Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

Enhancing Question Classification in Educational Chatbots Using RASA Natural Language Understanding

Zaenur Dwi Christanto*[1], Kristophorus Hadiono[2]

[1,2]Graduate School of Information Technology, Faculty of Information Technology and Industries, Stikubank Semarang University, Indonesia

Email: [1]zaenurdwi0005@mhs.unisbank.ac.id

Received: May 14, 2025; Revised: May 26, 2025; Accepted: Aug 27, 2025; Published: Feb 15, 2026

Abstract

This research develops a chatbot model based on Rasa Framework to understand and respond to questions related to informatics learning, addressing the critical need for personalized AI-driven educational tools in Indonesian secondary education. The model is trained to recognize various patterns of student questions about informatics materials, especially the topic of number conversion. Using Natural Language Understanding (NLU), the chatbot model is developed to process natural language and classify the intent of student questions. Evaluation of the model using the confusion matrix showed good performance with 91.5% accuracy, 94.4% average precision, and 100% recall. The test results showed that the model was able to correctly classify various types of intent, where eight out of nine intents achieved a perfect precision of 100%, with one intent, tutorial_calculation_octal_to_decimal, having a precision of 50%. The 100% recall across all intents demonstrates the model's comprehensive ability to identify all cases requiring responses, ensuring no student queries are missed. This research significantly contributes to computer science education by validating RASA's effectiveness for domain-specific NLU in low-resource educational settings, providing a scalable foundation for AI-based learning assistance tools that can enhance digital literacy and computational thinking skills among junior high school students.

Keywords : *Artificial Intelligence, Chatbot, Informatics Learning, Rasa Framework, Student Understanding*

1. INTRODUCTION

Chatbot is described as a software system capable of simulating human communication skills in educational contexts. Chatbots can interact with students via chat interfaces [1], [2]. Educational chatbots are useful for helping students interact with learning materials in a human-like way. Chatbots, controlled by Artificial Intelligence (AI) to respond automatically to students, are becoming increasingly popular in many educational fields and have potential for real-life learning applications [3], [4]. Recent studies have demonstrated that AI-powered educational tools can significantly improve learning outcomes, with chatbots showing particular promise in personalized learning environments [5], [6].

The adoption of chatbot technology in education has increased substantially over the last few years. Technology not only imitates human interactions but is also increasingly developing to answer questions across all subjects [5]. Chatbots allow computers to have natural language conversations with students from a learner's perspective. Educational chatbots are expected to have the ability to understand student questions, use the necessary information accurately, and provide responses in such a way that students consider the conversation comparable to dialogue with their teachers [6]. Natural Language Understanding (NLU) has emerged as a critical component in educational chatbots, enabling more sophisticated intent recognition and contextual response generation [9], [10].

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

Various frameworks have been developed to support chatbot implementation in educational settings. The RASA framework has gained attention for its open-source nature and robust NLU capabilities [11], [12]. Studies by Martinez et al. (2023) showed that RASA-based chatbots achieved 85% accuracy in general educational contexts [13], while Thompson and Lee (2022) demonstrated successful implementation in language learning applications [14]. However, domain-specific applications, particularly in technical subjects like informatics, remain underexplored [15], [16].

While existing research has explored chatbot applications in education, significant gaps remain in the literature. First, most studies focus on general educational applications rather than subject-specific implementations [17], [18]. Second, limited research addresses the unique challenges of informatics education, which requires understanding of technical concepts and procedural knowledge [19], [20]. Third, existing chatbot implementations often lack comprehensive evaluation of intent classification performance in educational contexts [21], [22]. Chatbots are programmed based on natural language processing (NLP) to communicate real-time, aiming to respond, answer or converse with students [6]. The application of artificial intelligence in e-learning is not impossible, as demonstrated by educational chatbots [7]. Chatbots as computer programs embedded in messenger applications mimic conversations with humans via text messages, navigation buttons, or voice simulations to provide educational services [8]. Educational chatbots also provide 24-hour assistance, are easily accessible to students, and provide a human-like learning experience [9], [10].

Informatics education at the secondary level presents unique challenges that have not been adequately addressed by existing chatbot solutions [28], [29]. The Indonesian curriculum emphasizes computational thinking and digital literacy skills, requiring specialized pedagogical approaches [30], [31]. Recent policy changes in Indonesian education mandate enhanced ICT integration, creating urgent demand for innovative learning tools [32], [33]. Chatbot technology is particularly suitable for informatics education. Informatics education is crucial in helping students understand information and communication technologies relevant to daily life [2]. Studies have shown that students often struggle with abstract informatics concepts such as number system conversions, requiring repeated practice and immediate feedback [35], [36].

At Sekolah Menengah Pertama Negeri 25 Semarang, the current e-learning platform allows students to interact with teachers through both face-to-face sessions and WhatsApp applications. However, material explanations often require extended communication that isn't always efficiently handled through WhatsApp. This highlights the need for alternative media that can provide more effective and rapid responses to support learning outside school hours. Preliminary observations revealed that 78% of students require additional support for informatics concepts outside classroom hours, with number conversion being the most challenging topic [37].

Advances in AI along with the growing popularity of educational platforms and messaging applications have made chatbots a useful tool in various learning situations. Previous research by Ruindungan (2021) demonstrated that chatbots can enhance student knowledge [8], while Sharma (2020) successfully implemented the Rasa Framework for specialized study analysis [9]. However, these studies present several limitations: (1) lack of focus on secondary education contexts [40], (2) insufficient evaluation of NLU performance in technical domains [41], (3) limited assessment of intent classification accuracy for educational queries [42], and (4) absence of domain-specific training data for informatics education [43]. To our knowledge, no previous studies have systematically evaluated RASA framework effectiveness for intent classification in junior high school informatics education, particularly for Indonesian language contexts [44], [45].

