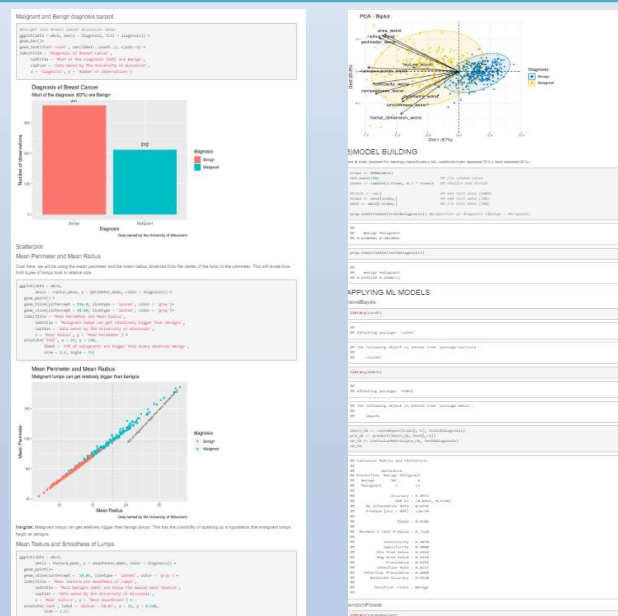
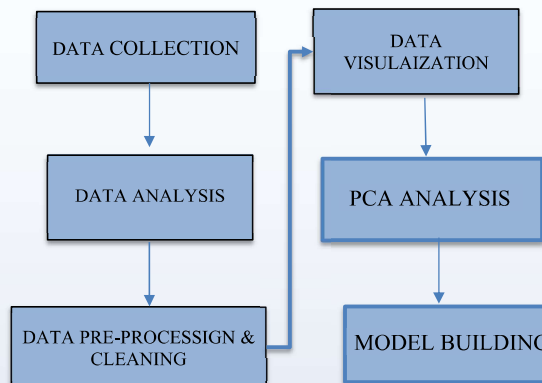


Breast cancer affects the majority of women worldwide, and it is the second most common cause of death among women. However, if cancer is detected early and treated properly, it is possible to be cured of the condition. Early detection of breast cancer can dramatically improve the prognosis and chances of survival by allowing patients to receive timely clinical therapy. In this paper, we explore machine learning models that can be applied to help increasing the accuracy of the diagnosis of breast cancer. Classic machine learning models including Naïve bayes, Rpart, Random Forest, AdaBoost, SVM and Tune SVM are tested on the Breast Cancer Wisconsin dataset. The performance of the study is measured with respect to accuracy, sensitivity, specificity, negative predictive value and positive predictive value. The results reveal that the SVM tune has obtained the highest accuracy of 98.83 whereas 94.15%, 95.25%, 97.66%, 97.75% and 98.2% are accuracies obtained by Naïve Bayes, RPart, Random Forest, AdaBoost and SVM respectively.

The main problem of the project is to detect breast cancer based on a set of features calculated from a digitized image of the Fine Needle Aspiration (FNA) of a breast mass from a patient. We present a diagnosis model using both traditional and deep learning machine learning models. The Wisconsin Breast Cancer dataset is obtained from a prominent machine learning database named UCI machine learning database. Through analysing the dataset and applying different ML models we try to get the best predicting model.



- RSTUDIO
- GGLOT2
- RPART
- DPLYR
- CARET
- ADA

