Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777196

DESIGNING A CONVERSATIONAL AI-TELEMARKETING CHATBOT SYSTEM FOR ADMINISTRATIVE CUSTOMER SERVICE

Hyeonji Kim [1], Yoosoo Oh [2,*]

[1]Dept. of Information and Communication Engineering, Daegu University, Gyeongsan-si, Republic of Korea

[2]School of AI, Daegu University, Gyeongsan-si, Republic of Korea

Abstract: With the advent of ChatGPT, the market for chatbot systems was revitalized. However, the existing chatbot system conducts scenario-based learning through vast open data. Therefore, generating an answer corresponding to a specific conversation takes time and effort. In this paper, we propose a chatbot system for civil complaint question and answer in a particular field. pa This per builds by converging topic modeling and sentence generation models to improve the performance of the chatbot system. In addition, it is made to enable natural conversation by learning daily conversation and emotional sentence datasets. Finally, in this paper, an individual database is provided to process the personal information of the civil complaint response service.

Keywords; Chatbot System, Topic extraction model, Generative Pre–Training (GPT), Database

1. Instruction

The AI chatbot is a technology that combines chat and interactive interfaces and refers to an interactive system that interacts with people based on machine learning or deep learning technology. [1] According to Market & Market, a market research institution, the global chatbot market is expected to grow at an average annual rate of 23.5 percent and reach 10.5 billion dollars (13.188 trillion won) by 2026. While demand for interactive services has increased due to the COVID-19 pandemic, complaints and responses have been delayed due to the closure of the complaint center space, causing great inconvenience to customer convenience. As demand for AI chatbots increases, interactive chatbots are being used in various fields and are developing into complaints business development and platforms. Many public institutions have recently used the complaint Q&A service as a chatbot service to assist with their work. [2] Public institutions currently use chatbot systems are at risk of management by manually storing customer personal information. Also, the existing chatbot system must generate appropriate answers when entering unlearned sentences. Therefore, this paper proposes AI-Telemarketing Chatbot, an interactive AI chatbot system for responding to customer complaints. The proposed AI-Telemarketing Chatbot is an interactive system

* Corresponding author: Yoosoo Oh, - Email: yoosoo.oh@daegu.ac.kr
- Address: (38453) 201, Daegudae-ro, Gyeongsan-si, Gyeongsangbuk-do, S.Korea

This research was supported by Daegu University Research Grant, 2018

combining topic modeling and autoregressive language models.

Topic modeling is a statistical model for discovering potential themes in a document. LDA, LSA, and KoBertopic are types of topic modeling.[3] Among the autoregressive language models, KoGPT-2 released by SKT is a model that trains Hangul data. KoGPT-2 is based on OpenAI's GPT-2 model.[4] This paper identifies the complaint category of the petitioner through topic modeling and generates and provides appropriate answers through KoGPT-2. This paper established an individual database to collect and process personal information in complaints automatically. The proposed system proposes a chatbot using a voice function to provide responding to complaints services. In addition, this paper learns specific fields, daily conversation data sets, and emotional data sets to provide natural response generation functions.

2. Related Research

Kim Hyun-Gu, Hwang Jong-Kook, and Hwang Jae-Seung [6] applied LSA and LDA to the paper on Korean Wind Engineering research topics and trends. They classified 20 topics using LSA and LDA. The top 10 LSA and LDA topics extracted were derived with 77% and 70% of topics, respectively. They confirmed that 36 LSA and 72 LDA were used as topic-composed words, excluding repetitive words. Therefore, we have verified that LSA has a higher similarity between topics than LDA in research and trend identification of papers in Korean Wind Engineering.[6]