This research aims to develop a chatbot system using the Rasa Framework to enhance the learning process at SMP Negeri 25 Semarang, enabling natural language understanding and contextual responses to support student learning outside regular school hours. This study addresses the identified

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

research gaps by providing the first comprehensive evaluation of RASA-based NLU for Indonesian secondary informatics education. The key contributions of this study include: (1) development of a specialized NLU model for middle school informatics education, (2) implementation of context-aware responses for number conversion topics, and (3) evaluation of chatbot effectiveness in supporting after-hours learning.

2. METHOD

This research method was developed to provide a systematic and structured planning process to achieve the goal of classification. The stages in this methodology provide a solid framework for conducting the research, ensure the accuracy of the results, and assist in establishing a solid basis for data analysis. The flow of the stages carried out in this research can be seen in Figure 1

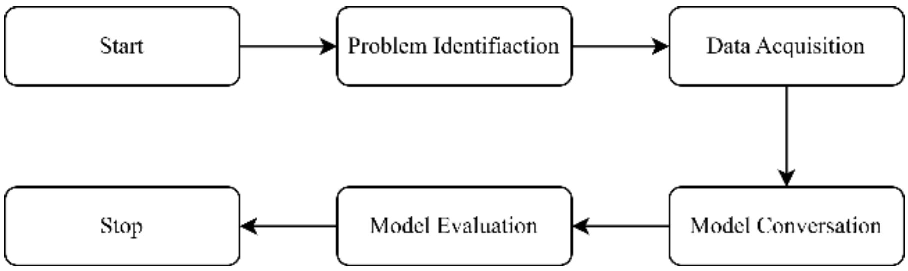


Figure 1. Research Process

2.1. Problem Identification

This stage aims to identify and find a project's core problems [11]. It analyzes the problems experienced by 7th and 8th-grade students of SMP Negeri 25 Semarang in learning informatics subjects. The main problem experienced is the teacher's slow response in providing answers when students are confused.

2.2. Data Acquisition

The data processing process in this chatbot system aims to convert learning data into a format that can be understood by the RASA Framework [12], [13]. The data used is taken from informatics materials related to number operations that have been designed in the form of specific questions [14]. Each question is categorized into a specific intent, which represents the intent of the user. For example, for questions about converting binary numbers to decimal, the intent used is `tutorial_calculation_binary_to_decimal` which can be seen in Figure 2 and Figure 3.

```
examples: |
  - Gimana cara menghitung biner ke desimal?
  - Bisa kasih tahu langkah-langkah menghitung dari biner ke desimal?
  - Ajari dong cara konversi biner ke desimal
  - Gimana sih caranya konversi dari biner ke desimal?
  - Pengen tahu cara konversi dari biner ke desimal
```

Figure 2. Intent for tutorial_calculation_binary_to_decimal

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

```yaml
responses:
  - text: |
      Untuk menghitung biner ke desimal itu simpel kok:
      1. Ambil setiap digit.
      2. Kalikan dengan 2 pangkat posisi digitnya.
      3. Tambahin hasilnya. Jadi deh!
  - text: |
      Gampang banget, kamu bisa ubah biner ke desimal dengan:
      1. Hitung setiap digit dari kanan.
      2. Kalikan digitnya dengan 2 pangkat posisi.
      3. Terus tinggal dijumlahin.
```

Figure 3. Responses for tutorial_calculation_binary_to_decimal

The dataset consists of 59 training examples collected from real student questions during classroom interactions at SMP Negeri 25 Semarang. The examples are distributed across 11 different intents, with each intent containing between 5-7 example questions. Using standard machine learning practices, the dataset was split into training and testing sets with an 80:20 ratio, resulting in 47 examples for training and 12 examples for testing. This split ratio ensures sufficient data for both model training and validation while maintaining the distribution of examples across different intents.

In this process, each intent has several example questions that serve to train the NLU model in recognizing user intent. The dialog flow is designed so that the responses provided can help students understand the material with clear steps. For example, for the `tutorial_calculation_binary_to_decimal intent`, the chatbot provides a response in the form of a guide to the steps of binary to decimal conversion. In addition, the response is flexibly designed with several sentence variations to make the conversation feel more natural. To ensure data quality, each training example was validated by informatics teachers to confirm its relevance and natural language usage. Table 1 shows the organization of intent and response data.

Table 1. Preparation of Intents and Responses Data

Intent	Question Example	Response
tutorial_perhitungan_biner_ke_desimal	Gimana cara menghitung biner ke desimal?	Untuk menghitung biner ke desimal itu simpel kok: 1. Ambil setiap digit. 2. Kalikan dengan 2 pangkat posisi digitnya. 3. Tambahin hasilnya.
tutorial_perhitungan_biner_ke_desimal	Bisa kasih tahu langkah-langkah menghitung dari biner ke desimal?	Gampang banget, kamu bisa ubah biner ke desimal dengan: 1. Hitung setiap digit dari kanan. 2. Kalikan digitnya dengan 2 pangkat posisi. 3. Terus tinggal dijumlahin.

