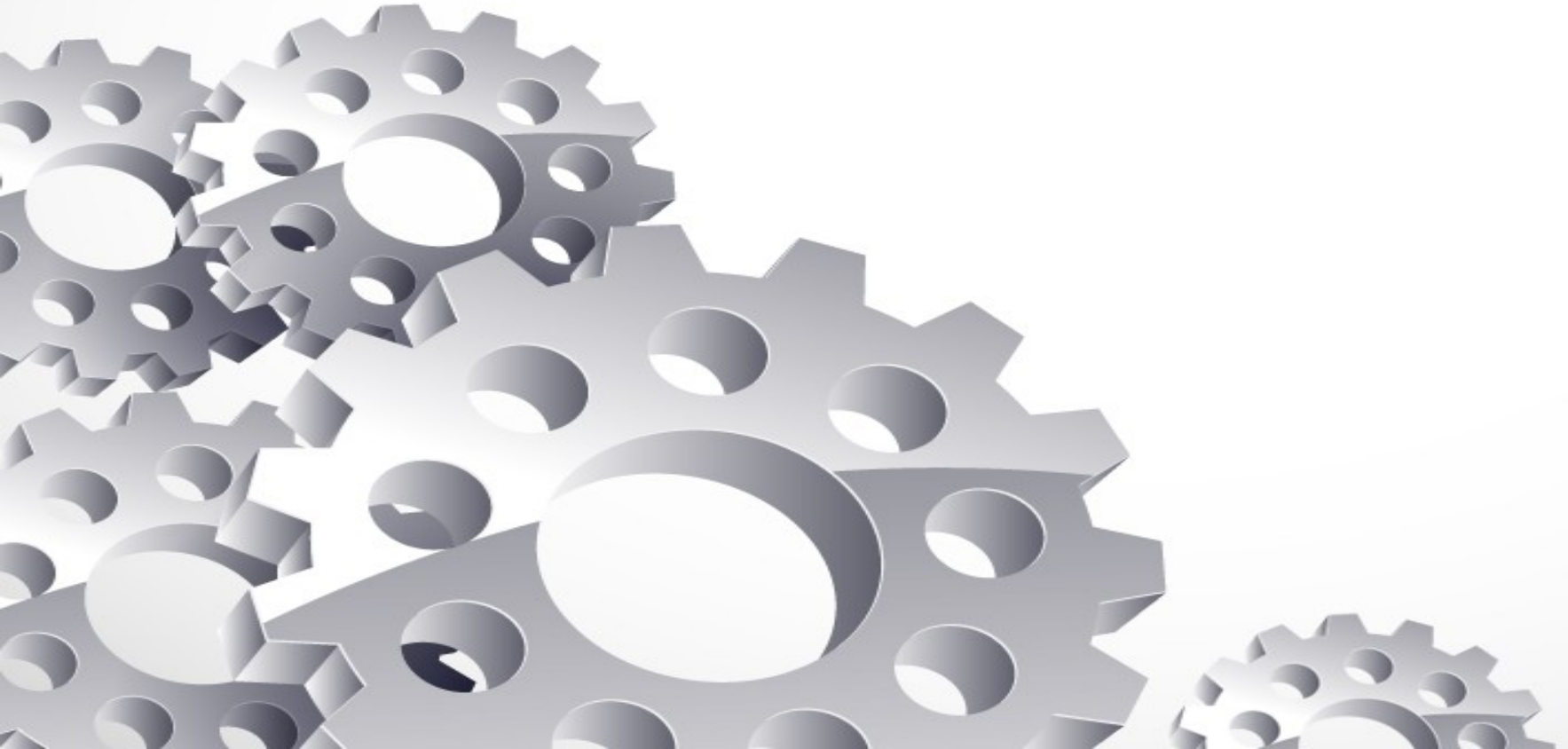
Lung Cancer Recognition Using CT Scan with NCA-XGBoosting & KNN



