International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol. 6 No. 38 September - October - 2025

Performance Evaluation Metrics for Machine Learning Classification Techniques In a Oncological Disease Prediction

Ms. Pratiksha S. Chavan
Department of Computer Science
Dr. D Y Patil Arts, Commerce and Science College Akurdi, Pune 44, INDIA
Corresponding Author –Ms. Pratiksha S. Chavan
DOI - 10.5281/zenodo.17315791

Abstract:

INDIA, Reported nearly 13, 24, 413 new cancer incidences in 2020. Prostate cancer, an oncological disease cases are also significantly increased with an increasing mortality rate. Accurate prediction tools are important in prostate cancer detection. Machine learning principles are best suited for developing modern applications in healthcare informatics and biomedical Engineering and sciences. The proposed paper discussed ML based classifier techniques for prediction of affected prostate cancer glands. Logistic regression decision tree, k nearest neighbor and naïve Bayes classifier are implemented for the prostate cancer patient data. Comparative analysis with the help of classification metrics is presented. The prediction accuracy using logistic regression is calculated as 91% which is better than the 73% for K nearest Neighbor classifier techniques is 73%.

Keywords: Machine learning, Image Segmentation, Oncology, PBCRs, HBCRs, IARS, Prostate cancer, Management

Introduction:

An introduction of machine learning based classification algorithms in oncological diseases aims at building efficient and accurate predictive modelling tools. These tools will not only help to predict the risk level of severity of cancers but also help in clinical and radiation oncology management [11] [14]. The predictive tools designed using ML techniques provide an optimal method in therapeutic and cancer patient management.

New incidences of male prostate cancer are significantly increasing across the globe. It ranks among the major cancer types found in men. The Population Based Cancer Registry (PBCRs) and Hospital Based Cancer Registry (HBCRs) which are aimed to keep records of all cancer incidences estimated that the prostate cancer incidences and deaths are increasing at a faster rate. The data published by International Agency on Cancer research (IARC) shows that 19292789 new cancer incidences of all types are reported in 2020 [1][3]. This data also shows that new 1414259 prostate cancer incidences reported which are more than reported in 2018 [2] [3].

India reported a significant rise in the number of cancer incidences in all regions of the country. A total of 13324413 new cancer incidences of all types of cancers are reported with 851678 deaths cases. Incidences of new prostate cancer patients also increase in major cities of India. There are 34540 new Prostate cancer patients reported in 2020. These patients are more than that of bladder 21096, Thyroid 20432, Gallbladder 19570 and 12642 Pancreas cancer patients [6]. Prostate cancer is a second leading cancer disease in Pune,

Delhi, Thiruvananthapuram and Kolkata and third in cities like Mumbai and Bangalore [5]. Prostate cancer ranks 12th in all cancer types found in India [6].

Fig. 1(a)

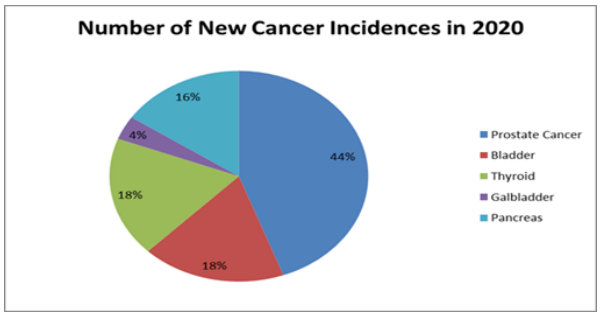

Fig. 1(b)

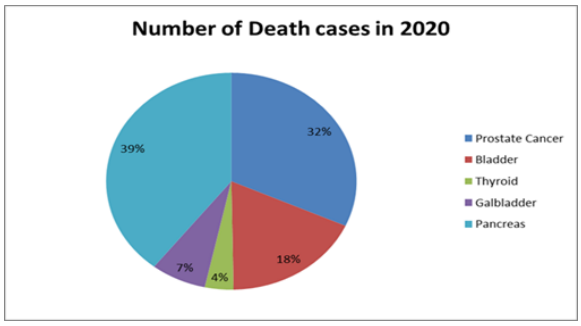

Fig.1 (a) (b) GLOBOCAN 2020 Statistic by WHO's IARC

Fig. 2 (a)

