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/

DAFTAR PUSTAKA

[1] The Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group, "Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (pcos)," *Human Reproduction*, vol. 19, no. 1, pp. 41–47, 1 2004. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/14688154/

[2] H. Teede, Chau Thien Tay, Joop S.E. Laven, Anuja Dokras, L. J. Moran, Terhi Piltonen, M. Costello, J. Boivin, L. M. Redman, J. Boyle, R. J. Norman, A. Mousa, and A. E. Joham, "Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome," *European journal of endocrinology*, vol. 189, no. 2, pp. G43–G64, 8 2023.

[3] World Health Organization, "Polycystic ovary syndrome," World Health Organization, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/polycystic-ovary-syndrome

[4] A. Z. Zhang, M. Sehrt, S. A. Ley, K. V. Ross, P. Phimphasone-Brady, and K. M. McKenney, "Gaps and opportunities in polycystic ovary syndrome care—a qualitative study," *F&S Reports*, vol. 6, no. 3, pp. 357–365, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666334125000947

[5] T. Baba, "Polycystic ovary syndrome: Criteria, phenotypes, race and ethnicity," *Reproductive Medicine and Biology*, vol. 24, no. 1, 1 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11751892/

[6] M. Ghaderzadeh, A. Garavand, and C. Salehnasab, "Artificial intelligence in polycystic ovary syndrome: a systematic review of diagnostic and predictive applications," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, 11 2025.

[7] J. Zhou, H. Li, S. Chen, Z. Chen, Z. Han, and X. Gao, "Large language models in biomedicine and healthcare," *npj Artificial Intelligence*, vol. 1, no. 1, 12 2025. [Online]. Available: https://www.nature.com/articles/s44387-025-00047-1

[8] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung, "Survey of hallucination in natural language generation," *ACM Computing Surveys*, vol. 55, no. 12, pp. 248:1–248:38, 2023.

[9] Z. Xu, S. Jain, and M. Kankanhalli, "Hallucination is inevitable: An innate limitation of large language models," arXiv.org, 1 2024. [Online]. Available: https://arxiv.org/abs/2401.11817

[10] I. Gunesli, S. Aksun, J. Fathelbab, and B. O. Yildiz, "Comparative evaluation of chatgpt-4, chatgpt-3.5 and google gemini on pcos assessment and management based on recommendations from the 2023 guideline," *Endocrine*, vol. 88, no. 1, pp. 315–322, 12 2024.

[11] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *AI Communications*, vol. 7, no. 1, pp. 39–59, 1994.

[12] H. Fauvillâtre, V. Auffret, M. Castro, F. Leclerc, H. Le Breton, M. Garreau, and P. Haigron, "Similarity measures and attribute selection for case-based reasoning in transcatheter aortic valve implantation," *PLOS ONE*, vol. 15, no. 9, p. e0238463, 9 2020.

[13] D. R. Wilson and T. R. Martinez, "Improved heterogeneous distance functions," *Journal of Artificial Intelligence Research*, vol. 6, pp. 1–34, 1997.

[14] I. Watson, "Case-based reasoning is a methodology not a technology," *Knowledge-Based Systems*, vol. 12, no. 5-6, pp. 303–308, 10 1999.

[15] S. K. Pal, S. Shiu, and Simon, *Foundations of Soft Case-Based Reasoning*. Wiley-Interscience, 2004.

[16] T. W. Liao, Z. Zhang, and C. R. Mount, "Similarity measures for retrieval in case-based reasoning systems," *Applied Artificial Intelligence*, vol. 12, no. 4, pp. 267–288, 6 1998.

[17] S. H. El-Sappagh and M. Elmogy, "Case based reasoning: Case representation methodologies," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 11, pp. 192–208, 2015. [Online]. Available: https://www.ijacsa.thesai.org

[18] J. Ramos-González, D. López-Sánchez, J. A. Castellanos-Garzón, Juan, and J. M. Corchado, "A cbr framework with gradient boosting based feature selection for lung cancer subtype classification," *Computers in Biology and Medicine*, vol. 86, pp. 98–106, 5 2017.

[19] C. Xu, W. Liu, Y. Chen, and X. Ding, "A supervised case-based reasoning approach for explainable thyroid nodule diagnosis," *Knowledge-Based Systems*, vol. 251, p. 109200, 9 2022.

[20] C. L. Bartlett, I. Bichindaritz, and G. Liu, "Predicting with confidence: A case-based reasoning framework for predicting survival in breast cancer," *The International FLAIRS Conference Proceedings*, vol. 34, no. 1, 4 2021.