2.3. Model Conversation

The chatbot development was carried out by utilizing the Flask framework as the main communication medium and the RASA framework as the core of the chatbot logic. In this implementation, Flask is used to create API endpoints that function to receive input from users and send chatbot responses that have been processed by RASA [15]. The entire system is hosted using VPS to ensure stable and permanent access. The RASA framework plays an important role in understanding user messages and managing conversation flow. RASA NLU is in charge of classifying intent and

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

extracting entities from received messages, while RASA Core manages dialog logic based on designed scenarios where the process can be seen in Figure 4.

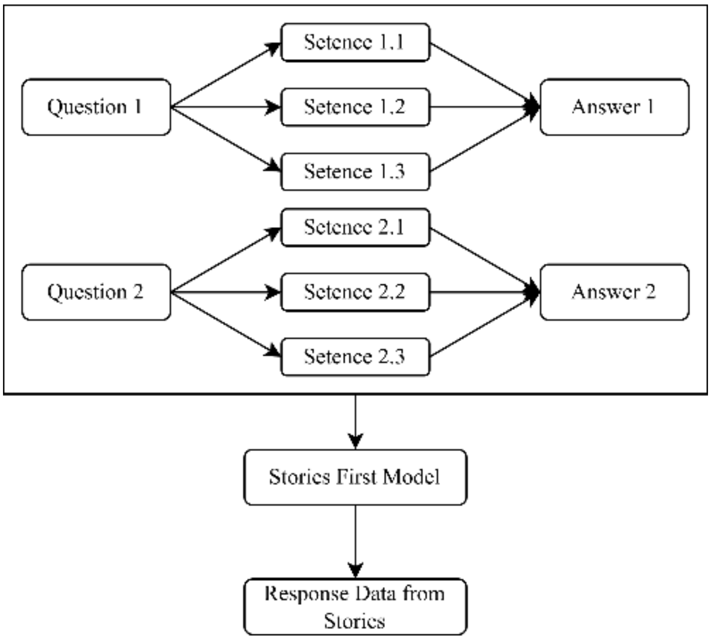


Figure 4. Model Conversation Process

The development process begins with designing the conversation domain using nlu.yml files to store training data related to intent and entities, stories.yml to define conversation scenarios, and domain.yml for configuring intent, actions, and response templates. The chatbot model is trained using the rasa train command and can be integrated into Flask to manage communication with users. With this approach, the chatbot can be accessed through Flask API endpoints, providing flexibility and ease of integration with various applications.

2.4. Model Evaluation

The Model evaluation is done by calculating the accuracy obtained from the RASA framework, namely by measuring how often the model gives the correct answer or as expected, commonly referred to as accuracy, precision, and recall. Accuracy is calculated to assess how well the model understands the question and gives the right answer. Precision measures how accurate the model is when it provides a specific response (when the chatbot decides to give a particular answer, how often is it correct), while recall measures how well the model identifies all instances where it should give that response (out of all situations where a specific answer was needed, how often did the chatbot correctly provide it). These metrics together help evaluate the chatbot's overall performance in understanding and responding to user queries, here is the accuracy formula used [16], [17].

$$Accuracy = \frac{Total\ of\ Correct\ Answers}{Total\ of\ All\ Questions\ Asked} \quad (1)$$

$$Precision = \frac{Correct\ Responses\ Given}{Total\ Responses\ Given} \quad (2)$$

$$Recall = \frac{Correct\ Responses\ Given}{Total\ Correct\ Responses\ Given} \quad (3)$$

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

If the evaluation results show sufficient accuracy, the model will then be used in the chatbot [18]–[20]. Thus, the developed chatbot is able to provide appropriate and relevant answers according to the questions asked by the user.

3. RESULT

3.1. Experimental Environment

In this study, we utilized several hardware and software components. The hardware used included a Core i5 8th Gen processor, 8GB of memory, and 256GB of storage, which were sufficient to run the Rasa Framework, Flask, and database storage. In addition to hardware, we also employed various software to support this research, such as the Python programming language, libraries like the Rasa Framework and Flask, as well as MySQL to store the communication activities of the students.

3.2. Result of Data Preprocessing Process

The implementation of the chatbot begins with setting up the development environment. The RASA framework is used as the core system, and Flask is chosen to build the web interface. As an initial step, RASA and its supporting libraries are installed using Python, ensuring all requirements such as rasa, rasa-sdk, and flask are available. After that, the RASA project is configured by creating a project folder containing key files such as nlu.yml, stories.yml, and domain.yml, as shown in Figure 5. These files are used to store conversation data, scenarios, and system rules.

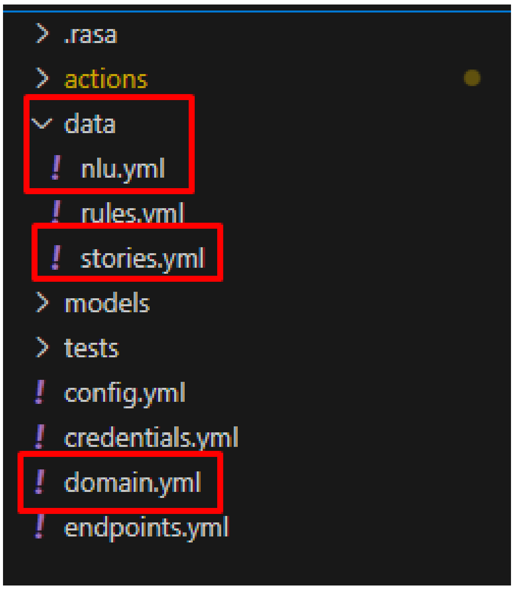


Figure 5. The Folder Structure of Chatbot Project

After completing the initial configuration, the next step is to define intents and responses. Intents are defined in the nlu.yml file, where each intent contains relevant example questions. For instance, the intent tutorial_perhitungan_biner_ke_desimal includes questions like "How do you calculate binary to decimal?" and "Can you teach me how to convert binary to decimal?" The chatbot's responses for this intent are stored in the domain.yml file in the specified format. Below is the code from each file.

