Amirkabir Journal of Mechanical Engineering

Amirkabir J. Mech. Eng., 55(8) (2023) 193-196
DOI: 10.22060/mej.2023.22001.7553

Designing and building a dialogue mechanism suitable for RoboPuppet with using deep inference learning

M. Amirkhani, H. Shahbazi*

Department of Mechanical Engineering, University of Isfahan, Isfahan, Iran

ABSTRACT: This research endeavors to construct a mechanism, blending text mining and natural language processing, to apply a deep learning dialogue and deep reasoning approach to "Puppet robot." Historically, tent dolls have been an ancient method of interacting with audiences, being directly managed by an operator. With breakthroughs in artificial intelligence and deep learning, it is now possible to reduce the dependence of tent dolls on operators, thereby enabling them to communicate intelligently with audiences. The robot, by identifying the audience's Persian speech, ascertains a fitting answer to their inquiries and broadcasts it in audible Persian. The dialogue mechanism, deeply ingrained in a deep learning algorithm, identifies the user's question and proffers a range of possible answers from the robot's dataset categories. Utilizing the highest probability, the category containing the user's question is identified, and responses to those questions are selected at random. Additionally, the Robo Tent Dialogue mechanism comprises several uncomplicated conditional sections that can furnish suitable responses to repetitive or inappropriate questions. Through diverse training and by altering parameters in the robot's deep learning model, using a 64-class dataset, results reveal that the application of technologically advanced, high-neuron layers outperforms multi-layers without detrimentally impacting the model's final accuracy.

Review History:

Received: Dec. 09, 2022
Revised: Jul. 19, 2023
Accepted: Oct. 03, 2023
Available Online: Nov. 27, 2023

Keywords:

Puppet robot
artificial intelligence
deep learning
neural network
intelligent response

1- Introduction

Lee et al. emphasized the importance of semantic analysis in English question classification [1]. Xiong et al. proposed DMN+, a model for question answering without fact labels [2]. Chang et al. discussed issues in voice-controlled robot development and introduced a new model [3]. A study [4] described a robot that can engage in natural conversations and was perceived as human-like. Another research [5] presented a superior multilingual ASR model supporting 53 languages. Lastly, [6] delved into the possibilities of BCI technology in the Metaverse, suggesting a brain-to-speech system using imaginary speech.

In the study, a human-robot dialogue interaction is designed using a fabric doll as shown in Figure1. The robot picks up human speech via its microphone and sends the audio to a computer through Bluetooth. This audio is converted into text by a conversational model which then determines the robot's audible response. This system promotes fluent two-way communication, leveraging the latest in machine dialogue technology. The robot's interaction is further enriched with features like servo motors for movement, a camera for the operator to view the user, and an ESP32 processor to manage these functions.

*Corresponding author's email: shahbazi@eng.ui.ac.ir

This article outlines various dialogue mechanism methods and then selects the most suitable method for tent placement based on the pros and cons of each. By adjusting parameters in the deep learning network model, like the number of layers, neural neurons, and activation functions, the most optimized model is developed and its results are showcased.

2- Methodology

This model has 64 different classes and can answer 64 different questions. The model architecture consists of two compressed layers and a Dropout with the random deletion rate of the first Dense layer neurons of 0.4.

Experiments with labeled questions identified an overfitting problem, causing the model to require exact matches and struggle with minor variations. Increasing question quantity under labels allows the model to focus on key sentence terms, enhancing label identification. Shuffling words within sentences didn't significantly boost learning. A method for question answering uses auxiliary text from which relevant questions and answers are derived, helping the model identify patterns between text and answers. Sajjad Ayoubi introduced a Persian-specific model, PersianQA, containing around 9,000 titles, each with 5 to 10 questions. A drawback is its 500-word text limit. For robot integration, shorter texts are stored in its memory. When queried, the robot seeks answers internally or, if needed, uses PersianQA.

Copyrights for this article are retained by the author(s) with publishing rights granted to Amirkabir University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit https://www.creativecommons.org/licenses/by-nc/4.0/legalcode.

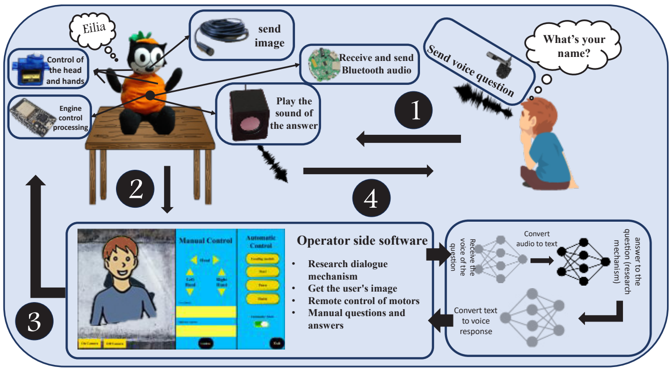

Fig. 1. The robot was designed and built to implement the research question and answer mechanism

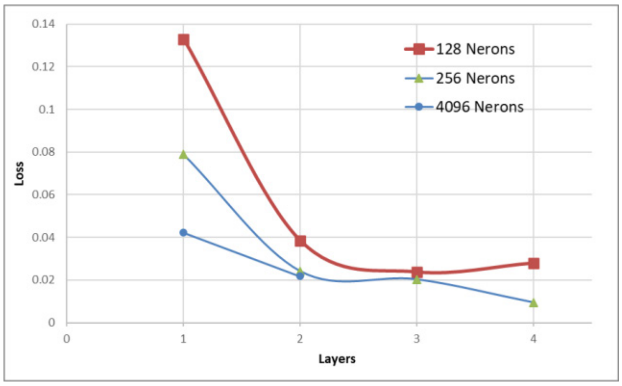

Fig. 2. Error diagram according to the number of layers in three values of 128, 256, and 4096 Dense network neurons

