Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Evaluating Machine Learning Classifiers for Prostate Cancer Diagnosis: A Comparative Study

Ravikumar R N[1], Sheikh Md. Jaimul Jakir[1], N. Sivakumar[1], Pappu Kumar Rai[1], Anand Singh[1], Mariappan Kandasamy[2]

[1]Dept. of Computer Engineering, Marwadi University, Rajkot, India
[2]School of Computer Science and Engineering, Chandigarh University, Punjab, India
Email: rnravikumar.cse@gmail.com

Prostate cancer is considered to be among the most common and deadly diseases all over the globe. In the present work, different machine learning models were applied to the prostate cancer database for diagnosis classification. The Kaggle dataset has 100 observations with 10 variables of which 8 are continuous variables and 1 is categorical. The traditional set of classification algorithms including k Nearest Neighbor (kNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest, Logistic Regression and Gradient Boosting, Ada Boost, and Extra Trees were used. The effectiveness of the models was assessed before and after hyperparameter optimization. When it came to iteratively calculating the accuracy the best models that were attained were only Logistic Regression followed by AdaBoost getting the best of 90% while the Accuracy stochastically varied for both models as seen in the following graphs. Other models had diverse performance with KNN, decision tree, SVM, random forest, and gradient boosting models ranging between 73% and 83 % accuracy. Overall, Naive Bayes had the lowest scores, but it had the same accuracy of 77% at all iterations. The outcomes of this analysis underline the ability of the proposed machine learning models in identifying the results of the prostate cancer diagnosis builds the foundation on which further research might be based; Thus, the Models that showed the best performance are AdaBoost and logistic regression.

Keywords: Machine learning, Classification, Ensemble methods, Gradient Boosting, Hyperparameter tuning, Prostate cancer.

1. Introduction

Prostate Cancer (PC) is one of the dangerous cancers in males and is the third leading cause of death from cancer. Clinical and biological research have been the main sources for the advancement of therapies and knowledge about the disease's features. The introduction of machine learning has profoundly transformed several facets of contemporary living, included healthcare [1]. Machine learning is widely used in healthcare to forecast patient outcomes, including the likelihood of survival. Although machine learning has the potential to be used for medical predictions, there are still some hurdles that need to be overcome. An important obstacle is the lack of medical data for public study [2]. The fig 1 shoes the types of PC.

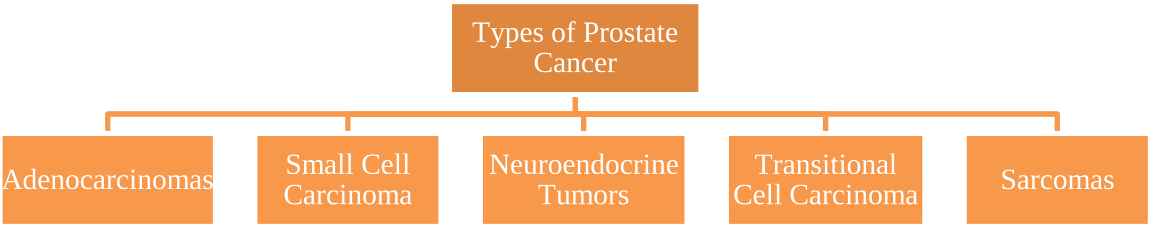


Fig. 1. Types of Prostate Cancer

Adenocarcinoma is the most frequent form of prostate cancer, arising specifically from the prostate glands responsible for prostate fluid, a component of the male semen. A Small Cell Carcinoma is a very aggressive cancer that forms from neuroendocrine cells present in the prostate and tends to proliferate and metastasize at a faster rate. Tumors of this kind arise from prostate neuroendocrine cells, are propelled by large cells of abnormal structure and may exhibit forceful growth. Transitional Cell Carcinoma is a rare type of cancer that generally arises from cells of the urethra and bladder but can also arise in the prostate, especially in the prostate ducts. Fig. 2 shows the stages of PC.

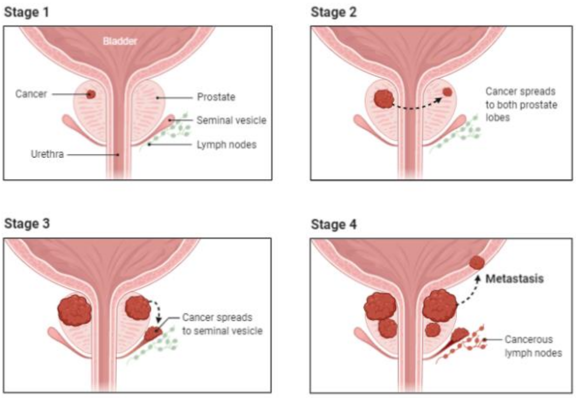


Fig. 2. Stages of Prostate Cancer (Biorender [3])

One of the major challenges in leveraging ML for healthcare, particularly for predicting patient survivability, is the lack of publicly available medical data. Additionally, while various machine learning algorithms can predict survival times, it remains unclear which techniques yield the most accurate results. Most of the research so far has relied on traditional machine learning methods, and there is a need for comprehensive evaluations to determine the best-performing models.

