International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I12P206

Volume 70 Issue 12, 42-58, December 2022

Original Article

Prostate Cancer Detection using Radiomics-based Feature Analysis with ML Algorithms and MR Images

M. N. Rajesh[1], B. S. Chandrasekar[2], S. Shivakumar Swamy[3]

[1]*Department of ECE, Jain (Deemed to be University), Bengaluru, Karnataka, India.*
[2]*Faculty of Engineering, Jain (Deemed to be University), Bengaluru, Karnataka, India.*
[3]*HealthCare Global, Bengaluru, Karnataka, India.*

[1]*Corresponding Author : rajeshmn.mn@gmail.com*

Received: 25 October 2022 Revised: 12 December 2022 Accepted: 16 December 2022 Published: 24 December 2022

Abstract *- The study is aimed to understand and predict prostate lesions identification and classifications on magnetic resonance images (MRI) of the prostate using Radiomics based texture features analysis while applying Machine Learning (ML) algorithms. This study includes retrospective MR Images of patients with Prostate Cancer (PCa) from HealthCare Global (HCG), totalling 76 patients. From 207 prostate MRI images, 109 texture features were extracted using the python PyRadiomics library. Three sampling methods and various ML feature classification techniques are utilised to balance the dataset to develop the best diagnostic models for assessing these models' accuracy. The discriminative capability of all the models was evaluated by receiver operating characteristics (ROC) analysis. Eight different texture feature-based predictive models are developed by running the cross combination of all the dataset balancing and classification methods to identify and classify PCa malignancy from benign. Based on the test group's results, the majority of the models performed better with a larger area under the curve (AUC) (>0.80) and higher accuracy (>0.80). Support vector machine (SVM) classifier with AUC of 0.9744, 0.9759 accuracy and extreme gradient boosting (XGB) classifier with AUC of 0.963, accuracy with 0.9639 were the best model within the eight models, followed by decision trees (DT), light gradient boosting machine (LGBM) and random forest (RF) models. Extracting texture parameters from MRI images and combining the Radiomics approach with ML models can classify prostate lesions effectively*

Keywords *- Prostate cancer, ROI Delineation, Feature extraction, Feature analysis, Machine learning.*

1. Introduction

PCa is projected to be the second primary form of cancer among men and the world's fifth leading reason of cancer deaths in the year 2020, with an estimated roughly 1.4 million PCa new cases and 375,000 mortalities [1]. According to estimates provided by GLOBOCAN 2018, the number of newly diagnosed cases of prostate cancer around the globe is 1,276,106, with the incidence being highest in developed countries like India. The GLOBOCAN 2018 projections for the United States place the number of newly diagnosed cases of PCa at 191,930 and the number of deaths attributable to the disease at 33,330 [2]. According to India's national cancer registries, the number of cases is anticipated to be 41,532 and is estimated to be 47,068 by 2025. PCa is the leading cause of death among males, surpassing lung cancer. The very diverse nature of PCa makes clinical care of the disease incredibly difficult and complex. Suppose a prostate-specific antigen (PSA) or an abnormal digital rectal exam (DRE) is discovered. In that case, the conventional method for diagnosing PCa is based on the performance of regular random biopsies with the assistance of transrectal ultrasound (TRUS). These techniques do not provide clear information on PCa sensitivity and specificity, and they offer insufficient details regarding the aggressiveness of the disease and its stage. Efforts are being made to develop risk stratification tools that will assist in combining serum PSA levels, Gleason grade, and assessments of the anatomic extent of primary tumours. These tools will be used for clinical decision-making and to optimise patient management.

Imaging, a non-invasive method of treating PCa, plays a critical part in managing diseases and determining the severity of the condition. Imaging techniques of all kinds and combinations of imaging techniques are increasingly being used in various clinical settings worldwide. These choices are made based on an evaluation of the imaging modality's assessment, availability, affordability, and regulatory restrictions. This has helped improve clinical decision-making by allowing for earlier disease detection, resulting in a slower rate of growth than in the past [3].

However, determining the best imaging modality to use is not only impossible but also contentious due to the

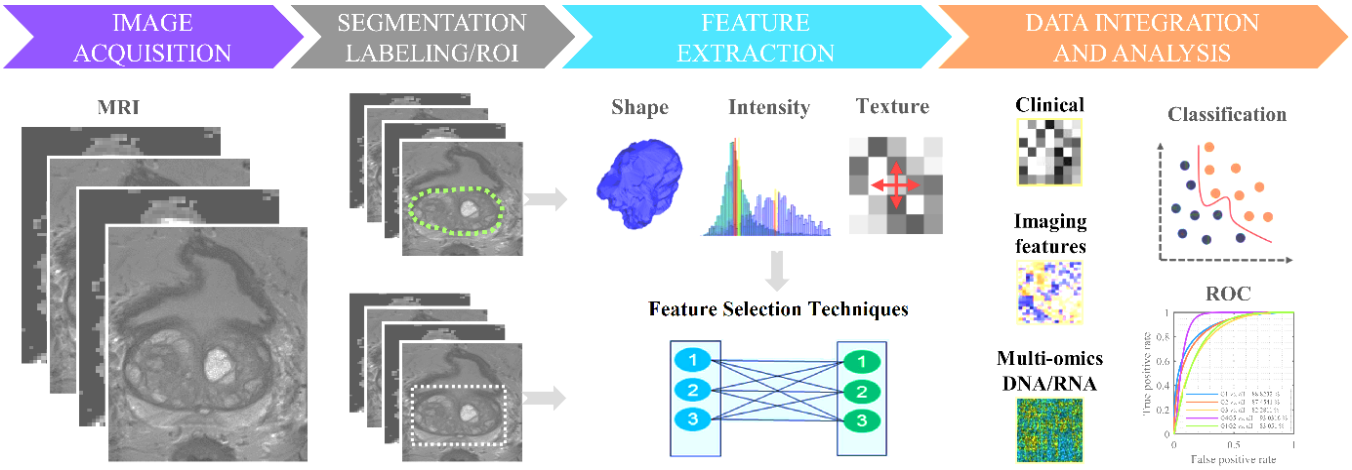


Fig. 1 Basic Pipeline of Radiomics in PCa [4]

numerous research conducted on modalities and their applications. Even though PCa imaging is important for managing patients, it has led to increased overdiagnosis and overtreatment.

1.1. Problem Statement & Challenge

The field of medicine known as "radiomics" is a new challenge that focuses on the quantitative aspects (radiomic features) extraction from radiology images that the radiologists and the application of this information for building a medical decision support system could not view. Radiomic features (like shape, intensity, wavelet or texture) are derived from radiology images (MRI, PET or CT scans), using sophisticated mathematical techniques, providing valuable information for tailored therapy. Radiomic features include:

The procedure of radiomics analysis related to detecting patterns to define the PCa grade group based on the Prostate Imaging Reporting and Data System (PI-RADS) is depicted in Figure 1. Firstly, a database of radiology images (MRI) is generated. Hence, standard images may be subjected to radiomics study with low bias [4]. This step is necessary to ensure accurate results. In most cases, the number of imaging characteristics should be either same as or lower than the sample volume. Secondly, segmenting images highlights portions of the image with PCa as a region of interest (ROIs). Segmentations can be achieved automatically (fully/semi) or manually. Then, extracting features stores the features of images (such as common features, histogram statistics, descriptors, in-depth features, textures and so on) in one or more distinct vectors for further analysis. After that, the radiomics feature has its predictive capability measured (i.e., the relative significance of various radiomics features). In conclusion, research models that utilise the earlier features of images for predicting the presence of PCa can be characterised by univariate analysis (such as the significance test, the Spearman correlation, and other similar tests) and multivariate analysis (such as regression and classification models). This final step needs to be carried out in a validation cohort of patients so that the newly developed radiomics model can be shown to have some generalizability [5].

In this work, the feature extraction and analysis of MRI PCa images for treatment prediction has been proposed based on the radiomics approach. The proposed radiomics model includes five stages: 1. MRI imaging (input), 2. ROI segmentation, 3. Extracting radiomics features from ROIs, 4. Data balancing, feature selection and reduction and 5. Analysis of the features using ML classification models (output). The remaining section of this article presents the following sections as a literature review of the related works on radiomics in the field of cancer diagnosis, the detailed representation of the proposed radiomics model, the experimental analysis, and the conclusions and future work.

2. Review of Related Works

2.1. Radiomics with ML for PCa Feature Extraction

The application of radiomics with ML algorithms presents numerous opportunities and benefits. Because the traditional interpretation of the image is based on radiologists' experience, the combination can reduce inter-individual variability while shortening the time needed for reporting, which could be an advantage for radiologists with less experience. A radiomics-based PCa feature extraction and analysis model was developed in [7] based on prostate MRI datasets. This model used various ML techniques, such as DT, RF, Gradient Boosted Tree, Ada Boost, k-Nearest Neighbor (kNN), and Naive Bayes (NB). PyRadiomics version 3.0 was used for image preprocessing and feature extraction in this scenario. The most accurate results were obtained using the Gradient Boosted Tree with the J48 DT model. The univariate statistical analysis was conducted to demonstrate the usefulness of the collected radiomics characteristics in differentiating prostate lesions.

A radiomics feature-based classification model was developed in [8] to distinguish PCa from other types of cancer using various ML methods. This study investigated the effectiveness of radiological assessment classes and the quantitative computer analyses of apparent diffusion coefficients (ADC) map. The highest Gleason Grade Group was used to annotate the index lesion in the MRI anatomical data collection (ADC) and the comparable histology slides. The volume of interest (VOI) and the normal-looking peripheral zone surrounding each lesion was assessed for each lesion. Radiomic analysis was utilised in the processing of VOIs. Principal component analysis, univariate analysis with sequential neural networks, SVM, and RF analysis were all utilised in classifying lesions according to their respective clinical relevance. An ML–based approach that used MR radiomics to detect clinically relevant PCa for improving the diagnostic performances of PI-RADS-v2 was proposed in [9]. This ML model used an SVM developed on the radial basis functions (RBF) kernels to distinguish between the present and absence of PCa optimally. It was suggested that the supervised ML known as RBF-SVM should be utilised for classification and regression analysis. Using SVM and recursive feature elimination (RFE), a preliminary diagnostic evaluation of feature selection was carried out.

In [10], an ML model based on radiomics features was constructed using PI-RADS-3's T2W lesions for identifying clinically relevant PCa. To consider the comparatively low frequency of PCa, the Synthetic Minority Oversampling Technique, often known as SMOTE, was utilised. On the T2W, multi-slice VOIs were delineated in the lesions of the PI-RADS-3 index. 107 radiomics characteristics, including histograms and textures, were recovered from the segmented lesions. To predict clinically relevant PCa, an RF classifier that takes the radiomics feature as the input has been trained and tested. Through the use of logistic regression (LR) with stepwise forward feature selection, the potential utility can be analysed that the integration of the radiomics RF classifier with prostate volume or PSA density could have for the prediction of clinically relevant PCa.

