Chatbot-Based Student Information Service in Indonesian Language

I Wayan Suasnawa, I Komang Wiratama, I Ketut Gede Sudiartha, I Gusti Ngurah Bagus Caturbawa, Anak Agung Ngurah Gde Sapteka and I Nyoman Eddy Indrayana
Department of Electrical Engineering, Politeknik Negeri Bali, Kampus Bukit Jimbaran, Badung, Indonesia

Keywords: Chatbot, Information Service, Rasa Framework, Academic Purpose.

Abstract: Academic services are the most important part in higher education. The services provided by the university will be a picture of the quality of the institution. Service is said to be of high quality if it meets the needs of its customers. The demands on higher education today are not only limited to the ability to produce good graduates measured by academic achievement alone, but the entire service program of higher education institutions, one of which is the academic services provided to students. Current technological developments can help universities in making it easier for students to get academic information. To support the need for accurate and up-to-date information related to academic information, it is possible to utilize an information technology-based system that can properly summarize data and display the information to students. In this study, chatbot development can be a solution in providing information and providing academic services. Chatbot developed using Rasa Core is a chatbot framework or open source chat framework for handling contextual conversations, used for machine learning-based conversation management. The results of the NLU evaluation show that the chatbot can understand user questions. This is indicated by a precision value of 0.955, a recall of 0.962 and an F1-Score of 0.962. Meanwhile, the evaluation of the model has an accuracy of 0.82, a precision value of 0.85 and an F1-Score of 0.85. This is a benchmark for evaluating chatbot performance in predicting a good response for users.

1 INTRODUCTION

Academic services are the most important part in higher education. The services provided by the university will be a picture of the quality of the institution. Service is said to be of high quality if it meets the needs of its customers. The demands on higher education today are not only limited to the ability to produce good graduates measured by academic achievement alone, but the entire service program of higher education institutions, one of which is academic services provided to students.

Current technological developments can help universities in making it easier for students to get academic information. To support the need for accurate and up-to-date information related to academic information, an information technology-based system is needed that can properly summarize the data and display the information to students. Chatbot development can be a solution in providing information and providing academic services.

Chatbot is a computer program designed to simulate an interactive conversation or communication to customers (humans) through the form of text, voice, and/or visuals. Conversations that occur between computers and humans are a form of response from programs that have been declared in the program database on the computer (Mashud, et al., 2019). In implementing a chatbot so that the system can respond dynamically to user queries, the use of Natural Language Processing plays a very important role, namely understanding user queries in natural language. Chatbot which is often known as Artificial Conversational Entity bot or Chatterbox is a computer program that is able to imitate human conversations using the NLP method (Zuraiyah, et al., 2019).

2 THEORY

2.1 Chatbot

Chatbot technology is one form of application with Natural Language Processing (NLP), NLP itself is one of the fields of Artificial Intelligence that studies

Suasnawa, I., Wiratama, I., Sudiartha, I., Caturbawa, I., Sapteka, A. and Indrayana, I.
Chatbot-Based Student Information Service in Indonesian Language.
DOI: 10.5220/0011753800003575
In *Proceedings of the 5th International Conference on Applied Science and Technology on Engineering Science (iCAST-ES 2022)*, pages 223-227
ISBN: 978-989-758-619-4; ISSN: 2975-8246

communication between humans and computers through natural language (Afrianto, et al., 2019). This application is famous for its automated conversational agents that run on computer programming or some kind of Artificial Intelligence (AI) interaction between the user and the machine with the intervention of Natural Language Processing. Chatbots have the potential to be called the most promising and sophisticated form of human-machine interaction (Battineni, et al., 2020). NLP has many purposes that can assist human communication, such as machine translation and assist human machine communication, such as with conversational agents and others (Aleedy, et al., 2019). Chatbot is a technology whose main purpose is to interact with human users by processing natural language input and generating relative output through a rule-driven machine or artificial intelligence engine (Indrayani, et al., 2020). Natural language processing uses tokenizing, filtering, and analysis stages and applies the knuth morris prrat algorithm (Amrizal, et al., 2019).

