VOLUME - 10, ISSUE - 07, JULY- 2021 • PRINT ISSN No. 2277 - 8160 • DOI : 10.36106/gjra

Original Research Paper | Health Science

MEDICARE AI-ASSISTED AI IN HEALTHCARE

Nitish Nagar	AIML Lambton College Toronto, Canada
Plaban Roy	AIML Lambton College Toronto, Canada
Saurabh Deswal	AIML Lambton College Toronto, Canada
Parth Pandya	AIML Lambton College Toronto, Canada
Dr. Aashna Bhardwaj*	BDS, MBA In Healthcare And Hospitality Management Delhi, India. *Corresponding Author
Parisa Naraei	PhD Cestar college of business, health and technology Toronto, Canada

ABSTRACT Artificial Intelligence is nowadays at the peak, and considering chatbots is one of the primary examples contributing to success. In the project paper, we implemented an instant diagnostic medical checkup for the patient, which can help doctors/healthcare providers, as this order saves time. By analyzing the symptoms, the Assist AI system will help determine the severity of the disease in real-time and will respectively guide the patient with the different stages – disease definition, precaution and recommendation.

KEYWORDS : healthcare, artificial intelligence, diagnosis, chatbots.

I. LITERATURE REVIEW

Wouldn't it be great to have a private doctor on your phone or laptop all the time? Well, we not there yet, but we are close enough with chat-based symptom checker applications. Their easy-to-use human-like conversations make them very easy to use and convenient[24].

Though chat-based symptom checker applications are more popular than ever before, there are few common drawbacks[23]; after reviewing top chat-based symptom checker applications, it was found that there are five common limitations, which are:

- Missing patient history
- Strict input requirements
- A long list of probing questions
- Not covering diverse health problems
- Missing follow-up treatment plan.

While building Medicare AI, we have reviewed these points and have worked to overcome the problem of missing follow-up treatment plans. Our application does disease classification, chronicity checking, severity checking and follow-up care path for the patients. This makes our application unique and competitive in the market.

Machine learning has the ability to provide physicians and hospital personnel with data-driven clinical decision support (CDS), potentially increasing revenue. Deep learning, a subset of AI that uses algorithms and data to provide automated insights to healthcare providers, is a subset of AI that identifies trends.

Example: Two most popular examples of AI in medicine and healthcare are Google Cloud's Healthcare app, which makes it easier for health organizations to capture, store, and access data, and IBM Watson's ability to pinpoint treatments for cancer patients.

II. INTRODUCTION

Why are chatbots as a technical challenge so fascinating? How is the healthcare industry revolutionized with the impact of artificial intelligence? The rise of automation has resulted in personal digital assistants developed with the increased computational power, the ease of accessibility of data, implementation of statistical analysis and much more[3]. The healthcare providers, the payers or also we can say the medical assistants have started leveraging the applications of AI in the medical field. Such AI-enabled tools determine and simplify the care for the patient and reduce unnecessary costs. Whenever a patient engages with the 'person standing by' on the virtual side of the machine, who may sound or represent as human but underneath, it's the intelligent conversational machine[2].

Integrating AI into the healthcare ecosystem has a number of advantages, including the ability to automate tasks and analyze large patient data sets in order to provide quality healthcare quicker and at a lower cost.

According to a report, Administrative duties account for 30% of healthcare expenses, according to Insider Intelligence. Some of these tasks, such as pre-authorizing insurance, following up on outstanding bills, and maintaining records, can be automated by AI to relieve healthcare professionals' strain and save them money.

AI is capable of analyzing large data sets, combining patient insights and resulting in predictive analysis. Obtaining patient insights quickly allows the healthcare ecosystem to identify crucial areas of patient care that need to be improved.[25]

AI is also used in wearable healthcare devices to better assist patients. FitBits and smartwatches, for example, employ AI to analyze data and warn users and their healthcare providers to potential health issues and risks. The ability to check one's own health through technology reduces professional workload and avoids unneeded hospital trips or remissions.[25]

Implementing an AI system for the early diagnosis and providing recommendations based on the severity of the disease could save a lot of money in terms of the doctor visit and the time wasted waiting at clinics. Therefore, in some situations, it becomes the dire need of the patient to have a well-versed idea about the disease description and recommendations based on that.

