International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 10 Issue 5, May-2021, Impact Factor: 7.957

Lung Cancer Detection Approach Using Data Mining Classifier Techniques

Nikitaben S. Prajapati[1], Khushboo S. Gandhi[2]

[1]Student of (M.E.), Computer Science and Engineering Department, Swaminarayan College of Engineering and Technology (115), Kalol, India
[2]Assistant Professor, Computer Science and Engineering Department, Swaminarayan College of Engineering and Technology (115), Kalol, India

ABSTRACT

Lung cancer by and large happens in both male and female due to wild development of cells within the lungs. This causes a genuine breathing issue in both breathe in and breathe out portion of chest. Cigarette smoking and detached smoking are the central donor for the cause of lung cancer as per world wellbeing organization. The mortality rate due to lung cancer is expanding day by day in adolescents as well as in ancient people as compared to other cancers. Subsequently it is highly necessary to require early safeguards at the introductory organize such that it's symptoms and impact can be found at early arrange for superior conclusion. In our paper we have analyzed various machine learning classifiers procedures to classify accessible lung cancer information in UCI machine learning store in to benign and dangerous. The input information is prepossessed and changed over in to double form followed by utilize of a few well known classifier strategy in python instrument to classify the information set in to cancerous and non cancerous. The comparison strategy reveals that the proposed Adaptive Booster classifier has brought about with a incredible exactness of higher accuracy than 81.25% and considered as the compelling classifier method for Lung cancer information forecast. In this research I have apply G-Booster classifier and result compare with the execution of the KNN,NB,J48 counterfeit neural arrange algorithms for execution was compared.

Keywords: KNN · ML · RBF · Lung cancer · ANN

INTRODUCTION

Lung cancer considers as the deadlier disease and a primary concern of high mortality in present world. Lung cancer affects human being at a greater extent and as per prediction it now takes 7th position in mortality rate index causing 1.5% of total mortality rate of the world [1]. Among surgical patients of early-stage lung cancer, there is a possibility of potentially advanced cancer. As such, the development of less invasive new techniques to diagnose tumor genesis of lung cancer contributes to improving the prognosis [2]. When the radiation dose increases it affects our body and once it decreases it affects the quality of image as well. So, through a comprehensive way it has developed to utilize the human healthcare one step ahead than before through computer diagnosis, thus able to detect cancer much earlier than before [3]. A better understanding of risk factors can help to prevent lung cancer disease.

The key to improve the survival rate is early detection using Machine learning techniques and if we can make the diagnosis process more efficient and effective for radiologists by using this, then it will be a key step towards the goal of improved early detection [4]. Lung cancer patients infected with SARS-CoV-2 were recently tracked in the TERAVOLT (Thoracic cancers international COVID-19 collaboration) registry study [6]. Patients with advanced NSCLC were found to have a higher risk of complications and 33% succumbed to complications from COVID-19. A smoking history was found to be an important predictor of developing complications from COVID-19[5]. These issues need to be resolved before performing the lung health condition identification else the system efficiency is less To ensure better performance of automatic lung disease identification process effectively, optimized and meta-heuristic approaches need to be introduced to analyzed further. In this manuscript, the automatic lung health condition analysis system is developed by applying the metaheuristic optimized neural networks. To maximize the system efficiency, the intelligent feature selection technique called Hilbert–Schmidt Independence Criterion based features selection process is applied. [6].

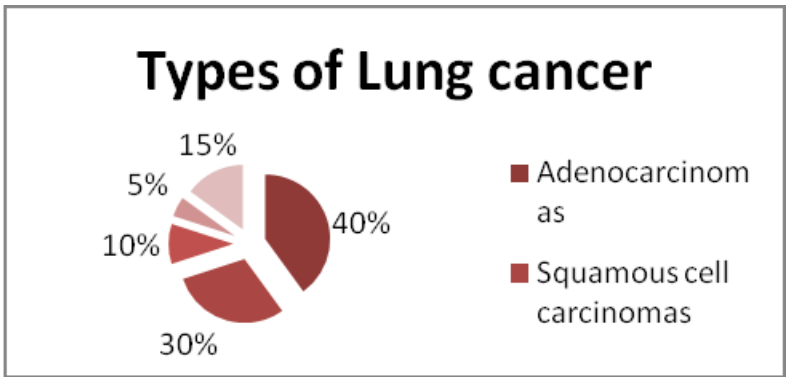

Figure 1: Types of lung cancer [7]

Table I: lung cancer cases and deaths over the globe [7]

Region	*Population*	*Cancer Cases*	*Deaths*
Asia	60%	66%	57.3%
Europe	9.0%	23.4%	20.3%
America	13.3%	21.0%	14.4%
Africa	17%	9-10%	7.3%

Table II: Lung Cancer Cases and Deaths In India [7]

Lung Cancer	*New Cases*	*Deaths*
Men	60%	66 %
Women	9.0%	23.4%
Both cases	13.3%	21.0%

RELATED WORK

Z. Zubi et al. (2014) extracted features from chest x ray images and used concept of back propagation neural network ethod to improve the accuracy [1]. Rashmee Kohad et al. (2015) used ant colony optimization with ANN and SVM to predict the accuracy of 98% and 93.2% respectively on 250 lung cancer CT images [1]. A brief description of the research papers based on Lung Cancer detection using different Machine learning algorithms are explained below:[4]

[1] Deals with the prediction of post-operative life expectancy in lung cancer patients using predictive data mining algorithms to compare algorithms such as Decision Tree, Naive Bayes and Artificial neural network..

