

Virtual Assistant for Telemedicine

Vridhhi Pai, Atharv Shiromani, and Shubham Kathuria

Abstract The advancement of technology in terms of low-cost, portable, hand-held devices like cell phones, bluetooth and WiFi permits users' mobility at an unparalleled level. The ability of these technologies to function collectively can be utilized to give rise to new tools in different fields thereby contributing to the proliferation of mankind. In the current scenario, ubiquitous computing is extending to the healthcare sector where increasing use of virtual assistance is becoming commonplace. This stems from the need to provide healthcare beyond the boundaries of user's location and availability of a medical professional. The reluctance of even technical enthusiasts regarding idea of relying entirely on an artificial system has become a perpetual obstacle in bringing this technology to the forefront. In this paper, we discuss the nuances of relying on a virtual pre-trained feedback system in telemedicine posing as a virtual assistant to help with a target user's medical condition- regular checkup and/or emergency. The assistant aims to implement pervasive healthcare by utilizing the convenience of handheld devices such as cellphones. The application can be utilized by people with common medical conditions and also aims to cater to users with chronic medical conditions in need of constant monitoring.

Index Terms—Virtual assistant, telemedicine, pervasive health care, ubiquitous computing

INTRODUCTION

The advancement of technology in terms of low-cost, portable, hand-held devices like cell phones, bluetooth and WiFi permits users' mobility at an unparalleled rate. The ability of these technologies to function collectively can be utilized to give rise to new tools in different fields thereby contributing to the proliferation of mankind. In the current scenario, ubiquitous computing is extending to the healthcare sector where increasing use of virtual assistance is becoming commonplace. However, every new venture is implicitly fated to deal with pitfalls that are resolved over time with extensive research and upgradations, finally leading to reliable systems with intended functionalities in place. The goal of 'pervasive health care' is to provide healthcare services to everyone at any time overcoming the constraints of place, time and availability of doctors, nurses and resident doctors. Handheld devices demonstrate great promise as point of care information devices. We aim to identify the scope of relying on a model/system that works entirely/nearly entirely on a virtual assistant in order to assess the possibility of making human-machine interaction optimized yet efficient. The scope of our work is in the field of telemedicine where the requirement is to utilize the strengths of a virtual voice assistant in interacting and discerning a user's vocal input and in the process identify the loopholes in the system to give way to research in making this idea a growing reality.

The benefits of pervasive healthcare show profound effects with specific users such as people with obesity, fluctuating blood pressure, heart conditions, diabetes, age-related regular health concerns and other common medical apprehensions which may require spontaneous attention and regular vigilance. Instead of having to make frequent uncalled for trips to the physician, having a personalized feedback system that is trained to understand and provide quality assistance at the right time becomes a valuable, cost-effective and convenient alternative.

We have attempted to come up with a feedback-based communication system in the form of a virtual assistant that is trained to deal with known medical conditions and is conveniently embedded in a portable PDA such as a cell phone

of the target user. The research challenges include: understanding and analyzing complex interactions, constructing and communicating logical and contextual responses, promptness in response, reliability in dealing with said medical condition and user convenience with the prototype.

The paper proceeds as follows: Section 2 of the paper describes the Related Work which essentially underscores the motivation for this research, Section 3 gives an insight into the methodology adopted in building and training the system, Section 4 describes the observed results, Section 5 is the Discussion section, Section 6 is a conclusive section that summarizes the paper along with describing the future prospects in related research.

RELATED WORK

Both, the use of virtual assistant and the concept of telemedicine have been around for longer than one would expect, thanks to ongoing research and developments on some profound ideas in the fields. One such pair of researchers, Shazia, and Imran [1] came up with an easy-to-set-up and low-cost method of 'Store-and-Forward' for telemedicine. In this, in order to deal with high response time, they developed the virtual system 'Clinic Decision Support System (CDSS)', deployed directly where the virtual telemedicine is needed. The CDSS is intelligent enough to diagnose a patient's condition and spontaneously prescribe a cure. In a case where the CDSS cannot answer a query, the CDSS immediately sends an e-mail to a medical expert (doctor) and when the response is received the CDSS knowledge-base is updated for the future queries. In this research paper, we not only report an NL-based CDSS that can answer NL queries but also present a complete architecture of a virtual telemedicine setup. The accuracy achieved with the designed system is 91.00%, based on the system's capacity to understand and give appropriate responses in a reasonable time. Dena Puskin, Barbara and Stuart presented a framework [2] of a telehealth system that was able to identify and understand the interaction between telemedicine services. Exploration of health information technology [2] (HIT) applications on local,