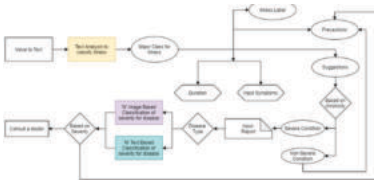


Fig. 1. Project Workflow.

The program aims to investigate the current system and development of add ons such as chatbots for the early diagnosis with the help of machine learning algorithms and techniques. It is a Django-python-based web application that would behave as the early diagnosis tool for health problems. The app contains three phases: an AI-based health physician system to interact with the patient with the help of voice-to-text conversion in the initial stage, perform the diagnose, and suggest a quick precautionary or recommendation measure to treat the problem[1]. Google's Speech-to-text API enabled us to convert the audio to text, supporting over 120 different languages[6]. We faced the challenge of working on various languages to train a machine learning model to classify the disease. Hence, we chose to go for English as the primary language for the classification.

The Naïve Bayes classifier works on the idealogy of determining the output label based on the probabilities of events. As in the industry level, significant companies use Naïve Bayes for the text classification, so we thought of going by the same on our disease classification dataset.

The decision tree machine learning algorithm is adapted to follow up a top-down approach for searching the correct disease classified by the system[5]. Decision trees are considered reliable and effective as they follow the high quality of the classification mechanism for decision-making [4]. The technique involves asking a series of questionnaires about the symptoms based on which the model decides the disease recommendation and preventive measures.

To predict the disease's classification, how severe the condition is, we have implemented and tested various machine learning models. The different diseases have their respective method of analysis which is determined by the predefined available datasets. These datasets include the parameters of influential factors about the disease. Primarily, for the initial development, we have focused on Hypothyroidism, Diabetes, Heart Disease and Acne. The first three modules work on machine learning analysis using the Random Forest algorithm, respectively.

In Hypothyroidism, the thyroid gland doesn't create or release the hormones into the bloodstream, which slows down the body's metabolism[7]. It will make the individual feel tired, gain weight and unable to tolerate cold temperatures.

For diabetes, it can be classified primarily in two categories chronic categories, diabetes type 1 and diabetes type 2, which lead to the excess of sugar (glucose) in the bloodstream, which is harmful to the organs[8].

Heart disease is one of the complex diseases with which a significant population of the world is suffering. Efficient identification and timing play a crucial role in the healthcare sector[9].

Skin diseases such as Acne problem is quite common even from the very initial age. The remedy is also an easy job, but getting the correct diagnosis and checkup can be very troublesome. Deep neural networks are nowadays in boom with image classification and object/pattern discovery methods[10].

We have implemented various machine learning models to classify the three diseases and concluded that Random Forest performs best in accuracy metrics as a machine learning model. And for the image classification, we chose a Convolutional Neural Network, a particular Deep Learning Neural Network. Hence, we adopted a novel approach to classify both text and image using separate models.

III. METHODOLOGY

The project is divided into three phases, based on its functionalities. This project uses different machine learning models at various stages.

A. Algorithms

In phase one, the project uses a Naive Bayes classifier. The strong point about the Naive Bayes classifier is that it is built based on Bayes Theorem with solid independence. Naive Bayes is one of the simplest models, but when integrated with Kernel Density Estimators (KDE), they can achieve higher accuracy levels.

In phase two, this project uses decision trees to make various classification-based decisions. A decision tree is a tree-like model that depicts decisions and their multiple symptoms.

In phase three, the project uses Random Forest Classifier. A Random Forest Classifier consists of several decision trees and outputs the class that is the mode of the class. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance.

