IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.19, No.3, July 2025, pp. 259~270
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: https://doi.org/10.22146/ijccs.95900

Development of Chatbot For Pre-Diagnosis and Recommendation of Anxiety Disorder Using Diet and Sentence Transformer Models

Angel Berta Desi Suryanti[1], Edi Winarko[*2]
[1,2]Department of Computer Science and Electronics, FMIPA UGM, Yogyakarta, Indonesia
e-mail:[1]angelbertadesisuryanti@mail.ugm.ac.id, *[2]**ewinarko@ugm.ac.id**

Abstrak

Penelitian sebelumnya tentang chatbot untuk pre-diagnosis dan rekomendasi anxiety disorder masih sebatas alat bantu terapi. Dengan membandingkan model NLU DIET dan LogisticRegressionClassifier, sistem chatbot ini dibuat untuk menghitung tingkat kecemasan menggunakan metode GAD-7, DASS dan STAIT/STAIS-5 bersama dengan model semantic similarity yaitu Sentence Transformer atau biasa disebut SBERT yang digunakan sebagai sistem rekomendasi.

Hasil pengujian klasifikasi intent untuk model NLU DIETClassifier dengan nilai akurasi sebesar 95% sementara LogisticRegressionClassifier sebesar 99%. Model Dialog mempunyai akurasi sebesar 68% yaitu menggunakan TEDPolicy. Pengujian dilakukan kepada 35 koresponden yang terdiri dari pelajar, mahasiswa dan pekerja secara acak. Dari hasil interaksi ini dari 35 koresponden dihitung nilai MAP model rekomendasi SBERT sebesar 30% dan 26% untuk 2 model yang digunakan yaitu indobert base dan paraphrase-multilingual-MiniLM-L12-v2. Serta nilai rata-rata kepuasan dan kinerja sistem chatbot sebesar 3.7 dari 5.

Penelitian ini dapat menjawab permasalahan yang ada bahwa sebuah prototype chatbot dapat digunakan untuk pre-diagnosis dan rekomendasi anxiety disorder dengan model NLU terbaik yaitu LogisticRegressionClassifier sebesar 99% dan model dialog dengan akurasi 68%. Sementara itu sistem rekomendasi masih mempunyai MAP yang kecil karena referensi data yang digunakan bukan data klinis yang valid. Hal ini dapat dijadikan sebagai ruang perbaikan untuk penelitian selanjutnya.

Kata kunci*— Chatbot, NLP, SBERT, Anxiety, Rasa*

Abstract

Previous research on chatbots for pre-diagnosis and recommendation of anxiety disorders has been limited to therapy aids. Comparing NLU DIET and LogisticRegressionClassifier models, this chatbot system calculates anxiety levels using GAD-7, DASS, and STAIT/STAIS-5 methods along with Sentence Transformer (SBERT) for semantic similarity.

Intent classification testing yielded 95% accuracy for NLU DIETClassifier and 99% for LogisticRegressionClassifier. The Dialog Model achieved 68% accuracy with TEDPolicy. Testing involved 35 randomly selected respondents, including students and workers. From their interactions, the SBERT recommendation model scored 30% MAP, 26% for the Indobert base and paraphrase-multilingual-MiniLM-L12-v2 models.

The average satisfaction and performance rating for the chatbot system was 3.7 out of 5. This research addresses the need for a prototype chatbot for pre-diagnosis and recommendation of anxiety disorders, with the best NLU model being LogisticRegressionClassifier at 99% accuracy and the dialog model at 68%. However, the

Received May 4th,2024; Revised July 19th, 2025; Accepted July 31th, 2025

recommendation system still has a low MAP due to the use of non-valid clinical data as references, suggesting room for improvement in future research.

Keywords— *Chatbot, NLP, SBERT, Anxiety, Rasa*

1. INTRODUCTION

Mental well-being is a fundamental necessity, encapsulated in the Latin phrase "*mens sana in corpore sano*," emphasizing the connection between a healthy mind and body. It's approached positively, focusing on happiness, life satisfaction, and love, and negatively, aiming to prevent conditions like anxiety and depression [1][2][3][4][5]. Assessing mental health aids in determining the need for counseling or consultation. Data from the Indonesian Suicide Prevention Association reveals a concerning number of unreported suicide attempts due to limited access to psychological care [6]. Early assessment of anxiety levels is crucial for timely intervention. While online counseling services offer convenience, their accessibility is hindered by costs and a lack of free pre-diagnosis options.

