

Abstract

In every nation, lung cancer is one of the major causes of death for both men and women. Because of its poor prognosis, lung cancer has a high mortality rate. With the use of picture recognition and data analytics, the computing sector is entirely automating it, and the medical sector is doing the same.

In order to detect lung cancer at an early stage and perhaps save many lives, this research aims to examine the accuracy ratio of two classifiers, Random Forest (RF) and Naive Bayes (NB). Basically, the informational indexes utilized as a part of this examination are taken from UCI datasets for patients affected by lung cancer. This paper's main focus is on the execution analysis of the classification algorithm's accuracy. The experimental results show that RF gives the best result with 97% and NB with 94%.

Introduction

When compared to other cancers, lung cancer is the second most frequent, affecting one in five men and one in nine women. Unfortunately, during the past few years, while lung cancer incidence has slowly decreased in males, it has been rapidly increasing in women. In 1940, just seven women out of every 100,000 acquired the illness; now, that number is 42. The cause of smoking is smoking, according to all the data. The number of cigarettes you smoke each day affects how long it takes to get cancer, according to one expert in the subject. However, research shows that giving up smoking does reduce the risk.

Small cell lung cancer (SCLC), sometimes known as oat cell cancer due to the cells' resemblance to oat grains, and non-small cell lung cancer are the two main kinds of lung cancer (NSCLC). The type of tumor found determines the disease's severity and available treatments. The importance of early discovery and timely treatment—typically surgery to remove the tumor—is underscored by the fact that many kinds of lung cancer develop and spread quickly and that the lungs are essential organs.

Related work

- [1] Y. Xie *et al.*, “Early lung cancer diagnostic biomarker discovery by machine learning methods,” *Translational Oncology*, vol. 14, no. 1, Jan. 2021, doi: 10.1016/j.tranon.2020.100907.
- [2] G. A. P. Singh and P. K. Gupta, “Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans,” *Neural Computing and Applications*, vol. 31, no. 10, pp. 6863–6877, Oct. 2019, doi: 10.1007/s00521-018-3518-x.
- [3] Dayananda Sagar College of Engineering, Institute of Electrical and Electronics Engineers. Bangalore Section, and Institute of Electrical and Electronics Engineers, *2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020) : conference proceedings : 5-7 March, 2020*.
- [4] C. H. Huang *et al.*, “A study of diagnostic accuracy using a chemical sensor array and a machine learning technique to detect lung cancer,” *Sensors (Switzerland)*, vol. 18, no. 9, Sep. 2018, doi: 10.3390/s18092845.

- [5] N. Banerjee, "Prediction Lung Cancer-In Machine Learning Perspective," 2020.
- [6] M. Imran Faisal, S. Bashir, Z. Sikandar Khan, and F. Hassan Khan, "An Evaluation of Machine Learning Classifiers and Ensembles for Early Stage Prediction of Lung Cancer."
- [7] D. M. Abdullah, "Lung cancer Prediction and Classification based on Correlation Selection method Using Machine Learning Techniques", doi: 10.48161/Issn.2709-8206.
- [8] "AN EXTENSIVE REVIEW ON LUNG CANCER DETECTION USING MACHINE LEARNING TECHNIQUES," *Journal of critical reviews*, vol. 7, no. 14, Jul. 2020, doi: 10.31838/jcr.07.14.68.
- [9] Q. Gu *et al.*, "Machine learning-based radiomics strategy for prediction of cell proliferation in non-small cell lung cancer," *European Journal of Radiology*, vol. 118, pp. 32–37, Sep. 2019, doi: 10.1016/j.ejrad.2019.06.025.
- [10] J. Pati, "Gene expression analysis for early lung cancer prediction using machine learning techniques: An eco-genomics approach," *IEEE Access*, vol. 7, pp. 4232–4238, 2019, doi: 10.1109/ACCESS.2018.2886604.
- [11] J. M. Luna *et al.*, "Predicting radiation pneumonitis in locally advanced stage II–III non-small cell lung cancer using machine learning," *Radiotherapy and Oncology*, vol. 133, pp. 106–112, Apr. 2019, doi: 10.1016/j.radonc.2019.01.003.