B. Dataset

Various datasets are being used in this project. In phase one, the project uses data for voice-to-text transformation using Google Translation web API and then implementing a Naïve Bayes model to classify the illness[11]. The medical speech dataset contains 8.5 hours of audio paired with symptoms of most common medical symptoms. It includes common symptoms such as knee pain or headache. The dataset contains both audios as well as the corresponding transcripts.

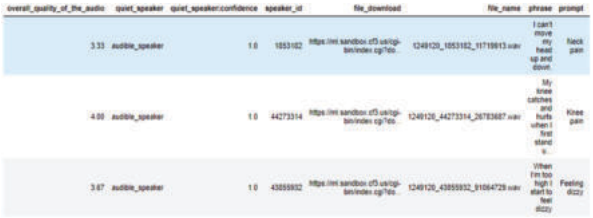


Fig. 2. Sample for medical speech dataset.

We used the symptoms dataset; the four significant datasets we have used are symptoms severity, symptoms description, and symptoms precaution. The symptoms precautions dataset contains the precautions associated with a disease. The symptoms description and symptom precaution dataset have the precautions needed to be taken by the patients, and the description provides the details about the disease.

	Disease	Description
0	Drug Reaction	An adverse drug reaction (ADR) is an injury ca...
1	Malaria	An infectious disease caused by protozoan para...
2	Allergy	An allergy is an immune system response to a f...
3	Hypothyroidism	Hypothyroidism, also called underactive thyroi...
4	Psoriasis	Psoriasis is a common skin disorder that forms...

Fig. 3. Sample for disease description dataset.

	Disease	Symptom_1	Symptom_2	Symptom_3	Symptom_4	Symptom_5	Symptom_6	Symptom_7	Symptom_8	Symptom_9	Symptom_10
0	Fungal infection	Itching	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN	NaN	NaN	NaN
1	Fungal infection	skin_rash	nodal_skin_eruptions	dischromic_patches	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	Fungal infection	Itching	nodal_skin_eruptions	dischromic_patches	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	Fungal infection	Itching	skin_rash	dischromic_patches	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	Fungal infection	Itching	skin_rash	nodal_skin_eruptions	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Fig. 4. Sample for diseases with respective symptoms.

	Symptom	weight
0	itching	1
1	skin_rash	3
2	nodal_skin_eruptions	4
3	continuous_sneezing	4
4	shivering	5

Fig. 5. Sample for symptom severity dataset.

The thyroid disease dataset contains Class, T3-Resin, Serum thyroxin, serum triiodothyronine, TSH, and AD_TSH. There are three class attributes, i.e. class 1, class 2, and class 3.

	Class	T3-Resin	Serum thyroxin	serum triiodothyronine	TSH	AD_TSH
0	1	113	9.9	3.1	2.0	5.9
1	1	127	12.9	2.4	1.4	0.6
2	1	109	5.3	1.6	1.4	1.5
3	1	105	7.3	1.5	1.5	-0.1
4	1	105	6.1	2.1	1.4	7.0

Fig. 6. Sample for thyroid dataset.

The diabetes dataset helps in predicting diabetes. The columns of the diabetes dataset are age, Gender, Polyuria, Polydipsia, sudden weight loss, weakness, Polyphagia, Genital thrush, visual blurring, itching, irritability, delayed healing, partial paresis, muscle stiffness, Alopecia, Obesity, and class.

	Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	Genital thrush	visual blurring	Itching	Irritability	delayed healing	partial paresis	muscle stiffness	Alopecia	Obesity	c
0	40	Male	No	Yes	No	Yes	No	No	No	Yes	No	Yes	No	Yes	Yes	Yes	Po
1	58	Male	No	No	No	Yes	No	No	Yes	No	No	No	Yes	No	Yes	No	Po
2	41	Male	Yes	No	No	Yes	Yes	No	No	Yes	No	Yes	No	Yes	Yes	No	Po
3	45	Male	No	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	No	No	No	Po

Fig. 7. Sample for diabetes dataset.