[2]Paper deals with comparative study of classification algorithm for detection of Brain Tumor. Using volumetric and location features overall accuracy rate was calculated based on 2 classification classes such as logistic regression and Quadratic Discriminate and 3 classification classes such as Linear SVM, Coarse Gaussian SVM, Cosine KNN and Complex and median tree.

[3]In this paper, different results are produced for each classifier on the lung cancer dataset obtained.The classifiers such as KNN,SVM,NN and Logistic Regression were implemented and corresponding accuracy rates were obtained. Support Vector Machine has the highest accuracy with 99.3%.The proposed method was applied to medical dataset which helped doctors to make more correct decision.

[4]Various segmentation algorithms were discussed which includes Naïve Bayes, Hidden Markov Model etc. Proper explanation is given about how and why various segmentation algorithms are used in detection of Lung tumor.

[5]Explained about how to create a basic flowchart for an algorithm which is used to detect brain tumor. Discussed about two types of data mining techniques and there classification methods.

1.Statistical methods- Naïve Bayes, SVM
2.Data compression methods-Decision tree, Neural Network
3.Discussed about various datasets.

- BRATS Dataset
- OASIS Dataset
- NBTR Dataset

Huang et al. 2020) [24] Created the deep convolution neural network architecture for analyzing and predicting the interstitial lung disease. During this process, radiological information are collected from patients who are processed with the help of two-stage transfer learning process. The efficiency of the system is evaluated using the benchmark dataset and the system ensures 98% accuracy while classifying the lung tumor. (ALzubi et al. 2019) [26] Maximizing the lung cancer prediction accuracy using ensemble weight-optimized neural network along with maximum likelihood boosting approach. Initially, the lung images are gathered and important features are extracted from the image.

Afterward, optimized features are selected with the help of integrated Newton Raphson maximum likelihood and minimum redundancy approach. Further classification process is performed using the ensemble and boosting approach. (Zhang et al. 2015) [28] Discussing urban air pollution and the impact of the health condition in developing countries. The main intention of the paper is to examine Outdoor air pollution because most of the developing countries generate pollution in huge amounts. This air pollution leads to create a different kind of diseases in developing countries. Based on the analysis, the respective recommendation is provided to minimize air pollution and reduces health disease effectively. The data is collected in China in 2014, which is examined using the multi-level mixed effect model.[6]

From the analysis, lung disease is the most commonly affected one because the pollutants directly affect the respiratory system. So, this manuscript focuses on examining the lung health condition using the metaheuristic optimization algorithm. The introduced algorithm helps to resolve the above Sect. 1discussed issues. Finally, the efficiency of the system is evaluated using experimental results.

Almost effective image-based ML systems are developed on SVM, linear regression, decision trees, KNN and so on. [7]

SVM: For prediction, regression and classification the most prominent method employed is SVM. It classifies the input data set by introducing a boundary called a hyper plane that separates the dataset into two parts. The favorable asset of SVM is, as SVM is a data-driven approach and feasible without a hypothetical scheme that produces an accurate classification. Particularly when the size of the sample is small. SVMs are broadly used for classification when the datasets are biomarkers, to predict and diagnose cancer, neurological and cardiology diseases.

ANN: ANN is a computational method that is compromised profoundly with interconnected processing components called neurons, which sort out information as feedback to external stimuli. It is accomplished in two modes, i.e. learning and testing. Learning is used to classify new input. Attesting phase, it receives an input signal from the network and computes it to generate an output. In diversified fields of healthcare, ANN approaches are useful for example in diagnosis and indicating breast cancer, lung cancer, and other ontology prediction, diagnostic systems, drug analysis, etc.

Recurrent Neural Network (RNN): This algorithm performs a similar function for all the elements in a sequence and the output relies on the information of previous outcomes. It uses a memory that stores the information of previous outcomes.

LITERATURE REVIEW

The novelty in classification relies on mapping input function to a certain output level. Various learning classifiers are described as Perception, Naïve Bayes, Decision Tree, Logistic Regression, K nearest neighbor, Artificial Network, Support Vector Machine [1].

Algorithms	Accuracy %
KNN	70
Naive Bayes	70
RBF	76
XGBM	82
D-TREE	70

Fig: 2 Accuracy table

K nearest - Neighbors Algorithm (KNN):
Knn classifier comes under lazy learning process in which training and testing can be realized on same data or as per the programmer's choice. In the process, the data of interest is retrieved and analyzed depending upon the majority value of class label assigned.[1]

KNN is one of many (supervised learning) algorithms used in data mining and machine learning, it's a classifier algorithm where the learning is based "how similar" is a data (a vector) from other. How it's working? The KNN is pretty simple, imagine that you have a data about colored ball.

