Research Article

The Health ChatBots in Telemedicine: Intelligent Dialog System for Remote Support

Maria V. Vasileiou (ID) **and Ilias G. Maglogiannis** (ID)

University of Piraeus, Department of Digital Systems, Piraeus, Greece

Correspondence should be addressed to Ilias G. Maglogiannis; imaglo@unipi.gr

The objective of the Health Bot is to analyze and elaborate on the implementation of an intelligent system that can support telemedicine services using natural language processing (NLP) technology. Health Bot is a robust, modularized, and user-friendly platform that aims to improve the patient's interactions with the health care system. Using NLP and speech recognition algorithms, the platform can analyze and classify free text and voice input data to symptoms. The classified data are used in the machine learning (ML) training process of the artificial intelligence (AI) models in order to predict the probability of the patient suffering from a specific disease and get alerted in case of disorders. The Health Bot is acting as a medical virtual agent in order to collect any information needed in the sense of a medical interview and provide medical assessments, close appointments with the doctors, and monitor/record the patient's health condition. In this paper, two models have been trained for the Covid-19 and heart disease cases with 98.3% and 82% accuracy, respectively.

1. Introduction

Living in the era of Internet of Things (IoT), Big Data, and AI analytics, telemedicine systems like wearable activity trackers, medical sensors [28], produce valuable health-related data that need to be consumed and analyzed by intelligent platforms using AI and ML technology. These data are vital for classification, prediction, and recommendation engines [1]. For the collection of patient's health data in a more user-friendly way, ChatBots have been introduced. The ChatBot evolution offered a breakthrough to the legacy questionnaire systems by making the interviewing and the symptoms collection process more user-friendly using NLP algorithms. NLP is the technology, that is, used to aid computers to understand human's natural language. ChatBots, by using AI and ML algorithms, can heavily contribute to the patient's interviewing process for the collection of health symptoms and the provisioning of self-care smart recommendations using well-trained ML models. By using telemedicine applications like ChatBots, patients have the ability to communicate from everywhere, provide their symptoms, and retrieve an improved diagnosis. In addition, healthcare professionals can have access to comprehensive and consolidated patient data. Also, the patient can easily book an appointment using human language via ChatBot.

There is a need for transformation of the methods that the health care system is using to interact with patients on multiple levels. From:

(i) how patient's sensitive data are securely stored and analyzed,

(ii) how the patient interacts with the health care system to get answers to his/her questions,

(iii) how the patient retrieves his/her diagnosis based on the provided symptoms using human language compatible interfaces,

(iv) how the patient books an appointment using human-like conversation.

The Health Bot has been developed to close this gap so that patients are able to communicate and extract answers regarding their health conditions using human language. The Health Bot provides the ChatBot interface for

extracting and recording health symptoms using NLP technology and by using classification algorithms to predict health disorders. The Health Bot is an AI software that can identify the intention of the patient's questions and lead to the correct conversation flow by using natural language intelligence.

Section 2 consists of the fundamentals of the ChatBot theory. Section 3 elaborates on the Health Bot platform architecture. Section 4 focuses on the model's implementation and evaluation. Section 5 provides an overall conclusion of the results and provides future work suggestions.

2. Background and Related Work

Every day huge amounts of user-generated content are produced either in voice or text format. The categorization, clean-up, and insights processes are able to be streamlined with the help of NLP.

2.1. Natural Language Processing. NLP is the parsing and semantic interpretation of human-generated text, allowing machines to learn, analyze, and understand the context. By using NLP algorithms, valuable insights regarding the contextual, behavioral, and sentiment segmentation of the data stream can be achieved. The input and output of an NLP system can be voice, text, and image. There are various techniques that are heavily used during the NLP process, including but not limited to: grammar induction [2], which is used to produce a formal grammar with no given context; lemmatization [3] to identify word's lemma according to the meaning within the context; morphological segmentation [4] task to split words into individual elements and recognize their class; part-of-speech tagging [5] technique that can identify words with similar grammatical properties; "bag of words" [6] to tokenize-vectorize words after split from sentences; and word embedding [7] algorithm to extract features of words with the same meaning based on semantic lookalike relationships and same vector space distance. The most important module of NLP is natural language understanding (NLU).

2.2. Natural Language Understanding. NLU is the subset of the understanding and comprehension part of NLP. It elucidates the concept of the input string and transforms the unstructured data into classified data assigning them to the appropriate intents. In order to distinguish the meaning, classify and conclude the correct intent from the provided input, specific techniques are used like sentiment and content analysis. The systems that heavily rely on NLP are the AI ChatBots.

2.3. AI ChatBots. AI ChatBot's (hereafter called "ChatBots" for simplicity) objective is to use any applicable technology in order to mimic the conversation among human beings, achieved by the NLP algorithms.

2.3.1. ChatBot Application Workflow. The main components are also depicted in Figure 1 workflow and are given as follows: (a) the web or application interface in order to retrieve the input data, (b) the NLP algorithms in order to analyze and segment the sequence of words or speech, (c) the classification of the contextual meanings to entities that lead to the flow selection (intents), and (d) the response (text or audio). Analyzing further the workflow, mobile users are sending a voice or text message to the respective connector. The messaging voice connector is converting speech to text for further processing using the natural language parser. The words are broken down into defined keywords via the natural language parser. Afterward, the conversational engine is trying to fetch the correct dialog according to the segmented keywords that lead to the correct intent. User's request is analyzed by the ChatBot in order to locate the intents and extract the entities. This process is the fundamental prerequisite in the ChatBot's kernel. The conversational ML engine, either has a predefined response to serve via the response engine or performs external requests to Webhooks to retrieve the response after the functional processing. The response engine is responsible to serve the outcome to the user's device.

2.3.2. Response Outliers. For the design, training, and optimization of the ChatBot, human intervention is vital and plays a key role. ChatBot is trained to respond with the right answer, but if the input request is not understood then it may respond with the wrong answer. At that point the retraining and the human intervention may take place, to identify the unknown words and assign them to the correct intents manually.

2.3.3. Intent Flow. The objective of the conversational ML engine (agent) is to analyze and segment the input data and proceed with the classification analysis of the existence of the keywords based on the intents' training phrases. The outcome is the accurate intent selection, the extraction of the feature values (actions parameters), the finalization of the response output (responses), and the continuation of the DialogFlow according to the intent's rules. The process flow is illustrated in Figure 2.