The Cleveland heart disease dataset is used to predict heart disease. This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that ML researchers to this date have used. The "goal" field refers to the presence of heart disease in the patient. It is integer-valued from 0 (no presence) to 4.

	age	sex	chestpain	restbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	1	145	233	1	2	150	0	2.3	3	0.0	6.0	0
1	67	1	4	160	286	0	2	108	1	1.5	2	3.0	3.0	2
2	67	1	4	120	229	0	2	129	1	2.6	2	2.0	7.0	1
3	37	1	3	130	250	0	0	187	0	3.5	3	0.0	3.0	0
4	41	0	2	130	204	0	2	172	0	1.4	1	0.0	3.0	0

Fig. 8. Sample for heart disease dataset.

C. *Tools*

- Jupyter Notebook: The Jupyter notebook provides an environment for interactive computing. Jupyter notebook makes it easier to visualize the data and the code more simply.
- Django Framework: Django is a high-level python-based framework that is extremely fast, highly scalable, and secure. It encourages rapid, hassle-free development as it is open source and free. Some famous applications using the Django framework are PBS, Instagram, Bitbucket etc.
- Pandas: Pandas is one of the more critical machine learning libraries as it helps in converting important file formats such as CSV, JSON, TSV, and SQL databases to pandas data frames. The pandas data frames give a better intuition and understanding of the data.
- Matplotlib: Matplotlib is a python library that is used for quality visualization. When matplotlib is integrated with Numpy, it performs well with extensive data.
- Scikit-learn: Scikit learn is a machine learning library built on Scipy, matplotlib, and Numpy. Scikit learn can be used in various Classification, Regression, Clustering, dimensionality reduction, model selection, data preprocessing, etc.
- Keras: Keras is a free and easy-to-use powerful python library used to develop and evaluate deep learning models. It is built on top of Tensorflow. Keras can be used with just a few lines of code. Just like other deep learning libraries, Keras also uses various layers to make optimal decisions.

D. *Programming Languages*

Python is a high-level, object-oriented programming language with high readability that helps attain high-level programming complexity. Python is a modular language and comes with the capability of code reuse. Python supports various high-level libraries and packages, making python the favourite programming language for data scientists and machine learning engineers.

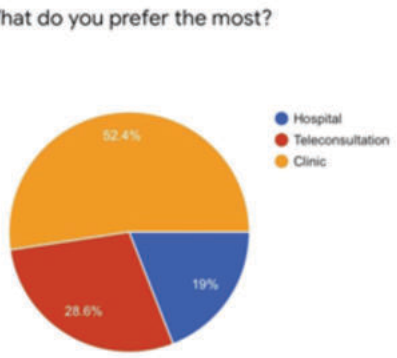
IV. RESULTS

The final system of the application is an add-on prototype of a medical assistant. Once the machine learning model classifies the illness/disease of the patient, it follows a module that prescribes the definition of the illness and the preventive measures one should take.

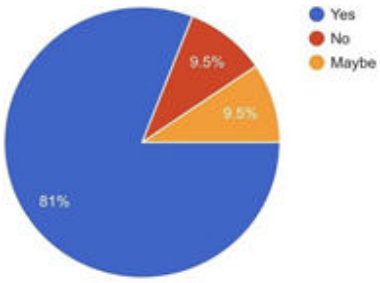
The system calculates the severity of the disease based on the patient's input for the number of days. It is determined as to how many days the patient is suffering from the disease. If the severity is below the threshold limit, the system describes the illness and its reason. Also, it explains the precautionary/ recommendation measures to be taken to minimize the effect of the disease. If the patient wants to analyze a particular disease with more detail-oriented inputs such as an image for the condition of disease (Acne), or lab report parameters as inputs, the system can perform it.

We have incorporated four diseases as mentioned earlier for the current system, and in future works, we will indeed scale the number. The system is highly dynamic, robust, and scalable as implementing or adding the supporting disease is manageable. The modules are interlinked but are independent of attaining the system scalability feature.

