Grounding Case-Based Reasoning (CBR) untuk Mitigasi Halusinasi Large Language Model (LLM) pada
Sistem Pendukung Keputusan Klinis Polycystic Ovary Syndrome (PCOS) Berbasis Kriteria Rotterdam
Diaz Astari Beningati, Prof. Dr.Eng. Ir. Igi Ardiyanto, S.T., M.Eng., IPM., ASEAN Eng., SMIEEE.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN PENGESAHAN ii
PERNYATAAN BEBAS PLAGIASI iii
HALAMAN PERSEMBAHAN iv
KATA PENGANTAR v
DAFTAR ISI vi
DAFTAR TABEL ix
DAFTAR GAMBAR x
DAFTAR SINGKATAN xi
INTISARI xii
ABSTRACT xiii
BAB I Pendahuluan 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 4
1.3 Tujuan Penelitian 5
1.4 Batasan Penelitian 5
1.5 Manfaat Penelitian 6
1.6 Sistematika Penulisan 6
BAB II Tinjauan Pustaka dan Dasar Teori 8
2.1 Tinjauan Pustaka 8
2.2 Dasar Teori 10
2.2.1 Diagnosis Domain: PCOS 10
2.2.2 Fenotipe PCOS 15
2.2.3 LLM dalam Diagnosis Klinis & Fenomena Halusinasi Intrinsik 15
2.2.4 Definisi Halusinasi 16
2.2.5 Faithfulness dan Factuality 16
2.2.6 Mitigasi Halusinasi dengan *Prompt Engineering* untuk *Reasoning* Klinis 16
2.2.7 *Case-Based Reasoning* (CBR) dalam Diagnosis Medis 17
2.2.7.1 Retrieve 18
2.2.7.2 Reuse 18
2.2.7.3 Revise 18
2.2.7.4 Retain 18
2.2.8 *Similarity Measure*/Ukuran Kesamaan Data Heterogen 19
2.2.8.1 *Heterogeneous Euclidean-Overlap Metric* (HEOM) 19
2.2.8.2 *Clinical Decision Tree* (CDT) 20

UNIVERSITAS GADJAH MADA
Grounding Case-Based Reasoning (CBR) untuk Mitigasi Halusinasi Large Language Model (LLM) pada
Sistem Pendukung Keputusan Klinis Polycystic Ovary Syndrome (PCOS) Berbasis Kriteria Rotterdam
Diaz Astari Beningati, Prof. Dr.Eng. Ir. Igi Ardiyanto, S.T., M.Eng., IPM., ASEAN Eng., SMIEEE.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id

2.2.8.3 *Hierarchical Weighted Heterogeneous Similarity Measure* (HWHSM) 20
2.2.9 Teori Ruang Metrik & Ketidaksesuaiannya 21
2.2.10 Matriks Evaluasi 22
2.2.10.1 *Retrieve-based* dan *Reuse-based* melalui *Leave-One-Out Cross-Validation* (LOOCV) 23
2.2.11 Evaluasi Halusinasi LLM 24
2.2.11.1 Uji Deterministik Proposisi 24
2.2.11.2 Uji McNemar 24
2.2.11.3 Koreksi Holm 25
2.3 Analisis Perbandingan Metode Penelitian Terdahulu 25
BAB III Metode Penelitian 28
3.1 Alat dan Bahan Tugas Akhir 28
3.1.1 Alat Tugas Akhir 28
3.1.2 Bahan Tugas Akhir 28
3.2 Metode Tugas Akhir 31
3.3 Alur Penelitian Tugas Akhir 31
3.3.1 *Preprocessing* dan *Feature Engineering* 33
3.3.2 Case-Based Reasoning (CBR) 37
3.3.2.1 Clinical Decision Tree (CDT) 37
3.3.2.2 HWHSM 38
3.3.2.3 *Reuse* dan *Mimic Gate* 39
3.3.2.4 Konfigurasi Pembanding 40
3.3.3 Leave-One-Out Cross-Validation (LOOCV) 40
3.3.4 Modul *Large Language Model* (LLM) 41
3.3.4.1 Model dan Konfigurasi 41
3.3.4.2 Strategi *Prompting* 41
3.3.4.3 Struktur *Prompt* dan Arsitektur 42
3.3.4.4 Keluaran Terstruktur 43
3.3.5 Evaluasi Halusinasi 44
3.3.5.1 Studi Pilot 44
3.3.5.2 Verifikasi Pernyataan 44
3.3.5.3 Metrik dan Uji Statistik 45
3.3.6 Uji Keterulangan dan Eksperimen Input Naratif 46
3.3.6.1 Uji Keterulangan 46
3.3.6.2 Eksperimen Input Naratif 46
BAB IV Hasil dan Pembahasan 48
4.1 Hasil Evaluasi *Retrieval* dengan LOOCV 48
4.1.1 Kriteria *Retrieve-based* 48