2.3.4. Conversational AI Tool. One of the most powerful platforms for ChatBot creation is Google's DialogFlow (GDF). GDF is Google's ChatBot platform empowered by NLP technology. It provides tools to improve application interoperability with users through text and voice chat with AI technology. The main feature of DialogFlow is "small talk," "multilingual agent support," "cross platform support," "fulfillment," "training," "agent creation and management," "entities," "intents," "integrations," "in-line code editor," "analytics."

2.4. Related Work: Literature Survey. Ahmed Fadhil [8] is analyzing the role of telemedicine and healthcare support for home-living elderly individuals by using ChatBots. Another interesting work has been conducted by Divya S, et al. [9] related to personalized diagnoses based on symptoms. V. Manoj Kumar [10] has designed a search engine mechanism around
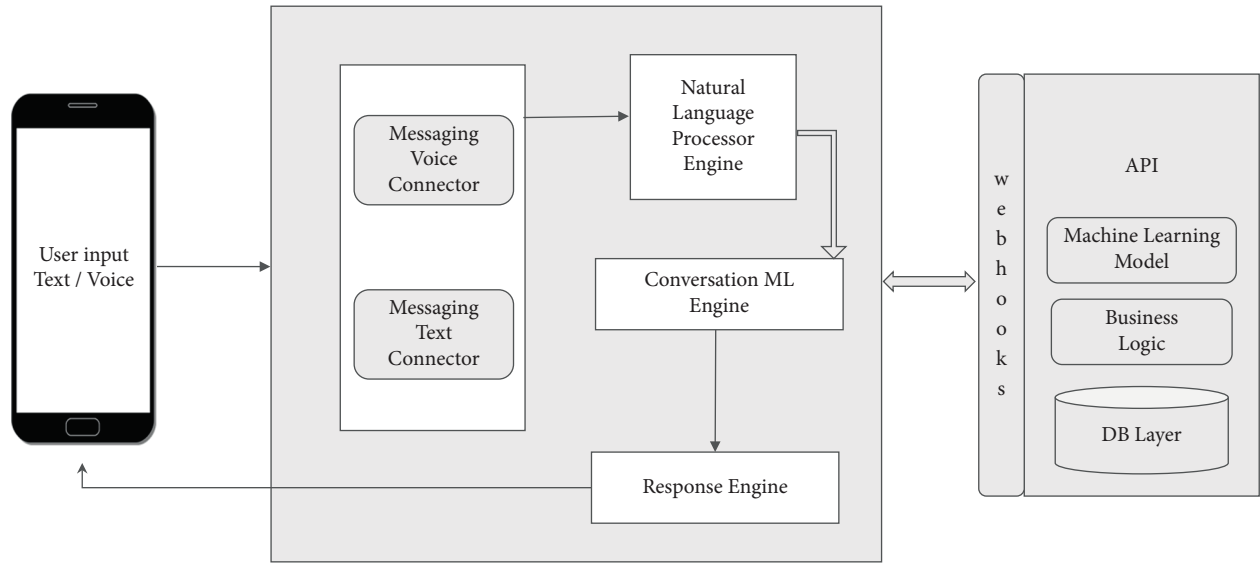
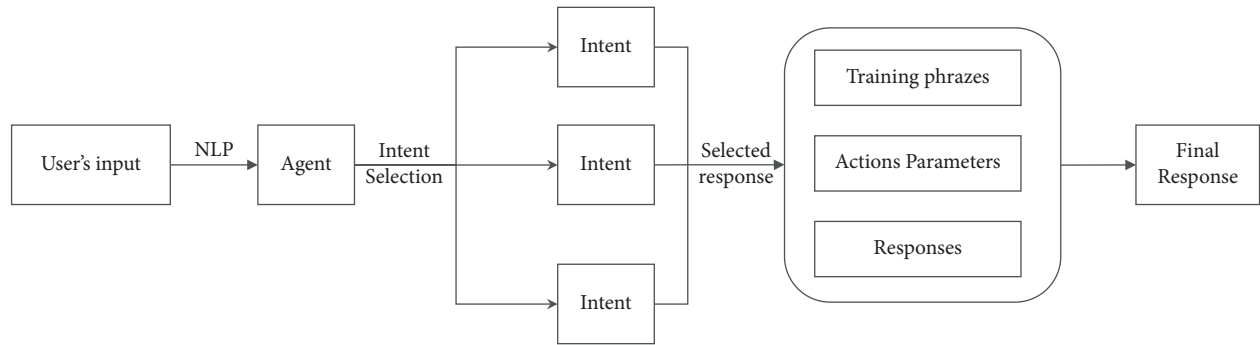FIGURE 1: ChatBot high-level design architecture.

FIGURE 2: Intent flow architecture.

the health context. Flora Amato, et al. [11] study elaborate on the effectiveness of e-health applications between humans and machines. Pereira J. and Díaz O. [12] analyzed the landscape of health ChatBots by focusing on three basic questions, what kind of diseases are the ChatBots encountering, what patient's skills do ChatBots aim for and who are the most interested ChatBot technology providers in the health sector. Automated medical bot is the subject of another interesting paper from Krishnendu Rarhi, et al. [13]. Battineni et al. designed and deployed a chatbot that can help patients living in remote areas by promoting preventive measures, virus updates, and reducing psychological damage caused by isolation and fear during the Covid pandemic [33]. In addition, S. P. Korres et al. provides solution of automated collection and storage of biosignals received from sensors that can help chatbot agent's AI training phase [30]. The objective of this paper is the design of a medical ChatBot that provides diagnosis and measures the seriousness of the diagnosis based on the symptoms. It is using AIML (artificial intelligence mark-up language) to detect human message patterns.

3. Design Methodology

The Health Bot platform is a modular system that facilitates the telemedicine ChatBot ecosystem requirements. The

objective of the platform is to enable the patients to interact with the system in human language format and extract valuable information regarding their health conditions. Also to predict the probability of suffering from specific diseases according to the provided symptoms. In addition, the patient can book an appointment with the doctor using the Health Bot virtual assistant and have an online interaction with the hospital institutes and the health experts.

The Health Bot architecture has been designed taking into consideration various aspects including, but not limited to.

