Logo

Description automatically generated

**Ain Shams University**  
**Faculty of Computer & Information Sciences**

**Computer Science Department**

Patient Triage

**July 2022**

Logo

Description automatically generated

**Ain Shams University**  
**Faculty of Computer & Information Sciences**

**Computer Science Department**

**By:**

Beshoy Hani [Computer Science]

Isaac Moris [Computer Science]

Mariv Ibrahim [Computer Science]

Youssef Nagy [Computer Science]

Patient Triage

*.*

**Under Supervision of:**

[Ayat Mohammed Naguib]

[Doctor],

Scientific Computing Department,

Faculty of Computer and Information Sciences,

Ain Shams University.

[Menna Allah]

[Teacher Assistant],

Computer Science Department,

Faculty of Computer and Information Sciences,

Ain Shams University.

# Acknowledgement

In successfully completing this project, many people have helped us. We would like to thank all those who are related to this project.

Primarily, we would thank God for being able to complete this project with success. Then We will thank Our supervisor Dr/ Ayat Mohammed Naguib, under whose guidance We learned a lot about this project. Her suggestions and directions have helped in the completion of this project.

Finally, we would like to thank our parents and friends who have helped us with their valuable suggestions and guidance and have been very helpful in various stages of project completion.

Special thanks for doctors and the medical staff at Ain-Shams University of Medicine and The Virtual Hospital staff.

# Abstract

The cultural transformation called digital health has been shaping the fundamental basics of healthcare since the beginning of the 21st century.

The Patient Triage system is used to accurately triage the patient to the clinic that matches their symptoms. It allows the use of English and Arabic speaker in a wide range of dialects.

Few platforms support building Arabic conversational agents for Arabic speakers. This creates barriers for none-bilingual speakers to benefit from such systems.

The system completely runs by voice. This allows even an illiterate person with a little or no computer knowledge to use the system.

It doesn't replace the physicians, instead provides accurate triaging of the patients to the clinics which supports and assists physicians with accuracy reaches 88%.

We will discuss the design, development, and testing of our system, and highlight the findings and limitations arose from the testing.

**Table of Contents**

[Acknowledgement i](#_Toc417741593)

[Abstract ii](#_Toc417741594)

[List of Figures iv](#_Toc417741595)

[List of Abbreviations v](#_Toc417741596)

[1- Introduction 1](#_Toc417741597)

[1.1 Motivation 1](#_Toc417741598)

[1.2 Problem Definition 1](#_Toc417741599)

[1.3 Objective 1](#_Toc417741600)

[1.4 Time Plan 1](#_Toc417741601)

[1.5 Document Organization 1](#_Toc417741602)

[2- Background 2](#_Toc417741603)

[3- Analysis and Design 3](#_Toc417741604)

[3.1 System Overview 3](#_Toc417741605)

[3.1.1 System Architecture 3](#_Toc417741606)

[3.1.2 System Users 3](#_Toc417741607)

[3.2 System Analysis & Design 3](#_Toc417741608)

[3.2.1 Use Case Diagram 3](#_Toc417741609)

[3.2.2 Class Diagram 3](#_Toc417741610)

[3.2.3 Sequence Diagram 3](#_Toc417741611)

[3.2.4 Database Diagram 4](#_Toc417741612)

[4- Implementation and Testing 5](#_Toc417741613)

[5- User Manual 6](#_Toc417741614)

[6- Conclusion and Future Work 7](#_Toc417741615)

[6.1 Conclusion 7](#_Toc417741616)

[6.2 Future Work 7](#_Toc417741617)

[References 8](#_Toc417741618)

# List of Figures

[Figure 1 Time Plan 2](#_Toc107522948)

[Figure 2 Similar Systems Comparison 6](#_Toc107522949)

[Figure 3 System Architecture 8](#_Toc107522950)

[Figure 4 Use Case Diagram 10](#_Toc107522951)

[Figure 5 Sequence Diagram 11](#_Toc107522952)

[Figure 6 Wireframe Welcome Screen 15](https://cisasuedu-my.sharepoint.com/personal/isaac2018170072_cis_asu_edu_eg/Documents/DocumentationTemplateWithCoverPage.docx" \l "_Toc107522953)