With chatbots or digital assistants, limited online services can be provided more easily to provide access to everyone, anywhere and anytime. Fulmer et al. [7] used the Tesschatbot for interventions on mental health issues with a randomized controlled trial method using scoring: GAD-7 score, Positive and Negative Affect Scales (PANAS), PHQ-9, and post-survey. In its development, Tewari et al. [8] created a chatbot application based on the Rasa framework (MentalEase) aimed at sentiment analysis using the Naive Bayes algorithm. This research still does not provide clear results. Meanwhile, similar research focused on adolescents was carried out by Stapleton et al. [9] using a randomized controlled trial and the Acceptance Commitment Therapy (ACT) method. The chatbot developed using Natural Language Processing (NLP) and integrated with Facebook Messenger is an ACT-based chatbot pilot project whose long-term use requires further study. Kaywan et al. [10] implemented Google DialogFLow (NLU), NodeJs integrated with Facebook Messenger and AWS hosting to create a chatbot that can detect the level of user depression with the results of an average level of satisfaction of 79% with an average no depression rate of 15.5%.

In Brazil, a study on the use of chatbots was carried out by evaluating three predetermined programs, namely less anxiety, less stress, and better mood. Daley et al. [11] implemented this chatbot in Ruby and JavaScript, whose conversations were designed by professional clinical psychologists and health experts. Research using the DASS, GAD-7, and PHQ-9 methods and also utilizing Cognitive Behaviour Therapy (CBT) obtained results that were sufficient to reduce symptoms of anxiety, stress, and depression, as seen from the decrease in Cohen's value. These results show the potential of chatbots to help reduce symptoms of mental health issues in general for Brazilian people.

CBT is a method commonly used in the development of chatbots for mental health issues; therefore, there are several similar studies, including He et al. [12], calculating the effectiveness of CBT clinically and non-clinically. Research based on randomized controlled trials ([7][13][14]) was conducted using the Rasa framework integrated with WeChat. While in Indonesia, chatbot-based psychologist consultation services have not been developed with complete functionality, like how to measure anxiety levels. Therefore, this paper develops an anxiety consultation service chatbot system using the Rasa framework to help everyone access an anxiety level assessment system and provide recommendations to users for the need to visit a psychologist using the Semantic Textual Similarity method.

2. METHODS

Previous research on chatbots for pre-diagnosis and recommendation of anxiety disorders has been limited to therapy aids that function to reduce anxiety, where respondents cannot undergo early screening to determine their anxiety level and receive recommendations on whether they need to consult a psychologist or not. This research assumes that: (1) The proposed chatbot built is only for consultations regarding general anxiety, post-traumatic disorder, and social anxiety disorder. (2) The recommendations from the chatbot are only twofold, whether there is a need for consultation with a psychologist or not.

2.1 Chatbot Architecture

The architecture of this prototype chatbot is shown in Figure 1, where the chatbot is integrated with Telegram BOT so that users can easily access the system. The use of NGROK for the chatbot was chosen because it provides convenience as an endpoint for the Rasa framework. Additionally, the system also utilizes MySQL as a data storage for calculations, validation, and session storage. Meanwhile, RabbitMQ is used for conversation tracking, acting as a conversation extractor to be processed and stored.