UNIVERSITAS GADJAH MADA
Grounding Case-Based Reasoning (CBR) untuk Mitigasi Halusinasi Large Language Model (LLM) pada
Sistem Pendukung Keputusan Klinis Polycystic Ovary Syndrome (PCOS) Berbasis Kriteria Rotterdam
Diaz Astari Beningati, Prof. Dr.Eng. Ir. Igi Ardiyanto, S.T., M.Eng., IPM., ASEAN Eng., SMIEEE.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

4.1.2 *Kriteria Rome-based* 48
4.1.3 Sensitivitas terhadap Nilai k 50
4.1.4 Analisis Kesalahan 50
4.1.5 Uji Urutan Data 51
4.1.6 Pembahasan 51
4.2 Hasil Evaluasi Halusinasi 52
4.2.1 Studi Pilot 52
4.2.2 Validasi Pemeriksa 53
4.2.3 Halusinasi Intrinsik Keseluruhan dan per Kelompok 53
4.2.4 Angka yang Tidak Didukung Sumber 54
4.2.5 Halusinasi per Jenis Proposisi 55
4.2.6 Kegagalan Penerapan Aturan Dua dari Tiga Kriteria 56
4.2.7 Perbandingan dengan CBR-*only* dan Label Diagnosis PCOS Dataset 57
4.2.8 Contoh Kasus 58
4.2.9 Pembahasan 59
4.3 Hasil Uji Keterulangan 61
4.4 Hasil Eksperimen Input Naratif 63
BAB V Kesimpulan dan Saran 65
5.1 Kesimpulan 65
5.2 Saran 66
DAFTAR PUSTAKA 68
LAMPIRAN L-1
L.1 Source Code Penelitian L-1
L.2 Aturan Pelabelan Basis Kasus L-1
L.3 Ambang Numerik dan Rujukan Klinis L-2
L.4 Contoh Keluaran Tiga Arsitektur untuk Pasien 149 ..... L-2
L.5 Contoh *Prompt* dan Keluaran LLM untuk Pasien 149 ..... L-4
L.5.1 *Prompt* CBR+LLM (Format Terstruktur) L-4
L.5.2 Bagian Data Pasien pada Format Naratif L-6
L.5.3 Keluaran JSON L-6

Grounding Case-Based Reasoning (CBR) untuk Mitigasi Halusinasi Large Language Model (LLM) pada
Sistem Pendukung Keputusan Klinis Polycystic Ovary Syndrome (PCOS) Berbasis Kriteria Rotterdam
Diaz Astari Beningati, Prof. Dr.Eng. Ir. Adha Imam Cahyadi, M.Eng., IPM., ASEAN Eng., SMIEEE.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR TABEL

Tabel 2.1 Rotterdam Consensus Criteria Revision [1] 11
Tabel 2.2 Pemeriksaan PCOS, Indikator, dan Relevansinya 11
Tabel 2.3 Metode Penelitian Terdahulu 26