Jung Sun-Woo, Choi Eun-Sung, Ahn Sun-Kyu, Kang Young-Jin, and Jung Seok-Chan [5] proposed a scenario-based AI voice chatbot system for museum guidance. The proposed AI voice chatbot system established a sentence intention classification model and proceeded with user sentence intention classification. The performance of the sentence intention classification model was derived using a softmax regression function and showed a classification accuracy of about 82%.[5] Wang Young-Min and Oh Dong-Han [15] designed the admissions counseling chatbot and analyzed the utility of simplifying admissions counseling tasks using chatbots. The admissions counseling chatbot was designed through Kakao Business' "*Kakao i Open Builder*." They collected repetitive and simple question data to proceed with their learning. The designed admissions counseling chatbot provides information based on what has been learned. Through the proposed system, the number of questions on the 2021 and 2022 entrance examination counseling boards was 121 and 83, respectively, down 28 times as of 2021 and down 31.4% as of 2021.[15]

Song Yoon-Kyung, Jung Kyung-min, and Lee Hyun [13] proposed BERGPT-Chatbot that can relieve negative emotions based on text. BERGPT-chatbot used pipeline algorithms to link KR-BERT and KoGPT2. They trained KR-BERT and conducted emotional classification. BERGPT-chatbot learned KoGPT2-chatbot using non-emotional data and existing emotional data after emotion classification. Using PPL measurement, we measured and compared the performance of the KoGPT2 chatbot with the system proposed by them. Chatbots modeled on KR-BERT and KoGPT2 connected by pipeline algorithms have performed better than ordinary KoGPT2 chatbots.[13]

Kim Hyun-Koo, Hwang Jong-Kook, and Hwang Jae-Seung confirmed that the performance of the theme model algorithm depends on the data used. Jeong Seon-woo proposed a chatbot system that derives scenario-based fixed questions and answers to classify sentence intentions. Song Yoon-Kyung, Jung Kyung-min, and Lee Hyun combined the emotion classification model and the sentence generation model to confirm that they performed better than the general sentence generation model. Wang Young-Min and Oh Dong-Han analyzed the usefulness of the examination counseling chatbot system that can reduce the physical requirements for entrance examination counseling work. However, the proposed entrance examination counseling chatbot cannot provide information by entering information not used for learning. In this paper, we select and use topic modeling appropriate for responding to customer complaint data through LSA, LDA, and KoBERT topic comparison analysis. In this paper, we apply a sentence generation model to generate appropriate answers to the conversation flow. This paper can also generate appropriate answers when entering information not used for learning.

3. Conversational AI Telemarketing Chatbot System for Customer Complaint Handling

This paper proposes an AI–Telemarketing Chatbot System. Fig.1 is a behavior diagram of the AI–Telemarketing Chatbot System. In this paper, we constructed a model using STT (Speech To Text) to convert user speech into text and TTS (Text to Speech) technology to convert text into speech. [5] The proposed system converts the petitioner's voice data into text through the Speech To Text (Speech) module. The topic extraction model can extract the topic of the petitioner's voice data content. The proposed system uses extracted topics and petitioner voice text data as input data for the sentence generation model. The sentence generation model generates a response corresponding to the input data. The AI–Telemarketing Chatbot creates the response text, which is delivered to the user in voice using Text To Speech (TTS) technology.