The main contributions of our work include to:

- Visualize the dataset and assess the performance of various ML classifiers on a prostate cancer dataset.
- Analyze model accuracy before and after hyperparameter tuning to determine its impact.
- Identify the most accurate models for PC diagnosis.
- Highlight the potential of ML models in predicting PC diagnosis results

The rest of the sections of the research are structured in the following manner: Section 1 examines the findings of prior studies on cancer detection and the ideas of the chosen machine learning techniques. Section 2 provides an in-depth description of the dataset used and introduces the assessment methodology. Section 3 shows the results. Section 4 and 5 shows future directions and conclusion.

2. Literature Review

A recent study compares four machine learning techniques—"k-Nearest Neighbors (KNN), Decision Tree (DT), Naïve Bayes (NB), and Support Vector Machine (SVM)”—for predicting prostate cancer survivability using the SEER (Surveillance, Epidemiology, and End Result) database. The paper highlights the effectiveness of these traditional machine learning algorithms alongside an “artificial neural network (ANN)”, demonstrating the superior accuracy of ANN in predicting patient survivability. The neural network achieved an 85.64% accuracy, outperforming other techniques. The work also reveals the difficulties inherent to making medical data public and the necessity to choose high impact features for increasing the accuracy of predictions. Future work is likely to use rent models and improve upon which features to pick to improve performance, especially for less frequent survival types [4].

This work is also aligned with the objectives of attempting to categorize prostate cancer and determine the potential risks factors leveraging on machine learning tools. Patient data samples from a hundred of prostate cancer patients were used to develop a prediction model based on the RF algorithm with the use of MCCV. Measures of accuracy, sensitivity, specificity as well as AUC were used. Sensitivity, specificity, PPV, NPV, accuracy, F1 score and AUC for RF model was 0.89, 0.84, 0.77, 0.93, 0.86, 0.83 and 0.88. From the aforementioned analysis, area, perimeter and texture proved to be the most significant risk factors. The RF algorithm helps to predict prostate cancer and the key risk factors as a result of this study may be useful in diagnosis and management of the disease [5].

In this work, MPI-PCA aims to construct classifiers for the recognition of PC using machine learning algorithms and prebiopsy data. The following four machine learning techniques were used: In this study, there are four models, namely Random Forest (RF), Artificial Neural Network (ANN), Least Squares Support Vector Machine (LS-SVM), and Support Vector Machine (SVM). The performance of the models was based on a sample of 1625 Chinese males who had prostate biopsies. The study shows that it is true that all paths could successfully detect seemingly distinctive PC situations: yet, ANN proved to offer greater accuracy and AUC value. In addition, RF proved to be more effective than the other techniques when it comes to selecting those PC instances that are benign, important, and inconsequential. The work shows the application of machine learning to PCs' diagnosis, particularly for different population groups such as the Chinese. Further analysis is still needed to enhance the rate of classification and additionally – they should be checked on the different population [6].

This consisted in developing robust deep learning models to classify skin cancer into its appropriate category. Such issues include handling of a large imbalanced class distribution and providing explanations of the decisions arrived at by them. From at about HAM10000 datasets and the improved CNNs the study differentiates seven skin cancers. Only optimization and activation functions are used to train models. Further, the methods such as Grad-CAM and Grad-CAM++ from explainable AI (XAI) are used for model interpretation. For the experimental evaluation with 121 images it has revealed a classification accuracy of 82% and loss accuracy of 0.47 also prove that deep learning can be applied for enhancing the skin cancer diagnosis [7].

The systematic review focuses on how ML and DL have been adapted for diagnosing PCa by using medical images. It points to the need to take PCa seriously because it is one of the major causes of malignancy and death among men across the world particularly in the developing world. The review compiles 77 studies and covers important aspects, including the trends in the field, the choice of dataset, and the method used. Major conclusions include the fact that majority of the research works stem from the United States, majority of the used datasets being MRI datasets, and majority of the used methods being transfer learning methods for the diagnosis of PCa. The study is in support of the future possibility of using ML and DL in the management and diagnosis of PCa, providing the researcher and clinical professionals significant discussions and suggestions [8].

This paper demonstrates a work on Pc prediction utilizing Supervised ML classification strategies. To address the clinical along with tumor stage attributes, a novel "Logistic Regression (LR)" classifier is thus introduced and utilized. American Joint Committee on Cancer (AJCC) and Tumor Node Metastasis (TNM) are features of the tumor stage, while clinical characteristics are BMI, age, history of cystitis infections and smoking. Even if our suggested diagnosis increases specificity (Sp) for the identification of PC by 4%, the value of accuracy and "positive prediction value (PPV)" is also higher compared to contemporary classifiers. 387 male patients, 188 of whom have Pc, make up the study's dataset. A comparison and implementation of several ML classifiers reveal that the suggested modified LR classifier performs better than the others, improving prediction by at least 3% [9].