[Figure 7 Wireframe Start Screen 15](https://cisasuedu-my.sharepoint.com/personal/isaac2018170072_cis_asu_edu_eg/Documents/DocumentationTemplateWithCoverPage.docx" \l "_Toc107522954)

[Figure 8 Wireframe Information and symptom Screen 16](https://cisasuedu-my.sharepoint.com/personal/isaac2018170072_cis_asu_edu_eg/Documents/DocumentationTemplateWithCoverPage.docx" \l "_Toc107522955)

[Figure Welcome Screen 17](https://cisasuedu-my.sharepoint.com/personal/isaac2018170072_cis_asu_edu_eg/Documents/DocumentationTemplateWithCoverPage.docx" \l "_Toc107522956)

[Figure Start Screen 18](https://cisasuedu-my.sharepoint.com/personal/isaac2018170072_cis_asu_edu_eg/Documents/DocumentationTemplateWithCoverPage.docx" \l "_Toc107522957)

[Figure Information and Symptoms Screen 18](https://cisasuedu-my.sharepoint.com/personal/isaac2018170072_cis_asu_edu_eg/Documents/DocumentationTemplateWithCoverPage.docx" \l "_Toc107522958)

# List of Abbreviations

|  |  |
| --- | --- |
| Abbreviation | What the abbreviation stands for |
| API | Application Programming Interface |
| APP | Application |
| EMR | Electronic Medical Record |
| GUI | Graphical User Interface |
| IT | Information Technology |
| STT | Speech to Text |
| TTS | Text to Speech |

# Introduction

## **1.1 Motivation**

Improving health makes life better and waiting lists are increasing dramatically in clinics and hospitals. AI-powered virtual doctor assistants can help in reducing queues and waiting lists in hospitals.

Focuses on people whose native language is Arabic which is not so common in the Artificial Intelligence industry compared to English.

## **1.2 Problem Definition**

The wrong patient triaging leads to a waste of time for both, the patient and the doctor, and unnecessary paid costs and increase the number of queues.

For example, many people suffer might suffer from mouth problems and go to a dental clinic then after the dentist examination and diagnosing, they discover it was a problem with stomach inflammation and that they need to reserve and go to another clinic.

## **1.3 Objective**

* Applying Knowledge to field of Medical Science and making the task of Physician easier our main purpose
* Help in getting immediate and proper medication.
* Help users to maintain a healthy lifestyle.
* Reducing infections caused by patients going to clinics.
* Enrichment of the Arabic language in technological and scientific communities.

## **1.4 Time Plan**

Chart, bar chart

Description automatically generated

Figure 1 Time Plan

## **1.5 Document Organization**

Chapter 2: Background:

This chapter includes detailed description of Digital health, scientific background, similar existing systems and the Technologies used in project.

Chapter 3: Analysis and design:

This chapter includes an overview over the whole system represented in figures and diagrams related to the project architecture and use cases and what the project exactly targets.

Chapter 4: Implementation and Testing:

This chapter includes all the main functions and how it operates inside the application, all the algorithms used and how we tested our application to make sure it works properly.

Chapter 5: User Manuel:

This chapter includes the Installation Guide alongside how to operate and run the program.

Chapter 6: Conclusion and Future Work:

This chapter shows all what we concluded through our journey and what can be further done to make our application more accurate and effective.

# Background

The broad scope of digital health includes categories such as mobile health (mHealth), health information technology (IT), wearable devices, telehealth and telemedicine, and personalized medicine.

From mobile medical apps and software that support the clinical decisions doctors make every day to artificial intelligence and machine learning, digital technology has been driving a revolution in health care. Digital health tools have the vast potential to improve our ability to accurately diagnose and treat disease and to enhance the delivery of health care for the individual.

Digital health technologies use computing platforms, connectivity, software, and sensors for health care and related uses. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device.

Digital tools are giving providers a more holistic view of patient health through access to data and giving patients more control over their health. Digital health offers real opportunities to improve medical outcomes and enhance efficiency.

These technologies can empower consumers to make better-informed decisions about their own health and provide new options for facilitating prevention, early diagnosis of life-threatening diseases, and management of chronic conditions outside of traditional health care settings. Providers and other stakeholders are using digital health technologies in their efforts to:

* Reduce inefficiencies,
* Improve access,
* Reduce costs,
* Increase quality, and
* Make medicine more personalized for patients.