3.1. Data Privacy. Data privacy and protection rules for GDPR compliance. The "patient's data" and the "Hospital application programming interface (API)" modules, as per Figure 3, are simulating the way that the hospitals should share the patient's data with health care 3rd party applications after the patient has given his consent. In order to protect the data exchange process, any information, that is, interchanged between the Hospital APIs and the ChatBot is encrypted using asymmetric encryption algorithms RSA/ECB/PKCS1PADDING. The private RSA key is protected and only the internal Hospital API layer has access to it. The
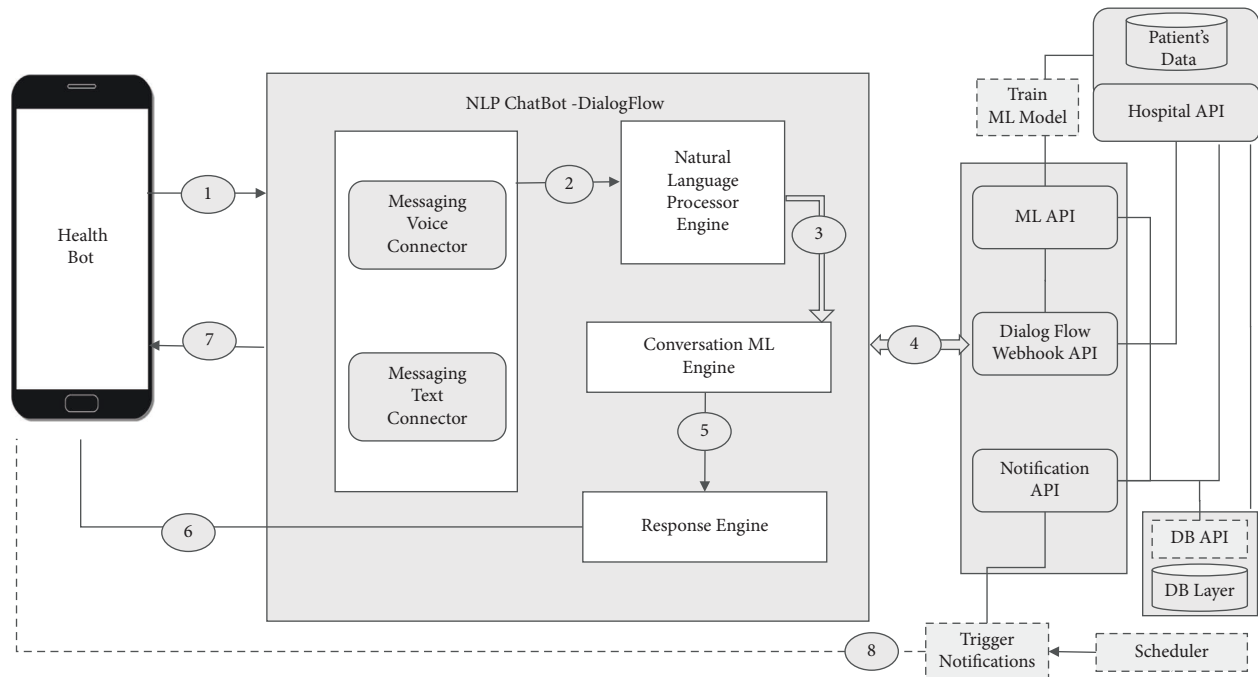
Figure 3: Health Bot architecture.

Health Bot application is using the public key to decrypt the request.

3.2. Platform Scalability and Availability.
Using distributed architecture with no single point of failure [14] for Health Bot platform hosting, Google cloud services have been used but not limited to firebase functions, database and hosting, big query ML, and AI cloud.

3.3. Fast Response times.
During the NLP process for entity identification and fetching the response from the Dialog-Flow Webhook APIs. The Health Bot platform, as depicted in Figure 3, is analyzed in the following subsections.

3.4. Health Bot DialogFlow

3.4.1. Health Bot User Interface (UI).
Health Bot User Interface (UI), is a vital component in the architecture for the interaction with the patient supporting voice and text-based conversational interface. It has been developed as a progressive web application (PWA) in order to facilitate the latest design methodologies and provide a native app-like user experience cross-device. There is a login authentication process in place in order to identify the patient, track his/her records and retrieve his/her health data via the Hospital API using his medical ID.

3.4.2. DialogFlow NLP.
DialogFlow NLP is the core module in the platform and consists of the NLP component that orchestrates the identification and conversion of the text input to intent classification in order to provide the most relevant response to the patient. It processes the text, and

voice input data using NLP algorithms and classifies the request to the respective intent via the intent classifier, which is trained to analyze the phrase and identify the entity keywords within that phrase. Once this process is successfully completed, the response management component is triggered to fetch the appropriate response.

The Health Bot has been trained based on the training phrase set of each intent in order to match the input data with relevant phrases that can lead to the correct intent selection. Each training phrase may have word reference to keywords that belong to specific entities that can act as custom vocabularies. The intent classifier consists of four main intent categories analyzed in the following section.

3.4.3. Intent Classifiers.
The demographic intent is used as the first step of the patient's interview. The objective is to collect the patient's personal data but not limited to home location, profession, marital status, age, and gender. These data are statistically interesting for identifying relations and building classifiers between age, gender, and location groups with symptoms and diseases. Each question for each of the demographic entities is mandatory, therefore, the bot requires to get a valid response before proceeding to the next question.

The symptom intent is responsible for collecting any known diseases that the patient is suffering from. This is a must step of the health interview and provides valuable insights into the patient's health status and the interpretation of the current symptoms. The symptom intent consists of sets of questions for different types of body anatomy clusters but not limited to eye, circulatory, ears, gastro, head, mental, musculo, nervus, nose, respiratory, skin, and urinary. The response to each question affects the overall health condition
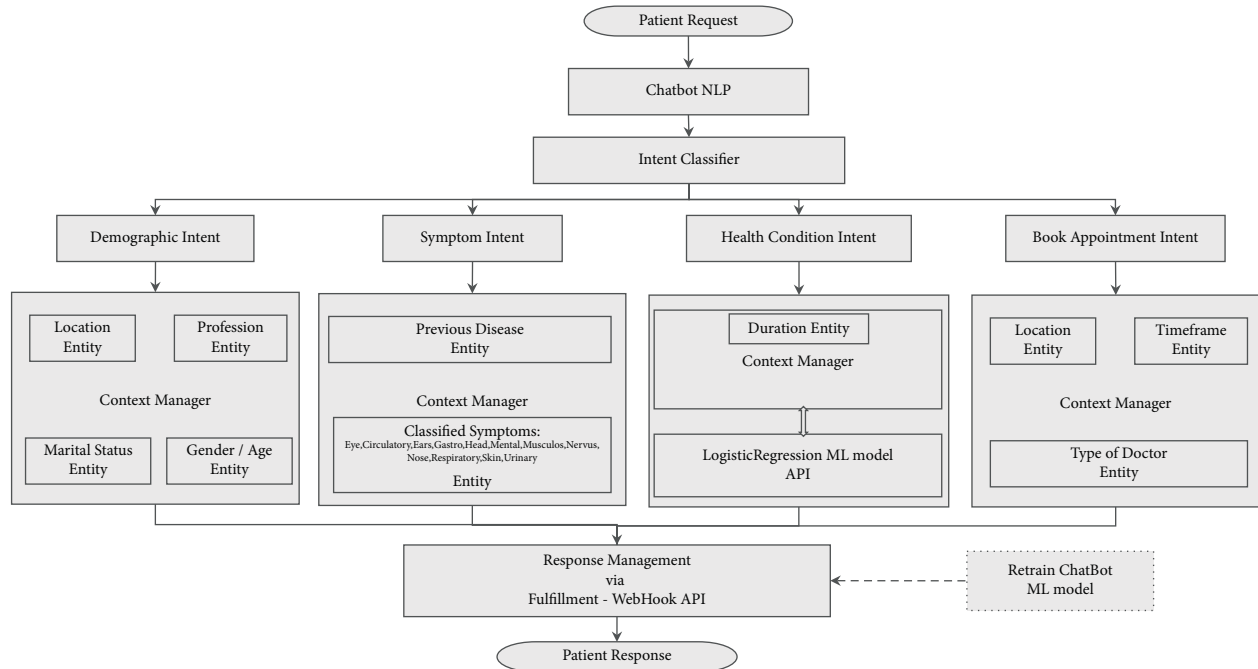
FIGURE 4: Health Bot conversation flow diagram.

