

Utilizing IndoBERT for Efficient University Department Chatbot Development

Najma Rafifah Putri Syallya¹, Anindya Aprilianti Pravitasari^{*2}, Afrida Helen³, Ni Gusti Ayu Putu Meyrasinta Susila⁴

^{1, 2, 4} Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, INDONESIA.

³ Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, INDONESIA.

(E-mail: najma20005@mail.unpad.ac.id, anindya.aprilianti@unpad.ac.id, helen@unpad.ac.id, gusti20002@mail.unpad.ac.id)

Keywords: Chatbot, IndoBERT, University, NLP.

ABSTRACT

In recent years, Natural Language Processing (NLP) techniques have gained significant traction across various applications, particularly in the development of intelligent chatbots. Chatbots have emerged as efficient tools for offering information and assistance to users, making them an ideal solution for universities and academic departments aiming to streamline their communication and support services [1]. This paper delves into the utilization of IndoBERT [2], a contextual language model designed for the Indonesian language, in creating an intent classification-based chatbot for a university department.

This study commences by introducing the concept of chatbots and highlighting their role in enhancing the user experience within university environments. Leveraging IndoBERT, which has been pre-trained on a substantial dataset of Indonesian text, allows us to harness the power of contextual language comprehension. This capability empowers the chatbot to understand nuanced queries and engage in meaningful and coherent conversations with users in the Indonesian language [3].

This study delves into the essential components of university chatbot, with a primary focus on intent recognition. Employing supervised learning and fine-tuning with domain-specific, limited data, we leverage IndoBERT to comprehend and respond to a broad spectrum of user inquiries, including those related to course information, credit details, faculty information, and other department-specific queries.

The experiment results demonstrate that IndoBERT achieves an accuracy rate of over 90% and effectively responds to the provided questions. This outcome substantiates that the chatbot serves as a valuable resource for both students and faculty, providing timely and precise information while concurrently reducing the administrative workload on university

staff.

Acknowledgment: The authors are grateful to the Research Center for AI and Big Data (AIDA) Universitas Padjadjaran and the Department of Statistics Universitas Padjadjaran which supports this research.

References:

- [1] Wolff, R., Hobert, S., Masuch, K., & Schumann, M.: ‘Chatbots at Digital Workplaces - A Grounded-Theory Approach for Surveying Application Areas and Objectives’. *Pac. Asia J. Assoc. Inf. Syst.* 2020. 12, 3. <https://doi.org/10.17705/1PAIS.12203>.
- [2] Koto, F., Rahimi, A., Lau, J. H., and Baldwin, T.: ‘IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP’. 2020. <http://arxiv.org/abs/2011.00677>
- [3] Richardson, B and Wicaksana, A.: ‘Comparison Of Indobert-Lite and Roberta in Text Mining for Indonesian Language Question Answering Application’. *International Journal of Innovative Computing, Information and Control*. 2022. 18. (6). pp. 1719–1734. <http://www.ijicic.org/ijicic-180604.pdf>