regional and national levels was the major emphasis of this research. Further and more specifically in the voice-based application, Imran and Abbas [3] introduce a newly designed rule-based framework that is able to read the English language text and extract its meanings after analyzing and extracting related information. After composite analysis and extraction of associated information, the designed system gives particular meaning to an assortment of speech-language text on the basis of its context. The designed system uses standard speech language rules that are clearly defined for all spoken languages like English, Urdu, Chinese, Arabic, French, etc. The designed system provides a quick and reliable way to comprehend speech language context and generate respective meanings. The current designed system incorporates the capability of mapping user requirements after reading the given requirements in plain text and drawing the set of speech language contents. In [4], the researchers investigate the task of building an open domain, conversational dialogue systems based on large dialogue corpora using generative models. Generative models produce system responses that are autonomously generated word-by-word, opening to the possibility for realistic, flexible interactions. In their work, they also show how performance can be improved by bootstrapping the learning from a larger question-answer pair corpus and from pre-trained word embeddings [5]. provides key insights into the construction and evaluation of interpersonal simulators—systems that enable interpersonal interaction with virtual humans. Using an interpersonal simulator, two studies were conducted that compare interaction with a virtual human to interactions with a similar real human. The specific interpersonal scenario employed was that of a medical interview. Medical students interacted with either a virtual human simulating appendicitis or a real human pretending to have the same symptoms. The medical students elicited the same information from the virtual and real human, indicating that the content of the virtual and real interactions was similar. However, participants appeared less engaged and insincere with the virtual human. These behavioral differences likely stemmed from the virtual human's limited expressive behavior. [6] introduces the 'SimCoach project' that aims to develop virtual human support agents to serve as online guides for promoting access to psychological healthcare information and for assisting military personnel and family members in breaking down barriers to initiating care. The SimCoach project that aims to address this need by developing virtual human support agents to serve as online guides for promoting access to psychological healthcare information and for assisting military personnel and family members in breaking down barriers to initiating the healthcare process. SimCoach will allow users to initiate and engage in a dialog about their healthcare concerns with an interactive VH. The SimCoach project is not conceived to deliver diagnosis or treatment or as a replacement for human providers and experts. Instead, SimCoach will aim to start the process of engaging the user by providing support and encouragement, increasing awareness of their situation and treatment options, and in assisting individuals, who may otherwise be initially uncomfortable talking to a "live" care provider, in their efforts to initiate care. Users can flexibly interact with these VHs by typing text, clicking on character

generated menu options. [7] describes how contextual information from other participants can be used to predict visual feedback and improve recognition of head gestures in human-human interactions. This paper addresses two main challenges: optimal feature representation using an encoding dictionary and automatic selection of optimal feature-encoding pairs. Multimodal integration between context and visual observations is performed using a discriminative sequential model (Latent-Dynamic Conditional Random Fields) trained on previous interactions. In [8], The virtual assistant runs on a computer to be able to receive human voice communications from a remote user interface and transmit a vocalization to a user interface. The virtual assistant application selects from amongst a plurality of word patterns which can communicate a response to the verbal query/instruction' the selection being based on a user-controllable event. The key features of the virtual assistant is its compatibility with a messaging server 62, such as Microsoft Exchange/Outlook. In [9], the authors Ahmed, Haque and Stamm have described their approach of a 'Wellness Assistant'- a system necessarily running as an integrated system with sensing capabilities, communication, event and information management which performs efficient data collection and in combination with a scheduler, provides time-sensitive health care alerts to the target users based on their condition(s).

METHOD

In order to study and test the problems of a virtual assistant all that was needed was to design the virtual assistant from scratch and train it for every possible situation that the user might put it test. Training was probably the most complex and tedious part of the project. At first, we thought of using some sought of predefined API's but the whole point of this project was to study and research the problems created by the virtual assistant and work upon them in order to reduce them and in some case list them out as potential research areas for future work. For the sole purpose of training of training our assistant we used an online cloud console name IBM BLUEMIX. This is console is very efficient in what it does and provides support for a wide variety of programming languages. We have developed our assistant entirely in java. The best part is that it is trained over cloud, so our assistant always keeps on learning as we train it. Four different services of IBM Bluemix have been used namely Language Translator, Text to Speech, Speech to Text, IBM Watson. Watson is the main service used to develop our assistant, the whole interaction process is summarised in Fig. 1.