scoring and the disease classification. The symptoms are evaluated and classified in order to provide an assessment regarding the patient's health condition, possible diseases, and if the patient should immediately visit a doctor. For the disease prediction process, both rule-based (refer to Appendix B), as well as ML prediction models, have been used.

The health condition intent is responsible for providing the patient's actual health status. The intent is matched with commands related to "what is my current health condition" or "how am I today?" Once the health condition indent is triggered, the Health Bot, by using the Webhook API, will fetch the health data from the Hospital's API (see in the next section). These data are provided to the ML model (see in the next section) which will respond to the health condition question with the respective prediction.

The book appointment intent is used in order for the patient to be able to book an appointment with the doctor using free text. It collects all the necessary information regarding the doctor's specialty, date, and time and proceeds with booking the appointment on behalf of the patient. The intent is triggered either by the patient by asking the Health Bot to book the appointment or by the Health Bot itself by judging according to the overall patient's health condition and prompting the patient to see a doctor as soon as possible.

The complete flow diagram is illustrated in Figure 4.

3.5. Health Bot Core. This section consists of modules that analyze the input data and extract the relevant response by using ML models and business intelligence.

The Hospital API is considered to provide patient's health data to external secure platforms via a web interface. Health Bot is consuming this API in order to train the ML models for providing more accurate results. The

identification of the patient's data is based on the medical ID unique identifier. Each patient should have his/her own medical ID as an identity to his/her health record. The Hospital API is communicating with the rest of the platform in a secure way. The payload data are encrypted using asymmetric encryption algorithms based on RSA/ECB/PKCS1PADDING technique. Health Bot's platform has only access to the public RSA key in order to decrypt the payloads. The private RSA key should be generated and safeguarded by the organization, that is, responsible for the health data collection and sharing.

The "train ML model" component is an asynchronous process, that is, used by the platform to train the ML models using the hospital's data and patient's provided input data. The trained model is used for health condition predictions and is saved to the Cloud AI so that the ML API can have access to it. Health Bot models have been fine-tuned using various different algorithms and datasets till the best outcome. The process is split into data preprocessing and normalization, model training and testing, model evaluation, prediction, and scoring as described in section 4.

The ML API is the interface that provides access to the trained ML model via RESTful HTTPs POST. The Health Bot platform is using Google Cloud. The trained models are stored in Google storage and exposed as API using the Cloud AI. Cloud AI is used to build, deploy, and manage ML models. The Health Bot platform is environment agnostic without dependencies.

The DialogFlow Webhook API is used by the DialogFlow NLP component in order to pass the patient's input dialog data and the classified intent to the Webhook and retrieve the produced response message after applying the respective business logic, that is, assigned to the specific intent's handler. For example, when the Health Bot has to respond to

TABLE 1: Covid-19 dataset attributes analysis.

Variable name	Variable type	Variable description
Medical ID	Unique identifier	The unique medical ID of the patient
fever	bit	If the patient has a fever (1 = true, 0 = false)
High fever	bit	If the patient suffers from a high fever (1 = true, 0 = false)
Cough	bit	If the patient has symptom of coughing (1 = true, 0 = false)
Shorten of breath target	bit integer	If the patient experienced shortness of breath the predicted attribute, diagnosis of Covid-19 (0 = absence, 1 = warning, 2 = quarantine)

the health condition intent question, "how am I today?," the Health Bot agent will trigger the specific handler of the Webhook API which will consume the Hospital API to retrieve the health data, then it will call the ML model to evaluate the data and provide the predicted value as the response to the patient.

The scheduler is a Crontab scheduler, that is, used to initiate the daily notification process in order to keep the patient informed of any disorder based on the daily hospital data feed that the Hospital API provides. If the Hospital API data for the specific medical ID deviates from the thresholds then an app notification is triggered via the Notification API.

The Notification API is listening to specific message queue topics. Once the scheduler publishes the message to the specific topic, the Notification API is triggered. It has two processes to follow. The first is to collect the daily health data for all the Health Bot subscribed patients using the Hospital API. The second is to invoke the ML API and predict the user's health condition by passing to the trained model the patient's daily data. If the ML prediction is abnormal the Notification API will send a notification to the respective user.

DB API is consumed from most of the processes in order to retrieve and store data in the following database collections. UserProfile which holds all the personal patient's metadata including the medical ID; HealthData which holds the health care retrieved data from the Hospital API categorized per patient per date; UserDialog which consists of the interview data that have been collected during the NLP dialog process; UserTokens which holds the specific device identifiers in order to send out the notification messages to the subscribed devices.

4. Evaluation and Results

There are two different models that have been trained for covering different needs. The first one is the logistic regression model, which is used for the prediction of Covid-19 disease, and the second one is another logistic regression model which is used for the prediction of heart diseases.

4.1. Covid-19 AI Model. The main objective of the model is to accurately predict if the user is suffering from Covid-19 according to the provided symptoms via the Health Bot interface. According to the World Health Organization [15], a person, that is, suffering from fever and experiencing cough, and shortness of breath has high probability of

TABLE 2: Covid-19 model training options.

Covid-19 model training options	
Max allowed iterations	20
Actual iterations	16
L1/L2 regularisation	0.00/0.00
Early stop	True
Min relative progress	0.01
Learn rate strategy	Line search
Line search initial learn rate	0.10

TABLE 3: Covid-19 model training overview.

