International Journal of Religion
2024
Volume: 5| Number 11 | pp. 7163 – 7171
ISSN: 2633-352X (Print) | ISSN: 2633-3538 (Online)
ijor.co.uk
DOI: https://doi.org/10.61707/n95fvk06

Building Conversational AI for Assamese Religious Text using Deep Learning

Surajit Sarma[1], Nabankur Pathak[2], Neeva Rani Phukan[3], Risheraj Baruah[4], Sonali Deka[5], Maya Devi Subba[6], Pinkumoni Deori[7] and Diganto Sahu[8]

Abstract

This paper focuses on the use of Deep Learning techniques to build an AI Conversational Agent for Assamese religious texts. Assamese is spoken by the people of Assam which is a state of North East India. It is a language that has unique semantic and syntactic features. In order to handle the religious text in Assamese language specialized Natural Language Processing (NLP) task is required. In this paper we have presented the Deep Learning approach using bi-LSTM model to handle Assamese religious text and the specialized NLP functions that were developed to process the Assamese text. This study also involved the construction of dataset for Assamese language religious text which has been presented in this paper. We have also presented the performance and accuracy testing result of the proposed model and compared it with other Deep Learning model using matrices such as precision, recall, and F1-score. Our proposed model has achieved a significant improvement, reaching an impressive accuracy of 89.99%.

Keywords: *Deep Learning, Conversational Agent, RNN, GRU, BiLSTM*

INTRODUCTION

Technological advancements have increased the demand for conversational agents in regional languages due to which research on this field has increased. Assamese is a language with unique semantic and syntactic features which is a dominant language in North East India. It is a low resources language in digital platform with handle full digital presence of its content and accessibility. Developing an AI conversational agent for Assamese religious text is a unique challenge as compared to other Indo-aryan language such as English, Hindi, Bengali or Marathi. Furthermore, the field of AI Conversational agent for Assamese religious text powered by deep learning remains unexplored, with minimal existing research.

AI technologies have improved the traditional Conversational Agents into an intelligent assistant that has the capability to understands the need and fulfilled them. The benefits of this Conversational Agents such has availability for 24/7, multitasking, serving multiple users and reduce operational costs has make numerous industries to adopted it to automate their customer services. A survey was conducted by Mindbrowser in the year 2020 over surveying 300 individuals

across aviation, e-commerce, education, and retail which reports how diverse industries are adopting and profiting from AI Conversational Agent for providing customer services. (See Figure 1). (Global Chatbot Trends Report, 2020, 2020)

[1] Department of Computer Science, Krishna Kanta Handiqui State open University, Guwahati, Assam, India, E-mail: suravsarma2007@gmail.com

[2] Department of Computer Science, Krishna Kanta Handiqui State open University, Guwahati, Assam, India

[3] Department of Assamese, Krishna Kanta Handiqui State open University, Guwahati, Assam, India

[4] Department of Computer Science, Gauhati University, Guwahati, Assam, India

[5] Department of Assamese, Krishna Kanta Handiqui State open University, Guwahati, Assam, India

[6] Department of Assamese, Krishna Kanta Handiqui State open University, Guwahati, Assam, India

[7] Department of Assamese, Krishna Kanta Handiqui State open University, Guwahati, Assam, India