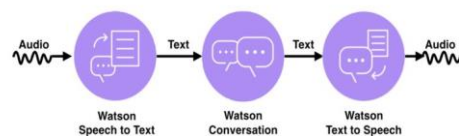


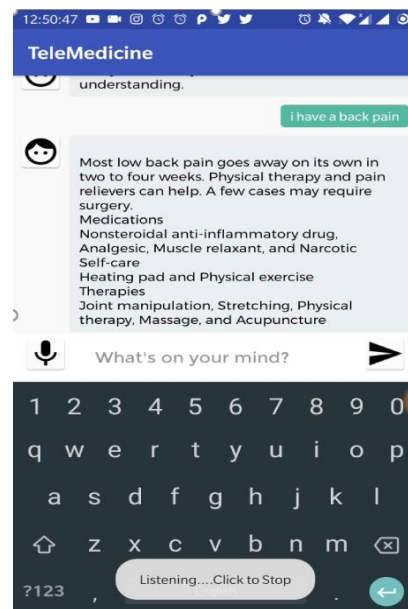
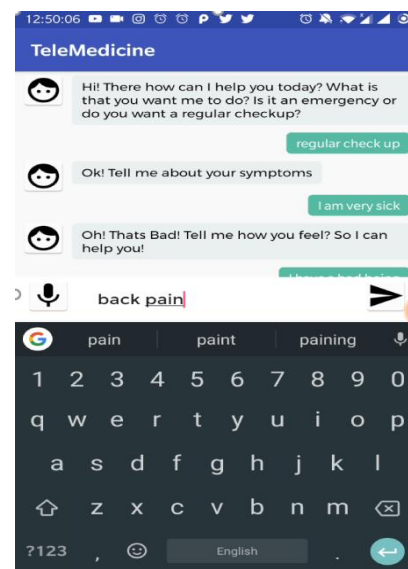
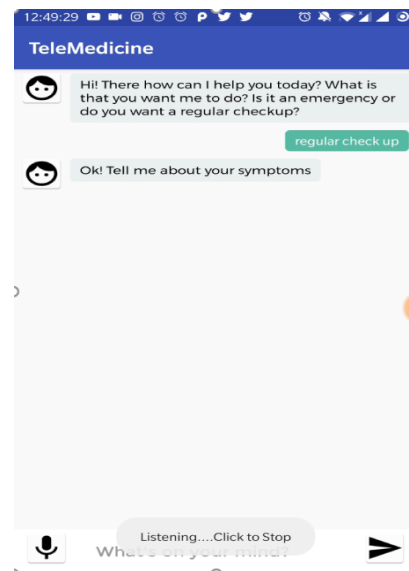
Fig. 1. Interaction procedure in IBM Watson Service

All the three services work in conjunction with each other. The whole process starts with the voice input of the user, the input is converted into text by speech to text service and then the inputs are matched to the knowledge of the Watson and if the

Watson has any reply for that particular question the response is sent back accordingly and then the text to speech functionality comes in action to speak out the response produced by Watson. The best part of all this is that the entire course of action is one to one and happens in real time so there is hardly any lag in the entire process. Going in deeper into the working of the Watson is where the actual design of our assistant lies. The Watson provides three sets of variables to control your assistant viz. Intents, Entities and Dialog. Intents are used to define the purpose of a user's input. All the request types that are supported by the assistant needs to be defined in the intents. Next are the entities, these are very important as these provides context to the assistant. These work in extension of your defined intents and store up values in the form of terms or objects that is somewhat related to your intents. Each value also contains synonyms that user might enter in order to refer to that entity and therefore the intent. Last is the dialog, this is the place where all the intents and entities are collectively combined to create a response for each of them. A user according to his/her liking can create responses that the assistant would answer in response to question asked by the user. I simple terms dialog can be defined as the branching conversation flow that defines responses to the defined intents and entities. With the above defined process, we trained our assistant to answer the expected user queries but the main complexity of the project was the part where the connection between the cloud and android application had to be made. Not only the android application had to be made to work in coalition with the cloud but also it had to be designed in such a way that it can display the types of responses created by our assistant. We developed around 2-3 applications where 2 applications were able to connect with Watson but out of those two only one was able to show the various types of response created. Further, in order to incorporate the aforementioned services into our application we need to define certain service credentials that could be used as an identity to our respective skill created.

RESULTS

A prototype application consisting of the virtual assistant fully equipped with Speech to Text and Text to Speech functionality has been developed which illustrates the scenario of a feedback-based health care assistance system needing no human intervention. This application exhibits the required abilities to predict the required health care services needed to tend to the target user's condition. Following is the glimpse of the prototype implementation of the application:



DISCUSSION

Our work primarily aimed to cater to a certain section of the population that composes of mainly two varieties of users: 1) those who have some chronic medical condition(s) and require constant, spontaneous feedback of their wellness and 2) people with generic, eventful medical situations that call for a quick, convenient communication system to help gather timely solutions beyond the obstacles of time, location and/or health care personnel's availability at the right time. The key contribution of this work is the ability of the system to learn contextual information regarding a particular medical condition through mere vocal input from the target user and apply that to construct a logical recommendation in no time. The results helped us conclude that not relying on a pretrained API helped and instead training the assistant from scratch to self-efficiently apply spontaneous prediction helped our work stand out from general recommendation systems that exist in this field of research today. Like every venture, the system that we came up with had to deal with some pitfalls in its functioning; since the application is trained on known conditions such as common cold symptoms, headaches, bowel system malfunction, diabetic symptoms, heart condition handling, there are possibilities that some subsets of the known symptoms may overlap for specific persons which might cause inconsistencies in the application's prediction results, since the assistant solely learns from and relies on communication with the users. This perpetual challenge faced by our feedback system has led us to want to augment our work with advance functionalities, which would effectively shape the future work in this project. Drawing from the ideas put across in [9], integrating the data collected from wearable technology equipped with sensors can help the assistant learn the physical implications of the conditions the users describe and thereby provide more refined feedbacks and alerts to the target users.