[21] A. Firomsa Wakuma, "Advanced tuberculosis diagnosis system: Integrating case-based reasoning with nearest neighbor algorithm," *The Indian journal of tuberculosis*, vol. 72, no. 4, pp. 506–512, 10 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/40975582/

[22] Fatemeh Rangraz Jeddi, Mansooreh Momen-Heravi, Razieh Farrahi, Ehsan Nabovati, H. Akbari, and M. E. khodabandeh, "Computer case-based reasoning simulation versus traditional lectures for medical interns teaching of diagnosis and antibiotic prescribing for acute respiratory infection: a comparative quasi-experimental study," *BMC Medical Education*, vol. 24, no. 1, 12 2024.

[23] E. M. Mustafa, M. M. Saad, and Lydia Wahid Rizkallah, "Building an enhanced case-based reasoning and rule-based systems for medical diagnosis," *Journal of Engineering and Applied Science*, vol. 70, no. 1, 11 2023.

[24] L. Qin, Q. Chen, X. Feng, Y. Wu, Y. Zhang, Y. Li, M. Li, W. Che, and P. S. Yu, "Large language models meet nlp: a survey," *Frontiers of Computer Science*, vol. 20, no. 11, 3 2026.

[25] Z. Arkaitz, "Natural language processing in the era of large language models," *Frontiers in artificial intelligence*, vol. 6, 1 2024.

[26] L. Dong, N. Yang, W. Wang, F. Wei, X. Liu, Y. Wang, J. Gao, M. Zhou, and H.-W. Hon, "Unified language model pre-training for natural language understanding and generation," Tech. Rep., 2019. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2019/file/c20bb2d9a50d5ac1f713f8b34d9aac5a-Paper.pdf

[27] B. Min, H. Ross, Elior Sulem, B. Veyseh, Thien Huu Nguyen, O. Sainz, E. Agirre, I. Heintz, and D. Roth, "Recent advances in natural language processing via large pre-trained language models: A survey," *ACM Computing Surveys*, vol. 56, no. 2, pp. 1–40, 6 2023.

[28] S. Maity and Saikia, Manob Jyoti, "Large language models in healthcare and medical applications: A review," *Bioengineering*, vol. 12, no. 6, p. 631, 6 2025. [Online]. Available: https://www.mdpi.com/2306-5354/12/6/631

[29] M. Zeba, A. Al Mamun, K. Jahan Tithee, D. Sutradhar, M. A. Khan Raiaan, S. Mukta, R. E. Mohamed, M. R. Islam, Y. Sebastian, M. Hussain, and S. Azam, "Mitigating hallucinations in healthcare llms with granular fact-checking and domain-specific adaptation," *Expert Systems with Applications*, p. 132966, 5 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417426018786

[30] S. Heo, S. Son, and H. Park, "Halucheck: Explainable and verifiable automation for detecting hallucinations in llm responses," *Expert Systems with Applications*, vol. 272, p. 126712, 5 2025.

[31] A. Saha, B. Gupta, A. Chatterjee, and K. Banerjee, "You believe your llm is not delusional? think again! a study of llm hallucination on foundation models under perturbation," *Discover Data*, vol. 3, p. 20, 2025. [Online]. Available: https://doi.org/10.1007/s44248-025-00041-7

[32] T. Savage, A. Nayak, R. Gallo, E. Rangan, and J. H. Chen, "Diagnostic reasoning prompts reveal the potential for large language model interpretability in medicine," *npj Digital Medicine*, vol. 7, p. 20, 2024. [Online]. Available: https://www.nature.com/articles/s41746-024-01010-1

[33] Y. Sonoda, R. Kurokawa, A. Hagiwara *et al.*, "Structured clinical reasoning prompt enhances llm's diagnostic capabilities in diagnosis please quiz cases," *Japanese Journal of Radiology*, vol. 43, no. 4, pp. 586–592, 2025. [Online]. Available: https://doi.org/10.1007/s11604-024-01712-2

[34] M. Ayoub, H. Zhao, L. Li, D. Yang, S. Hussain, and J. A. Wahid, "Structured clinical approach to enable large language models to be used for improved clinical diagnosis and explainable reasoning," *Communications Medicine*, vol. 6, p. 86, 2026. [Online]. Available: https://www.nature.com/articles/s43856-025-01348-x

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/

[35] D. Trujillo, D. Wang, N. Bahr, T. Yi-Jia Hsieh, B. Cho, G. Meckler, M. Hansen, C. Eriksson, K. S. Kim, S. Bedrick, X. Jiang, and J.-M. Guise, "A medically grounded llm agent-based tool to detect patient safety events in medical records: Identifying patient safety events with large language models," *medRxiv : the preprint server for health sciences*, p. 2025.12.16.25342438, 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/41445638/

[36] N. Wiratunga, R. Abeyratne, L. Jayawardena, K. Martin, S. Massie, I. Nkisi-Orji, R. Weerasinghe, A. Liret, and B. Fleisch, "Cbr-rag: Case-based reasoning for retrieval augmented generation in llms for legal question answering," arXiv.org, 2024. [Online]. Available: https://arxiv.org/abs/2404.04302