Prostate cancer is a significant health issue for older men, with rising incidence in China. Early detection via MRI is crucial but challenging. This study uses a deep learning "Convolutional Neural Network (CNN)" to classify prostate cancer from 10,056 diffusionweighted MRI images. The model achieved training and testing accuracy rates of 80.15% and 78.15%. This approach shows promise for aiding prostate cancer diagnosis and staging, with potential applications for other cancers. Future improvements could involve different MRI configurations, deeper networks, and additional image processing techniques [10].

A man’s health fatal disease which stands as the leading cause of cancer related death in North America is best detected early since it has high survival rates. The PSA tests and MRI are traditional ways of diagnosing the disease and most of the time lead to over diagnosis and biopsies. In this work, an example of detecting clinically relevant PCa employing deep CNNs in an automated pipeline is demonstrated, with the DWI being employed as a model input. The pipeline that does not required the definition of ROIs was applied to 427 patients (175 with PCa and 252 without); the achieved accuracy was high with the AUC of 0.87 at the slice level and 0.84 at the patient level. A potential strength in these methods is the strong and clinically relevant compared with the conventional ROI-based solutions that can provide higher accuracy and reliability for PCa detection [11].

This research aims to apply machine learning to enhance the preliminary stage diagnosis of cancer, including lung, prostate and breast cancer. Through the obtained data, the study assesses the effectiveness of different machine learning algorithms by examining their relevant indicators, including accuracy, precision, recall, as well as F1-score. The best performers models, in this case, are Simple LR model for lung cancer and Least Regression (LR) model for both the Prostate and Breast Cancers. These models were then incorporated in a Python Flask application to identify whether the user input suggests a benign or malignant disease. The study will help improve the early detection of cancer to decrease chances of death and treatment expenses [12].

In the paper, the authors provide a new approach to identifying the presence of prostate cancer in histopathological slides through deep learning. Through using a streaming implementation of convolutional layers on multi-scale, the final model can be trained directly on the high-resolution biopsy pictures simply using slide-level labels without necessary detailed annotations. This way gives nearly the same results as existing approaches and does not require much GPU memory. According to the research, there is an improved RT-qPCR detection of Proviral integration in prostate carcinoma meaning a higher potential for efficient and accurate prognosis [13].

In the presented work, the authors suggest an effective binary classification of prostate cancer using MRI, CT, PET and X-Ray based radiomic data. To enhance the classification performance, the proposed method employs linear SVM and ridge regression framework for feature selection. On the PROSTATEx dataset the derived model is above 90% accuracy when trained on particular subsets of features and 43.64 % when trained on the whole feature suite [15]. The work emphasizes feature selection as a means to improving diagnosis when it involves the use of machine learning algorithm. It also demonstrates the potential for accomplishing a higher level of classification specificity in the classification of prostate cancer. More evidence is needed to corroborate these findings and standardize the approaches used for feature selection in radiomics datasets [15]. The present study is concerned with the identification of biomarkers for early cancer diagnosis by applying gene expression profiles. It introduces an ECO- based gene selection method and a DSCNN-based classification model for the disease diagnosis. The developed model is 0.45% to 1.16% more accurate than existing methods in cross-dataset cases; 99.18% average accuracy. Such findings have been further supported by statistical tests which establish its significance [16].

This research develops a computerized tool for early cancer detection using gene expression analysis in breast, liver, lung, prostate, and renal cancers. It combines sparse autoencoder for dimensionality reduction and Remora Optimization for feature transformation, achieving an average balanced accuracy of 93.4% across 1027 patients. However, there are challenges like imbalanced precision and recall scores and increased computational time. Future work could explore alternative optimization techniques for improved cancer diagnosis through gene expression analysis [17].

This study presents GEDAAI-PCD, a new method using deep learning (LSTM-DBN) to accurately classify prostate cancer from gene expression data. It tackles challenges in hyperparameter tuning with EWHO optimization, showing superior performance. The goal is to improve prostate cancer precision medicine using AI [18]. The research compares LSTM and ResNet-101 deep learning algorithms for prostate cancer diagnosis to conventional approaches. Deep learning worked best, with ResNet-101 winning. Prostate cancer detection remains a crucial challenge in healthcare, with deep learning showing promise in improving diagnostic accuracy [19].

Raman spectroscopy (RS) is a valuable technique for molecular analysis, particularly in cancer diagnosis. While RS offers high specificity, its slow acquisition rate and limited depth information pose challenges. Integrating AI, specifically ML and DL, with RS can enhance diagnostic accuracy [20]. This review explores AI's recent applications in RS analysis for cancer grading, focusing on seven major cancer types. By summarizing current achievements and discussing methodologies, it aims to advance RS as a rapid and effective cancer screening tool. The review provides insights into overcoming challenges and encourages further research in RS and AI applications in oncology [21].

