IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 6, June 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906026 |

Lung Cancer Prediction using Gaussian Naive Bayes Algorithm

Varun Naik, Shaila H Koppad

Student, Department of MCA, RV College of Engineering, Bengaluru, Karnataka, India

Assistant Professor, Department of MCA, RV College of Engineering, Bengaluru, Karnataka, India

ABSTRACT: Lung cancer is the most common cause of cancer-related death worldwide.As a result, early identification, prediction, and diagnosis of lung cancer have become critical, as it speeds up and simplifies the clinical board that follows. Machine learning approaches have been used to improve the progression and treatment of malignant illnesses due to their accuracy. In this proposed system users can get fast guidance on their Lung Cancer disease using an intelligent online system. The Lung Cancer Disease Prediction web application is a project that aims to provide end-user support and online consultation. Two entities, Admin and Patient, have access to the proposed system. Admins must first log in using their valid login credentials to gain access after logging in successfully.Admins have access to all modules and can precisely perform/manage each task. Admins can control questions and answers for bots, as well as keep track of doctors' information by entering their information into the system. Admins can also manage hospital information, which is viewable by registered users. Admin has access to all of the registered users' information.Users must first register and then log in using their credentials. The lungs will be analyzed using the Machine Learning Gaussian Naive Bayes algorithm based on the user's input parameters to determine whether or not the user has lung disease. Gaussian Naïve Bayes algorithm is used here for the classification of success rate of the lung cancer prediction.

KEYWORDS: LCP: Lung Cancer, Trained data, Gaussian Naive Bayes algorithm.

This paper is divided into Introduction, literature survey, proposed model and results.

I. INTRODUCTION

Lung cancer is the deadliest disease and the leading cause of death in the modern world. Lung cancer is a type of cancer that starts in the lungs and spreads to the brain and other organs. Lung cancer is classified into two types. Lung cancer is divided into two types: non-small cell lung cancer and small cell lung cancer. The last orchestrate depicts the signs and symptoms of lung cancer. As a result, defining its first stage is challenging.Dyspnea,loss weigh, severe chest pain and dry cough are some of the symptoms that patients experience. When it comes to the development of cancer and its causes, doctors are emphasising the role of second hand smoke and smoking as major causes of lung cancer.Lung cancer is treated with surgery,immunotherapy, radiation therapy, and chemotherapy, among other things.Despite this, the lung cancer detection method is relatively ineffective because doctors will only be able to detect the disease once it has progressed to an advanced stage. As a result, early detection before the terminal stage is critical to lowering the mortality rate through effective control.

We can check the cancer patient and recommend a doctor using this system. There are two types of entities: admin and patient.Using credentials, the admin can log in. The bot's query and answer can be managed by the admin, who can also add and remove doctors.Admin can govern the hospital by adding and removing. The registered Patients can also be viewed by the Admin. Admin has the ability to provide feedback. Patients must first register and then log in using their credentials.

II. LITERATURE SURVEY

Goel, L and Bhatia, et al. In this [1] Lung cancer detection: a deep learning approach. Deep residual learning can be used to detect lung cancer from CT scans. We outline a set of preprocessing strategies for highlighting cancer-prone lung areas and extracting features with UNet and ResNet models.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906026 |

Therese, and Bhuvaneswari et al.[2] Detection of cancer in lung usinknn, classification using genetic algorithm. This focuses on lung cancer diagnosis at an early stage. For detection, using the Genetic K-Nearest Neighbour (GKNN) Algorithm to know the how nonparametric technique is implemented.

Kadir, T., Gleeson, F.[3] Lung cancer prediction using machine learning and advanced imaging techniques. In this paper tells about how the image will be classified using the machine learning technic and use of algorithm.

Lynch, C.M., et al.[4] Prediction of lung-cancer patient survival used supervised machine learning classification techniques. This study discusses how it is unclear which types of methodologies would give better predictive information in the case of lung cancer, and which data qualities should be employed to ascertain this information..