Group members information and Contribution

Person	Student ID	Contribution	Task
Likitha Dara	700743525	25%	Programming and development
Ayesha Farhana	700735341	25%	Data Management
Sabrina Shaik	700732583	25%	Programming and development
Prudhvi Mahesh M	700738978	25%	Research

Motivation

- Lung cancer is the most common type of disease among men, and it is the third most common type of malignancy among women.
- The early detection of lung cancer is critical to reducing the impact of the disease's death rate throughout the world. Because the symptoms of lung cancer are not recognized until the disease has progressed to an advanced stage, it is critical to predicting the disease at an earlier stage using any medical imaging techniques available.
- There is currently no specific type of diagnosis method available to prevent lung cancer in the clinical setting.
- The earlier and more rapid detection of cancer is a significant factor in the diagnosis, and this will increase the survival rate of cancer patients as a result of this.

Objectives

- Lung cancer can be diagnosed and detected using a variety of methods, including blood tests, radiology tests, endoscopic procedures, and biopsies.
- Using CT (Computerized Tomography) scanning, you can get a quick test result without any discomfort, and you can get information about the tumor's shape, size, and location.
- It is a three-dimensional image of the inside of the body created by an x-ray machine that takes multiple images of the same anatomical location from various angles.
- The details in the lung are more clearly visible as a result of this. A CT scan of the patient's chest provides detailed images of the patient's chest, allowing for more accurate detection of lung cancer.

Related work

Using machine learning methods, Reddy et al. [40] offer a model that can accurately diagnose the stages of lung cancer. The model incorporates the bagging ensemble technique to boost the prediction accuracy of K-NN, Decision Trees, and Neural Network topologies. Estimated results using the suggested model are more precise than those obtained using separate methods. Results from both the bagged and unbagged variants are compared. The accuracy ratings of 97% (Decision Tree), 94% (K-NN), and 96% (SVM) are all enhanced using the bootstrap aggregation approach (Neural Networks). The overall accuracy of the combined model is 0.98. The combined model is improved in accuracy by 3.33 percent.

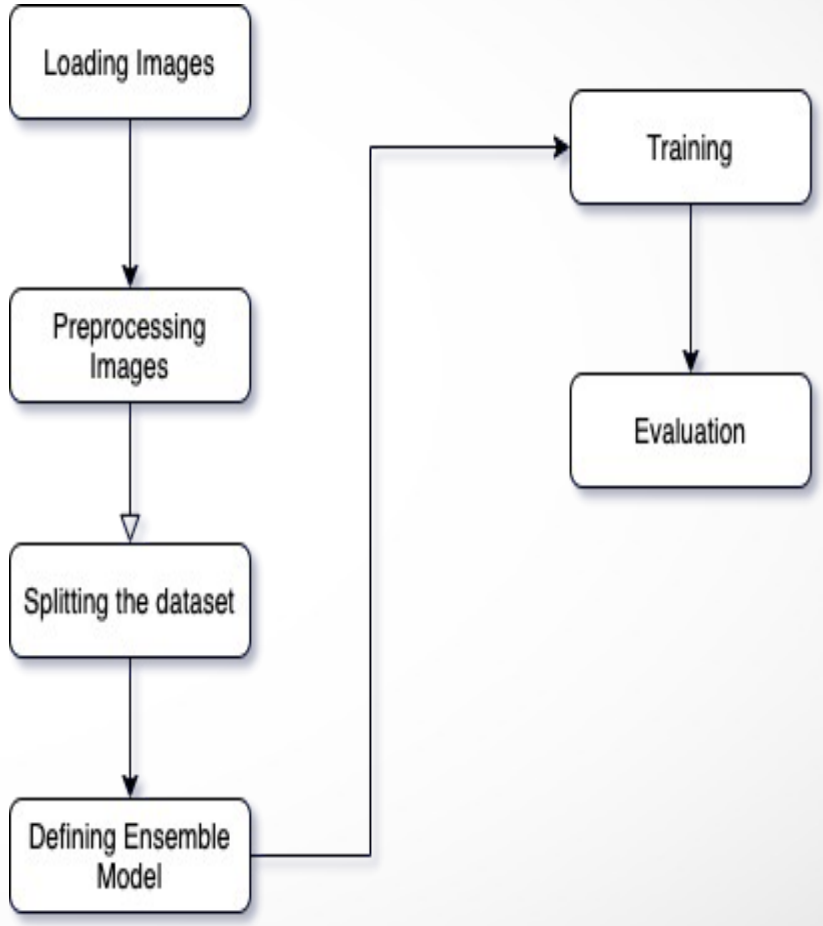
Machine learning classifiers, such as Multilayer perceptron (MLP), Nave Bayes, Decision Tree, Neural Network, Gradient Boosted Tree, and SVM, are recommended for use in the diagnosis of lung cancer by Faisal et al [12]. The UCI registry data is used to examine ensembles built using random forests and plurality voting to make predictions about lung cancer. All other individual and ensemble classifiers were determined to be inferior than Gradient Boosted Tree. When compared to other classifiers and ensemble methods, Gradient-boosted Tree had the highest accuracy (90%) in performance evaluations.