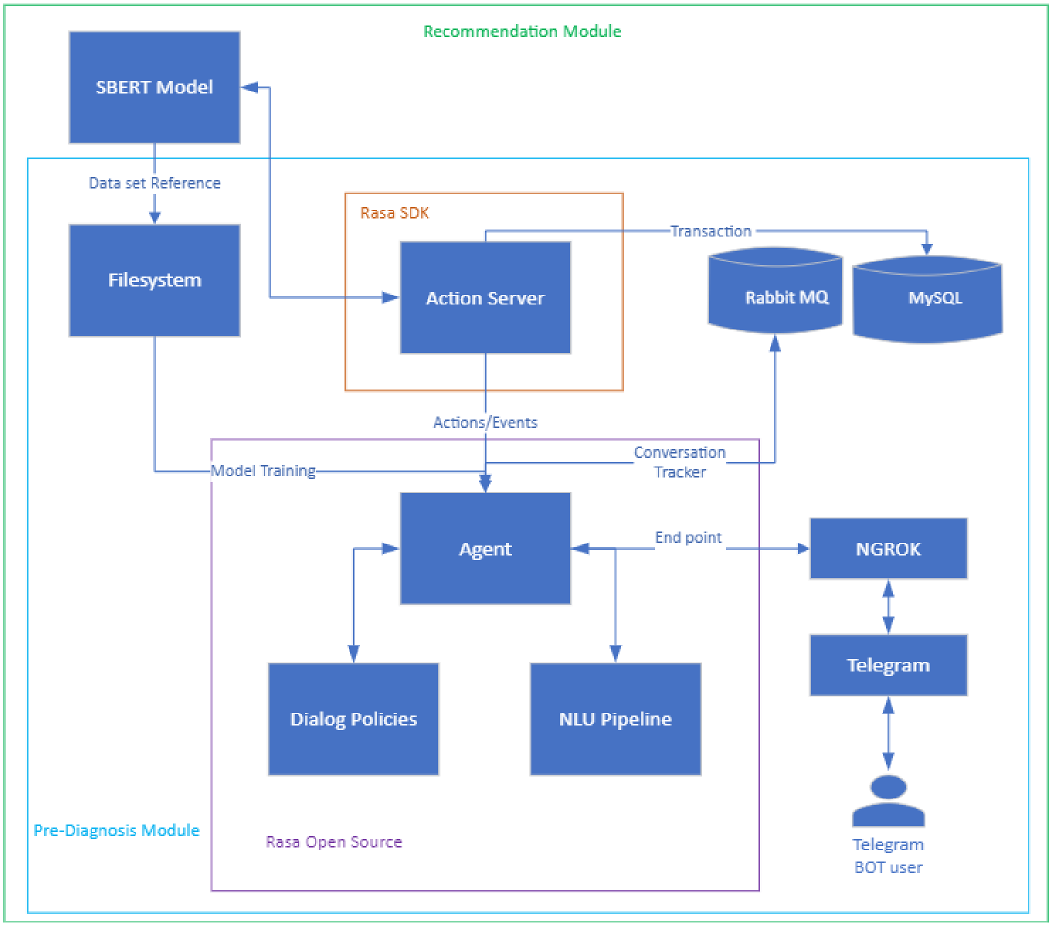

Figure 1 The architecture of the chatbot

This system consists of 2 main modules: pre-diagnosis & recommendation modules. Each modul involves direct interaction from the Telegram BOT user with the Rasa framework. Rasa Open Source functions as the NLU process, handling dialog interactions performed by the Agent, which acts as a message receiver or sender with input or output directly connected to Telegram through NGROK. Additionally, the Agent accesses NLU training models, extracts conversations for processing using RabbitMQ, and contacts Rasa SDK for custom-built actions tailored to each flow's needs. Only the recommendation flow uses the SBERT model, where the reference data is stored in the file system of the Rasa framework.

The program code is custom-written to suit the existing design needs. Specifically, in this research, all program code is written from scratch in Python, involving several configurations for intent classification training, domain setting, stories setting, action development, and event broker setting.