Pradeep, K and Naveen, N.[5] Healthcare analytics, lung cancer survivorship prediction based on performance utilising support vector machine classification approaches and naïve bayes algorithms. The technique for analysing clinical data and giving treatment to patients is described. The treatment is based on a review of clinical data from electronic health records.

G.A.P and Gupta, P.K et al [6] Performance analysis of machine learning based approaches for identification and classification of lung cancer . This paper tells about how the algorithm apply for predicting the lung cancer and usage of different algorithm.

Jason Brownlee [7] Naive Bayes for Machine Learning this paper tells about how Gaussian Naive Bayes algorithm implemented and their usage in prediction technique and also information about the algorithms.

Timor Kadir[8] Machine learning and advanced imaging techniques for Lung cancer detection. This paper tells about implementation of machine learning based model and how to improve the decision making for prediction.

Smita Raut1andShraddha Patil2 [9] Lung Cancer Detection Using Machine Learning Approach. This paper tells about matching the values using algorithms and feature extract from dataset and machine learning techniques.

Abbas Khosravi and Syed MoshfeqSalaken et al[10] Lung Cancer Classification in Deep Learned Features using on Low Population Dataset. This paper for how dataset can be classified using the different algorithm and performing the prediction of lung cancer.

From [1-10], the various technical details to build the system is understood and the important modules that can be developed is studied.

The next section gives the details of the proposed model.

III. PROPOSED MODEL

In this proposed model users can get fast guidance on their Lung Cancer disease using an intelligent online system in this approach. The Lung Cancer Disease Prediction web application is a project that aims to provide end-user support and online consultation. Two entities, Admin and Patient, have access to the proposed system.

Admins must login with appropriate login credentials. Once logged in, they can access all modules and accurately perform/manage each task. Admins can conduct tasks like managing training data. Admins can control bot questions and answers, and they can also keep track of doctors' information by enrolling their information into the system. Admin can also manage hospital information which is seen by registered users. Admin has access to all of the registered users' information. Users must first register and then log in using their credentials. The lungs will be analysed using the Machine Learning Gaussian Naive Bayes algorithm based on the user's input parameters to detect whether or not the user has any lung disease. Users may look for doctors and learn more about them so that they can contact them in an emergency. Users can read all of the hospital information, which will assist them in locating the nearest hospital with contact information. A chatbot will assist people in getting answers to their health-related questions via a computer system.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906026 |