The paper proposes a hybrid RNN-CNN method for predicting cancer types from gene expression data, outperforming existing methods. It reviews the significance of early cancer detection and the challenges in handling gene expression data. Using DL techniques, the paper addresses issues like high dimensionality and low sample size. Evaluation on two datasets validates the effectiveness of the proposed method in cancer classification. The paper explores cancer prognosis prediction models, particularly focusing on thyroid cancer survival. It utilizes machine learning techniques and the SEER database, aiming to improve accuracy and address data imbalance issues. The presented work can be seen as falling into a useful framework of progress in terms of survival prediction for thyroid cancer [22].

3. Methodology

3.1. Prostate Cancer Diagnosis Dataset:

The used data set contains 100 patient’s observations with 10 variables, in numerical and categoreal forms. Some of the fields in each record are “Radius, Texture, Perimeter, Area, Smoothness, Compactness, diagnosis_result, Symmetry, Fractal dimension”. The diagnosis_result can be Benign (B) or Malignant (M), it is a categorical regressed variable. It is mostly suited for applying and assessing the performance of ML techniques for the prognosis of the diagnosis of prostate cancer.

3.2. Dataset Description:

Table 1 Features in Dataset

Feature Name	Explanation	Data Type
id	Unique identifier for each record	int
diagnosis_result	Diagnosis result (M = Malignant, B = Benign)	str
radius	Mean radius of the tumor	float
texture	Mean texture of the tumor	float
perimeter	Mean perimeter of the tumor	float
area	Mean area of the tumor	float
smoothness	Mean smoothness of the tumor	float
compactness	Mean compactness of the tumor	float
symmetry	Mean symmetry of the tumor	float
fractal_dimension	Mean fractal dimension of the tumor	float

3.3. Model Training:

1. KNN is an easy, simple and a kind of supervised approach to be applied both for regression and classification problems.

2. The Decision Tree is actually a supervised learning method that is also used for classification and regression analysis.

3. Naive Bayes works as classifier using the Bayes theorem but it assumes that all features in a given data set are independent of each other.

4. VM stands out as a widely recognized model in supervised learning, commonly applied in classification and regression tasks.

5. Logistic Regression (LR) represents a linear model employed for binary classification to predict the likelihood of belonging to a specific class based on any input.

6. Random Forest operates as an ensemble approach that builds numerous decision trees during training and provides the class output, either the mode of classes for classification or mean prediction for regression.

7. Gradient Boosting is an ensemble method that trains models sequentially, with each subsequent model rectifying the errors of its predecessor.

8. AdaBoost, short for Adaptive Boosting, is a boosting technique that transforms a group of weak learners into a robust learner.

9. Extra Trees, or Extremely Randomized Trees, offers another ensemble method akin to Random Forest but with a heightened level of randomness in constructing individual trees.

3.4. Visualization:

These are the visualization Correlation Heatmap, plot distribution of numeric features and Pair plot for subset of features shown in fig 3, fig 4, fig 5.

