K14

by Trushna Parida

Submission date: 28-Mar-2023 05:16PM (UTC+0530)
Submission ID: 2048964037
File name: K14_Extended_abstract.pdf (93.65K)
Word count: 436
Character count: 2525

Breast Cancer Detection using Machine Learning

Priyambada Sahoo[1], Madhusmita Jena[2], Swayam Sarthak Rout[3] ,Aditya Ku. Samal[4],

Dr.Trushna Parida[5]

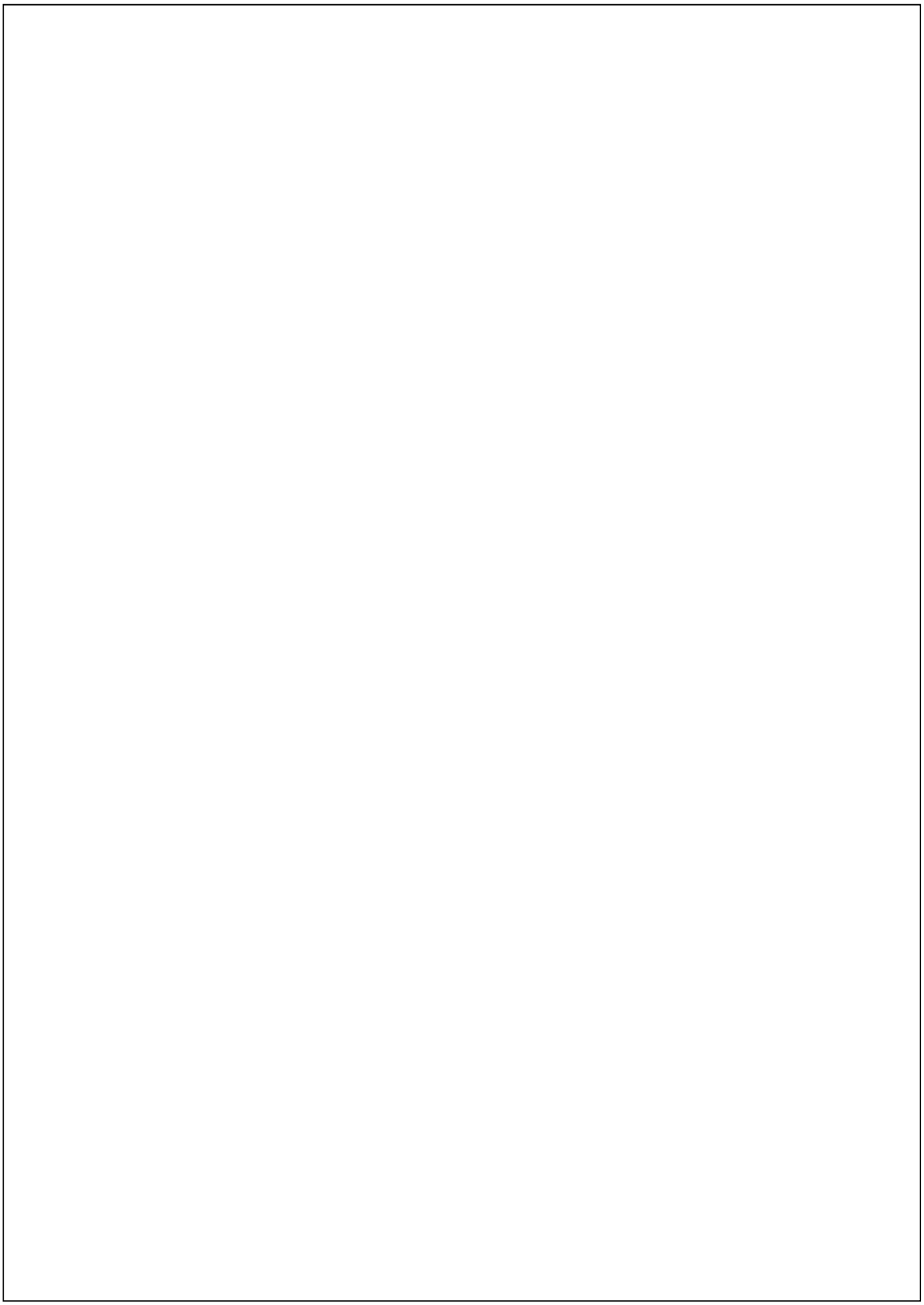
Institute of Technical Education and Research, Siksha 'O' Anusandhan (Deemed to be university)

Bhubaneswar, 751030, India

Abstract

Breast cancer is a severe disease that affects millions of people globally, and early detection is critical for successful therapy. The goal of this study is to develop a machine learning model to accurately predict breast cancer using the publicly available PIMA Indians dataset. The dataset includes mammography pictures, demographic information, family history, and lifestyle factors, and the study will use advanced machine learning algorithms such as Naive Bayes, Decision Trees, and Support Vector Machines (SVM) to analyze the dataset. The ML model classifies patients into high-risk groups, prioritizes additional screening, and recommends tailored treatment. SVM is the most accurate algorithm with 94.28% accuracy for detecting breast cancer. To further increase the accuracy of the results, a larger dataset can be used to train the model. The system will provide a trustworthy and useful tool for both patients and medical professionals. Patients can benefit from early detection of breast cancer as it can save them from the effort of visiting a medical center, consulting a doctor, and from various complications that occur if breast cancer remains untreated. Medical professionals can use this model to identify patients who are at high risk of developing breast cancer and provide personalized treatment recommendations, resulting in a better chance of successful therapy and reduced necessity for more intrusive operations. In conclusion, the development of a machine learning model for breast cancer prediction is vital for the early detection and successful therapy of the disease. The use of cutting-edge algorithms like Naive Bayes, Decision Trees, and SVM can provide accurate results and help to prioritize patients for further screening and tailored treatment recommendation. In the initial phase of our study, We collected and cleaned a cancer patient dataset, then divided it into a training and testing dataset. After training the algorithm on the training dataset, we used the output to predict on the testing dataset and compared different algorithm performances to obtain the final output.This project has limitations that need addressing, including the need for large, diverse datasets, consideration of potential biases in the training data ,and the influence of mammography image quality on model accuracy due to various factors like patient movement or imaging equipment. This study is an essential step towards developing a reliable tool for breast cancer prediction that can benefit both patients and medical professionals.

Keywords: K-Nearest Neighbors(KNN), Support vector machines(CNN), Random forest, Logistic regression, Mammography images



K14

ORIGINALITY REPORT

19%	11%	9%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.utkalbiotech.org Internet Source	4%
2	Submitted to The British College Student Paper	3%
3	Submitted to Valley Stream Central High School District Student Paper	3%
4	link.springer.com Internet Source	3%
5	Submitted to Midlands State University Student Paper	3%
6	www.pedrobalage.com Internet Source	2%
7	Xiangyi Kong, Qinheng Yang, Liquan Ouyang, Qiang Liu, Hao Dong, Zhongzhao Wang, Yi Fang, Jing Wang, Yinpeng Ren. "Incidence, Risk Factors, Prevention and Treatment of Postmastectomy Pain Syndrome in Breast Cancer: a Multicenter Study", Research Square Platform LLC, 2022	1%

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		