K-NN Algorithm in data mining

- Load the given data file into your program.
- Initialize the number of neighbour to be considered (i.e. value of 'K' must be odd).
- Now for each tuples (entries or data point) in the data file we perform:
- Calculate distance between the data point (tuple) to be classified and each data points in the given data file.
- Pick the first K entries from the sorted collection of data.
- Observe the labels of the selected K entries.
- For classification, return the mode of the K labels and for regression, return the mean of K labels.
- If we try to implement KNN from scratch it becomes a bit tricky however, there are some libraries like sklearn in python that allows a programmer to make KNN model easily without using deep ideas of mathematics.
- The result of research revealed that accuracy of KNN is 70%.
- Let's see code of the KNN:

```
from sklearn.neighbors import KNeighborsClassifier
neigh = KNeighborsClassifier(n_neighbors=2)
neigh.fit(X_train,y_train)
scores.append(neigh.score(X_test,y_test))
print("ACCURACY",scores[-1]*100)
precision_recall_fscore_support(y_test,neigh.predict(X_test), average='macro')
```

```
ACCURACY 70.58823529411765
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:3: DataConversionWarning:
  This is separate from the ipykernel package so we can avoid doing imports until
(0.6547619047619048, 0.6607142857142857, 0.6555555555555554, None)
```

Naive Bayes (NB):
Naive Bayes (NB) is a simple supervised function and is special form of analysis. It's a generative model and therefore returns probabilities. It's the opposite classification strategy of one Rule. All attributes contributes equally and

independently to the decision. Naive Bayes is based on bayes Theorem where Naïve Bayes classifier uses Bayesian network model p using the maximum a posteriori decision rule in Bayesian Setting. Naive Bayes works also on text categorization. NB affords fast model building and scoring and can be used for both binary and multi-class classification problems.

Bayes' theorem: Probability of event A given evidence B
Prob(A given B)=Prob(A and B)Prob(A); where:
A represents the dependent event: A target attribute
and B represents the prior event: A predictors attribute.

- The result of research revealed that accuracy of Naive Bayes is 70%.
- Let's see code of the Naive Bayes:

```
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(X_train,y_train)
scores.append(gnb.score(X_test,y_test))
print("ACCURACY",scores[-1]*100)
precision_recall_fscore_support(y_test,gnb.predict(X_test), average='macro')
```

```
ACCURACY 70.58823529411765
/usr/local/lib/python3.7/dist-packages/sklearn/naive_bayes.py:206: DataConversionWarning:
  y = column_or_1d(y, warn=True)
(0.637037037037037, 0.6666666666666666, 0.6301282051282051, None)
```

RBF Radial Basis Functions (RBF):
Radial basis function network comes under neural network that uses radial basis function as its threshold function.RBF network has advantage of easy of design and strong tolerance to input noises.

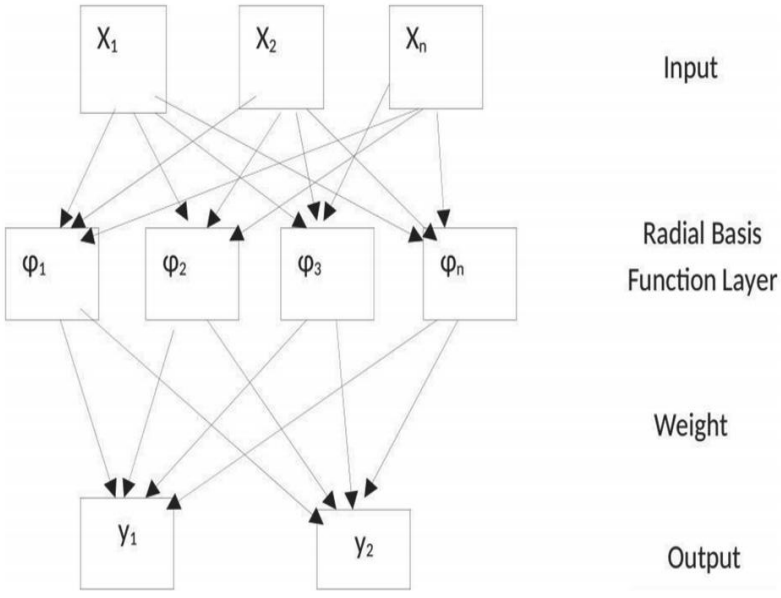

Fig: 3 Radial Basis Function Network

A simple concept behind the working of RBF is that you have to pay a small fee to change the fee for the unconfirmed transaction. Suppose you pay 1 satoshi/byte as your fee, which is the lowest fee for a transaction. Now, you pass on your transaction to the Bitcoin blockchain and wait for it to be successful. But even after hours and days, the transaction doesn't be successful and you do not have a single confirmation of the transaction.

