Implementation of the GPT-4 Language Model for Processing Users' Responses in Social-Psychological Services within Digital Communication Processes

Hryhorii Hnatiienko [1,†], Olena Prysiazhniuk [2,*,†], Anna Puzikova [2,†] and Olena Blyzniukova [2,†]

[1] *Taras Shevchenko National University of Kyiv, Volodymyrs'ka str, 64/13, Kyiv, 01601, Ukraine*
[2] *Volodymyr Vynnychenko Central Ukrainian State University, Shevchenka str, 1, Kropyvnytskyi, 25006, Ukraine*

Abstract

The article examines some key points of using artificial intelligence tools in digital communication processes in the field of social services for prompt social and psychological support of users. The problems of processing of social service users' responses to chatbots inquiries are analyzed. A way to process these results using the GPT-4 language model and the OpenAI API is proposed. The impact of the ChatGPT-4 temperature parameter value settings on the results of processing non-standard user responses with elements of fuzziness was studied. The proposed approach allows developers to integrate GPT-4 into their applications conveniently with the aim of automating the process of processing user responses and offloading operators of social support centers, which is relevant in the context of a personnel shortage in the Ukrainian labor market.

Keywords

Artificial Intelligence (AI), large language models (LLM), ChatGPT, digital communication, secure data processing, social and psychological services1

1. Introduction

Digitalization involves the total introduction of digital tools into communication processes at all levels and for various interaction channels. The practice of implementing AI-based chatbots has proven to be effective for 24/7 personalized client support and has effectively replaced operators, providing clients with answers in template situations. But in the process of processing the client's open-ended responses, it often turns out that, even when he is offered template answers, some people, due to situational circumstances, current emotional state, and due to experienced stress, may give answers outside the template options. Recognition to such answers required the intervention of a human operator to clarify the information that the client meant. In most cases, this happens without additional dialogue with the client, thanks to the inherent human ability to recognize fuzziness and ambiguity inherent in natural language and interpret information within a clearly defined framework [1].

The military actions in Ukraine led to an increase in the number of people affected by the war who need psychological and social support in various spectrums of life, namely, emotional, informational, instrumental, financial, legal, providing opportunities for training and self-development, social integration, and other forms of assistance. Existing centers of social and psychological support have a limited resource of operators, therefore, to work with this population

Information Technology and Implementation (IT&I-2024), November 20-21, 2024, Kyiv, Ukraine
* Corresponding author.
† These authors contributed equally.
g.gna5@ukr.net (H. Hnatiienko); elena_drobot@ukr.net (O. Prysiazhniuk); a.v.puzikova@cuspu.edu.ua (A. Puzikova); bllyus@yahoo.com (O. Blyzniukova)
0000-0002-0465-5018 (H. Hnatiienko); 0000-0002-7135-3124 (O. Prysiazhniuk); 0000-0002-6843-5583 (A. Puzikova); 0000-0001-8159-2391 (O. Blyzniukova)

category effectively, it is necessary to carry out operational electronic communication using artificial intelligence technologies, in particular, large language models (LLM). Large language models have created opportunities to develop chatbots that can support complex question and answer scenarios. But for many practical situations we still lack an understanding of how meaningfully a chatbot can process users' responses in digital communication processes, functioning as a human operator, i.e. simulating the activity of an operator [2].

In the research we analyze the possibilities and effectiveness of using the ChatGPT-4 language model in the processes of digital communication with clients of social support centers to solve problems:

- monitoring of customer needs in order to respond promptly to identified problems and provide professional support from specialists;
- reminding clients about timely completion of current tasks (in particular, about the terms of assigned services and coordination of relevant actions);
- providing recommendations regarding further referrals to specialists for the purpose of psychological support and social adaptation.

In order to determine the possibility and effectiveness of using ChatGPT to recognize and interpret users' open responses in digital communication processes experimentally, the researchers proposed as a working hypothesis to consider the ChatGPT model as a fuzzy system that can capture the uncertainty and ambiguity inherent to the natural language. In order to measure the overall effectiveness of ChatGPT in recognizing and interpreting users' responses, expert evaluation methods and statistical data processing techniques are used.

Overview of the capabilities of the ChatGPT language model for recognizing and processing users' responses in digital communication processes

Large language models are a field of artificial intelligence at the intersection of linguistics and computer science [3]. By learning from vast amounts of text data, language models can, in particular, interpret the text entered by the user and generate human-readable text in response. There are different types of pre-training architectures, including autoencoding models (e.g., BERT), autoregressive models (e.g., GPT), and encoder-decoder models (e.g., T5) [4]. GPT (Generative Pretraining of Transformers) is a type of neural network architecture useful in chatbots, which makes them particularly effective for imitating human conversations [5].

The GPT architecture was used as the basis for the development of OpenAI's pretrained artificial intelligence model ChatGPT. The ChatGPT model is designed for natural language communication and is built using sophisticated Natural Language Processing, supervised learning, and reinforcement learning to understand and generate text similar to human-generated text [6].