An ML model was developed in [11] to evaluate the effectiveness of the multiparametric MRI-based radiomics signatures to differentiate between medically significant PCa and insignificant PCa. For each patient, 819 radiomics attributes were collected from the mp-MRI scan. The SMOTE approach was used to ensure equitable representation of the small set in the training cohorts. A feature selection and radiomics signature-building approach using a minimum-redundancy maximum-relevance (mRMR) selection and the least absolute shrinkage and selection operator (LASSO) technique. The LR model was chosen for feature selection in the next step. Features from the candidate set with coefficients To create a radiomics signature that was not zero were chosen, and those features were then linearly merged.

A systematic and accurate ML-based model that included classification, statistical analysis and cross-validation was designed in [12] to identify the better-performing classification model for PCa risk stratifications using mpMRI-derived radiomics features acquired from a large set. A total of 110 radiomic features were extracted using the Gray-Level Difference, Co-occurrence Matrices (GLDM and GLCM), histogram analysis, and Fast Fourier transform (FFT)-based features of frequency. Using a Quadratic kernel-SVM (QSVM), 110 radiomic characteristics were analysed and interpreted. LR, linear, quadratic, cubic, and Gaussian kernel-based SVM, linear discriminant analysis (LDA), and RF models were used for classification. Radiomics analysis using ML models was performed on dynamic contrast-enhanced MRI images [13]. The variance threshold approach, the chosen k-best method, and the LASSO algorithm were utilised to lessen the number of feature dimensions. As the ML models were employed for evaluation, SVMs based on linear kernels and LR, RF, DT, and kNN models were used. This analysis concluded that the LR acquired better classification results when compared against various models. This model integrated the first enhanced phase of dynamic contrast-enhanced MRI with the most substantial enhanced phase and was found to be the most effective classification method.

2.2. Analysis of Research Gap

The analysis of radiomics features can provide information that can be used for diagnosis, clinical risk assessment, and treatment. Images produced from available imaging technologies do not require any further investigations, and with comparison to a biopsy, the entire tumour can be defined. The applications of AI, along with ongoing research into clinical interpretation of various computational models in medicine and PCa, will improve the detection of PCa tumours and the possible grades of those cancers and their classification. Radiomics, in conjunction with ML techniques, is being used to investigate the prospect of distinguishing between low and high-grade PCas, as well as to characterize tumours, evaluate risks, and plan treatments. This research was proposed to analyze and predict MRI prostate lesions identification and classifications using Radiomics-based texture features analysis while applying ML algorithms.

3. Proposed Model

This research proposes a feature analysis model for classifying Radiomics features based on PCa MR Images. These MRI images comprise 76 patients of about 207 from the HCG, Bangalore. Onco-Radiologists from HCG Bangalore were involved in identifying the ROI that is by marking prostate lesions on the Original MR Image. The image masks were created based on the marking provided on the Original Images. These image masks and the original images were used to extract features. Figure 2 represents the proposed model's workflow based on the Radiomics

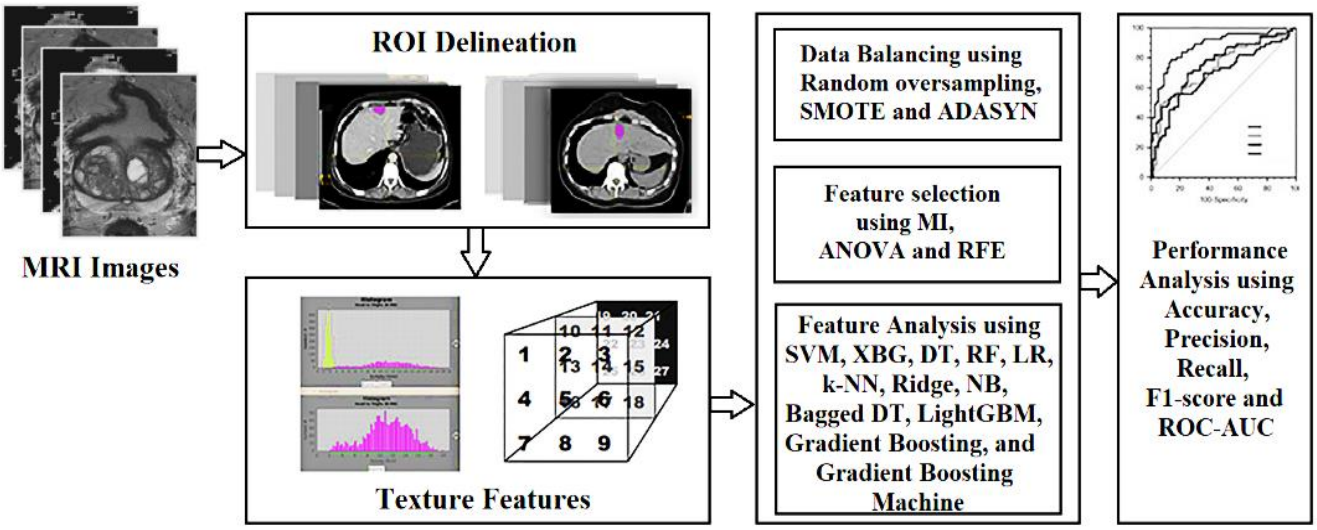


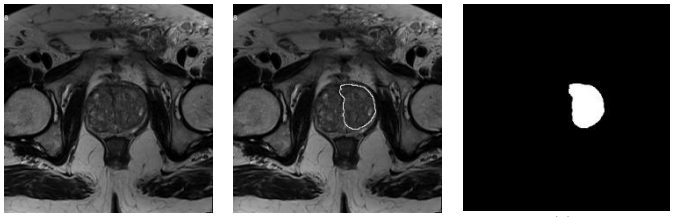
Fig. 2 Architecture of Research Model

technique. The proposed model's workflow includes the following stages: image acquisition, ROI delineation, feature extraction, data balancing, selection of features, and analysis of features. The research model works on the following steps.

MR Image Acquisition and Image Dataset Creation: A retrospective study of 76 patients with PCa summing to 207 prostate MRI images from HCG were included.

3.1. ROI Delineation

HCG Radiologists marked the Lesions' contour and provided if the lesion contoured was Benign or Malignant. Figure 3 shows the Original MR Image having lesions along with the image having lesions and the image mask.



(a) (b) (c)

Fig. 3 a) Original MR Image having lesions, b) Lesions in MR Image being marked for study, c) Image Mask

3.2. Feature Extraction and Feature Dataset Creation and Feature Normalization

Totally, 109 texture features are extracted using the python PyRadiomics Library which provides below 8 classes.

1. First-order statistics- The primary focuses are the pixel intensities distribution and the utilisation of elementary metrics for geometrical computing features.
2. Shape-based (2D) quantifies the region's or volume's overall geometric shape using a two-dimensional data representation. The subcategories are ROI size, compactness, sphericity, total volume, diameter, surface area, surface-to-volume ratio and flatness.
3. Gray Level Run Length Matrix- It is a matrix that displays the size of the sequential pixel with similar intensities. The subcategories under the GLRLMs were long and short run emphasis; grey level non-uniformity; run length non-uniformity; run percentage; high and less grey level run emphasis.
4. Shape-based (3D)- The shape and size descriptors of ROI in three dimensions are included in this category of characteristics, referred to as the "3D Shape Features." Consequently, these features were unaffected by the gray-level intensities distributions in ROI, the only way to compute them is by using the image that has not been produced from it and the mask.
5. Gray Level Cooccurrence Matrix- GLCM is a matrix that displays how often two intensity levels have happened in the two pixels separated by a particular distance. Autocorrelation; Correlation; Contrast; Homogeneity; Energy; Variance; Sum of Average; Inverse Difference Moment; Sum of Variance; Information Measures of Correlations; Difference of Variances; Dissimilarity; Cluster Prominences, Cluster Shade, Maximum Probability and Cluster Tendency are the subcategories of GLCM.
6. Neighboring Gray Tone Differences Matrix- NGTDM is implicated not with the pixel itself but rather with the intensities of the adjacent pixels. The NGTDM can break down into the subcategories of Contrast, Coarseness, Complexity, Busyness, and Strength.
7. Gray Level Size Zones Matrix- GLSZM is a tool for measuring areas of grey level in images. This type of configuration is meant to as a gray-level zone whenever there is a predetermined number of linked voxels that all

have the same intensity levels for the grey level. In contrast to GLCM and GLRLM, the GLSZM is not dependent on the rotation of the data and requires the construction of a single matrix to consider all the directions present in the ROI.

8. Gray Level Dependence Matrix- GLDM analyses the image's gray-level dependencies. This dependency was the total linked voxels within a certain distance of the central voxel and is dependent.

3.3. Feature Normalisation

The normalisation of features was done before selecting the features. Feature selection was made to choose the features most closely associated with the biopsy outcome to develop prediction models despite the related limited sample size of the patient and the higher-dimensional nature of the features. A feature normalising procedure was carried out to ensure that the features of radiomic collected from various locations fall within a range of values compared to one another [14]. Each feature vector $\bar{F}_i(k)$ is given a normalised value based on each site's S_k equation (1).

$$F_i(k) = \frac{\bar{F}_i(k) - \mu_i(k)}{\sigma_i(k)} \quad (1)$$

In this case, the mean denotes the symbol $\mu_i(k)$, whereas the mean absolute deviation $\bar{F}_i(k)$ is denoted by the symbol $\sigma_i(k)$ (considering all the samples c in site S_k). This procedure was carried out repeatedly for each unique site $S_{1,\ldots,4}$ so that all of the features contained within site would have a mean value of 0 and a mean absolute deviation value of 1.

3.4. Dataset Balancing

A procedure utilised to deal with uneven data is known as data balancing. Before applying a classifier to the dataset, the data must go through a process that balances the data. In brief, the following are the actions that this method takes on the data set:

- Under-sampling comprised sampling from a majority of classes to hold only a part of these points.
- The practice of duplicating some points from minority classes to extend the cardinality of the sample is known as oversampling.
- Developing synthetic points from the minority classes to improve their cardinality is known as the generation of synthetic data.

3.4.1. Random Oversampling

It consists of choosing samples randomly from a minority class with replacements and providing the training data samples with several copies of this sample. Because of this, it is feasible that a single sample could be chosen many times.