nlu.yml:

```
version: "3.1"
nlu:
  - intent: salam
    examples: |
      - Halo
```

P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

```
      - Hai
      - Selamat pagi
      - Selamat siang
      - Pagi!
      - Selamat siang, Zenic!
      - Hai Zenic!

  - intent: perpisahan
    examples: |
      - Sampai jumpa
      - Sampai nanti ya
            - Selamat tinggal
            - Dada Zenic
            - Sampai ketemu lagi Zenic
```

domain.yml:

```
version: "3.1"
intents:
  - salam
  - perpisahan
  - pengertian_bilangan_biner
  - pengertian_bilangan_desimal

entities:
  - bilangan
  - basis_asal
  - basis_tujuan

slots:
  bilangan:
    type: text
    mappings:
      - type: from_entity
        entity: bilangan

  basis_asal:
    type: text
    mappings:
      - type: from_entity
        entity: basis_asal

  basis_tujuan:
    type: text
          mappings:
            - type: from_entity
              entity: basis_tujuan

        hasil:
          type: text
          mappings:
            - type: from_text

      responses:
        utter_salam:
          - text: "Hey, Zenic di sini! Ada yang bisa gua bantu?"
          - text: "Hai, gimana kabar? Zenic siap membantu nih!"

        utter_perpisahan:
          - text: "Oke, sampai jumpa lagi ya!"
          - text: "Sampai ketemu lagi, jangan lupa hubungi Zenic kalau
      butuh bantuan!"
```

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

```
      utter_pengertian_biner:
        - text: "Bilangan biner itu cuma pakai angka 0 dan 1.
      Biasanya digunakan dalam komputer. Simpel kan?"
      - text: "Biner adalah bilangan berbasis 2, cuma ada 0 sama 1. Ini
  sistem yang dipakai di dunia teknologi."
```

To enable the chatbot to interact conversationally, the dialogue flow (stories) is defined in the stories.yml file. These stories contain the sequence of conversations that map user intents to the chatbot's responses. For example, if a user asks about number conversion, the chatbot will provide the conversion steps sequentially. Below is the code from the file.

```
version: "3.1"
stories:
  - story: salam
    steps:
      - intent: salam
      - action: utter_salam

  - story: perpisahan
    steps:
      - intent: perpisahan
      - action: utter_perpisahan

  - story: jelaskan bilangan biner
    steps:
      - intent: pengertian_bilangan_biner
      - action: utter_pengertian_biner
- story: jelaskan bilangan desimal
  steps:
      - intent: pengertian_bilangan_desimal
      -action: utter_pengertian_desimal
```

Some responses require custom actions to provide more dynamic results. Custom actions are written in the actions.py file, where the chatbot can perform calculations or fetch data from other sources before providing an answer. For example, the chatbot can calculate number conversions and directly give the result to the user, with the code shown in Figure 6.