Let's dwell on those characteristic properties of the ChatGPT model that led to its selection for solving the problem described in the research.

ChatGPT was developed using a two-step process involving unsupervised pre-training followed by supervised tuning [7]. This model was previously trained on a massive text corpus (which includes various sources such as books, articles, reviews, online chats, and human-generated data). This gave ChatGPT such a property as "understanding" the nuances of language, which in turn gave it the opportunity to generate quite accurate answers, even in the case of processing complex and ambiguous contexts [7]. After the pre-training phase, the model was fine-tuned using such further tasks as completion of the text, answering questions and conducting a dialogue [6]. Thus, the main advantage of this language model is its ability to 'understand' the context of input data and generate the correct result with a high degree of probability [8].

ChatGPT belongs to Large Language Models – general-purpose models that are designed to work with a wide range of tasks. According to the results presented in the research [8] ChatGPT demonstrated a significant level of proficiency in answering the questions of a multimodal 12-item exam. The authors of the research [9] point to the impressive performance of ChatGPT in various language tasks and tests, which established it as one of the leading language models in the world.

Along with significant positive results, researchers (including the OpenAI company) point to the problem of "hallutinations", which consists in the fact that the model can generate plausible, but unreliable or nonsensical information [10, 11]. Chat GPT developers pay attention to the model's sensitivity to settings in input phrases or retries of the same request. In other words, rephrasing a request can contribute to generating a more accurate and correct answer [10].

Among other problems of the ChatGPT model, the researchers note verbosity and abuse of repetetive phrases (the reason for this is bias in the training data, since the trainers of the GPT-3.5 language model preferred long comprehensive answers [10,12]), inaccuracy (current models usually "guess" what the user meant). According to the developers, ideally, the model would ask clarifying questions in response to an ambiguous user request [10, 12], bias, responsibility for the created content, transperancy, ethical issues regarding authorship, lack of creativity (the reason is that ChatGPT performs repetitive text generation, which is based on pre-loaded data), etc.

OpenAI's model bias measures are part of its commitment to foster a safe artificial intelligence ecosystem [13] and include API moderation to prevent and block dangerous content. As part of these activities, OpenAI engaged more than 50 experts in such areas as AI alignment risks, cyber security, biorisks, trust and security, and international security to test the model competitively to analyze the additional capabilities of the updated GPT-4 model. According to their conclusions, the behavior of the model was tested in high-risk areas, the assessment of which requires experience. Feedback and data from these experts were used to soften and improve the model [14].

Other problematic issues listed above, according to the authors, are not essential within the framework of using the model to achieve the goal of this work, and therefore are not considered in more detail.

It should be noted that the GPT neural architecture has been chosen as a basis for the development of the other language models, which also show good initial results. For example, the DialoGPT model, created for generating responses in the process of a dialogue, allows processing multiple inputs and generating highly personalized responses that are more relevant, meaningful, and consistent with the context, compared to other powerful frameworks [15].

Viewing ChatGPT as a fuzzy system that can capture the fuzziness and ambiguity inherent in natural language [16] allows this tool to be applied to analyze human texts or texts produced by generative artificial intelligence and use fuzzy logic to deal with the ambiguity of natural language and provide more flexible responses.

A detailed overview of the possibilities, prospects and potential of using ChatGPT in the areas of customer service, business operations and communications within the framework Industry 4.0, is given in the study [17]. In particular, among promising implementations of ChatGPT, a separate segment of ChatGPT applications for analysis and processing of clients' requests in communication processes stands out.

The analysis of publicly available sources shows insufficient attention and lack of research devoted to the use of the ChatGPT model to solve the problems of interpreting users' open responses.

The research task

The article discusses the results of using the ChatGPT-4 language model to process responses from users of social services at the Social Support Center (Kropyvnytskyi, Ukraine). To implement the research program effectively a series of studies was conducted. First, an express survey was conducted with the respondents of the approbation sample in the number of 138 people to identify persons in need of immediate psychological help. Questions were sent to the respondents. The

answer was intended to be open-ended, but typical templates provided expected positive ("yes"/1) and negative ("no"/2) answers. All answers were recorded in the array. Then, atypical users' responses were filtered from the response array. They went beyond the recommended template responses and accounted for 20.3% of the total number of processed responses. This array of atypical responses served as the source of raw data for the next two studies.

In the first study, the task of processing atypical users' responses was addressed to the ChatGPT-4 by formulating a corresponding request. This processing consisted of recognizing an atypical user response and assigning it to one of the response categories specified by the original template. To adjust the sensitivity of ChatGPT-4 to the recognition of respondents' answers, the adjustable parameter "temperature" was used.