3.4.2. SMOTE

It is an over-sampling approach developed in 2002. Instead of replacing the already present samples, samples for the minority class are generated synthetically. It contributes to the problem of over-fitting. The SMOTE algorithm is utilised so that the overfitting problem can be resolved and the accuracy can be improved. This methodology generates manual minority samples with the line segments that unite the minority sample and its 'k' minority class nearest neighbour. These line segments are called line segments combining the minority samples. The 'k' nearest neighbours have their neighbours chosen randomly, and the selection criteria are based on the required oversampling rate. One of the problems with the SMOTE algorithm is that it overgeneralizes the space occupied by the minority classes without taking into account the majority classes, which can lead to an increase in the amount of overlap between the classes [15].

3.4.3. Adaptive Synthetic Sampling (ADASYN) Oversampling

The methodology known as ADASYN is developed to produce data samples from minority groups following their distributions adaptively. Samples from minority classes that are more difficult to learn are utilised to generate additional synthetic data than samples from classes that are simple to understand. This helps to lessen the learning bias initially established due to an imbalance in the data distribution. The SMOTE approach produced the same count of synthetic samples for the entire minority classes. In contrast, in the ADASYN approach, density distribution was utilized to automatically decide the count of synthetic samples that need to be produced for all the minority classes samples. It contrasts with the SMOTE approach, which produces the exact count of synthetic samples for all the minority classes. The algorithm performs its primary objective, assigning weights for various minority classes of samples to produce varying quantities of synthetic data for all the samples [16].

3.5. Feature Selection

The feature selection method is a preprocessing type that identifies a particular issue's essential characteristics. In medical applications, feature selection has proven to be effective. It not only helps to minimise the number of dimensions but also provides information on the factors contributing to developing a disease. One of the methods that can be utilised in the process of dimensionality reduction is known as feature selection. In this method, only the pertinent characteristics are chosen, while the unnecessary and redundant ones are rejected. A reduction in the dimensionality of the inputs can increase performance in one of two ways: either by reducing the learning speed and the complexity of the model or by boosting the capacity for generalisation and classification accuracy. In addition to lowering the total cost of the measurement, selecting appropriate characteristics might increase comprehension of the issue [26].

3.5.1. Mutual Information

MI was successfully endorsed in filter features-selection approaches for assessing the relevance of the feature subsets in the target variables prediction and the redundancy related to other variables. This is accomplished by comparing the predicted values of the target variable to the predicted values of the other variables. The MI is an index that measures the degree to which random variables statistically depend on one another. Intuitively, the MI quantifies the degree to which knowing the values of one variable decreases the amount of uncertainty associated with the other variable. In contrast to other indices, such as the Pearson correlation coefficients, the MI could consider both linear and non-linear dependencies unaffected by differentiable and invertible transformations of random variables. In addition, the MI is an invariant index.

3.5.2. Analysis of Variance

It is a method that is both very easy to use and very effective for determining whether or not there is a difference in means among groups. It has been utilized for feature selection since it possesses the advantages listed below. Initially, it can withstand significant deviations from the majority of its assumptions. Second, it is effective while performing the analysis of the relationship that exists between the two variables. Thirdly, it is possible to apply it successfully even when the number of observations in each group differs. Last, it is simple to generalize it to more than two groups while maintaining the same level of Type 1 error [18].

3.5.3. Recursive Features Elimination

RFE is an example of the wrapper-type algorithm for selecting features. It represents the various ML algorithms presented and utilized in the model's core, that RFE wrapped it, and that it was used to determine characteristics. Contrarily to this, filter-based features selection assigns a score to all the features before choosing those features that have the higher (or lower) scores. RFE further uses filters-based feature selection internally. It finds the features subset by initiating with every feature in the data set of training and then deleting the features until the appropriate number of features is left successfully [19].

3.6. Feature Analysis and Cancer Identification using Classification Algorithms

3.6.1. SVM

The goal of this classifier was to create a model that could precisely predict the classes of validation data or unknown data comprising only characteristics, as shown in equation (2). The SVM kernels are often utilised in mapping non-linearly individual input into feature space instances of higher dimensional. These feature space instances are made up solely of features [25].

$$f(x) = sign(\sum_{i \in SV} \alpha_i^* y_i . k(x_i, x_{SV}) + b^*) \quad (2)$$

In this context, "kernel function" refers to $k(x_i, x_{SV})$ the equation for the kernel k of the RBF described in Equation 3.

$$k(x, y) = exp\left(\frac{\|x-y\|^2}{2\sigma^2}\right) \quad (3)$$

3.6.2. DT

Classifiers based on DTs are powerful, efficient, and widely used data mining and knowledge discovery methods. They are used to examine huge and complicated datasets to locate helpful patterns. This field is very important since it enables the extraction of knowledge and modelling from massive volumes of data. Theorists and practitioners are constantly searching for new ways to make the process simpler, less expensive, more accurate, and more efficient. DTs are extremely useful tools in various fields, such as data mining, machine learning, information extraction, text mining, and pattern identification.

3.6.3. RF

RFs are essentially an ensemble of many different DTs. RF is a classifier that considers the average count of DTs implemented for different dataset subsets provided. It supports increasing the dataset's predicted accuracy. The RFs do not rely on one DT but instead consider the prediction from all the trees and determine the last outcome based on the trees with the largest count of votes for its prediction. The greater the count of trees, the higher the accuracy obtained and the prevention of overfitting issues [22].

3.6.4. kNN

kNN is an approach for ML that determines the proximity of two instances based on the Euclidean distance between them. K-NN makes predictions about the class labels assigned to various instances by determining their lowest Euclidean distances from other samples [24]. The Euclidean distance was determined by considering all of the features as dimensions and is provided in equation (4).

$$d(x_i, x_j') = \|x_i - x_j\|^2 = \sum_{k=1}^{d} (x_{ik} - x_{jk})^2 \quad (4)$$

3.6.5. LR

LR is a supervised classification technique in ML. A logistic function is applied here to model the dependent variable. It is employed to forecast the likelihood of the target variables. The dependent or target variable is dichotomous, indicating that there are just two distinct categories of viable options. It is one of the simplest ML methods, and it could be applied to a wide variety of classification issues, including the detection of spam, the prediction of diabetes, the detection of cancer, and so on [20].

3.6.6. LGBM

LGBM is a GB model developed on DTs. It aims to enhance the model's effectiveness while minimising the

memory volume it needs. It uses two approaches: exclusive feature bundling (EFB) and gradient-based one-side sampling, both of which fulfil the histograms-based algorithm's limitations, which were mainly utilised in every GBDT model [16].

3.6.7. Ridge Classifier

This classifier approaches the problem as if it were a regression problem and trains itself accordingly using {-1,1} labels for binary data and multiple numbers for non-binary data. This classification algorithm depends on the subspace assumption, which asserts that training samples of a particular class fall on a linear subspace. A new test sample to a category will also be characterised as a linear combination of the relevant class training samples.

3.6.8. NB

This classifier is a probabilistic based on the Bayesian posterior probability distribution. It ensures that there is an independent relationship between each of the attributes by using conditional probability. The NB models are a set of classification algorithms derived from Bayes' theorem. It is not an algorithm but a group of algorithms where each shares a general premise: all pairs of classified features are independent of one another. Only binary data can be modelled successfully using the multivariate Bernoulli NB model. The Bernoulli model performs significantly better than the multinomial model in situations where the dataset size is limited [23].

3.6.9. Bagged DT

DT is a method that can be used to organise difficult challenges into a hierarchy of more manageable challenges. The problem of individual DTs having the propensity to overfit the data is solved using bootstrap-aggregated (bagged) DTs, which incorporate the findings of several different DTs. This method was chosen because it enhances generalisation, lessens the impact of overfitting, and is the methodology utilised in this research. A bootstrap sample of the data is used in conjunction with an ensemble of DTs to generate a bagged DT. This method is similar to an RF algorithm since it chooses a random subset of predictors to apply at each decision split.

3.6.10. Gradient Boosting Classifier (GBM)

This approach is utilised to develop regression and classification models to optimise the model's learning process. These models were generally non-linear and more commonly referred to as regression or decision trees. GB is a technique that is used to develop these models. The process of modelling the group of unreliable prediction models, i.e., regression DTs involves gradually and sequentially incorporating new learners into the mix. It is composed of nodes and leaves that, when combined, produce results that are predictive according to the decision nodes. Individual regression trees were not very good models; however, when considered as the group, the accuracy of the regression trees significantly increased.

3.6.11. XGB

This model is a well-known GB technique (ensemble) that has boosted performance and speed in tree-based ML algorithms (sequential DTs). Boosting techniques are a subcategory of Ensemble Learning, and XGB is one of those techniques. Learning through ensemble includes a count of compiling predictors, a multiple model framework that aimed to increase prediction accuracy. Through the utilisation of boosting technique, issues caused by earlier models were aimed to be rectified through succeeding models using the integration of some weights to the frameworks.

3.6.12. GBM

The learning functionality in GBMs successively adapts newer models to provide a more accurate estimate of the response variable. The most important aim of this approach was to create the newer base learners to have the highest possible correlation with the negative gradients of the loss functions linked with the full ensemble. The loss functions that are utilized could be chosen randomly. Hence, to provide a better understanding, consider that if the error functionality was the common squared-error loss, then the learning technique will result in sequential error-fitting [21].

4. Experimental Results and Discussion

In this section, the experimentation of the proposed models performs, and the results are carried out. 120 features were available from the eight classes of radiomics features discussed above. However, in this research, only 109 features are extracted using the proposed model and the remaining features are avoided, related to Shape-2D, 3D and boundary classes. These features were extracted by using the python PyRadiomics Library. Results were obtained in three stages to understand the importance and improvements in accuracy and ROC-AUC values as represented following.

4.1. Stage 1: Classification on Unbalanced Feature Dataset and without Feature Selection

This research model's experimentation is carried out in three stages. The first stage includes the performance analysis of the Unbalanced Feature Dataset without Feature Selection. The second stage comprises the performance analysis of the Balanced Feature Dataset without Feature Selection. Finally, the third stage consists of the performance analysis of the Balanced Feature Dataset and Feature Selection.

