WEEKLY REPORT

ON

**BREAST CANCER DIAGNOSIS USING MACHINE LEARNING AND DEEP LEARNING ALGORITHM**

#### *Submitted to*

**NMAM INSTITUTE OF TECHNOLOGY, NITTE**

(An Autonomous Institution under VTU, Belagavi)

*In partial fulfillment of the requirements for the award of the*

Degree of Bachelor of Engineering

in

Computer Science and Engineering

#### *by*

**Elton Pereira 4NM20CS068**

**Pranav P Shetty 4NM20CS133**

**Ryan D’souza 4NM20CS148**

**Shashank KT 4NM20CS164**

Under the guidance of

**Ms.Jayapadmini Kanchan**

**Assistant Professor**

Dept. of CSE, NMAMIT, NITTE



WEEKLY REPORT

|  |  |
| --- | --- |
| **DATE** | **ACTIVITES** |
| WEEK 1: 1/09/2023  to  7/09/2023 | Researched about the project and decided on the title of the project. Advancing Breast Cancer Diagnosis: Leveraging Machine Learning and Deep Learning Algorithms for Accurate and Timely Detection. To Enhance Precision and Promptness in Detection for Improved Patient Outcomes and Treatment Strategies. |
| WEEK 2: 8/09/2023  to  14/09/2023 | Created the synopsis of the project. Through the fusion of machine learning and deep learning techniques, this project aims to revolutionize breast cancer diagnosis. By optimizing algorithms for precision and timeliness, the goal is to elevate patient outcomes. This research seeks to refine treatment strategies by facilitating early and accurate detection, ultimately enhancing the overall effectiveness of combating breast cancer. |
| WEEK 3: 15/09/2023  to  21/09/2023 | Reviewed literature paper 1  Challenges defining mammogram features hinder accuracy. Deep learning aids crucial feature identification, hampered by limited datasets. MRI-CT fusion offers promise. Pre-trained CNNs, notably InceptionV3, excel in breast cancer detection.  (Author: Zeyad Q. Habeeb, Branislav Vuksanovic, and Imad Q. Al-Zaydi)  Reviewed literature paper 2  Various methods enhance breast cancer detection: data mining aids statistical analysis from cancer logs, promising for future research. Deep learning, ensembles, and CNNs ensure precise diagnosis and risk prediction advancement.  (Author: V. Durga Prasad Jasti,1Abu Sarwar Zamani,2K. Arumugam,3Mohd Naved) |
| WEEK 4: 22/09/2023  to  28/09/2023 | Reviewed literature paper 3  Detecting macrocalcification in dense breast tissue is challenging due to white pixel similarity. Machine learning aids cancer diagnosis, enhancing traditional methods, especially in pattern recognition.  (Author: Prannoy Giri and K. Saravanakumar )  Reviewed literature paper 4  The study emphasizes ML, particularly deep learning, automating with BreaKHis. Techniques like transfer learning, CNNs analyzed. Bibliometrics highlight global trends, stressing automated breast cancer diagnosis' importance.  (Author: Shubhangi A. Joshi, Anupkumar M. Bongale Dr, Arunkumar M. Bongale Dr) |
| WEEK 5: 29/09/2023  to  5/10/2023 | Reviewed literature paper 5  Research explores mammography techniques: enhancement (adaptive histogram equalization, difference pics, region growth), segmentation, aiding lesion detection. SVM, perceptron, Bayesian classifiers ensure accurate breast cancer diagnosis.  (Author: **Sushovan Chaudhury**,1Alla Naveen Krishna,2Suneet Gupta,3K. Sakthidasan Sankaran,4Samiullah Khan,5Kartik Sau,6Abhishek Raghuvanshi,7and **F. Sammy)** |
| WEEK 6: 6/09/2023  to  12/10/2023 | Reviewed literature paper 6  An SVM, K-NN, RF-based automated mammogram classifier; RF excels, stressing pre-processing importance. Future goal: fully autonomous system aiding radiologists in interpretation.  (Author: V Harvind Viswanath, Lorena Guachi-Guachi and Saravana Prakash Thirumuruganandham) |
| WEEK 7: 13/10/2023  to  19/10/2023 | Reviewed literature paper 7  Breast cancer detection explored via data mining, image processing, ML. Algo comparison for benign/malignant classification, emphasizing error estimation, model selection for early prevention.  (Author: Neela A G, S Gayathri, Jayashree K) |
| WEEK 8: 20/10/2023  to  26/10/2023 | Reviewed literature paper 8  An ensemble voting ML approach achieves 98.50% precision in breast cancer analysis, using only 16 features.  (Author: Sri Hari Nallamala, Pragnyaban Mishra, Suvarna Vani Koneru) |
| WEEK 9: 27/10/2023  to  3/11/2023 | Reviewed literature paper 9  Study explores breast cancer prediction: ML (SVM, Random Forest), DL (CNN, ANN) on Wisconsin dataset. SVM, RF reach 96.5%, CNN, ANN surpass with 97.3%, 99.3%, emphasizing DL's superior efficacy.  (Author: Monika Tiwari, Rashi Bharuka, Praditi Shah, Reena Lokare) |
| WEEK 10: 4/10/2023  to  10/11/2023 | Reviewed literature paper 10  Study assesses ML in breast cancer detection from digitized fine needle aspirate images. Deep learning promising but hindered by limited data, conversion challenges for CNN. Recommendations: future enhancements with expanded data.  (Auhor: Pradeeba.R, Mrs.C.Clement Sherlin) |
| WEEK 11: 11/11/2023  to  15/11/2023 | Implementation and producing report and ppt for phase one of project.  The initial phase of the project involves the practical implementation of machine learning and deep learning algorithms for breast cancer diagnosis. The generation of a comprehensive report and a visually engaging PowerPoint presentation outlining the methodologies, findings, and initial insights obtained, setting the groundwork for subsequent research phases. |
| WEEK 12: 20/11/2023  to  c 27/11/2023 | In the modified dataset, two attributes, marital status, and race, have been removed. This alteration aims to simplify the dataset by eliminating these features, potentially reducing complexity and focusing the analysis on other relevant factors. |
| WEEK 13: 1/1/2024  to  8/1/2024 | Following mentor guidance, logistic regression was excluded from model comparison. Although commonly used, its removal suggests a focus on alternative models, possibly to explore different algorithmic approaches or prioritize interpretability over logistic regression's linearity assumption. |
| WEEK 14: 9/1/2024  to  15/1/2024 | SVM (Support Vector Machine) was introduced as a machine learning algorithm in the model comparison. SVM is renowned for its efficacy in classification tasks, especially in scenarios with complex decision boundaries. Its addition broadens the analysis, leveraging SVM's strengths in handling high-dimensional data and non-linear relationships. |
| WEEK 15: 16/1/2024  to  22/1/2024 | KNN (K-Nearest Neighbors) was incorporated as a machine learning algorithm. KNN is valuable for its simplicity and effectiveness in both classification and regression tasks, particularly in scenarios where data points exhibit local relationships or clusters. Its inclusion enriches the analysis, offering a versatile approach to pattern recognition. |
| WEEK 16: 23/1/2024  to  29/1/2024 | After evaluating multiple machine learning models, their performances were compared. The model with the highest accuracy was identified, indicating its superior predictive capability among the tested algorithms, thus providing valuable insight into the most effective approach for the task at hand. |
| WEEK 17: 30/1/2024  to  5/2/2024 | Incorporating a metric system alongside accuracy provided a standardized measure for evaluating model performance. Additionally, representing accuracy through graphical visualization enhanced comprehension, enabling stakeholders to intuitively grasp and compare the effectiveness of different machine learning models. |
| WEEK 18: 13/2/2024  to  18/2/2024 | Phase two of the project entails executing machine learning and deep learning algorithms for breast cancer diagnosis. A detailed report and visually appealing PowerPoint presentation will be created, summarizing methodologies, findings, and preliminary insights, laying the foundation for future research phases. |