Problem statement

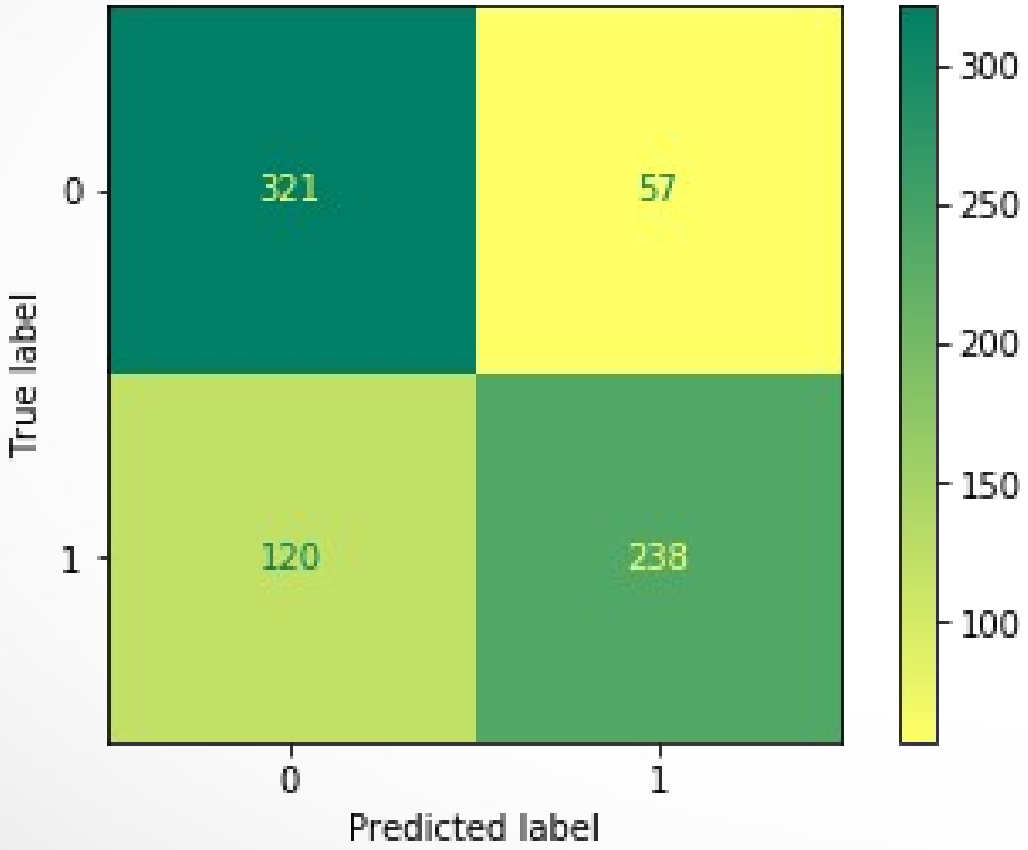
- The early detection of lung cancer is critical to reducing the impact of the disease's death rate throughout the world. Because the symptoms of lung cancer are not recognized until the disease has progressed to an advanced stage.
- It is critical to predicting the disease at an earlier stage using any medical imaging techniques available.