The use of artificial intelligence technology has made chatbots more advanced, including natural language processing and machine learning so as to provide accurate results when interacting with bots (Ayanouz, et al., 2020). The chatbot developed uses natural language processing so that the system can understand user queries in natural language (Elcholiqi, et al., 2020). Chatbots are able to communicate with website visitors and chatbots can be optimized in communication (Herwin, et al., 2019). From the results of this study it can be concluded that NLP has unique features with an excellent communication approach (Shruthi, et al., 2020). Chatbot application to 10 examiners, the result of the level of suitability of answers with user input is 84% (Khoirunisa et al., 2020). These findings suggest that the NLP/ML method can be used to be able to differentiate stroke features from big data groups for clinical investigations and related research (Ong, et al., 2020).

2.2 Natural Language Processing

Natural language processing (NLP) is a programming technique where computers can understand and provide output in the form of human language or simply facilitate communication between humans and machines. The purpose of NLP is to provide an appropriate answer or response based on machine understanding of the meaning of human language (Herwin, 2019 - Ong, et al., 2007). The use of NLP has been applied in various fields of human life. This is because NLP is easier to use as a computer interface than learning the language of computer commands. Elements in natural language processing are parser, lexicon, understander, knowledge base, and generator. The parser is the part that identifies each word. Lexicon is a collection of words recognized by the program. The understander is the part that determines the meaning of a sentence. Knowledge base is a knowledge base that contains words and phrases. Generator is the output that is generated based on the input that has been processed.

2.3 Rasa.AI Framework

Rasa.ai is an open source machine learning framework for text-based or spoken intelligent conversation. Rasa.ai can understand user input, hold conversations with users and connect with communication platforms and APIs. Rasa.ai works on two main components namely Rasa NLU and Rasa Core. Rasa NLU is an open source natural language processing tool, used for intent classification and entity extraction in conversations, then using machine learning to pick up patterns and generalize to invisible sentences. Rasa Core is an open source chatbot framework for handling contextual conversations, used for machine learning-based conversational management.

Rasa.ai is an open source machine learning framework for building AI assistants and contextual chats. Rasa.ai also has a user interface platform namely Rasa X. Rasa X is a tool designed for use that helps software developers to build, improve and deploy AI assistants supported by the rasa framework. Rasa.ai works on two main components namely Rasa NLU and Rasa Core. Rasa Core is an open source chatbots framework for handling contextual conversations, used for dialog management to hold conversations and decide what to do next. Rasa NLU is an open source natural language processing tool, used for intent classification and entity extraction in conversations, then using machine learning to pick up patterns and generalize to invisible sentences.

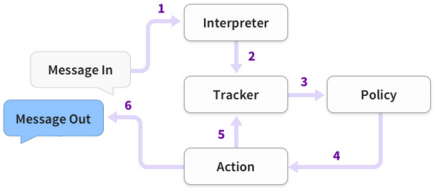

Figure 1: Architecture Message Handling of Rasa.

3 RESEARCH METHODOLOGY

3.1 Data Collection

In this study, a chatbot was developed to provide information related to questions that are often asked by students to study program managers. The data needed to build a chatbot in this study are sample dialogs and intents. Frequently asked questions by students are the datasets used to develop Chabot for academic services. The data contains data on questions that are often asked by students along with answers from the study program manager. Before entering the processing stage, the initial processing of the question data is carried out, the intent and entity definitions are carried out. The definition of intent consists of identifying the name of the intent and categorizing or labeling the question sentence based on the name of the intent that has been defined, as shown in Figure 2. Besides the NLU data as a dataset, a list of responses or answers that can be given by the chatbot is also prepared, as shown in Figure 3 .

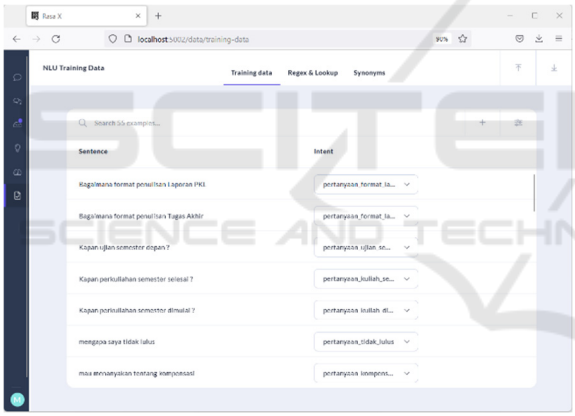

Figure 2: Data NLU.