Patients and consumers can use digital health technologies to better manage and track their health and wellness-related activities.

The use of technologies, such as smart phones, social networks, and internet applications, is not only changing the way we communicate, but also providing innovative ways for us to monitor our health and well-being and giving us greater access to information. Together, these advancements are leading to a convergence of people, information, technology, and connectivity to improve health care and health outcomes.

Scientific Background:

The problem arose when people started to make wrong decisions and triaging to themselves and others and it wasted the patient and the doctor time and effort which might leads to a lot of impact in other people lives.

A lot of solutions were applied to solve such problems like going to a family doctor or someone who’s specialized in knowing to which clinic should someone go into in order to get the recovery needed.

From applications to other solutions were also applied like chatbots but they were all limited to a certain language or by a limited database and usually these applications are private to certain hospitals and sectors.

Similar Systems:

Ada Health:

It provides a personal health companion app that uses AI and machine learning to help people to understand and manage their health. Developed by doctors and scientists, the app knows thousands of symptoms and conditions, from a common cold to rare diseases.

The Company builds a detailed picture of a person's personal health over time. At its core lies a medical reasoning engine that also supports doctors by providing earlier health information and informing clinical decision making.

Babylon Health:

It is a fully integrated mobile health service providing access to doctors. With the Babylon app, the user can talk to a GP via phone or video call, ask medical questions via text service and monitor health with a tracking system.

Amwell:

Previously known as American Well is a company providing a telehealth application that connects and enables providers, insurers, patients, and innovators to deliver access to healthcare. It offers health system modules and health plan programs in urgent care, pediatrics, therapy, menopause nutrition and counseling, population health management, telepsychiatry, pregnancy and postpartum nutrition, pregnancy and postpartum care, breastfeeding support, menopause care, and chronic disease management services. The company also provides telemedicine equipment, including telemedicine carts, peripherals, and kiosks.

Enlitic:

uses recent advances in machine learning to make medical diagnostics faster, more accurate, and more accessible. The Company's mission is to provide the tools that allow physicians to fully utilize the vast stores of medical data collected today, regardless of what form they are in - medical images, doctors' notes, structured lab tests, and so forth. The Company is working with the top research hospitals and medical device manufacturers to make this happen - they are providing the data, Enlitic is providing the process and algorithms.

HealthTap:

A company developing an online telemedicine platform which provides virtual care with immediate access to doctors. It offers a freemium service enabling users to confidentially ask a doctor, assess symptoms with AI-driven checker, get personalized recommendations and treatment plan, browse doctor-answered questions, and access a health profile integrated with health records

A picture containing calendar

Description automatically generatedComparison Between them:

Figure 2 Similar Systems Comparison

**Technologies used in the project:**

**Speech-to-Text: Automatic Speech Recognition | Google Cloud**

Accurately convert voice to text in over 125 languages and variants by applying Google's powerful machine learning models with an easy-to-use API.

**Translation API**

Translation API Basic uses Google’s neural machine translation technology to instantly translate texts into more than one hundred languages. Translation API Advanced offers the same fast, dynamic results you get with Basic and additional customization features. Customization matters for domain- and context-specific terms or phrases, and formatted document translation.

**PyCharm**

An integrated development environment used in computer programming, specifically for the Python language.

**TensorFlow**

Free and open-source software library for dataflow and differentiable programming across a range of tasks.

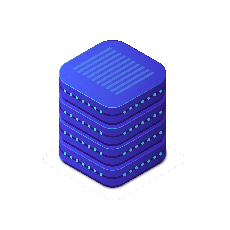
**Python**

An interpreted, high-level, general purpose programming language.

# Analysis and Design

## **3.1 System Overview**

### **3.1.1 System Architecture**

Graphical user interface, timeline

Description automatically generated

Electronic

Medical

Record

Figure 3 System Architecture

1. **Google Cloud Speech-To-Text API**
   * Takes user symptoms audio file to google to process and transform it into text (string).
2. **Data Pre-processing** 
   * Takes the text and starts processing and analyses it by generalizing similar words using stemming and lemmatization steps, removing stop-words then Extracting symptoms.
3. **Database**
   * The dataset feds the model. It grows through the data collected from doctors and patients.
4. **Machine Learning Model**
   * Learns about the symptoms and mapped clinics in order to map the given symptoms using decision Tree algorithms.
5. **Electronic Medical Record**
   * To save user information, symptoms, and mapped clinic. It allows clinicians to see a larger number of patients through better access to comprehensive patient histories that include clinical data

### **3.1.2 System Users**

1. ***Intended Users:***

Patients and anyone who have a symptoms ad wants to get a triaging on which clinic to go to.