Work flow

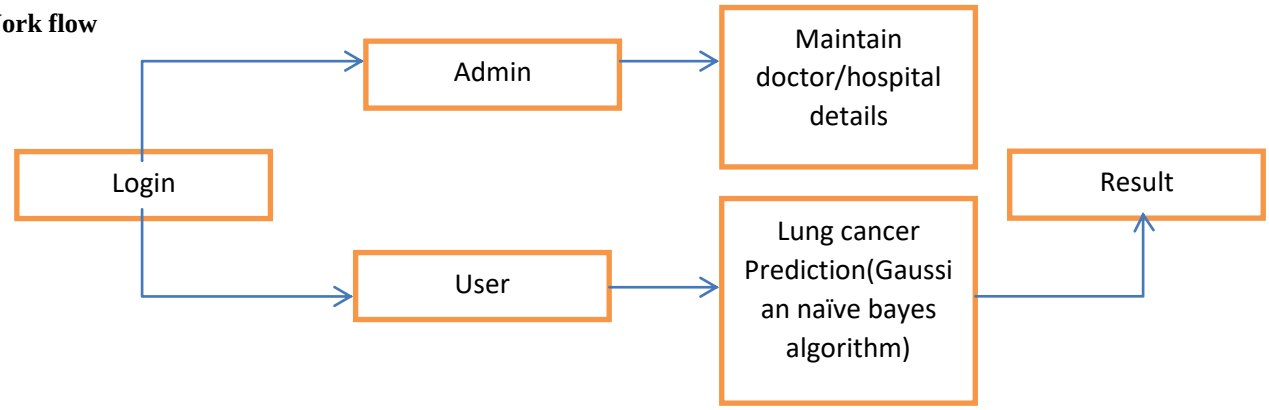

Figure: 1 Block diagram

3.1 Modules and Description

In this there are mainly two modules admin and patient

3.1.1 Admin

- **Login:** Admin can login using their credentials
- **Manage QnA for Chatbot:** The bot's health related questions and answer can be managed by the admin
- **Manage Doctors:** Admin add the doctor, and also manage the doctors detail
- **Manage Hospitals:** Admin add the hospital, and also manage the hospital details
- **View Patients:** Admin view all the registered patient details
- **Feedback:** admin view user feedback and sentiment ratings

3.1.2 Patient

- **Register:** Patient must provide basic registration information for create appropriate login credential
- **Login:** Patient valid credential used for login
- **Lungs analysis::** The lung will be analyzed using Machine Learning Gaussian Naive Bayes algorithm used on the user's input parameters to determine whether or not user has lung cancer
- **View doctors:** Patient view doctor details and contact number
- **View hospital:** Patient view hospital details and contact number.
- **Health Chatbot:** AChatbot will assist people in getting answer to their health related questions via computer programs
- **Write Feedback:**A feedback given by the registered users with sentiment analysis technique, in order to notalert the admin about feedback rating.

.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906026 |

id	Age	PersistentCough	ChestPain	Hoarseness	CoughingBlood	ShortnessofBreath	PersistentTiredness	LossOfAppetite
1	30	5	5	5	5	5	5	5
2	31	5	5	5	5	5	5	5
3	32	5	5	5	5	5	5	5
4	33	5	5	5	5	5	5	5
5	34	5	5	5	5	5	5	5
6	35	5	5	5	5	5	5	5
7	36	5	5	5	5	5	5	5
8	37	5	5	5	5	5	5	5
9	38	5	5	5	5	5	5	5
10	39	5	5	5	5	5	5	5
11	40	5	5	5	5	5	5	5
12	41	5	5	5	5	5	5	5
13	42	5	5	5	5	5	5	5
14	43	5	5	5	5	5	5	5

Figure 2: Trained Data set

The trained data set will be download from kaggle and it has 18 rows. It involve the symptoms of lung cancer patient and their records. Using this data set we will predict the patient have a lung cancer or not.

