prevenIA: a Chatbot for Information and Prevention of Suicide and other Mental Health Disorders

Pablo Ascorbe[1], María Soledad Campos[2], César Domínguez[1], Jónathan Heras[1] and Ana Rosa Terroba-Reinares[1,3]

[1]*Departamento de Matemáticas y Computación, Universidad de La Rioja*
[2]*Unidad de Salud Mental Espartero, Logroño, La Rioja*
[3]*Fundación Rioja Salud*

Abstract

Suicide and mental disorders are major health and social issues worldwide; therefore, a simple access to reliable sources of information that can mitigate these problems is instrumental for people suffering them and for their relatives. The goal of this project is the development of a Whatsapp chatbot that will provide resources and advices to help in suicide prevention. In addition, the chatbot will detect sensitive users and derive them to a team of specialists. Initially, the focus of the projet is on suicide prevention and in the Whatsapp platform, but it will be extended to other mental disorders and other platforms to reach as many people as possible. In conclusion, this project aims to disseminate reliable information about suicide and mental disorders in a format that is understandable and easily accessible by the general population.

Keywords

Chatbot, Semantic Search, Question Answering, Suicide, Mental disorders

1. Introduction

According to data published on 19 December 2022 by the National Statistics Institute (INE), 4,003 people died by suicide in Spain in 2021, an average of 11 people per day. In addition, suicide among children under 15 years old has doubled [1]. For each completed suicide, it is estimated that there are around 20 suicide attempts; for each attempt, around 14 more people have thought about committing suicide; and for each suicide, at least 6 people who survive the deceased have been directly affected by this death [2]. Therefore, due to these data, the World Health Organisation (WHO) identifies suicide and attempted suicide as one of the most serious health problems that can affect people, which is why the WHO recommends to all its members that it should be addressed as a priority [2].

On 12 March 2014, the Health and Social Services Commission of the lower house in the Spanish Parliament approved, unanimously by all the groups, a non-legislative proposal regarding the development of a National Suicide Prevention Plan by the Spanish health, educational and social institutions in accordance with the directives of the European Union and international organisations. Since then, several suicide prevention plans have been developed in some Autonomous Regions (see, for example, those of La Rioja [3], the Canary Islands [4], and Navarre [5]).

Prevention plans propose different interventions targeting different audiences (general population, health professionals, or media, among others) [6]. Interventions aimed at the general public include raising awareness among the general public, creating contacts for help, or setting up training programmes.

It is precisely in this context that the present project is framed. Specifically, the project aims to create a Whatsapp chatbot that provides reliable information on suicide prevention, in principle, in the autonomous region of La Rioja. The chatbot will also be able to detect risk situations and refer them to 112 or to a support team with specialised professionals in the Health Service of La Rioja (SERIS). In addition, the chatbot will have to communicate in a clear and adapted way to the user using it.

Initially, suicide prevention will be the focus of this project; but, in addition, there are also other mental disorders (eating disorders, for example) that have only worsened with the SARS-CoV-2 coronavirus pandemic [7]. Therefore, the development of this chatbot can help people suffering from these conditions by providing information directly to the affected person and their families.

The rest of this article is organised as follows. We begin with a section that reviews the chatbot-style tools that have been described in the literature in the context of mental health and discusses the advantages and disad-

vantages of these tools. Then, we explain the phases into which the development of the chatbot will be divided, as well as the team that will carry them out. Finally, we draw some conclusions about the hypothetical results that are intended to be obtained with this work.

2. Conceptual framework

A chatbot, or conversational assistant, is a software application that simulates a conversation with a person by providing automatic responses, and from whose application it is possible to obtain some information or some kind of action [8]. Chatbots are currently being used in a wide range of fields, including health in general [9] and mental health in particular [10]. In fact, the use of chatbots in mental health is present in the very origins of these tools in the 1960s, a period in which what is considered to be the first chatbot, called ELIZA, was developed. This chatbot made it possible to simulate a conversation with a psychologist in a psychotherapy session [8].