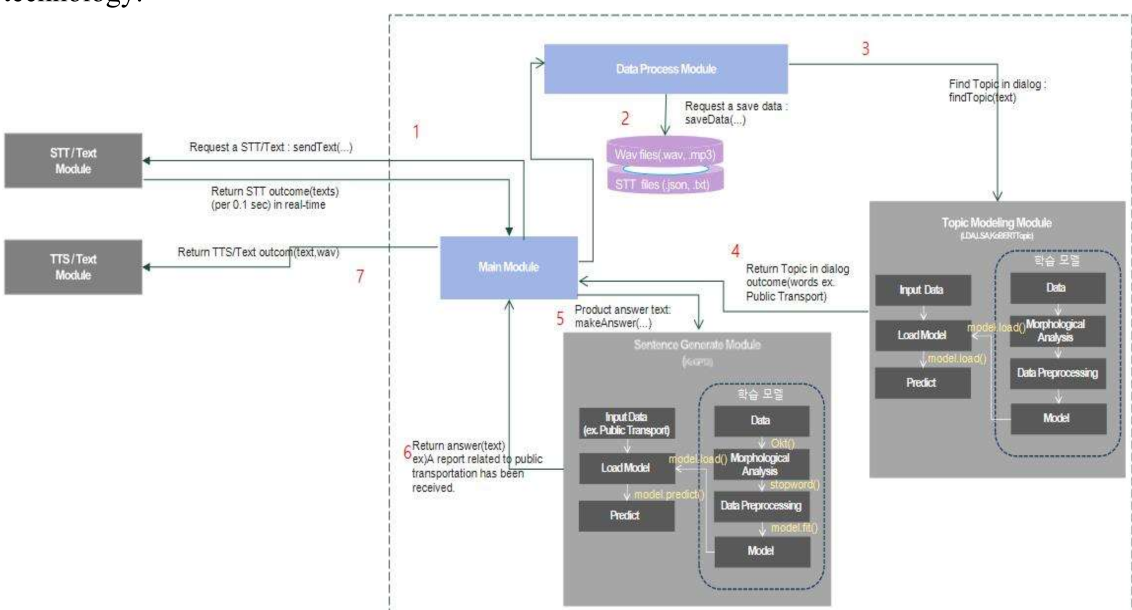


Figure 1. Diagram of our proposed AI-Telemarketing Chatbot System

3.1 Topic Modeling Module

The AI–Telemarketing Chatbot System used the LSA algorithm to construct a topic modeling module. Latent Semantic Analysis (LSA) reduces dimensions and considers the possible meaning of words using the SVD matrix delivered to the TF-IDF matrix.[3] In this paper, topic modeling learning was conducted using AI Hub complaint (call center) question-and-answer data. The proposed system used 20,000 data from 81,205 categories from Dasan Call Center's public transportation, living sewerage, and general administrative inquiries among AI Hub complaint (call center) question-and-answer data. Before learning the data, Stopword processing was performed except for region names, exclamations, and other things, among the question-and-answer data of AI Hub complaints (call centers), as shown in Table 1. This paper conducted morpheme analysis using Konlpy's 0kt library after data preprocessing.

Table 1**.** Result of StopWord Preprocessing (*Translated from Korean to English, Our system used Korean.*)

	Before PreProcessing Stopword	After Preprocessing StopWord
1	Is there a subway in Gongju City?	Is there a subway
2	Then what time is the last bus in Pohang?	Then what time is the last bus
3	What time is the first bus in Changwon?	What time is the first bus

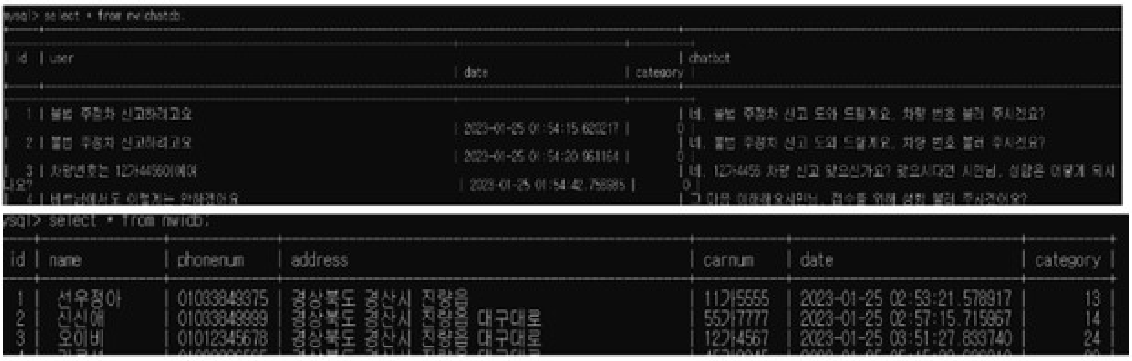
Table 2 shows the output of LSA algorithm learning results in the order in which the weight values are significant. For example, looking at the top five words printed, topics related to public transportation, such as city buses, public transportation, and subways, were extracted.