- The result of research revealed that accuracy of RBF is 76%.
- Let's see code of the RBF:

```
from sklearn.svm import SVC
clf=SVC(kernel='rbf')
clf.fit(X_train, y_train)
scores.append(clf.score(X_test,y_test))
print("ACCURACY",scores[-1]*100)
precision_recall_fscore_support(y_test,clf.predict(X_test), average='macro')
```

```
ACCURACY 76.47058823529412
/usr/local/lib/python3.7/dist-packages/sklearn/utils/validation.py:760: DataConversionWarning:
  y = column_or_1d(y, warn=True)
/usr/local/lib/python3.7/dist-packages/sklearn/metrics/_classification.py:1272: UndefinedMetric
  _warn_prf(average, modifier, msg_start, len(result))
(0.519047619047619, 0.5773809523809524, 0.544973544973545, None)
```

eXtreme Gradient Boosting (XGBM):

XGBoost is an algorithm. It has been dominating applied machine learning. XGBoost is an implementation of gradient boosted decision trees. Although, it was designed for speed and performance. Basically, it is a type of software library. That you can download and install on your machine. Then have to access it from a variety of interfaces.

XGBoost algorithm working with Main Interfaces:

- C++, Java and JVM languages.
- Julia.
- Command Line Interface.
- Python interface along with integrated model in scikit-learn.
- R interface as well as a model in the caret package.

- The result of research revealed that accuracy of XGBoost is 82%.
- Let's see code of the XGBoost:

```
from xgboost import XGBClassifier as XGBM
model = XGBM(learning_rate=0.01,max_depth=3,n_estimators=5)
model.fit(X_train,y_train)
scores.append(model.score(X_test,y_test))
print("ACCURACY",scores[-1]*100)
precision_recall_fscore_support(y_test,model.predict(X_test), average='macro')
```

```
ACCURACY 82.35294117647058
/usr/local/lib/python3.7/dist-packages/sklearn/preprocessing/_label.py:235: Dat
  y = column_or_1d(y, warn=True)
/usr/local/lib/python3.7/dist-packages/sklearn/preprocessing/_label.py:268: Dat
  y = column_or_1d(y, warn=True)
(0.7592592592592592, 0.75, 0.7440476190476191, None)
```

Decision Tree (DT):

Decision tree uses supervised learning technique to build a model which is in the form of a tree data structure(set of nodes arranged in hierarchical fashion) .Initially, entropy of parent is calculated. Then Information gain is calculated by subtracting weighted sum of entropy of children from entropy of parent.

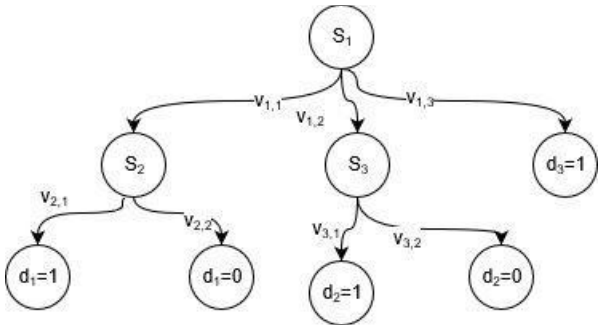

Fig.4. GENERAL VIEW OF DECISION TREE

A decision tree is a supervised learning algorithm that works for both discrete and continuous variables. It splits the dataset into subsets on the basis of the most significant attribute in the dataset. How the decision tree identifies this attribute and how this splitting is done is decided by the algorithms. In the decision tree, the dataset is divided into homogeneous and non-overlapping regions. It follows a top-down approach as the top region presents all the observations at a single place which splits into two or more branches that further split. This approach is also called a greedy approach as it only considers the current node between the worked on without focusing on the future nodes. Once a decision tree is built, many nodes may represent outliers or noisy data. Tree pruning method is applied to remove unwanted data. This, in turn, improves the accuracy of the classification model. Some of the decision tree algorithms include Hunt's Algorithm, ID3, CD4.5, and CART. Overfitting happens when a decision tree tries to be as perfect as possible by increasing the depth of tests and thereby reduces the error. This result in very complex trees and leads to overfitting. Pruning is the method of removing the unused branches from the decision tree. Some branches of the decision tree might represent outliers or noisy data.

There are two ways of pruning the tree:

1) Prepruning: In this approach, the construction of the decision tree is stopped early. It means it is decided not to further partition the branches. The last node constructed becomes the leaf node and this leaf node may hold the most frequent class among the tuples.

2) Postpruning: This method removes the outlier branches from a fully grown tree. The unwanted branches are removed and replaced by a leaf node denoting the most frequent class label. This technique requires more computation than prepruning, however, it is more reliable.