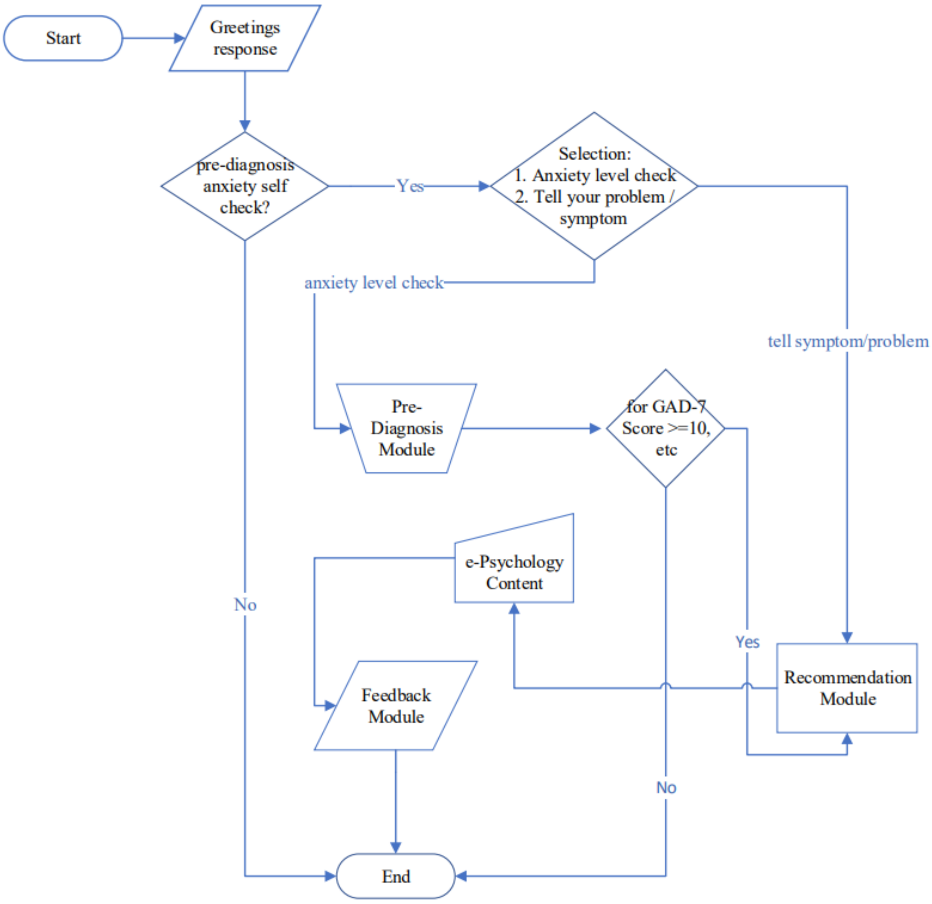

Figure 2 Overall flowchart diagram

As shown in Figure 2, the overall flowchart diagram of the chatbot system is depicted. There are several modules built into this system, namely: Pre-Diagnosis module, and Recommendation module. If the user chooses the Pre-diagnosis self-check, it means the respondent enters the main flow where they are given 2 response options/selections, either entering the Pre-Diagnosis module or the second option to describe the perceived problems/symptoms, which directly enters as input into the Recommendation module. The output of this Pre-Diagnosis module is the measurement of anxiety level performed through intent classification using the NLU classifier model. Conversely, if the user does not choose the Pre-diagnosis self-check, their session will end with a thank you response, and the user will not enter the Pre-Diagnosis flow or Recommendation flow.

Users access @Koncoku and start a session with intents such as: "/start", "hi", "hello" which are received by Rasa via HTTP request, thus predicted by the NLU model and processed as the greet intent activating the welcome path stories. The Rasa server will respond with 2 options, "Yes" to proceed to the selection option (anxiety level check and tell your problem/symptom) and "No" to stop, here the Rasa server will provide the appropriate response for each option. Broadly speaking, the chatbot system will use intents, stories, and responses (from the user) for the overall flow of the chatbot.

The configuration of these selection options is located in the story: check topic health with intent: health and action: action_health_greet. Inside this action class, there are 2 options, "Check Anxiety Level" and "Tell Your Problem". Each of these options has values configured in the payload in the form of buttons to be sent to the chatbot system from Telegram. Both values are trained in their NLU model with consecutive intents being anxietylevel and vent. Next, when the Recommendation flow is successfully completed, it will enter the Feedback flow.

2. 2 Pre-diagnosis module

The Pre-Diagnosis module is a functional block designed to calculate or measure the anxiety level experienced by the user. The measurement methods employed include GAD-7, DASS, and STAIT/STAIS-5, which are commonly used in previous research studies. This module also involves data training for the intent classification of two predefined classifier models, namely the DIETClassifier and LogisticRegressionClassifier, to extract values for the measurement process.

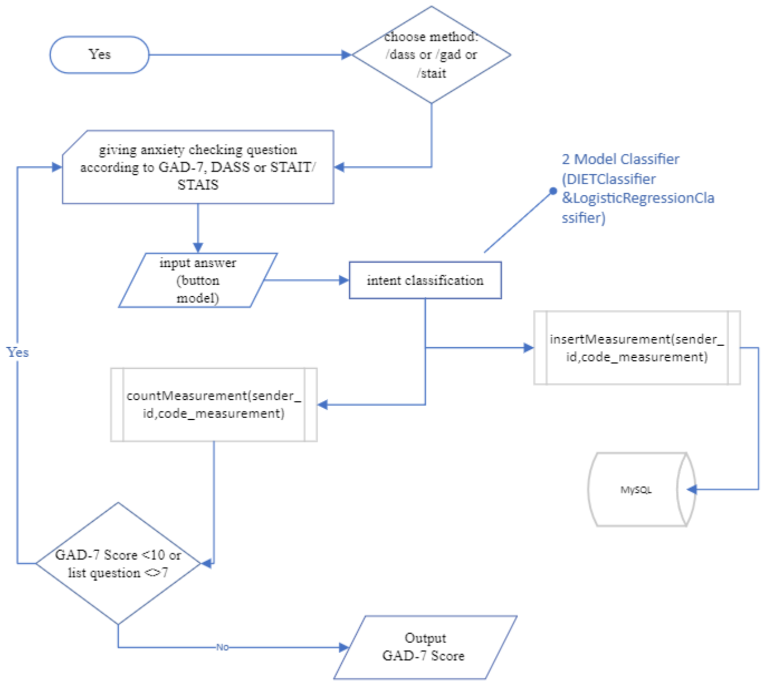

Figure 3 Pre-Diagnosis Flow

Meanwhile, Figure 3 explains how the calculation of anxiety level is done in this chatbot system. Each answer chosen by the correspondent will be processed by the NLU model, thereby being recognized for each intent representing questions from the anxiety disorder assessment using GAD-7, DASS, and STAIT/STAIS-5.

