Diagnosing malignancy of breast masses using Machine Learning

Theme09 - Introduction to Machine Learning

Performing an exploratory data analysis, creating a model using machine learning algorithms and creating a Java Wrapper around the model

Vincent Talen

389015

BFV3

October 6, 2022

Dave Langers (LADR)

Bart Barnard (BABA)

Diagnosing malignancy of breast masses using Machine Learning

Theme09 - Introduction to Machine Learning

Performing an exploratory data analysis, creating a model using machine learning algorithms and creating a Java Wrapper around the model

Vincent Talen
389015
Bioinformatics
Institute for Life Science & Technology
Hanze University of Applied Sciences
Dave Langers (LADR)
Bart Barnard (BABA)
October 6, 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

EDA Exploratory Data Analysis
FNA Fine Needle Aspirate

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 performing full biopsies, which are invasive surgical procedures. An alternative method is to take fine needle aspirate (FNA) fluid samples and then have the acting physician try to determine the diagnosis by looking with a microscope at the cell nuclei. The problem with this, and why they thus mainly performed full biopsies, was that this gave mixed results because it was highly subjective and depended a lot on the skill of the physician. But because diagnosing the breast masses by taking FNAs and examining them is much less invasive than the full biopsies it was desired to make this process faster and improve the correctness and objectivity of the diagnosis process. To achieve this goal microscopic images were digitized of the FNA samples and processed into features, machine learning techniques were then used to create a diagnostic model.

1.1 Objective

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

1.2 Theory

??

2 Materials and Methods

Link to GitHub Repository of Part 1: Analysis:
https://github.com/Vincent-Talen/Project-Machine-Learning-Part1_Analysis

Link to GitHub Repository of Part 2: Java Wrapper:
https://github.com/Vincent-Talen/Project-Machine-Learning-Part2_Java-Wrapper

2.1 Materials

??

2.2 Existing Methods

??

2.3 Developed Methods

??

3 Results

3.1 Data

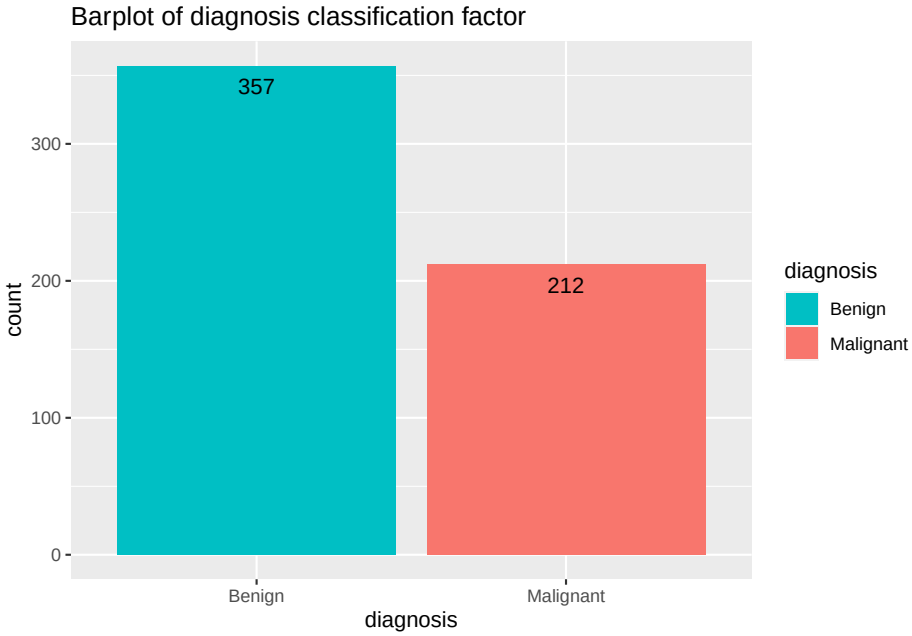

Figure 1: Barplot showing distribution of diagnosis classification labels

4 Discussion and Conclusion

4.1 Discussion

??

4.2 General conclusion and perspective

??

5 Project proposal for Minor

??

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.

Appendices

Appendix A: '???'