There are several recent literature reviews on the use of chatbots in mental health [9, 10, 11, 12] and also on the use of artificial intelligence methods in aspects related to suicide [13]. These reviews highlight aspects where chatbots can be useful in this area. On the one hand, a chatbot can give access to virtual services to certain people who would avoid using a face-to-face service, either because the latter is overburdened, because they cannot afford it, or to avoid the stigma attached to certain people with mental health problems. On the other hand, the anonymity offered by chatbots allows some people, especially the younger ones, to seek information about their doubts or freely express their feelings and problems; feelings that they are not comfortable to be shared to other human beings [10, 13, 14]. Furthermore, both people who use these chatbots [12] and mental health professionals [15] have a positive perception and opinion of them. However, although it is emphasised that these systems can help the professional in some aspects, they are never intended to replace them [15].

In a literature review carried out in 2022 by Valizadeh and Parde [9] on the application of chatbots in health, 70 studies were identified; and 22 of them correspond to different pathologies related to mental health. Among these pathologies are depression, anxiety, phobias or addictions; however, none of these studies were related to suicide. After that review was published, the designed of a chatbot for the detection of suicidal ideation has been proposed [14]. This detection is done through a natural language processing model, called BERT, retrained on a database obtained from a Reddit subnetwork, called Reddit Suicide Watch [16]. If ideation is detected, the user is asked for permission to send help; if permission is not given, the chatbot will continue chatting with the user and will express concern for their well-being.

It is important to understand that chatbots also have their limitations. Human communication has nuances that are complicated for artificial intelligence to understand, and inappropriate or unclear responses must be avoided [10, 15]. In addition, the non-verbal language that is often decisive in this context is lacking in chatbots [8]. A special care aspect is to provide an adequate reaction once an emergency situation has been detected [12]. All these limitations must be treated with particular care in a subject as sensitive as suicide risk [15].

A very noteworthy aspect of the literature reviews on the use of chatbots in mental health is that most of the studies have been conducted in English-speaking populations, and there is a notable absence of works for Spanish-speakers [9, 11, 12, 13]. An exception is the work by Romero et al. [8] wherein the basis for the design of a chatbot with psychological assessment functions is presented. Research into the personalisation of chatbots in order to provide answers to different types of users is also highlighted as an interesting and little-studied aspect. In particular, the complexity of the language could be adapted to the level required by the user [12]. Finally, the uses of machine learning methods that stand out include the classification and detection of people potentially at risk of suicidal behaviour, but there is no evidence of studies that involve providing information, for example to family members, about suicidal behaviour [13].

3. Work plan

The development of the project is divided into the following phases.

3.1. Phase 1. Informative chatbot

In this first phase, a chatbot will be developed on the WhatsApp platform. Through natural language processing techniques, the chatbot will be able to automatically provide information related to suicide, for example, to prevent it or to deal with mourning situations. To achieve this aim, it will be necessary to build a corpus of information sources using the materials available on websites such as MenteScopia[1], prevensuic[2], or the prevention plans created by the Autonomous Regions.

Once the corpus has been built, models will be combined to perform semantic searches and to answer questions [17]. The former will be used to extract the context that will be used by the latter to answer users' questions. One of the biggest challenges in this context is that both types of models are mainly used in the context of the English language, so it will be necessary to use either

---

[1]https://psynal.eu/mentescopia/
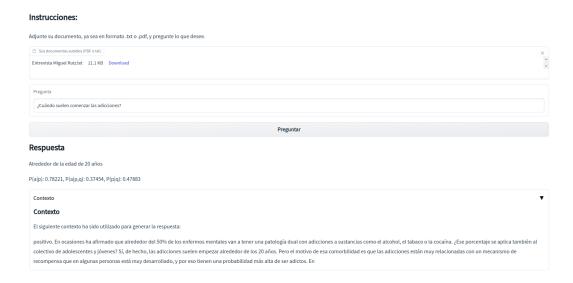[2]https://www.prevensuic.org/

Figure 1: Demo for obtaining answers from mental health documents.

translation models or to train new models. A small application developed with Gradio is already available (see Figure 1) which allows answers to be obtained from a series of documents identified in an initial search.