Iter	Training loss	Evaluation loss	Learn rate	Duration (sec)
0	0.3219	0.3239	0.2	2.31
1	0.2598	0.2644	0.4	2.99
2	0.1972	0.2044	0.8	2.37
3	0.1553	0.164	1.6	2.33
4	0.1276	0.1349	3.2	2.65
5	0.1004	0.1041	6.4	2.54
6	0.0741	0.0762	12.8	3.11
7	0.066	0.0626	12.8	2.25
8	0.0587	0.0594	3.2	2.71
9	0.0548	0.0539	6.4	2.66
10	0.0508	0.0518	12.8	2.29
11	0.0471	0.0447	6.4	2.5
12	0.0447	0.044	3.2	2.31
13	0.0428	0.0418	6.4	2.29
14	0.0394	0.0385	12.8	2.68
15	0.035	0.0326	25.6	2.88

suffering from Covid-19. Based on this analysis, an emulated version of the dataset, consisting of 2100 rows, has been produced to train and test the Covid-19 model, using the attributes as per Table 1.

The classification algorithm that has been selected is the logistic_regression. Out of the 2,100 total rows, 20% (420) is used for testing and the rest 80% (1,680) for training and evaluation. The configuration options for the model's training are described in Table 2.

The model was initially scheduled for 20 iterations having an initial learning rate set at 0.2. At the 16th iteration, the model did not manage to exceed the lower boundary of the relative progress (i.e., the improvement of model's accuracy since the previous iteration) which was set to 0.01. As a result, the training process came to an early stop. Table 3 illustrates the quantitative and quality analysis per iteration. The average of the iteration process is 2.55 seconds. The iteration strategy in order to improve the model's occurrence
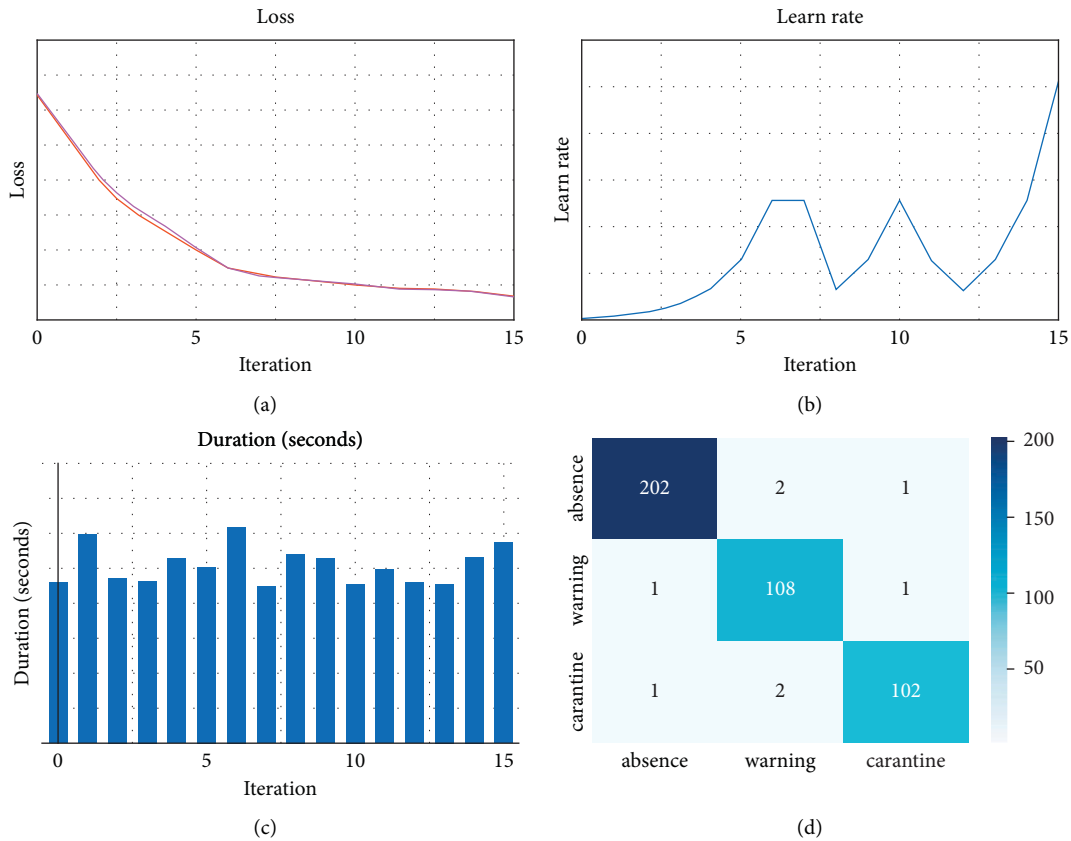
FIGURE 5: The AI model training graphs. (a) Loss function graph. (b) Learn rate graph. (c) Iterations' duration graph. (d) Confusion Matrix.
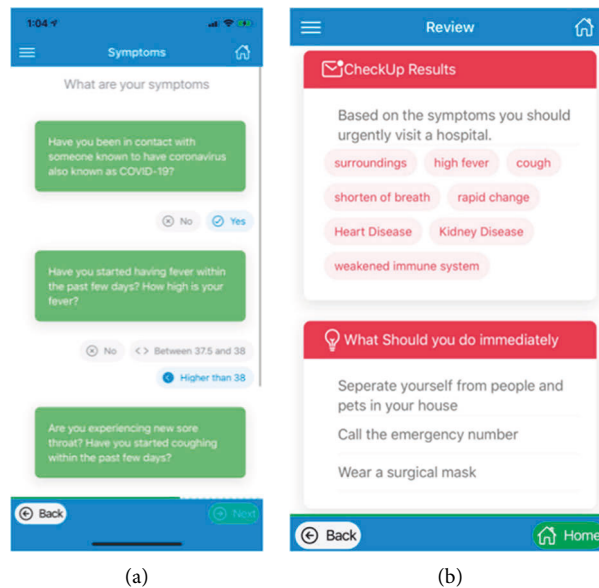

FIGURE 6: Covid-19 presentation layer. (a) Questionnaire. (b) Prediction.

was to double the learning rate per iteration till the training loss started to become bigger than the evaluation loss. At that point, the learning rate was starting to under doubling. The mean log loss is 0.0350. The exact parameters that have been used can be found in Appendix C.