Table 1. Performance Analysis of Feature Classification ML Models

Feature Classification Model	Accuracy	Recall	Precision	F1-Score	ROC-AUC
SVM	0.8077	0.81	0.65	0.72	0.5
XGB	0.7692	0.77	0.7	0.73	0.5143
DT	0.7692	0.77	0.77	0.77	0.6286
RF	0.8077	0.81	0.65	0.72	0.5
kNN	0.8077	0.81	0.65	0.72	0.5
Ridge	0.8077	0.81	0.65	0.72	0.5
LR	0.7885	0.79	0.65	0.71	0.4881
LGBM	0.8824	0.88	0.88	0.88	0.8823
NB (Bernoulli)	0.75	0.75	0.72	0.73	0.5405
Bagged DT	0.7885	0.79	0.72	0.74	0.5
GBC	0.8077	0.81	0.65	0.72	0.5
GBM	0.8077	0.81	0.65	0.72	0.5

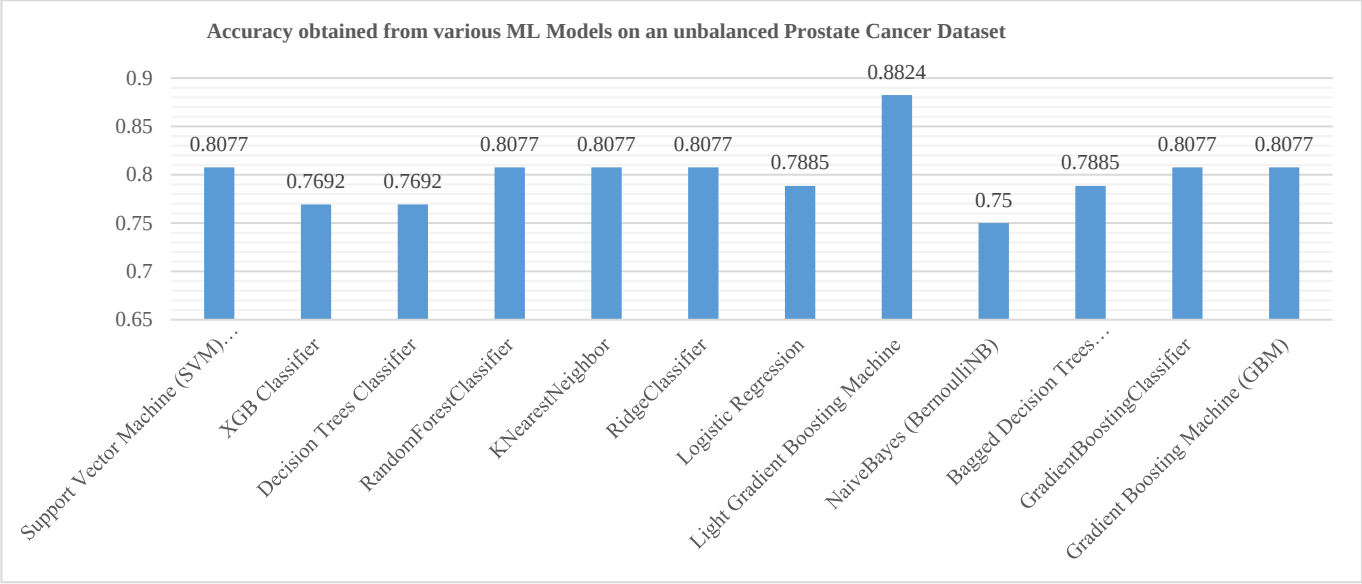


Fig. 4 Accuracy Comparison of ML Models on Unbalanced Dataset without Feature Selection

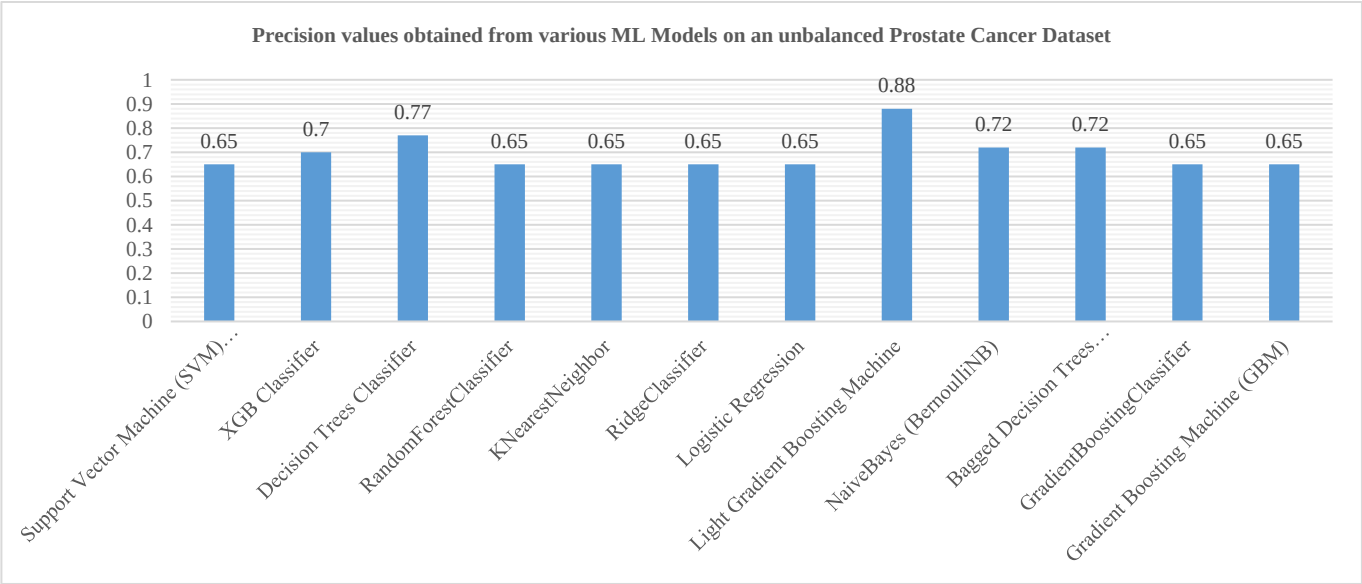


Fig. 5 Precision Comparison of ML Models on Unbalanced Dataset without Feature Selection

Table 1 represents the proposed classification model's performance analysis based on the unbalanced feature dataset without feature selection. For this analysis, twelve different ML models were used. Each model's performance was evaluated in terms of accuracy, precision, recall, f1-score and ROC-AUC. Figure 4 represents the accuracy performance of the ML models based on an unbalanced dataset without feature selection techniques. In this comparison, the LGBM model achieved the highest accuracy, with 88.24% accuracy than other ML models. Most models, like SVM, RF, kNN, Ridge, GBC and GBM, obtained similar accuracy with 80.77% accuracy. NB achieved the least performance with 75% accuracy.

Figure 5 represents the precision performance of the ML models based on an unbalanced dataset without feature selection techniques. In this comparison, the LGBM model achieved the highest precision value than other ML models, with 88%. Most models like SVM, RF, kNN, Ridge, LR, GBC and GBM obtained similar precision values of 65%. The DT classifier obtained the next better performance than LGBM with 72% precision.

The recall performance of the ML models based on an unbalanced dataset without feature selection techniques is displayed in figure 6. In this comparison, the LGBM model achieved the highest recall value than other ML models, with 88%. Most models like SVM, RF, kNN, Ridge, GBC and GBM obtained similar precision values of 81%. The Bagged DT and LR classifiers obtained the next better performance than LGBM with 79% recall.

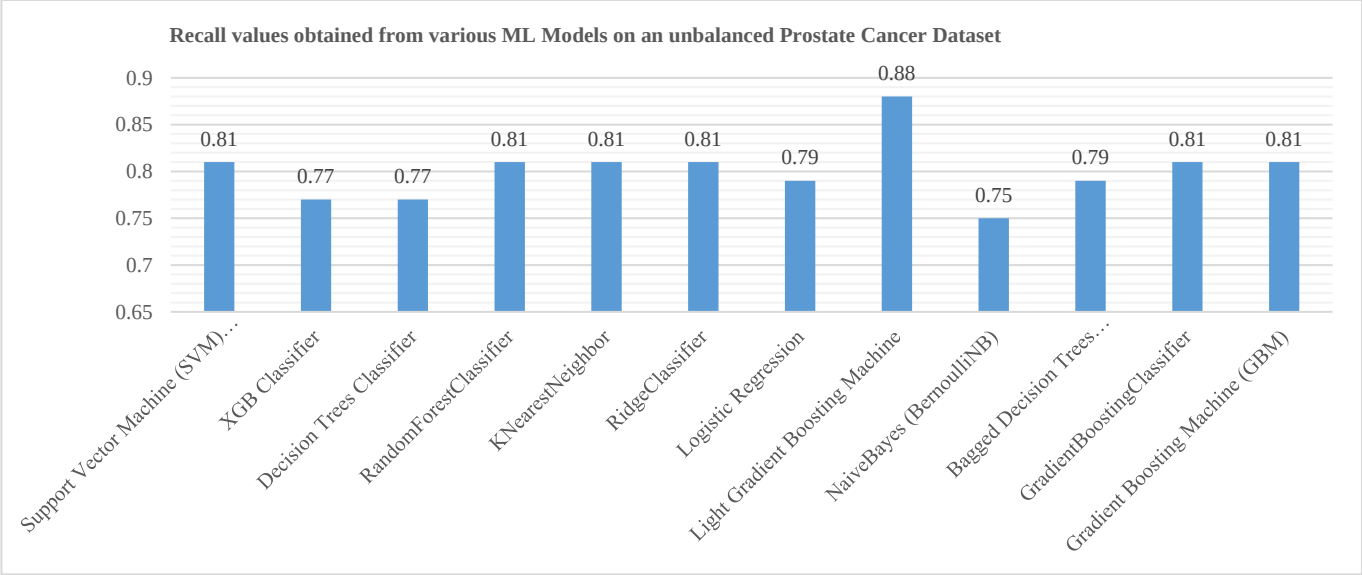


Fig. 6 Recall Comparison of ML Models on Unbalanced Dataset without Feature Selection

