Diagnosing malignancy of breast masses using Machine Learning

Theme09 - Introduction to Machine Learning

Performing an EDA and creating a Machine Learning Model in a Java Wrapper

Vincent Talen

389015

BFV3

October 5, 2022

Dave Langers (LADR)

Bart Barnard (BABA)

Diagnosing malignancy of breast masses using Machine Learning

Theme09 - Introduction to Machine Learning
Performing an EDA and creating a Machine Learning Model in a Java Wrapper

Vincent Talen

389015

Bioinformatics

Institute for Life Science & Technology

Hanze University of Applied Sciences

Dave Langers (LADR)

Bart Barnard (BABA)

October 5, 2022

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Tincidunt lobortis feugiat vivamus at augue eget arcu dictum. Arcu ac tortor dignissim convallis aenean et. Sed vulputate odio ut enim blandit volutpat maecenas. Ut diam quam nulla porttitor massa. Ipsum dolor sit amet consectetur adipiscing elit duis. Phasellus faucibus scelerisque eleifend donec pretium. Varius duis at consectetur lorem donec massa sapien. Eget dolor morbi non arcu risus quis varius. Id semper risus in hendrerit gravida rutrum quisque. Pellentesque habitant morbi tristique senectus et. Ut etiam sit amet nisl. Egestas fringilla phasellus faucibus scelerisque eleifend donec pretium vulputate sapien. Nibh tellus molestie nunc non blandit massa enim. Viverra mauris in aliquam sem fringilla ut. Mollis aliquam ut porttitor leo a diam. Sodales ut etiam sit amet nisl purus in mollis nunc. Pellentesque habitant morbi tristique senectus et netus et malesuada fames.

Table of Contents

List of Abbreviations

GCD   Greatest Common Divisor
EDA   Exploratory Data Analysis
LCM   Least Common Multiple

List of Figures

List of Tables

1 Introduction

Diagnosing if breast masses/lumps are benign or malignant, and thus if they are breast cancer, was done by full biopsies, which are invasive surgical procedures. So the researchers wanted a faster and less invasive way to diagnose breast cancer. A promising technique was to take fine needle aspirate fluid samples from the breast mass and then look at the cells under a microscope, the diagnosis would then be based on the characteristics and contextual features of the cells that could visually be seen. This gave mixed results because it was highly subjective and depended a lot on the skill of the physician.

To make this process faster, improve the correctness and objectivity of the diagnosis process, image processing and machine learning techniques were used.

- breast lump or mass
- sample of fluid
- fine needle aspirate
- placed on a glass slide
- stained to highlight nuclei
- microscope creates image
- image is loaded digitally
- the boundaries of nuclei are curve-fitted with the program Xcyt
- ten features are computed
- mean, standard error and extreme values are computed for each image

1.1 Goal

1.2 Research Question

Can breast mass malignancy be assessed with >95% accuracy, using visual features of nuclei boundaries computed from fine needle aspirate images?

1.3 Theory

??

2   Methods

https://github.com/Vincent-Talen/Project-Machine-Learning-Part1_Analysis
https://github.com/Vincent-Talen/Project-Machine-Learning-Part2_Java-Wrapper

3 Results

3.1 Data acquisition and preprocessing

??

3.2 Validation and performance

??

3.3 Research findings

??

4 Discussion and Conclusion

4.1 Discussion

??

4.2 General conclusion and perspective

??

5 References

Mangasarian, Olvi L., W. Nick Street, and William H. Wolberg. 1995. "Breast Cancer Diagnosis and Prognosis via Linear Programming." *Operations Research* 43 (4): 570–77. https://doi.org/10.1287/opre.43.4.570.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Street, W. Nick, W. H. Wolberg, and O. L. Mangasarian. 1993. "Nuclear feature extraction for breast tumor diagnosis." In *Biomedical Image Processing and Biomedical Visualization*, edited by Raj S. Acharya and Dmitry B. Goldgof, 1905:861–70. International Society for Optics; Photonics; SPIE. https://doi.org/10.1117/12.148698.

6 Appendices

Appendix A: '???'