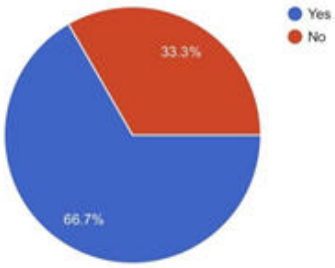
- Design: Based on a literature review and data above, we developed a questionnaire to assess patient's knowledge and views on artificial intelligence in healthcare.
- Sample size: 100
- Data collection tool: questionnaire



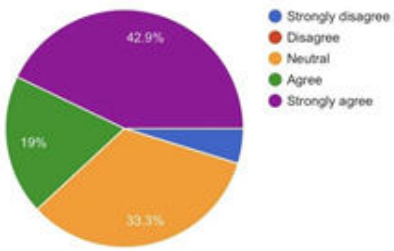
Are you familiar with the term Artificial Intelligence and it's applications in healthcare?



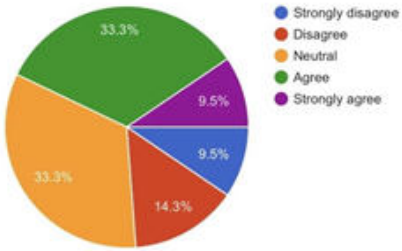
Have you ever tried Tele-consultation before?



Do you agree that artificial intelligence has useful applications in the medical field?



Do you agree that the diagnostic ability of AI is superior to the clinical experience of a human doctor?



V. DISCUSSION AND CONLUSIONS

As stressful as it can be, group works are beneficial in the long run. Our team has closely paralleled the dynamics of serving a project as we do in the workplace settings. We formed the group, keeping in mind the different traits of the individual member. We as a team contributed with our skills, experience and personality to the project.

AI-based medical assist will never be able to replace the traditional physician-patient relationship. However, we can relieve the doctors and care providers from repetitive work so that the doctors can do what they do the best to provide care to the patients and let the machine do the mundane job.

We can perform the implementation of the system as an add-on for the current system of Epic System's EHR model. And as stated by the Epic Development Labs, the team needs the frontiers to work on inventive ideas. "These projects give me the freedom and flexibility to pursue things that could be valuable to the codebase and end-users, but might have been prioritized behind other features on the to-do list." – Tim Bahls, Software Developer at Epic.

The primary goal is to save the efficient time and money of doctors and somehow for patients as well so that for minor cases, the early diagnosis and treatment can be performed. Also, it is to sort out the patients who need a quick visit to doctors using our AI assistant initial screening, thus providing care to the patients who need attention.

As for the current system front-end UI, we implemented the Django framework to build the application. The system needs some improvement as in:

- Calling dynamic valued inputs: The symptom list is dynamic and varies concerning the disease or illness classified. It will enable the accessibility to provide input for the symptom that whether the patient is suffering from it or not.
- Implementation as an API call: There are multiple cloud services available freely and with the paid versions. The use of these cloud services depends on the application and the design. We can host the developed application over a cloud platform and provide it as a service over the web itself as an API.

The implementation of voice to text module currently supports the only language, ie. English, for the disease classification based on the text input. As this input is available from the patient's voice and it can be in any of the languages, we need to perform the translation to English and then perform the data analysis. Or we can build machine learning models for all the languages and then train them respectively for the classification, but that won't be a feasible solution to develop.

