipun mahasiswa menganggap sistem penting, sebagian fitur atau alur penggunaan masih dirasakan kurang intuitif.
- 2) Kemanfaatan sistem (*Usefulness*) pada item tertentu seperti U3 perlu evaluasi lebih lanjut, terutama terkait apakah fungsi sistem benar-benar mendukung kebutuhan akademik pengguna.
- 3) *Item-item* dengan selisih negatif yang kecil menunjukkan performa yang sudah baik. Atribut-atribut tersebut harus dipertahankan kualitasnya karena dianggap sudah memenuhi harapan mahasiswa.
- 4) IPA juga dapat digunakan untuk membuat peta kuadran (*Quadrant Analysis*) yang memvisualisasikan prioritas pengembangan. Pada penelitian ini, gap analysis sudah cukup menggambarkan mana aspek yang perlu perhatian segera.

4.4 Integrasi Islam

Integrasi nilai-nilai Islam dalam penelitian ini bertujuan untuk memberikan landasan normatif dan etis terhadap hasil evaluasi usability SIAKAD berbasis web. Pendekatan ini menegaskan bahwa pengembangan dan evaluasi teknologi informasi tidak hanya berorientasi pada aspek teknis dan empiris, tetapi juga harus

selaras dengan nilai-nilai keislaman yang menekankan kemanfaatan, kemudahan, keadilan, serta pengembangan ilmu pengetahuan. Nilai-nilai tersebut sebagaimana telah diuraikan pada Bab I menjadi dasar dalam mengaitkan hasil penelitian dengan prinsip-prinsip ajaran Islam.

Dalam konteks kemanfaatan (*usefulness*), hasil evaluasi *usability* menunjukkan bahwa sistem SIAKAD memiliki peran penting dalam membantu mahasiswa dalam menjalankan aktivitas akademik, seperti pengisian Kartu Rencana Studi (KRS), akses Kartu Hasil Studi (KHS), serta pengelolaan informasi akademik. Temuan ini sejalan dengan prinsip Islam yang menekankan pentingnya memberikan manfaat bagi sesama manusia. Rasulullah SAW bersabda:

خَيْرُ النَّاسِ أَنْفَعُهُمْ لِلنَّاسِ

"Sebaik-baik manusia adalah yang paling bermanfaat bagi orang lain" (HR. Ahmad)

Hadis ini menunjukkan bahwa nilai kebermanfaatan merupakan indikator utama dalam menilai kualitas suatu perbuatan. Dalam penjelasannya, Imam Al-Munawi dalam kitab *Faid al-Qadir Syarh al-Jami' al-Shaghir* menyatakan bahwa manusia terbaik adalah mereka yang mampu memberikan kontribusi nyata bagi kehidupan orang lain, baik dalam bentuk materi maupun non-materi, termasuk dalam bentuk layanan dan kemudahan akses informasi. Dengan demikian, sistem SIAKAD yang mampu memberikan manfaat nyata bagi mahasiswa dapat dikatakan telah memenuhi prinsip kemaslahatan (*maslahah*) dalam Islam.

Selanjutnya, dalam aspek kemudahan penggunaan (*ease of use*) dan kemudahan pembelajaran (*ease of learning*), hasil penelitian menunjukkan bahwa tingkat kemudahan sistem menjadi faktor penting dalam meningkatkan pengalaman pengguna. Hal ini sejalan dengan prinsip Islam yang menekankan kemudahan dan tidak mempersulit dalam setiap urusan. Allah SWT berfirman dalam Al-Qur'an:

يُرِيدُ اللَّهُ بِكُمُ الْيُسْرَ وَلَا يُرِيدُ بِكُمُ الْعُسْرَ

“Allah menghendaki kemudahan bagimu dan tidak menghendaki kesukaran bagimu” (QS. Al-Baqarah: 185)

Menurut Ibnu Katsir dalam kitab tafsirnya *Tafsir al-Qur'an al-'Azhim*, ayat ini menunjukkan bahwa seluruh syariat Islam dibangun di atas prinsip kemudahan (*taysir*) dan menghindari kesulitan (*ta'sir*). Dalam konteks sistem informasi, prinsip ini dapat diimplementasikan melalui desain antarmuka yang sederhana, navigasi yang jelas, serta proses interaksi yang tidak membebani pengguna. Oleh karena itu, sistem SIAKAD yang mudah digunakan dan mudah dipelajari mencerminkan implementasi nilai *taysir* dalam teknologi.