Table 2**.** Result of Trained LSA (*Translated from Korean to English, Our system used Korea n.*)

Topic	Result of extract result
1	['**City buses**', 'is', 'most', 'different', 'every route', 'city bus', 'is', 'faster']
2	['question', 'if', 'contact', '**Public Transportation**', 'consultation', 'here', 'center', 'hello']
3	['Do you have', 'maybe', '**Subway**', 'let's see', 'in a moment', 'once', 'check', 'thank you']

3.2 Answer Generation Model

The AI–Telemarketing Chatbot System constructed an answer generation model with KoGPT-2 algorithm fine-tuning. In this paper, we conducted the learning using AI Hub's Question and Answer data and Emotional Dialogue corpus data. In this paper, we use three data from the question-and-answer data of the Complaint (Call Center) related to public transportation, living sewerage, and general administration of Dasan Call Center. This paper uses 20,000 out of 81,205 complaints (call centers) question and answer data. The proposed system uses 5,000 additional data corresponding to human text one and system text one among AI Hub's emotional dialogue corpus data. The proposed system performed stopword preprocessing to remove the collected data, such as region names and symbols. After data preprocessing, this paper conducted morpheme analysis through Konlpy's 0kt morpheme analyzer. The proposed sentence generation system conducted a pre-trained model of KoGPT-2 learning using 25,000



data collected. The hyperparameters used during model learning are batch_size=64, max_len=256, learning_rate=3e-5, and epoch=50. Table 3 results from generating answers through the model learned in the pre-trained model, KoGPT-2.

Table 3**.** Result of Trained LSA (*Translated from Korean to English, Our system used Korea n.*)

Topi c	Example of Customer Complaint Content	Answers generated by the learning model
1	I want to report illegal parking	Yes, I can help you report illegal parking
2	Is there a subway in Pohang	I will look into it and let you know
3	I am so angry that the car is blocking the alley. Please take action.	I understand that feeling. When you are angry, how about taking a deep breath first

In this paper, emotional dialogue corpus data are also used for learning, enabling the generation of appropriate responses to conversations other than topics corresponding to customer complaints.

3.3 Voice Model, Database

In this paper, we use the Google Cloud API to apply STT and TTS modules to provide voice services. The proposed system used the STT module to convert the petitioner's voice (user) into text data. The answer text generated through the sentence generation model is provided by voice using the TTS module.[14]

This paper established a database to collect and manage personal information collected when receiving complaints. In this paper, we built a database using Mysql 8.0.32. Mysql supports multi-threading in open-source relational databases. Mysql supports all applications and is compatible with the Python programming language.[7] In this paper, we established a personal information collection database and a conversation content collection database between civil petitioners and counselors. Figure 2 shows the database items created for illegal parking.

The Personal Information Collection database contains columns by name, address, vehicle number, phone number, date, and category. The conversation content collection database consists of columns based on the content of the complainant's conversation, the content of the counselor's conversation, the date, and the category. This paper saves the collected database as an Excel file to enhance access to information.

The proposed system provides a UI similar to Fig.3. When the user presses the start consultation button, the system starts and collects information in the database in real time. The proposed system outputs the conversation between the user and the AI–Telemarketing Chatbot System as text.