- The result of research revealed that accuracy of DT is 70%.
- Let's see code of the DT:

```
from sklearn import tree
clf1 = tree.DecisionTreeClassifier(max_depth=3)
clf1.fit(X_train,y_train)
scores.append(clf1.score(X_test,y_test))
print("ACCURACY",scores[-1]*100)
precision_recall_fscore_support(y_test,clf1.predict(X_test), average='macro')
```

```
ACCURACY 70.58823529411765
(0.5079365079365079, 0.5357142857142857, 0.5210084033613446, None)
```

PROPOSED MODEL

Data analysis processwas carried using bothweka tool of version 3.6 and Jupiter platform in python tool [1], Weka is an open source tool used for classification, clustering, regression and data visualization.Weka generally supports input file either in .csv or .arff extension format. Classificationn mostly occurs with splitting of 80% trainning and 20% testing [6]. But In Weka tool our analysis process was carried out with 10 fold cross validation with selected classifier technique to obtain an output of interest [8]. Weka is a user friendly visualization tool with we have tested various classifier technique and its output performance.[1]

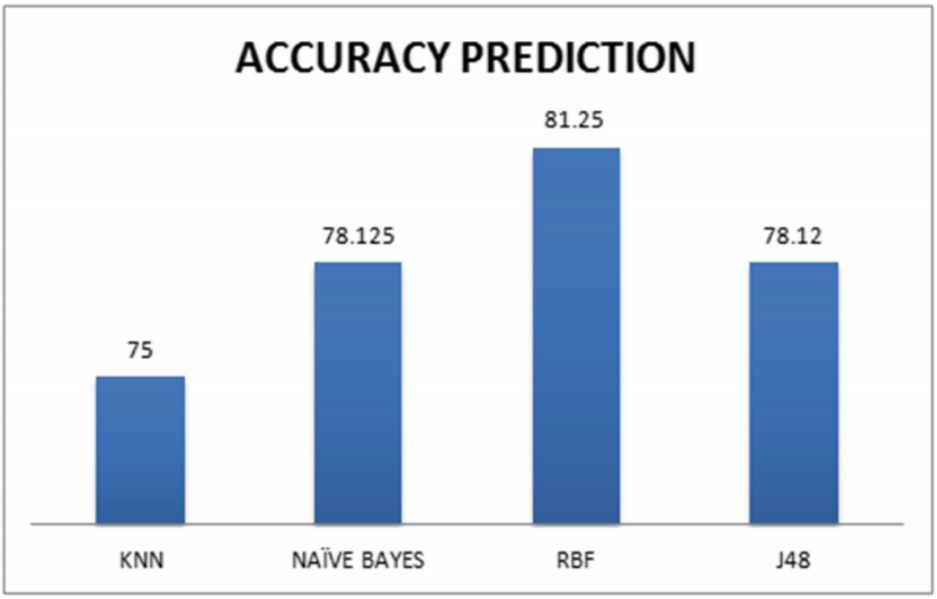

Fig. 5. Accuracy graph.

Positive and False Negative both have a value of 3 each. The output result of various classifiers used in Weka tool on lung cancer data is represented in below table. Generally in confusion matrix Accuracy, Recall, Precision and F-Measure are the key process parameter for classification [4]. Classification accuracy is the measure of number of correct prediction made out from total number of prediction. These parameters depend on some specific outcome[1].

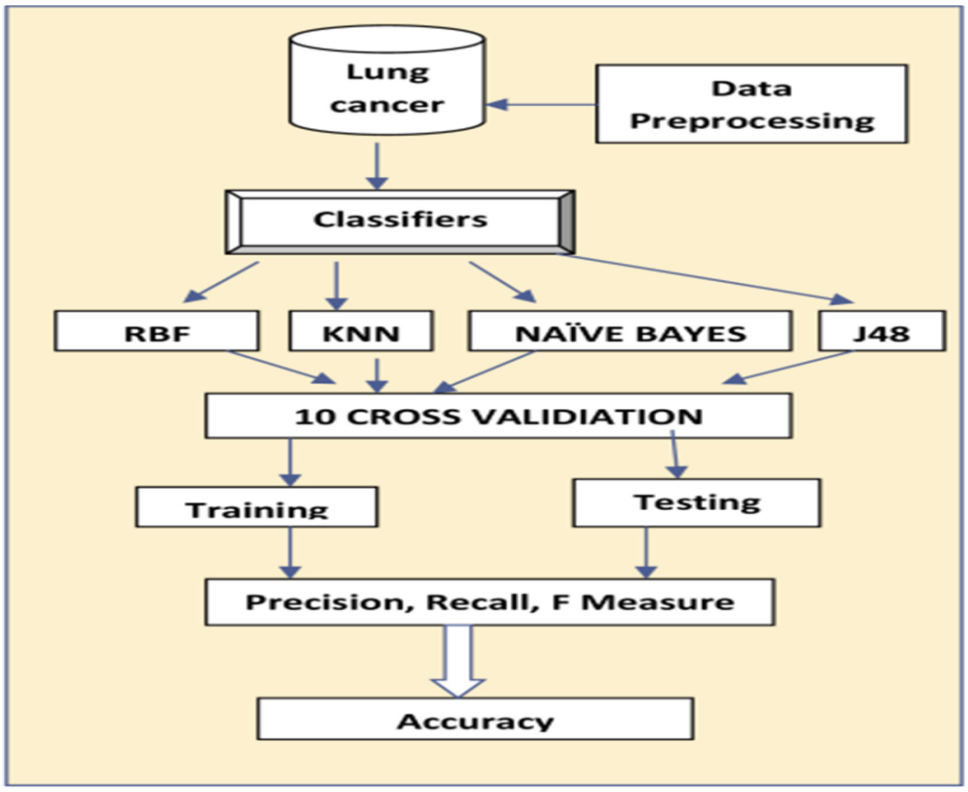

Fig. 6. Process flow of various classifiers in Weka tool.

```
y_train.hist()
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f54cc7a8f90>
```