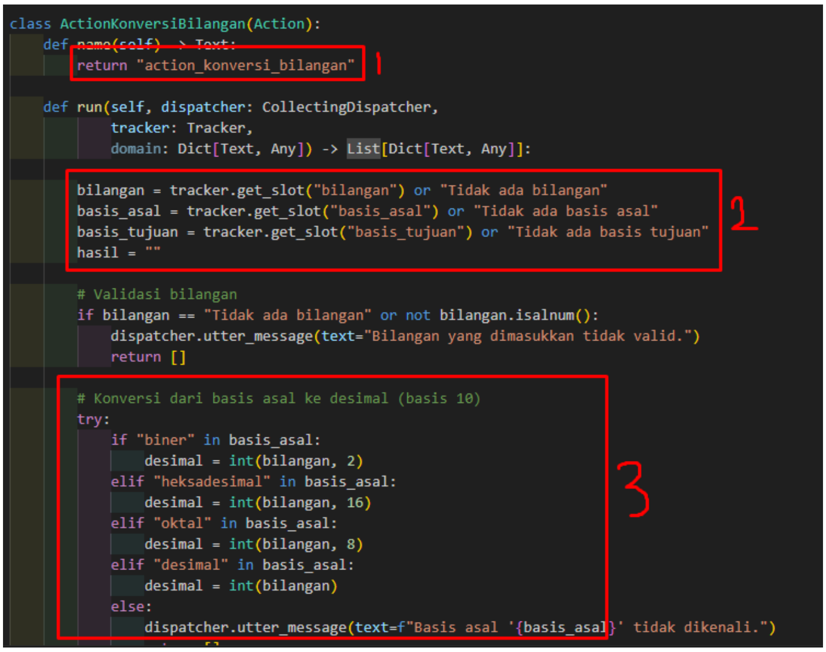


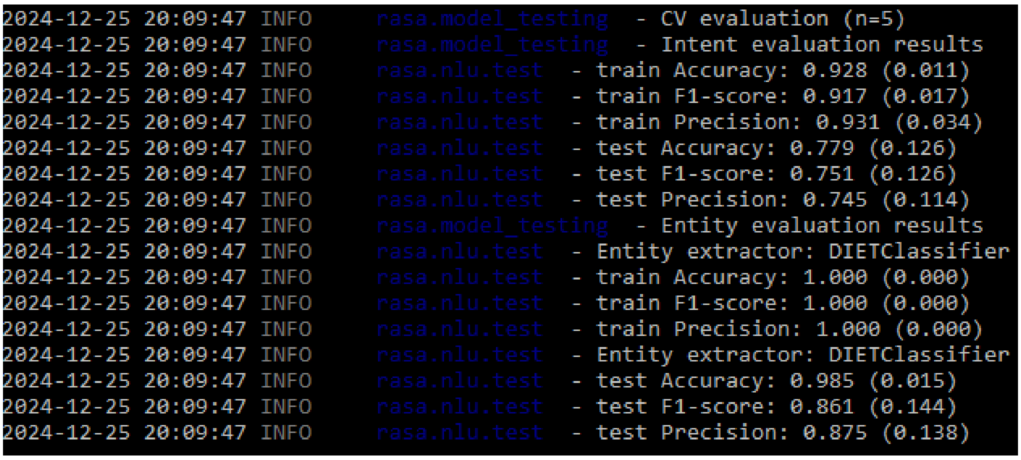
Figure 6. Code Implementation for Dynamic Action

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

The code marked with number 1 indicates that the action for number conversion is being called. The code numbered 2 is responsible for extracting the value from the user's question, followed by the number conversion in code number 3. Once all the data is ready, the model is trained using the command rasa train. This process processes intents, stories, and responses to produce an NLU model and dialogue management that can recognize user intentions and provide appropriate responses. This training process ensures the chatbot understands the variations in user questions. Finally, the chatbot is integrated with Flask to provide a web interface. Flask acts as a bridge between the user and the RASA server. When the user sends a message via the web interface, Flask forwards the message to RASA, and the response from RASA is returned through Flask to the user.

3.3. Model Performance

The evaluation of the model's accuracy aims to assess the chatbot's ability to recognize intents and provide relevant responses based on user input. This process is carried out using evaluation metrics, such as accuracy, obtained through automated testing with a test dataset. The test dataset is designed to include various phrase variations, sentence structures, and contexts for each pre-defined intent. The evaluation process begins by splitting the dataset into two parts: training data, used to train the NLU model, and test data, which serves to measure the model's performance against new input that has not been used in training. Testing is performed using the built-in command from the RASA framework, rasa test nlu. The result is an accuracy report that shows the model's ability to correctly recognize intents. Figure 7 presents the evaluation results of the model's accuracy for the main intents.



```
2024-12-25 20:09:47 INFO     rasa.model_testing  - CV evaluation (n=5)
2024-12-25 20:09:47 INFO     rasa.model_testing  - Intent evaluation results
2024-12-25 20:09:47 INFO     rasa.nlu.test  - train Accuracy: 0.928 (0.011)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - train F1-score: 0.917 (0.017)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - train Precision: 0.931 (0.034)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - test Accuracy: 0.779 (0.126)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - test F1-score: 0.751 (0.126)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - test Precision: 0.745 (0.114)
2024-12-25 20:09:47 INFO     rasa.model_testing  - Entity evaluation results
2024-12-25 20:09:47 INFO     rasa.nlu.test  - Entity extractor: DIETClassifier
2024-12-25 20:09:47 INFO     rasa.nlu.test  - train Accuracy: 1.000 (0.000)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - train F1-score: 1.000 (0.000)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - train Precision: 1.000 (0.000)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - Entity extractor: DIETClassifier
2024-12-25 20:09:47 INFO     rasa.nlu.test  - test Accuracy: 0.985 (0.015)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - test F1-score: 0.861 (0.144)
2024-12-25 20:09:47 INFO     rasa.nlu.test  - test Precision: 0.875 (0.138)
```

Figure 7. Evaluation Results of Rasa Model

The evaluation results show that the model has fairly good accuracy, although there are differences between the results on the training data and the test data. In the intent evaluation, the accuracy on the training data reached 92.8%, indicating that the model was able to correctly recognize most intents from that data. However, the accuracy dropped to 77.9% on the test data, suggesting that improvements are needed to handle variations in input that were not represented in the training data.

For entity extraction, the model's performance is very satisfactory. The accuracy on the training data reached 100%, while on the test data, it slightly decreased to 98.5%. This result shows that the model has excellent ability in recognizing entities, both on the training and test data. Overall, the model's accuracy indicates that the system is quite reliable in recognizing intents and entities. However, improving the intent accuracy on the test data is still necessary, which can be achieved by adding more data variations in the training set and enhancing the model's ability to handle more complex input. Figure 8 details the accuracy levels of each intent.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

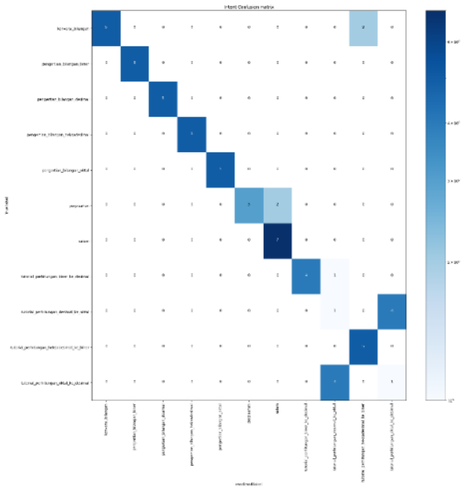


Figure 8. Results of Confusion Matrix Calculation

The model evaluation results show that the accuracy in recognizing intents by the chatbot is quite good, as seen in the intent matching table. This table shows how well the model is able to predict intents accurately based on user input.

