Design of a Chatbot to Assist the Elderly

Stefano Valtolina(✉) and Mattia Marchionna

Department of Computer Science, Università degli Studi di Milano, Milan, Italy
stefano.valtolina@unimi.it, mattia.marchionna@studenti.unimi.it

Abstract. Nowadays, conversational agents are solutions that can provide a highly valuable addition to the existing healthcare services to assist the elderly in following their care plans and gradually changing adverse patterns of behaviour. Nevertheless, the development of a conversation agent in the healthcare domain presents several technical, design and linguistic challenges. In our paper, we describe a chatbot conversing with elderly persons, with age-related problems. Charlie, the name of our chatbot, has been designed to provide the elderly with companionship through innovative strategies based on gamification, active notifications, and promotion of self-compassion that can be explored for preventive mental healthcare. Moreover, Charlie can be used to monitor meaningful or anomalous situations that can affect aged people. To specify these situations in the paper we describe a web application that enables medical assistants and relatives to create rules that depend on data gathered by Charlie such as the number of required news, followed tips, completed games, but also biometrical data such as the number of steps made during the day, burned calories or the number of hours of sleep. To help caregivers in creating these rules, we are studying how to endow our web app with a recommendation service for predicting recurring anomalous situations and, accordingly, provide suggestions on which characteristics the elderly will have to change to improve their lifestyle.

Keywords: End-user development · Virtual assistants · CUI – conversational interface in healthcare · AI for HCI

1 Introduction

Since the COVID-19 pandemic outbreak, healthcare systems have been subjected to enormous stress and people all across the sector are looking for digital solutions to help ease the pressure. Specifically, due to this pandemic, the elderly are increasingly alone and many are beginning to establish empathic relationships with human and virtual telephone operators. Loneliness, especially in old age, begins to become a serious problem. A research study commissioned by Ipsos, Comieco and Symbola [1] claims as 60% of those who suffer from it today are in a particularly fragile age range. The result is confirmed by another study [2] carried out by the Enuan, a startup that deals with artificial intelligence applied to voice and chat channels. This study states how from March 2020 onwards, especially during the months of lockdown and forced smart working, the most empathic interactions have increased by over 30%. For supporting health facilities

D. Fogli et al. (Eds.): IS-EUD 2021, LNCS 12724, pp. 153–168, 2021.
https://doi.org/10.1007/978-3-030-79840-6_10

that are overburdened by the task of assisting a lot of patients in complying with their care plans, and for helping the elderly to deal with this lockdown and isolation that are disrupting their social life, this paper aims at studying a personal chatbot for elderly able to act as a medical consultant, friend, assistant and entertainer.

In detail, the first research question that this paper wants to answer concerns the value of involving a virtual assistant in helping older people not to feel alone, to be engaged in entertaining activities, to manage their medicines and to establish remote connections with their relatives. To this aim, in Sect. 2 we present several works that motivated us to use a conversation agent in our context of the study. Nevertheless, we identified how some works propose solutions that miss in addressing pre-emptive care for strengthening mental health or improving the quality of life, without necessarily assuming diagnosed disorders, as happening for the elderly.

For this reason, Sect. 3 describes our study that starts by investigating the problems relating to the elderly, their needs and the existing technological solutions and then proposes a conversational user interface (CUI) that is software that runs simple and structurally repetitive tasks inside a messaging application. Our idea is to design a conversation to share knowledge and emotion that can be affordable for non-technical people. The challenge is to understand which communication characteristics and interaction strategies encourage their use, specifically for the elderly. Specifically, we want to investigate the level of acceptability by senior users of a combination of assistance, entertainment, gamification and self-compassion strategies to evaluate the better solutions that can support the elderly in their daily routines.

A second research question we try to answer in Sect. 4 is how to train our chatbot to monitor relevant and critical situations that can affect the lives of our dear seniors. After observing the elderly's daily activities, caregivers, that is health assistants or family members need to identify critical behaviours that need to be avoided or regulated. For doing this, we designed a web application that allows us to define some rules that are automatically triggered when specific situations take place. This web app aims at providing non-professional users (health assistants or relatives) with an environment for creating rules using an easy and visual language, familiar to them. Those rules are used for monitoring events related to the elderly's habits and are automatically triggered to notify behaviour that needs to be corrected. To provide help to the health assistants and relatives in their rules creation task, we provided the web app with an Artificial Intelligence (AI) module to make predictions and to compute suggestions helpful to monitor the elderly's behaviours. The main feature of this AI module is its capability of describing the reasons for the computed predictions, by computing the smallest change to the feature values that increases the prediction. Such explanations are important since they not only allow us to understand our dear seniors' conditions but also shall ensure that caregivers can trust the output of the automatic learner, which would be otherwise viewed as a "black-box". Finally, Sect. 5 reports conclusions and future works.

2 Elderly Assistance with a Virtual Agent

Chatbots [3] are commonplace in the online retail space, but they are also emerging in the healthcare sector [4]. In this area, well-designed chatbots can help to communicate more efficiently with a non-expert target audience [5, 6]. Several studies [3, 7–10] and commercial systems [11–13] highlight how virtual agents can help patients to deal with common issues, such as remembering to take medicines, doing some exercises, following a specific diet. Other works [14, 15] focused on how hospitalized medical patients would respond to a conversational agent that would provide empathic support. Similarly, King et al. in [16] describe an agent that aims at delivering personalized physical activity advice to older adults which the goal of increasing their brisk walking. Finally, a last important study presented in [17] reviews existing substantial scientific works concerning medical chatbots from a behaviour change perspective. Specifically, the authors identify problems of acceptance of medical chatbots in society and the use of chatbots to change harmful behaviour. They mainly suggest how, to design an effective medical chatbot, is useful to insert chatbots in messaging apps, establish an understandable agreement process between the user and the chatbot and implement emotionality of the chatbot's responses. Likewise, [6] investigates which functionalities text-based chatbots can provide to benefit human interactions and what are the challenges and strategies associated with them not only in the medical domain.

