Digital Medicine Project Synopsis

Discrimination of benign and malignant mammographic masses based on BI-RADS attributes and the patient's age.

Name: Rohit Sarkar
Branch: Computer Science Engg.
Reg. No. : 160905220

Abstract

In 2019, 271,270 cases of Breast Cancer are expected in the United States. Worldwide, Breast Cancer is the leading type of cancer in women accounting for 25% of all cases. Mammography is the most effective method for breast cancer screening available today. However, the low positive predictive value of breast biopsy resulting from mammogram interpretation leads to approximately 70% unnecessary biopsies with benign outcomes. This project aims to provide a computer aided diagnosis (CAD) system to reduce the high number of unnecessary breast biopsies. This system will help physicians in their decision to perform a breast biopsy on a suspicious lesion seen in a mammogram or to perform a short term follow-up examination instead. A comparative study between various Machine Learning models will be carried out to find an optimal CAD system. The prediction will be based on BI-RADS attributes and the patients age.

Dataset Used

Mammographic Mass Dataset from the UCI Machine Learning Repository.
The data set contains a BI-RADS assessment, the patient's age and three BI-RADS attributes together with the ground truth (the severity field) for 516 benign and 445 malignant masses that have been identified on full field digital mammograms collected at the Institute of Radiology of the University Erlangen-Nuremberg between 2003 and 2006. Each instance has an associated BI-RADS assessment ranging from 1 (definitely benign) to 5 (highly suggestive of malignancy) assigned in a double-review process by physicians.

Expected Outcome

An optimal Machine Learning model that predicts the severity of a mammographic mass based on BI-RADS attributes and thus helping physicians in prevent unnecessary biopsies. Data Visualisations that give insight into the relation between the BI-RADS attributes, the patients and the severity of mammographic masses.

Related Research

1. Shikhman R, Keppke AL. Breast, Imaging, Reporting and Data System (BI RADS) [Updated 2019 Jan 19]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2019 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459169/
2. S. A. Ludwig, “Prediction of Breast Cancer Biopsy Outcomes Using a Distributed Genetic Programming Approach,” Proceedings of the 1st ACM International Health Informatics Symposium, ACM, pp. 694-699,New York, 2010.

3. Miao, Julia H., Kathleen H. Miao, and George J. Miao. "Breast cancer biopsy predictions based on mammographic diagnosis using Support Vector Machine learning." Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Bioinformatics 5.4 (2015): 1-9.