Chatbots meet eHealth: automatizing healthcare

Flora Amato, Stefano Marrone, Vincenzo Moscato,
Gabriele Piantadosi, Antonio Picariello, and Carlo Sansone

DIETI - University of Naples Federico II,
via Claudio 21, 80125 Napoli (Italy)
{flora.amato,stefano.marrone,vincenzo.moscato,
gabriele.piantadosi,picus,carlo.sansone}@unina.it

Abstract. The aim of this work is to investigate the effectiveness of novel human-machine interaction paradigms for *eHealth* applications. In particular, we propose to replace usual human-machine interaction mechanisms with an approach that leverages a chat-bot program, opportunely designed and trained in order to act and interact with patients as a human being. Moreover, we have validated the proposed interaction paradigm in a real clinical context, where the chat-bot has been employed within a medical decision support system having the goal of providing useful recommendations concerning several disease prevention pathways. More in details, the chat-bot has been realized to help patients in choosing the most proper disease prevention pathway by asking for different information (starting from a general level up to specific pathways questions) and to support the related prevention check-up and the final diagnosis. Preliminary experiments about the effectiveness of the proposed approach are reported.

Keywords: eHealth, Big Data, Deep Learning, Watson, Decision Support System, Prevention Pathways

1 Introduction

The recent advances of technologies for data processing and analytics have radically changed the healthcare giving rise to digital healthcare solutions, promising to transform the whole healthcare process to become more efficient, less expensive and higher quality. In the context of *eHealth*, numerous flows have generated and will continue to produce a huge amount of information showing *Big Data* features from several sources such as electronic medical records (EMR) systems, mobilized health records (MHR), personal health records (PHR), mobile health care monitors, genetic sequencing and predictive analytics as well as a large array of biomedical sensors and smart devices.

One of the most interesting challenge for modern eHealth applications is to provide intelligent recommender systems, which leveraging the different kinds of available data and the related knowledge, are able to support people in making decisions in a large variety of scenarios: from doctors that have to formulate

D. Impedovo and G. Pirlo (Eds.), Workshop on Artificial Intelligence with Application in Health, Bari, Italy, November 14, 2017.

the correct diagnosis, to patients that can be periodically motivated to undergo themselves to a prevention visit.

As an example, medical image processing usually needs to elaborate huge quantities of data under tight computational, temporal and privacy constraints [1,2]. In addition, patients records have to properly be stored in order to maintain historical information of patients over the time with the aim of improving further diagnosis.

It is clear that traditional databases and knowledge management systems are not appropriate to handle such variety of data, since each single managed information usually may have a structure that differs from that of the other ones. Moreover, in a real clinical scenario, those records are produced at a fast pace requiring suitable storing technology and reactive elaboration infrastructures.

The need of such smart environments with the described features pushes to search solutions in the eHealth that adopt big data technologies.

As well known, the term eHealth indicates all the healthcare practices supported by electronic elaboration and remote communications [3], even if several definitions were so far proposed [4]. Some authors consider the eHealth as an extension of the classical health procedures to informatics and/or digital processing [5], others definitions uses the eHealth paradigm to group all the healthcare procedures delivered via the Internet [6,7] or via mobile devices [8], but the most general and recognized meaning considers under the eHealth paradigm all the services or systems laying on the edge of healthcare and information technology.

However, automatizing healthcare according to the eHealth paradigm, requires to face several problems, among which it is worth to mention:

- Data sensitivity of medical records (privacy must be guaranteed) [1,2];
- Operational time comparable with clinical constraints [9,10];
- Massive data handling [11];
- Context scalability, since some sort of distributed elaboration and modularity should be applied [1,2].

The discussed requirements led to study, design and realise systems, infrastructures and architectures based on big data analytics techniques and technologies ensuring at the same time computational efficiency, scalability, upgradeability and reduced costs.

If, from one hand, the new data processing and analysis technologies have supported the diffusion of the eHealth paradigm, from the other one, they have given rise to the new *patient trustiness* problem, referring to the feeling of interacting with a dehumanized entity when she/he comes to filling forms or answering to a given set of questions by using a predefined given set of answers. It is worth mentioning that this issue might make the patient more frustrated since she/he is not able to fully express her/his symptoms, worries and pains, as when he/she interacts with a human physician. Such problems are pushing researchers to explore and define new knowledge representation models and interaction paradigms in order to fulfil the rising need for a more suitable form of human-machine interaction.