1. ***User Characteristics***

The users do not need much experience as the application can be used by voice. So even illiterate with a little computer knowledge will be able to use it.

## **3.2 System Analysis & Design**

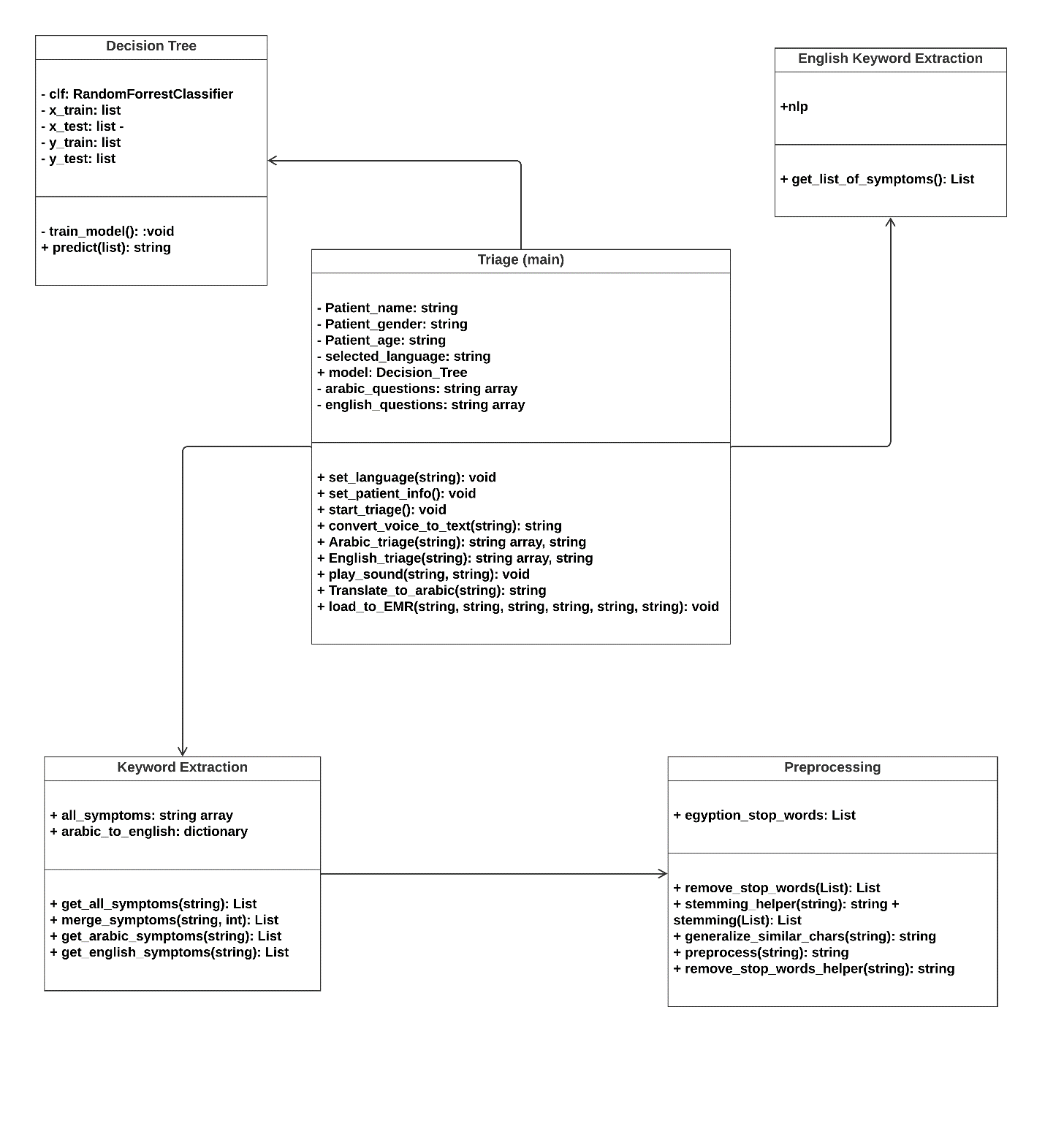
### **3.2.1 Use Case Diagram**

Diagram, schematic

Description automatically generated

Figure 4 Use Case Diagram

### **3.2.2 Class Diagram**



**Description of The Diagram:**

* + **Triage (main) Class:** The main class that is responsible for triaging the patient using the functions provided by the rest of classes.
  + **Decision Tree Class:** The model which is responsible for predicting the appropriate clinic based on symptoms entered by the patient.
  + **Keyword Extraction Class:** Extracts the keywords and symptoms from the Arabic text in order to send them to the model to predict the clinic.
  + **English Keyboard Extraction:** Extracts keywords and symptoms from the English text to send them to the model.
  + **Preprocessing Class:** Helper Class that perform NLP operations to help extracting keywords.

### **3.2.3 Sequence Diagram**

Diagram

Description automatically generated

Figure 5 Sequence Diagram

# Implementation and Testing

## **4.1 Description of the Functions of the system:**

### 4.1.1 Preprocessing Class:

**remove\_stop\_words(List of Arabic phrases):** It takes list of Arabic phrases and removes any stop word from each of the phrases (using predefined list of Arabic stop words) and returns the filtered list.

**generalize\_similar\_chars(word):** It takes one Arabic word and generalize similar characters in it for example: “ة” gets transformed to “ه” and returns the modified word.

**stemming (List of Arabic phrases):** It takes list of Arabic phrases. And stems each word in each phrase using ISRIS stemmer and returns the stemmed list of phrases.

**preprocess (List of Arabic phrases):**  It takes list of Arabic phrases. And removes stop words, stems the phrases and generalize the characters and return a list of preprocessed phrases.

### 4.1.2 KeyWordsExtraction Class :

**get\_all\_symptoms(csv file path):** It takes a csv file path and loads all the English symptom names in it with all of its Arabic translations and puts it in a map in the format of map[Arabic translation]=English symptom.