3.2 Conversation Modeling

Chatbot development requires the preparation of basic knowledge that represents the domain of the chatbot. The domain is required as part of the learning environment of the chatbot. This domain includes intent types, actions, and sentence templates for speech responses to user messages. At this stage, the initial domain design is carried out based on data from questions that are often asked by students. The domain specifies the training data that will generate

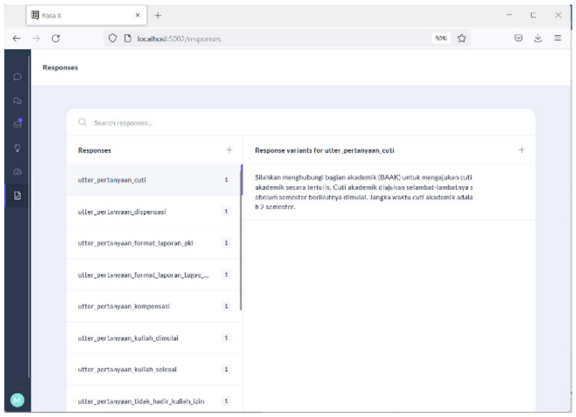

Figure 3: Data Responses to build dialogs.

the model for the chatbot. The training data for the chatbot consists of NLU training data and dialogue training. The quality of this training data can be continuously improved so as to produce a chatbot model that can respond well to messages and information needs from users. This Conversation Modeling in the Rasa Framework is done by formulating Rules and Stories.

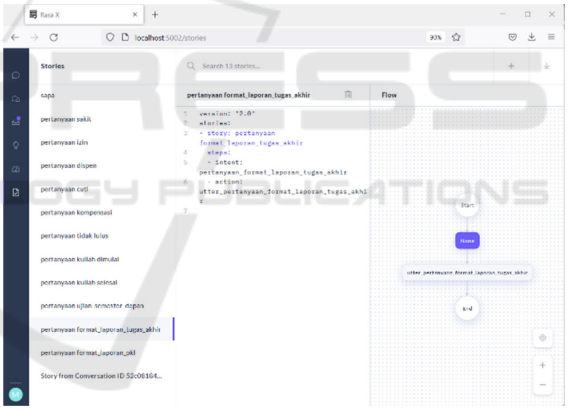

Figure 4: Conversation Modeling using Stories.

3.3 Interactive Learning

The use of rules and stories provides knowledge for chatbots to form the required conversation model. To get more natural conversation results, interactive learning can be done. With interactive learning will provide the possibility of variations of questions that can be given by students. From the variety of questions, it is expected that the responses given by the chatbot are also more in line with minimal errors. This interactive learning process is as shown in Figure 5.

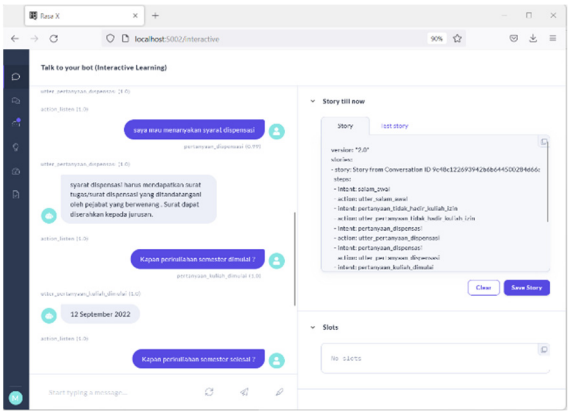

Figure 5: Conversation Modeling using Stories.

4 RESULTS AND DISCUSSION

The system chatbot is built according to the Sense Open Source design and framework. In addition, at this stage also use Rasa as a tool for developing conversations with real users. The chatbot development process on a local computer using Rasa Open Source consists of the main processes, namely: framework initialization and configuration, NLU data definition, chatbot response definition, dialog data definition, training and testing. The configuration process involves defining the language type, pipeline specifications and policies.

The training process uses machine learning algorithms specified in the policies section. This section defines the machine learning process that will be used to process received text messages and response text messages to users. The chatbot system is built to serve messages from real users. To that end, the implementation of the system interface uses Sense X to assist with model implementation and test live conversations with users. Users can directly interact with the chatbot via a link created by the administrator of Rasa X. Further implementation, the chatbot system can be connected with other chat applications such as Telegram and others. The appearance of a chatbot that shows a human character can also affect the quality of its implementation. Before being used by end users, the implementation of the chatbot should be tested in a local environment, to ensure that the training process, the resulting model and the implementation of the system are error free. At this stage, training data validation is also carried out to ensure the training data has the correct structure.

