BREAST CANCER DETECTION
PROJECT PROPOSAL

Abdelrahman Karawia
Department of Computer Science
Mansoura University, FCIS
Mansoura

Proposal Report

March 2, 2022

Abstract: **T**his project aims to detect or to be more specific to classify the given data (Radiology or Numerical Medical Results for a breast tumor) which will be given as input to the program is benign or malignant tumor. The problem of tumor classification for its malignancy level has been studied deeply and has a dozen of applications here and there for everyone who's trying to improve a model which has a very reasonable level of predictability. There are enormous number of algorithms to be used in machine learning for such a classification problem and this would a great challenge to choose between, not only this we have also to figure out which parameters for every algorithm we chose (Tuning Process). We will move through two process or steps, the first one is to use different classical machine learning algorithms and observe the differences between them and try to choose the relatively best one among them, then we will turn into neural network model and try to figure out the best number of hidden layers and activation functions to use in this problem. All of this would be in a program with a quite good GUI and the user will have the capability to input either Radiology or Numerical Medical Results and get the results immediately.

1. MOTIVATION

Breast Cancer is considered as the most common type of cancer between females and the first cause of cancer death among them In Egypt, it constitutes 33% of female cancer cases and more than 22,000 new cases diagnosed each year. This is expected to rise **exponentially** over the next years given the enlarging population [1]. **Early Detection and Diagnosis**, it's not just a combination of two words it's the key for treatment and the only way to be far away from the danger zone [2].

Therefore, it's important to develop an effective benign or malignant breast tumor classifier which will be based on Deep and Machine learning Algorithms, it could be of great help to automatically detect and locate tumor tissue cells and to speed up the process. This way one would be able to overcome the dependence on the pathologist which would be especially useful in regions where no experts are available.

2. METHODS

Implementing a deep neural network on large dataset (Radiology or Numerical Medical Results for a breast tumor). ANNs are at the very core of Deep Learning. They are versatile, powerful, and scalable, making them ideal to tackle large and highly complex Machine Learning tasks, such as classifying billions of images [3]. There is now a huge quantity of data available and for our case here in breast cancer dataset as well to train neural networks, and ANNs frequently outperform other ML techniques on very large and complex problems.

Implementing Different machine learning algorithms on the same dataset mentioned above like:

- Logistic Regression
- K- Nearest Neighborhood (KNN)
- Random Forest Model
- Support Vector Machine (SVM)

Which they are the almost the most important algorithms used in classification problems but here we must check the four algorithms using different evaluations methods known and choose the best one for our program to be the one.

2.1 Libraries & Modules

- NumPy
- SciPy
- Matplotlib
- Pandas
- Scikit-Learn
- Seaborn

References

[1] A. H. Abdelaziz, "Breast Cancer Awareness among Egyptian Women," *Clinical Oncology Department, Faculty of Medicine, Ain Shams University,* vol. 17, no. 1-8, 2021.

[2] J. M. A. G. S. L. A. N. a. A. J. C. E. DeSantis, "Breast cancer statistics, 2017, racial disparity in mortality by state," *Cancer Journal for Clinicians,* vol. 67, no. 6, 2017.

[3] A. Géron, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, Sebastopol: O'Reilly Media, Inc, 2019.