As possible solution to this issue, we propose the chat-like conversation model, an internet-based communication paradigm that relays its foundations in the social network era. Healthcare, and in particular eHealth, can benefit from this model since it strongly requires a human-like interaction schema. A *chatbot* (also known as a talkbot or chatterbot) is an artificial entity able to autonomously hold a conversation via message exchange. With progresses in machine and deep learning, nowadays chatbot can be very reliable and able to provide automatic and adaptive human-like conversation behaviour, getting involved in several application fields including customer services or data collection.

The aim of this paper is to propose HOLMeS (*Health On-Line Medical Suggestions*), a novel eHealth recommendation system that leverages a chatbot to emulate human physician in a clinical environment, in order to overcome the mentioned limitation of biased interaction between the user and the software. The cluster-computing facility is provided by Databricks where a cluster of servers allows cluster-computing over the Spark framework. The chat-bot application is implemented by the Watson Conversation Service, designed and trained via the Bluemix platform.

As motivating example of our work, we can implement the following scenario. A clinical center offers several prevention pathways but, so far, it provides the service only when the patient books a de-visu visit. Thus, the patient has to physically reach the physician in the clinic infrastructure.

In the novel scenario, the clinic offers prevention services, through its social interface (such as Facebook, Instagram, Telegram or WhatsApp) using autonomous chatbots. It follows that a user can interact with the chatbot autonomously. The Figure 1 depicts an ideal automatic bot interaction that the medical centre could provide.

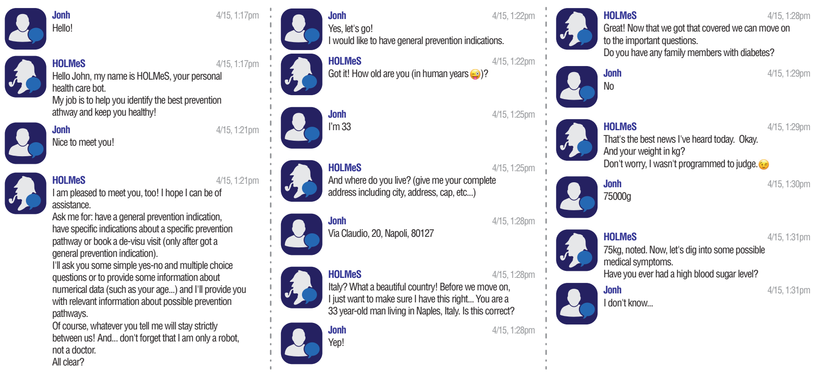

Fig. 1. An ideal interaction with a medical chatbot.

The paper is organised as following. Our proposal is presented in Section 2 where we detail each module of HOLMeS and provide an overview of the complete architecture. A specific application case has been then implemented in Section 3. Finally, the obtained implementation is discussed in the Section 4, where we also draw some conclusions and present the open issues.

2 HOLMeS: Health On-Line Medical Suggestions

In this work we present HOLMeS (Health On-Line Medical Suggestions), a medical recommendation system designed to autonomously interact with the user by understanding natural language in a chat and acting as a human physician. It is made of different modules to provide several advanced eHealth services through an intuitive chat application (Figure 2).

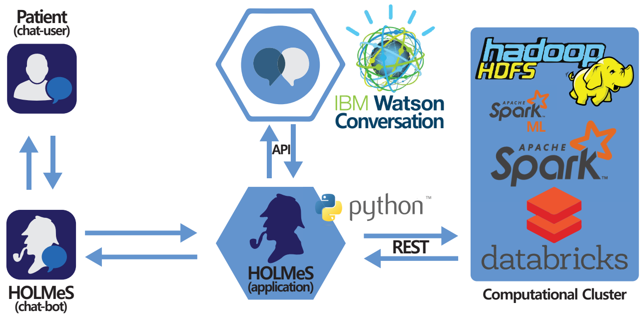

Fig. 2. HOLMeS System main modules with interaction paradigm: On bottom-centre the Application core; On the left the Chat-Bot (bottom) and the patient (top); On top-centre the IBM Watson Conversation logic; On the right the Computational Cluster providing storage, computing and machine learning skills.

The HOLMeS system is composed by the following components.

HOLMeS Application is the core of the HOLMeS system, implements the operational logic, orchestrates modules communications and functionalities. Developed in Python, it interacts with the user through the chat-bot, interpreting patient request by means of the Watson Conversation API.