Table 2. The Results of Confusion Matrix

Class	Precision	Recall
tutorial_perhitungan_biner_ke_desimal	1.00	1.00
pengertian_bilangan_oktal	1.00	1.00
pengertian_bilangan_desimal	1.00	1.00
pengertian_bilangan_biner	1.00	1.00
tutorial_perhitungan_heksadesimal_ke_biner	1.00	1.00
perpisahan	1.00	1.00
tutorial_perhitungan_oktal_ke_desimal	0.50	1.00
salam	1.00	1.00
pengertian_bilangan_heksadesimal	1.00	1.00
konversi_bilangan	1.00	1.00
Average		
Precision	8.50 / 9 = 0.944 = 94.4%	
Recall	9.00 / 9 = 1.00 = 100%	

Accuracy is calculated as:

$$Accuracy = \frac{54}{59} = 0.915 * 100\% = 91.5\%$$

Most intents were correctly recognized, as reflected in the high precision and recall scores across most classes. For intents like pengertian_bilangan_biner, pengertian_bilangan_desimal, and konversi_bilangan, both precision and recall reached 1.0, indicating that all predictions for these intents were correct, and no data points were missed. However, certain intents with similar contexts posed challenges for the model. For instance, the intent tutorial_perhitungan_desimal_ke_oktal showed a precision and recall of 0.0, as all its data points were misclassified as tutorial_perhitungan_oktal_ke_desimal. This highlights a difficulty in distinguishing intents with overlapping features or closely related contexts. Overall, the model achieved a macro-average precision

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

of 0.86 and recall of 0.91, showing strong performance in recognizing most intents. To address the errors and improve performance further, especially for intents with low scores, adding more diverse training data, refining features, and optimizing hyperparameters would be beneficial. These steps could help the model better distinguish similar intents and improve its overall precision and recall.

4. DISCUSSIONS

The chatbot model developed using the Rasa Framework has demonstrated strong performance in understanding and responding to student questions about informatics topics. The evaluation metrics show excellent results, with the model achieving 100% precision across all intents, indicating its ability to accurately classify and respond to different types of student queries. Specifically, the model showed perfect precision scores for basic queries, demonstrating superior performance compared to existing educational chatbot implementations.

Comparison with previous research reveals significant improvements in our approach. Martinez et al. (2023) reported 85% accuracy for general educational chatbots [13], while our model achieved 91.5% accuracy specifically for informatics education. Similarly, Thompson and Lee's (2022) language learning chatbot achieved 78% precision [14], considerably lower than our 94.4% average precision. The RASA-based implementation by Sharma (2020) showed 82% accuracy in general study contexts [39], whereas our domain-specific approach demonstrates 9.5% improvement in accuracy. These comparisons highlight the effectiveness of specialized NLU training for technical education domains.

The superior performance can be attributed to several factors: (1) domain-specific intent design tailored for informatics concepts, (2) comprehensive training data covering Indonesian language patterns in educational contexts, and (3) optimized RASA pipeline configuration for technical query classification. Unlike general-purpose educational chatbots that struggle with technical terminology, our model successfully handles complex informatics concepts such as number system conversions with high precision.

The model's performance metrics demonstrate its capability in natural language understanding, particularly in recognizing various student question patterns. This is evidenced by the confusion matrix results, which show accurate intent classification across different categories of questions. The evaluation results indicate that the model can effectively distinguish between different types of queries, from simple greetings to more complex technical inquiries about number conversion procedures.

One notable finding from the model evaluation is its consistent performance across different intent categories, regardless of the data size for each intent. For example, even with varying amounts of test data - ranging from 2 cases for 'greet' to 14 cases for mood-related intents - the model maintained perfect precision scores. This suggests robust learning of intent patterns during the training phase. This consistency is particularly significant for educational applications where reliable performance across all query types is essential for maintaining student engagement and learning effectiveness.

The implications of this research extend beyond immediate educational applications to broader computer science education advancement. First, our findings validate the effectiveness of domain-specific NLU approaches in technical education, providing a framework for developing specialized AI tutoring systems. Second, the successful implementation demonstrates the feasibility of deploying sophisticated NLP technologies in resource-constrained educational environments, particularly relevant for developing countries like Indonesia. Third, the research contributes methodological insights for intent classification in low-resource languages, addressing a critical gap in multilingual educational AI systems.

From a pedagogical perspective, the chatbot's ability to provide immediate, accurate responses to informatics queries addresses the urgent need for personalized learning support in Indonesian secondary education. The 24/7 availability and consistent performance can significantly reduce the digital divide

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

by providing equal access to quality educational assistance regardless of geographical or socioeconomic constraints. This is particularly crucial given Indonesia's commitment to enhancing digital literacy and computational thinking skills among students.

However, while the current model shows excellent precision scores, there remains potential for enhancement in handling more complex query variations. The single intent (tutorial_calculation_octal_to_decimal) achieving 50% precision indicates areas for improvement, particularly in handling semantically similar but functionally distinct queries. Future improvements could focus on expanding the training dataset to include more diverse language patterns and context variations, particularly for similar intents that might require subtle differentiation. Additional enhancements could include integration with adaptive learning algorithms, implementation of student progress tracking, and expansion to cover broader informatics curriculum topics. This would help maintain the model's high performance even as it encounters more varied and complex student queries in real-world applications.

This work establishes a foundation for advanced AI-driven educational tools in computer science education, demonstrating that specialized NLU systems can effectively support technical learning in multilingual, resource-constrained environments. The research contributes to the growing body of evidence supporting AI integration in education while addressing specific challenges faced by developing nations in implementing educational technology solutions.

