Creating an Application for Lung Cancer Prediction using PEGA Tool

[1]Chenagani Venkatesh, [2]Chinthalapati Sai Prathyusha, [3]Martha Sreeja,[4]Charupalli Sunil Kumar
[1,2,3,4] Department of Electronics and Computer Engineering
Sreenidhi Institute of Science and Technology
Yamnampet, Ghatkesar, Hyderabad, India.

Abstract— **Using an app to appeal our brand to millennial users and maybe broaden our consumer base can help us grow our business. Additionally, one can give our brand a modern twist that will help us to advertise our business at a reasonable cost. In this project, creating an application using pega where patient can consult the doctor online and analyze his disease and also he can order the medicines online. Along with consulting the doctor, we predicting the lung cancer which is known as lung carcinoma, is a malignant lung tumour characterised by uncontrolled cell proliferation in the tissues of the lungs. This tumour can spread beyond the lung by metastasizing into nearby tissue or other parts of the body. In affluent countries, lung cancer is one of the main causes of mortality, and early identification is difficult. In recent decades, diagnosing and treating lung cancer has been one of humanity's most challenging undertakings. Early tumour detection will continue to save a great number of lives on a daily basis all around the world, we predicting cancer using various machine learning algorithms like KNN, decision tree, logistic regression**

Keywords—PEGA, metastasis, KNN, Decision tree.

I. INTRODUCTION

People Cancer is one of the most common causes of non-accidental mortality. Lung cancer has been proven to be the leading cause of cancer death in both man and women around the world. The death rate can be minimised if people seek early diagnosis so that professionals can provide appropriate therapy within a reasonable time frame. Cancer occurs when a group of cells undergoes uncontrollable development and loses balance, resulting in malignant tumours that infiltrate surrounding tissues. Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLLC) are two types of lung cancer (SCLC). There is a distinction between non-small cell lung cancer and small cell lung cancer in terms of diagnosis and treatment. Lung cancer classification, on the other hand, is difficult. The classification of cancer is now based on the subjective interpretation of histological and clinical data. The location of the tumor's origin also influences its classification. Clinical information can be lacking at times, and most cancers lack morphologic characteristics that are necessary for classification. Tobacco usage is the leading cause of lung cancer. More than 60% of new lung cancer cases are diagnosed in nonsmokers or former smokers. Lung cancer is uncommon before the age of 40, with rates increasing until the age of 80, after which they begin to decline.

Machine learning allows a system to solve a problem using its own learning processes. Unsupervised learning, supervised learning, and reinforcement learning are the three types of machine learning. Under the umbrella of supervised learning, two processes are identified: classification and regression. Input data is analysed and sorted into groups during the classification process. Weka was used to carry out the proposed task. In the Weka tool, algorithms such as j48, KNN, Naive Bias, and RBF are applied, and a comparative analysis is created in the end.

In this study, a Convolutional neural network is used for classifying the input image and displays the traffic sign in the form of text and alerts with a voice note.

Using the Pega platform, the suggested work predicts disease using several machine learning methods. Pega is a low-code application development platform for building applications. We're leveraging the Pega platform here, where users can anticipate their disease, arrange an appointment with a doctor, consult with him online, and get their medications directly from the app. We employ routing options here so that when a patient makes an appointment, it is routed to the doctor, who looks at his or her schedule and approves the appointment. Once the appointment is authorised, it is routed to the payment, which is then approved by the financial team. The patient communicates with the doctor over the internet, and they discuss the situation and tell him/her about his/her medical report and what precautions and medications he/she should take

II. LITERATURE SURVEY

AAsunitha, A Brindha and S Indirani propsed a lung cancer detection by using the techniques of image processing, here they developed the model using PS), Genetic optimization and SVM algorithm, they proposed the results of feature extraction and feature selection [1].

Manasee Kurkure and Anuradha Thakare proposed a automated system for lung cancer detection using evolutionary approach This optimization system enables assistant doctors to detect pulmonary nodules in their early stages. Manual interpretations take a long time and are quite important. The genetic algorithm (GA) aids in the identification of genes that aid in the classification of patient lung cancer status and has a high predictive performance [2].

Radhanaath patra proposed the prediction of lung cancer using machine learning using various classification techniques like KNN, RBF, J48, ANN and he proposed the prediction using weka tool and he found that he got about

81.25% accuracy using RBF classifer[3].

