**Title:** Breast Cancer Detection using Classification Algorithms

**Abstract:**

Disease Diagnosis is a major application of Big Data Analytics in health care. Machine Learning algorithms are widely used for disease diagnosis especially for breast cancer. My objective through this project is to apply three classification algorithms on a Breast Cancer Dataset to classify whether the tumor is benign or malign. The dataset has 699 observations with 10 variables. Algorithms will classify each data as benign or malign. I am considering this dataset as of now, I am also looking for a better data set simultaneously.

**Domain of Study:** Machine Learning algorithms for Disease Diagnosis

**Algorithms:** I am planning to apply classification algorithms – Logistic Regression, Decision Trees, Random Forests and Support Vector Machine on a dataset and compare their performance.

**Data Source:**

<http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>

**References:**

<https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>

<https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>

<https://www.digitalocean.com/community/tutorials/how-to-build-a-machine-learning-classifier-in-python-with-scikit-learn>

<https://machinelearningmastery.com/machine-learning-in-python-step-by-step/>

<https://medium.freecodecamp.org/the-hitchhikers-guide-to-machine-learning-algorithms-in-python-bfad66adb378>