In the next Figures 5(a)–5(c), the training overview data are illustrated per loss, duration, and learn rate charts for better view ability. The following chart depicts that the loss was decreasing per iteration. A rapid loss was achieved during the 8th and the 9th iteration, which caused the

TABLE 4: Cleveland dataset attributes analysis.

Variable name	Variable type	Variable description
Age	int64	Age of individual in years
Sex	int64	Gender (1 = male, 0 = female)
Cp	int64	Chest pain type (1 = typical angina, 2 = atypical angina, 3 = non-anginal pain, 4 = asymptomatic)
Trestbps	int64	Resting blood pressure in mm Hg
Chol	int64	Serum cholesterol in mg/dl
Fbs	int64	Fasting blood sugar level >120 mg/dl (1 = true, 0 = false)
Rest ECG	int64	Resting electrocardiographic results (0 = normal, 1 = having ST-T, 2 = hypertrophy)
Thalach	int64	Maximum heart rate achieved
Exang	int64	Exercise-induced angina (1 = yes, 0 = no)
Oldpeak	float64	ST depression induced by exercise relative to rest
Slope	int64	The slope of the peak exercise ST segment (1 = upsloping, 2 = flat, 3 = downsloping
Ca	int64	Number of major vessels (0–4) colored by fluoroscopy
Thal	int64	Thalassemia is an inherited blood disorder that affects the body's ability to produce hemoglobin and red blood cells.
Target	int64	(3 = normal, 6 = fixed defect, 7 = reversible defect the predicted attribute, diagnosis of heart disease (0 = absence, 1, 2, 3, 4 = present)

TABLE 5: Algorithmic model comparison.

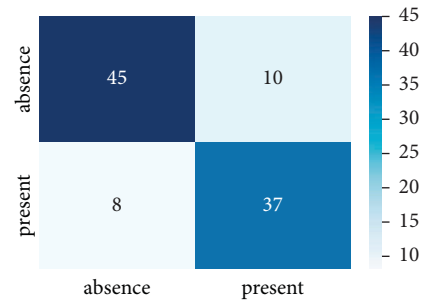| Model	MAE	Accuracy	
		Test (%)	Train (%)
Logistic regression	0.18	82	86.7
SVC	0.44	56	100
Naive bayes	0.2	80	87.68
Decision tree	0.22	78	100
Random forest	0.19	81	100

FIGURE 7: Confusion matrix of the selected logistic regression model.

learning rate to be downgraded accordingly for avoiding overfitting. On the other hand, from the 13th iteration, the learning rate was consistently increased in order to further improve the accuracy. The duration per iteration was lying in the range of 2.25 and 3.11 seconds. The model's accuracy on the train set is 99.5% and 98.3% on the test set and precisely can predict if the patient has to stay in quarantine and separate himself from the outside world. Figure 5(d) provides the confusion matrix metrics, the y-axis consists of the actual labels, whereas the x-axis is from the predicted ones.

Figures 6(a) and 6(b) are partially illustrating the Health Bot Covid-19 front end. In the presentation layer, the model is triggered once the patient's input is provided and responds with the level of emergency.

For further reference regarding the questionnaire that has been used to collect the patients' Convid-19 symptoms refer to Appendix A.
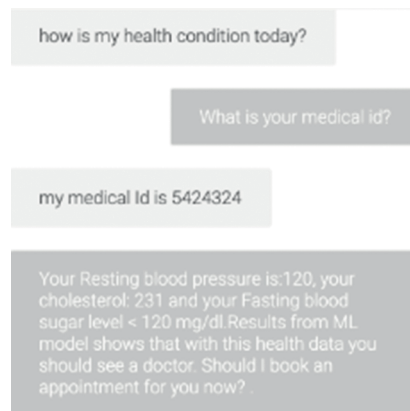
4.2. Heart Disease AI Model. The main objective of the model is to accurately predict if the user is suffering from heart disease. The model has been trained and tested using the Cleveland heart dataset [16] from UCI (Table 4) taking into account several contributing risk factors. The provided data are from 303 real clinical cases. In order to conclude the best heart disease ML model, various algorithms have been assessed according to their accuracy including but not limited to SVM, naive bayes, logistic regression, decision tree, random forest regression algorithms parameterized as per Appendix C.

FIGURE 8: Heart disease presentation layer.

Table 5 contrasts the models that have been trained using the sklearn [17] Python module. The test and the train set proportion was 33% (100)/67% (203), respectively. According to the mean absolute error as well as the scoring of each model against the test set, logistic regression appeared to be the most performing model and has been selected to be used by the Health Bot with 82% accuracy. In

TABLE 6: Covid-19 questionnaire.

Question	Phrazes values	Entities
Have you been in contact with someone known to have coronavirus also known as Covid-19?	Yes, No	Surroundings type: informative
Have you started having a fever within the past few days? How high is your fever?	No, between 37.5 and 38, lower than 38, higher than 38	Fever, high fever type: feature
Are you experiencing a new sore throat? Have you started coughing within the past few days?	Yes, No	Cough type: feature
Have you experienced shortness of breath the past few days	Yes, No	Shorten of breath type: feature
Are your symptoms getting worse very fast? Do you see a rapid change?	Yes, No	Rapid change type: informative
Do you have any conditions such as congestive heart disease, kidney disease, liver disease, or lung disease?	No, heart, kidney, liver, lung	Heart disease, kidney Disease, Liver disease, lung Disease, Type: informative
Do you have a weakened immune system such as patients that have had a transplant or are receiving chemotherapy	Yes, No	Weakened immune system type: informative

TABLE 7: Questionnaire for eyes category.

Question	Phrazes values	Symptom entities	Diagnose entities
Do you experience any of the following symptoms?	Do not know, blurry vision, trouble seeing at night, seeing colors as faded, increased sensitivity to glare, halos surrounding lights, double vision in the affected eye, need for frequent changes in prescription glasses	Blurry vision night vision fade colors glare sensitivity halos surrounding lights double vision	Cataracts
Do you experience any of the following symptoms?	Do not know, pink or red color in the white of the eyes, tearing, itchy or scratchy feeling in the eye, swelling, burning or irritation, crusting of the eyelids or lashes, especially in the morning, thick discharge that makes the eyes stick together	Red eyes tearing itch swelling irritation burning eyelids crusting discharge from the eye	Conjunctivitis (pink eyes)
Do you experience any of the following symptoms?	Do not know, double vision, decreased depth perception, jittery vision, eye strain, headache	Double vision depth perception jittery vision eye strain headache	Strabismus
Do you feel your eyes are dry with a sense of feeling like there is something in your eye?	Yes, No	Eyelids dry eyes	Dry eyes
Do you feel that your eyes are burning, that you have puffy eyelids, especially in the morning, stringy discharge in the eyes, or widened blood vessels in the outer layer of the eye?	Yes, No	Burning eyes puffiness of the eyelids	Allergic conjunctivitis
Do you experience any of the following symptoms?	Do not know, patchy blind spots in your side (peripheral) or central vision, often in both eyes, tunnel vision, bad headaches, eye pain, nausea, vomiting, blurry vision, seeing halos around lights	Blind spots bad headaches pain in the eye nausea vomiting blurry vision halos around lights	Glaucoma

the confusion matrix as presented in Figure 7, the y-axis consists of the true labels, whereas the x-axis from the predicted ones.

