Final Project Document

Duyen Ho, Michael Garbus, Christina Hu

10/28/2021

Table of Contents

## Project Description and Summary (NEED GOAL, APPROACH, and CONCLUSION AND THIS NEEDS TO BE ONE PAGE?!!?)

The goal of this project is to find the most accurate model using a variety of statistical model building techniques in order to predict medical outcomes relating to breast cancer. We will use the BRCA Multi-Omics (TCGA) data from Kaggle, and predict outcomes for the PR.Status, ER.Status, HER2.Final.Status, and histological.type. For our approach, we will first perform the necessary data cleaning operations, then [method 1] and [method2] to build classification models for PR.Status and [method1] and [method2] for histological.type, using classification error and AUC respectively as the evalution criteria. Next, we will use [method] to build a model with the goal of accurately predicting **all four** outcomes. Finally, [INSERT CONCLUSION HERE]