Selain itu, aspek kepuasan pengguna (*satisfaction*) yang diperoleh dari hasil *USE Questionnaire* juga dapat dikaitkan dengan konsep ihsan dalam Islam, yaitu melakukan sesuatu dengan sebaik-baiknya. Kepuasan pengguna menunjukkan bahwa sistem tidak hanya berfungsi secara teknis, tetapi juga mampu memberikan pengalaman yang baik. Dalam hal ini, Al-Ghazali dalam karyanya *Ihya' Ulumuddin* menjelaskan bahwa kualitas suatu amal tidak hanya diukur dari terpenuhinya kewajiban, tetapi juga dari kesempurnaan dan keindahan dalam

pelaksanaannya. Dengan demikian, sistem yang memberikan kepuasan kepada pengguna dapat dikatakan telah mendekati konsep ihsan dalam implementasinya.

Hasil analisis menggunakan IPA menunjukkan adanya beberapa atribut *usability* yang berada pada kuadran prioritas perbaikan (*Concentrate Here*), yaitu atribut dengan tingkat kepentingan tinggi namun kinerja yang masih rendah. Temuan ini mencerminkan perlunya upaya perbaikan berkelanjutan yang melibatkan berbagai pihak. Dalam Islam, konsep ini sejalan dengan prinsip amar ma'ruf nahi munkar sebagaimana firman Allah SWT:

وَالْمُؤْمِنُونَ وَالْمُؤْمِنَاتُ بَعْضُهُمْ أَوْلِيَاءُ بَعْضٍ يَأْمُرُونَ بِالْمَعْرُوفِ وَيَنْهَوْنَ عَنِ الْمُنْكَرِ

“Dan orang-orang yang beriman, laki-laki dan perempuan, sebagian mereka adalah penolong bagi sebagian yang lain. Mereka menyuruh kepada yang ma'ruf dan mencegah dari yang munkar.” (QS. At-Tawbah: 71)

Menurut Al-Qurtubi dalam *Al-Jami' li Ahkam al-Qur'an*, ayat ini menekankan pentingnya kerja sama dalam memperbaiki kondisi sosial dan sistem yang ada. Dalam konteks penelitian ini, hasil IPA dapat dijadikan dasar kolaborasi antara pengembang sistem dan pengguna dalam meningkatkan kualitas SIAKAD secara berkelanjutan.

Selanjutnya, penggunaan RMT dalam penelitian ini berfungsi untuk memastikan validitas dan reliabilitas instrumen pengukuran secara objektif. Pendekatan ini dapat dikaitkan dengan prinsip keadilan ('adl) dalam Islam. Allah SWT berfirman:

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya.”
(QS. Al-Baqarah: 286)

Menurut Fakhruddin al-Razi dalam *Mafatih al-Ghaib*, ayat ini menunjukkan bahwa setiap beban atau penilaian harus disesuaikan dengan kapasitas individu secara adil dan proporsional. Dalam model Rasch, probabilitas respons individu ditentukan oleh perbedaan antara kemampuan responden dan tingkat kesulitan *item*, sehingga mencerminkan prinsip keadilan dalam pengukuran. Dengan demikian, penggunaan RMT dalam penelitian ini tidak hanya memberikan validitas ilmiah, tetapi juga selaras dengan nilai keadilan dalam Islam.

Lebih lanjut, integrasi nilai Islam dalam penelitian ini juga mencerminkan pentingnya ilmu pengetahuan sebagai sarana untuk meningkatkan kualitas kehidupan manusia. Allah SWT berfirman dalam QS. Al-‘Alaq ayat 1–5 yang menekankan perintah membaca dan belajar. Dalam tafsirnya, Sayyid Qutb dalam *Fi Zhilal al-Qur’an* menjelaskan bahwa wahyu pertama ini merupakan landasan utama peradaban Islam yang berbasis ilmu pengetahuan. Dalam konteks penelitian ini, penggunaan metode ilmiah seperti USE *Questionnaire*, IPA, dan RMT merupakan bentuk implementasi dari perintah tersebut dalam bidang teknologi informasi.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa hasil penelitian ini tidak hanya memberikan kontribusi dalam bidang evaluasi *usability* sistem informasi, tetapi juga menunjukkan keselarasan dengan nilai-nilai Islam yang mencakup kemanfaatan, kemudahan, kepuasan, keadilan, kerja sama, serta