In general, these works focus is on what should be "fixed", e.g. they suggest moving more, changing diet habits, and how to mitigate depressive symptoms, and suicidal tendencies and the target is the person with these problems [18]. They miss in addressing preemptive care for strengthening mental health or improving the quality of life, without necessarily assuming diagnosed disorders, as happening for the elderly.

The main research question we try to answer in this paper aims at studying how to design a chatbot able to adapt its functionality, interface, personality, information access, and content to increase its communication skills to an individual or a category of individuals. Moreover, we want to investigate the efficacy of a combination of assistance, entertainment, gamification and self-compassion strategies to evaluate the better solutions that can support the elderly in their daily routines. In summary, we aim to understand if it is possible to use chatbots not only to fill a void for lonely pensioners but to help them for increasing the perception of their quality of life as well.

3 Design Strategies to Develop a Chatbot for Assisting the Elderly

Understanding users' background and collecting as many features about them as possible helps narrow the bot focus and provide personalized services. For this reason, we focused our study on the design of a virtual agent, named Charlie, able to act as a caregiver but also as a friend of people aged 60 years and above. Taking into account the results presented in the literature [19–23] we designed a set of functionalities to bring alive our chatbot. A chatbot able to develop easy connections with older users, and for this, we developed it as an empathetic, sensitive, sociable and friendly robot.

Some studies (e.g. [24]) suggest avoiding defining a chatbot default gender. Indeed overly humanized agents could create a higher expectation on users, which eventually leads to more frustration when the chatbot fails [6]. Despite the name borrowed from nice characters of films and comics, we designed Charlie as a bot (as shown in Charlie's icons in Figs. 1, 2 and 4).

About Charlie's choice of age, we decided to give it child traits. This because children and the elderly can in some ways be very similar. They both need to be looked after, taken by the hand, helped to cheer up or relax. Such context of use should feel friendly since Charlie invites the elderly into an intimate one-on-one chat space [23]. Moreover, Charlie talks a lot about himself but he also needs to be very good at listening to what others have to say. Charlie is so used as a sort of psychiatric counselling to change aspects of user habits and lifestyle, such as the habit of taking medicine or drinking waters more often [25, 26].

From a technical point of view, the chatbot agent is designed and developed by using Dialogflow [27]. DialogFlow is a natural language processing (NLP) platform provided by Google Corporation that can be used to build conversational applications and experiences on multiple platforms (e.g. Facebook, Messenger) or devices (e.g. Google Home. To create a customized web interface for Charlie, the interactions are carried out through a service built in *node.js* [28]. Charlie's functionalities have been designed to allow him to know the user's preferences. To do it, Charlie triggers specific Google Cloud functions [29] that are used to save the user's preferences, intentions and actions in a database of the Firebase platform [30].

For example, Charlie can trigger small talks about sports or the world of celebrities (Fig. 1). Since the user expresses the intention to find in-depth this argument, Charlie will save the preference in the internal database so that the subsequent interactions will comply with it. Moreover, as depicted in Fig. 2, Charlie can also send healthy tips to the user every day and suggests to her/him how to follow them. Adopting a gamification strategy, in the evening Charlie will ask the user if she/he has followed the advice or not and if yes, she/he will receive a "bot-coin" (a sort of recognition/prize). In exchange for recognition and rewards, this approach aims at helping users to follow a healthy life.

Charlie can also help if the user is subject to cognitive impairment and memory reduction. To do it, he can also ask the user if she/he needs help to remember something and helps her/him to fix a timetable (Fig. 2 – Second screenshot).

Another Charlie's functionality is used to provide the user with short quizzes on preferred topics to stimulate her/him and to satisfy her/his need for entertainment. When a quiz is activated, the user can select the number and difficulty of the questions (Fig. 3). Once the quiz is finished, Charlie will display the result obtained, with the correct and wrong answers (Fig. 3 – Second screenshot). In this way, Charlie can save the number of quizzes carried out during the day and statistics on wrong and correct answers in the Firebase database.

Finally, Charlie can ask the user for helping him in solving some riddles (Fig. 4 – First screenshot) or he can talk about an anecdote about his life that can generate self-compassion in the user (Fig. 4 – Second screenshot). These anecdotes see him as the protagonist and they are unfortunate or unpleasant events. According to some studies [12], this would lead to self-compassion in the user, a stimulus to reflect on one's life and a sense of identification. First of all, Charlie involves the user by asking her/him for an opinion on the matter and, if the user wishes, she/he can also share a similar episode. This would then help develop self-compassion and fill the need to be heard and understood.

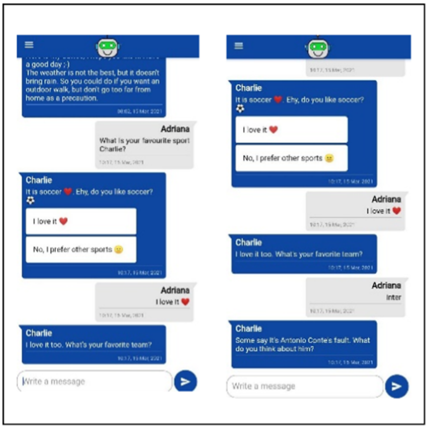

Fig. 1. Example of engagement interaction strategy provided by Charlie.