These answers will be converted into a code_measurement value, which will be stored in the MySQL database. The function InsertMeasurement(sender_id, code_measurement), where sender_id represents the Telegram session_id, will be called in each action. Subsequently, the function countMeasurement(sender_id, code_measurement) is used to calculate the total score, which will be used to determine whether to proceed to the next question or to the final total score.

2. 3 Recommendation Module

This module serves as a recommendation model used to complement the functionality of this chatbot system. Input from symptoms will be processed through the prediction of 2 intents, namely mood_great and mood_unhappy. Figure 4 explains the difference in the SBERT models used, specifically the SBERT model for "happy" used for the intent mood_great and the SBERT model for "anxiety" for the intent mood_unhappy. The SBERT model utilizes training data, commonly referred to as reference data. The final result obtained from this flow is that the user receives a recommendation on whether they should meet with a psychologist or not, and simultaneously provides e-Psychology content information from the action_intent_sad. Meanwhile, for the intent mood_happy, the user will be offered the pre-diagnosis flow as a repetition to ensure that the correspondent remains consistent in their answers.

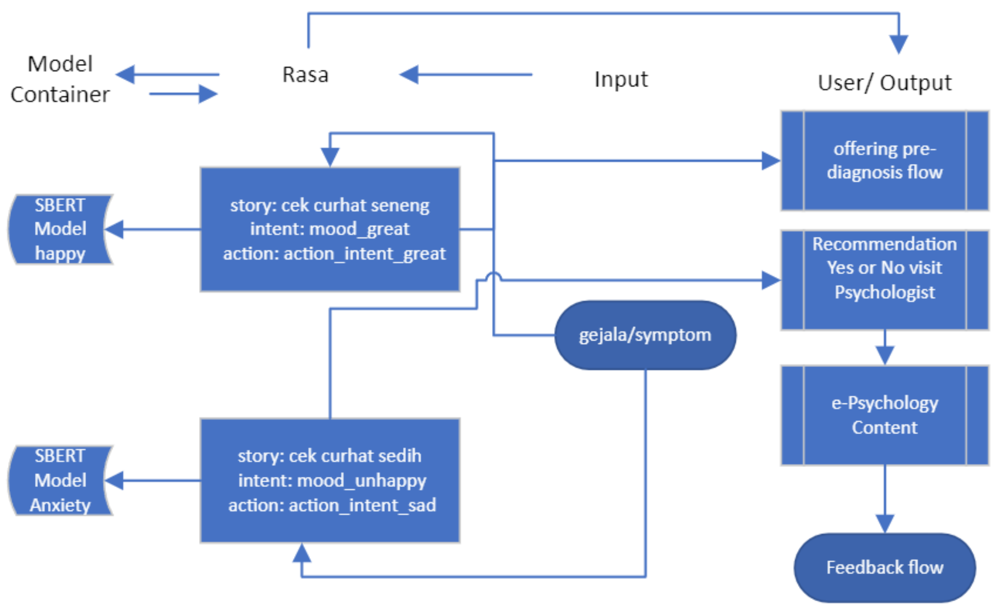

Figure 4 Recommendation Flow

The recommendation model that will be used is the library available on Hugging Face - SentenceTransformer which is also called Sentence-BERT (SBERT) i.e. the indobert base model is a method that is trained using the built from the pre-trained Bert model [15] and the paraphrase-multilingual-MiniLM-L12-v2 model that is already available in the library.

2. 4 Evaluation Model

This testing step is carried out through several stages, namely for the NLU model, recommendation model and feedback/survey from the chatbot which is carried out with a sample of 36 correspondents who randomly try the chatbot system openly.

1. NLU Model

DIETClassifier and LogisticRegressionClassifier test results for classifier and dialog models are in the form of precision, recall, accuracy, and f1score values and confusion matrix tables that can be obtained by running the command: rasa test. Testing is also done by including screenshots of the conversation results of each flow in accordance with the stories, intent and actions.

2. Recommendation Model

This model is tested manually using direct test data from correspondents stored in the database in the form of symptom input. The test is carried out by validating the results of the recommendation model, namely the Y or N labeling, whether or not they match the groundtruth. The evaluation results will be in the form of precision, recall, accuracy and Mean Average Precision (MAP).

3. Feedback/Survey

This test uses a questionnaire with the level of satisfaction using the Chatbot service. The calculation only uses the average value of each question from the questionnaire.