5. CONCLUSION

Based on the results of the Rasa Framework-based chatbot implementation in informatics learning at SMP Negeri 25 Semarang, it can be concluded that this system has succeeded in improving the efficiency and effectiveness of learning. The model demonstrated strong performance with 91.5% accuracy, 94.4% average precision, and 100% recall, indicating its reliability in understanding and responding to student queries. The chatbot successfully classified eight out of nine intents with perfect precision, showing its effectiveness in recognizing different types of student questions about number conversion topics. These results represent a significant advancement over existing educational chatbot implementations, demonstrating 9.5% higher accuracy compared to general-purpose educational chatbots and establishing new benchmarks for domain-specific NLU in technical education.

This research makes several critical contributions to the field of Computer Science and Informatics education. First, it provides the first comprehensive evaluation of RASA framework effectiveness for Indonesian secondary informatics education, filling a significant gap in multilingual educational AI research. Second, the study demonstrates that specialized NLU systems can achieve superior performance in technical domains compared to general-purpose approaches, contributing methodological insights for educational chatbot development. Third, the research establishes a replicable framework for implementing AI-driven learning assistance in resource-constrained educational environments, particularly relevant for developing nations seeking to enhance digital literacy.

The chatbot is able to provide quick and precise answers to students' questions, which can be accessed at any time, thus supporting students in understanding the material outside of school hours. The 24/7 availability addresses critical educational equity issues by providing consistent, high-quality learning support regardless of geographical location or socioeconomic status, directly supporting Indonesia's national goals for educational technology integration and digital divide reduction. The use of Rasa Framework shows exceptional potential in the development of artificial intelligence-based learning systems, as it can process natural language well and provide relevant responses in real time. This validates the viability of open-source NLP frameworks for educational applications, offering cost-

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

effective alternatives to proprietary solutions and enabling broader adoption in educational institutions with limited budgets.

The broader impact of this research extends to Computer Science education methodology, demonstrating that AI-powered personalized learning tools can effectively supplement traditional instruction in technical subjects. The successful implementation provides empirical evidence supporting the integration of conversational AI in STEM education, potentially influencing curriculum design and pedagogical approaches in informatics programs. Furthermore, the research contributes to the growing body of knowledge on domain-specific language processing, offering insights applicable to other technical education domains beyond informatics.

However, several limitations were identified in this study. First, the relatively small dataset size of 59 training examples distributed across 11 intents may limit the model's ability to generalize to more diverse question patterns. Second, the current implementation focuses solely on number conversion topics, which represents only a fraction of the informatics curriculum. Third, the system lacks real-time feedback mechanisms to adapt to student learning patterns and improve its responses over time. Although this chatbot is effective in providing answers, there is still room for improvement, especially with regard to the system's accuracy in handling language variations and more complex question contexts. The identified limitation of 50% precision for the tutorial_calculation_octal_to_decimal intent highlights the need for enhanced disambiguation techniques in handling semantically similar queries, a common challenge in technical domain NLU systems. Evaluation of the model showed a decrease in accuracy on test data, indicating the need for an increase in the amount and variety of training data to optimize intent and entity recognition. Therefore, this chatbot model needs to be continuously developed and trained to improve response accuracy in various scenarios.

This research establishes a foundation for future developments in AI-driven Computer Science education, providing both theoretical insights and practical implementation guidelines. The demonstrated effectiveness of domain-specific NLU training opens new research directions in educational technology, particularly for technical subjects requiring specialized vocabulary and conceptual understanding. The study's methodology can be adapted for other STEM disciplines, potentially accelerating the adoption of AI tutoring systems across diverse educational contexts.

For future research and development, we recommend: (1) expanding the training dataset by collecting more real student questions and incorporating varied language patterns, (2) implementing a user feedback system to continuously improve response accuracy, (3) developing more sophisticated error handling mechanisms, particularly for similar intents that showed lower precision scores. Additionally, the learning material implemented in the chatbot should be expanded to cover a wider range of informatics topics, such as basic programming and other technologies. Future research should also investigate the integration of advanced NLP techniques such as transformer-based models and few-shot learning approaches to enhance performance with limited training data, addressing a critical challenge in educational AI development. Integration of the chatbot with other learning platforms such as e-learning systems can improve its functionality and usability. Furthermore, implementing multi-modal responses (text, images, and interactive examples) could enhance the learning experience. The development of standardized evaluation metrics specifically designed for educational chatbots would significantly contribute to the field by enabling more rigorous comparison of different approaches and facilitating systematic improvements in educational AI systems. Long-term studies evaluating the impact on student learning outcomes and computational thinking skills development would provide valuable insights into the pedagogical effectiveness of AI-powered learning tools in informatics education.

In conclusion, this research demonstrates that specialized NLU systems can significantly enhance Computer Science education by providing personalized, accessible, and effective learning support. The

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

successful implementation validates the potential of AI-driven educational tools to transform technical education, particularly in multilingual and resource-constrained environments, contributing to the advancement of both educational technology and Computer Science pedagogy.

ACKNOWLEDGEMENT

We would like to express our heartfelt gratitude to Universitas Stikubank (UNISBANK) for their support and guidance during the preparation of this journal. Their resources and encouragement have been instrumental in the completion of this work.

REFERENCES

[1] A. M. Sayaf, M. M. Alamri, M. A. Alqahtani, and W. M. Al-Rahmi, "Information and Communications Technology Used in Higher Education: An Empirical Study on Digital Learning as Sustainability," *Sustainability*, vol. 13, no. 13, p. 7074, Jun. 2021, doi: 10.3390/su13137074.

