CREATING SMART HEALTH SERVICES USING NLP TECHNIQUES

Aldina Avdić¹*,
Ulfeta Marovac¹,
Dragan Janković²

¹ Department for Technical Sciences,
State University of Novi Pazar,
Novi Pazar, Serbia
² Faculty of Electronic Engineering -
University of Nis,
Nis, Serbia

Abstract:
The rapid growth of the urban population and the development of information
and communication technologies have led to the creation of a new concept - a
smart city. The smart city uses information and communication technologies to
improve the quality of life of its citizens. One of its most important components
is smart health. Also, heaps of textual data are created every day, in medical
information systems, then there are the government documents, and citizens’
comments, reviews, etc. This paper examines the possibility of using text min-
ing techniques and natural language processing to create smart health services.
Three services intended for these purposes have been proposed, designed, and
implemented: a service for answering patients’ questions, an information board
for visualizing data about an epidemic, and automatic processing of question-
naires and psychological tests.

Keywords:
Smart Health Services, Natural Language Processing, Text Mining.

INTRODUCTION

The constant progress of information and communication technologies
(ICT) has enabled their application in almost all segments of individuals’
lives to provide many services which make life easier for citizens. Modern
technological infrastructure (Internet, smartphones with built-in sensors,
smart devices in the household) is available today to most of the popula-
tion. Techniques for processing large amounts of data, collective intel-
ligence, and increasing the number of inhabitants in urban areas enabled
the creation of a smart city, which is efficient in terms of energy con-
sumption, transport, administration, learning, economic development,
and other aspects [1]. It is based on Big Data [2] and IoT technologies
[3]. Within the smart city, there are smart transport, smart agriculture,
smart administration, smart education, smart healthcare, etc. Smart
health uses ICT technologies to create services which contribute to the
facilitation of medical procedures and better health of the population [4].
The difference between e-health and smart health is in the flow of data. For example, creating electronic reports using the medical information system is an example of e-health, but enabling communication with doctors, and dashboard with analysis based on data in medical information systems (where the data flow is from citizens to smart city, and from smart city to citizens), are examples of using smart health.

Large amounts of data are collected daily, in medical information systems, then there are documents related to the medical domain (instructions for medicines, various forums containing instructions of experts, etc.).

Some information from patients can be collected using crowdsourcing (the technique of acquiring knowledge through the voluntary participation of users) [5]. EHRs [6] are electronic medical records created and stored in medical information systems and contain information about the patient’s examination. They contain structural fields, such as name, identification number, date of examination, location of the institution where the examination was performed, diagnosis code, name of the diagnosis, but also contain a non-structural part consisting of physician observations that cannot be expressed through structural fields, physician notes, diagnoses, laboratory results, therapies, etc. The documents also contain textual data, as the free-answer questions in the questionnaires for patients. Unstructured data must be organized and structured in a way that allows their analysis. This requires sophisticated statistical and linguistic techniques.

The research question of this paper is whether and how NLP techniques can be used to create smart health services. The motivation for this research is the creation of smart health services based on NLP which would be useful for patients and citizens in a smart city. For the realization of this research, the method of description was applied to the existing smart health services based on NLP techniques. Three services are proposed and using the modeling method and use case diagrams are shown. Details of the implementation and a screenshot of the demo systems that implement these services are given.

The paper is organized as follows. The second section presents the application of NLP techniques for medical purposes through related research. This is followed by a proposal and a description of NPL-based health services, and the implementation of the service is given. In the end, the conclusion and directions of further research are given.

2. RELATED WORK

This section presents research and examples of the use of tools based on NLP techniques for improving health. One of the ways to use the processing of textual data entered by users is the detection of depression [8].

The most popular medical text marking systems are CTAKEs and CLAMP systems [9] [10]. The identification of medical terms in texts written by patients, using crowdsourcing, is presented in [11]. One approach to correcting errors in the free text of medical reports is presented [12].