**merge\_symptoms(words,** **maximum length)** It takes an Arabic phrase and maximum length and returns a list of all continuous phrases of length less than or equal the maximum length.

**get\_arabic\_symptoms(text):** It takes an Arabic text and returns a list of the Arabic symptoms extracted from the text (fuzzywuzzy is used to get the similarity between the text and the symptoms).

**get\_english\_symptoms(text):**  It takes an Arabic text and returns a list of the Arabic symptoms extracted from the text then returns a list of the English translation of the Arabic symptoms using a predefined map.

### 4.1.3 KeywordsEnglishExtraction Class:

**\_init\_\_ ():** it is used to load spaCy large dictionary and the english symptoms

**get\_list\_of\_symptoms (text):** It takes an English text and returns a list of the English symptoms extracted from the text (spaCy is used to get the similarity between the text and the symptoms).

### 4.1.4 Decision\_Tree Class :

**Decision\_Tree ():** Define Decision Tree Classifier with hyperparameters and train the model.

**Random\_Forest():** Define Random Forest Classifier with hyperparameters and train the model.

**GBC():** Define Gradient Boosting Classifier with hyperparameters and train the model.

### 4.1.5 main Class :

**\_\_init\_\_ ():** Sets the defaults and load the data.

**set\_patient\_info ():** Get the patient gender, name and age.

**start\_triage ():** Get the voice convert it to text get the symptoms and get the clinic name from the pretrained model and convert the clinic name to speech (the function works on both Arabic and English).

**convert\_voice\_to\_text (language):** It takes the language and convert the speech to text using speech\_recognition Library.

## **4.2 Algorithms Implemented:**

### **4.2.1 Symptom Extraction steps:**

1. Take the voice form the patient and convert it to text.
2. Extract all substrings from the text that contain one, two or three words.
3. For every substring, find the symptom with the highest similarity using (fuzzywuzzy library).
4. IF this similarity > threshold then take that symptom.

## **4.3 Testing:** split the dataset to 60% training and 40% testing

## **4.4 UI Design and Wireframes**

Graphical user interface, application

Description automatically generatedThe UI Design goes as following:

Figure 6 Wireframe Welcome Screen

Welcome screen consists of:

1. 2 radio buttons to choose the language from Arabic or English.
2. Button to go to the next page.