In the second study, to access the effectiveness of ChatGPT-4 in recognizing atypical responses given by respondents outside of the instructions, the researchers provided peer review. For this, the question, the original answer of the respondent and the result of processing (recognition) of the answer by the GPT-4 chat were provided to the expert. The expert's task is to assess whether ChatGPT-4 recognized the user atypical response correctly and assigned it to the appropriate category of template responses. The research team consisted of an employee of the Social Support Center, a psychologist and a linguist. As a result of the expertise atypical users' responses were identified within the given template ("yes"/1 or "no"/2). Four categories of atypical answers were highlighted such as "explicitly positive", "implicitly positive", " explicitly negative " and "implicitly negative". The performance of ChatGPT-4 recognition of atypical responses was evaluated by comparing the processed response with a similar result obtained by an expert evaluation

ChatGPT settings

ChatGPT-4 has several parameters, the setting of which in accordance with the given task can significantly affect the result [18, 19].

For the problem considered in the research, the temperature parameter is relevant, which is used to control the degree of randomness or unpredictability of the model's responses in the context of ChatGPT. It determines how risky or conservative the responses will be depending on the value provided, which can range from 0 to 1. According to the documentation, values for the temperature parameter set in the range (0.2-0.5) tune the model to more predictable and conservative responses. This means that it will choose the most likely options more often, which makes the answers more accurate and with fewer errors. The value of the temperature parameter in the range (0.7-1.0) adjusts the model to more diverse and creative responses. In other words, the model will choose less likely options, which can lead to unexpected and original results. Thus, a temperature value closer to 0 makes the model more predictable, and closer to 1 makes the model more experimental.

It should be noted that optimizing the temperature parameter in GPT models is a powerful technique for developers to improve Human-ChatGPT collaboration. By experimenting with different temperature settings, developers can adapt ChatGPT to perform a variety of tasks. Experimentation is a key concept, as the optimal temperature setting may vary depending on the specific use cases of ChatGPT and the task requirements

An example of software processing of users' responses

Generative models of artificial intelligence, in particular ChatGPT-4, generate answers based on prompts, the detailed and thoughtful formulation of which contributes to obtaining accurate and relevant answers [20]. The formulation of a prompt according to the strategies for achieving better results [21] should include the following elements such as context, instructions, input data, output indicator. The prompt for ChatGPT-4, formulated in the context of the problem under consideration, is shown in Figure 1 and further is the value of the variable $condition, which is mentioned in the code section below.

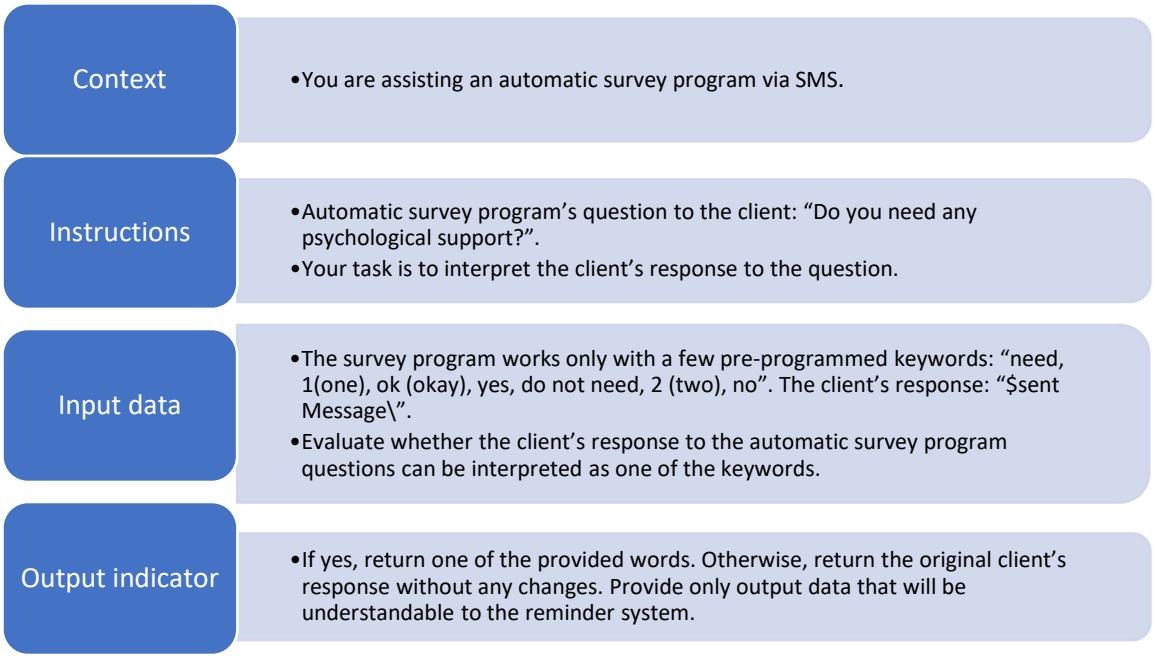

Figure 1: ChatGPT-4 prompt.

