**CSE523: Machine Learning**

**Project Progress Report**

**Breast Cancer Detection**

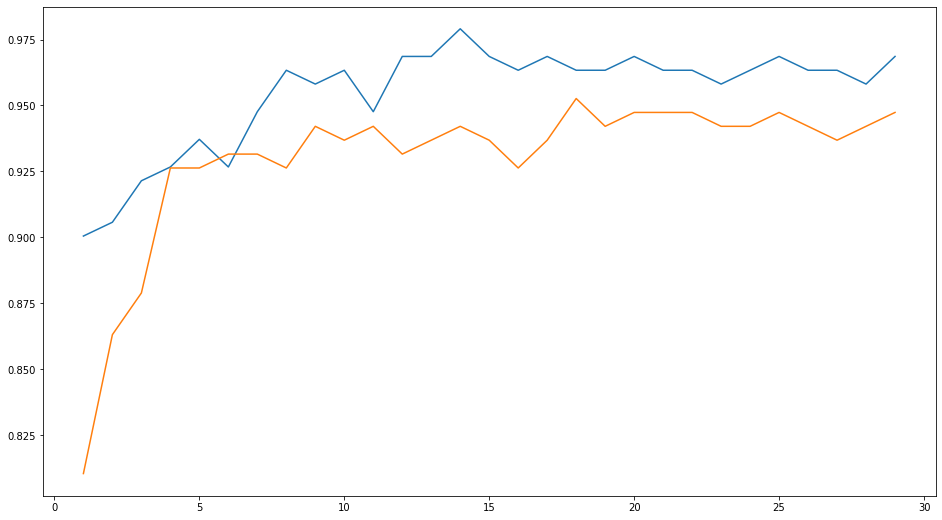
**Group:  The Curators**

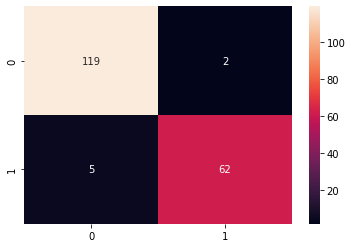
| **Name** | **Roll number** |
| --- | --- |
| **Harsh Patel** | **AU1940114** |
| **Kavan Desai** | **AU1940126** |
| **Sarthak Bharad** | **AU1940176** |

**Work done this week:**

* Worked on Feature importance defining the feature importance.
* We worked on implementing random forest which provides some exclusive best features which are the most important feature to be considered explicitly.
* Using random forest we got an accuracy of 97%
* The features were reduced to 14 from total of 29 features
* We performed Recursive feature elimination (RFE) with random forest Basically, it uses one of the classification methods (random forest in our example), assigning weights to each of features. Whose absolute weights are the smallest are pruned from the current set features. That procedure is recursively repeated on the pruned set until the desired number of features. RFE is a wrapper-style feature selection algorithm that also uses filter-based feature selection internally.

**Outcomes of the work done:**

****



**LINK:-** [**GoogleCollab**](https://colab.research.google.com/drive/1-mlAUkQt--MynKE7K-UQZdrDHUvlg2xK#scrollTo=BxjoFfZ_KJcS)

**Future work:**

We will work on random forest and will work more on feature importance to reduce dimensionality. We might throw some light on PCA if possible.