Figure 2. Database about Illegal Parking

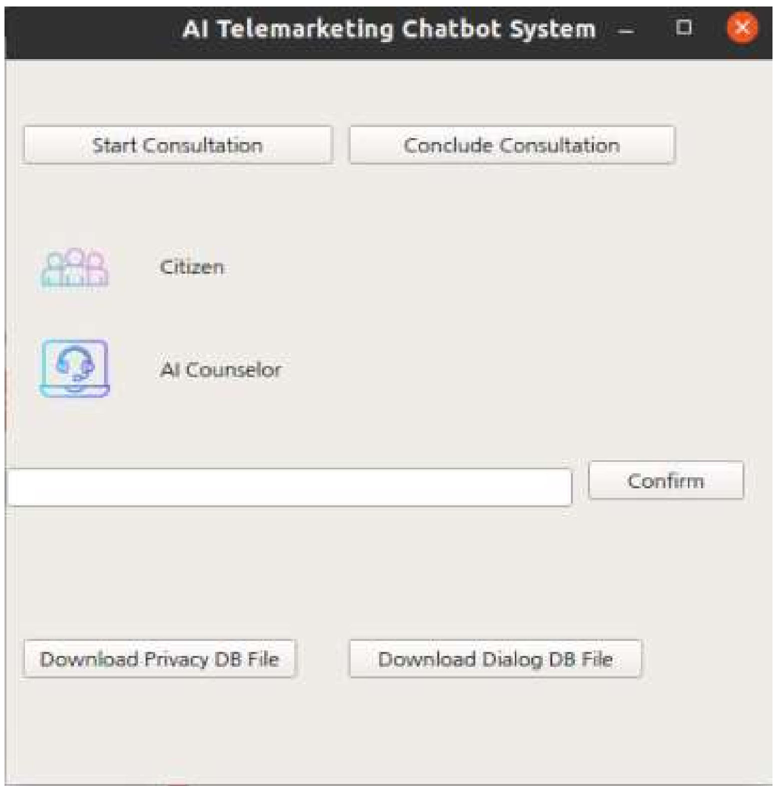


Figure 3. UI of our AI-Telemarketing Chatbot System

4. Experimental Result

In this paper, we evaluated topic modeling and chatbot systems. Topic modeling conducted a comparative analysis of LSA, LDA, and KoBertopic. Latent Dirichlet Allocation (LDA) is a method of extracting a topic by estimating the probability that a particular word exists in a topic and that a particular topic exists in a document as a combination probability.[8] KoBERTopic is a model that allows BERTopic TF-IDF to be applied to Korean data. KoBERTopic is a topic model technology that uses the BERTopic model to create clusters while maintaining essential words.[8]

This paper evaluated topic modeling algorithms using Recall, Precision, F1-Score, and Perplexity (PPL) evaluation methods. The recall indicates the proportion of actual topics found by the model among topics.[9] Recall measures how much the actual topic has been reproduced, and the closer it is to 1, the better the performance. Precision is an indicator of how accurately the predicted value of the model is predicted.[9] F1-Score is a harmonious combination of two methods, Recall and Precision, and is an indicator of a harmonious average and is expressed as a single numerical value.[9] Perplexity quantifies the degree of confusion. Perplexity interprets that the more candidates for words to be considered when creating sentences are confusing, and the lower the result value, the better the performance.[10] Figure 4 is a graph of the measurement result. Figure 5 is the result of graphing the performance evaluation results when 30 Epoch repeatedly performs each topic modeling algorithm. Table 4 is a table that measures the time taken to derive the topic modeling results. Recall, Precision, F1-Score, and PPL evaluation results show that LDA performs less than LSA or KoBERTopic. On the other hand, there is little difference in the performance of LSA and KoBERTopic. However, LSA algorithms derive better performance results than KoBERTopic in Recall, Precision, F1-Score, and PPL evaluation results. In addition, as shown in Table 4, it can be seen that KoBERTopic takes about 1,000 times more than LSA.