Finally, in this phase, the accessibility of the chatbot will be taken into account. Thus, easy read and plain language techniques will be used to allow users to find the information they need, understand what they find and use that information. The work done in the context of CLARA-MeD [18] for medical text simplification will be particularly useful for this. Moreover, the chatbot will include the necessary functionality to communicate with users in written or oral form. In the latter, it will be necessary to use models such as Whisper [19], which will allow the chatbot to transcribe the audios sent by the user; or fairseq S^2, to synthesise voice from text [20].

It is important to emphasise that at the beginning of the conversation the users will be informed that they are interacting with a machine, and will always have an option to choose to speak to a person via a specialised service such as 112.

3.2. Phase 2. Preventive chatbot

In this phase, the chatbot developed in Phase 1 will be extended to automatically detect possible risks. If a sensitive situation is detected, the user will be redirected to 112 or to a support team with specialised professionals in SERIS.

For the moment, it is not intended to build an open text chatbot that can replace psychologists, psychiatrists or specialised staff working in community help services such as the "*Teléfono de la Esperanza*". Instead, it is intended to locate risk cases by means of surveys based on clinical interviews developed by psychologists and psychiatrists [21]. In addition, the chatbot will make it easier for users to create help resources such as "My safety plan" [3].

3.3. Phase 3. Extension to other platforms

The chatbot developed in the previous two phases will be implemented on other platforms in addition to WhatsApp in order to reach the widest possible audience. These platforms include Facebook chat, messaging apps such as Telegram, or platforms such as Discord.

3.4. Phase 4. Extension to other disorders

With all the experience from the previous phases, the chatbot will be enhanced to not only provide information on suicide, but also on other mental health disorders such as eating and behavioural disorders.

4. The working team

The project described in this paper is carried out by a multidisciplinary team combining experts in computer science, psychiatry and linguistics. The technical development of the project is carried out within the Computer Science Group of the University of La Rioja. In this sense,

different members of the group collaborate in different tasks of the work plan.

María Soledad Campos, a doctor in psychiatry from the Espartero Mental Health Unit in La Rioja, is also collaborating with the project and will contribute her expert vision for the compilation and review of materials related to suicide as well as clinical interviews present in the literature. She will also be in charge of validating the correctness of the answers provided by the chatbot.

Finally, we are working with Ana Rosa Terroba Reinares, PhD in Linguistics and specialist in health communication at SERIS. Her contribution will be essential for the chatbot to communicate clearly with users, in addition to contributing her knowledge in the generation of the corpus that feeds the chatbot.

5. Conclusions

Mental health is a serious problem in today's society, and the use of natural language processing and artificial intelligence tools, such as the one we intend to build with the chatbot of this project, can support patients and families. However, the aim is not to replace the human role played by agents such as psychiatric teams or organisations such as the Telefono de la Esperanza, but to complement it with user-friendly tools which provide reliable information.

Acknowledgments

References

[1] Instituto Nacional de Estadística, Defunciones según la causa de muerte Año 2021, Technical Report, 2022. URL: https://www.ine.es/prensa/edcm_2021.pdf.

[2] WHO, Suicide worldwide in 2019: global health estimates, 2021.

[3] Rioja Salud, Plan de prevención del suicidio en La Rioja, 2019. URL: https://www.riojasalud.es/files/content/ciudadanos/planes-estrategicos/PLAN_PREVENCION_CONDUCTA_SUICIDA_DEF.pdf.

[4] Servicio Canario de Salud, Programa de prevención de la conducta suicida en Canarias, 2021. URL: https://www3.gobiernodecanarias.org/sanidad/scs/content/3f5ce57d-1085-11ec-bfb0-874800d2c074/PPCSC.pdf.