The work deals with creating an application using the requirements given by the client using pega platform which is aloww code application development [4]. By using this platform we create any type of application without writing the code. In this prediction studio is used for predicting the results by using machine learning classifiers.

The work deals with the methods like Decision Tree, Naive Bayes, and Artificial Neural Network to predict post-operative life expectancy in lung cancer patients using predictive data mining algorithms. On the aforesaid algorithms, a stratified 10-fold cross-validation comparison study was performed, and accuracy was calculated for each classifier [12].

III. METHODS USED

3.1 PEGA:
Pega is a low-code application development which is a tool used to create various applications. Pega platform consists of various studios like App studio, Dev Studio, Admin Studio and Prediction Studio [4].

App studio is used to track development changes in the context of a feature. The features of an agile-development methodology assist you in planning and implementing your application's capabilities. For example, you could add a feature to your app that allows users to request new training modules [6].

Developers can use Dev Studio to expand the App Studio-configured application behaviour. Learn how to utilise Dev Studio to customise individual rules, integrate rules into a finished user interface or process, and expand the reach of rules to other applications throughout an enterprise [7].

Pega Platform's Admin Studio is a platform where you can execute administrative and operational functions. Give people thePegaRULES:SysOpsAdministratororPegaRULES:SysOpsO bserver role to have access to Admin Studio. Admin Studio has taken the position of the System Management Application. [8].

Prediction Studio is a specialised workspace for data scientists to manage advanced analytics model life cycles. Predictive models, text models (for categorization and entity extraction), and self-learning adaptive models are among the model types accessible in Prediction Studio [9].

3.2 Logistic Regression:
The method of modelling the probability of a discrete result given an input variable is known as logistic regression. The most frequent logistic regression models have a binary outcome, which might be true or false, yes or no, and so forth. It represents true/ false by displaying binary numbers such as "0's" and "1's" . In this project '1' represents people having cancer and '2' represents people not having caner [23].

3.3 KNN:
The KNN algorithm is a sort of supervised machine learning method that may be used to solve both classification and regression predicting problems. However, in industry, it is

mostly employed to solve classification and prediction problems [19].

Knn's Pseudo Code

- Take the input.
- Consider the value of k at the start.
- Separate the data from the train and the test.
- For all training data points to achieve the required target iteration
- Calculate the distance between each row of training data and the test data. (The best method is to use Euclidean Distance.)
- Arrange the distances calculated in ascending order based on the distance values.
- Consider the sorted value's top k value.
- Locate the label for the Majority class.
- Obtain the target class

3.5 Decision tree:
Decision Tree is a supervised learning technique that may be used to solve both classification and regression problems, however it is most commonly employed to solve classification issues. Internal nodes represent dataset attributes, branches represent decision rules, and each leaf node provides the conclusion in this tree-structured classifier. [19].

IV. PROPOSED SYSTEM
The contribution of this project will be three folds First is, predicting the cancer using various classification algorithms by taking smokes, area, alcohol as the input dataset and analyzing the result using graphical format second is consulting the doctor online by booking the appointment and submitting the medical report to the doctor and knowing the present health status, third is to order the medines online if he is diagnoised with any of the health issue if needed any medical assistance doctor will suggest the patient to join the hospital or not

4.1 creating an operator ID
In this tool an operator ID is created where a casetype created belongs to that operator and adding various routing options based on the requirements [4].
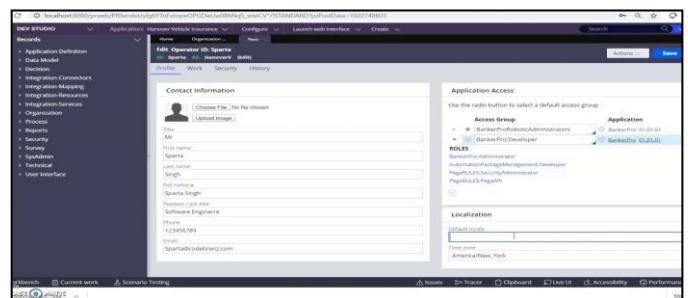

Fig 1: creating an operator

4.2 Creating a casetype
In this platform a case type is created with the requirements needed in the application. In this project the requirements for creating the casetype are patient details, booking an