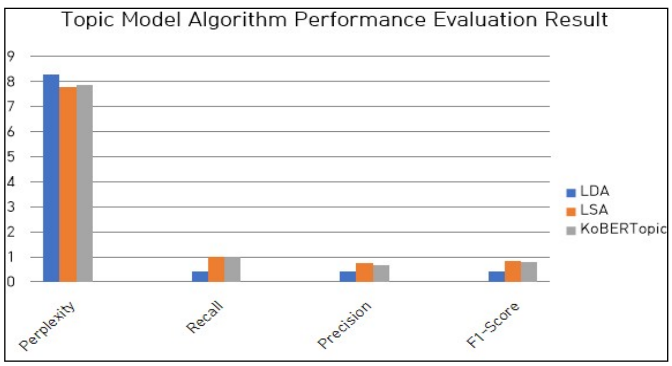


Figure 4. Topic Modeling Evaluation Result

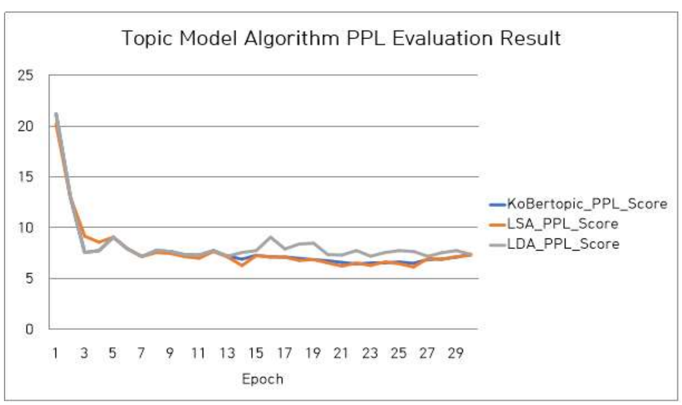


Figure 5. PPL(Perplexity) Result of each topic model algorithm

Table 4. Time taken to output Topic Modeling

Model Name	Time(sec)

LDA	0.015644
LSA	0.029104
KoBERTopic	49.658073

In this paper, we compared and analyzed the proposed AI telemarketing chatbot system and two commercial chatbot systems. This paper used Danbi and Naver Clova chatbot programs as commercial chatbots. This paper used the same data as the proposed AI telemarketing chatbot system for the commercial chatbot learning we use. In this paper, Coherence and BLEU Core measurements were performed to evaluate the performance of the chatbot system. Consistency measures similarity between sentences using cosine similarity measurement methods.[12] The BLEU score is measured for words in sentences with sequential information. The BLEU score measures the probability of the next word appearing for a word, and the higher the value, is better.[11] Figures 6 and 7 graphs represent cosine similarity and BLEU score measurements for the chatbot system. While the proposed system can generate appropriate responses to unlearned conversations, commercial chatbots cannot. Therefore, the measurement results show that the proposed AI telemarketing chatbot system performs better than other commercial chatbots.

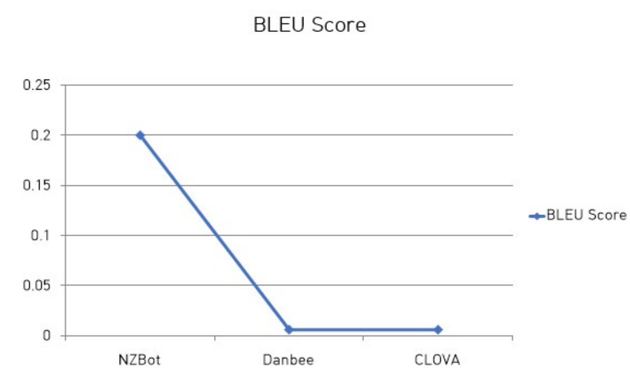


Figure 6. Chatbot's BLEU Score

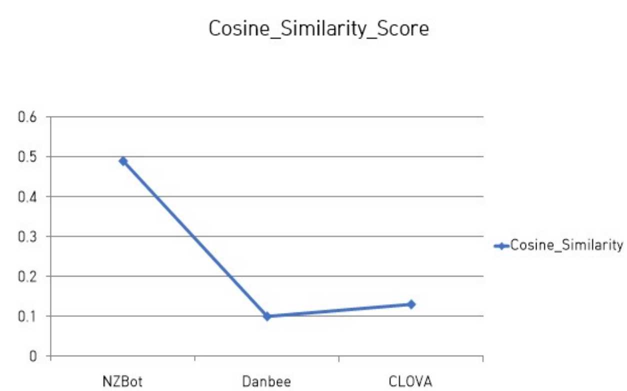


Figure 7. Chatbot's Cosine Similarity Score

5. Conclusion

In this paper, we compared and analyzed various topic modeling and sentence generation models. In this paper, we proposed an AI telemarketing chatbot system by selecting the highest-performance algorithm according to the analysis results. The proposed system further learns an emotional dataset for sentence generation models, enabling proper response generation for unlearned conversations. In addition, this paper is easy to collect and manage personal information through database construction.

The current proposed model has a heavy capacity of about 4GB. Therefore, future research will build a lightweight model through KoGPT-2 fine-tuning. We are currently constructing a sentence generation model using KoGPT-2. Future research will improve the model's performance through comparative analysis and application with various sentence generation algorithms such as GPT-4. The proposed evaluation method has different performance evaluation results depending on the number of test data sets. Therefore, we will collect more test data sets in future studies and analyze the evaluation results.