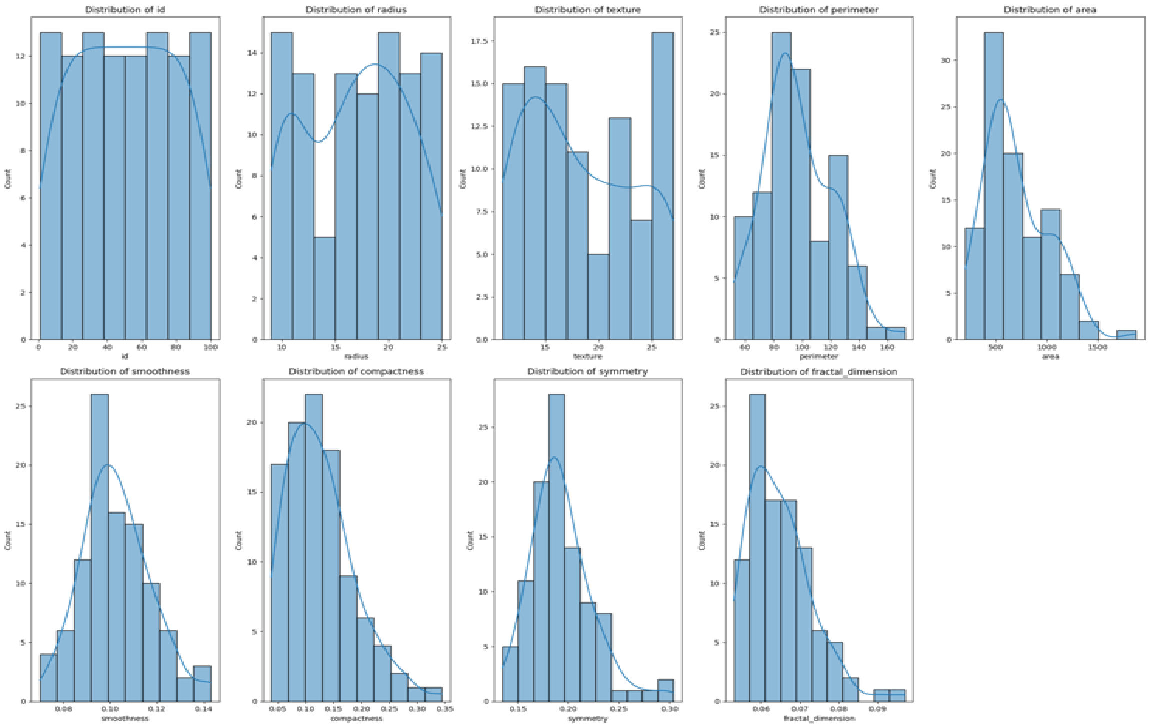
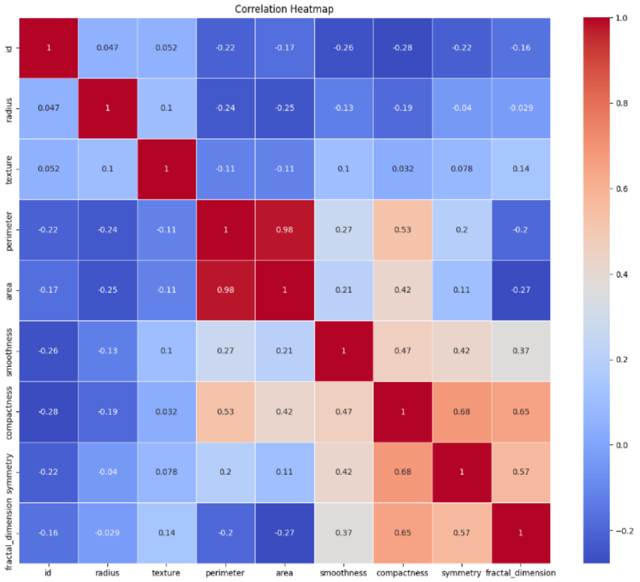


Fig. 3. Plot distribution of numeric features



	id	radius	texture	perimeter	area	smoothness	compactness	symmetry	fractal_dimension
id	1	0.047	0.052	-0.22	-0.17	-0.26	-0.28	-0.22	-0.16
radius	0.047	1	0.1	-0.24	-0.25	-0.13	-0.19	-0.04	-0.029
texture	0.052	0.1	1	-0.11	-0.11	0.1	0.032	0.078	0.14
perimeter	-0.22	-0.24	-0.11	1	0.98	0.27	0.53	0.2	-0.2
area	-0.17	-0.25	-0.11	0.98	1	0.21	0.42	0.11	-0.27
smoothness	-0.26	-0.13	0.1	0.27	0.21	1	0.47	0.42	0.37
compactness	-0.28	-0.19	0.032	0.53	0.42	0.47	1	0.68	0.65
symmetry	-0.22	-0.04	0.078	0.2	0.11	0.42	0.68	1	0.57
fractal_dimension	-0.16	-0.029	0.14	-0.2	-0.27	0.37	0.65	0.57	1