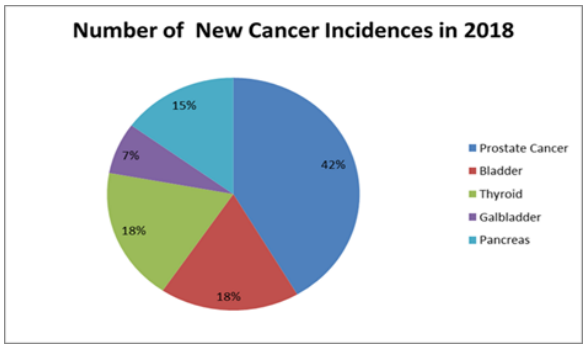

Fig. 2 (b)

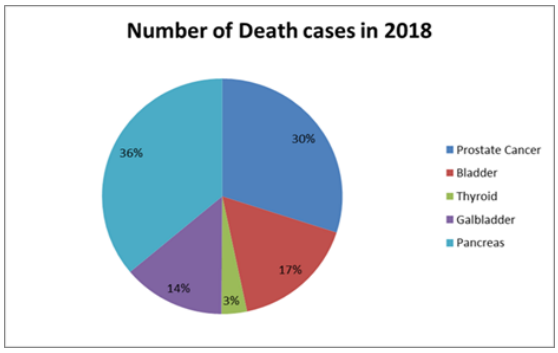

Fig.2(a) (b). GLOBOCAN 2018 Statistic by WHO's IARC

Prostate Gland Anatomy And Prostate Cancer:

The reproductive system of male has seminal vesicles, Bulbourethral and Prostate accessory glands. These glands are responsible for transportation of semen and production of structural protein and alkaline solution that are needed in the formation of spermatophores [7].

Prostate gland, a fibromuscular gland lies in the front of the rectum area and surrounding urethra [7] [8]. Prostate gland anatomy describes this gland in different lobes and three zones as shown in fig.3.

- Lateral lobe
- Median lobe
- Anterior lobe
- Posterior lobe
- Central zone
- Transition zone
- Peripheral zone

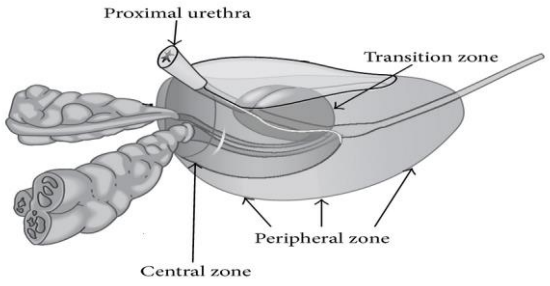

Fig.3. Section of Prostate Gland Anatomy zone Representation

Prostate gland size changes with an increase in men's age. The prostatitis condition in the prostate gland occurs by infection. When gland cell size changes abruptly, prostate cancer begins to form. This cancer spreads very slowly and starts causing symptoms for years and years.

Digital rectum examination [10], prostate specific antigen [9] Imaging test and prostate biopsy tests are used for possible cancer symptoms examination and diagnosis in male patients.

Traditional prostate cancer testing tools sometimes become less accurate and digital rectum analysis and biopsy may cause infection, bleeding, and pain due to external instrument use.

The new approach has been introduced in cancer diagnosis using advanced radiographic tests and machine learning techniques. With the incursion of artificial intelligence techniques in spine related surgery, a new approach has been adopted to create clinical analytics tools and diagnostic instruments. ML and AI technique makes clinical predictions and decision-making tasks easier [12] [13]. These techniques also play an important role in transforming medical imaging which incorporates medical radiology.

Newly designed diagnostic imaging systems are utilizing the capabilities of AI and ML techniques. Specially designed Algorithms help performance tuning for imaging systems [15-18] and facilitate the quantification and detection of various possible clinical scenarios.AI techniques not only show impressive performances in accurately identifying the cancer region in prostate gland in male reproductive system.