[8] Faculty of Engineering, Assam down town University, Guwahati, Assam, India

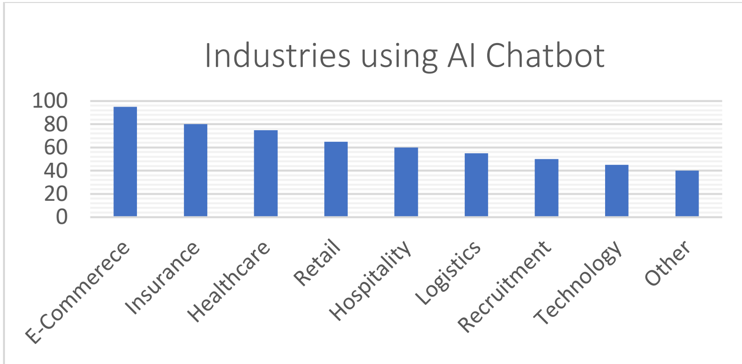


Figure 1: Industries benefiting from AI chatbots

Figure 1 reveals the percentage of e-commerce usages of conversational agents, or chatbots, more to deliver customer service as compared to other industries. Other industries that uses conversational agents includes insurance, healthcare, retail, hospitality, logistics, recruitment, and even the tech sector. The rush in chatbot popularity and demand has fuelled new research efforts focused on improving their design, development, and evaluation.

As per the Investopedia (2022) and IBM (2023), the machine learning and the natural language processing is the main technique used by most of the AI Conversational Agent to perform conversation between humans and machines. Conversational Agents are now days found in almost many portable devices, websites and software such as voice assistants, chatbot apps, phone etc. It is also known as virtual assistant. Researched by Caldarini (2022) defines that the conversational AI are the complicated programs that helps in human conversation to provide automated online assistance and guidance. Costa (2018) added that the chatbots or Conversational Agents are the sophisticated software programs that has the capability to provide assistance and can able to interact with multiple users simultaneously. Further, Madhumitha's (2019) in her researched study highlights the increasing demands of conversational agent or chatbot, which is also known as chatterbots or Human-Computer Interaction (HCI) systems that simulate human conversation.

The primary motivation for developing an AI Conversational Agent in Assamese is to address the digital divide affecting Assamese speakers. By creating an AI Conversational Agent capable of understanding and responding in Assamese, we aim to provide equitable access to information and services, empower users, and contribute to the growth of Assamese language technology. This project is driven by the potential to improve the lives of millions of Assamese speakers while advancing natural language processing research in underrepresented languages.

The main objective of this research are as follows:

- To develop a corpus for Assamese Religious text.
- To design and implement a robust natural language processing (NLP) and deep learning model capable of understanding and generating appropriate responses in Assamese.
- To evaluate the performance of the developed model for the given dataset.

REVIEW OF LITERATURE

Deep learning technologies are mostly applied to create AI Conversational Agent or chatbot in Indo-Aryan languages. One of the good example is the Bengali chatbot that was created by Masum(2021).This chatbot provide replies to the general knowledge question using a transformer model. Masum had developed a big dataset of Bengali question and answers in order to train the model. This Conversational Agent or chatbot has achieved 85.0 BLEU on the QA data and has shown a good performance in tackling sequence related problems. M. Kowsher (2019) has developed another Bengali Conversational Agent using machine learning and BNLP for the educational domain. Abhay Chopde(2022) has came up with an new innovative way of developing

Conversational Agent that can learn from different kind of information and even can speak. In the medical field, Qiming Bao (2020) recommended a hybrid model incorporating a knowledge graph and a text similarity technique to understand and answer difficult health questions.

Most AI Conversational Agent or chatbot that are available in market were developed for English language and few of them extend to Indo-Aryan language such as Bengali, Hindi, Urdu or others. Assamese is functionally and structure different form English and other Indo-Aryan languages, so the models or technique used by this language don't work for Assamese. So, creating an AI Conversational Agent for Assamese religious text was a challenging task because there wasn't much information or helpful tools available in the digital world.

People have been trying developed AI Conversational Agent that can understand and find information in Assamese religious text. It is very hard because there aren't many tools available in digital platform to help with this. One of the study proposed an IR system that utlize the "Vector Space" model and Assamese wordnet to enhance search through query expansion with synonyms (Barman, 2013). Other researchers have also made contributions to Assamese language. Rita Chakraborty(2015) developed a system that could answer questions asked in Assamese by understanding the language's structure. M. P. Bhuyan (2019) created a IR system that used a list of Assamese words and Wikipedia to find information. Although this tool worked well for some searches, it struggled with complex questions. To improve it, the researcher suggested creating a bigger database, looking for information from more websites, and using a better method to rank the results. These studies show that there's potential to develop better search tools for Assamese and that this would be helpful for both people who use the language and those who study it.