HOLMeS Chat-Bot is the agent designed to make patient feel more comfortable, by interacting with it by a chat. Based on deep learning, it is designed in order to understand and to adapt to several interaction schema, ranging from formal writing to more handy ones. It is the HOLMeS System entry point and interacts with the user to let her/him choose the required service. It is also intended to kindly ask the user for required information (such as age, height, weight, smoking status, and so on) just as a human physician would behave.

IBM Watson (with its Conversation APIs) is the service used to establish a written conversation, simulating human interactions. Its main features include text mining and natural language processing by means of deep learning approaches.

Computational Cluster implements the decision making logic. It uses the Apache Spark cluster executed over the Databricks infrastructure, in order to be enough fast and scalable to be effectively used in a very big clinical scenario, ensuring response time comparable to that of a human physician. It uses machine learning algorithms from Spark ML library while the storage service is delegated to Hadoop HDFS.

2.1 The Chat-Bot Interaction Skills

HOLMeS Chat-Bot module was designed to mimic human behaviour in order to let patients feel more comfortable, overcoming the biasing of a machine interaction. It uses the IBM Watson Conversation APIs to be able to hold a complete automatic chat with the user by understanding the natural language.

Watson Conversation service identifies the user intents and the concerning entities, thanks to its ability to process structured and unstructured contents producing and elaborating up to 4 TeraBytes of data.

During a chat session between the patient and the user interface of our application, there are a broad variety of utterances that could be employed by the user aiming the same purpose.

In particular Watson Conversation service provides high-level dialogue flow design to allow writing down the entire conversation example-by-example.

An *intent* (tagged by a #) is a purpose or goal expressed in a user input, such as answering a question or requiring for a service. By recognizing the intent expressed in a user input, Watson can route into the correct dialogue flow for responding to it. An *entity* (tagged by a @) represents a class of object or a data type that is relevant to a user purpose. By recognizing the entities that are mentioned in the user input, Watson can choose the specific actions to take to fulfil an intent. The *dialog* uses the intents and entities that are identified in the user input, plus context from the application, to interact with the user and ultimately provide a useful response.

For example, the response might be the answer to a question such as: "Can you give me information about the medical centre where you work for?". In that case, the intent might be requesting for information and the entity is something regarding the clinic infrastructure. The chatbot might consider that the entity is not specific enough to return an exhaustive response, therefore, it might ask the user for more input such as "About what of our three centres want to know more? We are in Naples, Rome and Milan!". This flow of questions and answers, triggered by the intents and the entities, represents part of the dialogue design (see Figure 3).

In the next Section a specific case has been evaluated. The results are depicted in the Figure 4 reporting a real usage of the HOLMeS system in the described case study showing that the chat-bot service can successful held a full conversation with the user, also providing the final diagnosis.

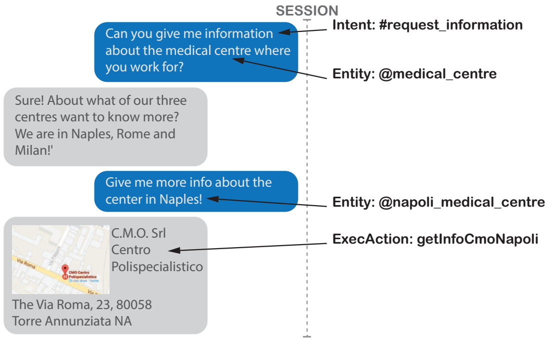

Fig. 3. Watson at work in a real medical-like chat experience.

3 A Case Study

With the aim of applying HOLMeS (Health On-Line Medical Suggestions) to a real medical recommendation case, we signed a collaboration with a the C.M.O. Srl clinical centre. It has provided us a dataset of clinical records of all the patients undergoing to disease prevention pathways with the related evaluation.

The goal, analysed jointly with the medical and managerial staff of the clinical centre, was to provide machine learning based insights for all the prevention pathways and make available them as chatbot services.

We consider splitting the prevention pathway evaluation into two main steps. In a first level interview, HOLMeS evaluates the pathway likelihood relying only on the information the patient can provide without any medical consult or further clinical investigation. Information such as age, sex, the address of residence, familiarity or historical medical conditions are questioned during the interaction with the chatbot. When all the first level investigation are accomplished, a bar plot reporting the likelihood of each prevention pathways is sent to the user. The prevention pathways that most fit the user data are, then, suggested for a further investigation. The user is informed about the risk of the pathology and the HOLMeS propose a new virtual interview by adding further medical-based information. To make the medical data available, the system proposes the user for a specialistic medical consult or to get back with the required clinical investigation.

