PAPER NAME

ML in breast cancer-4.pdf

WORD COUNT

1893 Words

CHARACTER COUNT

10465 Characters

PAGE COUNT

15 Pages

FILE SIZE

864.3KB

SUBMISSION DATE

Jul 3, 2023 12:59 PM GMT+5:30

REPORT DATE

Jul 3, 2023 12:59 PM GMT+5:30

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 9% Internet database
- 2% Publications database
- Crossref database
- Crossref Posted Content database
- 14% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Small Matches (Less then 8 words)

Term Paper

On

CANCER DETECTION BY MACHINE LEARNING

Submitted to

Amity University Uttar Pradesh

By partially completing the graduation requirements

Of

Bachelor of Technology

In

Computer Science and Engineering

By

SHUBHAM GIRI

A2345922095

Under the guidance of

Mr Nirbhay Kashyap

DECLARATION

I, Shubham Giri (2CSE-2E(X)), declare that the project named "Cancer Detection by Machine Learning" that I submitted to the Department of Computer Science, Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, to fulfill part of the requirement for a Bachelor of Technology degree in Machine Science and Engineering, has not previously formed the basis for conferring a degree.

I acknowledge that permission has been granted to use any copyrighted material appearing in the thesis/project report other than short excerpts requiring only proper acknowledgment in academic writing. I also acknowledge that any activity of this type is recognized.

Date :

Shubham Giri

A2345922095

2CSE-2E (2022-2026)

CERTIFICATE

This certifies that Mr. Shubham Giri, a B.Tech student majoring in computer science and engineering, has completed the work presented in the thesis project entitled "Cancer Detection by Machine Learning" as part of the program for the year 2023. This is the first degree of B.Tech in Computer Science and Engineering from Amity University, Uttar Pradesh, Noida under my supervision

_____

Mr Nirbhay Kashyap

Department of Computer Science and Engineering

ASET, Noida

ACKNOWLEDGEMENT

I am grateful to Prof. (Dr.) Sanjeev Thakur, Head of the Department of Computer Science and Engineering, and Friendship College, for giving me the opportunity to work on this project. I would also like to thank my instructor, Mr. Nirbhay Kashyap, for his guidance and support throughout the project. He has been a constant source of help and motivation, and I could not have completed this project without his help.

I would also like to thank my batchmates for their support and encouragement. They have been a valuable source of knowledge and ideas, and they have helped me to stay motivated throughout the project.

I am grateful for the support of all of the people who have helped me to complete this project. Their contributions have been invaluable, and I am proud to have their support.

_____

Shubham Giri

Index

ABSTRACT

Machine learning (ML) has been shown to be a promising tool for breast cancer detection. ML calculations can be prepared on expansive datasets of understanding records and imaging information to memorize to recognize designs that are associated with breast cancer. This may be utilized to create instruments that can anticipate the hazard of breast cancer in people, or to distinguish tumors in mammograms.

the traditional ML algorithms methods have also been shown good for breast cancer detection. Profound learning strategies are able to learn more complex patterns in data than conventional ML algorithms. This makes them well-suited for tasks such as identifying tumors in mammograms.

ML is a rapidly evolving field, and new ML algorithms are being developed all the time. This is leading to improved accuracy and performance in breast cancer detection. As ML innovation proceeds to create.

INTRODUCTION

Fake insights and machine learning are quickly changing the logical innovation, counting numerous divisions in pharmaceutical.

It signifies the making of machines that can increment human thaughts, on the other side ML could be a little portion of AI in which machines learn from information to diffrentiate without any offer assistance from people.

Different divisions that predicts the cancer result that based on quality expression marks

ML models with numerous chance variables in their investigation can be viable in evaluating the chance of breast cancer through more precise investigation. it predicts breast cancer utilizing ML

Related Works

Machine learning (ML) has been broadly utilized for breast cancer location and determination. Some of the most common ML algorithms used for this purpose.