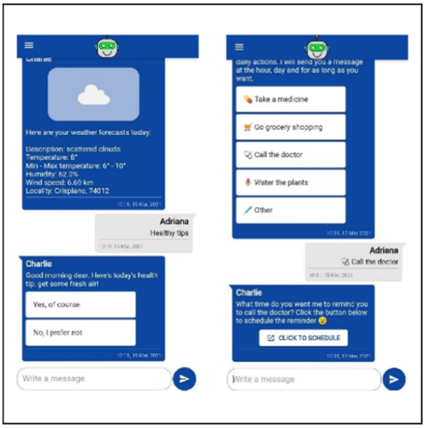

Fig. 2. Example of a healthy tip that Charlie can give to the user. In the second screenshot, Charlie asks the user if she/he needs help to remember something and helps her/him to fix a timetable.

3.1 A Preliminary Evaluation of Charlie's Functionalities

To evaluate Charlie's personality and the level of acceptability by senior users, we conducted a test involving four students of the bachelor degree in Computer Science for New Media communications and two students of master degree in Computer Science at the University of Milano. Each student contacted a couple of aged relatives, who after signed an informed consent, have been involved in the evaluation of Charlie's personality and acceptance.

All participants (12 in total, aged 60 through 70 years old) live in pairs or alone, and have a quite good skill and inclination in using technology (all have a smartphone and use it to chat, see videos on YouTube, browse the Web). We asked testers to interact with the agents daily in their homes for a week. After that, students carried a set of interviews and accordingly defined an affinity diagram [31] (Fig. 5) to group keywords and recurring themes that have meanings considering the experimented interactions.

The adopted structured interview is based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model, a model based on TAM [32], that has been tested extensively in various fields and promises to be a great tool for analyzing users acceptance of health technology [33–35]. In [36] authors demonstrate how this model can effectively measure older users' perceptions and their level of acceptance in using healthcare applications. According to this study, we designed specific questions to measure the strength of one's intention to perform a specified behaviour [37] and questions to investigate how much users considered easy to use Charlie.

Specifically, we provided senior users with questions for measuring the evoking anxious or emotional reactions they experimented with when they used our chatbot, for measuring their trust in Charlie in terms of safety and reliability, and for measuring how much physicians' opinions can influence their perception of technology as useful.

Although the number of interviews makes it impossible to present a statistically significant analysis, we carried out a qualitative study that allowed us to define a set of keywords or themes that are used to assess the perception of Charlie's personality and his level of acceptability by users. In details, when it comes to Charlie's personality, testers described him as a young, cheerful, active, friendly and smart assist (Fig. 6). The perceived age of Charlie varied, but most of the participants agreed that he had the innocence of childhood regardless of age.

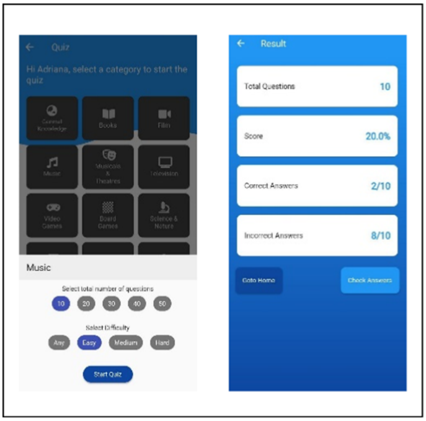

Fig. 3. By using this screenshot, the user can select the number of questions and the difficulties of the quiz. The second final screenshot is used by Charlie to visualize correct and wrong answers.

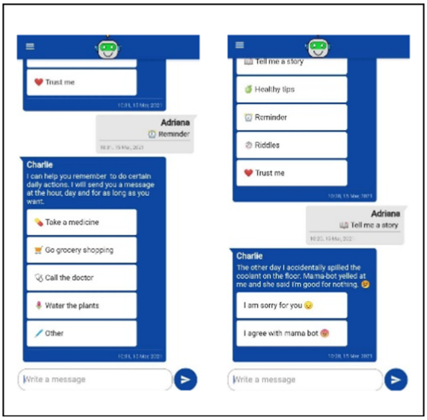

Fig. 4. In the first screenshot, Charlie presents a riddle and in the second, he presents an example of a self-compassion strategy.

In the evaluation of the core of Charlie's personality, more people mentioned that the chatbot was efficient and helpful, while fewer people described the chatbot as imaginative or creative. Moreover, in some cases, people perceived Charlie's chatbot to be a robot rather than a human. Participants reported that how the chatbot used words and phrases contributed to their impression of the chatbot's personality. They reported that they felt the chatbot was warm-hearted, energetic, or cheerful. Meanwhile, a few participants mentioned that the number of words per message and the number of messages that the chatbot sent at a time led to perceptions that the chatbot was pushy and compulsive. Another important cue upon which the research participants when judging the chatbot's personality was the visual cues. In particular, emojis played a crucial role in affecting participants' judgment of the chatbot's gender and characteristics. Several interviewees mentioned that the emojis made them feel that the chatbot was cheerful, friendly, and approachable. Meanwhile, the emojis also affected interviewees' perceptions of the chatbot's gender. Two participants mentioned that the emojis that Charlie used remembered them that Charlie is a robot. In general, the affinity diagram (Fig. 5) highlight how Charlie is considered polite, smart, charming, helpful and reliable.