When the user is ready to a more specific pathway evaluation, it can ask to HOLMeS for a second-level interview (specific for a given pathway), during which HOLMeS will ask about crucial information to improve the reliability of the likelihood evaluation.

3.1 Dataset

The C.M.O. centre provides disease preventive healthcare for 13 different diseases (Alzheimer, Cardiovascular diseases, Large Bowel, Diabetes, Geriatrics issues, Gynecology, Obesity, Diabetic Foot, Lung diseases, Prostate diseases, Psychological Obesity, Breast cancer, Thyroid diseases) and, at the study execution time, a total of 16733 patients prevention records have been collected. Each patient contains a positive or negative ground-truth indicating whether its prevention pathway leads to a positive diagnosis or not.

3.2 Chatbot dialogue design

For the described problem, HOLMeS was intended to be able to handle four possible use-case scenarios:

1. Providing general information about itself or the affiliated medical centre.
2. Collecting general patient information in order to provide general prevention pathways indications among different diseases.
3. Collecting detailed patient information (clinical, examination results and so on) in order to evaluate the probability of needing a specific disease prevention partway.
4. Book a *de-visu* examination with the affiliated medical centre.

In order to fulfil these four possible uses cases, Watson conversation service has been designed to recognise the following intents: *#greetings* to handle the initial conversation preamble; *#book* to describe actions as reserving a *de-visu* examination; *#get* to ask for receiving something such as prevention pathways indications or information about the centre; *#put* to catch the intent of giving the required information to the system.

Moreover, several entities, useful to contextualise the above intents, has been described. The entities are formalised by synonyms. More equivalent formulations of an entity are provided, the more precise will be the subject recognition. Some designed entities are in the following list: @HOLMeS, @HOLMeS_functionalities, @clinical_centre, @address, @de_visu_examination, @specific_patway_examination, @general_patways_examination, @age, @sex, @birthday, @height

Intents and entities are then combined to achieve a fully automated conversation flow by using the 'dialog design toolbox' of the Bluemix platform.

3.3 The Machine Learning algorithms

Even if able to understand natural language, the chatbot alone is not able to provide any kind of medical advice. We specifically design machine learning algorithms using Spark as cluster-computing framework deployed over the Databricks infrastructure. Spark provides data storage, implicit data parallelism and fault-tolerance features. The Spark.ML library provides all the data-preparation functionalities and the machine learning algorithms to train the Random Forest models using the collected training data. The same Spark.ML library is, then, used to achieve the final classification for each diseases prevention pathway.

When a general-level evaluation is required, the algorithm will produce a histogram graph containing the disease occurrence probability per each of the disease available in the dataset. This result will be stored in the database and can be retrieved by the physician when a *de-visu* examination will occur or can be used by the patient to choose the specific disease to be evaluated in the second working modality.

Using our machine learning-based approach to prevention pathway assessment we obtain 74.65% of Area Under ROC Curve (AUC) when first-level features are used to assess the occurrence of different prevention pathways. When disease specific features are added, HOLMeS shows 86.78% of AUC achieving a more specific prevention pathway evaluation.

3.4 General functionalities

As stated in the previous section, the HOLMeS core orchestrates the data flows, from and to the patient, through the chat-bot interactions. The interaction between the user and the chatbot occurs without any further elaboration by the HOLMeS application that observes the data flow, memorising the information of interest. When interaction requires elaborating machine learning algorithms or accessing to the data storage, the core system contacts the cluster for accomplishing the specific task.

In order to fulfil the four possible uses cases (as described above) core application needs to catch the following intents:

- The user wants information about the affiliated medical centre.
- The user asks for general information about the system and its functionalities.
- The patient desires to obtain general-level evaluation about the available prevention pathways.
- The patient desires to obtain detailed indications about a specific disease prevention pathway (only after a general survey has been carried out).
- The patient wants a *de-visu* examination (only after a general survey has been carried out).

For example, to meet the second use-case scenario, after the greeting preamble is completed, HOLMeS recognises the *#get* user intent combined with the *@general_patways_examination* entity. Such interaction yield to the data collecting chat flows with the aim of achieving a general-level prevention pathway evaluation. Figure 4 shows the chatbot interaction and a real result graph according to the information provided in the test case.