Let's consider an example of implementing the interaction of the application with the OpenAI API (application programming interface) for recognizing the text entered by the user.

A message to the chatbot user could be as following:

```
$sentMessage = 'Do you need any psychological support? Press 1 to confirm or 2 to decline.'
```

In the simplest case, to process the received results, you can create two arrays for standard positive and negative users' responses, to provide further processing of an uncertain result in another way (for example, sending a message to a human operator to contact the customer personally to clarify the result). Below is a sample code for handling users' responses.

```
if (in_array(strtolower($response), ['need', '1', 'ok', 'okay', 'yes'])) {
    $this->set('confirmation', 'User confirmed.');
}
elseif (in_array(strtolower($response), ['do not need', '2', 'no'])) {
    $this->set('confirmation', 'User did not confirm.');
}
else {
    $this->set('confirmation', 'Unable to determine user response.');
}
```

The use of the GPT-4 language model provides the possibility of additional analysis of the entered text in order to confirm the positive or negative response of the user.

Contacting OpenAI directly is done using the method getOpenAIResponse():

```
public function getOpenAIResponse(string $sentMessage, string $message, string $triggerWords): ?string
```

This method takes the following parameters as input:

```
$sentMessage – the message that was sent to the user;
$message – user's response;
$triggerWords – a list of possible user's responses.
```

The method body contains a request to OpenAI, the parameters of which are the name of the language model 'gpt-4', as well as an array with messages. The first element of the array is a message in which the context is specified, that is, a brief description of the situation with further instructions. The second element of the array is a message containing the user's response, which is recognized by OpenAI. Below is a code snippet that describes these steps:

```
  $response = OpenAI\Completion::create([
      'model' => 'gpt-4',
      'messages' => [
  [
      'role' => MessageRole::SYSTEM,
      'content' => $condition,
  ],
  [
      'role' => MessageRole::USER,
      'content' => $message,
  ],
      ],
      'temperature' => 0.5,
  ]);
```

The implementation of processing the response and returning its text part, or the null value, can be as follows:

```
  $data = json_decode($response);
          if (!empty($data->error)) {
              throw new \Exception($data->error->message ?? 'Unknown
error');
          }
          return $data->choices[0]->message->content ?? null;
```

The results

As a result of the express survey, whether the user needs urgent psychological help, approximately 23.9% of the answers (33 out of 138) turned out to be atypical. The answers that did not correspond to the proposed template (“yes”/1 or “no”/2), i.e., contained fuzziness and ambiguity (which was 28 responses, 20.3%), were selected among them and tested for recognition in ChatGPT-4. The answers with an unknown result such as "Maybe", "I don't know" (5 answers) were not used in the testing.

The performance evaluation of ChatGPT-4 recognition of atypical user responses took place in two versions, in Ukrainian and English. The impact of the temperature parameter settings on the quality of chat processing of users’ responses was also studied. The analysis was carried out with the temperature parameter, in the following modes: 0, 0.3, 0.5, 0.8 and 1.

Table 1 presents the atypical users’ responses, which are divided into four categories, the response recognition results from ChatGPT-4, and the expert evaluation of the recognition performance, given in terms of “+” (recognized) and “-“ (not recognized). Responses were processed with the temperature parameter value of 0.

When the temperature parameter value increased to 0.5, there was a qualitative jump in the ability of ChatGPT-4 to recognize two more atypical responses from users from the "Implicit positive responses" category. However, when the temperature parameter value was further increased to 1, there were no qualitative changes in the processing of users’ responses by ChatGPT-4. It should be noted considering the unexpected result which was obtained when the temperature parameter value was changed from 0.8 to 1 during the recognition of the implicitly negative (according to experts)

answer “I’m not sure it’s necessary” that there was a jump from “refusal” to “agreement”. According to the authors, similar results of response recognition require the further involvement of qualified specialists such as psychologists and linguists to provide consultations to the human operator to work with the client in order to clarify his needs.

Table 1
Results of processing ChatGPT-4 with Temperature=0 parameter value of atypical users’ responses.

Atypical users’ responses	Recognition result from Chat GPT-4	Expert evaluation
1. Positive responses		
1.1. Explicitly positive responses		
I need help	Need	+
I would not refuse help	Yes	+
I would like to get a consultation from a psychologist	Yes	+
I think I really need it	Need	+
I think that is so	Yes	+
Could try	Yes	+
Help would be in time now	Need	+
1.2. Implicitly positive responses		
Looks like that	Looks like that	-
Most likely because I feel sick	Yes	+
It seems so, I feel very bad	Yes	+
Most likely	Need	+
I can try	I can try	-
Most likely yes	Yes	+
It is entirely possible	Yes	+
2. Negative responses		
2.1. Explicitly negative responses		
Unnecessary	Do not need	+
I don’t need it	Do not need	+
I don’t want it	Do not need	+
I appreciate it, but I don't want it	No	+
I definitely don't want it	Do not need	+
I don't have time for that	Do not need	+
I am not one of those people who need help	Do not need	+
2.2. Implicitly negative responses		
I think I am fine	Do not need	+
I think I’m OK	Do not need	+
I'm afraid I won't be able to see a doctor	I'm afraid I won't be able to see a doctor	-
I don’t think so	No	+
I believe that I can do it myself	Do not need	+
I don't trust psychologists	No	+
I'm not sure it's necessary	I'm not sure it's necessary	-

The corresponding results are compiled according to the temperature parameter and are presented in Table 2.