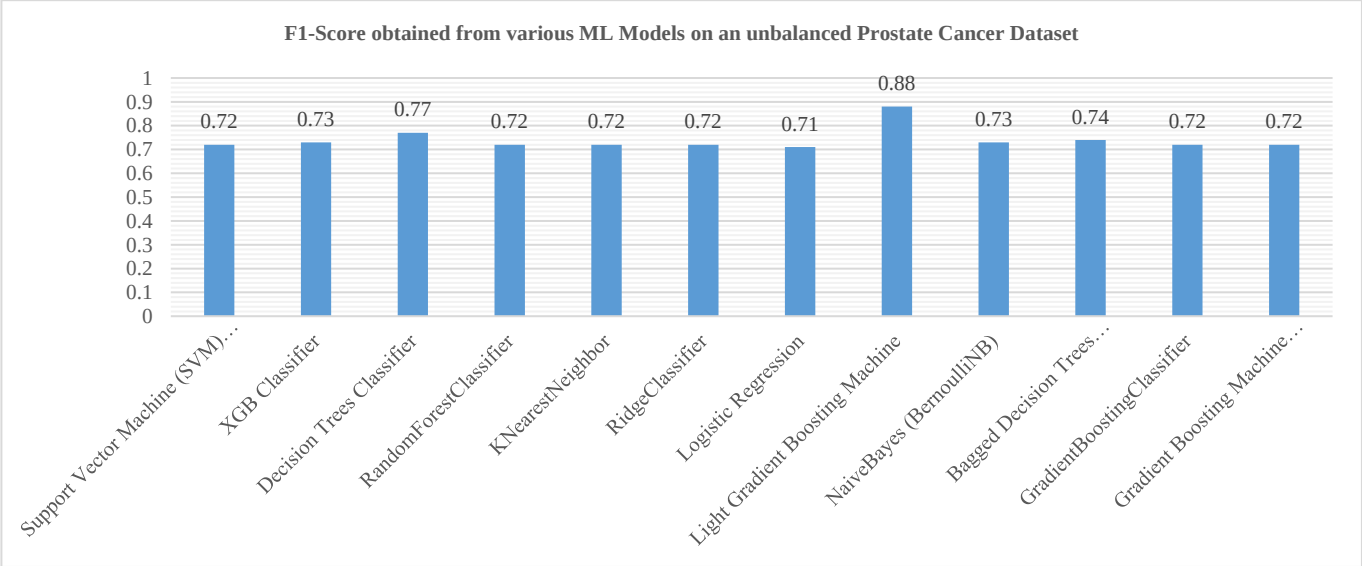


Fig. 7 F1-score Comparison of ML Models on Unbalanced Dataset without Feature Selection

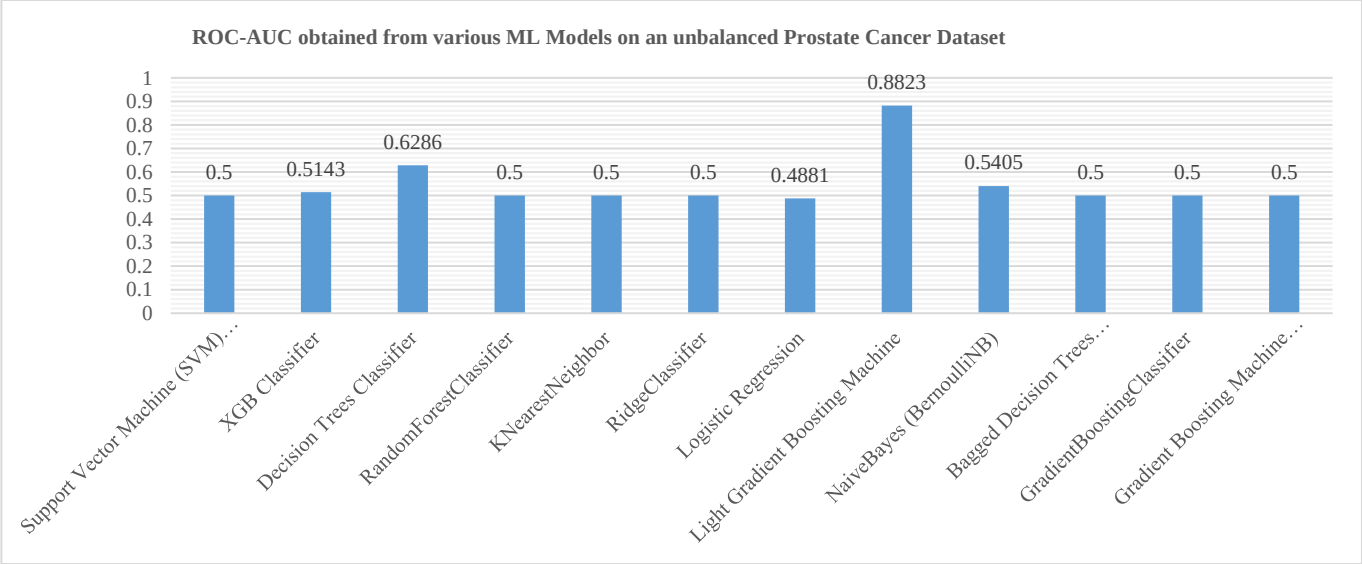


Fig. 8 ROC-AUC Comparison of ML Models on Unbalanced Dataset without Feature Selection

Figure 7 displays the f1-score performance of the ML models based on an unbalanced dataset without feature selection techniques. In this comparison, the LGBM model achieved the highest f1-score value than other ML models, with 88%. Most models, like SVM, RF, kNN, Ridge, GBC and GBM, obtained similar precision values of 72%. The Bagged DT classifier obtained the next better performance than LGBM with 74% precision, and the LR model achieved the least performance with 71%.

The ROC-AUC performance of the ML models based on an unbalanced dataset without feature selection techniques is displayed in figure 8. In this comparison, the LGBM model achieved the highest ROC-AUC value than other ML models with 0.8823. Most models like SVM, RF, kNN, Ridge, Bagged DT, GBC and GBM obtained similar ROC-AUC values with 0.5. The DT classifier obtained the next better performance than LGBM with a ROC-AUC value of 0.6286. LR model achieved the least performance with a ROC-AUC value of 0.4881.

4.2. Stage 2: Classification Done on Balanced Dataset but without Feature Selection

In this stage 2 classification, the dataset was balanced using data balancing techniques like ADASYN oversampling, random oversampling and SMOTE. The performance of the ML models was evaluated based on these data balancing techniques applied individually to all the ML models. In this stage 2, only eight feature classification models were used, as represented in the following sections.

4.2.1. ADASYN Over-Sampling Technique for Dataset Balancing

The below steps were followed while adopting ADASYN over-sampling technique:

- 167 Malignant Image Cases and 42 Benign Images Cases were upsampled to 173 Malignant Image Cases and 165 Benign Images Cases there by Balancing the Classification use case.
- 75% of the cases were used for training, and 25% were used for validation.

Table 2 represents the performance analysis comparison of ML models evaluated based on the balanced dataset using the ADASYN method without feature selection. Eight ML classifiers, such as SVM, XGB, DT, RF, kNN, Ridge, LR and LGBM, were used in this stage 2 evaluation. In this comparison, the RF and XGB models have performed better than the other models.

Table 2. Performance Analysis of ML Models based on Balanced Dataset using ADASYN Over Sampling without Feature Selection

Feature Classification Model	Accuracy	Recall	Precision	F1-Score	ROC-AUC
SVM	0.5412	0.54	0.54	0.54	0.5418
XGB	0.8824	0.88	0.88	0.88	0.8823
DT	0.7294	0.73	0.73	0.73	0.7287
RF	0.8824	0.88	0.89	0.88	0.8832
k-NN	0.6353	0.64	0.67	0.63	0.6379
Ridge	0.8235	0.82	0.83	0.82	0.8239
LR	0.8118	0.81	0.82	0.81	0.8126
LGBM	0.7647	0.76	0.81	0.76	0.7625

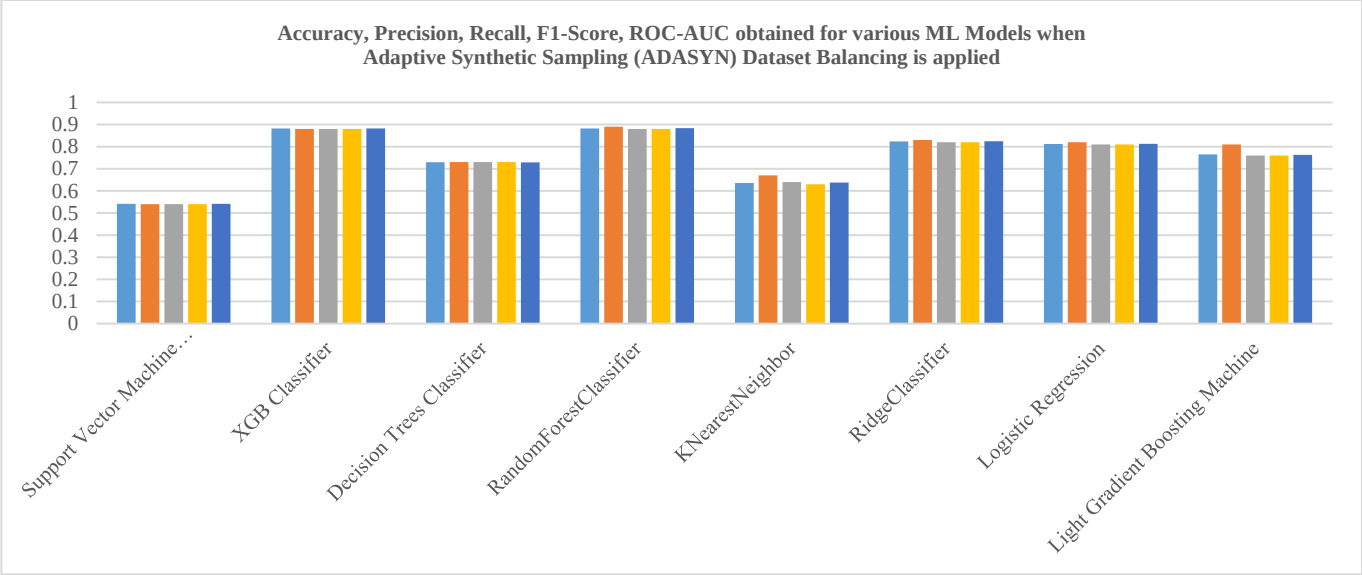


Fig. 9 Performance Comparison of ML Models on Balanced Data set using ADASYN without Feature Selection

Table 3. Performance Analysis of ML Models based on Balanced Dataset using Random Over Sampling without Feature Selection

Feature Classification Model	Accuracy	Recall	Precision	F1-Score	ROC-AUC
SVM	0.9759	0.98	0.98	0.98	0.9744
XGB	0.9398	0.94	0.94	0.94	0.9417
DT	0.9518	0.96	0.95	0.95	0.9545
RF	0.9398	0.95	0.94	0.94	0.9432
kNN	0.8298	0.83	0.81	0.81	0.8191
Ridge	0.759	0.77	0.76	0.76	0.764
LR	0.8193	0.86	0.82	0.82	0.8281
LGBM	0.9398	0.95	0.94	0.94	0.9432

Both these RF and XGB models have similar accuracy values of 88.24%, recall and f1-score of 88%. Besides that, the RF model scores higher in precision value with 89%, whereas the XGB achieved 88% precision. The ROC-AUC value of the RF was 0.8832 and 0.8823 for XGB. The best-performed model was RF and the XGB, and the least-performed model was SVM, with an overall 54% score in all the parameters. Figure 9 indicates the graphical chart comparing performance analysis based on the balanced dataset using ADASYN without feature selection.

4.2.2. Random Over-Sampling Technique for Dataset Balancing

Below steps were followed while adopting the Random over-sampling technique:

- 167 Malignant Image Cases and 42 Benign Images Cases were upsampled to 165 Malignant and 165 Benign Images Cases there by Balancing the Classification use case.
- 75% of the cases were used for training, and 25% were used for validation.