The results of the performed analysis have shown very comforting outcomes even if it is necessary to carry out more detailed tests with real users to be able to give clearer and more truthful information in the middle and long period. Unfortunately, due to the current situation linked to the COVID 19 pandemic, these tests are still to be carried out.

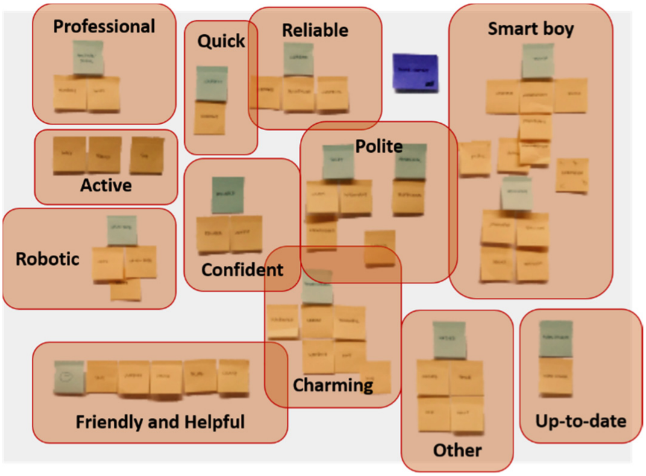

Fig. 5. Affinity diagram to sort cues that shape Charlie's personality.

"Boy, 6 years old, playful, enterprising kid." (tester n. 4)

"Constructive and cheerful, outgoing and[has] child-like attitude towards daily routines." (tester n. 2)

"He would be a playful boy" (tester n.5)

Fig. 6. Examples of answers about Charlie's personality of participants at the test.

4 A EUD Solution for Training the Virtual Agent

4.1 EUD Strategies for Designing Chatbots

Charlie's functionalities described in the previous section allow him to act as a friend and a caregiver for the elderly. In addition to supporting the dear seniors in their daily activities, Charlie has the task to monitor her/his behaviours to detect eventually anomaly situations and warn the health assistant or relatives accordingly. To define what to monitor we designed a web application that can be used by health assistants or relatives themselves. The idea is to adopt an end-user development (EUD [38]) strategy to involve them in the specification of rules that best fit their dear seniors [39, 40]. EUD represents the ideal approach for empowering non-technical users making them developers of rules [41, 42] that will be at the base of the chatbot design. In the EUD approach, this problem can be faced through a design environment where visual entities representing the conditions that need to be connected graphically to define the sequence of operations for monitoring specific situations. Visual strategies typically used for modelling Event-Condition-Action rules can be described through the most famous systems that apply them: IFTTT, Atooma, and Yahoo's Pipes. In [43–45] authors discuss how the first two design strategies support users without programming knowledge to define their context-dependent applications. Specifically [46] describes how these EUD strategies allow users to define sets of desired behaviours in response to specific events. This is made mainly through rules definition-wizards. Rules can be typically chosen among existing ones or can be tweaked through customization.

Based on the results of these studies, we devised a new solution specifically addressed to enable non-technical users in designing rules to monitor the elderly's behaviours and gathering data according to spatio-temporal dimensions.

4.2 EUD Strategies to Train Charlie

The web app has been designed on a previous application, SmartFit [43] specifically developed for allowing coaches and athletic trainers of non-professional sports teams to monitor and analyze data regarding the fitness and well-being of the athletes.

Designed on SmartFit, our web app adopt EUD techniques for supporting non-experts in computer science in designing rules. In detail, it enables caregivers and relatives in capturing a set of measurements describing their dear's behaviour over a period, typically a day. Such measurements concern activities carried out by the elderly such as the number of solved quizzes or riddles, how many she/he retrieved news on her/his favourite topics, how many Charlie's suggestions or stories she/he have followed and listened to. We are also endowing the elderly with an electronic bracelet for gathering biometric data such as the heartbeat, the quality of sleep (hours of sleep, number of awakenings for the night, and minutes of restless sleep), the burned calories and physical activities (number of steps or kilometres walked).

Based on these gathered data, the web app aims at supporting non-expert users in the composition of rules based on an ECA (Event-Condition-Action) paradigm. Adopting this paradigm, users can specify conditions and temporal operations for implementing the rules. For example, Fig. 7 presents a dashboard on which the caregivers can visualize the set of defined rules to monitor her/his dear senior's behaviour. The conditions can be composed by using simple drop-down menus to combine groups of statements connected through AND/OR operators. The order of the conditions can be changed just by dragging and dropping the statements into the right position.

The strategy used to create rules extends the IF-THIS-THEN-THAT approach and supports the definition of time dimensions that are exploited for expressing more articulated rules. The time dimension is used for creating rules by using temporal operators that point out temporal correlations among relevant events. In detail, users can use a set of temporal operators such as *opt* ∈ *{before, after, when}* to specify if she/he wants to monitor events that happen according to specific time constraints. For example, as depicted in Fig. 8 the user can specify to monitor if her/her dear has had a difficult night (e.g. less than 5 h of sleep with at least 3 awakenings) when she/he did not take all the medicines the day before.