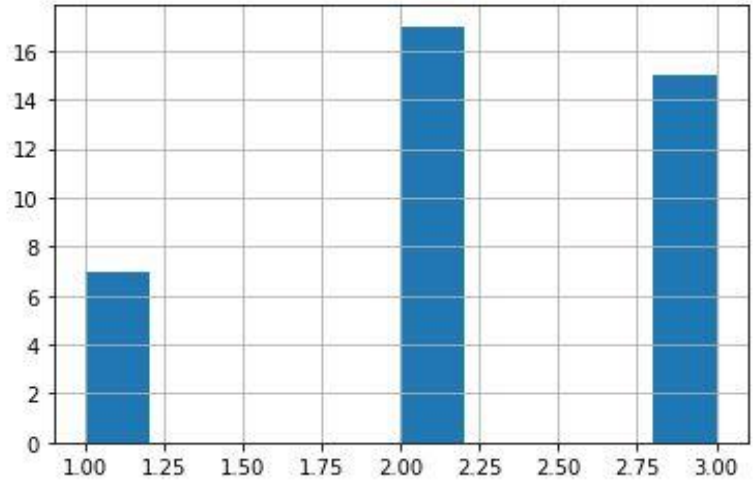

Fig: 7 Class Imbalance problem

```
y_train.hist()
```

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f54cc6b90d0>]],
      dtype=object)
```

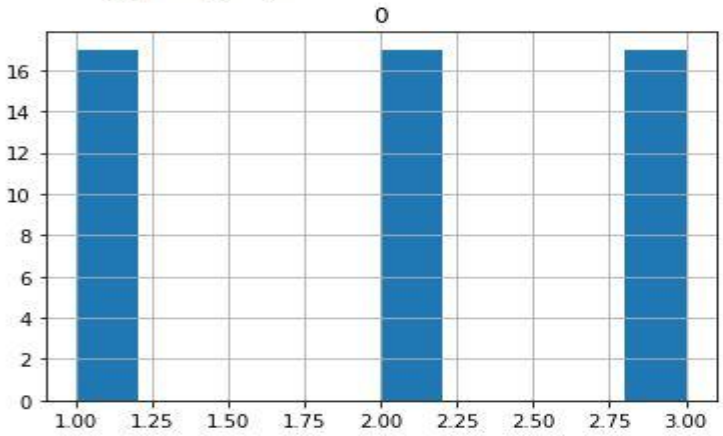

Fig: 8 Class Imbalance problem Solver

CLASS IMBALANCE PROBLEM SOLVER:

```
from imblearn.over_sampling import RandomOverSampler, SMOTE
import numpy as np
from sklearn.metrics import precision_recall_fscore_support
sm = SMOTE(random_state = 42)

X_res, y_res = sm.fit_resample(X_train, y_train)

X_train = pd.DataFrame(np.array(X_res))
y_train = pd.DataFrame(y_res)
X_test = pd.DataFrame(np.array(X_test))
```

PREDICTIONS:

```
#prediction
print("y test",y_test.values)
print("Naive Bayes",gnb.predict(X_test))
print("KNN",neigh.predict(X_test))
print("RBF",clf.predict(X_test))
print("XGBM",model.predict(X_test))
print("D-TREE",clf1.predict(X_test))
```