3. RESULTS AND DISCUSSION

The trial was conducted by 35 correspondents who contacted the @Rencangku_Bot account or KoncoKu through the Telegram channel. Using NGROK, the chatbot was connected from Telegram to the Rasa framework server deployed on a local machine. Connected sessions in the chatbot are recorded in the Session table. The stored information includes session_id, name, and the time connected with action_hello_world, which is the main entry point for the chatbot flow.

The chatbot response when the "greet" intent is detected is an introduction explaining what Koncoku is, its purpose, and the limitations of its usage. In terms of the UI, during the mobile version testing, it appears as shown in Figure 5.

Figure 5 Welcome greeting

This flow offers a commonly used questionnaire method for assessing anxiety symptoms. Tobin & House [16] state that anxiety is something experienced by anyone, including children and adults, where symptoms can be observed in their responses to the surrounding environment, such as in family, social, educational, work, and personal development settings. These symptoms can manifest biologically, involving neurological and neurohormonal aspects, leading to feelings of tension, inability to relax, worry, panic symptoms, and fear. Rapee et al. [17] state that individuals with Generalized Anxiety Disorder (GAD) typically have concerns about abstract and uncertain matters. The recently experienced symptoms can provide a general assessment, which is used as a calculation factor in GAD-7, DASS, and STAIT/STAIS-5.

Out of the 35 correspondents who participated in this trial, 30 correspondents took part in the questionnaire. The obtained scores range from minimum, moderate, to high levels. These results will then be validated according to the specified rules. Each measurement has its minimum value: for GAD-7, it is 10; for DASS, it is 4; and for STAIT/STAIS-5, it is 24. When any calculation reaches its minimum value, the intent will switch to the next flow, which is the recommendation flow.

The GAD-7 measurement consists of 7 questions based on symptom assessment over the past 2 weeks. The questions have weights ranging from 1 to 4 when the results of the GAD-7 questionnaire trial happened, where the questioning stops, and the recommendation action continues as it has reached its minimum value. The use of buttons is a system limitation to focus on training data for intents and flows that need to be maintained because these intents are used to obtain values for each question asked. Figure 6.6 is the chatbot's response inviting the correspondent to express the problems they are currently facing when the questionnaire result reaches its minimum value.

The predicted intent is "mood-unhappy," where there are 51 training data accommodating possible responses from correspondents expressing symptoms or current issues. Various responses have been observed during the trial process. By examining the SBERT model's results, it can be calculated how accurate the recommendations are using a minimum similarity_score of 0.8.

Table 1 Comparison of intent classification model

Model NLU	Accuracy	Precision	Recall	F1-Score
DIETClassifier	0.95	0.91	0.92	0.91
LogisticRegressionClassifier	0.99	0.99	0.99	0.99
SklearnIntentClassifier	0.26	0.23	0.10	0.12

According to Bunk et al. [18], DIETClassifier is considered the most effective model for intent modeling, achieving approximately 90% accuracy compared to other models like BERT and Glove. However, in this study, an accuracy of 95% was obtained for intent training without the use of hyperparameters. Meanwhile, the LogisticRegressionClassifier model yielded an accuracy of 99% without fine-tuning. Both the first and second model configurations demonstrated satisfactory intent prediction results, as shown in Table 1.

As for the negative configuration selected, it only achieved an accuracy value of 26%. The use of this type of model provides a clearer comparison, emphasizing that the model selection in the Rasa framework is crucial for chatbot performance. In this model, no intent is comprehensible other than the "mood_happy" intent. Table 2 shows the confidence rate results for intents in the first configuration model, where 14 intents received training results that did not align with the expected intents. For example, when the input "dass90" from the chatbot payload should lead to the DASS questionnaire flow, the result is the "mood_great" intent with a confidence rate of only 12%.

Table 2 *Confidence Rate Intent* Model 1

Training Data	*Expected Intent*	*Intent Prediction*	*Confidence*
baik	greet	mood_great	0.78
apik	greet	mood_great	0.82
great	greet	mood_great	0.71
mulai dass	mulaidass	masukan	0.20
dass90	dass9	mood_great	0.12
dass91	dass9	mood_great	0.25
dass92	dass9	mood_great	0.11
dass93	dass9	stait2	0.15
stait	stait	startstait	0.29
stata91	stait9	deny	0.15
stata92	stait9	dass12	0.14
stata93	stait9	deny	0.13
stata94	stait9	deny	0.18
ya survey	yaSurvey	dass13	0.23
tidakSurvey	tidakSurvey	dass12	0.15
tidak survey	tidakSurvey	stait10	0.13
ok	thankyou	information	0.05

The results of the recommendation model evaluation are taken from a sample of n = 29 from the results of the previously conducted trials. Evaluation is carried out for the SBERT anxiety model by comparing the results of model 1 and model 2 shown in Table 3.