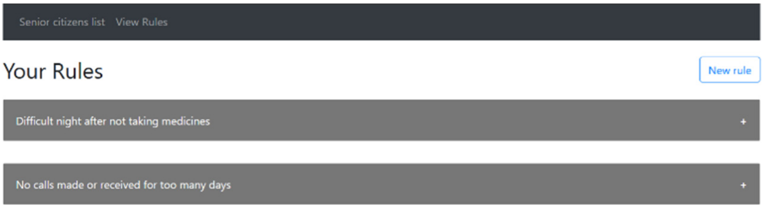

Fig. 7. Screenshot of the dashboard used to visualize the defined rules to monitor the elderly's behaviour

In Sect. 1 of this Fig. 8, the user indicates the name of the rule and the message that will arise when the conditions will be met. In Sect. 2, the user specifies a rule that is used to check if the aged person has slept less than 5 h AND that the number of awakenings for the night has been more than 3 AND that the day before the monitored senior has not taken her/his medicines; or she/he has taken less than a given number of medicines (e.g. 5) that the user can specify using the proper label. Finally, in Sect. 3 the user indicates the temporal correlation that exists between the conditions expressed in Sect. 2.

As soon as a rule is defined, it is implemented as a new Google Cloud function that by accessing the data stored in the database of the Firebase platform allows Charlie to check if the rules will be verified. When the conditions of the rule will be met, a notification (expressed in Sect. 1 of Fig. 8) will be raised both in the web app to warn caregivers and in the bot to alarm the user.

4.3 Intelligent Suggestions of Rules

By using specific rules, caregivers can detect relevant and significant events that may affect their dears' life and that depend on data gathered by Charlie.

Unfortunately, the problem with this system is that the caregiver is often "lost in the data sea" and does not know which parameters she/he has to check for monitoring the senior's behaviour. To help caregivers in the monitoring task, we are extending our web app with machine learning techniques that enable predictions based on counterfactual explanations [47, 48], to compute suggestions about the better rules to adopt.

Essentially, suggestions are provided as rules, which describe a change in the behaviours. The idea is to endow the web app with functionality to compute predictions and the consequent rules, by extracting knowledge from the data describing the elderly's behaviours. In particular, we are investigating a machine learning approach that produces a set of readable and understandable IF-THEN rules, which are easily interpretable and follow a similar pattern to human thinking.

To a first approximation, an initial step is to identify recurrent behaviours that can be used to classify the elderly's lifestyle through a set of labels, such as "depression", "sense of loneliness", "mood disorders", "anxiety", "normal" according to Geriatric Depression Scale [49–51]. To do it, we need to choose the so-called "right data" [52] by selecting the most informative (discriminating) features while discarding redundant information. To identify the most informative features, we are analysing the data collected during the tests performed with aged relatives of our university students (as described in Sect. 3.1). During the monitoring weeks involving 12 elderly, the collected data daily by Charlie

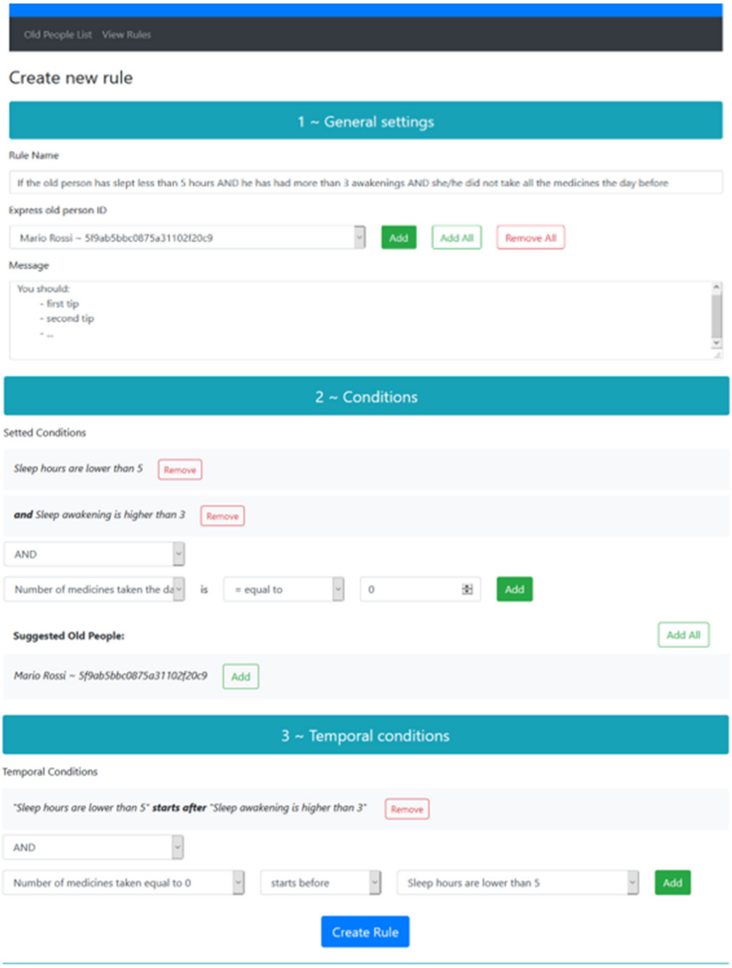

Fig. 8. Workflow used to specify a new rule. In Sect. 1, the user indicates the name of the rule and the message that will arise when the conditions will be met. In Sect. 2, by using AND OR operators, the user specifies the condition to monitor. Finally, in Sect. 3 the user indicates the temporal correlation that exists between the conditions expressed in Sect. 2. In the depicted case, the user adds a new temporal condition specifying to check if the monitored senior has slept less than 5 h **after** that she/he has had more than 3 awakenings AND that **the day before** she/he has taken less than 5 medicines.

were: The number of completed quizzes, statistics on wrong and correct answers, the number of news and stories requested or told by the user, the number of times the user forgot the take medicines or to call her/his relatives (in case she/he scheduled a reminder about it). With the caregivers' help, we are working for recruiting other seniors whose behaviours can fit the labels used to define the lifestyles. In this way, we can train our recommendation service and evaluate better prediction models.