This research aims to build an AI chatbot that can understand and respond in Assamese. Since there wasn't enough data available for this, we created our own dataset. We used a combination of Bi-LSTM and attention to build our main model and trained it with our new dataset. To see how well our proposed model worked, we compared it to other popular AI models such as FFDN, RNN, LSTM, GRU, Bi-GRU, Bi-LSTM and measured their performance using specific criteria such as recall, precision, and F1 score. In conclusion, we scrutinized and compared the outcomes of these models to identify the most effective one for our specific context.

METHODOLOGY

In this section we presented the design methodology of our proposed system. The designed methodology is divided into six part. (See Figure 2)

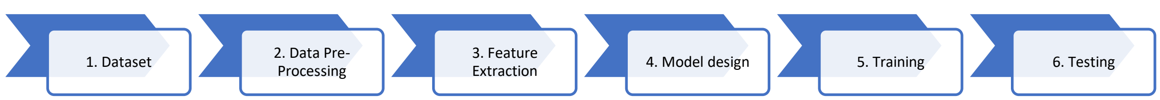


Figure 2: Proposed System Design Methodology

Dataset

For this research study we have developed our own dataset for Assamese Religious text. We have gathered data form diverse sources. The primary source of the data is taken form questionnaires and interviews. The secondary source of data is websites, social media platforms, books etc. The non-Assamese and English data were translated into Assamese language by a dedicated team of translators. We have collected 2165 question and their respective reply. A snapshot of the dataset is provided in Table 1:

Table 1: Snapshot of Dataset

Tag	Query	Response
ধৰ্ম	অসমৰ প্ৰাথমিক ধৰ্মীয় প্ৰথা কি?	অসমৰ প্ৰধান ধৰ্মীয় প্ৰথা বৈষ্ণৱ ধৰ্ম, বিশেষকৈ শ্ৰীমন্ত শংকৰদেৱে প্ৰচাৰ কৰা ভক্তিৰ ৰূপ।
বিহু	অসমীয়া সংস্কৃতিত বিহুৰ তাৎপৰ্য কি?	অসমীয়া নৱবৰ্ষ আৰু কৃষি চক্ৰ বিশেষকৈ বীজ চপোৱাৰ বতৰ উপলক্ষে পালন কৰা বিহু অসমৰ আটাইতকৈ গুৰুত্বপূৰ্ণ উৎসৱ।
শংকৰদেৱ	শ্ৰীমন্ত শংকৰদেৱ কোন আছিল?	অসমৰ নৱবৈষ্ণৱী আন্দোলনৰ প্ৰতিষ্ঠাপক হিচাপে গণ্য কৰা পঞ্চদশ শতিকাৰ এজন সাধু-পণ্ডিত শ্ৰীমন্ত শংকৰদেৱ।
নামঘৰ	অসমীয়া ধৰ্মত নামঘৰ কি?	এটা নামঘৰ হৈছে অসমত বৈষ্ণৱ ধৰ্মৰ চৰ্চাৰ কেন্দ্ৰীয় প্ৰাৰ্থনা ভৱন আৰু সামূহিক কেন্দ্ৰ।
সত্ৰ	অসমীয়া সমাজত সত্ৰৰ ভূমিকা কি?	সত্ৰ হৈছে অসমৰ ধৰ্মীয়, সাংস্কৃতিক, শৈক্ষিক কাম-কাজৰ কেন্দ্ৰ হিচাপে কাম কৰা মঠৰ অনুষ্ঠান।
ভক্তি	ভক্তি আন্দোলনে অসমীয়া ধৰ্মক কেনেদৰে প্ৰভাৱিত কৰিছিল?	শ্ৰীমন্ত শংকৰদেৱৰ দৰে ব্যক্তিৰ নেতৃত্বত ভক্তি আন্দোলনে একক দেৱতা বিষ্ণুৰ প্ৰতি ভক্তি (ভক্তি)ৰ ওপৰত গুৰুত্ব আৰোপ কৰিছিল আৰু অসমত জাতিগত প্ৰভেদ নাকচ কৰিছিল।