Based on the mean absolute error and the scoring of each model against the test set, logistic regression proved to be the most performing model. Therefore is the one that has been used by the Health Bot for predicting heart disease cases.

Figure 8 is partially illustrating the Health Bot heart disease section using free text commands. The intent is identified with NLP methodology. In the presentation layer, the model is triggered once the patient's input is provided and responds with the diagnosis.

5. Conclusion

The Health Bot has been implemented to provide a different perspective in the current way that healthcare interviews, symptom collection, and diagnosis take place. The Health Bot provides an intuitive web and app

interface that helps the patients to easily interact with the platform. Patients are enabled to use human language to communicate with the platform which makes the process more engaging and user-friendly. This is achieved by the well-trained NLP agent, that is, taking care of the free text classification and the flow dialog process. The Health Bot modules can be shared via APIs so that other third-party platforms can use them.

In the background, well-trained ML models and algorithms are running to provide the diagnosis using two models for the Covid-19 and heart disease cases with 98.3% and 82% accuracy, respectively.

It is vital to understand that by digitizing the health sector and using the patient's data for AI purposes, more powerful predictive ML models will be developed and more diseases will be identified and prevented on time by analyzing the patient's symptoms. [18] Therefore, Hospitals should share their data in an anonymized and secure way and help third-party platforms, such as the Health Bot, to extend their functionality and improve the health sector.

Future work will focus on training the NLP algorithms in more languages and vocabularies. This will help the patient's experience of using the platform and improve the dialog interaction. An additional improvement for the platform is to connect any available health data source via APIs and IoT data stream in order to enrich the patient's historical data and improve the algorithms accuracy and scoring.

Appendix

A. Covid-19 Questionnaire

The process of the DialogFlow agent is to identify the correct intent, which in this case is the Covid intent, and use the follow-up questions to complete the symptoms questionnaire (presented in Table 6), map it to the correct entities (symptoms) and provide the list of symptoms to the model to respond with the respective diagnosis. The mapped entities to the questions may be an actual feature of the model's training dataset or being collected for more in-depth analysis after the data are provided to the doctor for final evaluation (informative type). Some of the answers may be repeating (e.g., yes, no) but the agent can distinguish the meaning, as each question is part of its own context.

B. Rule-Based diagnosis

In case of the generic symptoms, there is a segmentation in order to apply more targeted questionnaires based on the affected human body's area (e.g., eyes, skin, head, back, thigh-knee-lower, leg-foot, arm-upperarm-elbow-forearm-hand, chest-heart, abdomen-upper abdomen, abdomen-middle abdomen, abdomen-lower abdomen, and throat). The patient's input phrase can consist of more than one human body's area categories, therefore, will get the respective questionnaires for all the different categories. Each matched symptom entity provides a specific weight regarding the potential diseases which are calculated before the diagnosis via DialogFlow's Webhook response handler. The data that have been used are manually generated and not from an official source. A sample of the questionnaire is presented in Table 7.

C. Environmental set up information of the ML models

This section provides extensive information regarding the parameters that have been used during the model's training process.

3.1 Covid-19 AI model

Logistic regression.

  (i) Max allowed iterations = 20
  (ii) Actual iterations = 16
  (iii) L1 regularisation = 0.00
  (iv) L2 regularisation = 0.00
  (v) Early stop = true
  (vi) Min relative progress = 0.01
  (vii) Learn rate strategy = Line search
  (viii) Line search initial learn rate = 0.10

3.2 Heart Disease AI Model

Logistic regression

based on sklearn logistic regression classifier [19]. The parameters that have been used are given as follows:

  (i) penalty function is l2′ Euclidean distance
  (ii) dual = False
  (iii) tol = 0.0001
  (iv) C = 1.0
  (v) fit_intercept = True
  (vi) intercept_scaling = 1
  (vii) class_weight = None
  (viii) random_state = None
  (x) solver = 'lbfgs'
  (xi) max_iter = 100
  (xii) multi_class = 'auto'
  (xiii) verbose = 0
  (xiv) warm_start = False
  (xv) n_jobs = None
  (xvi) l1_ratio = None

SVC

based on sklearn support vector classification [20]. The parameters that have been used are given as follows:

  (i) C = 1.0
  (ii) kernel = 'rbf'
  (iii) degree = 3
  (iv) gamma = 'scale'
  (v) coef0 = 0.0
  (vi) shrinking = True
  (vii) probability = False

(viii) tol = 0.001
(ix) cache_size = 200
(x) class_weight = None
(xi) verbose = False
(xii) max_iter = -1
(xiii) decision_function_shape = 'ovr'
(xiv) break_ties = False
(xv) random_state = None

Naive bayes

based on sklearn Gaussian naive bayes classifier [21]. The parameters that have been used are given as follows:

(i) priors = None
(ii) var_smoothing = 1e−09

Decision tree

based on sklearn decision tree classifier [22]. The parameters that have been used are given as follows:

(i) criterion = 'gini'
(ii) splitter = 'best'
(iii) max_depth = None
(iv) min_samples_split = 2
(v) min_samples_leaf = 1
(vi) min_weight_fraction_leaf = 0.0
(vii) max_features = None
(viii) random_state = None
(ix) max_leaf_nodes = None
(x) min_impurity_decrease = 0.0
(xi) class_weight = None
(xii) ccp_alpha = 0.0

Random forest

based on sklearn exhaustive search over specified parameter values for random forest estimator classifier [23]. The parameters that have been used are given as follows:

(i) estimator = 'Random forest classifier()'
(ii) param_grid = {'n_estimators': [200, 300]}
(iii) scoring = None
(iv) n_jobs = None
(v) refit = True
(vi) cv = 3
(vii) verbose = 0
(viii) pre_dispatch = '2 × n_jobs'
(ix) error_score = nan
(x) return_train_score = False

Data Availability

The dataset used to support the findings of the study can be obtained at Public Dataset D. Dua, E. Karra Taniskidou, 2017, UCI Machine Learning Repository, University of California, School of Information and Computer Science, Irvine, CA (http://archive.ics.uci.edu/ml/datasets/Heart+Disease).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. P. Erdeniz, I. Maglogiannis, A. Menychtas, A. Felfernig, T. Ngoc, and T. Tran, "Recommender systems for IoT enabled m-health applications," in *Proceedings of the AIAI 2018: Artificial Intelligence Applications and Innovations*, pp. 227–237, Springer, Cham, May 2018.

