Project Title: Analysis of Breast Cancer

Group Members: Benjamin Barnett, Yaniv Bronshtein, Krishna Piratla

For our project, we would like to evaluate the characteristics of benign and malignant breast tumors. To accomplish this, we would like to study the computed features from digitized images of the fine needle aspiration of breast mass cell nuclei, including radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. We plan on training a number of different models such as logistic regression, linear discriminant analysis, random forest, support vector machine, and the multi-layer neural network to test performance on our split data. We would also like to evaluate the ROC curves for each method and try different threshold values (such as 0.2, 0.5, and 0.8) to compare classification accuracy. This project aims to find a model with high prediction performance and answer the key question: How can we best distinguish between benign and malignant breast tumors?

Data Source: https://www.kaggle.com/uciml/breast-cancer-wisconsin-data