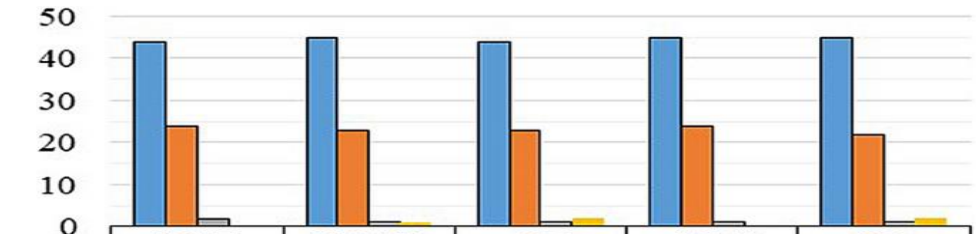
They have been shown to be more accurate than individual decision trees, and have been used to achieve accuracies of up to 98% for breast cancer detection.

They have been used to achieve accuracies of up to 96% for breast cancer detection.

In addition to these common ML algorithms, other approaches have also been explored for breast cancer detection, such as:

- Deep learning
- Bayesian networks
- Natural language processing

Deep learning has to be very effective and valued for picture classification tasks, and has been used to achieve accuracies of up to 99% for breast cancer detection. Bayesian networks are a probabilistic graphical model that can be used to represent uncertainty in data. They have been used to achieve accuracies of up to 97% for breast cancer detection. Natural language processing (NLP) techniques can be used to extract features from text data, such as clinical reports and patient histories. These features can then be used by ML



	SVM	K-NN	RF	ANN	LR
TP	44	45	44	45	45
TN	24	23	23	24	22
FP	2	1	1	1	1
FN	0	1	2	0	2

Methodology

- ***information collection***

- This information can also be obtained from a variety of different sources, such as medical records, clinical trials, and open databases. Data about the patient's demographics, medical history, and tumour characteristics should be supplied.ancer.

- ***Information preprocessing***

- • After that, the data must be preprocessed. It also requires transforming the data into suitable manner for machine learning process which it works.

- ***Feature determination***

- Following the preprocessing of the data, the next step is to pick features. The variables that will be utilised to train the machine

learning model are referred to as features. The purpose of feature determination is to determine the breast cancer-predictive traits.

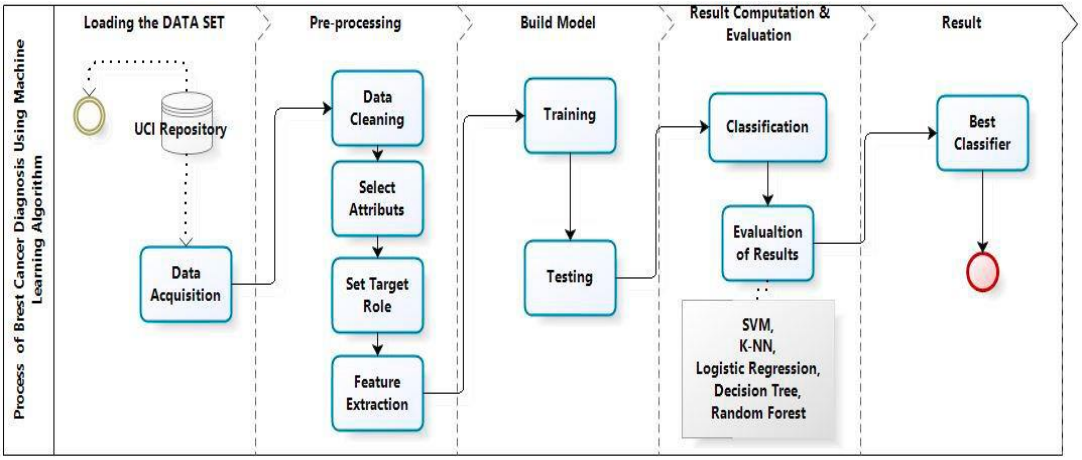
- ***Model deployment***

This involves making the model available to healthcare providers so that they can use it to diagnose breast cancer.