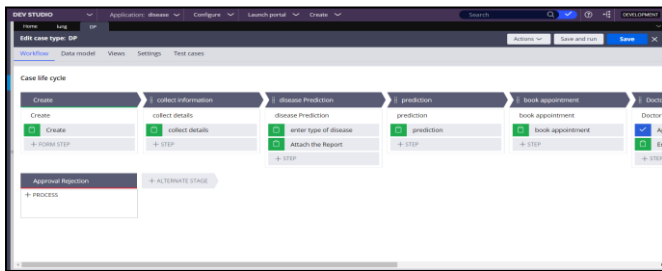appointment with doctor, analyzing the report, booking medicines online [10].
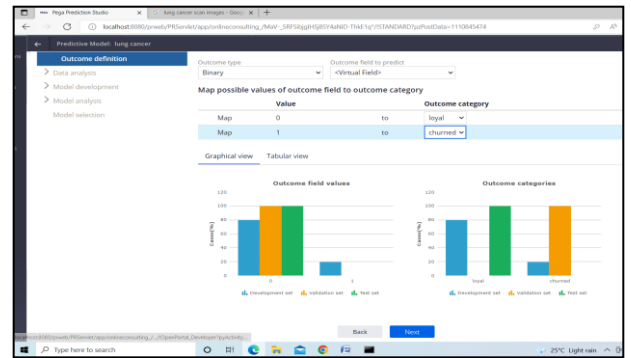

Fig 2: Creating a casetype

4.3 Routing

Routing is the process of assigning tasks to a specified user. When a patient makes an appointment, the case is routed to the doctor, who will confirm the appointment. If the patient requires any medications, the case is then routed to the pharmacist, who will then route the case to the delivery team, who will bring the medicines to the patient. patient [24].

4.4 Data Model:

The data model specifies the different types of data in your application, such as patient records and physician profiles, as well as their relationships. In our application, a data model is constructed to contain the information about the patient, the doctor, the pharmacist, and the ailment that the patient has been diagnosed with. [11]

Fig 3. Data Model creation

4.5 Creating UI:

In any application, the user interface (UI) is the point of interaction and communication between humans and computers. Display screens, keyboards, mice, and the appearance of a desktop are all examples of this. It can also refer to the way a user interacts with a programme or a website [28].

4.6 Prediction studio

In this application algorithms like logistic regression, KNN and decision tree are using in this we split the dataset and classification of dataset based on the trained test cases. We import the algorithm into prediction studio for the analysis of results

Fig 4: analysis of result

V. DESIGN FLOW

The flow chart of this project design is shown below in fig 5 and fig 6. As fig 5 shows the design of pega platform to create an application, and fig 6 shows the design flow of the cancer prediction based on the given input dataset.
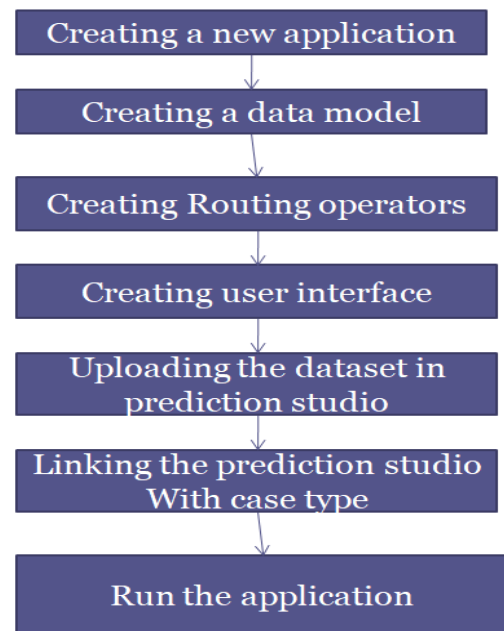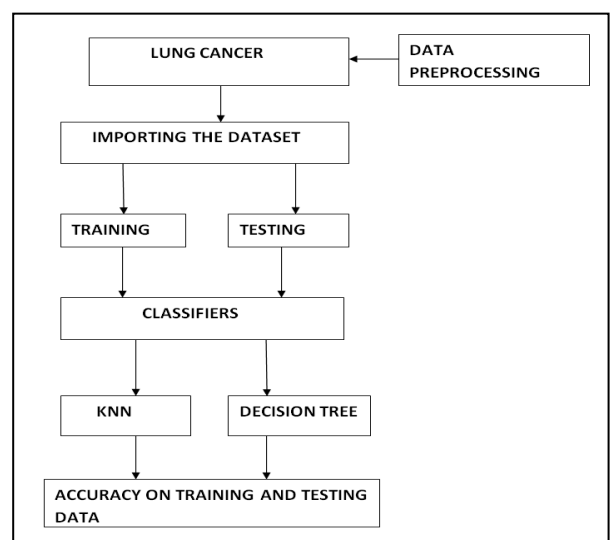

Fig 5. Pega platform design flow

Fig 6. Process flow of ML Algorithms

VI. RESULTS

After experimenting, we have come up with some results. Patient can predict his disease earlier and he can consult the doctor online by booking the appointment with doctor and he knows about his health status and if he need any medical assistant doctor will be suggest him to go to the hospital or else if it was not emergency doctor will give some medicines to the patient. Patient can order the medicines online and get the medicines to home by the delivery team.