Table 3 Comparison of Anxiety Recommendation Results

Model Rekomendasi	Accuracy	Precision	Recall	F1-score	MAP
indobert base	0.72	0.83	0.41	0.54	0.30
paraphrase-multilingual-MiniLM-L12-v2	0.62	0.57	0.33	0.41	0.26

The greatest accuracy value is owned by model 1 at 72% compared to model 2 which is at 62%. While the largest precision value is owned by model 1 which is 83% followed by 57% by model 2. Following the recall value and f1-score for model 1 is greater at 41% and 54% compared to model 2 which is 33% and 41%. Meanwhile, the MAP values of the two models have almost the same magnitude, respectively, model 1 is 30% and model 2 is 26%.

So, the MAP value is 0.30 or 30%. This result indicates that the SBERT model used is only able to retrieve relevant information by 30%. The use of training data influences the results because out of the 700 data used, contextual cleansing or clinical data validation by the relevant team, namely psychologists or health consultants, has not been conducted. The use of a threshold for the similarity score of 0.8 limits the model from obtaining more classification of anxiety symptoms because, from the sample data, only 4 inputs have a value >= 0.8.

The chatbot feedback flow is the end session of the chatbot. The intent used is 'yaSurvey' after the recommendation flow is skipped. During the trial, respondents who follow the entire flow will be able to enter this feedback flow. There are 21 out of 35 respondents who provided feedback. Questionnaire 1 contains the question 'Can this chatbot help you to perform early screening for symptoms of anxiety?' which obtained an average score of 3.5 out of 5. Questionnaire 2 contains the question 'Is this chatbot easy to use?' and received an average score of 3.7 out of 5. Questionnaire 3 contains the question 'Would you recommend this chatbot to your friends?' and obtained an average score of 4 out of 5. Figure 6 displays the feedback flow ending with a 'Thank you' response

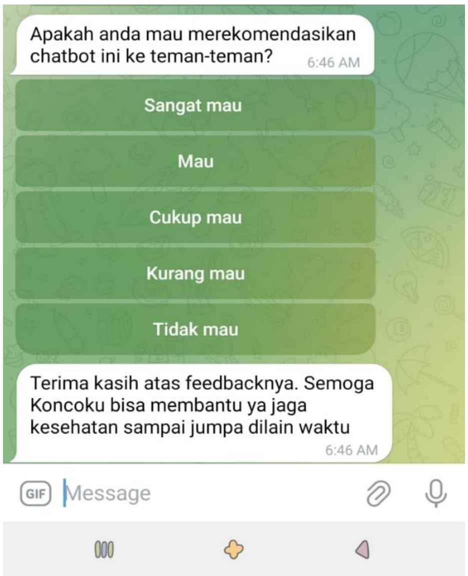

Figure 6 Feedback questionaire

From the results of the trial, it can be seen that respondents are quite enthusiastic about using the chatbot from several aspects such as utility, ease of use, and interest in recommending it to others with scores of 3.5, 3.7, and 4, respectively. These results can be used as a measure of the chatbot's performance. At the same time, it answers that this research can enhance the performance of the chatbot in complementing its function as a tool for early screening of anxiety disorders that can be easily accessed online, free of charge, and without registration, focusing solely on the objectivity values of each respondent.

4. CONCLUSIONS

The results of this study are able to answer where previous research on chatbots for pre-diagnosis and recommendations for anxiety disorder is still limited to therapeutic aids can be overcome with a prototype chatbot model for pre-diagnosis of anxiety disorder by calculating anxiety levels and also a recommendation whether to consult or not. This model is accommodated in the form of the @Koncoku chatbot model on the Telegram platform.

The NLU DIETClassifier model with 95% accuracy and LogisticRegressionClassifier with 99% accuracy are able to provide a good response to 59 existing intentions. Meanwhile, the MAP value in the SBERT model for model 1 indobert base and model 2 paraphrase-multilingual-MiniLM-L12-v2 is only at 30% and 26%, which is not enough to be able to clinically analyze the symptoms of anxiety disorder because it is limited to reference data that is not in accordance with the previous plan, namely data taken from Kaggle only.

The favorable responses from the correspondents for usability, convenience and interest were 3.5, 3.7 and 4 on a scale of 1-5, respectively, indicating that this type of chatbot is capable of adding functionality to the development of chatbots for pre-diagnosis and recommendation services.