Fig. 4. Correlation Heatmap

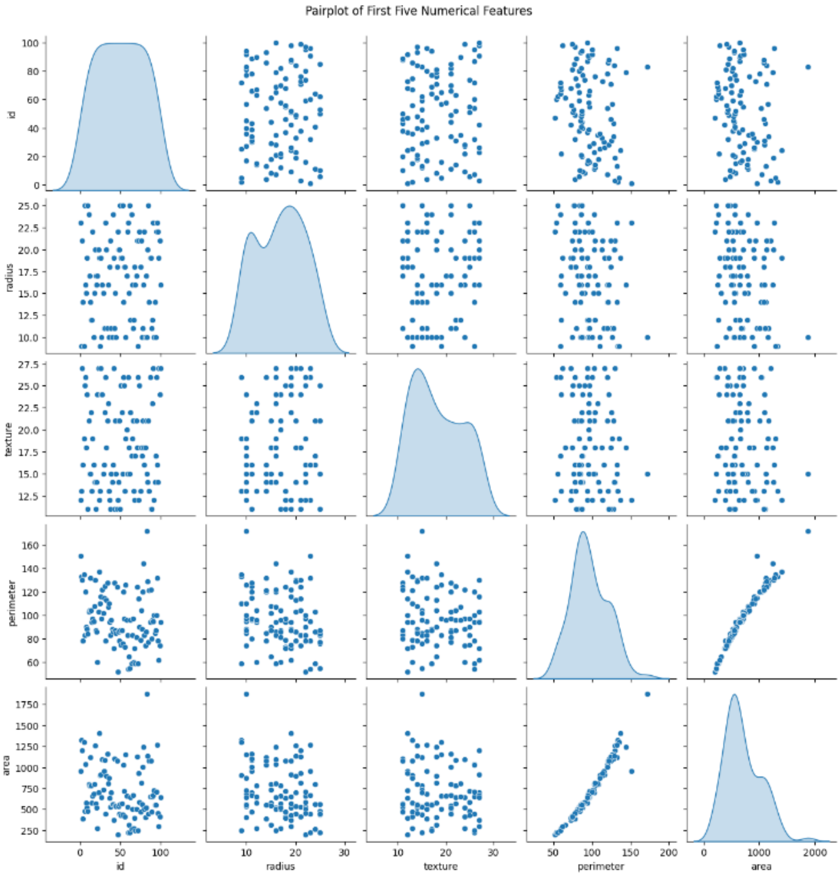


Fig. 5. Pair plot for a subset of features

3.5. Model Evaluation:

Key indicators for evaluating classification models consist of precision, recall, and F1 score. Precision represents the proportion of correct positive forecasts out of all positive predictions. It gauges the accuracy of identifying positive instances. Within your code setting, precision is computed for each classifier and detailed in the classification report.

On the other hand, recall, also known as sensitivity or true positive rate, is the ratio of true positive predictions to the total number of actual positive instances. It quantifies the effectiveness of recognizing real positive samples by the model. Precision and recall are harmonic means, and the F1 score strikes a balance between both. It is helpful for assessing a model's overall performance, particularly in the case of unbalanced datasets.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (1)$$

$$\text{Precission} = \frac{\text{TP}}{\text{TP+FP}} \qquad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (3)$$

$$\mathrm{F1-Score} = \frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \qquad (4)$$

These metrics are crucial in assessing the suitability of a classification model.

4. Results

The accuracy scores of all the models are plotted using Bar Plot and it shows that before Hyper Parameter tuning the top maximum accuracy 87 obtained by Logistic regression. And accuracy 83 obtained by SVM, AdaBoost. Accuracy 80 obtained by KNN, accuracy 77 obtained by DT, RF, NB and GB. Accuracy 70 obtained by ET respectively as shown in Table 2.

After Hyper Parameter Tuning maximum accuracy 90 obtained by LR AND AB. 83 obtained by SVM. Accuracy 80 obtained by KNN. Accuracy 77 obtained by DT, NB, GB. Accuracy 73 obtained by RF and ET as shown in Table 3. The below figure 4 shows the comparison of before and after tuning performances.

Table 2 Before Hyper Parameter Tuning

Classifier	Class	Precision	Recall	F1-score	Support	Accuracy
KNN	B	0.62	0.62	0.62	8	0.80
	M	0.86	0.86	0.86	22	
DT	B	0.55	0.75	0.63	8	0.77
	M	0.89	0.77	0.83	22	
SVM	B	0.67	0.75	0.71	8	0.83
	M	0.90	0.86	0.88	22	
RF	B	0.56	0.62	0.59	8	0.77
	M	0.86	0.82	0.84	22	
LR	B	0.75	0.75	0.75	8	0.87
	M	0.91	0.91	0.91	22	
NB	B	0.55	0.75	0.63	8	0.77
	M	0.89	0.77	0.83	22	
GB	B	0.56	0.62	0.59	8	0.77
	M	0.86	0.82	0.84	22	
AB	B	0.71	0.62	0.67	8	0.83
	M	0.87	0.91	0.89	22	
ET	B	0.45	0.62	0.53	8	0.70
	M	0.84	0.73	0.78	22	

Table 3 After Hyper Parameter Tuning

Classifier	Class	Precision	Recall	F1-score	Support	Accuracy
KNN	B	0.62	0.62	0.62	8	0.80
	M	0.86	0.86	0.86	22	
DT	B	0.56	0.62	0.59	8	0.77
	M	0.86	0.82	0.84	22	
SVM	B	0.67	0.75	0.71	8	0.83
	M	0.90	0.86	0.88	22	
RF	B	0.50	0.62	0.56	8	0.73
	M	0.85	0.77	0.81	22	
LR	B	0.86	0.75	0.80	8	0.90
	M	0.91	0.95	0.93	22	
NB	B	0.55	0.75	0.63	8	0.77
	M	0.89	0.77	0.83	22	
GB	B	0.56	0.62	0.59	8	0.77
	M	0.86	0.82	0.84	22	
AB	B	0.86	0.75	0.80	8	0.90
	M	0.91	0.95	0.93	22	
ET	B	0.50	0.62	0.56	8	0.73