Machine Learning In Prostate Cancer Detection:

Researchers with the help of the latest machine learning techniques are developing advanced methods for detection and analyzing prostate cancer [9].

ML techniques can be best suited with imaging tests that are used for prostate cancer diagnosis. Radiography such as CT and MRI are the x-ray imaging testing methods that easily defines all prostate gland anatomical details [11]. This method provides details about lesion location, structural tissue changes in gland and size of the prostate gland.

A. The logistic regression:

Logistic Regression, a ML based predictive analytic algorithm which uses probability statistics. Binary and multi linear logistic regression are the two types of the regression used for data classification.

Conversion of data samples into its statistical probabilities, sigmoid functions are used in logistic regression given by

$$f(x) = \frac{1}{1+e^{-x}} \qquad \text{------ (1)}$$

Following steps are required to do logistic regression in classification process

- Data preprocessing
- Logistic regression function fitting to training dataset
- Test result prediction
- Test result visualization

B. K Nearest Neighbor Classifier

- Selection of no. of neighbor
- Euclidian distance calculation between no. of neighbors
- Counting no. of data points for every data
- Assign new data points

Proposed Work:

An aim of the proposed paper is to implement the four ML based classifiers for

prediction of the prostate cancer type from the given prostate cancer patient data. The cancer data is a CSV based file that has information about prostate cancer results that are either malignant or benign and geometric information about cancer prostate gland.

Logistic regression and K- nearest Neighbor (KNN) classifiers are implemented to predict the results based on the information provided in the given data.

Following essential steps performed for the prediction process using logistic regression and KNN Classifier techniques

- Load the dataset
- Feature extraction from the given dataset
- Data collection for training and testing purpose
- Train the model on the training data
- Prediction on the basis of testing data

Observations:

The values of classification metrics are calculated by implementing the proposed classifiers for the prostate cancer prediction. The prostate cancer result is either Benign or Malignant. The benign prostate cancer is represented by 0 and malignant by 0. The values of the metrics for each classifier are given in respective tables.

Table 1: Classification Metrics for Logistic Regression

Accuracy	Precision		Recall		F1-Score	
91%	+Ve	-Ve	Sensitivity	Specificity	+Ve	-Ve
	0.91	0.90	0.95	0.82	0.93	0.86

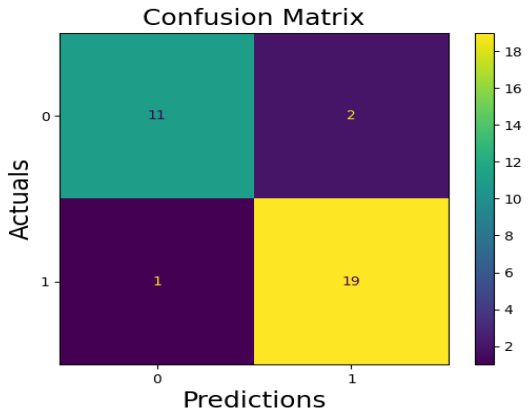

Fig 6 Confusion Matrix for Logistic Regression

TABLE 3 Classification Metrics for KNN Classifier

Accuracy	Precision		Recall		F1-Score	
73%	+Ve	-Ve	Sensitivity	Specificity	+Ve	-Ve
	0.84	0.57	0.73	0.73	0.78	0.64

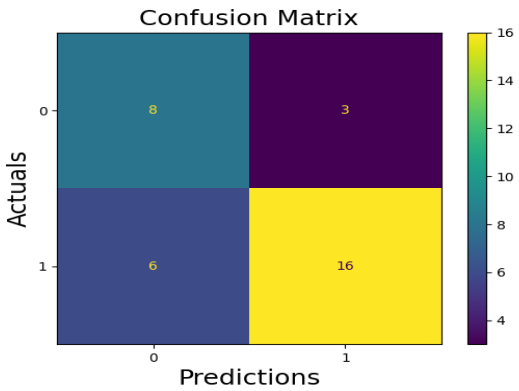

Fig.6. Confusion Matrix for KNN Classifier

Conclusion:

The comparative analysis between logistic regression and KNN based ML classifiers is presented. The numerical data related to prostate cancer gland is used for this comparative analysis. Logistic regression classifier (LR) is accurate to 91 % in predicting the nature of prostate cancer either malignant or benign. The accuracy for KNN is only 73% . The ability to capture the all positive samples called as True Positive Rate (TPR) or sensitivity for given classifiers Logistic regression and KNN are 94% and 83% respectively. Thus the prostate cancer prediction can be efficiently performed by using logistic regression technique.