Reference

[1] Kim, Sung-geun., Shin, Min-chul., and Kang, Ju-young. 2018. "Introduction to chatbot technology and case analysis." information and communications 35.2, no. 8 (2018): 21-28.

[2] Ahn, Sung-hoon. 2022. “Exploring Ways to Build a Chatbot System for Effective Guidance of University Undergraduate Operation.” Research on Creative Information Culture 8, no.3 (2022): 145-152.

[3] Kim, Sun-wook., and Yang, Kiduk. 2022. "Topic Model Augmentation and Extension Method using LDA and BERTopic." Journal of the Society for Information Management 39, no.3 (2022): 99-132. doi: 10.3743/KOSIM.2022.39.3.099.

[4] Lee, Hyo-min. 2022. "Development of emotional chatbot system through multi-head generation of GPT" Graduate School of Yonsei University, a master's degree thesis in Korea, 2022 Seoul

[5] Jeong, Seon-woo., Choi, Eun-sung., Ahn, Sun-kyu., Kang, Young-jin., and Jeong, Seok-chan. 2022. "Implementation of Scenario-based AI Voice Chatbot System for Museum Guidance." Journal of the Korean Big Data Society 7, no.2 (2022): 91-102.doi: 10.36498/kbigdt. 2022.2.91.

[6] Kim Hyun-koo, Hwang Jong-kook and Hwang Jae-seung. "Topic Modeling of The Journal of the Wind Engineering Institute of Korea using LSA and LDA." Korean Wind Engineering Journal 24, no.3 (2020) : 113-120. doi: 10.37109/weik.2020.24.3.113.

[7] Jeong, Jae-woo. 2001, "Building an Internet Survey Analysis System Using Apache PHP Mysql and Application of Webmining." Graduate School of Hanyang University, Seoul, 2001

[8] Kim Sun-wook and Yang, Kiduk. "Topic Model Augmentation and Extension Method using LDA and BERTopic." Journal of the Society for Information Management 39, no.3 (2022): 99-132. doi: 10.3743/KOSIM.2022.39.3.099

[9] Nindam, Somsawut., Manmai, Thong-oon., Sung, Thaileang., Wu, Jiahua., and Lee, Hyo Jong. 2022. "Human Activity Classification Using Deep Transfer Learning." Proceedings of the Korea Information Processing Society Conference, vol. 29, no. 2(2022): 478-480,

[10] Lee, Daeyoung., and Lee, Hyunsook. "Exploring methods for determining the appropriate

number of topics in LDA: Focusing on perplexity and harmonic mean method," educational evaluation study 34, no.1 (2021): 1-30.

[11] Cho, Junghyun., Jung, Hyunki., Park, Chan-young., and Kim, Yu-seop. "An Autonomous Assessment of a Short Essay Answer by Using the BLEU." Academic Conference of the hci society of Korea (2009): 606-610.

[12] Hwang, Chi-gon., Yoon, Chang-pyo., and Yoon, Dae-yeol. 2021. "Sentence Similarity Analysis using Ontology Based on Cosine Similarity." Journal of the Korean Society of Information and Communication Conference 25, no.1 (2021): 441-443.

[13] Song, Yun-kyung., Jung, Kyung-min., and Lee, Hyun. 2021. "A BERGPT-chatbot for mitigating negative emotions" Journal of the Korean Computer Information Society, 26, no.12 (2021): 53-59.doi: 10.9708/jksci.2021.26.12.053

[14] Mo, Shi-eun., Yang, Hye-in., Cho, Eun-bi., and Yoon, Jong-ho., 2022. "AI voice phishing prevention solution using Open STT API and machine learning." Proceedings of the Korea Information Processing Society Conference, vol. 29, no. 2, 1013-1015, 2022.

[15] Wang, Young-min., and Oh, Dong-han. 2022. "Chatbot Design and Efficiency analysis through the use of chatbot service on the Entrance Counseling." Journal of the Digital Content Society 23, no. 9 (2022): 1787-1794.