The performance analysis comparison of ML models evaluated based on the balanced dataset using a random over-sampling method without feature selection was presented in table 3. In this comparison, the SVM model has performed better in all parameters than the other models. The SVM model obtained 97.59% accuracy, precision, recall and f1-score with 98%. The ROC-AUC score of the SVM model was 0.9744. Compared to the SVM's performance, the following best performance was obtained by DT with 95.18% accuracy and overall 95% performance in all the parameters. Models like XGB, RF and LGBM have obtained similar and close performances.

The best-performed model was SVM, and the least-performed model was Ridge, with an overall 76% score in all the parameters. Figure 10 represents the graphical plot of the performance analysis comparison based on the balanced dataset using random over-sampling without feature selection.

4.2.3. Synthetic Minority Over-sampling Technique (SMOTE) for Dataset Balancing

The below steps were followed while adopting SMOTE over-sampling technique:

- 167 Malignant Image Cases and 42 Benign Images Cases were upsampled to 165 Malignant and 165 Benign

Images Cases there by Balancing the Classification use case.

- 75% of the cases were used to train, and 25% were used to Validate.

Table 4 represents the performance analysis comparison of ML models evaluated based on the balanced data set using SMOTE method without feature selection. In this comparison, the XGB model has performed better in all parameters than the other models. The XGB model has obtained 89.16% accuracy, 90% precision, recall and f1-score with 89%. The ROC-AUC score of the XGB model was 0.8963. Compared to the XGB's performance, the following best performance was obtained by LGBM with 87.95% accuracy and, overall, 88% performance in all the parameters. The RF model has obtained close performance correlated to LGBM. The best-performed model was XGB, and the least-performed model was SVM, with 61% scores in all the parameters. Figure 11 indicates the graphical chart comparing performance analysis based on a balanced dataset using SMOTE without feature selection.

4.3. Stage 3: Classification of Balanced Feature Dataset and with Feature Selection

In this final stage of classification, the feature analysis and cancer identification were performed using the feature classification ML models like SVM, XGB, DT, RF, kNN, Ridge, LR and LGBM. The following table represents the performance analysis of these eight ML models after balancing the dataset and feature selection.

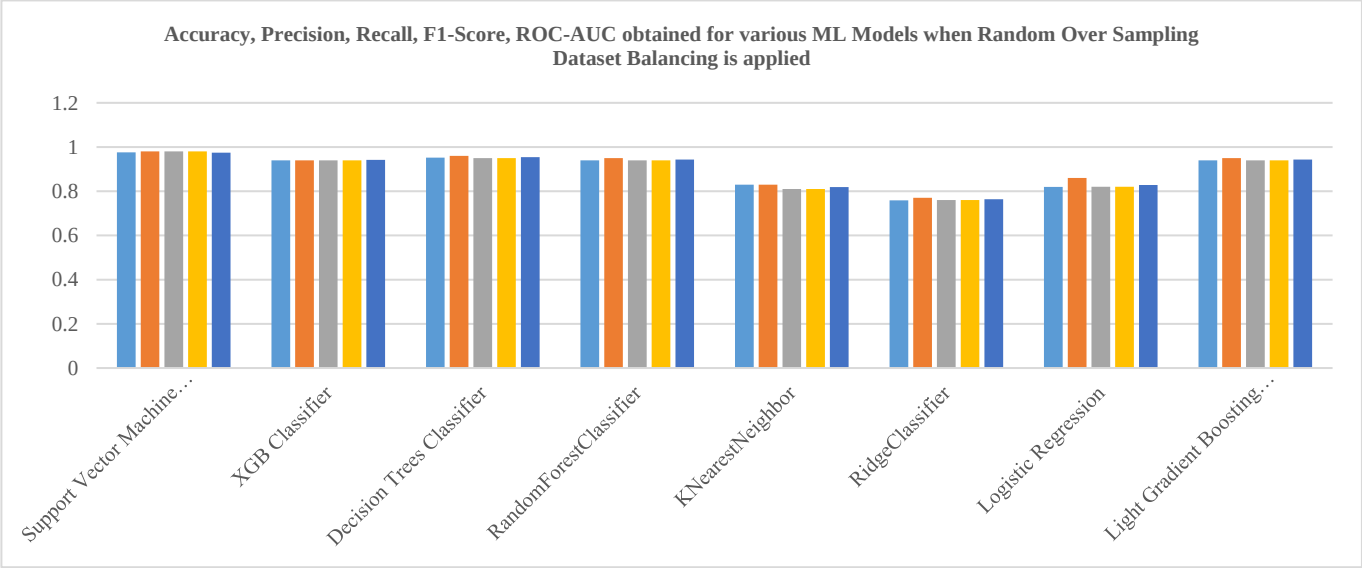


Fig. 10 Performance Comparison of ML Models on Balanced Data set using Random Oversampling without Feature Selection

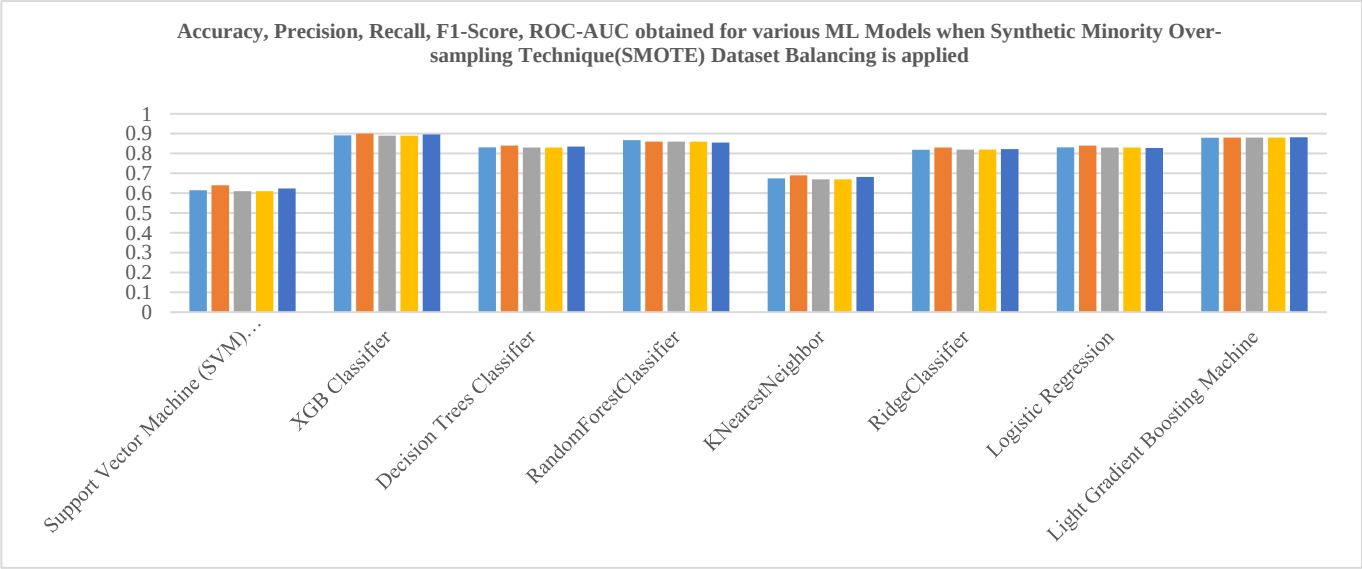


Fig. 11 Performance Comparison of ML Models on Balanced Data set using SMOTE without Feature Selection

Table 4. Performance Analysis of ML Models based on Balanced Dataset using SMOTE without Feature Selection

Feature Classification Model	Accuracy	Recall	Precision	F1-Score	ROC-AUC
SVM	0.6145	0.61	0.64	0.61	0.6233
XGB	0.8916	0.89	0.9	0.89	0.8963
DT	0.8313	0.83	0.84	0.83	0.8351
RF	0.8675	0.86	0.86	0.86	0.8552
kNN	0.6747	0.67	0.69	0.67	0.6815
Ridge	0.8193	0.82	0.83	0.82	0.8223
LR	0.8313	0.83	0.84	0.83	0.8276
LGBM	0.8795	0.88	0.88	0.88	0.882

Table 5. Performance Analysis of ML Models after Balancing the Dataset and Features Selection

New Name	Accuracy	Recall	Precision	F1-Score	ROC-AUC
SVM-RoS-A	0.9759	0.98	0.98	0.98	0.9744
SVM-RoS-MI	0.9759	0.98	0.98	0.98	0.9744
LGBM-RoS-RFE	0.9639	0.96	0.96	0.96	0.9645
DT-RoS-A	0.9518	0.95	0.96	0.95	0.9545
DT-RoS-MI	0.9518	0.95	0.96	0.95	0.9545
RF-RoS-A	0.9518	0.95	0.95	0.95	0.9531
XGB-RoS-A	0.9398	0.94	0.95	0.94	0.9432
XGB-RoS-MI	0.9277	0.93	0.94	0.93	0.9318
DT-RoS-RFE	0.9277	0.93	0.94	0.93	0.9318
LGBM-RoS-MI	0.9157	0.92	0.93	0.92	0.9205
LGBM-RoS-A	0.9157	0.92	0.92	0.92	0.919
RF-AoS-RFE	0.9176	0.92	0.92	0.92	0.9175
XGB-RoS-RFE	0.9157	0.92	0.92	0.92	0.9161
RF-RoS-MI	0.8916	0.89	0.9	0.89	0.8963
RF-RoS-RFE	0.8916	0.89	0.9	0.89	0.8963
LGBM-AoS-A	0.8941	0.89	0.9	0.89	0.8948
XGB-SoS-MI	0.8916	0.89	0.89	0.89	0.8919
RF-AoS-A	0.8824	0.88	0.89	0.88	0.8832
RF-SoS-RFE	0.8795	0.88	0.89	0.88	0.882
RF-SoS-MI	0.8795	0.88	0.88	0.88	0.8794
XGB-SoS-RFE	0.8675	0.87	0.87	0.87	0.8648
LGBM-SoS-RFE	0.8434	0.84	0.85	0.84	0.8479
XGB-AoS-RFE	0.8471	0.85	0.85	0.85	0.8477
RF-SoS-A	0.8434	0.84	0.84	0.84	0.8427
kNN-RoS-A	0.8313	0.83	0.88	0.83	0.8409
kNN-RoS-MI	0.8313	0.83	0.88	0.83	0.8409
LGBM-AoS-MI	0.8353	0.84	0.84	0.83	0.8361
LGBM-AoS-RFE	0.8353	0.84	0.84	0.84	0.8358
DT-AoS-MI	0.8353	0.84	0.84	0.84	0.8353
DT-SoS-RFE	0.8313	0.83	0.84	0.83	0.8351
XGB-SoS-A	0.8313	0.83	0.84	0.83	0.8336
LGBM-SoS-MI	0.8313	0.83	0.83	0.83	0.8322
DT-AoS-RFE	0.8235	0.82	0.83	0.82	0.8242
DT-AoS-A	0.8235	0.82	0.83	0.82	0.8239
SVM-SvSoS-MI	0.8209	0.82	0.83	0.82	0.8236
RF-AoS-MI	0.8235	0.82	0.82	0.82	0.8236
DT-SoS-A	0.8072	0.81	0.82	0.81	0.8138
XGB-AoS-A	0.8118	0.81	0.82	0.81	0.8128
DT-SoS-MI	0.7952	0.8	0.82	0.79	0.8024
XGB-AoS-MI	0.8	0.8	0.8	0.8	0.8004
SVM-AoS-MI	0.7882	0.79	0.79	0.79	0.7882