[5] Gobierno de Navarra, Prevención y actuación ante conductas suicidas, 2014. URL: https://www.educacion.navarra.es/documents/27590/548485/Suicidio.pdf/b5374981-511a-40ed-82c5-7c74bc23b049.

[6] T. Sufrate-Sorzano, E. Jiménez-Ramón, M. E. Garrote-Cámara, V. Gea-Caballero, A. Durante, R. Júarez-Vela, I. Santolalla-Arnedo, Health plans for suicide prevention in Spain: a descriptive analysis of the published documents, Nursing Reports 12 (2022) 77–89.

[7] D. Crespo Amaro, P. Vázquez Herrera, G. Velázquez Basterra, M. S. Campos Burgui, Salud mental y covid-19, Zubía (2021) 131–140.

[8] M. Romero, C. Casadevante, H. Montoro, Cómo construir un psicólogo-chatbot, Papeles del Psicólogo 41 (2020) 27–34.

[9] M. Valizadeh, N. Parde, The AI doctor is in: A survey of task-oriented dialogue systems for healthcare applications, in: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2022, pp. 6638–6660.

[10] A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. S. Kashavan, J. B. Torous, Chatbots and conversational agents in mental health: a review of the psychiatric landscape, The Canadian Journal of Psychiatry 64 (2019) 456–464.

[11] T. Zhang, A. M. Schoene, S. Ji, S. Ananiadou, Natural language processing applied to mental illness detection: a narrative review, NPJ digital medicine 5 (2022) 46.

[12] A. A. Abd-Alrazaq, M. Alajlani, N. Ali, K. Denecke, B. M. Bewick, M. Househ, Perceptions and opinions of patients about mental health chatbots: scoping review, Journal of medical Internet research 23 (2021) e17828.

[13] S. Ji, S. Pan, X. Li, E. Cambria, G. Long, Z. Huang, Suicidal ideation detection: A review of machine learning methods and applications, IEEE Transactions on Computational Social Systems 8 (2020) 214–226.

[14] J. X. Chan, S.-L. Chua, L. K. Foo, A two-stage classification chatbot for suicidal ideation detection, in: International Conference on Computer, Information Technology and Intelligent Computing (CITIC 2022), Atlantis Press, 2022, pp. 405–412.

[15] C. Sweeney, C. Potts, E. Ennis, R. Bond, M. D. Mulvenna, S. O'neill, M. Malcolm, L. Kuosmanen, C. Kostenius, A. Vakaloudis, et al., Can chatbots help support a person's mental health? Perceptions and views from mental healthcare professionals and experts, ACM Transactions on Computing for Healthcare 2 (2021) 1–15.

[16] S. Ji, C. P. Yu, S.-f. Fung, S. Pan, G. Long, Supervised

learning for suicidal ideation detection in online user content, Complexity 2018 (2018).

[17] A. Esteva, A. Kale, R. Paulus, K. Hashimoto, W. Yin, D. Radev, R. Socher, Covid-19 information retrieval with deep-learning based semantic search, question answering, and abstractive summarization, NPJ digital medicine 4 (2021) 68.

[18] L. Campillos Llanos, A. R. Terroba Reinares, S. Zakhir Puig, A. Valverde, A. Capllonch-Carrión, Building a comparable corpus and a benchmark for Spanish medical text simplification (2022).

[19] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, I. Sutskever, Robust speech recognition via large-scale weak supervision, arXiv preprint arXiv:2212.04356 (2022).

[20] C. Wang, W.-N. Hsu, Y. Adi, A. Polyak, A. Lee, P.-J. Chen, J. Gu, J. Pino, fairseq Sˆ 2: A scalable and integrable speech synthesis toolkit, arXiv preprint arXiv:2109.06912 (2021).

[21] D. V. Sheehan, Y. Lecrubier, K. H. Sheehan, P. Amorim, J. Janavs, E. Weiller, T. Hergueta, R. Baker, G. C. Dunbar, et al., The mini-international neuropsychiatric interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10, Journal of clinical psychiatry 59 (1998) 22–33.