3- Discussion and Results

The random Q&A approach was selected for the tent's Q&A system, enhancing the robot's conversational ability. The neural network uses Fully Connected and Dropout layers to prevent data overlap. The SGD algorithm, aimed at minimizing the cost function reflecting the difference between predicted and actual labels, is employed. The optimizer's primary parameters include a learning rate of 0.01, a factor of , and momentum of 0.9. The cost function, cross-entropy, targets categorization tasks. The objective is to minimize this function to better match predictions to actual labels. The network's effectiveness was assessed through 10 tests over 100 training sessions across 64 to 78 categories. Figure 2 showcases error rates and the effect of layer increases, respectively. Training sessions took between 80 to 1200 milliseconds each.

According to this form, increasing the number of layers in the number of neurons above 128 has no effect on educational accuracy. This can be due to the small number of Q&A classes. Also, with the increase in the number of educational neurons, accuracy increases. In network architecture with the low number of neurons, increasing layers do not necessarily increase accuracy. According to the results, the random Q&A network architecture, with the number of neurons 4096 and a dense training layer, reached 0.042 and 0.98.

To improve this model, well-regarded evaluation metrics like Precision, Recall, and F1 -Measure are employed. Table 1 compares the evaluation of a random Q&A mechanism with other models from the literature. Analysis indicates that the unique feature of the random Q&A mechanism in Robo Tent is its support for Persian, unlike other models that focus on English. This specific focus on Persian enhances its capability to process and comprehend the language semantically. Table1. Comparing evaluation criteria of random question and answer dialogue mechanism

Experiments adjusted the question-answer mechanism for 64 classes in audio mode, aiming to reduce unanswered questions. The strategy allows the robot to respond even if uncertain, enhancing user interaction, as continuous answers, even if potentially erroneous, boost user engagement. Table 1 reveals significant results for the Dialogue and Random Q&A mechanism. The model excelled in Precision at 85%, confirming its accuracy, and in Recall at 90.65%, indicating its capacity to address a vast array of questions. The Robo Tent model scored 87.73% in the F1-Measure, representing a balance between Precision and Recall. The survey suggests the Robo Tent model's efficacy in providing accurate answers and covering numerous questions. The F1-Measure affirms its balanced performance. The model processes a sentence input, outputting a numerical value between 0 and 5, indicating the detected emotion.

Table 1. Comparing evaluation criteria of random question and answer dialogue mechanism

Evaluation models/criteria	Precision %	Recall %	F1-measure %
Ref [7]	51.50	80.41	55.9
Ref [8] no T5 method	62.62	72.50	70.75
Ref [8] with T5 method	77.78	77.75	77.78
Ref [1] MultinomialNB method	78	71	74
Ref [1] BernoulliNB method	73	75	74
Ref [21] Logistic regression method	83	82	82
Ref [21] LinearSVC method	80	80	80
Ref [21] The hybrid model method	83	83	83
This article	85	90.65	87.73

4- Conclusions

In this article, the mechanism of dialogue mechanisms was designed and implemented for a tental robot. Among these mechanisms, the question diagnosis method was examined by random answers and the feelings of emotions. All questions and answers have been implemented in Farsi and the user can communicate with the Robo Tent. The error is 0.998 and the error is 0.042. Therefore, by identifying the user's question, the class is selected by the deep learning model and the random response from the set of answers intended for that class is shown as an output.

References

[1] E.H. Almansor, F.K. Hussain, O.K. Hussain, Supervised ensemble sentiment-based framework to measure chatbot quality of services, Computing, 103 (2021) 491-507.

[2] C. Xiong, S. Merity, R. Socher, Dynamic memory networks for visual and textual question answering, in: International conference on machine learning, PMLR, 2016, pp. 2397-2406.

[3] P. Chang, S. Liu, D.L. McPherson, K. Driggs-Campbell, Learning visual-audio representations for voice-controlled robots, in: 2023 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2023, pp. 9508-9514.

[4] T. Minato, K. Sakai, T. Uchida, H. Ishiguro, A study of interactive robot architecture through the practical implementation of conversational android, Frontiers in Robotics and AI, 9 (2022) 905030.

[5] A. Conneau, A. Baevski, R. Collobert, A. Mohamed, M. Auli, Unsupervised cross-lingual representation learning for speech recognition, arXiv preprint arXiv:2006.13979, (2020).

[6] S.-H. Lee, Y.-E. Lee, S.-W. Lee, Toward imagined speech based smart communication system: potential applications on metaverse conditions, in: 2022 10th International Winter Conference on Brain-Computer Interface (BCI), IEEE, 2022, pp. 1-4.

[7] J.C. Vasquez-Correa, J.C. Guerrero-Sierra, J.L. Pemberty-Tamayo, J.E. Jaramillo, A.F. Tejada-Castro, One system to rule them all: A universal intent recognition system for customer service chatbots, arXiv preprint arXiv:2112.08261, (2021).

[8] J.J. Bird, A. Ekárt, D.R. Faria, Chatbot Interaction with Artificial Intelligence: human data augmentation with T5 and language transformer ensemble for text classification, Journal of Ambient Intelligence and Humanized Computing, 14(4) (2023) 3129-3144.

HOW TO CITE THIS ARTICLE

M. Amirkhani, H. Shahbazi, Designing and building a dialogue mechanism suitable for RoboPuppet with using deep inference learning , Amirkabir J. Mech Eng., 55(8) (2023) 193-196.

DOI: 10.22060/mej.2023.22001.7553