Data Pre-Processing

We created an Assamese Natural Language processing library (ANLP) which encompasses all the required functions for our research study such as word tokenization, Assamese stop word removal, Assamese word stemming etc. With the help of all this function we filter the data to enhance the conversational agent's understanding and responsiveness in Assamese.

Features Extraction of Assamese Religious Text

For the feature's extraction of religious text in Assamese language we have adopted the Bag of Words (BOW) technique. It is most widely used technique and is suitable for this research study.

Model Design

We have selected the bi-LSTM model along with attention mechanism to developed our proposed model. The design of the proposed system is presented in the "Figure 3".

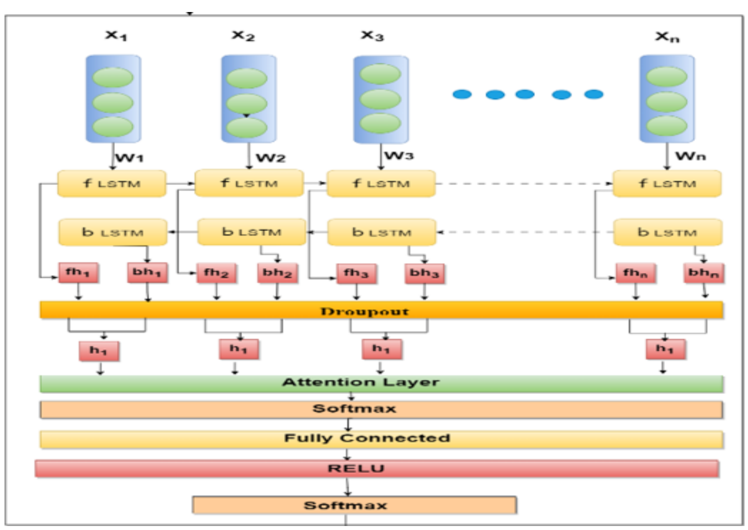


Figure 3: Proposed System Design

Bi-LSTM: Bi-LSTM model is can handle sequence data such as sentence. It can process the data form both forward and backward direction. It is better in information capturing and learning form the processed data.

The main role of dropout layer in the proposed system is to exclude some units in the network to counteract the overfitting. This layer is employed on the output of the bi-directional LSTM layer before it undergoes the attention mechanism. Attention layer focused on the most pertinent segments of the bi-LSTM layer's output. It processed the concatenation of forward and backward hidden state by undergoing a linear transformation by tanh activation function. The role of first softmax function here is to normalize the scalar values of each hidden state by yielding attention weight. The final component of the network is the fully connected layer, responsible for carrying out the classification task. The fully connected layer learns to map from the larger context vector to the smaller class vector. In simpler terms, it serves as the network segment that makes the ultimate decision regarding the classification task. After the fully connected layer, there's a ReLU function. This makes sure that the result is always positive and adds a bit of complexity to the network. The second softmax function is used as the neural network's final activation function in the proposed model. The goal is to confirm that the output result of the network is converted into a probability distribution over the expected result classes. This function makes sure that the total of the output values is 1 and that they fall within the range of 0 to 1.

Training

To enhance the proposed model's training, we employ the Adam optimizer, a sophisticated variant of standard gradient descent. This method accelerates learning by incorporating momentum and adaptively adjusting the learning rate for each parameter. It has been found that using Adam optimizer for our task yields better results than using other optimizers like AdaBoost or RMSProp. We have also made arrangements to avoid overfitting. To mitigate, the same, the optimizer is provided with model-specific information, including learning rate and weight decay.

