Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240570

Towards a Chatbot for Medical Diagnosis Based on Patient Symptoms

Kisito KABORE[a], Omar SAWADOGO[a,1], Yaya TRAORE[a] and Julie THIOMBIANO[b]
[a]*Joseph KI ZERBO University, Burkina Faso*
[b]*Nazi BONI University, Burkina Faso*
ORCiD ID: Omar SAWADOGO https://orcid.org/0009-0008-6399-5697

Abstract. The advent of artificial intelligence has positively transformed many areas of our lives, including the medical field. In this article, we propose the development of a medical diagnosis chatbot based on patients' symptoms, using artificial intelligence as an innovative solution. The aim of this tool is to provide doctors with a preliminary diagnosis based on the symptoms presented by patients. Our proposal involves first building models to be used for predicting diseases, and then using a llama-inspired model to generate the medical report and prescription. The chatbot system was implemented using Python, HTML, CSS and JavaScript, with FastAPI for the RESTful API and PostgreSQL for data management. The results highlight the potential of AI to improve healthcare services, especially in areas where resources are limited.

Keywords. AI, Machine Learning, Chatbot, Medical diagnosis

1. Introduction

Burkina Faso's health sector is facing major challenges as a result of rapid population growth. The shortage of doctors (about 1 doctor per 10,000 inhabitants) and the inadequacy of medical infrastructure, especially in rural areas, are key elements of this challenge. These conditions underscore the need for innovative approaches to improve disease diagnosis within healthcare services, where artificial intelligence (AI) can provide the physician with a preliminary diagnosis of the patient based on his or her symptoms, before conducting further tests for confirmation.

Several studies have shown AI's potential in health, particularly in identifying and predicting illnesses. For example, Rayan Alanazi [1] used advanced machine learning techniques such as convolutional neural network (CNN) and nearest neighbor (KNN) algorithms to automatically extract features and predict diseases, achieving an accuracy rate of 96%. Similarly, Rinkal Keniya et al [2] introduced a disease prediction system using various algorithms, with the weighted KNN algorithm achieving an impressive accuracy of 93.5%. Patil et al [3] applied machine learning to disease prediction on the basis of user-provided symptoms, with the decision tree algorithm showing the highest accuracy at 97%. Another study [4] focused on data mining machine learning techniques to improve disease prediction using models like Random Forest, Decision Tree, and LightGBM to help healthcare professionals diagnose early. Furthermore, in [5], the use

[1] Corresponding Author: Omar SAWADOGO; E-mail: sawomar226@gmail.com.

of Long Short-Term Memory (LSTM) recurrent neural networks allowed the integration of patient symptom history into the model to further improve prediction accuracy. Based on this work, we propose to develop a medical diagnostic chatbot as a promising way forward. Using symptom data collected from physicians and consultation records, this chatbot can provide a reliable initial assessment and guide medical treatment. In addition to the provision of early diagnostic information, this chatbot paves the way for reliable virtual health assistants capable of generating preliminary reports and prescriptions. The rest of this paper has the following structure: Section 2 presents our methodology for developing a medical diagnosis chatbot based on patient symptoms. Section 3 presents the results of our chatbot implementation. Section 4 concludes this article and presents some perspectives for future research.

2. Methods

The figure 1 shows our methodology for designing the medical chatbot. The method is divided into two main phases: construction of a model to predict diseases and using the model in combination with an llma2 model to generate a medical report.