4 Conclusion

The aim of this paper was to introduce HOLMeS (Health On-Line Medical Suggestions), a system designed to improve the eHealth paradigm by using a chatbot to simulate human interaction in medical contexts. Based on deep learning techniques, the chatbot is able to overcome the limitation of classical human-machine interaction, thus removing bias and allowing the patient to a freer and natural

Fig. 4. Conversing with HOLMeS: Asking for general indications about the available prevention pathways.

communication. HOLMeS design allows to adapt it to different clinical scenarios and medical task since the human interaction workflow is independent from the specific task the bot was assigned to. This allows to easily extend HOLMeS capability by designing new knowledge-base and defining the proper intent and entities to make HOLMeS able to ask for required information. Moreover, as described, HOLMeS was designed to be modular, where each concern is associated with a given module. This allows designing and developing each piece separately, in order to better fit specific requirements such as scalability, speed, availability, and so on. It follows that thanks to its design and structure, HOLMeS can fulfil many of the desired characteristics described in Section 2. To demonstrate HOLMeS abilities, in Section 3 we introduced a disease prevention application that makes use of HOLMeS to collect patients' data, provide booking and general information services. Future works will focus on improving the human-like behaviours and interaction paradigm by making HOLMeS able to best adapt to the user need and characteristics (age, gender, social class, education, etc). Finally, since Watson supports different languages (among which English, French and Italian), we are planning to spread HOLMeS in different countries for helping in the medical knowledge sharing.

Acknowledgements

This work has been carried out within a research collaboration between the University of Naples and DATALIFE[1]. In particular, we would like to thank CMO[2] - one of the partner of the DATALIFE research consortium - for the availability of the dataset that allow us to perform the experimentation of such paper.

[1] www.datalife.life
[2] http://www.cmocentropolispecialistico.it/

References

1. Gabriele Piantadosi, Stefano Marrone, Mario Sansone, and Carlo Sansone. A Secure OsiriX Plug-In for Detecting Suspicious Lesions in Breast DCE-MRI. In *Algorithms and Architectures for Parallel Processing*, pages 217–224. Springer International Publishing, 2013.
2. Gabriele Piantadosi, Stefano Marrone, Mario Sansone, and Carlo Sansone. A secure, scalable and versatile multi-layer clientserver architecture for remote intelligent data processing. *Journal of Reliable Intelligent Environments*, 1(2-4):173–187, dec 2015.
3. Vincenzo Della Mea. What is e-health: The death of telemedicine? *Journal of Medical Internet Research*, 3(2):e22, 2001.
4. Hans Oh, Carlos Rizo, Murray Enkin, and Alejandro Jadad. What is ehealth?: a systematic review of published definitions. *J Med Internet Res*, 7(1):e1, 2005.
5. JC Healy. Implementing e-health in developing countries: Guidance and principles. *ICT Applications and Cyber security Division (CYB). Policies and Strategies Department.[monografía en Internet]. Bureau for Telecommunication Development International Telecommunication Union*, 2008.
6. Marion J Ball and Jennifer Lillis. E-health: transforming the physician/patient relationship. *International journal of medical informatics*, 61(1):1–10, 2001.
7. Gunther Eysenbach and Thomas L Diepgen. The role of e-health and consumer health informatics for evidence-based patient choice in the 21st century. *Clinics in dermatology*, 19(1):11–17, 2001.
8. John Odonoghue and John Herbert. Data management within mhealth environments: Patient sensors, mobile devices, and databases. *Journal of Data and Information Quality (JDIQ)*, 4(1):5, 2012.
9. Stefano Marrone, Gabriele Piantadosi, Roberta Fusco, Antonella Petrillo, Mario Sansone, and Carlo Sansone. A Novel Model-Based Measure for Quality Evaluation of Image Registration Techniques in DCE-MRI. In *2014 IEEE 27th International Symposium on Computer-Based Medical Systems*, pages 209–214. IEEE, IEEE, may 2014.
10. Gabriele Piantadosi, Stefano Marrone, Roberta Fusco, Antonella Petrillo, Mario Sansone, and Carlo Sansone. Data-driven selection of motion correction techniques in breast DCE-MRI. In *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pages 273–278, Torino, Italy, may 2015. IEEE.
11. Stefano Marrone, Gabriele Piantadosi, Roberta Fusco, Antonella Petrillo, Mario Sansone, and Carlo Sansone. Breast segmentation using fuzzy c-means and anatomical priors in dce-mri. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pages 1472–1477. IEEE, 2016.