Table 2
ChatGPT-4 processing results with Temperature>0 parameter values of unidentified atypical users' responses.

Atypical user's responses	Temperature=0,3. Recognition result from Chat GPT-4.	Temperature=0,5. Recognition result from Chat GPT-4.	Temperature=0,8. Recognition result from Chat GPT-4.	Temperature=1. Recognition result from Chat GPT-4.
1. Positive responses				
1.2. Implicitly positive responses				
Looks like that	Looks like that	Yes	Need	Yes
I can try	I can try	Ok	Yes	Yes
2. Negative responses				
2.2. Implicitly negative responses				
I'm afraid I won't be able to see a doctor	I'm afraid I won't be able to see a doctor	I'm afraid I won't be able to see a doctor	I'm afraid I won't be able to see a doctor	I'm afraid I won't be able to see a doctor
I'm not sure it's necessary	I'm not sure it's necessary	Do not need	Do not need	Ok

Figure 2 presents a comparison of ChatGPT-4 recognition performance of user responses in the "Positive responses" category at different chat temperature modes. Similar information for the "Negative responses" category is presented in Figure 3.

It can be seen from the graphs that the influence of temperature settings on the performance of processing user' responses by ChatGPT-4 takes place recognizing clearly positive responses. When the temperature increases, there is an unambiguously positive dynamic of its work efficiency, since already at a temperature parameter value of 0.5 all such responses were recognized. On the other hand, processing implicitly negative responses by ChatGPT-4, changing the temperature settings shows less dynamics of impact on its work performance. In this case, as the temperature increased, the number of recognized responses increased, but unrecognized responses also remained (Figure 3).

Psycholinguistic studies testify that the processing of negative statements is more difficult than positive statements [22, 23], regardless of the fact whether the negation is explicit (for example, not) or implicit (for example, forget) [24, 25]. The authors of the study explain the results by referring to the cognitive processing of negative statements, namely the increased cognitive load and difficulty of processing the negation. It is noted that a statement containing an implicit negation can cause an additional cognitive load which is necessary to understand its meaning. Language-based artificial intelligence tools such as OpenAI's ChatGPT have limitations in performing complex reasoning tasks [26, 27]. Although these models can interpret most queries and contexts, they occasionally face limitations of understanding while dealing with ambiguous or contextually complex queries, which we believe includes processing of implicit negative responses from users.

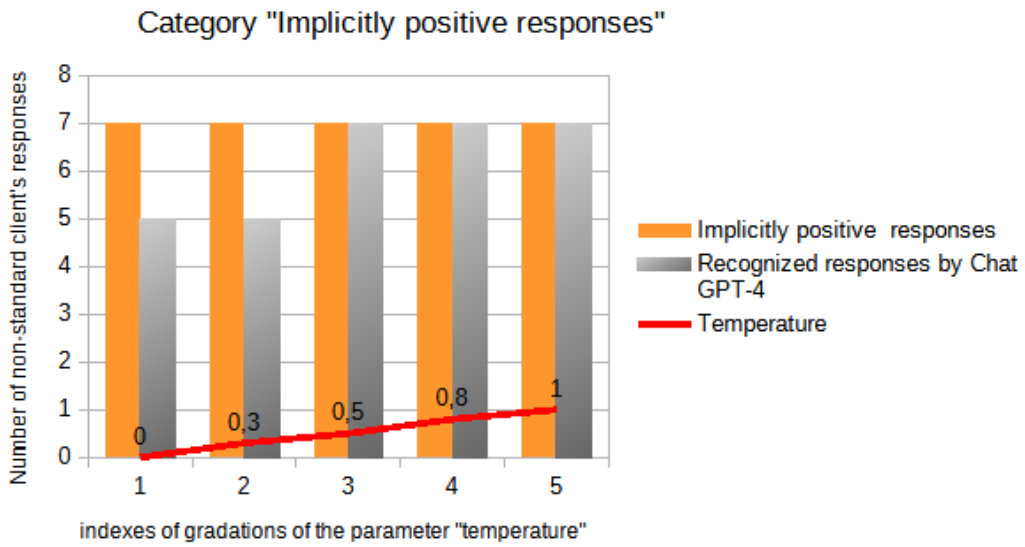

Figure 2: Productivity of recognition users' implicitly positive responses under different temperature conditions.