```
y test [2 2 3 2 2 3 1 3 1 2 3 2 2 2 3 3 3]
Naive Bayes [2 1 3 2 2 3 2 3 1 1 3 3 2 3 3 3 3]
KNN [2 2 3 2 2 3 2 3 1 2 2 3 3 1 3 3 3]
RBF [2 2 3 2 2 2 2 3 2 2 3 3 2 2 3 3 3]
XGBM [2 2 3 2 2 3 3 3 1 2 3 3 1 2 3 3 3]
D-TREE [2 2 3 2 2 3 2 2 2 2 3 3 1 2 3 3 3]
```

CONCLUSION

In this paper we have shown that with RBF classifier the accuracy is found to be 81.25% on lung cancer data. So In the analysis it can be predicted that with suitable feature selection method and integrated approach with other supervised learning process and modified functional approach in RBF, accuracy will be further improved. As a future work, along with these techniques stage identification of lung cancer also helps to predict the stage of cancer the patient is suffering from. In earlier times, the doctor has to do multiple tests in order to detect whether a given patient has lung cancer or not . But this was a very time consuming process. In a diagnosis sometimes a patient has to undergo unnecessary check-ups or different tests to identify the disease of lung cancer.

REFERENCES

[1]. Radhanath Patra, "Prediction of Lung Cancer Using Machine Learning Classifier" Electronics Science, Berhampur University, Berhampur, Odisha, India.

[2]. Noriyoshi Sawabata1,2 · Yoshiyuki Susaki1 · Takahito Nakamura1 · Takeshi Kawaguchi2 · Motoaki Yasukawa2 ·Shigeki Taniguchi2, "Cluster circulating tumor cells in surgical cases of lung cancer" 24 December 2019 / Accepted: 29 January 2020 The Japanese Association for Thoracic Surgery 2020.

[3]. Varsha Prakash, "Survey on Lung Cancer Detection Techniques" Dept. of Computer Science and Engineering, LBS Institute of Technology for Women, Trivandrum, India.

[4]. Radhika P R, Rakhi.A.S.Nair "A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms" Department of Computer Science and Engineering Department of Computer Science and Engineering Amrita Vishwa Vidyapeetham, Amritapuri, India.

[5]. Mariam Alexander1 · So Yeon Kim2 · Haiying Cheng2, "Management of Non‐ Small Cell Lung Cancer" Springer Science+Business Media, LLC, part of Springer Nature 2020.

[6]. Ahed Abugabah1 • Ahmad Ali AlZubi2 • Feras Al-Obeidat1 • Abdulaziz Alarifi2 • Ayed Alwadain, "Data mining techniques for analyzing healthcare conditions of urban space-person lung using meta-heuristic optimized neural networks" Springer Science+Business Media, LLC, part of Springer Nature 2020.

[7]. Syed Saba Raoof, M A. Jabbar, Syed Aley Fathima, "Lung Cancer Prediction using Machine Learning: A Comprehensive Approach" Department of Computer Science & Engineering Vardhaman College of Engineering Hyderabad, India.

[8]. Alcantud, J.C.R., Varela, G., Santos-Buitrago, B., Santos-Garcia, G., Jimenez, M.F.: Analysis of survival for lung cancer resections cases with fuzzy and soft set theory in surgical decision making. PLoS ONE 14(6), e0218283 (2019).

[9]. Bhuvaneswari, P., Therese, A.B.: Detection of cancer in lung with k- nn classification using genetic algorithm. Procedia Mater. Sci. 10, 433–440 (2018).

[10]. Jakimovski, G., Davcev, D.: Using double convolution neural network for lung cancer stage detection. Appl. Sci. 9(3), 427 (2019).

[11]. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.

[12]. Desitter I, Guerrouahen BS, Benali-Furet N, Wechsler J, Jänne PA, Kuang Y, et al. A new device for rapid isolation by size and characterization of rare circulating tumor cells. Anticancer Res. 2011;31:427–41.

[13]. Bottos A, Hynes NE. Cancer: staying together on the road to metastasis. Nature. 2014;514(7522):309–10. olly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. Oncotarget. 2016;7:27067–84.

[14]. Chunran, Yang, Wang Yuanvuan, and Guo Yi. "Automatic Detection and Segmentation of Lung Nodule on CT Images." 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE, 2018.

[15]. Li, Yunfei,"Lung Nodule Segmentation Using Pleural Wall Shape." 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS). IEEE, 2018.

[16]. Deshmukh, Shraddha, and Swati Shinde. "Diagnosis of lung cancer using pruned fuzzy min-max neural network." 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT). IEEE, 2016.

[17]. M. Tanino, H. Takizawa and S. Yamamoto, "A detection method of ground glass opacities in chest X-ray CT images using automatic clustering techniques".

[18]. H. Tang, D. R. Kim, and X. Xie, ''Automated pulmonary nodule detection using 3D deep convolutional neural networks,'' in Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI), Washington, DC, USA, Apr. 2018, pp. 523–526.

[19]. P. Monkam, S. Qi, M. Xu, H. Li, F. Han, Y. Teng, and W. Qian, ''Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification in CT images,'' IEEE Access, vol. 7, pp. 5564–5576, 2018. doi: 10.1109/ACCESS.2018.2889350.

[20]. B. Wu, Z. Zhou, J. Wang, and Y. Wang, ''Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction,'' in Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI), Apr. 2018, pp. 1109–1113.

[21]. Huang, Wenkai, and Lingkai Hu. "Using a Noisy U-Net for Detecting Lung Nodule Candidates." IEEE Access 7 (2019): 67905-67915.