Proposed solution

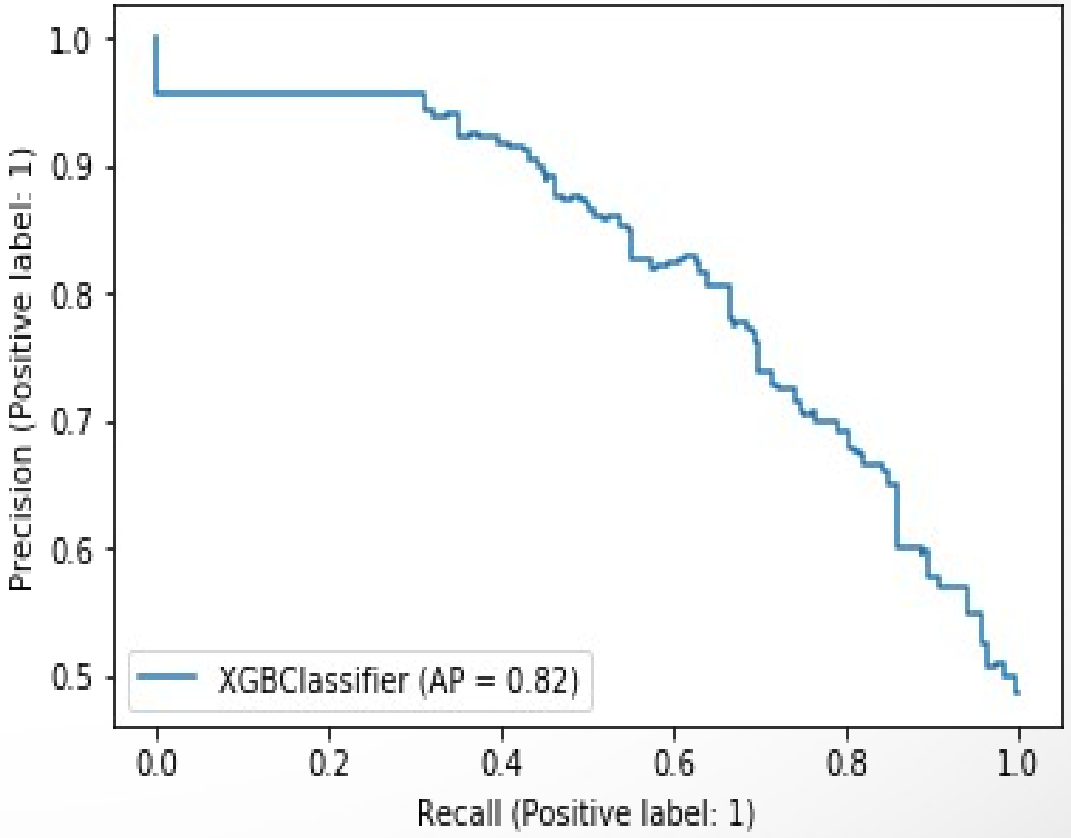
- The goal of this work is to propose a classification methodology for lung cancer that can be used automatically at the onset of the disease.
- In this project we are implementing lung Cancer Recognition Using CT-Scan with NCA-XG Boosting and KNN.
- The dataset that was used is Lung Image Database Consortium (LIDC). The Lung Image Database Consortium (LIDC) archive collects CT scan images taken from actual patient cases.
- A thoracic CT clinical scan image and an associated XML file are included in each object in the collection.
- https://drive.google.com/file/d/1AZ2UQJToA3J8k8h0XJWD12-v2w8p9H2Q/view?usp=share_link