3. SMART SERVICES BASED ON NLP TECHNIQUES

This section describes services based on NLP. The source of data for the implementation of these services are documents written in natural languages, citizens’ answers obtained by crowdsourcing and EHRs from medical information systems. The service whose input data are documents is used to answer patients’ questions. The service whose input data are the answers of citizens obtained by filling out questionnaires is the automatic processing of psychological tests. The service that processes the data entered in the EHRs displays data on the state of the epidemic. Use case diagrams, a description of the services, and finally presentations of their implementation are given.

3.1. ANSWERING THE PATIENTS’ QUESTIONS

The smart health service for automatic answering of questions and search of medical documents should provide the user to enter a question and get an answer based on the similarity with the sentences in medical documents entered in the system by experts. The capabilities of this service are shown by the use case diagram used in Figure 1.
Figure 1. Use case diagram for the service for the automatic answering of citizens’ questions

This service should help the citizen to get an answer to his question quickly at any time. There are three types of content in the system: questions, formal documents, and expert answers. Questions and documents are grouped based on keywords. When the user asks a new question, its similarity with the questions and documents in the selected group is considered. Preparation of documents for processing, normalization, extraction of keywords, the grouping of the documents and questions as well as finding the answer to the question are NLP techniques necessary for the realization of this service [13].

3.2. AUTOMATIC PROCESSING OF QUESTIONARIES

Psychological tests often limit the respondent by offering him the possibility of choosing some of the offered options without entering a free answer because it is complicated for analysis. For the realization of the service for automatic processing of questionnaires and psychological tests, it is necessary to collect information about the mental state, as well as about specific reactions to appropriate stressful situations.

The questions are expressed in the form of standardized psychological tests as well as fields for describing a stressful situation. The frequency of similar stressful situations, as well as evoked feelings in comparison with the results obtained from psychological tests, can more clearly indicate the existence or non-existence of a corresponding psychological problem. A person best describes his mental state in his mother tongue, so the analysis of these tests is adapted to people who speak Serbian. The complexity of processing texts in the Serbian language is an additional challenge. This service can be applied to various types of questionnaires. As an example, a psychological test is given intended to provide support to women during pregnancy and the postpartum period who may have some specific psychological problems. The use case diagram of the service is given in Figure 3 and its realization in Figure 4. The text data that users enter is in a format that is not suitable for direct processing and needs to be processed using NLP techniques to be machine-readable. Descriptions of stressful situations are compared with those in the database, to obtain the results of the questionnaire. To compare descriptions as word vectors, their pre-processing and normalization are required [5].
3.3. DASHBOARDS WITH INFORMATION ABOUT AN EPIDEMIC

At the time of an epidemic, people need to have information about the situation in their place from verified sources, to act following it. The lack of such information can lead to the spread of news from non-objective sources, unscrupulous behavior of citizens, but also the spread of panic. Epidemic control is one of the topics in the field of smart health within smart cities. Based on the reports that are stored daily in medical information systems, various information on the epidemiological situation in cities can be extracted. Examples of realization are given based on EHRs from Health Centre Nis, from the period of the measles epidemic. The service is designed to work in real-time, directly takes data from EHRs, and provides many details, such as the prevalence of the virus by age structure, health facilities, and an overview of the number of infected in the appropriate time interval [15] [16] [17]. This service enables two functionalities based on NLP techniques (Figure 5):

- **a)** Report on the most common symptoms. Using this report, the citizen would be more informed about the disease and could recognize the symptoms and in that case, contact a doctor. This report is created based on data obtained by entering in an unstructured field for physician comments from EHRs. With adequate processing of this text, frequently occurred symptoms can be extracted and shown on the diagram (Figure 6).

- **b)** Questionnaire on the patient’s current health condition (symptom presence/absence) to check whether to report to the medical station for treatment or not.
4. CONCLUSION

A lot of textual data is created daily in medical information systems, and NLP techniques can be applied to gain new knowledge that will be used to create smart health services. Three use cases of NLP-based smart health services for each input type of text data (document, EHRs, and patient’s responses) are presented. In further research, we will improve these and create new services that use NLP methods, such as finding and correcting errors when entering a non-structural part in the EHRs.

5. ACKNOWLEDGEMENTS

This paper is partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under projects III44007 and ON 174026.

REFERENCES