[22]. Moradi, Pouria, and Mansour Jamzad. "Detecting Lung Cancer Lesions in CT Images using 3D Convolutional Neural Networks." 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA). IEEE, 2019.

[23]. Survey of Intelligent Methods for Brain Tumor Detection-IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 5, No 1, September 2014

[24]. Ada, RajneetKaur, " A Study of Detection of Lung Cancer Using Data Mining Classification Techniques " International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3, March 2013.

[25]. Survey of Intelligent Methods for Brain Tumor Detection-IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 5, No 1, September 2014

[26]. D. Vinitha, Dr.Deepa Gupta, and Khare, S., "Exploration of Machine Learning Techniques for Cardiovascular Disease", Applied Medical Informatics, vol. 36, pp. 23–32, 2015.

[27]. Sathyan H, Panicker,J.V,,"Lung Nodule Classification Using Deep ConvNets on CT Images",9th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2018

[28]. Tang MS, Wu XR, Lee HW, Xia Y, Deng FM, Moreira AL, Chen LC, Huang WC, Lepor H (2019) Electronic-cigarette smoke induces lung adenocarcinoma and bladder urothelial hyperplasia in mice. Proc Natl Acad Sci USA 116(43):21727–21731.

[29]. Midha A, Dearden S, McCormack R (2015) EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). Am J Cancer Res 5(9):2892–2911 Lung Cancer Symptoms, Types, Causes, Treatment & Diagnosis, https://www.medicinenet.com/lung_cancer/article.html.

[30]. ICMR-National Institute of Cancer Prevention & Research (ICMRNICPR). The International Agency for Research on Cancer (IARC). Latest global cancer data.

[31]. M.Gomathi, Dr.P.Thangaraj. 2010. A Computer-Aided Diagnosis System for Detection of Lung Cancer Nodules using Extreme Learning Machine. International Journal of Engineering Science and Technology. Vol. 2(10), 2010, 5770-5779.

[32]. Awai, K., Murao, K., Ozawa, A., Komi, M., Hayakawa, H., Hori, S., & Nishimura, Y. 2004. Pulmonary Nodules at Chest CT: Effect of Computer-aided Diagnosis on Radiologists' Detection Performance. Radiology, 230(2), 347352. doi:10.1148/radiol.2302030049.

[33]. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60– 88. doi:10.1016/j.media.2017.07.005.

[34]. M. Havaei., A. Davy., D. Warde-Farley., A. Biard., A. Courville., Y. Bengio., C. Pal, P.-M. Jodoin, H. Larochelle. 2017. Brain Segmentation With Deep Neural Networks. Medical image analysis, 35 (2017) 18-31.

[35]. H. Chen., Q. Dou., X. Wang., J. Qin & P.-A. Heng. 2016. Mitosis Detection in Breast Cancer Histology Images via Deep Cascaded Networks, AAAI2016, pp. 1160-1166.

[36]. W. Li., F. Jia & Q. Hu. 2015. Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. Journal of Computer and Communications, 3 (2015) 146.

[37]. Z. Tian., L. Liu & B. Fei. 2017. Deep Convolutional Neural Network for Prostate MR Segmentation. SPIE Medical Imaging, International Society for Optics and Photonics2017, pp. 101351L-101351L-101356.

[38]. Alsiddiky, A., Awwad, W., Bakarman, K., Fouad, H., Mahmoud, N.: Magnetic resonance imaging evaluation of vertebral tumor prediction using hierarchical hidden Markov random field model on Internet of Medical Things (IOMT) platform. Measurement 159, 107772 (2020).

[39]. Mahmoud, N.M., Fouad, H., Alsadon, O., Soliman, A.M.: Detecting dental problem related brain disease using intelligent bacterial optimized associative deep neural network. Cluster Comput. (2020).

[40]. Al Kheraif, A.A., Alshahrani, O.A., Al Esawy, M.S.S., Fouad, H.: Evolutionary and Ruzzo-Tompa optimized regulatory feedback neural network based evaluating tooth decay and acid erosion from 5 years old children. Measurement 141, 345–355 (2019).

[41]. Huang, S., Lee, F., Miao, R., et al.: A deep convolutional neural network architecture for interstitial lung disease pattern classification. Med. Biol. Eng. Comput. (2020).

[42]. Selvanambi, R., Natarajan, J., Karuppiah, M., et al.: Lung cancer prediction using higher-order recurrent neural network based on glowworm swarm optimization. Neural Comput. Appl. (2018).

[43]. ALzubi, J.A., Bharathikannan, B., Tanwar, S., Manikandan, R., Khanna, A., Thaventhiran, C.: Boosted neural network ensemble classification for lung cancer disease diagnosis. Appl. Soft Comput. 80, 579–591 (2019).

[44]. Zhang, J., Day, D.: Urban air pollution and health in developing countries. In: Nadadur, S., Hollingsworth, J. (eds.) Air Pollution and Health Effects Molecular and Integrative Toxicology. Springer, London (2015).

[45]. Liu, C., Ma, Q., Xu, J.: Multi-label feature selection method combining unbiased Hilbert-Schmidt independence criterion with controlled genetic algorithm. In: Cheng, L., Leung, A., Ozawa, S. (eds.) Neural Information Processing. ICONIP 2018. Lecture Notes in Computer Science. vol. 11304, Springer, Cham (2018).

[46]. Baskar, S., Shakeel, P.M., Kumar, R., Burhanuddin, M.A., Sampath, R.: A dynamic and interoperable communication framework for controlling the operations of wearable sensors in smart healthcare applications. Comput. Commun. 149, 17–26 (2020).