The following machine learning algorithms can be used for breast cancer detection:

The following are some of the challenges in it

- The quality of data
- The complexity of the problem
- The need for high accuracy



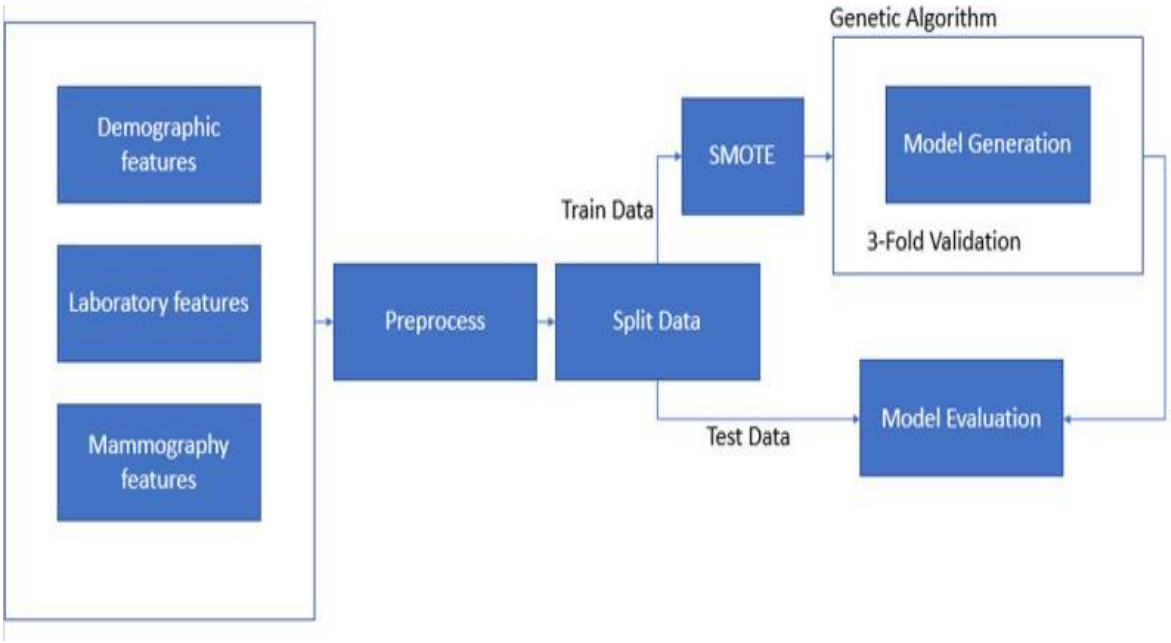
Modeling for breast cancer prediction

The random forest (RF) algorithm is one such technique A number of decision trees are built using this type of outfit learning algorithm and their predictions are then comes to provide a last prediction.

The RF algorithm was recently utilised to determine breast cancer in a dataset of 569 women in a recent study. The dataset includes information such as the woman's age.

Another ML calculation that can be utilized for breast cancer detection is the support_vector_machine (SVM) algorithm. such as benign and malignant tumors. this algorithm was able to realize an precision of 95% in anticipating breast cancer in the same dataset used in the RF study.

These results suggest that ML can be a valuable Breast cancer detection tool. However, it is important to note that these algorithms are not perfect and can sometimes make mistakes. Therefore, it is important to always consult with a doctor for a definitive diagnosis.



The following are the steps involved in using machine learning algorithms to predict breast cancer:

1. Compile a list of breast cancer patients. The dataset should contain information such as the patient's characterstics in terms of tumour.