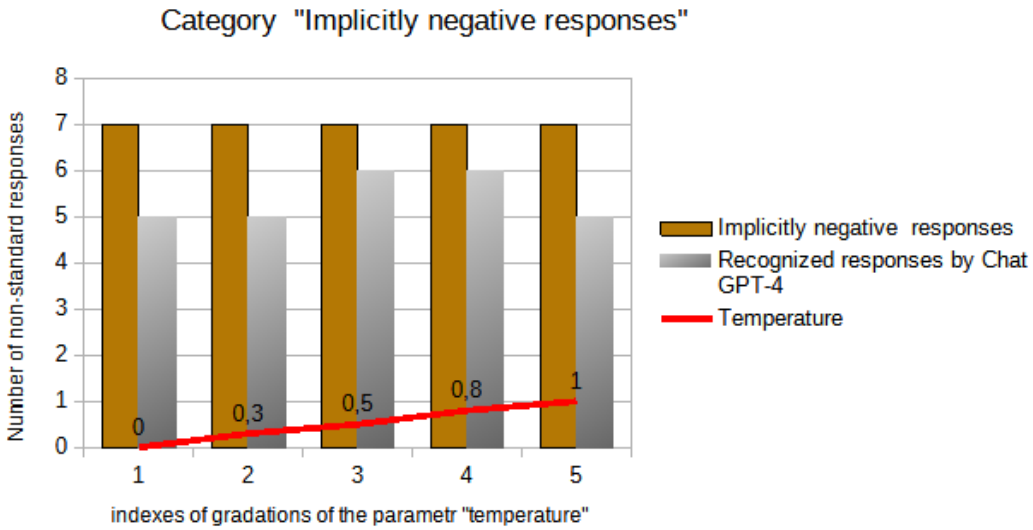

Figure 3: Productivity of recognition users' implicitly negative responses under different temperature conditions

Thus, without using the GPT-4 language model in the conducted experiment, 23.9% of responses should be processed by the operator, while their additional processing using GPT-4 with the appropriate setting of the temperature parameter allows to redirect to the operator only 5.07% of the respondents' responses.

It should be pointed out that the results of testing the effectiveness of ChatGPT-4 processing of similar user responses "originally" provided in the Ukrainian language are fully consistent with the results given in Tables 1-2.

Issues of secure data processing

Discussing the use of chatbots with artificial intelligence in various spheres of public life, we cannot ignore the significant risks associated with the processing of personal data, in particular those that may constitute a violation of human rights and freedoms. Thus, due to alleged privacy violations related to the use of ChatGPT, the Italian Data Protection Supervisory Authority issued a decision

on March 31, 2023. To restrict the use of the specified chatbot in public administration and business [28]. Such actions of the Italian national body for the protection of personal rights are related to the fact that ChatGPT had "suffered a data breach on March 20 concerning users' conversations and payment information of subscribers to the paid service".

In order to ensure the proper regulation of social relations related to the use of artificial intelligence, to prevent the use of these technologies for groundless interference in private and family life, to systematize risks in the specified area of legal regulation, in June 2023, the European Parliament and the Council of the European Union adopted "Act on Artificial Intelligence" [29]. The specified act of the European Union proposes to establish obligations for the owners of information systems based on generative artificial intelligence technologies to prevent illegal copying of author's content, illegal collection and subsequent profiling of information about individuals.

Among the main cyber problems associated with chatbots such as ChatGPT, work [30] notes the reduction of barriers for cybercriminals, which include social engineering attacks (inclining victims to disclose confidential information), phishing attacks (sending malicious links or messages), identity theft, data leakage, etc.

OpenAI company constantly takes measures to keep the artificial intelligence ecosystem safe, which include API moderation to prevent and block dangerous content. The developers of the ChatGPT Enterprise version state that customer tips and company data are not used to train OpenAI models, AES 256 is used for data encryption at rest, and TLS 1.2+ protocol is used during transmission. ChatGPT Enterprise is also compatible with the SOC 2 standard [31].

Therefore, developers of chatbots using ChatGPT need to provide security measures and access controls to prevent unauthorized access to the system. Regarding the fulfillment of the requirements to prevent illegal copying of author's content, illegal collection and subsequent profiling of information about individuals, the authors suggest using ChatGPT only for the narrow task of interpreting ambiguous customer responses. At the same time, the protection of personal data of users is carried out using the necessary technologies to ensure the security of personal data. This is encryption to protect sensitive data, in particular, hashing is applied to user IDs (ID) and passwords stored in the database, making them impossible to compromise in the event of a data breach. Access to the ChatGPT API occurs through a secure HTTPS connection, which provides the necessary level of data security during their transmission

Conclusions and discussion

In the article the authors have analyzed the possibilities and performance of the language model GPT-4 for the tasks of the automating the processing of responses of social and psychological services users through digital means of communication. The conclusions of the conducted research show that the results of processing non-standard users' responses by the chatbot are quite high and correlate well with expert assessments regarding the content and correctness of recognition. It was found that the impact of GPT-4 temperature settings on the processing performance of implicitly positive users' responses has positive dynamics with increasing temperature values. It was also established that for implicitly negative users' responses, the impact of GPT-4 temperature settings on processing performance remains uncertain and needs further research in collaboration with linguists and psychologists. It should be noted that according to the research results, for the most accurate recognition of non-standard users' responses, it is recommended to use the temperature parameter value in the range of (0.5-0.8). But we should also not ignore the results in which there is a qualitative jump from one plane (for example, negative) to another (for example, positive).