1

11111

Graphical user interface, application, icon

Description automatically generated

Figure 7 Wireframe Start Screen

2

Start Screen which consists of:

a start button for when the user is ready to start using the application.

Chart, scatter chart

Description automatically generated

Figure 8 Wireframe Information and symptom Screen

Information and symptom Screen:

Has one start button when hit the virtual doctor starts talking to the user in the language he chose in the welcome screen.

the user is asked to answer some personal information and currently it needs his/her Name, Age and Gender to save it in the application (EMR) Electronic Medical Record dataset for later use.

Then the virtual doctor asks for the symptoms he has. After the user answers the Virtual doctor gives a triaging for the patient to a clinic that the patient has to go to.

# User Manual

**Installation Guide:**

1. Download The project from Github Link:

<https://github.com/BeshoyHani/GP> or from the flash drive provided with the documentation.

1. Download and install pycharm or any python IDE and open the project from it.
2. Install all the project dependencies and requirements through inside “Requirements.txt” file.
3. Run The program and follow the instructions.

**Graphical user interface, application

Description automatically generatedUser Guide:**

Welcome Screen:

The first screen in our application gives the user

The choice to choose his preferred language and currently our application supports only two languages as shown in figure ‘4’ and they are Arabic and English Language.

Figure Welcome Screen

Graphical user interface, application, icon

Description automatically generatedStart Screen:

The second screen waits for the user

to be fully ready to start his journey with our application.

Figure Start Screen

Information and Symptoms Screen:

Chart, scatter chart

Description automatically generated

In the following figure, it shown that the application started listening to the user.

At first the user is asked to answer some personal information and currently it needs his/her Name, Age and Gender to save it in the application (EMR) Electronic Medical Record dataset for later use.

Then the application asks the patient about the symptoms felt.

After that the application starts triaging and tell the patient to which clinic to go to.

Figure Information and Symptoms Screen

# Conclusion and Future Work

## **6.1 Conclusion**

Patient Triaging is the initial task to direct the patient into the right direction of being correctly diagnosed and recovery and our project automates this process alongside saving the medical record of the patient.

Our project starts from taking some basics patient information such as: Name, Age and Gender to save them into his/her Medical Record.

The next step of our program is asking the patient about what she/he feels and then the program starts to listen to what the patient would say.

After the patient finishes, the program takes the audio file and sends it to Google’s Speech-To-Text API which helps us to get the record of the patient in an unstructured text format that can be processed and analyzed using our preprocessing functions mentioned before in our implementation section.

The preprocessing step is crucial and vital step in our program it’s aimed to filter out all unnecessary text and remove stop words which doesn’t helps us in Triaging.

We take the patient text and transform it into pre-defined symptoms using stemming and lemmatization which gives us the root of the words and therefore the right symptoms.

After Preprocessing and extracting the symptoms, the program starts to predict the clinic the patient should go to in order to receive the right diagnosis.

The system reaches an accuracy of 88%.

The program saves the patient information and the prediction for later use.

## **6.2 Future Work**

There are a lot of future work to be made to improve both the performance and the accuracy.

Performance Wise:

* Better Arabic Natural Language Processing Algorithms in the future would lead to decrease to the processing time of text analysis.
* Enhanced Arabic Speech-To-Text APIs and service to process the speech faster

Accuracy Wise:

* Better and Improved dataset with more symptoms and its related clinics.
* More Samples and Real data obtained.
* Usage of EMR (Electronic Medical Record) to improve the results and accuracy.
* Help from top doctors and nurses.

# References

* Hammoud J., Vatian A., Dobrenko N., Vedernikov N., Shalyto A., Gusarova N. (2021) New Arabic Medical Dataset for Diseases Classification. In: Yin H. et al. (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2021. IDEAL 2021. Lecture Notes in Computer Science, vol 13113. Springer, Cham. <https://doi.org/10.1007/978-3-030-91608-4_20>
* Alshammari, N.; Alanazi, S. An Arabic Dataset for Disease Named Entity Recognition with Multi-Annotation Schemes. Data 2020, 5, 60. <https://doi.org/10.3390/data5030060>
* Health Information in Arabic <https://medlineplus.gov/languages/arabic.html>
* Wootton R. Telemedicine BMJ 2001; 323 :557 doi:10.1136/bmj.323.7312.557