|  |  |  |
| --- | --- | --- |
|  | **Department of Computer Science and Engineering**  Bangladesh University of Business and Technology (BUBT) | BUBT |

**CSE 498: Literature Review Records**

|  |  |
| --- | --- |
| **Student’s Id and Name** | **Name:** Sm Raziur Rahman Pushon, **ID:** 19202103214 |
| **Project Title** | Deep Learning in Healthcare: Breast Cancer Detection and Classification using Image Processing and CNN |
| **Supervisor Name & Designation** | **Name:** Khan Md. Hasib, **Designation:** Assistant Professor, Department of CSE, BUBT |
| **Course Teacher’s Name & Designation** | **Name:** Khan Md. Hasib, **Designation: :** Assistant Professor, Department of CSE, BUBT |

|  |  |
| --- | --- |
| **Aspects** | **Paper # 02 (Title)** |
| **Title / Question**  (What is problem statement?) | Machine Learning Algorithms For Breast Cancer Prediction And Diagnosis. |
| **Objectives / Goal**  (What is looking for?) | The purpose of this study is to examine and evaluate the efficacy of several machine learning algorithms for the detection and prognosis of breast cancer. The research specifically seeks to discover the best method in terms of precision, accuracy, and other performance measures. It concentrates on five machine learning algorithms: K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, Decision Tree (C4.5), and Support Vector Machine (SVM). The ultimate goal is to support the creation of precise and trustworthy methods for early detection and prediction of breast cancer, as these processes have a substantial bearing on patient treatment and survival. |
| **Methodology/Theory**  (How to find the solution?) | Using machine learning methods, the paper seeks to detect and diagnose breast cancer. The process involves obtaining the Breast Cancer Wisconsin Diagnostic dataset, preprocessing the data, choosing five machine learning algorithms (SVM, Random Forest, Logistic Regression, Decision tree, and KNN), building models, evaluating them using various metrics, comparing the results, and choosing the most effective algorithm based on accuracy and precision. Implementation and assessment were done by the authors using Python and the Scikit-learn module. |
| **Software Tools**  (What program/software is used for design, coding and simulation?) | Python programming is used to carry out the work outlined in the paper under the Anaconda environment. The machine learning algorithms are created and assessed using the Python Scikit-learn module. |
| **Test / Experiment**  How to test and characterize the design/prototype? | Performance measures including Confusion Matrix, Accuracy, Precision, Sensitivity, F1 Score, and Area Under the Receiver Operating Characteristic Curve (AUC) are used to assess and characterize machine learning models. The optimal algorithm for breast cancer prediction and diagnosis is determined using these measures, which are also used to evaluate and compare the models. The models' performance is evaluated using a separate testing dataset (25% of the data) after they have been tested on a training dataset (75% of the data). Each machine learning algorithm's efficiency and accuracy are evaluated using the outcomes of these measures. |
| **Simulation/Test Data**  (What parameters are determined?) | The performance of multiple machine learning algorithms for breast cancer prediction and diagnosis is evaluated in this research using a variety of parameters and metrics. The confusion matrix, accuracy, precision, sensitivity, F1 score, and area under the receiver operating characteristic curve (AUC) are some of these characteristics. Five machine learning algorithms' performances are compared using these metrics in order to determine which is the most accurate and precise. |
| **Result / Conclusion**  (What was the final result?) | According to the paper's conclusion, of the five machine learning algorithms analyzed, the Support Vector Machine (SVM) had the highest efficiency, precision, and accuracy. With regard to the Wisconsin Breast Cancer Diagnostic dataset, SVM specifically achieved an accuracy of 97.2%, precision of 97.5%, and an AUC of 96.6%. In the context of this dataset, it performed better than the other algorithms, making it the most useful for breast cancer detection and prediction. |
| **Obstacles/Challenges**  (List the methodological obstacles if authors mentioned in the article) | Methodological barriers aren't mentioned in the article specifically.. |
| **Terminology**  (List the common basic words frequently used in this research field) | Breast Cancer; Prediction; Diagnostic , SVM; Random Forest; Logistic regression ; C4.5; k-NN; Accuracy; Precision. |
| **Review Judgment**  (Briefly compare the objectives and results of all the articles you reviewed) | Both publications focus on the use of machine learning algorithms for the identification and diagnosis of breast cancer. In the first article, three algorithms are compared, and SVM performs the best; in the second, five methods are evaluated, and SVM is found to have the highest accuracy. |
| **Review Outcome**  (Make a decision how to use/refer the obtained knowledge to prepare a separate and new methodology for your own research project) | With a systematic approach, I can build a new method for my research project using the knowledge I learned from the publications. Prior to anything else, it's critical to understand the articles' goals, procedures, findings, and conclusions. The application of the information learned to the objective of my project is then analyzed. I must apply the information by taking project requirements and variable considerations into account. |