The proposed approach to processing the results of users' responses allows developers to integrate GPT-4 into their applications conveniently, using the capabilities of this model effectively. Automation of the process of processing users' responses allows to relieve the operators of social service centers, which is relevant in the conditions of personnel shortage in the Ukrainian labor market. However, it is important to note that these tools do not completely replace the involvement

of qualified professionals, instead, they serve as additional tools in the field of providing social and psychological services to users and contribute to the overall digitalization of the process.

Let's note that the solution to the research task formulated above could be accomplished by creating a neural network and its further training [32], which requires significant costs, both financial (paying the work of programmers) and time (it is necessary to spend a certain amount of time on training the network). Using alredy trained GPT-4 language model allows to reduce or to eliminate some costs significantly.

The obtained research results can be used by the developers of the GPT-4 model in order to optimize and improve the quality of text recognition during weighted learning to adjust the weight of errors in rare or important cases [33], as well as to improve the techniques used to analyze the emotional state, mood and intonation in the text.

Acknowledgements

The work was carried out on an initiative basis. We express our gratitude to the Social Support Center (Kropyvnytskyi, Ukraine) for the opportunity to conduct the research and providing data for processing and analysis.

We wish to thank the anonymous reviewers for their comments as they improved the quality of this paper.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] O. Prysiazhniuk, O. Blyzniukova, Application of Fuzzy Approach in Modeling of Psychodiagnostic Decision Support Systems for One Class of Tasks, in: Proceedings of the 2th. symposium on Intelligent Solutions, IntSol '2021, Kyiv - Uzhhorod, Ukraine, 2021, pp. 11–20. URL: http://ceur-ws.org/Vol-3106/Paper_2.pdf.

[2] V. Goar, N. Yadav, P. Yadav, Conversational AI for Natural Language Processing: An Review of ChatGPT, International Journal on Recent and Innovation Trends in Computing and Communication, 11 (2023), 109–117. doi: 10.17762/ijritcc.v11i3s.6161.

[3] F.C. Kitamura, ChatGPT is Shaping the Future of Medical Writing but Still Requires Human Judgment, Radiology 307 (2023). doi: 10.1148/radiol.230171.

[4] Z. Du, Y. Qian, X. Liu, M. Ding, J. Qiu, Z. Yang, and J. Tang, General Language Model Pretraining with Autoregressive Blank Infilling, in: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 320–335. doi: 10.18653/v1/2022.acl-long.26.

[5] J. Gravel, M. D'Amours-Gravel, E. Osmanlliu, Learning to Fake It: Limited Responses and Fabricated References Provided by ChatGPT for Medical Questions, Mayo Clinic proceedings: Digital Health 1 (2023), 226-234. doi: 10.1016/j.mcpdig.2023.05.004.

[6] K.I. Roumeliotis, N.D. Tselikas, ChatGPT and Open-AI Models: A Preliminary Review, Future Internet 192 (2023). doi: 10.3390/fi15060192.

[7] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, Language Models Are Unsupervised Multitask Learners, 2019. URL: https://hayate-lab.com/wp-content/uploads/2023/05/61b1321d512410607235e9a7457a715c.pdf.

[8] T. Susnjak, CHATGPT: The end of online exam integrity?, J. Education Sciences 656 (2024). doi: 10.3390/educsci14060656.

[9] M. Chen, J. Tworek, H. Jun, Q. Yuan, H.P. de Oliveira Pinto, J. Kaplan, H. Edwards, Y. Burda, N. Joseph, G. Brockman, et al., Evaluating large language models trained on code, 2021. doi: 10.48550/arXiv.2107.03374.