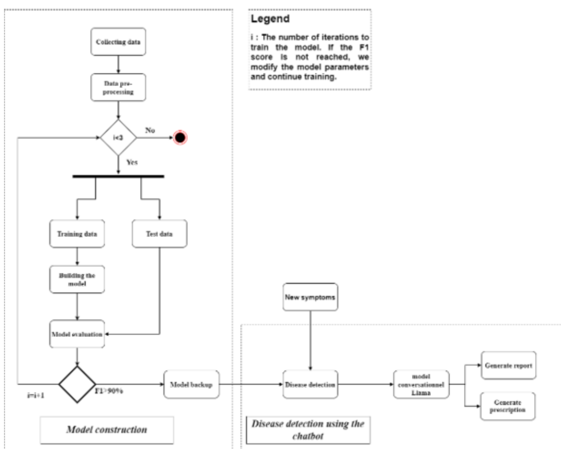

Figure 1. *Methodology flowcard.*

2.1. Construction of the model to predict disease

This step includes:

- Data collection: represents the initial step in the machine learning process. It involves gathering the necessary data to train the model. The data we collected were extracted from the patient consultation register at one of the country's healthcare centers. The nature of the collected data encompasses gender, age, dominant/constant signs/symptoms, and the diagnosis made by the physician. Our collected data span twelve diseases and include 54 symptoms.
- Data preprocessing**:** Once data have been collected, they must undergo preprocessing before being utilized to train a machine learning model. Preprocessing may involve handling missing values, detecting and addressing

outliers, converting categorical data into numerical formats, and restructuring the collected data, among other tasks. The primary goal of preprocessing is to refine the data to ensure its effective application in training a machine learning algorithm.

- Data split: After preprocessing, the data are typically partitioned into training and test sets. The training set serves to train the model, whereas the test set is employed to assess the model's performance. This segregation guarantees that the model can generalize its predictions to new, unseen data beyond the training phase.
- Model construction: At this stage, the machine learning model is constructed utilizing the training set. A variety of machine learning algorithms may be employed, including linear regression, decision trees, random forests, support vector machines (SVMs), among others.
- Model evaluation: Upon constructing the model, it is assessed utilizing the test set to measure its accuracy, robustness, and reliability. Various evaluation metrics may be applied, including precision, recall, F1 score, and accuracy, among others.
- Save model: After its evaluation, the model is preserved for future utilization. This approach enables the reuse of the model without the necessity of reconstructing it from the beginning each time

2.2. Using Model

Finally, the model is deployed to make predictions on new data. In the research we have conducted, beyond the predictions generated by our model, the outcomes are input into a conversational model, Llama2, to produce a consultation report along with prescriptions and recommendations for the patient.

3. Results

The implementation of the disease detection chatbot system utilizes various programming languages, including Python, HTML, CSS, and JavaScript. Python is predominantly employed for back-end processing, whereas HTML, CSS, and JavaScript are dedicated to constructing the chatbot's user interface. FastAPI is chosen for developing the RESTful API due to its capabilities in providing automatic documentation, type validation, and high performance. Furthermore, Jupyter Notebook and PyCharm serve as the primary tools for design and development. Lastly, PostgreSQL is adopted as the database management system for storing and managing the chatbot's data. The solution we have developed integrates a chatbot to facilitate the medical diagnosis process. Once on the platform, the chatbot guides the doctor through a structured process to gather the necessary information. The first step (Figure 2) is to collect the patient's age. If the age is valid, the chatbot then asks for the patient's gender. Checks are carried out to ensure the validity of the data entered. Error messages are displayed if necessary. Once the age and gender have been confirmed, the doctor is invited to select the patient's symptoms, as well as the results of any paraclinical examinations if available. This data is then submitted to our artificial intelligence for analysis. The model (Figure 3) performs an analysis of the symptoms presented by the patient and identifies the most likely illnesses that the patient could have, accompanied by a medical report and a medical prescription. This approach provides effective assistance with medical diagnosis, using innovative technologies to support healthcare professionals in their practice.