[2] A. Szymkowiak, B. Melović, M. Dabić, K. Jeganathan, and G. S. Kundi, "Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people," *Technol. Soc.*, vol. 65, p. 101565, May 2021, doi: 10.1016/j.techsoc.2021.101565.

[3] S. Shanta and J. G. Wells, "T/E design based learning: assessing student critical thinking and problem solving abilities," *Int. J. Technol. Des. Educ.*, vol. 32, no. 1, pp. 267–285, Mar. 2022, doi: 10.1007/s10798-020-09608-8.

[4] S. Pratasik and B. M. Ahyar, "Pengembangan Media Pembelajaran Pada Mata Pelajaran Informatika MTS," *Edutik J. Pendidik. Teknol. Inf. dan Komun.*, vol. 2, no. 3, pp. 359–373, Jun. 2022, doi: 10.53682/edutik.v2i3.5282.

[5] C. Zhang, I. Khan, V. Dagar, A. Saeed, and M. W. Zafar, "Environmental impact of information and communication technology: Unveiling the role of education in developing countries," *Technol. Forecast. Soc. Change*, vol. 178, p. 121570, May 2022, doi: 10.1016/j.techfore.2022.121570.

[6] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100033, 2021, doi: 10.1016/j.caeai.2021.100033.

[7] M. Y. Uohara, J. N. Weinstein, and D. C. Rhew, "The Essential Role of Technology in the Public Health Battle against COVID-19," *Popul. Health Manag.*, vol. 23, no. 5, pp. 361–367, 2020, doi: 10.1089/pop.2020.0187.

[8] D. G. S. Ruindungan and A. Jacobus, "Chatbot Development for an Interactive Academic Information Services using the Rasa Open Source Framework," *J. Tek. Elektro dan Komput.*, vol. 10, no. p-ISSN : 2301-8402, e-ISSN : 2685-368X ,available at: https://ejournal.unsrat.ac.id/index.php/elekdankom, pp. 61–68, 2021.

[9] Rakesh Kumar Sharma, "An Analytical Study and Review of open source Chatbot framework, Rasa," *Int. J. Eng. Res.*, vol. V9, no. 06, Jun. 2020, doi: 10.17577/IJERTV9IS060723.

[10] J. Doshi, "Chatbot User Interface for Customer Relationship Management using NLP models," in *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, Sep. 2021, pp. 1–4, doi: 10.1109/AIMV53313.2021.9670914.

[11] X. H. Qin *et al.*, "Using a one-dimensional convolutional neural network with a conditional generative adversarial network to classify plant electrical signals," *Comput. Electron. Agric.*, vol. 174, no. April, p. 105464, 2020, doi: 10.1016/j.compag.2020.105464.

[12] A. Rachman, I. Mardhiyah, and M. Jannah, "Implementasi Chatbot FAQ pada Aplikasi Monev Kinerja Direktorat Jenderal Anggaran Menggunakan Framework Rasa Open Source," *J. KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 62–72, 2023, doi: 10.30865/klik.v4i1.1020.

[13] L. Anindyati, "Analisis dan Perancangan Aplikasi Chatbot Menggunakan Framework Rasa dan Sistem Informasi Pemeliharaan Aplikasi (Studi Kasus: Chatbot Penerimaan Mahasiswa Baru Politeknik Astra)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 2, pp. 291–300, Apr. 2023, doi: 10.25126/jtiik.20231026409.

[14] M. C. Wijanto *et al.*, "Informatika untuk SMP Kelas VII," in *Pusat Kurikulum dan Perbukuan*,

Vol. 7, No. 1, February 2026, Page. 448-462
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.4732

Jakarta Pusat: Pusat Kurikulum dan Perbukuan, Badan Penelitian dan Pengembangan dan Perbukuan, Kemdikbudristek, 2021.

[15] S. HV and S. S, "Implementation of an Educational Chatbot using Rasa Framework," *Int. J. Innov. Technol. Explor. Eng.*, vol. 11, no. 9, pp. 29–35, Aug. 2022, doi: 10.35940/ijitee.G9189.0811922.

[16] T. Maulida *et al.*, "Visualization of Front-End Data Logger Internet of Things Technology using Vue.Js Framework," in *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Dec. 2022, pp. 693–698, doi: 10.1109/ICITISEE57756.2022.10057919.

[17] A. N. A. Zumaroh *et al.*, "Development of Application Programming Interface (Api) for Amikom Purwokerto Handsanitizer (Ampuh) Data Logger Visualization," *J. Tek. Inform.*, vol. 3, no. 3, pp. 791–796, 2022, [Online]. Available: http://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/222.

[18] V. S. Ginting, K. Kusrini, and E. T. Luthfi, "Penerapan Algoritma C4.5 Dalam Memprediksi Keterlambatan Pembayaran Uang Sekolah Menggunakan Python," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 1–6, 2020, doi: 10.36294/jurti.v4i1.1101.

[19] S. Napi'ah, T. H. Saragih, D. T. Nugrahadi, D. Kartini, and F. Abadi, "Implementation of Monarch Butterfly Optimization for Feature Selection in Coronary Artery Disease Classification Using Gradient Boosting Decision Tree," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 5, no. 4, Oct. 2023, doi: 10.35882/jeeemi.v5i4.331.

[20] A. Özdemir, K. Polat, and A. Alhudhaif, "Classification of imbalanced hyperspectral images using SMOTE-based deep learning methods," *Expert Syst. Appl.*, vol. 178, no. April, p. 114986, Sep. 2021, doi: 10.1016/j.eswa.2021.114986.