M	0.85	0.77	0.81	22

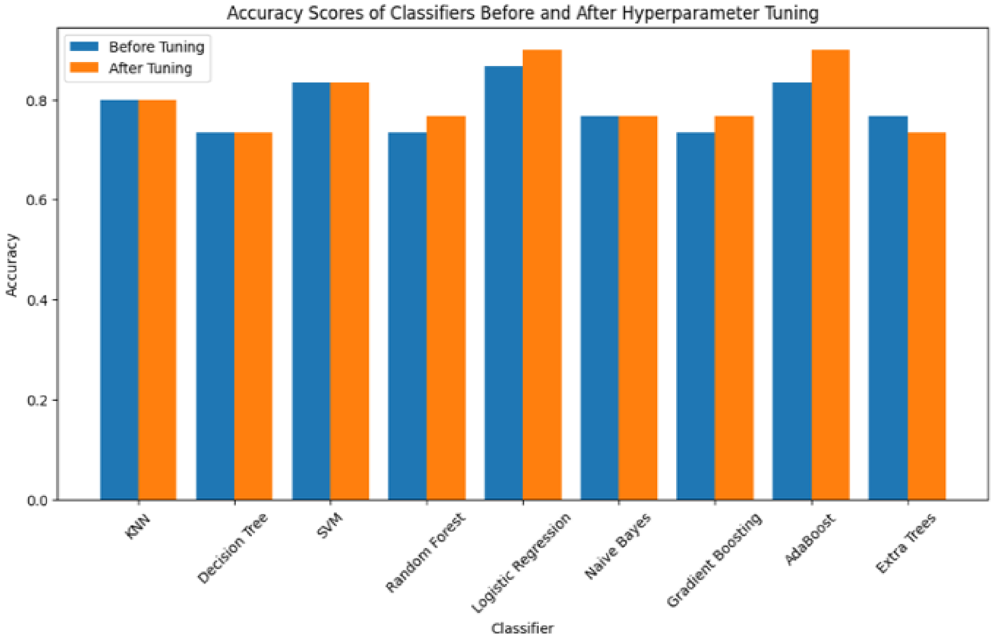


Fig. 4. Result Graph Before and After Hyper Parameter Tuning

5. Future Directions

The fairly high accuracy of different models implies that the features used in this study can be used to predict the trend of prostate cancer incidence with high accuracy. It would be beneficial for the future work to try with the larger and more diversified dataset in order to make the result as generalized as possible. Someday, it could be useful to investigate more detailed models as well as other as yet unnoticed characteristics that would make the predictions more accurate and lessen effect of bias in the sample. Some of the other measures that should be used in the analysis include cross validation and hyper parameter tuning. Besides, ideas like ensemble handling with methods like bagging, boosting, or feature learning handling with CNN, could be employed to increase the depth of pattern search within the data.

6. Conclusion

Therefore, accurate and reliable projection of trends of prostate cancer incidence rates is important in planning and functioning of care facilities. Concerning this study, we used several machine learning techniques to forecast the trend in the PRRs using the dataset acquired. The classifiers utilized are KNN, DT, SVM, RF, LR, NB, GB, AdaBoost and Extra Trees. In the bar plot of the accuracy score, we can observe that before hyperparameter tuning, Logistic Regression had a maximum accuracy of 87. There was a better achievement of accuracy of 83% for both SVM and AdaBoost. Decision Tree's accuracy was 77, Random Forest, Naive Bayes and Gradient Boosting were also 77, KNN was 80. Further, Extra Trees with 70% accuracy furnishes the conclusion. After tuning hyperparameters, it was discovered that the best accuracy of 90 belonged to two models, Logistic Regression and AdaBoost. SVM provided the accuracy of 83 %, KNN which upheld the accuracy of 80, while Decision tree, Naive Bayes and Gradient boosting provided the accuracy of 77%. Random Forest and Extra Trees models had an accuracy of 73%. Finally, overall, the study shows that the prognosis is correct regarding the trends of prostate cancer, which is encouraging for clinical trials.

Acknowledgement

This work is an outcome of Student Research Internship Training (SRIT) program from Department of Computer Engineering, Marwadi University, and collaborative work from Chandigarh University.

Notes on Contributors

The following author's contribution as follows: RN, SJ – Initial draft, LR, Implementation;

NS – Methodology, Review; PR, AS – Results formulation, Diagrams, Review; RN – Supervision.

