**Predicting Breast Cancer**

**Data Report**

**By**

**Yeshwanth**

**The Seattle Grace Hospital**

**Executive Summary**

This report is for the better understanding of “Predicting breast cancer”. This document provides a clear picture about how to find if the tumors is malignant or benign in human breast tissue. These are the requirements of Grace Hospital. They are looking for a model that can help them to identify the breast cancer but with a less invasive model. This model presents the exact information the hospital requires for predicting breast cancer. All the data we require is already provided by the hospital. They need a new technology to identify tumors by using data analyzing techniques.

The model which we designed uses random forest method. The new technology is totally reliable as it is providing almost accurate results.

**Machine Learning:**

This model is done by creating machine learning which can predict malignant or benign tumors to identity breast cancer.

**About the Model:**

This model is based on random forest model which is very reliable and is a machine learning model which can predict the tumors almost accurately.

This model randomly chooses some data and then it verifies them for the malignant or benign tumors based upon different variables. After one set of variables are done it selects other set of variables and then it chooses few more set of variables. This process continues till it verifies all the variables and then it provides results for all the variables it examined.

Thus this model provides with best possible way of predicting the tumors.

The success rate of this model is very high. It can be used with the latest technology by the grace hospital. Its accuracy rate is more than 90 percent.

This model is a comparison between RandomForestRegressor and RandomForestClassifier.

Both the models have been performed and compared. It turns out that the RandomForestRegressor model is the best way to identify if the tumor is malignant or benign.As it has an accuracy of more than that of the classifier.

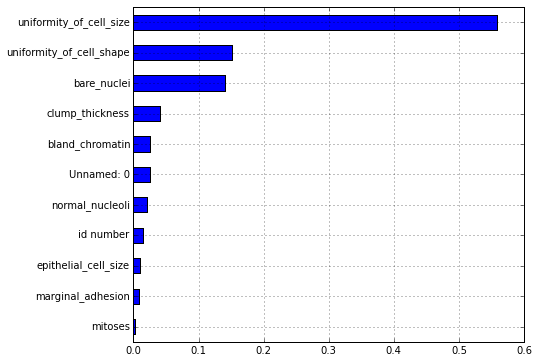
**Questions by the lawyers:**

1. The model is almost accurate and it can predict malignant tumors and benign tumors up to an accuracy of more than 90 percentage.
2. There is a less than 10 percent risk that this model will not identify if a tumor is malignant or benign.

**Variables:**

The following variables can be used to predict the tumors. These variables make a huge difference when identifying the tumors.

1. Uniformity\_of\_cell\_size
2. Uniformity\_of\_cell\_shape
3. Bare\_nuclei
4. Dump\_thickness



**Conclusion:**

This model is very invasive and it is can predict the malignant tumors based on the scan results that are provided by the