The statistical analyses that we are carrying out by using ANOVA [53] aim at computing the p-values to check if the feature distribution in different classes have the same mean, against the alternative hypothesis that they have different means. The final goal is to keep only features for which the feature distribution in different classes have different means, that is, we want to keep the features for which the p-value allows us to discard the null hypothesis at 5% significance level, that is $p < 0.05$ (95% confidence). Of course, to obtain more meaningful results we need to complete the current data collection to monitor the elderly's behaviour. Moreover, to increase the number of features from which to select more discriminative ones, our idea is also to endow the elderly with an electronic bracelet for gathering biometric data such as the heartbeat, the quality of sleep, burned calories and physical activities. Unfortunately, gathering these data is experiencing considerable delays for reasons due to the CODIV pandemic. Anyway, we designed an analysis protocol as described below.

4.4 Protocol for Counterfactual Explanations

Once the ANOVA test will be completed, a second step will aim at calculating the inter-features relationships to check if discriminative features are highly correlated that is if the redundant information is included in the selected feature set. To remove such redundant information, we will compute the pairwise Pearson linear correlation coefficient between each pair of feature vectors. For each couple of features, of which we obtain an absolute Pearson correlation coefficient ≥ 0.5, we will remove the less discriminative feature that is the feature for which the corresponding p-value is higher.

At this point, the next step of our protocol of analysis will deal with the multiclass classification problem by using an internal Leave-One-Out cross-validation (LOOCV) procedure based on ECOC models, which have been shown to improve the performance of multiclass classifiers. Rather than limiting the choice of algorithms or adapting the algorithms for multi-class problems, ECOC approach is to reframe the multi-class classification problem as multiple binary classification problems. Two common methods that can be used to achieve this include a KNN, and an SVM classifier. Our idea is to employ a greedy method that will sort the selected features according to their *p-values* (from the smallest to the highest), and, for each classifier, chooses the first features which allow minimizing the classification error computed through LOOCV.

Once the best predictive model will be detected, the data that Charlie is collecting will be used to compute prescriptive suggestions for improving the elderly's behaviour, based on the computation of the so-called counterfactual explanations. One of the main goals of our study is the ability to explain the computed predictions. To open the so-called "black-boxes", a great deal of research work have been recently devoted to the development of automatic techniques for explaining the predictions computed by machine learning methods, and the growing literature about this subject has indeed generated a novel field of research called "interpretable machine learning" [47, 48]. Among the different techniques for explaining the classifiers' decision, counterfactual explanations are sometimes preferred because they explain the output of the classifier for the specific input point.

Counterfactual Explanations are defined as statements taking the form [48]: "*Score p was returned because variables V had values (v1, v2, …) associated with them. If V instead had values (v1', v2', …), and all other variables had remained constant, score p' would have been returned*". In the field of pattern classification, counterfactual explanations explain how to change the values of the input data to obtain the desired classification result. Therefore, by computing counterfactual explanations we may provide prescriptive suggestions. In our context of use, thinking in counterfactuals terms requires imagining how the elderly's acceptable behaviour (labelled as "normal") contradicts the observed undesired situations such as "depression", "sense of loneliness", "mood disorders", "anxiety". To do it we have to compute the smallest changes to the feature values of the current prevision to transform them into values that characterize the "normal" status. For example for a person labelled as "anxious", we need to understand how to mitigate this situation. A possible suggestion could be: If you want to avoid anxiety attacks you will have to increase your physical activity (the number of steps taken during the day has to be more than 1000 units and the calories burned more than 200 kilocalories), to have daily social interaction with your relatives (ask Charlie to schedule a daily reminder to call your children), and to solve more than 3 quizzes that Charlie will be glad to offer you.

Based on these counterfactual explanations, we are working to endow our web application with functionality to suggest to caregivers new rules able to check anomalous situations and to provide suggestions on how to remedy them.

5 Conclusions

In this paper, we have presented Charlie as an empathetic, sensitive, sociable and friendly child robot that provides the elderly with interactive activities based on s gamification, active notifications, and promotion of self-compassion. According to the analysis of discussions carried out with the elderly involved in the preliminary test, we can conclude that Charlie is considered by users polite, smart, charming, helpful and reliable. This solution can offer a pool of interactive strategies that can relieve the state of loneliness the aged people live. The final goal is to support pre-emptive care for improving the elderly's quality of life, avoiding necessarily "fixing" something but offering assistance and companionship without assuming diagnosed disorders. The chatbot has been also presented during a hackathon organized by Facebook and Funka aimed at encouraging students from Europe to create innovative digital solutions to increase social inclusion. Charlie, after being submitted to a jury made up of representatives of organizations for disabled people, policymakers and industries, was among the top three finalists.[1]

Then the paper describes our studies on how to allow caregivers, such as medical assistants or relatives, to implement in Charlie the possibility to monitor their dear seniors' behaviours. To this aim, we designed a web application that enables caregivers to create rules for detect relevant and significant events that depend on data such as the number of required news, followed tips, completed games, suggestions to Charlie on how to solve his problems, but also biometrical data that can be achieved by using

[1] http://www.euroblind.org/events/facebook-funka-accessibility-hackathon https://www.funka.com/en/about-funka2/news/en/facebook-funka-accessibility-hackathon-winner/.

smart wristbands such as the number of steps made during the day, burned calories or the number of hours of sleep.