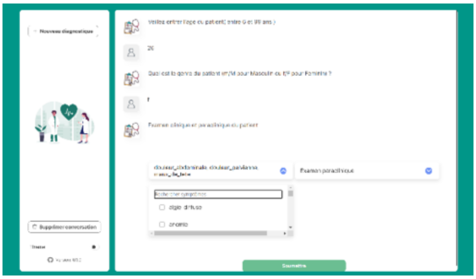

Figure 2. *Conversation between the doctor and the bot*

Figure 3. *Predicted disease and generation of the report*

4. Discussion

The model was evaluated based on metrics used to assess machine learning models. These are accuracy, precision, recall, and the F1-score. The formulas for the evaluation metrics used in assessing the performance of machine learning models are given by:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad \text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \qquad \text{F1-Score} = \frac{Precision * Recall}{Precision + Recall} \quad (2)$$

Where:

- TP (True Positives) is the number of correct positive predictions.
- TN (True Negatives) is the number of correct negative predictions.
- FP (False Positives) is the number of incorrect positive predictions.
- FN (False Negatives) is the number of incorrect negative predictions.

Overall, the comprehensive evaluation of the machine learning models based on various metrics revealed that the model trained with the ExtraTreeClassifier algorithm showed superior performance. Compared to other algorithms such as RandomForestClassifier, GradientBoostingClassifier, GaussianNB and LogisticRegression, ExtraTreeClassifier achieved better results in terms of accuracy, precision, recall and F1 score. These results suggest that the ExtraTreeClassifier model has an exceptional ability to make accurate predictions while maintaining good generalizability to unseen data, as confirmed by its cross-validation score. Consequently, in the context of future predictions, the choice of ExtraTreeClassifier is justified by its superior overall performance on all evaluated measures, providing a robust and reliable solution to our problem statement. The results highlight the potential of AI to improve healthcare services, especially in resource-constrained settings.

Table 1. Table of algorithm with borders

Algorithm	F1 score	Recall	Precision	Accuracy
RandomForest	92%	91%	96%	94%
ExtraTree	97%	96%	99%	98%
GradientBoost	90%	90%	95%	90%
GaussianNB	90%	90%	95%	90%
LogisticRegression	87%	86%	90%	88%

5. Conclusions

This paper explores the application of large language models (LLMs) to develop a medical diagnostic chatbot based on patients' symptoms. First, we developed a model that combines various algorithms such as RandomForestClassifier, ExtraTreeClassifier, GaussianNB, and LogisticRegression to predict diseases. We then integrated a LLama-inspired model to generate personalized medical reports based on the information provided by patients and the predicted diseases. The results obtained with our methodology demonstrate that the implementation of intelligent chatbots, such as the one we developed, could significantly improve the quality of healthcare and accessibility of medical services for all. In our future work, we plan to integrate association rules to further explore the correlation between diseases and symptoms. This approach could lead to a better understanding of the interactions between different symptoms, helping to refine the accuracy of the diagnosis provided by our medical chatbot.

References

[1] Alanazi R. Identification and Prediction of Chronic Diseases Using Machine Learning Approach. J Healthc Eng. 2022 Feb;2022:2826127. doi: 10.1155/2022/2826127.

[2] Keniya R, et al. Disease Prediction From Various Symptoms Using Machine Learning. SSRN Electron J. 2020. doi: 10.2139/ssrn.3661426.

[3] Arumugam K, Naved M, Shinde PP, Leiva-Chauca O, Huaman-Osorio A, Gonzales-Yanac T. Multiple disease prediction using Machine learning algorithms. Mater Today Proc. 2023;80:3682-5. doi: 10.1016/j.matpr.2021.07.361.

[4] School of Computer Science And Engineering, Vellore Institute of Technology, Vellore (Tamil Nadu), India, et al. Symptoms Based Multiple Disease Prediction Model using Machine Learning Approach. Int J Innov Technol Explor Eng. 2021 Jul;10(9):67-72. doi: 10.35940/ijitee.I9364.0710921.

[5] Gaurav K, Kumar A, Singh P, Kumari A, Kasar M, Suryawanshi T. Human Disease Prediction using Machine Learning Techniques and Real-life Parameters. Int J Eng. 2023;36(6):1092-8. doi: 10.5829/IJE.2023.36.06C.07.