Acknowledgment:

I would like to express my gratitude to Dr. Mohan Waman Principal, Computer Science, Dr. D Y Patil ACS College, Pune INDIA for their valuable guidance for providing Prostate Medical Image Database.

References:

[1] Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal; "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries"; CA: A Cancer Journal for Clinicians, Volume 71, Issue 3, 04 February 2021

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, A. Jemal; "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries"; CA: A Cancer Journal for Clinicians, Volume 70, Issue 4, 06 April 2020

[3] International Agency for Research on Cancer and Global Cancer Observatory; https://gco.iarc.fr/

[4] Krishnamoorthy Hariharan, Venugopal Padmanabha; "Demographic and Disease characteristics of prostate cancer in India"; Indian Journal of urology, Volume 32, Issue 2, Pages: 103-108, 2016

[5] Prameesha Pereraa, Ayam Guptac, Nidhi Shuklab, Mamta Nehrac, Mukesh Sharmad, Sneha Mishrad ,Vikram Singh Chauhand, et al; " Revisiting Prostate Cancer in India"; Preprints , 26 November 2018

[6] Prashant Mathur, Krishnan Sathishkumar, Meesha Chaturvedi, Priyanka Das, et al; "Cancer Statistics, 2020: Report From National Cancer Registry Programme, India"; Global Oncology, American Society of Clinical Oncology, 16 July 2020

[7] Omesh Singh, Srinivasa Rao Bolla; "Anatomy, Abdomen and Pelvis, Prostate"; StatPearls Publishing, January 2021

[8] Anil Bhavsar, Sadhna Verma; "Multiparametric MRI in Prostate Cancer"; BioMed Research International, vol. 2014, Article ID 728539, 9 pages, 2014

[9] Michael Obana, Henry O'Lawrence; "Prostate Cancer Screening: PSA Test awareness among adult males"; J Health Hum Serv Adm. 2015

[10] Leen Naji, Harkanwal Randhawa, Zahra Sohani; "Digital Rectal Examination for Prostate Cancer Screening in Primary Care: A Systematic Review and Meta-Analysis"; Analysis of Family Medicine, 149-154, March

[11] Mohammed AliKazem; "Predictive models in cancer management: A guide for clinicians"; Elsevier, The Surgeon, Volume 15, Issue 2, Pages 93-97, April 2017

[12] Vincenzo Valentini, Nicola Dinapoli, Andrea Damiani; "The future of predictive models in radiation oncology: from extensive data mining to reliable modeling of the results"; Future Oncology, March 2013

[13] Derek J Van Booven, Manish Kuchakulla, Raghav Pai, Fabio S Frech, Reshna Ramasahayam, Pritika Reddy, Madhumita Parmar, Ranjith Ramasamy, Himanshu Arora1; "A Systematic Review of Artificial Intelligence in Prostate Cancer"; Dovepress Journal, Research and Reports in Urology, 2021

[14] Ronan Thenault, Kevin Kaulanjan,Thomas Darde, Nathalie Rioux Leclercq, Karim Bensalah, Marie Mermier, Zine-eddine Khene, Benoit Peyronnet, Shahrokh Shariat, Benjamin Pradere , Romain Mathieu; "The Application of Artificial Intelligence in Prostate Cancer Management—What Improvements Can Be Expected? A Systematic Review"; Mdpi, Applied Science, 2020

[15] Lidia Alcala Mata , Juan Antonio Retamero, Rajan T. Gupta, Roberto García Figueras, Antonio Luna; "Artificial Intelligence–assisted Prostate Cancer Diagnosis: Radiologic-Pathologic Correlation"; RadioGraphics, Vo1ume 41, Number 6, 1 October 2021