[10] Introducing ChatGPT, 2022. URL: https://openai.com/index/chatgpt/.
[11] J. Ziwei, Y. Tiezheng, X. Yan, L. Nayeon, I. Etsuko, F. Pascale, Towards Mitigating Hallucination in Large Language Models via Self-Reflection, in: Findings of the Association for Computational Linguistics: EMNLP, Singapore, 2023, pp.1827-1843. doi: 10.18653/v1/2023.findings-emnlp.123.
[12] L. Gao, J. Schulman, J. Hilton, Scaling Laws for Reward Model Over optimization, in: Proceedings of the 40th International Conference on Machine Learning, Honolulu Hawaii USA, 2023, pp. 10835-10866. URL: https://dl.acm.org/doi/10.5555/3618408.3618845.
[13] OpenAI, OpenAI Charter, 2023. URL: https://openai.com/charter/.
[14] OpenAI, GPT-4, 2023. URL: https://openai.com/index/gpt-4-research.
[15] Y. Zhang, S. Sun, M. Galley, Y.-C. Chen, C. Brockett, X. Gao, J. Gao, J. Liu, B. Dolan, DIALOGPT: Large-Scale Generative Pre-training for Conversational Response generation, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 2020, pp.270-278. doi: 10.18653/v1/2020.acl-demos.30.
[16] A. Mukherjee, ChatGPT: a Fuzzy System that talks Like a Human, J. of Mathematical Sciences & Computational Mathematics 251 (2024). doi: 10.15864/jmscm.5303.
[17] M. Javaid, A. Haleem, R. P. Singh, A study on ChatGPT for Industry 4.0: Background, potentials, challenges, and eventualities, Journal of Economy and Technology 1 (2023) 127-143. doi: 10.1016/j.ject.2023.08.001.
[18] Setting Parameters in OpenAI, URL: https://www.codecademy.com/article/setting-parameters-in-open-ai.
[19] GPT-4 Technical Report, 2023. URL: https://arxiv.org/pdf/2303.08774.
[20] N. Gouws-Stewart, The ultimate guide to prompt engineering your GPT-3.5-Turbo model, 2024. URL: https://masterofcode.com/blog/the-ultimate-guide-to-gpt-prompt-engineering.
[21] Six strategies for getting better results, 2024. URL: https://platform.openai.com/docs/guides/prompt-engineering/six-strategies-for-getting-better-results.
[22] H. Hnatiienko, V. Snytyuk, N. Tmienova, O. Voloshyn, Application of expert decision-making technologies for fair evaluation in testing problems, in: Proceedings of the 20th. annual workshop on Information Technologies and Security, ITS '20, Kyiv, Ukraine, 2020, pp. 46–60. URL: https://ceur-ws.org/Vol-2859/paper5.pdf.
[23] H. Hnatiienko, Choice Manipulation in Multicriteria Optimization Problems, in: Proceedings of the 19th. annual workshop on Information Technologies and Security, ITS '20, Kyiv, Ukraine, 2019, pp. 234–245. URL: https://ceur-ws.org/Vol-2577/paper19.pdf.
[24] J. Maciuszek, M. Polak, M. Sekulak, There is no item vs. I wish there were an item: Implicit negation causes false recall just as well as explicit negation, PLoS ONE 14 (2019). doi: 10.1371/journal.pone.0215283.
[25] M. Xiang, J. Grove, A. Giannakidou, Semantic and pragmatic processes in the comprehension of negation: An event related potential study of negative polarity sensitivity, Journal of Neurolinguistics 38 (2016) 71–88. doi: 10.1016/j.jneuroling.2015.11.001.
[26] Y. Liu, T. Han, S. Ma, J. Zhang, Y. Yang, J. Tian, H. He, A. Li, M. He, Z. Liu, Z. Wu, L. Zhao, D. Zhu, X. Li, N. Qiang, D. Shen, T. Liu, B. Ge, Summary of ChatGPT-Related research and perspective towards the future of large language models, J. Meta-Radiology 1 (2023). doi: 10.1016/j.metrad.2023.100017.
[27] S. Finch, J. Choi, ConvoSense: Overcoming Monotonous Commonsense Inferences for Conversational AI, J. Transactions of the Association for Computational Linguistics, 12 (2024), 484-506. doi: 10.1162/tacl_a_00659.
[28] OpenAI's ChatGPT Chatbot Blocked in Italy Over Privacy Concerns, 2023. URL: https://www.euronews.com/next/2023/03/31/openais-chatgpt-chatbot-banned-in-italy-by-watchdog-over-privacy-concerns.
[29] Artificial intelligence regulation in the EU, 2023. URL: https://multimedia.europarl.europa.eu/en/audio/-ai_EPBL2107202301_EN.

[30] G. Sebastian, Do ChatGPT and Other AI Chatbots Pose a Cybersecurity Risk?: An Exploratory Study, International Journal of Security and Privacy in Pervasive Computing 15 (2023). doi: 10.4018/IJSPPC.320225.
[31] OpenAI, Introducing ChatGPT Enterprise, 2023. URL: https://openai.com/index/introducing-chatgpt-enterprise/.
[32] Z. Wu, Q. She, C. Zhou, Intelligent Customer Service System Optimization Based on Artificial Intelligence, Journal of Organizational and End User Computing, 36 (2024). doi: 10.4018/JOEUC.336923.
[33] A. Voloshin, G. Gnatienko, E. Drobot, A Method of Indirect Determination of Intervals of Weight Coefficients of Parameters for Metricized Relations Between Objects, Journal of Automation and Information Sciences, 35 (2003), 25-30. doi: 10.1615/JAutomatInfScien.v35.i3.30.