[2] E. Brill, "Automatic grammar induction and parsing free text: a transformation-based approach," in *Proceedings of the workshop on Human Language Technology*, pp. 259–265, ACL, March 1993.

[3] J. Kanis and L. Skorkovská, "Comparison of different lemmatization approaches through the means of information retrieval performance," in *Proceedings of the 13th international conference on Text, speech and dialogue TSD'10*, pp. 93–100, Springer, Berlin, Heidelberg, 2010.

[4] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, San Diego, CA, USA, May 2015.

[5] O. Owoputi, B. O'Connor, C. Dyer, K. Gimpel, and N. Schneider, *Part-of-speech tagging for Twitter: Word Clusters and Other Advances*, Carnegie Mellon University, Technical Report CMU-ML-12-107, 2012.

[6] S. Cucerzan, "Large-scale named entity disambiguation based on Wikipedia data," in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, Prague, Czech Republic, June 2007.

[7] R. Lebret and R. Collobert, "Word Embeddings through Hellinger PCA," in *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Gothenburg, Sweden, April 2014.

[8] F. Ahmed, "Beyond patient monitoring: conversational agents role in telemedicine & healthcare support for home-living elderly individuals," 2018, http://arxiv.org/abs/1803.06000.

[9] F. Habib, G. Shakil, S. Iqbal, and S. Sajid, "Self-diagnosis medical chatbot using Artificial intelligence," in *Proceedings of the Second International Conference on Smart Energy and Communication. Algorithms for Intelligent Systems, Springer, Singapore*, January 2021.

[10] V. M. Kumar, A. Keerthana, M. Madhumitha, S. Valliammai, and V. Vinithasri, "Sanative chatbot for health Seekers," *International Journal Of Engineering And Computer Science*, 2016.

[11] F. Amato, S. Marrone, V. Moscato, G. Piantadosi, A. Picariello, and C. Sansone, ∗ *Chatbots Meet eHealth: Automatizing Healthcare*, 2017.

[12] J. Pereira and O. Díaz, "Using health chatbots for behavior Change: a mapping study," *Journal of Medical Systems*, vol. 43, no. 5, p. 135, 2019.

[13] K. Rarhi, A. Bhattacharya, A. Mishra, and K. Mandal, "Automated medical chatbot," *SSRN Electronic Journal*, 2017.

[14] K. Jeferry, G. Kousiouris, D. Kyriazis et al., "Challenges emerging from future cloud application Scenarios," *Procedia Computer Science*, vol. 68, pp. 227–237, 2015.

[15] WHO, "COVID-19 advice for the public," 2019, https://www.
who.int/emergencies/diseases/novel-coronavirus-2019/covid-
19-vaccines/advice.

[16] D. Dua and E. Karra Taniskidou, *UCI Machine Learning
Repository*, University of California, School of Information
and Computer Science, Irvine, CA, 2017, http://archive.ics.
uci.edu/ml/datasets/Heart+Disease.

[17] scikit-learn, "scikit-learn," 2020, https://scikit-learn.org/
stable/.

[18] I. Maglogiannis, "Towards the Adoption of open source and
open access Electronic health record systems," *Journal of
Healthcare Engineering*, vol. 3, no. 1, pp. 141–162, 2012.

[19] scikit-learn, "Scikit-learn Logistic regression," 2022,
https://scikit-learn.org/stable/modules/generated/sklearn.
linear_model.LogisticRegression.html.

[20] scikit-learn, "scikit-learn SVC," 2022, https://scikit-learn.org/
stable/modules/generated/sklearn.svm.SVC.html.

[21] scikit-learn, "Scikit-learn Gaussian naive Bayes," 2022, https://
scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.
GaussianNB.html.

[22] scikit-learn, "Scikit-learn decision tree," 2022, https://scikit-
learn.org/stable/modules/generated/sklearn.tree.
DecisionTreeClassifier.html.

[23] scikit-learn, "Scikit-learn random forest," 2022, https://scikit-
learn.org/stable/modules/generated/sklearn.model_selection.
GridSearchCV.html.

[24] M. Vasileiou, *Implementation of Intelligent System to Support
Remote telemedicine Services Using Chatbots technology*,
University of Piraeus, Department of Digital Systems, 2020,
https://dione.lib.unipi.gr/xmlui/handle/unipi/13494.

[25] M. Carl and A. Way, *Recent Advances in Example-Based
Machine Translation*, Springer, Netherlands, 2003.

[26] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On
the properties of Neural machine translation: Enco-
der–Decoder approaches," in *Proceedings of the SSST-8,
Eighth Workshop on Syntax, Semantics and Structure in
Statistical Translation*, Association for Computational Lin-
guistics, Doha, Qatar, October 2014.

[27] J. Chorowski and N. Jaitly, "Towards better decoding and
language model integration in sequence to sequence models,"
2016, https://arxiv.org/abs/1612.02695.

[28] D. Chousiadas, A. Menychtas, P. Tsanakas, and
I. Maglogiannis, "Advancing Quantified-self applications
Utilizing Visual data analytics and the Internet of Things," in
*Proceedings of the AIAI 2018: Artificial Intelligence Applica-
tions and Innovations*, pp. 263–274, Springer, Cham, May
2018.

[29] S. Hochreiter and J. Schmidhuber, "Long short-term mem-
ory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[30] S. P. Korres, A. Menychats, P. Tsanakas, and I. Maglogiannis,
"A low-cost IoT-based health monitoring platform enriched
with social networking facilities," in *Proceedings of the IEEE
International Conference on Pervasive Computing and Com-
munications Workshops (PerCom Workshops)*, IEEE, Athens,
Greece, March 2018.

[31] L. H. Sarah and H. David, "One-hot Encoding," in *Digital
Design and Computer Architecture*, 2022.

[32] R. Rojas, *Neural Networks - A Systematic Introduction*,
Springer-Verlag Berlin, Heidelberg, 2022.

[33] G. Battineni, N. Chintalapudi, and F. Amenta, "AI chatbot
design during an Epidemic like the Novel Coronavirus,"
Healthcare, vol. 8, no. 2, p. 154, 2020.