[37] G. Liu, Y. Zhang, X. Liu, and Q. Yao, "Case-based reasoning enhances the predictive power of llms in drug-drug interaction," arXiv.org, 2025. [Online]. Available: https://arxiv.org/abs/2505.23034

[38] W. Gan, M.-S. Dao, and K. Zettsu, "Case-based reasoning augmented large language model framework for decision making in realistic safety-critical driving scenarios," arXiv.org, 2025. [Online]. Available: https://arxiv.org/abs/2506.20531

[39] S. Guo, C. Deng, Y. Wen, H. Chen, Y. Chang, and J. Wang, "Ds-agent: Automated data science by empowering large language models with case-based reasoning," arXiv.org, 5 2024. [Online]. Available: https://arxiv.org/abs/2402.17453

[40] R. Yang, "Casegpt: a case reasoning framework based on language models and retrieval-augmented generation," *arXiv (Cornell University)*, 7 2024.

[41] B. Devranoglu, T. Gurbuz, and O. Gokmen, "Chatgpt's efficacy in queries regarding polycystic ovary syndrome and treatment strategies for women experiencing infertility," *Diagnostics*, vol. 14, no. 11, p. 1082, 5 2024.

[42] Y. Zhu, D. Luo, X. Shen, Q. Shi, S. Zhang, F. Ye, and N. Kong, "Application of chatgpt-based artificial intelligence in the diagnosis and management of polycystic ovary syndrome," *BMC Medical Informatics and Decision Making*, vol. 25, p. 271, 7 2025.

[43] M. Venkatesh, C. Umayal, P. Chitra, B. Kiruthika, and A. Dharini, "Pcos-gpt: A rag enhanced llm for personalized ultrasound based pcos diagnosis," in *2025 IEEE 17th International Conference on Computational Intelligence and Communication Networks (CICN)*, 2025, pp. 599–605.

[44] A. S. Jamil, S. K. Alalaf, N. G. Al-Tawil, and T. Al-Shawaf, "A case-control observational study of insulin resistance and metabolic syndrome among the four phenotypes of polycystic ovary syndrome based on rotterdam criteria," *Reproductive Health*, vol. 12, p. 7, 2015. [Online]. Available: https://doi.org/10.1186/1742-4755-12-7

[45] E. Carmina and R. A. Lobo, "Comparing lean and obese pcos in different pcos phenotypes: Evidence that the body weight is more important than the rotterdam phenotype in influencing the metabolic status," *Diagnostics*, vol. 12, no. 10, p. 2313, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9600591/

UNIVERSITAS
GADJAH MADA

Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

[46] R. Dadachanji, A. Patil, R. Joshi, and S. Mukherjee, "Elucidating the impact of obesity on hormonal and metabolic perturbations in polycystic ovary syndrome phenotypes in indian women," *PLoS One*, vol. 16, no. 2, p. e0246862, 2 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0246862

[47] L. Rasquin and J. V. Mayrin, "Polycystic ovarian disease," Nih.gov, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK459251/

[48] A. Oroan, P. Minz, A. Kujur, and B. R. Ekka, "Evaluation of thyroid function status among diagnosed cases of polycystic ovarian syndrome in a tribal-preponderant state in india," *Cureus*, vol. 17, no. 5, 5 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12143108/

[49] M. Al-Chalabi and I. Alsalman, "Physiology, prolactin," National Library of Medicine, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK507829/

[50] National Institution of Health (NIH), "Evidence-based methodology workshop on polycystic ovary syndrome executive summary," National Institution of Health (NIH), Tech. Rep., 12 2012. [Online]. Available: https://prevention.nih.gov/sites/g/files/mnhszr241/files/2018-06/FinalReport.pdf

[51] M. Basma Bentaiba-Lagrid, L. Bouzar-Benlabiod, S. H. Rubin, T. Bouabana-Tebibel, and M. R. Hanini, "A case-based reasoning system for supervised classification problems in the medical field," *Expert Systems with Applications*, vol. 150, p. 113335, 7 2020. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417420301603

[52] G. Gan, C. Ma, and J. Wu, *Data Clustering*. SIAM, 1 2007.

[53] A. Levy, B. R. Shalom, and M. Chalamish, "A guide to similarity measures and their data science applications," *Journal of Big Data*, vol. 12, no. 1, 7 2025.

[54] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, vol. 12, no. 2, pp. 153–157, 1947.

[55] P. Kottarathil, "Polycystic ovary syndrome (pcos)," Kaggle Dataset, 2020. [Online]. Available: https://www.kaggle.com/prasoonkottarathil/polycystic-ovary-syndrome-pcos

[56] A. D. Association, "Standards of medical care in diabetes," *Diabetes care*, vol. 28 Suppl 1, pp. S4–S36, 1 2005. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/15618112/

[57] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979.

[58] World Health Organization, "Obesity: Preventing and managing the global epidemic. report of a WHO consultation," World Health Organization, Geneva, WHO Technical Report Series 894, 2000.