pengembangan ilmu pengetahuan. Integrasi ini memperkuat bahwa pengembangan teknologi, khususnya dalam lingkungan akademik, seharusnya tidak hanya berorientasi pada efisiensi dan efektivitas, tetapi juga pada nilai-nilai etis dan spiritual yang memberikan manfaat yang lebih luas bagi masyarakat.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis usability menggunakan pendekatan USE *Questionnaire* yang diintegrasikan dengan *Importance Performance Analysis* (IPA), dapat disimpulkan bahwa tingkat *usability* Sistem Informasi Akademik (SIKAD) berbasis web di Universitas Islam Negeri Maulana Malik Ibrahim Malang berada pada kategori yang cukup baik, namun masih memerlukan optimalisasi pada beberapa atribut prioritas. Hal ini ditunjukkan oleh sebaran kuadran IPA, dimana terdapat 12 *item* pada Kuadran I (prioritas utama untuk perbaikan), 2 *item* pada Kuadran II (pertahankan kinerja), 13 *item* pada Kuadran III (prioritas rendah), dan 3 *item* pada Kuadran IV (berlebihan). Dominasi *item* pada Kuadran I dan III menunjukkan bahwa masih terdapat kesenjangan antara tingkat kepentingan dan kinerja sistem, khususnya pada aspek utilitas dan kemudahan penggunaan. Secara keseluruhan, hasil ini mengindikasikan bahwa meskipun sistem telah memenuhi kebutuhan dasar pengguna, peningkatan kualitas layanan tetap diperlukan untuk mencapai tingkat *usability* yang optimal. Adapun visualisasi sebaran atribut secara rinci dapat dilihat pada Gambar 4.4.

Selanjutnya, berdasarkan validasi menggunakan *Rasch Measurement Theory* (RMT), dapat disimpulkan bahwa pengukuran hubungan antara tingkat kepentingan dan kinerja atribut dalam analisis IPA memiliki tingkat validitas yang baik. Hal ini dibuktikan dengan tidak ditemukannya *item* yang mengalami *misfit*,

sehingga seluruh *item* kuesioner dinyatakan mampu mengukur konstruk *usability* secara konsisten. Namun demikian, terdapat 5 responden yang terindikasi *misfit* dan dikeluarkan dari analisis karena memiliki nilai *Outfit* dan *Infit Mean Square* di luar batas yang dapat diterima, masing-masing sebesar (1,76;1,75), (2,13;2,08), (1,60;1,60), (2,17;2,16), dan (1,91;1,90). Pengecualian responden ini bertujuan untuk menjaga kualitas estimasi model agar tetap memenuhi asumsi Rasch. Dengan demikian, integrasi antara IPA dan RMT dalam penelitian ini terbukti mampu menghasilkan pengukuran yang valid dan reliabel dalam mengevaluasi *usability* sistem. Untuk informasi lebih rinci mengenai analisis kesesuaian *item* dan responden, dapat dilihat pada Gambar 4.2 dan Gambar 4.3.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa rekomendasi yang dapat dipertimbangkan untuk pengembangan penelitian selanjutnya agar menghasilkan temuan yang lebih komprehensif dan mendalam.

Penelitian berikutnya disarankan untuk memperluas cakupan responden dengan melibatkan lebih banyak program studi, jenjang pendidikan, atau bahkan institusi yang berbeda. Hal ini bertujuan untuk meningkatkan generalisasi hasil penelitian serta memperoleh gambaran *usability* sistem yang lebih representatif terhadap berbagai karakteristik pengguna.

Pengembangan instrumen penelitian juga dapat dilakukan dengan menambahkan variabel lain yang relevan, seperti aspek kualitas layanan, pengalaman pengguna (*user experience*), atau faktor teknis sistem seperti kecepatan akses dan stabilitas server. Integrasi metode evaluasi lain di luar USE

Questionnaire dan IPA dapat memberikan sudut pandang tambahan dalam menilai kualitas sistem secara menyeluruh.

Pada sisi analisis data, penelitian selanjutnya dapat mempertimbangkan penggunaan model pengukuran lain sebagai pembandingan terhadap RMT, sehingga diperoleh validasi silang yang memperkuat akurasi hasil pengukuran. Selain itu, analisis longitudinal juga dapat dilakukan untuk melihat perubahan tingkat *usability* sistem dari waktu ke waktu setelah dilakukan perbaikan berdasarkan hasil evaluasi.