Ridge-AoS-A	0.7529	0.75	0.76	0.75	0.7536
LGBM-SoS-A	0.747	0.75	0.75	0.75	0.7497
Ridge-SoS-RFE	0.747	0.75	0.75	0.75	0.7453
kNN-AoS-MI	0.7412	0.74	0.77	0.73	0.7431
kNN-SoS-A	0.7108	0.71	0.73	0.71	0.7171
Ridge-SoS-A	0.6988	0.7	0.71	0.7	0.7028
LR-RoS-A	0.6867	0.68	0.72	0.69	0.6958
LR-RoS-MI	0.6867	0.69	0.72	0.68	0.6958
LR-AoS-A	0.6941	0.69	0.7	0.69	0.6946
LR-AoS-MI	0.6941	0.69	0.69	0.69	0.6941
LR-AoS-RFE	0.6941	0.68	0.68	0.68	0.6941
LR-RoS-RFE	0.6747	0.67	0.71	0.67	0.6844
Ridge-AoS-RFE	0.6824	0.68	0.68	0.68	0.6819
Ridge-RoS-A	0.6627	0.66	0.63	0.66	0.6629
LR-SoS-A	0.6627	0.66	0.66	0.66	0.6613
Ridge-RoS-RFE	0.6506	0.64	0.66	0.63	0.6559
SVM-SoS-MI	0.6506	0.65	0.65	0.65	0.6486
Ridge-AoS-MI	0.6353	0.6	0.65	0.6	0.637
LR-SoS-RFE	0.6386	0.64	0.64	0.64	0.6363
Ridge-RoS-MI	0.6145	0.61	0.61	0.61	0.6131
kNN-SoS-MI	0.6024	0.6	0.62	0.6	0.609
LR-SoS-MI	0.6024	0.6	0.6	0.6	0.6006
SVM-SoS-A	0.5783	0.58	0.59	0.58	0.5833
kNN-AoS-A	0.5765	0.58	0.58	0.57	0.5781
Ridge-SoS-MI	0.5663	0.57	0.57	0.56	0.5705
SVM-AoS-A	0.5529	0.55	0.56	0.55	0.554
SVM-SvSoS-A	0.6418	0.64	0.41	0.5	0.5

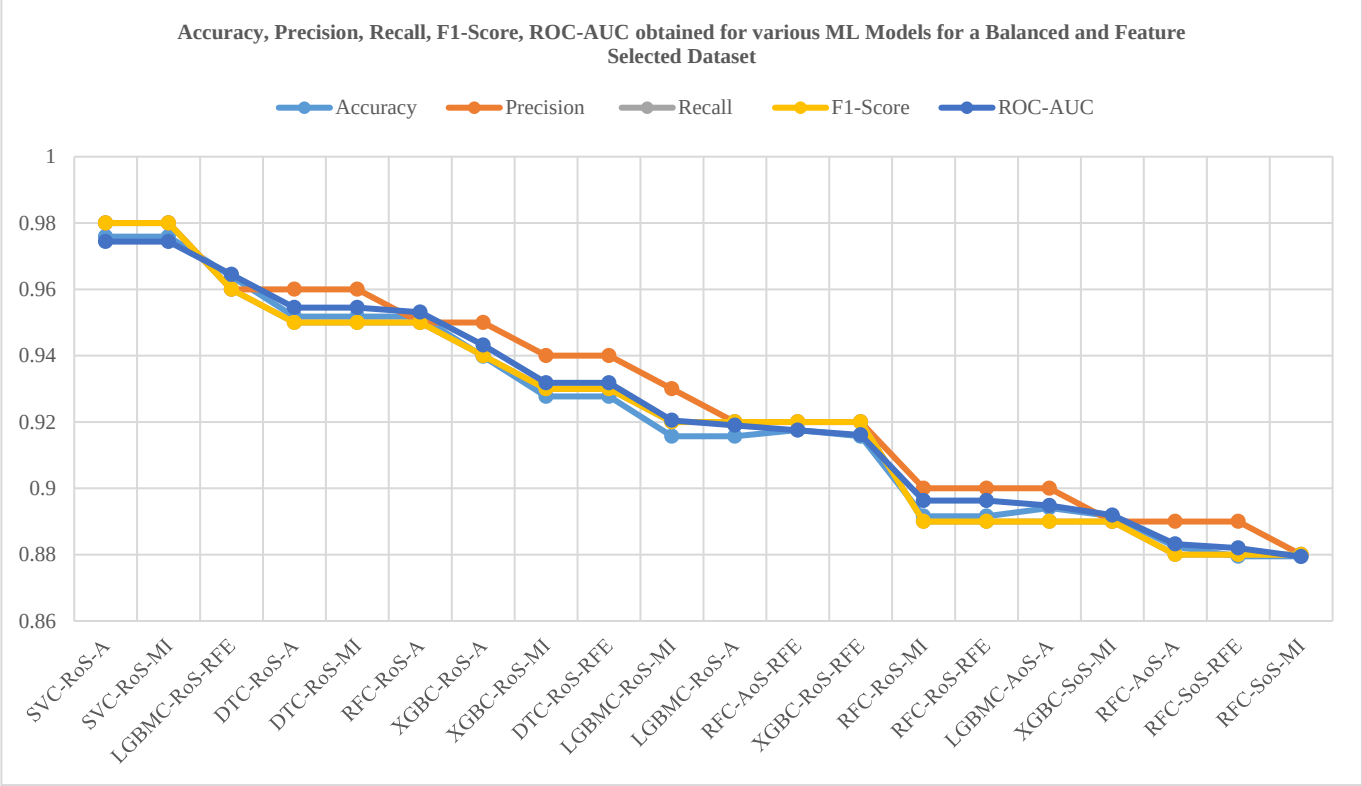


Fig. 12 Performance Comparison of ML Models on Balanced and Features Selected Dataset

Table 6. Comparison of Performance Analysis of ML Models with Existing Models

Feature Classification Model	Accuracy	Recall	ROC-AUC
GBT-HO-BFE [7]	0.767	0.435	0.774
RF-HO-BFE [7]	0.822	0.391	0.730
LR-LASSO [13]	0.90	0.86	0.93
SVM-LASSO [13]	0.82	0.86	0.90
DT-RoS-A [Proposed]	0.9518	0.95	0.9545
LGBM-RoS-RFE [Proposed]	0.9639	0.96	0.9645
SVM-RoS-A/MI [Proposed]	0.9759	0.98	0.9744

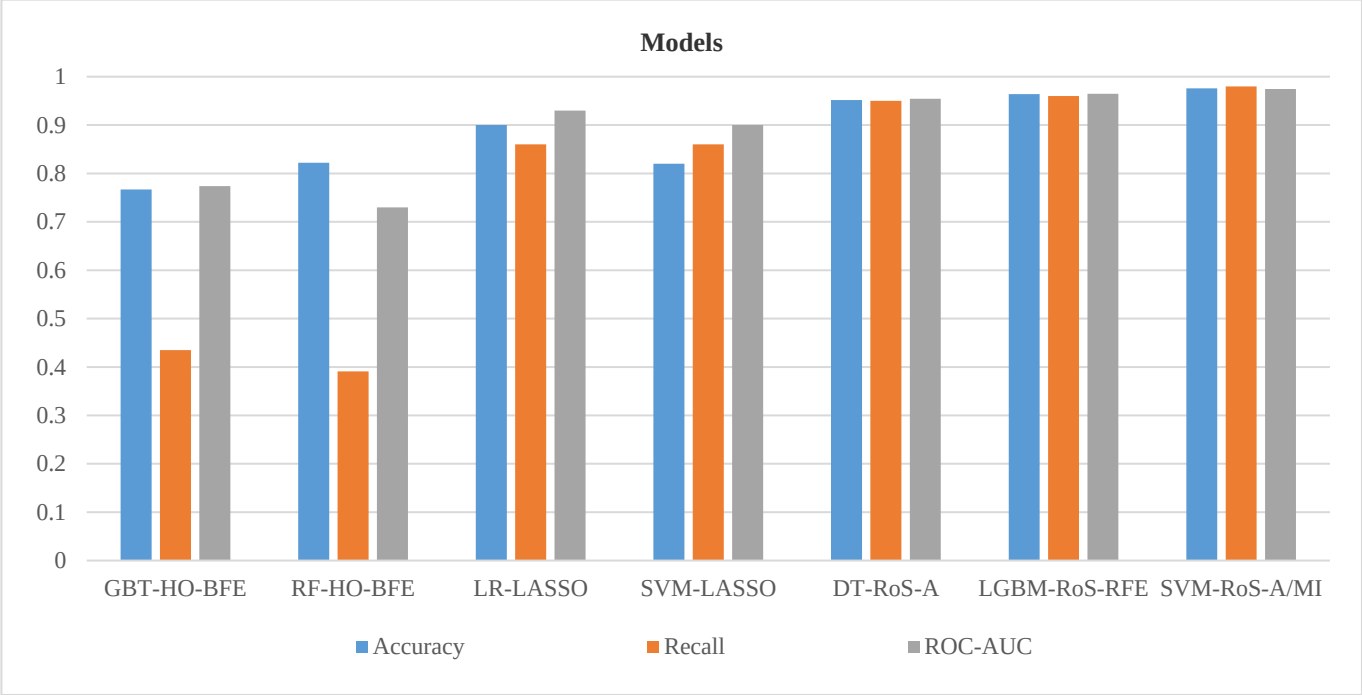


Fig. 13 Performance Comparison of Proposed ML Models with Existing Models

In the table 5, the performance analysis was ordered and arranged based on the models that obtained the scores higher to lower. Figure 12 was plotted based on the top 20 results obtained from the different ML models evaluated on the balanced dataset with features selection. Based on these top 20 results, the SVM model achieved the highest performance with random oversampling using ANOVA and MI feature selection methods, which obtained 97.59% accuracy, 98% precision, recall and f1-score and ROC-AUC value of 0.9744. The LGBM model with random oversampling and RFE has obtained 96.39% accuracy and 96% overall performance. The top 20 results include the models like SVM, LGBM, DT, and RF.