|  |  |
| --- | --- |
| WEEK 19: 23/2/2024  to  29/2/2024 | The mentor recommended refining use case relationships, such as adding or removing cases, and optimizing message exchanges in sequence diagrams to ensure clarity and alignment with system requirements, facilitating better understanding of system interactions and functionality.  **Top of Form**  **Bottom of Form** |
| WEEK 20: 1/3/2024  to  6/3/2024 | Eliminating SVM algorithm from breast cancer diagnosis models aims to streamline computational complexity, focusing on deep learning architectures for improved accuracy and feature extraction, potentially enhancing diagnostic performance and facilitating more efficient model training and deployment processes. |
| WEEK 21: 7/3/2024  to  13/3/2024 | Testing the breast cancer diagnosis model with KNN, ANN, and Random Forest algorithms ensures comprehensive evaluation of predictive performance, leveraging diverse methodologies for robust assessment of accuracy, sensitivity, and specificity, enhancing confidence in model reliability and clinical applicability and for proper accuracy. |
| WEEK 22: 14/3/2024  to  20/3/2024 | Optimizing the training and testing split in breast cancer diagnosis involves carefully partitioning data into training and testing sets, ensuring model generalizability, minimizing overfitting, and accurately assessing performance, vital for reliable and clinically relevant machine learning and deep learning models. |

|  |  |
| --- | --- |
| WEEK 23: 24/3/2024  to  29/3/2024 | Deepening comprehension of both code and dataset in our breast cancer diagnosis project involves thorough exploration, analysis, and interpretation of algorithms and data, fostering insight into model behavior and performance to inform further refinement and optimization efforts. Preparing for the presentation of our project. |
| WEEK 24: 30/3/2024  to  2/4/2024 | Preparing our report for the expo entails condensing complex project details into a concise, compelling narrative that highlights the problem, methodology, results, and significance. We aim to effectively communicate our breast cancer diagnosis project's innovation and impact, engaging and informing the audience at the expo. |
|  |  |
|  |  |

**Team Members Guide Name**

1) Elton Pereira. (4NM20CS068) **Ms.Jayapadmini Kanchan**

2) Pranav P Shetty (4NM20CS133)

3) Ryan D’Souza. (4NM20CS148)

4) Shashank KT. (4NM20CS164)