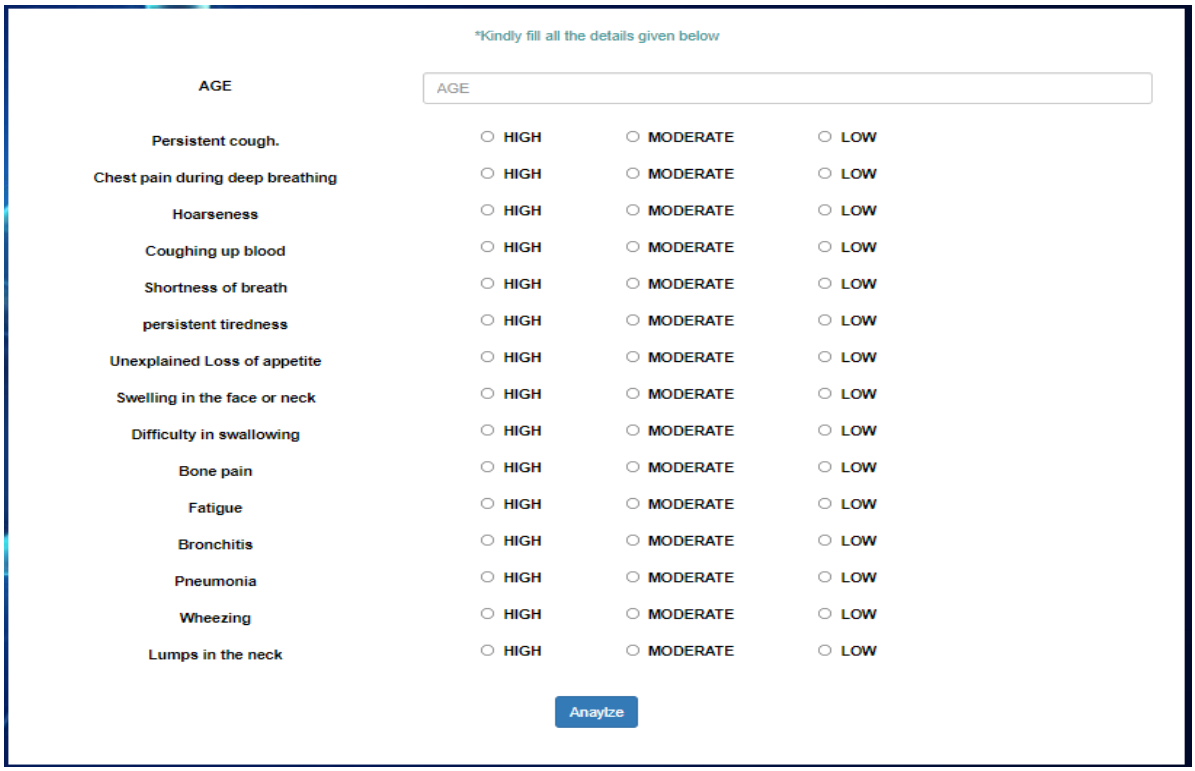

Figure 3: Lung cancer analysis

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906026 |

Figure 3 will describe that the user inputs their symptoms for predicting that he has lung cancer or not. The user must enter the 17 fields and each field have three values (High, Moderate, Low), The patient selects the details after selecting click the analyze, the values will be matched the trained data and the result will be displayed on the same window.