This research provides a lot of room for improvisation both from the measurement method, NLU model and primary training data from the Psychology Unit or related parties. It is also expected to explore other intentions that can support consultation services. In particular, the use of the SBERT model must use clinical data so that it can produce recommendation outputs that are in line with expectations.

REFERENCES

[1] M. Clark, V. Gritsenko, J. S. Bonnici, T. Marinova, A. Reznik, and R. Isralowitz, "Psychology Student Attitudes and Beliefs Toward Cannabis for Mental Health Purposes: a Cross National Comparison," International Journal of Mental Health and Addiction , vol. 19, pp. 1866–1874, 2021, doi: https://doi.org/10.1007/s11469-020-00280-6.
[2] J. Pretsch, B. Flunger, and M. Schmitt, "Resilience predicts well-being in teachers, but not in non-teaching employeess," Social Psychology of Education, vol. 15, pp. 321–336, 2012, doi: https://doi.org/10.1007/s11218-012-9180-8.
[3] A. Asmika, H. Harijanto, and N. Handayani, "Prevalensi depresi dan gambaran stressor psikolososial pada remaja sekolah menengah umum di wilayah kotamadya Malang," Jurnal Kedokteran Brawijaya, vol. 24, no. 1, pp. 15–21, 2008, doi: https://doi.org/10.21776/ub.jkb.2008.024.01.2.
[4] M. Coles, " Adolescent Mental Health Literacy: Young People's Knowledge of Depression and Social Anxiety Disorder," Journal of Adolescent Health, vol. 58, no. 1, pp. 57–62, 2016, doi: https://doi.org/10.1016/j.jadohealth.2015.09.017.
[5] I. A. Wani and B. Singh, " Effect of Islamic psycho-spiritual therapy in managing craving, withdrawal symptoms, and mental health problems among cannabis users," Ment Health Relig Cult, vol. 22, no. 7, 2019, doi: https://doi.org/10.1080/13674676.2019.
[6] S. , Onie and A. V. , et al. Daswin, "Suicide in Indonesia in 2022: Underreporting, Provincial Rates, and Means (or a shorter version if needed)," 2022.
[7] R. Fulmer, A. Joerin, B. Gentile, L. Lakerink, M. Rauws, and others, "Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial," JMIR Ment Health, vol. 5, no. 4, p. e9782, 2018.
[8] A. Tewari, A. Chhabria, A. S. Khalsa, S. Chaudhary, and H. Kanal, "A survey of mental health chatbots using NLP," in Proceedings of the International Conference on Innovative Computing & Communication (ICICC), 2021.
[9] A. Stapleton, J. Lavelle, and L. McHugh, "Chatbot-delivered acceptance and commitment therapy with adolescents: A pilot randomized controlled trial," 2022.
[10] P. Kaywan, K. Ahmed, A. Ibaida, Y. Miao, and B. Gu, "Early detection of depression using a conversational AI bot: A non-clinical trial," PLoS One, vol. 18, no. 2, p. e0279743, 2023.
[11] K. Daley, I. Hungerbuehler, K. Cavanagh, H. G. Claro, P. A. Swinton, and M. Kapps, "Preliminary evaluation of the engagement and effectiveness of a mental health chatbot," Front Digit Health, vol. 2, p. 576361, 2020.
[12] Y. He, L. Yang, B. W. BE, S. Z. BE, C. Qian, and T. T. BE, "Mental health chatbot for young adults with depressive symptoms: a single-blind, three-arm, randomized controlled trial.," J Med Internet Res, 2022.
[13] M. C. Klos, M. Escoredo, A. Joerin, V. N. Lemos, M. Rauws, and E. L. Bunge, "Artificial intelligence–based chatbot for anxiety and depression in university students: pilot randomized controlled trial," JMIR Form Res, vol. 5, no. 8, p. e20678, 2021.
[14] S. Greer et al., "Use of the chatbot 'vivibot' to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial," JMIR Mhealth Uhealth, vol. 7, no. 10, p. e15018, 2019.
[15] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Apr. 2019. [Online]. Available: https://arxiv.org/abs/1908.10084
[16] R. M. Tobin and A. E. House, DSM-5® Diagnosis in the Schools. Guilford Publications, 2015.

[17] R. M. Rapee, H. J. Lyneham, E. S. Sburlati, and C. A. Schniering, Evidence-Based CBT for Anxiety and Depression in Children and Adolescents A Competencies Based Approach. Wiley, 2014.
[18] T. Bunk, D. Varshneya, V. Vlasov, and A. Nichol, "Diet: Lightweight language understanding for dialogue systems," *arXiv preprint arXiv:2004.09936*, 2020.