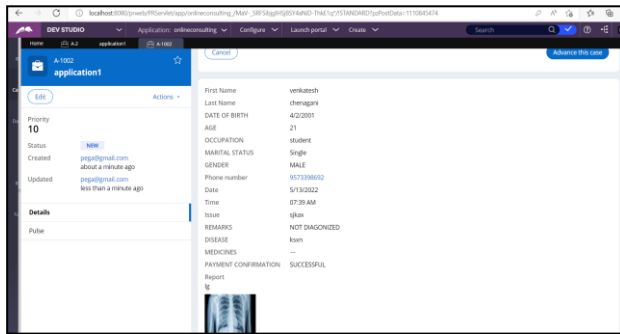In this project we are using pega platoform to create the application and also to predict the cancer
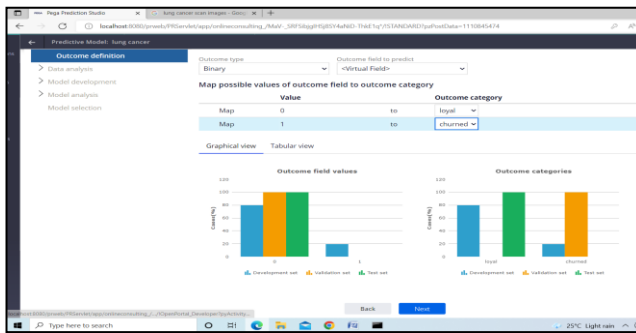

Fig 7: Details of patient

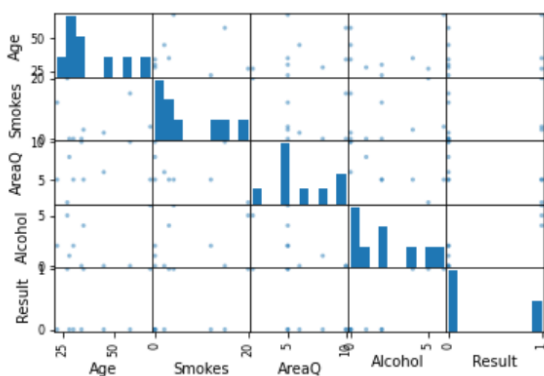Fig 8: Prediction results using prediction studio
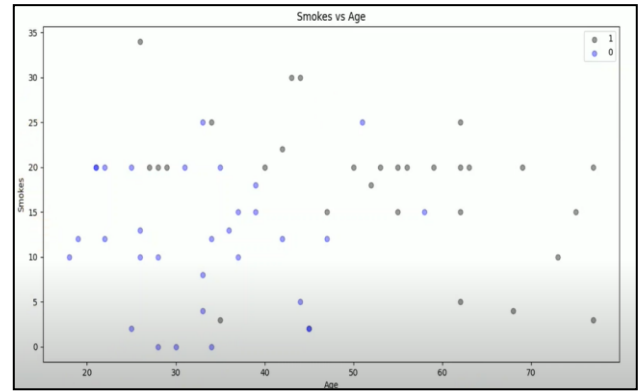