This experiment utilized an Intel Core i5-10400 CPU operating at 2.90GHz, featuring a 6-core, 12-thread architecture and 8 GB of RAM. Deep learning models were developed using Python 3.8.2 and PyTorch 1.9.0.

Testing

Dataset Splitting: The prepared dataset was partitioned into training and testing subsets using Python's Scikit-learn library. We have opted for the general splitting ratio for our work, starting with a 80:20 split. The 80/20 split was employed, allocating 80% of the data for model training and the remaining 20% for evaluation.

Model performance was assessed through accuracy testing, mean square error testing and a comprehensive evaluation of precision, recall, and F1-score. To establish a comparative benchmark, the proposed model was contrasted against six other deep learning architectures: FFDN, RNN, LSTM, GRU, Bi-LSTM, and Bi-GRU. The testing is discussed in brief in the following sections: -

Accuracy Testing: This testing basically tests how a model can predict the right answer. This is obtained by dividing the total number of correct predictions by total number of predictions. The formula for accuracy testing is as follows:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{1}$$

Mean Square error (MSE): This can be defined as the measure of difference between the expected and predicted value. The mathematical formula for MSE is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

where:

- n is the number of observations
- y_i is the actual value of the variable being predicted for the ith observation
- $\hat{y}_i$is the predicted value of the variable for the i^{th} observation

Precision: This is the measure of ratio of correct positive predictions to total positive predictions. The formula is given below:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{3}$$

Where,

TP=True Prediction

FP= False Prediction

Recall: Recall is defined as the measures of the proportion of actual positive cases correctly identified. The formula for recall is:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{4}$$

F1: F1 score is a balanced measure of a model's accuracy, considering both precision and recall. The formula for F1 score is:

$$\text{F1 Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{5}$$

RESULT

Accuracy Testing

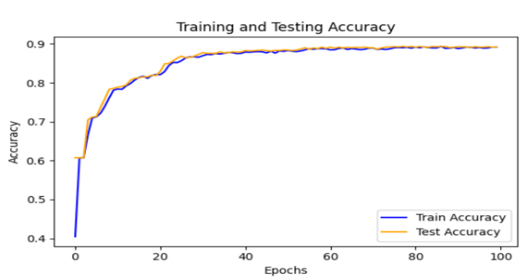
Table 2: Accuracy Table of all the models

Model	Accuracy
FFDN	89.23
RNN	89.74
LSTM	89.17
GRU	89.78
Bi-LSTM	89.69
Bi-GRU	89.84
Proposed Model (Bi-LSTM+Attention)	89.99

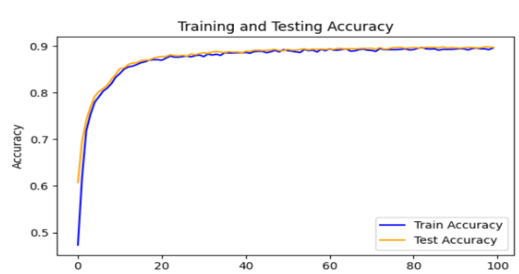
A comparative analysis of different neural network models is provided in Table 2. Using accuracy as the evaluation metric, the proposed model demonstrated superior performance, correctly classifying 89.99% of instances.

Based on the accuracy metric, the proposed model demonstrated superior performance compared to the other models evaluated. The LSTM model, on the other hand, yielded the poorest results.

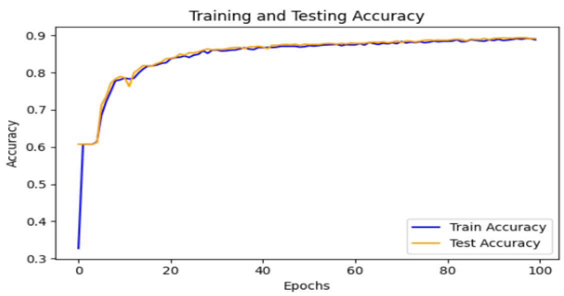
The model's learning process is tracked by the training accuracy graph, which represents its performance on the data used to adjust its parameters. The testing accuracy graph, conversely, shows how well the model performs on new, unseen data. A model with strong generalization capabilities is characterized by high and closely aligned training and testing accuracy curves, as depicted in Figure 4.



