**Abstract**

The use of natural language processing (NLP) methods and their application to developing conversational systems for health diagnosis increases patients’ access to medical knowledge the proposed system is an artificial intelligent that aids the doctor in diagnosing diseases through listening to conversation between doctor and patient in order to specify the accurate diagnose and estimate the proper treatment with the help of the medical history of the patient to avoid any conflict in medication that may cause harm.

**Introduction**

Remote diagnosis systems are becoming increasingly popular and accurate, with enormous advantages such as cost-effectiveness, fast and reliable decision support for medical diagnostics and treatment and prevention of disease, illness, injury, and other physical and mental damages in human beings. The rise in remote health services (or telehealth) offered by healthcare institutions coincided with the evolution of assisted living systems and environments, aiming to widen the possibility for older and disadvantaged people to access appropriate healthcare services and thus improve their health status and clinical outcome. With the increase in the innovation of medical technologies, there is a need to adopt medical expert systems that will oversee and control diagnosis and treatment processes Medical diagnostic processes carried out with the aid of computer-related technology which is on the rise daily have improved the experience and capabilities of physicians to make an effective diagnosis of diseases while employing novel signal processing techniques for analysis of patients’ physiological data and deep neural networks for decision support. With the rise of the artificial intelligence (AI) techniques, the intelligent doctor-assistant have appeared as a promising direction to emphasize the accuracy of diagnosing diseases & do deep analysis on the health records and symptoms of the patient to provide reliable health state of patients

The natural language processing (NLP) & technology can serve as an interaction between computers and humans using linguistic analysis and deep learning methods to obtain knowledge from an unstructured free text. The NLP systems have shown their uniqueness and importance in the areas of information retrieval mostly in the retrieval and processing of large amount of unstructured clinical records and return structured information by user-defined queries. In general, the NLP system is aimed at representing explicitly the knowledge that is expressed by the text written in a natural language. There are few applications of the NLP techniques in diagnosing diseases despite the enormous amount of text-based information, which can be retrieved from patients’ self-narrations The main challenges addressed by the application of NLP for medical records are flexible formatting, structure without sentences, missing expected words and punctuation, unusual parts of speech (POS), medical jargon, and misspellings Linguistic structures such as coreferences make medical texts difficult to be interpreted Moreover, unique linguistic entities such as medical abbreviations make the inference of knowledge from medical texts much harder.

Applying Machine learning algorithms, especially SVM, have shown promising results in classifying free text & make a decision to assist medical experts in their diagnosis would serve as a boost in successfully improving healthcare services through effective analysis of narrative text of symptoms provided by a patient.

The proposed system serve as a doctor-assistant during the conversation with the patient by combining the NLP and machine learning algorithms to aid the doctor in the diagnosing and treatment process.

* **Risks**

|  |  |
| --- | --- |
| Risk | Effect |
| * Patient may not know how to use application * Patient may speak unclearly * Patient may describe the disease in the wrong way * Patient may enter incorrect data | Which may lead to incorrect diagnosis |
| User does not have internet | Which leads to the program not working |
| Patient speaks in an unsupported language | Which leads to the failure of the program |
| Patient describes an untrained disease for the program | Program can’t detect the disease |
|  |  |
|  |  |

* **Constraints**
  + The user should have the application on his device
  + The user must have Internet on his device
  + Internet must always run when using application
  + The user should speak in a clear voice
  + The user should talk about symptoms in a clear and direct way
  + The user should know how to deal with the application

***related work:***

A large portion of the world's population does not have access to adequate healthcare. We approach this problem from two angles: prevention and treatment. Prevention by monitoring; Diagnosis as part of treatment. The monitoring part is still in the idea stage, while we have made some progress on the diagnostic aspect. While we present our view of machines that aid in both monitoring and diagnosis, the focus of this paper is more on the diagnostic part.