In the design of the web app, we dealt with the problem that rarely users know which parameters to follow for better assisting their dear seniors. To help them in the monitoring task, we need to use machine-learning techniques that enable predictions and to do it, we are studying a method, based on counterfactual explanations, to compute suggestions for improving the elderly's quality of life. Suggestions that are then transformed into rules integrated into the chatbot intents for extracting knowledge from the collected measurements describing the elderly's behaviour. While creating these rules, users must specify what happens when a particular set of conditions is met/not met by defining a list of actions to be performed, such as by specifying that a warning needs to be sent via direct messages. In this way, we hope to help caregivers and relatives in the definition of what Charlie needs to monitor for taking under control their dear's behaviours and attitudes.

Despite the limitations due to the pandemic situation, at the moment we are planning several user tests to, on the one hand, gather data for training our machine learning model and on the other hand to test the of use Charlie and our web application in real contexts of use. In this way, we will have the possibility to investigate the outcomes of the application of our method based on counterfactual explanations.

References

1. https://agcult.it/a/22645/2020-07-23/italiani-e-solitudine-symbola-nuove-tecnologie-ma-anche-lettura-e-impegno-civico-gli-antidoti-piu-diffusi. Accessed Apr 2021
2. https://www.enuan.com/. Accessed Apr 2021
3. Chakrabarti, C., Luger, G.F.: A semantic architecture for artificial conversations. In: The 6th International Conference on Soft Computing and Intelligent Systems, and the 13th International Symposium on Advanced Intelligence Systems, pp. 21–26. IEEE (2012)
4. Jain, A.M.D., Daniel, D., Fraser, H., Saravanakumar, S., Nair-Hartman, A.: The emergence of value-based health: how healthcare is using technology to create insights, enhance efficiency, and improve patient outcomes. IBM Institute for Business Value (2019). http://ibm.co/value-based-health
5. Rapp, A., Curti, L., Boldi, A.: The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots. Int. J. Hum. Comput. Stud. **151**, 102630 (2021)
6. Chaves, A.P., Gerosa, M.A.: How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design. Int. J. Hum. Comput. Interact., 1–30 (2020)
7. Muppirishetty, P., Lee, M.: Voice user interfaces for mental healthcare: leveraging technology to help our inner voice. In: 3rd ACM Conference on Computer-Supported Cooperative Work and Social Computing, CSCW (2020)
8. Valério, F.A.M., Guimarães, T.G., Prates, R.O., Candello, H.: Chatbots explain themselves: designers' strategies for conveying chatbot features to users. SBC J. Interact. Syst. **9**(3), 61–79 (2018)
9. Valério, F.A., Guimarães, T.G., Prates, R.O., Candello, H.: Here's what i can do: chatbots' strategies to convey their features to users. In: Proceedings of the Xvi Brazilian Symposium on Human Factors in Computing Systems, pp. 1–10 (2017)

10. Bickmore, T.W., Caruso, L., Clough-Gorr, K.: Acceptance and usability of a relational agent interface by urban older adults. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1212–1215 (2005)
11. Bickmore, T.W., Picard, R.W.: Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput. Hum. Interact. (TOCHI) **12**(2), 293–327 (2005)
12. Lee, M., Ackermans, S., van As, N., Chang, H., Lucas, E., IJsselsteijn, W.: Caring for Vincent: a chatbot for self-compassion. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)
13. Smith, J.: GrandChair: conversational collection of grandparents' stories. Doctoral dissertation, Massachusetts Institute of Technology (2000)
14. Bickmore, T.W., Mitchell, S.E., Jack, B.W., Paasche-Orlow, M.K., Pfeifer, L.M., O'Donnell, J.: Response to a relational agent by hospital patients with depressive symptoms. Interact. Comput. **22**(4), 289–298 (2010)
15. Bickmore, T.W., Schulman, D., Sidner, C.L.: A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. J. Biomed. Inform. **44**(2), 183–197 (2011)
16. King, A., Bickmore, T., Campero, I., Pruitt, L., Yin, L.X.: Employing "virtual advisors" to promote physical activity in underserved communities: results from the COMPASS study. Ann. Behav. Med. **41**, S58 (2011)
17. Gentner, T., Neitzel, T., Schulze, J., Buettner, R.: A Systematic literature review of medical chatbot research from a behavior change perspective. In: 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 735–740. IEEE (2020)
18. Melia, R., et al.: Mobile health technology interventions for suicide prevention: systematic review. JMIR mHealth uHealth **8**(1), e12516 (2020)
19. Nass, C., Moon, Y.: Machines and mindlessness: Social responses to computers. J. Soc. Issues **56**(1), 81–103 (2000)
20. Reeves, B., Nass, C.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People. Cambridge University Press, Cambridge (1996)
21. De Angeli, A., Johnson, G.I., Coventry, L.: The unfriendly user: exploring social reactions to chatterbots. In: Proceedings of The International Conference on Affective Human Factors Design, London, pp. 467–474 (2001)
22. Brahnam, S., De Angeli, A.: Gender affordances of conversational agents. Interact. Comput. **24**(3), 139–153 (2012)
23. Lee, S., Choi, J.: Enhancing user experience with conversational agent for movie recommendation: effects of self-disclosure and reciprocity. Int. J. Hum. Comput. Stud. **103**, 95–105 (2017)
24. Neff, G., Nagy, P.: Automation, algorithms, and politics| talking to bots: symbiotic agency and the case of tay. Int. J. Commun. **10**, 17 (2016)
25. Oh, K.J., Lee, D., Ko, B., Choi, H.J.: A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In: 2017 18th IEEE International Conference on Mobile Data Management (MDM), pp. 371–375. IEEE (2017)
26. Lee, D., Oh, K.J., Choi, H.J.: The chatbot feels you-a counseling service using emotional response generation. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 437–440. IEEE (2017)
27. Dialogflow API. https://cloud.google.com/dialogflow/es/docs/reference/rest/v2-overview. Accessed Apr 2021
28. OpenJS Foundation: About Node.js. https://nodejs.org/en/about/. Accessed Apr 2021
29. https://cloud.google.com/functions. Accessed Apr 2021
30. https://firebase.google.com/. Accessed Apr 2021