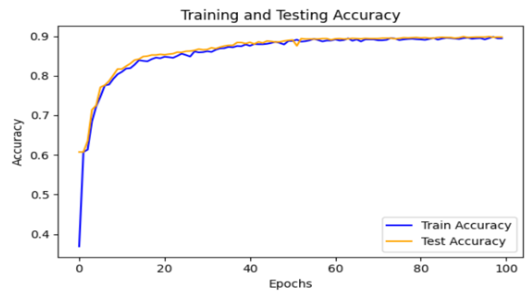
FFDN



RNN



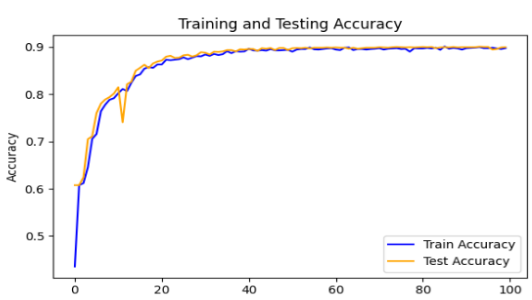
LSTM



GRU

Bi-LSTM

Bi-GRU



Proposed model (Bi-LSTM + Attention)

Figure 4: Training and Testing Graph of all the models

Mean Square Error

A comparative analysis of neural network models based on mean squared error (MSE) is presented in Table 3. The proposed model demonstrated superior performance with the minimum MSE value of 5.67, signifying the most accurate predictions.

Table 3: Mean Square Error Table of all the models

Models	Mean Square Error
FFDN	6.51
RNN	6.15
LSTM	6.92
GRU	6.25
Bi-LSTM	6.21
Bi-GRU	5.85
Proposed Model (Bi-LSTM+Attention)	5.63

Based on the MSE metric, the proposed model demonstrated the best performance, making it the most suitable model for this regression problem.

Performance Testing

Table 4: Precision, Recall, F1 Score Table of all the models

Model	Precision		Recall		F1 Score	
	Macro	Weighted	Macro	Weighted	Macro	Weighted
FFDN	0.78	0.88	0.71	0.88	0.73	0.87
RNN	0.95	0.9	0.79	0.91	0.84	0.88
LSTM	0.78	0.88	0.63	0.88	0.68	0.87
GRU	0.89	0.9	0.81	0.91	0.84	0.88
Bi-LSTM	0.95	0.92	0.79	0.91	0.82	0.88
Bi-GRU	0.95	0.92	0.81	0.91	0.85	0.88
Proposed Model (Bi-LSTM+Attention)	0.95	0.91	0.81	0.91	0.85	0.88

Table 4 presents a performance comparison of various neural network models across precision, recall, and F1-score metrics for a classification task.

The proposed model, along with RNN, Bi-LSTM, and Bi-GRU, achieved the highest macro precision (0.95), indicating strong overall class-wise accuracy. In terms of weighted precision, Bi-LSTM, Bi-GRU, and the proposed model excelled (0.92), effectively considering class imbalance.

For macro recall, GRU, Bi-GRU, and the proposed model demonstrated superior performance (0.81), capturing a significant portion of relevant instances for each class. Weighted recall metrics revealed GRU, Bi-LSTM, Bi-GRU, and the proposed model as top performers (0.91), effectively addressing class imbalance.

The proposed model and Bi-GRU shared the highest macro F1-score (0.85), demonstrating a balanced trade-off between precision and recall across classes. Furthermore, RNN, GRU, Bi-LSTM, Bi-GRU, and the proposed model achieved comparable weighted F1-scores (0.88), indicating good overall performance while considering class distribution.