Tabel 3.1 Atribut *Case Base* (data mentah, 44 atribut)..... 29
Tabel 3.2 Atribut Hasil *Feature Engineering*..... 36
Tabel 3.3 Persebaran Phenotype 37
Tabel 3.4 Struktur CDT dan Bobot Level 38
Tabel 3.5 Konfigurasi ukuran kemiripan pada evaluasi *retrieval*..... 40
Tabel 3.6 Elemen strategi *prompting* dan dasarnya. 42
Tabel 3.7 Struktur *prompt*. 43

Tabel 4.1 Sensitivitas dan spesifisitas per fenotipe pada $k = 1$ (dalam persen). 48
Tabel 4.2 Hasil evaluasi *retrieval* pada $k = 5$ (dalam persen)..... 49
Tabel 4.3 *Recall* dan F1 per fenotipe pada $k = 5$ (dalam persen). 49
Tabel 4.4 *Confusion matrix* fenotipe $\text{HWHSM}_{\text{CDT}}$ pada $k = 5$. 51
Tabel 4.5 Pengaruh urutan data terhadap $\text{HWHSM}_{\text{CDT}}$ pada $k = 5$..... 51
Tabel 4.6 Tingkat halusinasi LLM-*only* dan CBR+LLM ($n = 541$, dalam persen, CI 95%)..... 53
Tabel 4.7 Tabel kontingensi uji McNemar LLM-*only* dan CBR+LLM ($n = 541$)..... 54
Tabel 4.8 Angka yang tidak terdapat pada teks sumber..... 55
Tabel 4.9 Tingkat halusinasi intrinsik per proposisi (dalam persen). 56
Tabel 4.10 Kesalahan kesimpulan pemenuhan kriteria menurut jumlah kriteria sebenarnya. 57
Tabel 4.11 Kinerja kesimpulan antar-arsitektur (dalam persen). 57
Tabel 4.12 Presisi dan *recall* per fenotipe (dalam persen). 58
Tabel 4.13 Rerata $\pm$ simpangan baku antar-run (tiga run, dalam persen). 62
Tabel 4.14 Nilai p_{Holm} uji McNemar LLM-*only* vs CBR+LLM pada setiap run.. 62
Tabel 4.15 Perbandingan format masukan terstruktur dan naratif (dalam persen). 63
Tabel 4.16 Uji McNemar format terstruktur vs naratif dengan koreksi Holm. ... 63

Tabel L.1 Aturan pelabelan basis kasus..... L-1
Tabel L.2 Ambang numerik (*threshold*) dan rujukan klinis..... L-2
Tabel L.3 Data masukan pasien 149. ..... L-3
Tabel L.4 Lima kasus termirip HWHSM untuk pasien 149 (masukan CBR-*only* dan bagian *Kasus serupa* pada *prompt* CBR+LLM). L-3
Tabel L.5 Keluaran per proposisi LLM-*only* dan CBR+LLM untuk pasien 149. Tanda $*$ = bertentangan dengan label basis kasus. L-4

Grounding Case-Based Reasoning (CBR) untuk Mitigasi Halusinasi Large Language Model (LLM) pada
Sistem Pendukung Keputusan Klinis Polycystic Ovary Syndrome (PCOS) Berbasis Kriteria Rotterdam
Diaz Astari Beningati, Prof. Dr. Ir. Lukito Edi Nugroho, M.Sc., IPM., ASEAN Eng., SMIEEE.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR GAMBAR

Gambar 2.1 Algoritma Diagnosis PCOS [2] 14
Gambar 3.2 Alur kerja penelitian. Kotak bergaris putus-putus menunjukkan eksperimen tambahan, sedangkan panah putus-putus menunjukkan acuan modul kriteria dan audit pemeriksa. 32
Gambar 4.3 Akurasi *reuse* fenotipe terhadap nilai k. 50