Fig 9: Classifers
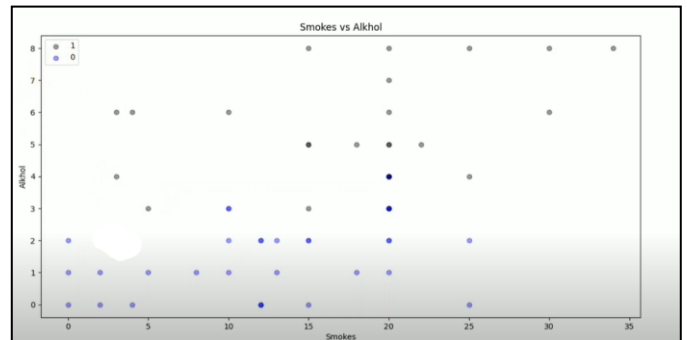

Fig 10: Smokes vs Age
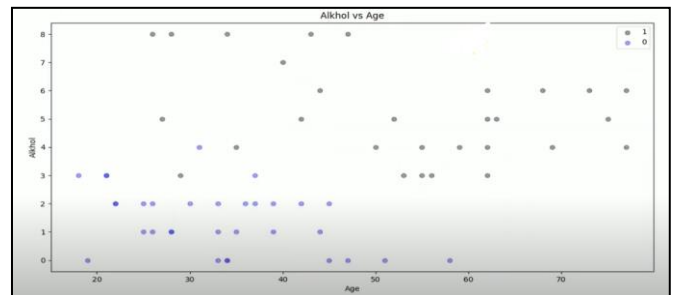

Fig 11: Smokes vs Alcohol

Fig 12: Alcohol vs Age

VII. CONCLUSION

This is an application built using pega platform which is a low-code application. This application helps the patient to predict the disease earlier and he can also consult the doctor online by booking the appointment with the doctor and analyzes his disease and he can order his medicines online and get the delivery to home by the delivery partner.

REFERENCES

[1] A.Asuntha ,A.Brindha , S.Indirani , Andy Srinivasan ,Lung cancer detection using SVMalgorithm and optimization technique Journal of Chemical and Pharmaceutical Sciences -2016.

[2] Manasee Kurkure, Anuradha Thakare ,"Introducing automatedsystem for Lung CancerDetectionusingEvolutionaryApproach",InternationalJourna lOfEngineeringAndComputerScienceISSN:2319-7242Volume 5Issue,5May2016.

[3] RadhanathPatra,"PredictionofLungCancerUsingMachineLearning Classifier",CCIS,volume1235,19 July2020.

[4] PEGAplatformforApp creationhttps://academy.pega.com

[5] Lowcodedevelopmentinpegahttps://academy.pega.com/topic/low-code-application-development/v2/in/29936/30706

[6] AppStudioinpegahttps://academy.pega.com/topic/app-studio/v1

[7]   DevStudioinpegahttps://academy.pega.com/topic/dev-studio/v1

[8]   Prediction   Studio   in
      Pegahttps://academy.pega.com/module/prediction-studio-overview/v1

[9]   Adminstudiocreationforstoringdatainpegahttps://academy.pega.com/top
      ic/admin-
      studio/v1#:~:text=Admin%20Studio%20is%20a%20workspace,config
      uration%20options%20for%20system%20resources.

[10]]Case   life   cycle   creation   for   case   study   in
      pegahttps://academy.pega.com/topic/case-life-cycle-design/v2

[11] Datamodelcreationforstoringdatainpega"https://academy.pega.com/topi
      c/data-types/v1"

[12] KwetisheJoroDanjuma,PerformanceEvaluationofMachineLearningAlg
      orithms in Post-operative Life Expectancy in the Lung Cancer
      Patients,ModibboAdama UniversityofTechnology,April2015.

[13] LinaChato, Shahram Latifi, "Machine Learning and Deep
      LearningTechniques
      toPredictOverallSurvivalofBrainTumorPatientsusingMRIImages",IEEE
      ,2017.

[14] ZehraKarhan ,TanerTunç ,"Lung CancerDetection and Classification
      withClassificationAlgorithms"IOSR-
      JCEVolume18,Issue6,Ver.IIINov.-Dec.2016.

[15] K.SenthilKumar,1K.Venkatalakshmi,and K.Karthikeyan,"LungCancer
      Detection UsingImage Segmentation by means of Various
      Evolutionary Algorithms",NCBI,8Jan2019.

[16] DecisionTreeClassificationAlgorithm,https://www.javatpoint.com/mac
      hine-learning-decision-tree-classification-algorithm

[17] K-Nearest   Neighbor(KNN)   Algorithmfor   Machine
      Learning,

[18] https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-
      learning

[19] MachineLearning-Logistic
      Regression,https://www.tutorialspoint.com/machine_learning_with_pyt
      hon/machine_learning_with_python_classification_algorithms_logistic
      _regression.htm

[20] Understanding   Logistic
      Regression,https://www.geeksforgeeks.org/understanding-logistic-
      regression/

[21] LogisticRegression   Supervised   Learning   Algorithm
      forClassification,https://www.analyticsvidhya.com/blog/2021/05/logisti
      c-regression-supervised-learning-algorithm-for-classification/

[22]   https://academy.pega.com/topic/routing-work/v1