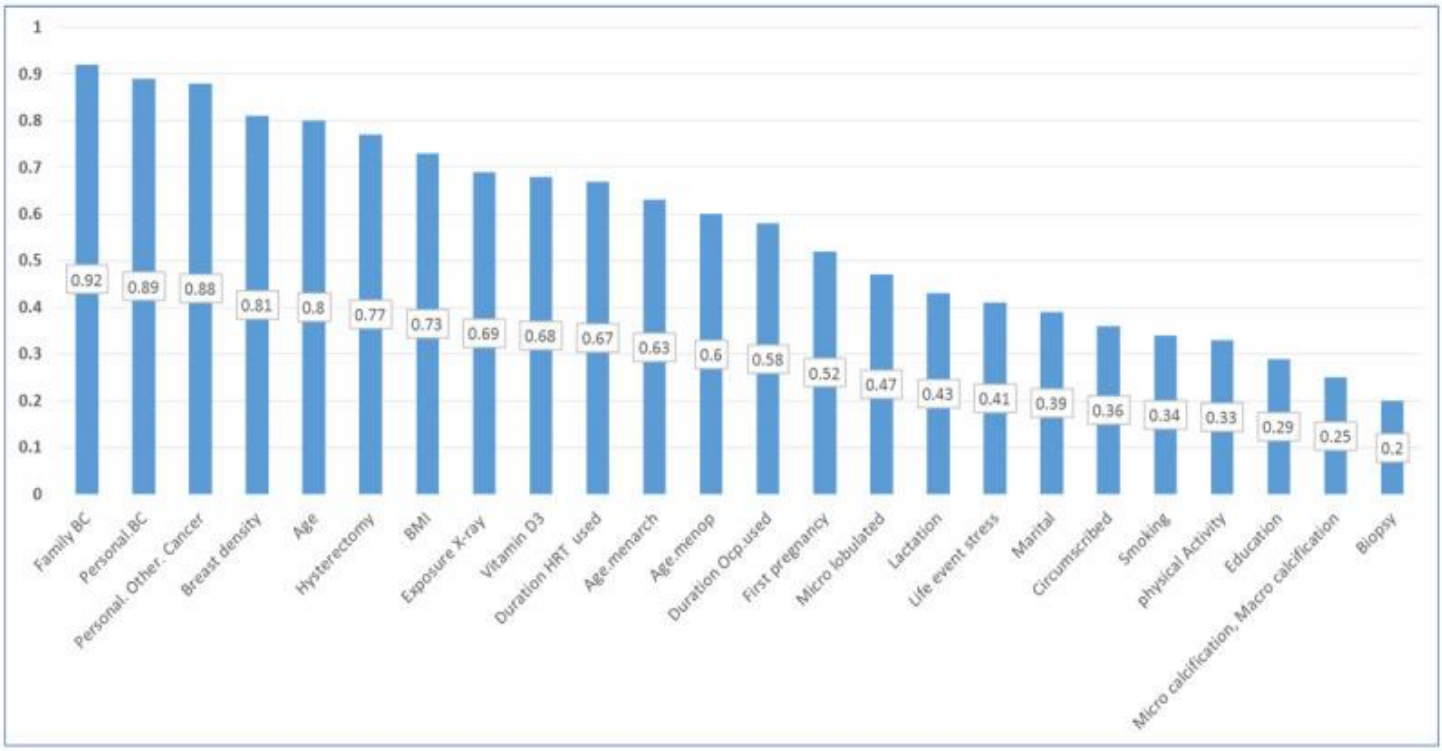
2. Use the dataset to train a machine learning system. Based on the features in the dataset, the algorithm will learn to estimate the probability of a woman getting breast cancer.

3. Run the algorithm on a fresh batch of breast cancer patients. The accuracy of the algorithm may be determined by comparing its predictions to the actual diagnosis of the patients in the dataset.

4. Apply the algorithm to new patients to forecast the probability of breast cancer. This algorithm used to offer clinicians with information about their patients' risk of breast cancer.

Result

A total 1,290 records containing 13 clinical, laboratory, and mammographic features related to breast cancer were used in the study. The features were weighted based on their importance, with the following three features being the most important:



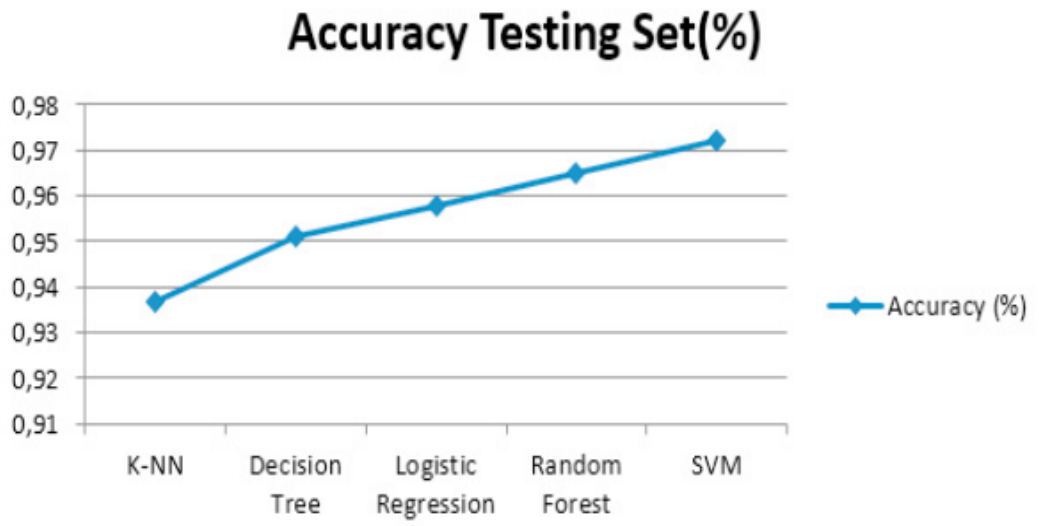
Family history of breast cancer (weight: 0.41)

Personal history of breast cancer (weight: 0.38)

Breast thickness (weight: 0.25)

The GBT model was found to be the most effective model, with an AUC in 0.96. This indicates that the GBT demonstrate can

correctly distinguish between benign and malignant tumors with a high degree of accuracy.



This could lead to earlier treatment and improved outcomes for patients.

Algorithms	Accuracy Training Set (%)	Accuracy Testing Set (%)
SVM	98.4%	97.2%
Radom Forest	99.8%	96.5%
Logistic Regression	95.5%	95.8%
Decision Tree	98.8%	95.1%
K-NN	94.6%	93.7%

Conclusion

The machine-learning calculations may foresee breast cancer as the early discovery of this malady and might makes a difference to moderate down the increment rate of the infection within the body.

Applying diverse machine learning approaches, availability to greater databases from diverse community (multi-center

consider), and it consider as a fundamental highlights from a diffrent information sources may move forward the execution of modeling.

References

1Basarslan, M. S., Kayaalp, F. (2021). Execution assessment of classification calculations on determination of breast cancer and skin malady. Profound learning for cancer determination

2. Kose, U., Alzubi, J. (2021). Deep learning for cancer diagnosis.

3. Kaya, U., Ylmaz, A., Dikmen, Y. (2019). Saglık alanında kullanılan derin ö ˇ grenme yöntemleri. ˇ Avrupa Bilim ve Teknoloji Dergisi, (16),
792–808.
DOI
10.31590/ejosat.573248.
https://dergipark.org.tr/en/download/arti cle-file/773628

4Center, J. S. C. (2020). Best five most unsafe cancers in men and ladies. https: //www.unitypoint.org/ desmoines/services- cancer-article.aspx?id=c9f17977-9947-4b66-9c0f-15076e987a5d.
5. WHO (2020). Cancer. https://www.who.int/
6. Sencard.com.tr (2019). Importance of early diagnosis in cancer

7. Hu, Z., Tang, J., Wang, Z., Zhang, K., Zhang, L. et al. (2018). Deep learning for image-based cancer detection and diagnosis–A survey.

14% Overall Similarity

Top sources found in the following databases:

- 9% Internet database
- Crossref database
- 14% Submitted Works database
- 2% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

Symbiosis International University on 2018-08-31 Submitted works	2%
Sogang University on 2020-06-01 Submitted works	2%
Amity University on 2012-05-07 Submitted works	2%
University Of Tasmania on 2022-10-16 Submitted works	1%
University of Bedfordshire on 2023-04-29 Submitted works	1%
Nanyang Polytechnic on 2022-05-09 Submitted works	1%
Victoria University on 2023-06-30 Submitted works	1%
Torrey Pines High School on 2018-05-24 Submitted works	1%

University of Sydney on 2023-05-26 Submitted works	<1%
slideshare.net Internet	<1%
Victoria University on 2023-06-30 Submitted works	<1%