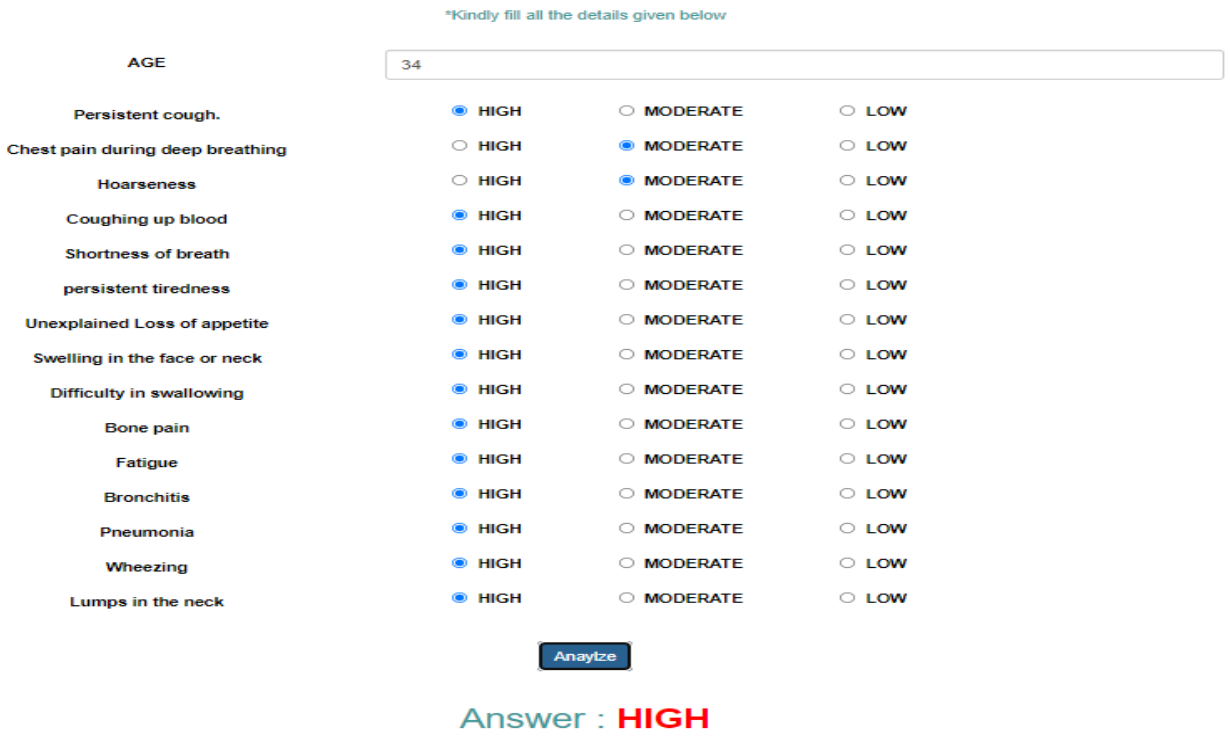

Figure 4 Lung Cancer prediction

In figure 4 the result will be displayed if it High he has lung cancer in the final stage. If it shows Moderate it will first or the second stage, it displays Low then the patient not have a lung cancer.The Gaussian Naive Bayes algorithm is used to predict lung cancer disease.Accuracy of the result depends on training data set. The proposed lung cancer disease prediction system identification system has an accuracy of 82%.

Using Gaussian Naive Bayes algorithm the accuracy rate:

Table 1. Prediction Accuracy

SL. No	Number of data use in percentage	Prediction accuracy
1	100%	82%
2	75%	85%

IV. CONCLUSION

The Gaussian Naive Bayes algorithm is used in this study to demonstrate the construction of a machine learning technique for lung cancer prediction. Lung cancer prediction application an efficient online software tool may now completely automate lung cancer prediction, which has long been considered as a mandatory burden at medical offices and healthcare facilities. The advantages of using this technology benefit everyone engaged in the scheduling process because administrators and users can execute jobs more quickly and easily.The algorithm looks for information in a historical database of lung diseases. Using patient health data to identify lung disease will aid in the long-term saving of

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906026 |

human lives. The death rate can be dramatically decreased if the condition is diagnosed early and treatment is provided as soon as possible.This system can be further improved and expanded.

REFERENCES

1. Bhatia, S., Sinha, Y., Goel, L.: Lung cancer detection: a deep learning approach. In: Bansal, J.C., Das, K.N., Nagar, A., Deep, K., Ojha, A.K. (eds.) Soft Computing for Problem Solving. AISC, vol. 817, pp. 699–705. Springer, Singapore (2019).
2. Bhuvaneswari, P., Therese, A.B.: Detection of cancer in lung with k- nn classification using genetic algorithm. Procedia Mater. Sci. 10, 433–440 (2015)
3. Kadir, T., Gleeson, F.: Lung cancer prediction using machine learning and advanced imaging techniques. Transl. Lung Cancer Res. 7(3), 304 (2018)
4. Lynch, C.M., et al.: Prediction of lung cancer patient survival via supervised machine learning classification techniques. Int. J. Med. Inform. 108, 1–8 (2017)
5. Pradeep, K., Naveen, N.: Lung cancer survivability prediction based on performance using classification techniques of support vector machines, c4. 5 and naive bayes algorithms for Procardia computer science and healthcare analytics 412–420 (2018)
6. Singh, G.A.P., Gupta, P.K.: Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. Neural Comput.Appl. 31(10), 6863–6877 (2018).
7. K. Bhatia, B. Chhabra and M. Kumar paper "Lung Cancer Prediction in machine learning technique ," published in 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2020
8. Timor Kadir Lung cancer prediction using machine learning and advanced imaging techniques June 2018Translational Lung Cancer Research 7(3):304-312 DOI:10.21037/tlcr.2018.
9. Smita Raut1, Shraddha Patil2, Gopichand Shelke3 [8] Lung Cancer Detection Using Machine Learning Approach | Volume 6 || Issue 1 || January 2021 || ISSN (Online) 2456-0774 international journal of advance scientific research
10. Syed MoshfeqSalaken, Abbas Khosravi, Amin Khatami, SaeidNahavandi, Mohammad Anwar Hosen “Lung Cancer Classification Using Deep Learned Features on Low Population Dataset” 2017 IEEE 30th Canadian Conferen

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details