4.5. Performance Analysis Comparison

The performance of the proposed models was compared with the existing techniques derived from the literature survey for validation.

As shown in table 6, the proposed model's performance was compared with the existing models related to the PCa classification. The compared models are derived from the related works discussed in the literature. According to the comparison, the top three performed proposed research models are used in this comparison, which is better in terms of accuracy, recall and ROC-AUC parameters compared to the other models like Gradient Boosted Tree using hold-out (HO) and backward feature elimination (BFE) (GBT-HO-BFE) & RF-HO-BFE [7] and LR-LASSO & SVM-LASSO [13]. Based on the obtained results, it is clear that the SVM is the best ML model, Random sampling is the best data balancing model, and ANOVA and MI are the best feature selection models in this research.

5. Conclusion

A feature analysis model was proposed in this research for the classification of Radiomics features based on PCa MR Images. These MRI images comprise 76 patients of about 207 from the HCG, Bangalore. Onco-Radiologists from HCG Bangalore were involved in identifying the ROI that is by marking prostate lesions on the Original MR Image. The image masks were created based on the marking

provided on the Original Images. These image masks and the original images were used to extract features. The proposed model's workflow includes the following stages: image acquisition, ROI delineation, feature extraction, data balancing, selection of features, and analysis of features. ADASYN, SMOTE and random over-sampling methods were used for data balancing. For feature selection, ANOVA, MI and RFE methods were used to select the Radiomics features extracted from PCa MRI images. Different ML models such as SVM, XGB, DT, RF, kNN, Ridge, LR, LGBM, NB, Bagged DT, GBC and GBM were used for feature analysis. The performance analysis was carried out in three different stages. The feature analysis was performed on an unbalanced dataset without feature selection, a balanced dataset without feature selection and a balanced dataset with feature selection using the ML models. Based on the obtained results, the SVM ML model implemented with random oversampling and ANOVA and MI was the best feature analysis model in this research. In future, the limitation of this research can be solved by experimenting with more data samples for improved performance. Some conventional models should be included to validate the performance of the proposed model.

Authors' Contributions

All authors contributed to the Study, Design, Data Collection, MRI Feature Analysis and in writing or revising this manuscript. All authors have read and approved this version of the manuscript to be published and agree to be responsible for all aspects of the work.

Availability of Data and Materials

The PCa MR Image dataset used and analysed during the current study is a retrospective MR image available at HealthCare Global, Bengaluru, Karnataka, India.

Declarations

Ethics approval and consent to do research were obtained from the ethics committee of HealthCare Global, Bengaluru, Karnataka, India, using retrospective MR Images.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest concerning this article's research, authorship, and publication.

References

[1] Hyuna Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021. *Crossref*, http://doi.org/10.3322/caac.21660

[2] Freddie Bray et al., "Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394-424, 2018. *Crossref*, http://doi.org/10.3322/caac.21492

[3] Ahmad Chaddad et al., "Magnetic Resonance Imaging-Based Radiomic Models of Prostate Cancer: A Narrative Review," *Cancers*, vol. 13, no. 3, p. 552, 2021. *Crossref*, http://doi.org/10.3390/cancers13030552

[4] Giuseppe Cutaia et al., "Radiomics and Prostate MRI: Current Role and Future Applications," *Journal of Imaging*, vol. 7, no. 2, p. 34, 2021. *Crossref*, http://doi.org/10.3390/jimaging7020034

[5] Femke C.R. Staal et al., "Radiomics for the Prediction of Treatment Outcome and Survival in Patients with Colorectal Cancer: A Systematic Review," *Clinical Colorectal Cancer*, vol. 20, no. 1, pp. 52-71, 2021. *Crossref*, http://doi.org/10.1016/j.clcc.2020.11.001

[6] Parnian Afshary et al., "From Hand-Crafted to Deep Learning-based Cancer Radiomics: Challenges and Opportunities," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 132-160, 2019. *Crossref*, http://doi.org/10.1109/MSP.2019.2900993

[7] Leandro Donisi et al., "A Combined Radiomics and Machine Learning Approach to Distinguish Clinically Significant Prostate Lesions on a Publicly Available MRI Dataset," *Journal of Imaging*, vol. 7, no. 10, p. 215, 2021. *Crossref*, https://doi.org/10.3390/jimaging7100215

[8] Simon Bernatz et al., "Comparison of Machine Learning Algorithms to Predict Clinically Significant Prostate Cancer of the Peripheral Zone with Multiparametric MRI Using Clinical Assessment Categories and Radiomic Features," *European Radiology*, vol. 30, no. 12, pp. 6757–6769, 2020. *Crossref*, https://doi.org/10.1007/s00330-020-07064-5

[9] Jing Wang et al., "Machine Learning-Based Analysis of MR Radiomics Can Help to Improve the Diagnostic Performance of PI-RADS V2 in Clinically Relevant Prostate Cancer," *European Radiology*, vol. 27, no. 10, pp. 4082–4090, 2017. *Crossref*, https://doi.org/10.1007/s00330-017-4800-5

[10] Stefanie J. Hectors et al., "Magnetic Resonance Imaging Radiomics-Based Machine Learning Prediction of Clinically Significant Prostate Cancer in Equivocal PI-RADS 3 Lesions," *Journal of Magnetic Resonance Imaging*, vol. 54, no. 5, pp. 1466-1473, 2021. *Crossref*, https://doi.org/10.1002/jmri.27692

[11] Xiangde Mina et al., "Multi-Parametric MRI Based Radiomics Signature for Discriminating Between Clinically Significant and Insignificant Prostate Cancer: Cross-Validation of a Machine Learning Method," *European Journal of Radiology*, vol. 115, pp. 16-21, 2019. *Crossref*, https://doi.org/10.1016/j.ejrad.2019.03.010

[12] Bino Varghese et al., "Objective Risk Stratification of Prostate Cancer Using Machine Learning and Radiomics Applied to Multiparametric Magnetic Resonance Images," *Scientific Reports*, vol. 9, pp. 1-10, 2019. *Crossref*, https://doi.org/10.1038/s41598-018-38381-x

[13] B. Liu et al., "Prediction of Prostate Cancer Aggressiveness with a Combination of Radiomics and Machine Learning-Based Analysis of Dynamic Contrast-Enhanced MRI," *Clinical Radiology*, vol. 74, no. 11, pp. 896.e1-e896.e8, 2019. *Crossref*, https://doi.org/10.1016/j.crad.2019.07.011

[14] Valentina Brancato et al., "Evaluation of a Multiparametric MRI Radiomic-Based Approach for Stratification of Equivocal PI-RADS 3 and Upgraded PI-RADS 4 Prostatic Lesions," *Scientific Reports*, vol. 11, no. 643, pp. 1-10, 2021. *Crossref*, https://doi.org/10.1038/s41598-020-80749-5

[15] Anjana Gosain, and Saanchi Sardana, "Handling Class Imbalance Problem using Oversampling Techniques: A Review," *International Conference on Advances in Computing, Communications and Informatics*, pp. 79-85, 2017. *Crossref*, https://doi.org/10.1109/ICACCI.2017.8125820

[16] Jingmei Liu, Yuanbo Gao, and Fengjie Hu, "A Fast Network Intrusion Detection System Using Adaptive Synthetic Oversampling and Light GBM," *Computers & Security*, vol. 106, no. 102289, 2021. *Crossref*, https://doi.org/10.1016/j.cose.2021.102289

[17] V.P. Amadi, N.D Nwiabu, and V. I. E. Anireh, "Case-Based Reasoning System for the Diagnosis and Treatment of Breast, Cervical and Prostate Cancer," *SSRG International Journal of Computer Science and Engineering*, vol. 8, no. 8, pp. 13-20, 2021. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V8I8P103

[18] Hui Ding et al., "Identification of Bacteriophage Virion Proteins with the ANOVA Feature Selection and Analysis," *Molecular Bio Systems*, vol. 10, no. 8, pp. 2229-2235, 2014. *Crossref*, https://doi.org/10.1039/c4mb00316k

[19] Ke Yana, and David Zhang, "Feature Selection and Analysis on Correlated Gas Sensor Data with Recursive Feature Elimination," *Sensors and Actuators B: Chemical*, vol. 212, pp. 353–363, 2015. *Crossref*, https://doi.org/10.1016/j.snb.2015.02.025

[20] R. Surendiran et al., "Exploring the Cervical Cancer Prediction by Machine Learning and Deep Learning with Artificial Intelligence Approaches," *International Journal of Engineering Trends and Technology*, vol. 70, no. 7, pp. 94-107, 2022. *Crossref*, https://doi.org/10.14445/22315381/IJETT-V70I7P211

[21] Navoneel Chakrabarty et al., "Flight Arrival Delay Prediction Using Gradient Boosting Classifier," *Emerging Technologies in Data Mining and Information Security*, pp. 651–659, 2018. *Crossref*, https://doi.org/10.1007/978-981-13-1498-8_57

[22] C Narmatha, and Prasad M Surendra, "A Review on Prostate Cancer Detection using Deep Learning Techniques," *Journal of Computational Science and Intelligent Technologies*, vol. 1, no. 2, pp. 26-33, 2020. *Crossref*, https://doi.org/10.53409/mnaa.jcsit20201204

[23] S. Manimurugan, "IoT-Fog-Cloud model for Anomaly Detection Using Improved Naive Bayes and Principal Component Analysis," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-10, 2021. *Crossref*, https://doi.org/10.1007/s12652-020-02723-3

[24] S. Rinesh et al., "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms," *Journal of Healthcare Engineering*, vol. 2022, pp. 1-9, 2022. *Crossref*, https://doi.org/10.1155/2022/2761847

[25] T. Rajendran, K.P. Sridhar, and S. Deepa, "Performance Analysis of Fuzzy Multilayer Support Vector Machine for Epileptic Seizure Disorder Classification Using Auto Regression Features," *The open Biomedical Engineering Journal*, vol. 13, pp. 103-113, 2019. *Crossref*, https://doi.org/10.2174/1874120701913010103

[26] Jorge R. Vergara, and Pablo A. Estévez, "A Review of Feature Selection Methods Based on Mutual Information," *Neural Computing and Applications*, vol. 24, pp. 175–186, 2014. *Crossref*, https://doi.org/10.1007/s00521-013-1368-0