

# Diagnosing Breast Cancer using Julia

---

Dipta Paul, Md Tarikul Islam

[dpaul143008@bscse.uiu.ac.bd](mailto:dpaul143008@bscse.uiu.ac.bd), [mdtarikulislam6174@gmail.com](mailto:mdtarikulislam6174@gmail.com)

## Problem Description

Breast cancer is one of the major causes of the high number of women's death. Therefore, early diagnosing breast cancer can promote timely clinical treatment to patients. Moreover correct classification of the benign tumor can prevent patients from undergoing unnecessary treatments. Thus, diagnosing breast cancer and accurate prediction of malignant and benign tumors is an important research issue. In this work, we have diagnosed breast cancer on the [Breast Cancer Wisconsin\(Diagnostic\) Dataset](#) which is also available in the UCI Machine Learning Repository.

## Dataset Description

---

| Dataset Name        | Breast Cancer Wisconsin(Diagnostic) Dataset |
|---------------------|---|
| Source              | <a href="#">Kaggle</a>                      |
| Number of Instances | 569   |
| Features            | 31  |
| Feature Type        | Numeric and Categorical                     |
| Missing Values      | No  |
| Class Value         | 2   |
| Learning Type       | Classification                              |
| Class Instances     | Benign => 357   Malignant => 212            |

## Proposed Solution

---

The proposed solution uses a **random forest classifier** to diagnose the breast cancer.

**Parameters:**  $n\_trees = 100$ ,  $n\_subfeatures = 20$ ,  $max\_depth = 7$

# Results

## Result Table

|  |           |
|--|-----------|
| <b>Accuracy</b>                                    | 0.964912  |
| <b>Sensitivity or Recall or True Positive Rate</b> | 0.936508  |
| <b>Specificity or True Negative Rate</b>           | 0.981481  |
| <b>AUC</b>   | 0.958995  |
| <b>Gmean</b>                                       | 0.968731  |
| <b>Precision</b>                                   | 0.967213  |
| <b>F1_score</b>                                    | 0.951613  |
| <b>False Positive Rate</b>                         | 0.0185185 |

## ROC Curve