References

1. Chen, S., Zhang, Y., Li, X., & Wang, Z. (2022). Machine learning-based models enhance the prediction of prostate cancer. Frontiers in Oncology, 12, 941349. https://doi.org/10.3389/fonc.2022.941349 (Journal Article)
2. Li, R., Zhu, J., De Zhong, W., & Jia, Z. (2022). Comprehensive evaluation of machine learning models and gene expression signatures for prostate cancer prognosis using large population cohorts. Cancer Research, 82(9), 1832–1843. https://doi.org/10.1158/0008-5472.CAN-21-3074 (Journal Article)
3. https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/staging.html (Web Link)
4. Wen, H., Li, S., Li, W., Li, J., & Yin, C. (2018). Comparison of four machine learning techniques for the prediction of prostate cancer survivability. In 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), IEEE, Dec. 2018. (Conference Paper)
5. Özhan, O. (2022). Machine learning approach for classification of prostate cancer based on clinical biomarkers. Journal of Cognitive Systems, Dec. 2022. (Journal Article)
6. Wang, G., Teoh, J. Y.-C., & Choi, K.-S. (2018). Diagnosis of prostate cancer in a Chinese population by using machine learning methods. Annual International Conference IEEE Engineering in Medicine and Biology Society, 2018, 1–4. (Conference Paper)
7. Mridha, M. M. U. J. S. S. K. A. M. F. M. K. (2023). An interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system. IEEE Access, 11. (Journal Article)
8. Olabanjo, O., Akindele, A. O., & Adedeji, A. (2023). Application of machine learning and deep learning models for prostate cancer diagnosis using medical images: A systematic review. Analytics, 2023. (Journal Article)
9. Ismail, M. A. M., T. H. K., V. A. P. R., & M. Ismail. (2020). A machine learning classification

technique for predicting prostate cancer. In 2020 IEEE International Conference on Electro Information Technology (EIT), Chicago, IL, USA. (Conference Paper)

10. Liu, Y., & An, X. (2017). A classification model for prostate cancer based on deep learning. In 2017 10th International Congress on Image and Signal Processing, Biomedical Engineering and Informatics (CISP-BMEI), IEEE, Oct. 2017. (Conference Paper)
11. Yoo, S., Gujrathi, I., Haider, M. A., & Khalvati, F. (2019). Prostate cancer detection using deep convolutional neural networks. (Unpublished Paper) (Conference Paper)
12. Verma, A., Shah, C. K., Kaur, V., Shah, S., & Kumar, P. (2022). Cancer detection and analysis using machine learning. In 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), IEEE, Sep. 2022. (Conference Paper)
13. Pinckaers, H., Bulten, W., van der Laak, J., & Litjens, G. (2021). Detection of prostate cancer in whole-slide images through end-to-end training with image-level labels. IEEE Transactions on Medical Imaging, 40(7), 1817–1826. https://doi.org/10.1109/TMI.2021.3068256 (Journal Article)
14. A. B. M. M. Varan. (2023). Enhancing prostate cancer classification by leveraging key radiomics features and using the fine-tuned linear SVM algorithm. IEEE Access, 11. (Journal Article)
15. Sethi, K. S., S. K. R., & B. K. P. S. (2024). Long short-term memory-deep belief network-based gene expression data analysis for prostate cancer detection and classification. IEEE Access, 12. (Journal Article)
16. Matte, S., D. L. F. & J. G. M. Varan. (2024). Prostate cancer detection using deep learning and traditional techniques. JoCNV, 2(1), 1–8. https://doi.org/10.1109/JoCNV.2024.3045873 (Journal Article)
17. Conforti, P. M., Lazzini, G., Russo, P., & D’Acunto, M. (2024). Raman spectroscopy and AI applications in cancer grading: An overview. IEEE Access, 12, 54816–54852. https://doi.org/10.1109/ACCESS.2024.3046991 (Journal Article)
18. Thakur, T., Batra, I., Malik, A., Ghimire, D., Kim, S.-H., & Sanwar Hosen, A. S. M. (2023). RNN-CNN based cancer prediction model for gene expression. IEEE Access, 11, 131024–131044. https://doi.org/10.1109/ACCESS.2023.3067384 (Journal Article)
19. Alhashmi, S. M., Al-Ghamdi, A. F., & Alshehri, M. A. (2024). Survival analysis of thyroid cancer patients using machine learning algorithms. IEEE Access, 12, 61978–61990. https://doi.org/10.1109/ACCESS.2024.3084261 (Journal Article)
20. S. S. Nigatu, P. C. R. Alla, R. N. Ravikumar, K. Mishra, G. Komala and G. R. Chami, "A Comparitive Study on Liver Disease Prediction using Supervised Learning Algorithms with Hyperparameter Tuning," 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, 2023, pp. 353-357, doi: 10.1109/InCACCT57535.2023.10141830.
21. R. R N, S. Jain and M. Sarkar, "SMOTE and Hyperparameter Optimization: A Dual Machine Learning Strategy for Enhancing Coupon Recommendation in Vehicular Contexts," 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 2023, pp. 1-6, doi: 10.1109/SMARTGENCON60755.2023.10442306.
22. M. Habib, P. Rajah, R. R. N, S. Nadarajan, S. S. Betgeri and M. Itagi, "Machine Learning Classifiers and Hyperparameter Tuning to Enhance Prediction of ASD: An Automated Django Application," 2024 2nd World Conference on Communication & Computing (WCONF), RAIPUR, India, 2024, pp. 1-7, doi: 10.1109/WCONF61366.2024.10692172.