Penelitian mendatang juga disarankan untuk mengkaji lebih dalam atribut-atribut yang berada pada kuadran prioritas utama dalam pemetaan IPA, baik melalui pendekatan kuantitatif lanjutan maupun melalui metode kualitatif seperti wawancara atau *focus group discussion*. Pendekatan tersebut dapat memberikan pemahaman yang lebih detail mengenai akar permasalahan yang dirasakan pengguna.

Dengan pengembangan pada aspek cakupan responden, instrumen, metode analisis, serta pendekatan evaluasi yang lebih beragam, penelitian selanjutnya diharapkan mampu menghasilkan model evaluasi *usability* yang semakin komprehensif, adaptif, dan aplikatif dalam mendukung peningkatan kualitas SIAKAD berbasis web.

DAFTAR PUSTAKA

- AlSalem, T. S., & AlShamari, M. A. (2023). Assessing Interactive Web-Based Systems Using Behavioral Measurement Techniques. *Future Internet*, *15*(11), 365.
- Andrich, D., & Marais, I. (2019). *A course in Rasch measurement theory: Measuring in the educational, social and health sciences*. Springer.
- Andry, J. F., Christianto, K., & Wilujeng, F. R. (2019). Using Webqual 4.0 and importance performance analysis to evaluate e-commerce website. *Journal of Information Systems Engineering and Business Intelligence*, *5*(1), 23–31.
- Aryadoust, V., Ng, L. Y., & Sayama, H. (2021). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing*, *38*(1), 6-40.
- Aryadoust, V., Tan, H. A. H., & Ng, L. Y. (2019). A Scientometric review of Rasch measurement: The rise and progress of a specialty. *Frontiers in Psychology*, *10*, 2197.
- de Campos, T. P., Damasceno, E. F., & Valentim, N. M. Usability and User Experience Questionnaire Evaluation and Evolution for Touchable Holography.
- Esmailpour, J., Aghabayk, K., Vajari, M. A., & De Gruyter, C. (2020). Importance–Performance Analysis (IPA) of bus service attributes: A case study in a developing country. *Transportation Research Part A: Policy and Practice*, *142*, 129-150.
- Fadhilah, F., & Meuthia, R. F. (2022). Measuring usability of academic information system using USE *Questionnaire*: Case study of Padang State Polytechnic. *Malaysian Journal of Business and Economics*, *9*(2), 79-91.
- Hamzah, M. L., Rahmadhani, R. F., & Purwati, A. A. (2022). An integration of Webqual 4.0, Importance Performance Analysis and Customer Satisfaction Index on E-Campus. *Journal of System and Management Sciences*, *12*(3), 25-50.
- Hariyanto, D., Triyono, M., & Köhler, T. (2020). Usability evaluation of personalized adaptive e-learning system using USE questionnaire. *Knowledge Management & E-Learning*, *12*(1), 85–105.
- Hendra, S., & Arifin, Y. (2018). Web-based usability measurement for student grading information system. *Procedia Computer Science*, *135*, 238–247.

- Joung, J., & Kim, H. M. (2021). Approach for importance–performance analysis of product attributes from online reviews. *Journal of Mechanical Design*, *143*(8), 081705.
- Kilis, B. M., Mamonto, J., Legesan, H., & Sagai, Z. (2022). Usability Evaluation of The Android Operating System Using Use Questionnaire. *International Journal of Information Technology and Education*, *1*(2), 16-23.
- Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface*, *8*(2), 3–6.
- Martilla, J. A., & James, J. C. (1977). Importance-Performance Analysis. *Journal of Marketing*, *41*, 77–79.
- Melin, J., Bonn, S. E., Pendrill, L., & Lagerros, Y. T. (2020). A questionnaire for assessing user satisfaction with mobile health apps: Development using Rasch measurement theory. *JMIR mHealth and uHealth*, *8*(5), e15909.
- Moulaei, K., Moulaei, R., & Bahaadinbeigy, K. (2024). The most used questionnaires for evaluating the usability of robots and smart wearables: A scoping review. *Digital Health*, *10*, 20552076241237384.
- Phadermrod, B., Crowder, R. M., & Wills, G. B. (2019). Importance-performance analysis based SWOT analysis. *International Journal of Information Management*, *44*, 194–203.
- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 020111.
- Rasch, G. (1980) [1960]. Probabilistic models for some intelligence and attainment tests. *University of Chicago Press*.
- Sukmasetya, P., Setiawan, A., & Arumi, E. R. (2020). Usability evaluation of university website: A case study. *Journal of Physics: Conference Series*, *1517*(1), 012071.