Results

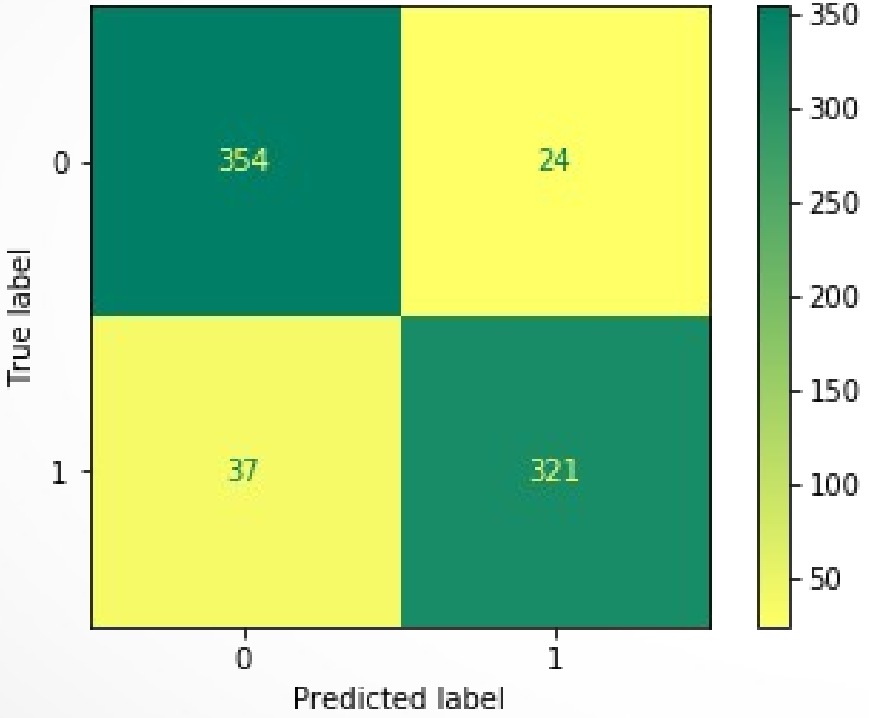


Confusion Matrix of NCA-XG Boosting

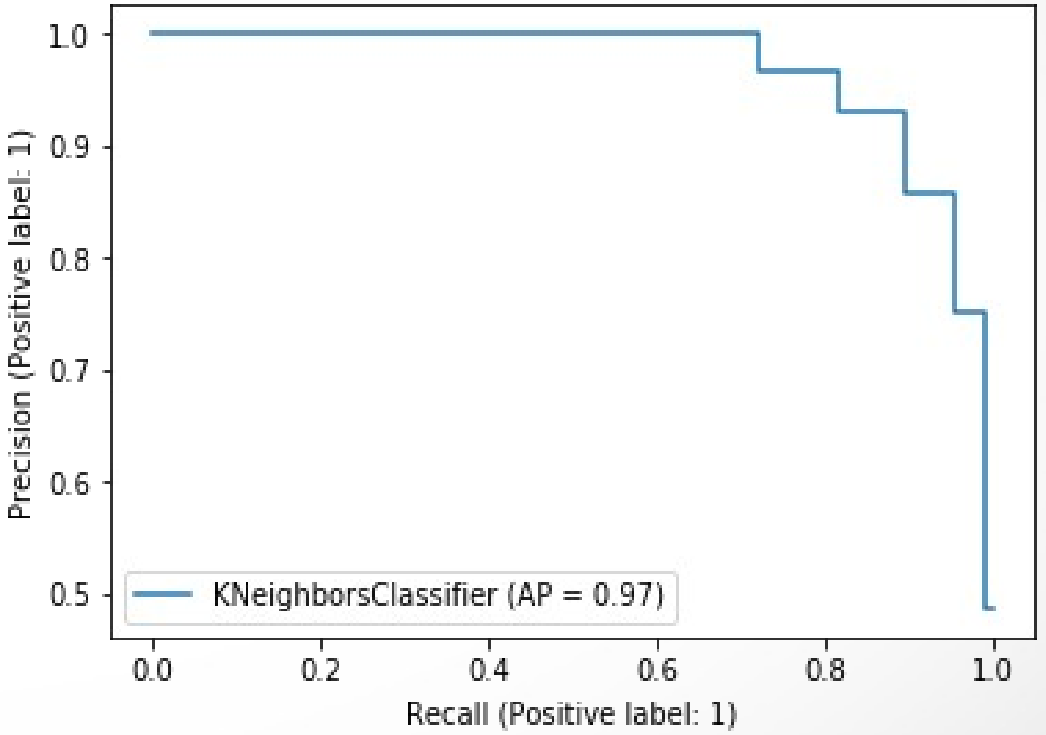


RoC Curve of NCA-XG Boosting

Classification Report of KNN

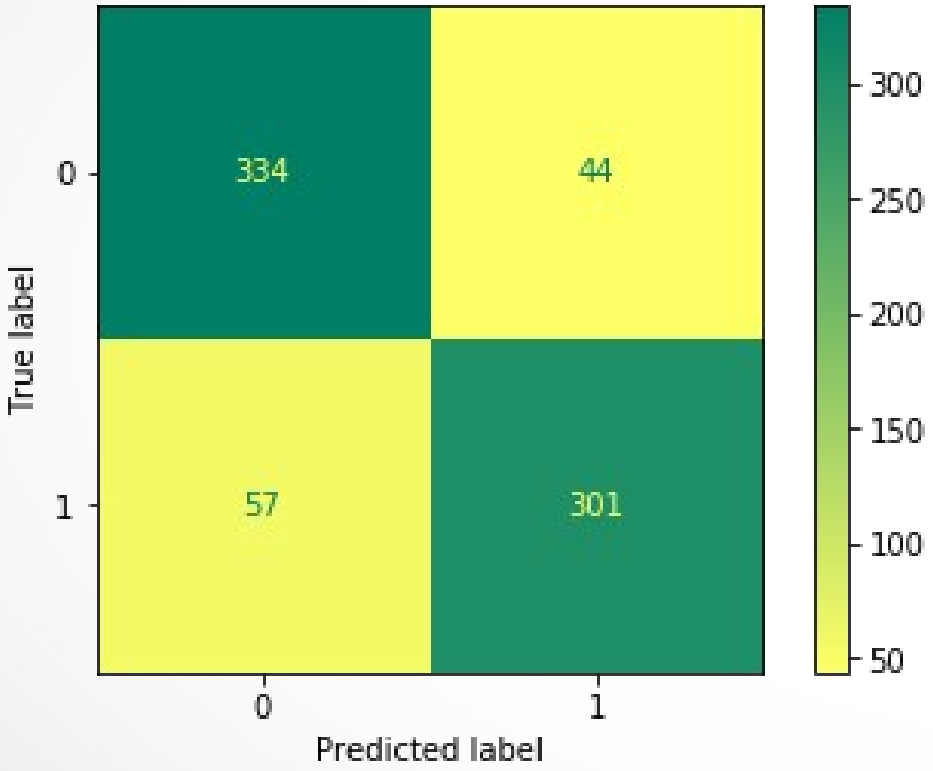


Confusion Matrix of KNN

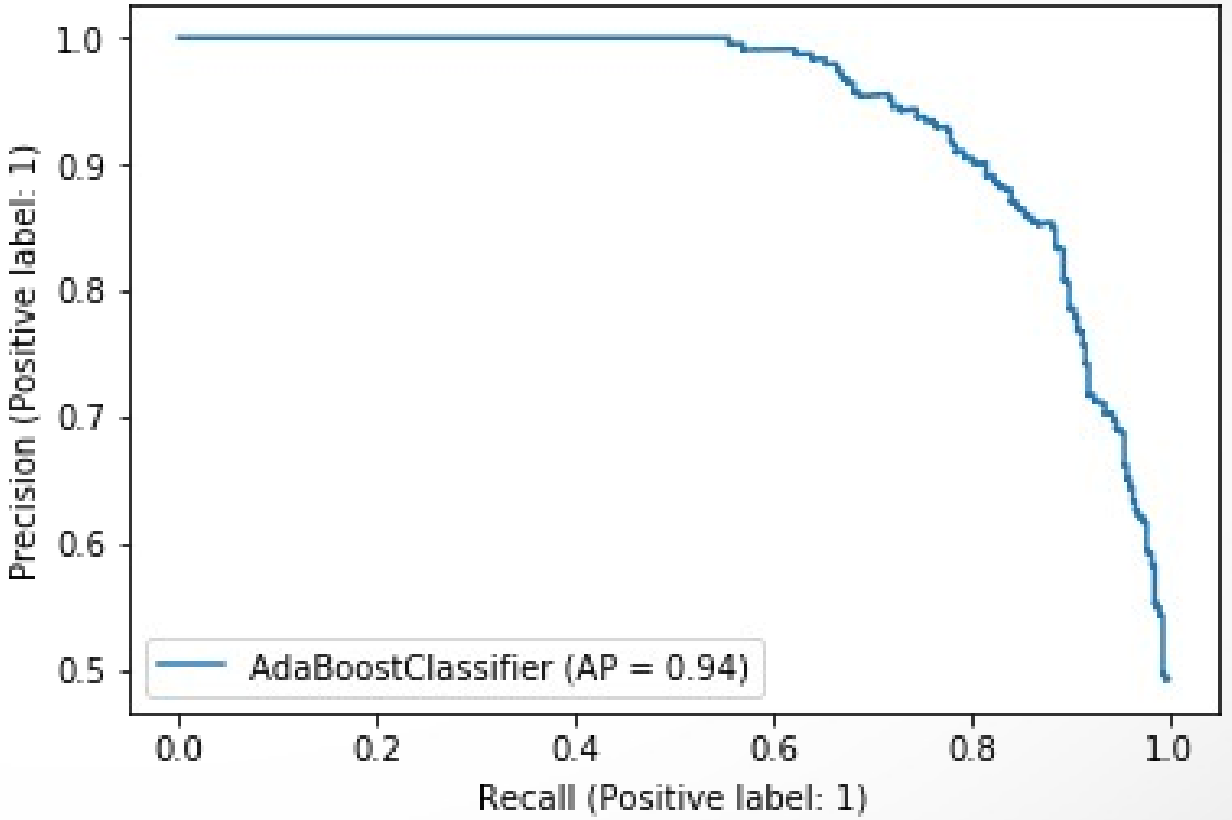


RoC Curve of KNN

Results of AdaBoost



Confusion Matrix of Ada Boosting



RoC Curve of Ada Boosting

References

- R. Navid, A. Mohsen, K. Maryam et al., “Computer-aided diagnosis of skin cancer: a review,” Current Medical Imaging, vol. 16, no. 7, pp. 781–793, 2020.
- [9]Cancer. Accessed: Apr. 30, 2021. [Online]. Available: https:// en.wikipedia.org/wiki/Cancer
- [11]S. F. Khorshid and A. M. Abdulazeez, “BREAST CANCER DIAGNOSIS BASED ON K-NEAREST NEIGHBORS: A REVIEW,” PalArch’s J. Archaeol. Egypt/Egyptology, vol. 18, no. 4, pp. 1927–1951, 2021.
- [33]N. O. M. Salim and A. M. Abdulazeez, “Human Diseases Detection Based On Machine Learning Algorithms: A Review,” Int. J. Sci. Bus., vol. 5, no. 2, pp. 102–113, 2021.
- S. Lakshmanaprabu, S. N. Mohanty, K. Shankar, N. Arunkumar, and G. Ramirez, “Optimaldeep learning model for classification of lung cancer on CT images,” Future Generation Computer Systems, vol. 92, pp. 374–382, 2019.