CONCLUSION AND FUTURE WORK

In this paper, we describe the details of a healthcare service application implemented as a virtual feedback system using a virtual assistant which integrates Speech to Text and Text to Speech functionalities to incorporate pervasive healthcare to some extent. We have accomplished creating the feedback system but our future research challenges include incorporating video-based conversations involving image processing capabilities for the assistant to be able to help with skin conditions, allergies, wounds, infections and other such physical conditions of target users to be able to extend the application's capability to first aid care. We also plan to include biosensors in our work to be able to not only help patients with chronic conditions but also meticulously monitor them to be able to predict certain health-related emergencies and avoid medical potential fatal medical emergencies. This work contributes to the concept of pervasive healthcare by attempting to provide health care regardless of the physical location of the user, availability/non-availability of a professional medical practitioner and somewhat bringing about cost-effectiveness as an added advantage.

REFERENCES

[1] Dena Puskin, Barbara and Stuart, [2006] "Telemedicine, Telehealth, and Health Information Technology", An ATA Issue Paper, The American Telemedicine Association, May 2006(5) (PDF) *Virtual Telemedicine Using Natural Language Processing*
[3] Bajwa, I. (2006). *A Rule Based System for Speech Language Context Understanding*. [online] Citeseerx.ist.psu.edu.

[2] Serban, I., Sordoni, A., Bengio, Y., Courville, A. and Pineau, J. (2019). *Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models*. [online] arXiv.org.
[3] Rajj, A., Johnsen, K., Dickerson, R., Lok, B. and Cohen, M. (2019). *Comparing Interpersonal Interactions with a Virtual Human to Those with a Real Human - IEEE Journals & Magazine*.
[4] Konstantas D, van Halteren A, Bults R, Wac K, Widya I, Dokovsky N, Koprnikov G, Jones V, Herzog R., "Mobile patient monitoring: the MobiHealth system.", *In Stud Health Technol Inform*. 2004; 103:307-14.
[4] Maglaveras, N, "Citizen Health System: telehealth homecare.", *Stud Health Technol Inform*. 2003; 92:117-25.
[5] Maglaveras, N.; Chouvarda, I.; Koutkias, V.G.; Gogou, G.; Lekka, I.; Goulis, D.; Avramidis, A.; Karvounis, C.; Louridas, G.; Balas, E.A; "The citizen health system (CHS): a Modular medical contact center providing quality telemedicine services", *IEEE Transactions on Information Technology in Biomedicine*, Volume 9, Issue 3, Sept. 2005 Page(s):353 – 362
[6] van Halteren A, Konstantas D, Bults R, Wac K, Dokovsky N, Koprnikov G, Jones V, Widya I., "MobiHealth: ambulant patient monitoring over next generation public wireless networks.", *Stud Health Technol Inform*. 2004; 106:107-22.
[7] P. J. Hu, P. Y. K. Chau, O. R. Liu Sheng, and K. Y. Tam, "Examining the technology acceptance model using physician acceptance of telemedicine technology," *Journal of Management Information Systems*, vol. 16, no. 2, pp. 91–112, 1999.
[8] B. R. Mitchell, J. G. Mitchell, and A. P. Disney, "User adoption issues in renal telemedicine," *Journal of Telemedicine and Telecare*, vol. 2, no. 2, pp. 81–86, 1996.
[9] R. L. Bashshur, J. H. Sanders, and G. W. Shannon, Eds., *Telemedicine: Theory and Practice*. Springfield, IL: Thomas, 1997.
[10] R. L. Bashshur, J. H. Sanders, and G. W. Shannon, Eds., *Telemedicine: Theory and Practice*. Springfield, IL: Thomas, 1997.
[11] N. Saranummi, I. Korhonen, M. van Gils and S. Kivisaari, "Barriers limiting the diffusion of ICT for proactive and pervasive health care," in *Proc. of the IX MEDICON*, Pula, Croatia, 2001.
[12] S. I. Korhonen, J. Lötjönen, M. Sola, and M. Myllymäki, "IST Vivago—an intelligent social and remote wellness monitoring system for the elderly," in *Proc. 4th Annu. IEEE EMBS Special Topic Conf. Information Technology Applications in Biomedicine (ITAB 2003)* Birmingham, U.K., Apr. 24–26, 2003, pp. 362–365.