31. Lucero, A.: Using affinity diagrams to evaluate interactive prototypes. In: Abascal, J., Barbosa, S., Fetter, M., Gross, T., Palanque, P., Winckler, M. (eds.) INTERACT 2015. LNCS, vol. 9297, pp. 231–248. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22668-2_19
32. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **27**, 425–478 (2003)
33. De Veer, A.J., Peeters, J.M., Brabers, A.E., Schellevis, F.G., Rademakers, J.J.J., Francke, A.L.: Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv. Res. **15**(1), 1–9 (2015)
34. Liu, C.F., Tsai, Y.C., Jang, F.L.: Patients' acceptance towards a web-based personal health record system: an empirical study in Taiwan. Int. J. Environ. Res. Public Health **10**(10), 5191–5208 (2013)
35. Kohnke, A., Cole, M.L., Bush, R.: Incorporating UTAUT predictors for understanding home care patients' and clinician's acceptance of healthcare telemedicine equipment. J. Technol. Manag. Innov. **9**(2), 29–41 (2014)
36. Cimperman, M., Brenčič, M.M., Trkman, P.: Analyzing older users' home telehealth services acceptance behavior—applying an extended UTAUT model. Int. J. Med. Inform. **90**, 22–31 (2016)
37. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: Extrinsic and intrinsic motivation to use computers in the workplace 1. J. Appl. Soc. Psychol. **22**(14), 1111–1132 (1992)
38. Koch, M.: End-user development. Wirtschaftsinformatik **48**(6), 455 (2006). https://doi.org/10.1007/s11576-006-0107-x
39. Petre, M., Blackwell, A.F.: Children as unwitting end-user programmers. In: Proceeding ofthe IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2007), pp. 239–242 (2007)
40. Fischer, G., Giaccardi, E., Ye, Y., Sutcliffe, A., Mehandjiev, N.: Meta-design: a manifesto for end-user development. Commun. ACM **47**(9), 33–37 (2004)
41. Costabile, M.F., Mussio, P., Parasiliti Provenza, L., Piccinno, A.: End users as unwitting software developers. In: Proceedings of the 4th International Workshop on End-User Software Engineering, pp. 6–10. ACM, New York (2008)
42. Barricelli, B.R., Valtolina, S.: A visual language and interactive system for end-user development of internet of things ecosystems. J. Vis. Lang. Comput. **40**, 1–19 (2017)
43. Valtolina, S., Barricelli, B.R.: An end-user development framework to support quantified self in sport teams. In: Paternò, F., Wulf, V. (eds.) New Perspectives in End-User Development, pp. 413–432. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60291-2_16. ISBN 9783319602905
44. Ghiani, G., Manca, M., Paternò, F., Santoro, C.: Personalization of context-dependent applications through trigger-action rules. ACM Trans. Comput. Hum. Interact. **24**(2), 33 (2017)
45. Desolda, G., Ardito, C., Matera, M.: Empowering end users to customize their smart environments: model, composition paradigms and domain-specific tools. ACM Trans. Comput. Hum. Interact. **24**(2), 52 (2017)
46. Caivano, D., Fogli, D., Lanzilotti, R., Piccinno, A., Cassano, F.: Supporting end users to control their smart home: design implications from a literature review and an empirical investigation. J. Syst. Softw. **144**(2018), 295–313 (2018)
47. Molnar, C.: Interpretable machine learning. Lulu.com (2020). https://christophm.github.io/interpretable-ml-book/
48. Wachter, S., Mittelstadt, B., Russell, C.: Counterfactual explanations without opening the black box: automated decisions and the GDPR. Harv. JL & Tech. **31**, 841 (2017)
49. http://www.minddisorders.com/Flu-Inv/Geriatric-Depression-Scale.html. Accessed Apr 2021

50. Holmén, K., Ericsson, K., Winblad, B.: Quality of life among the elderly: state of mood and loneliness in two selected groups. Scand. J. Caring Sci. **13**(2), 91–95 (1999)
51. Gerino, E., Rollè, L., Sechi, C., Brustia, P.: Loneliness, resilience, mental health, and quality of life in old age: a structural equation model. Front. Psychol. **8**, 2003 (2017)
52. Fisher, R.A:. XV.—the correlation between relatives on the supposition of Mendelian inheritance. Earth Environ. Sci. Trans. Roy. Soc. Edinb. **52**(2), 399–433 (1919)
53. Kajdanowicz, T., Wozniak, M., Kazienko, P.: Multiple classifier method for structured output prediction based on error correcting output codes. In: Nguyen, N.T., Kim, C.-G., Janiak, A. (eds.) ACIIDS 2011. LNCS (LNAI), vol. 6592, pp. 333–342. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-20042-7_34