In conclusion, the proposed model consistently ranks among the top performers across all evaluation metrics, showcasing its effectiveness for this classification task.

DISCUSSION

Selecting the optimal model hinges on the chosen evaluation metric. Different metrics prioritize distinct aspects of performance, often involving trade-offs. Accuracy, while simple, overlooks class imbalance and misclassification costs. MSE excels at regression but is unsuitable for classification. Precision, recall, and F1-score are classification metrics, but their interpretations vary based on class weighting.

Ultimately, the best model aligns with specific problem objectives and conditions. Our proposed Bi-LSTM + Attention model emerges as the superior choice due to its exceptional performance across multiple metrics. It minimizes MSE, maximizes accuracy, and consistently outperforms competitors in precision, recall, and F1-score. These findings collectively support the conclusion that Bi-LSTM + Attention is the most effective model for this task.

CONCLUSION

This research aimed to bridge the gap in Assamese natural language processing by developing a robust AI Conversational Agent. To achieve this, a novel Bi-LSTM with attention model was designed and trained on a meticulously curated Assamese religious text dataset. Complementing this, an Assamese NLP library was constructed, including essential tools like a stemmer and stop word remover. Comparative analysis against six other models underscored the superior performance of our proposed model, achieving an impressive accuracy of 89.99%. This study represents a pioneering effort in Assamese AI Conversational Agent development, given the paucity of resources and tools for this language. We anticipate that our work will stimulate further research and innovation in this domain.

REFERENCES

Abhay Chopde, M. A. (2022). Chatbot Using Deep Learning. Chopde, Abhay and Agrawal, Mohit, Chatbot Using Deep Learning (June 19, 2022). Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2022, Available at SSRN: https://ssrn.com/abstract=4140506 or http://dx.doi.org/10.

Barman, A. K. (2013). WordNet Based Information Retrieval System for Assamese. In 2013 UKSim 15th International Conference on Computer Modelling and Simulation (pp. pp. 10-12). IEEE. https://doi.org/10.1109/UKSim.2013.90.

Costa. (2018). Conversing with personal digital assistants: On gender and artificial intelligence. Journal of Science and Technology of the Arts , 10(6), 59-72.

(2020). Global Chatbot Trends Report - 2020. Mindbrowser.

Guendalina Caldarini, S. J. (2022). A Literature Survey of Recent Advances in Chatbots. Information , 13(1).

IBM. (2023). Chatbots: What they are and why they're important. Retrieved from IBM: https://www.ibm.com/topics/chatbots

Investopedia. (2022). Chatbot. Retrieved from Investopedia: https://www.investopedia.com/terms/c/chatbot.asp

Jia, J. (2003). The study of the application of a keywords-based chatbot system on the teaching of foreign languages. arXiv .

M P Bhuyan, R. P. (2019). Information Retrieval In Assamese Using Wordnet & Assamese Wikipedia. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH , 8(9), 01-10.

M. Kowsher, F. S. (2019). Doly: Bengali Chatbot for Bengali Education. 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), (pp. pp. 1-6). Dhaka, Bangladesh.

Madhumitha, S. K. (2019). Interactive chatbot using AIML. Ijana , 1(1), 62-66.

Masum, A. K. (2021). Transformer Based Bengali Chatbot Using General Knowledge Dataset. arXiv , preprint arXiv:2111.03937.

Qiming Bao, L. N. (2020). HHH: An Online Medical Chatbot System based on Knowledge Graph and Hierarchical Bi-Directional Attention. ACSW '20: Australasian Computer Science Week 2020. Melbourne.

Rita Chakraborty, S. K. (2015). Modelling a Question-Answering System using Structured Representation of Assamese Text. International Journal of Artificial Intelligence & Applications , 6(6),01-10,DOI:10.5121/ijaia.2015.6601.