REFERENCES

[1] Bulla, Chetan & Parushetti, Chinmay & Teli, Akshata & Aski, Samiksha & Koppad, Sachin. (2020). A Review of AI Based Medical Assistant Chatbot. 2. 1-14. 10.5281/zenodo.3902215.
[2] Mohamud Mohamed, M., & Zhuopeng, W. (2020). Artificial Intelligence Healthcare Chatbot System. IJARCCE. https://ijarcce.com/wp-content/uploads/2020/03/IJARCCE.2020.9223.pdf
[3] Conversational Agents in Health Care: Scoping Review and Conceptual Analysis. (2020, August 1). PubMed Central (PMC). https:// www. ncbi. nlm. nih. gov/pmc/articles/PMC7442948/
[4] Podgorelec, V., Kokol, P., Stiglic, B., & Rozman, I. (2002, November 1). *Decision Trees: An Overview and Their Use in Medicine*. ResearchGate. https:// www. researchgate. net/ publication/ 11205595 Decision Trees An Overview and Their Use in Medicine
[5] Bae J. M. (2014). The clinical decision analysis using decision tree. Epidemiology and health, 36, e2014025. https:// doi. org/ 10. 4178/ epih/ e 2014025
[6] Using the Speech-to-Text API with Python |. (2018). Google Codelabs. https:// codelabs. developers.google.com/codelabs/cloud-speech-text-python3#0
[7] *Hypothyroidism: Symptoms, Causes, Treatment & Medication*. (2016). Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/12120-hypothyroidism#:%7E:text=Hypothyroidism%20is%20a%20common%20condition,unable%20to%20tolerate%20cold%20temperatures.
[8] *Diabetes - Symptoms and causes*. (2020, October 30). Mayo Clinic. https://

www.mayoclinic.org/diseases-conditions/diabetes/symptoms-causes/syc-20371444

[9] *Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare*. (2020). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/9112202

[10] M. S. Junayed et al., "AcneNet - A Deep CNN Based Classification Approach for Acne Classes," 2019 12th International Conference on Information & Communication Technology and System (ICTS), 2019, pp. 203-208, doi: 10.1109/ICTS.2019.8850935.

[11] Aiyappan, J. (2019, September 8). *Naive Bayes Classifier for Text Classification - Analytics Vidhya*. Medium. https://medium.com/analytics-vidhya/naive-bayes-classifier-for-text-classification-556fabaf252b

[12] Exploring Tech Frontiers with Epic Development Labs. (2020). Epic. https://www.epic.com/epic/post/exploring-tech-frontiers-epic-development-labs

[13] OBrien, M. S. (2006). Implementation of the EPIC Electronic Medical Record/Physician Order-Entry System. Journal of Healthcare Management, 51(5), 338–343. https://doi.org/10.1097/00115514-200609000-00011

[14] Park, A. (2017, October 6). How Robots Are Changing the Way You See a Doctor. Time. https://time.com/4967153/artificial-intelligence-machine-learning/

[15] Medical Speech, Transcription, and Intent. (2019, February 24). Kaggle. https://www.kaggle.com/paultimothymooney/medical-speech-transcription-and-intent

[16] Thomas Davenport, The potential for artificial intelligence in healthcare. (2019,June6).NCBI.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6616181/

[17] D. Paul, Artificial intelligence in drug discovery and development. (2021, October,21).PubMedCentral(PMC).https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7577280/

[18] C. Park, Artificial Intelligence in Health Care: Current Applications and Issues. (2020, November 2). PubMed Central (PMC). https:// www. ncbi. nlm. nih. gov/ pmc/articles/PMC7606883/

[19] B. Pact, Doctors as patients: a systematic review of doctors' health access and the barriers they experience. (2008, July 1). PMC. https:// www. ncbi. nlm. nih. gov/ pmc/ articles/PMC2441513/

[20] repository, U. C. I. M. (2012). UCI Machine Learning Repository: Heart Disease Data Set. UCI: Machine Learning Repository. https:// archive. ics. uci. edu/ ml/ datasets/ Heart+Disease/

[21] repository, U. C. I. M. (2001). UCI Machine Learning Repository: Thyroid Disease Data Set. UCI: Machine Learning Repository. http:// archive. ics. uci. edu/ ml/ datasets/Thyroid+Disease

[22] *Medical Speech, Transcription, and Intent*. (2019b, February 24). Kaggle. https://www.kaggle.com/paultimothymooney/medical-speech-transcription-and-intent

[23] You, Yue & Gui, Xinning. (2020). Self-Diagnosis through AI-enabled Chatbot-based Symptom Checkers: User Experiences and Design Considerations.

[24] Richens, J.G., Lee, C.M. & Johri, S. Improving the accuracy of medical diagnosis with causal machine learning. Nat Commun 11, 3923 (2020). https:// doi.org/10.1038/s41467-020-17419-7