The Open Source Flavor framework has provided a feature to automatically evaluate the chatbot model. The types of tests carried out include NLU testing, dialogue testing and conversation testing by actual users. The test data comes from conversational data collected by Rasa and then validated by the system developer. The conversation data captured by Rasa is used as the test data needed to execute NLU tests and dialogs automatically. The test results will generate reports on the accuracy and precision of chatbot conversations.

5 CONCLUSIONS

The conclusion of this study is that the creation of a student information service chatbot has been made by identifying questions that are often asked by students, including questions about academic regulations, lecture implementation, academic guidelines. The chatbot has been able to work according to the conversation flow that has been arranged, starting with an opening message, then students can ask questions and the chatbot will provide responses based on the model that has been generated during training. The results of the chatbot test are carried out by evaluating the accuracy and precision, it is found that the rules that have been run have been able to provide a good response. The results of the NLU evaluation show that the chatbot can understand user questions. This is indicated by a precision value of 0.955, a recall of 0.962 and an F1-Score of 0.962. Meanwhile, the evaluation model has an accuracy of 0.82, a precision value of 0.85 and an F1-Score of 0.85. This is a benchmark for evaluating chatbot performance in predicting a good response for users.

REFERENCES

Mashud, M., & Wisda, W. (2019). Aplikasi Chatbot Berbasis Website sebagai Virtual Personal Assistant dalam Pemasaran Properti. Inspiration: Journal Teknologi Informasi Dan Komunikasi, 9(2), 99.

Zuraiyah, T. A., Utami, D. K., & Herlambang, D. (2019). Implementasi Chatbot pada Pendaftaran Mahasiswa Baru Menggunakan Recurrent Neural Network. Journal Ilmiah Teknologi Dan Rekayasa, 24(2), 91–101.

Afrianto, I., Irfan, M. F., & Atin, S. (2019). Aplikasi Chatbot Speak English Media Pembelajaran Bahasa Inggris Berbasis Android. Komputika: Jurnal Sistem Komputer, 8(2), 99–109.

Battineni, G., Chintalapudi, N., & Amenta, F. (2020). AI Chatbot Design during an Epidemic like the Novel Coronavirus. Healthcare, 8(2), 154.

Aleedy, M., Shaiba, H., & Bezbradica, M. (2019). Generating and Analyzing Chatbot Responses using

Natural Language Processing. International Journal of Advanced Computer Science and Applications, 10(9).

Indrayani, L. ., Amalia, R. M., & Hakim, F. Z. M. (2020). Emotive Expressions on Social Chatbot. Jurnal Sosioteknologi, 18(3), 509.

Amrizal, V., Munandar, A., &Arini, A.-. (2019). Identifikasi Matan Hadits Menggunakan Natural Language Processing Dan Algoritma Knuth Morris Pratt Berbasis Web. Journal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi, 5(2), 56.

Ayanouz, S., Abdelhakim, B. A., & Benhmed, M. (2020). A Smart Chatbot Architecture based NLP and Machine Learning for Health Care Assistance. Proceedings of the 3rd International Conference on Networking, Information Systems & Security.

Elcholiqi, A., & Musdholifah, A. (2020). Chatbot in Bahasa Indonesia using NLP to Provide Banking Information. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 14(1), 91.

Herwin, H. H. (2019). Super Agent Chatbot "3S" Sebagai Media Informasi Menggunakan Metoda Natural Language Processing(NLP). Jurnal Teknologi Dan Open Source, 2(1), 53–64.

Shruthi J., Suma Swamy. (2020). A prior case study of natural language processing on different domain. International Journal of Electrical and Computer Engineering (IJECE), 10(5), 4928.., & Swamy, S. (2020). A prior case study of natural language processing on different domain. International Journal of Electrical and Computer Engineering (IJECE), 10(5), 4928.

Khoirunisa, R. (2020). Penggunaan Natural Language Processing Pada Chatbot Untuk Media Informasi Pertanian. Indonesian Journal of Applied Informatics, 4(2), 55.

Ong , C. J., Orfanoudaki, A., Zhang, R., Caprasse, F. P.M., Hutch, M., Ma, L., ... Bertsimas, D. (2020). Machine learning and natural language processing methods to identify ischemic stroke, acuity and location from radiology reports. PLOS ONE, 15(6).