The ultimate goal of the Diagnostics Project is to put medical diagnostics in the hands of millions of underprivileged people without reaching a doctor, which is made possible thanks to the proliferation of portable smart devices including wearable devices that can instantly connect to the enormous computing power in the cloud. The project aligns with seeing the web as the ubiquitous computer. A Guide to the Concept of Embedded Ideas has also been published, using Semantic Web Concepts for a Restaurant Find app previously. This project expands these ideas and uses machine learning / text mining techniques to solve a critical problem affecting a large segment of the population - those related to medical diagnosis and broadening that insight to include health monitoring. The paper will present some preliminary findings in this direction and conclude with future directions.

Innovative medical technologies are developing day by day, as there is an important need for integrated medical expert systems (ESs) that will help to effectively manage and control diagnosis and treatment processes. These systems, with new approaches, have improved the experiences and capabilities of physicians to make the diagnosis of diseases. In this work, an integrated medical ES called Expert Doctor Verdis (Ex-Dr Verdis) is developed, which combines an advanced medical information system containing various medical services supported by information technologies, with ES capabilities in a single system. This system is also one kind of decision support system. Implementation of this system is applied for vertebral column diseases. Ex-Dr Verdis is a strong decision support tool with 94% sensitivity, 71% specificity, 87% positive, and 86% negative predictive values for the diagnosis of vertebral diseases. In addition to its facilities of medical information, Ex-Dr Verdis, with a sharing platform, provides physicians with the opportunity to share and discuss their own patients, cases, experiences, and expert knowledge with other colleagues. This integrated medical ES can be used in all hospital services, such as hematology, neurology, or cardiology, by adding new expert modules for other diseases.

Information Extraction (IE) refers to the automatic extraction of concepts, entities and events

as well as their relations and associated attributes from free text. A recent review of clinical IE applications notes the increasing interest to NLP but lists only 25 IE systems

which were used multiple times, outside the labs where they were created. Isolated attempts exist

to apply IE in the context of EHR processing in frameworks for semantic search, for instance

SemEHR deployed to identify contextualized mentions of biomedical concepts within

EHRs in a number of UK hospitals .We mention the following research prototypes

as experimental developments, based on some sort of IE: reports about

a system extracting textual medical knowledge from heterogeneous sources in order to integrate

it into knowledge graphs; describes a machine learning system

that annotates radiology reports and extracts concepts according to a model covering most clinically

significant contents in radiology; (Jackson et al., 2018) presents the information extraction

and retrieval architecture CogStack, deployed in the King’s College Hospital. CogStack has functionality

to transform records into de-identified text documents and applies generic clinical IE

pipelines to derive additional structured data from free texts. Most of the successful systems listed above work for clinical narratives in English. All major resources, ontologies and terminology classifications like UMLS4 and MESH5 are available in English. The comprehensive ontology SNOMED CT6 was developed initially in English and then translated to other languages.

**project plan**

|  |  |  |
| --- | --- | --- |
| Activity | Proceeding activity | Duration |
| Converting voice of Conversation to Text  (A) | None |  |
| Proposing Natural Language Processing mode to  Make text preprocessing  (B) | None |  |
| Testing Natural Language Processing mode to  Make text preprocessing  (C) | B |  |
| Extracting Symptoms from Text with NLP Model  (D) | A, B, C |  |
| proposing machine learning model classification illness  (E) | None |  |
| Testing machine learning model classification illness  (F) | E |  |
| proposing machine learning model to prediction Appropriate treatment  (G) | None |  |
| Testing machine learning model to prediction Appropriate Treatment  (H) | G |  |
| BACK END AND HOSTING  (I) | A, B, C, D, E, F, G, H |  |
| proposing android app for doctor  (J) | A, B, C, D, E, F, G, H, I |  |
| Connect doctor app with HOSTING  (K) | A, B, C, D, E, F, G, H, I, J |  |
| proposing android app for patient for record Previous illnesses  (L) | None |  |
| Testing apps and make optimizations  (M) | A, B, C, D, E, F, G, H, I, J, K, L |  |