Early stage breast cancer prediction

Yasser - Keziah - Mariama

Data splitting

Features selection

Training set

128

70%

Test set

56

30%

Using lasso

153

Features remaining

Models

1 Linear SVC

3 Nu SVC

5 Logistic Regression CV

2 Gradient Boosting Classifier

4 Adaboost Classifier

Voting

GBC		77%
LRCV		74.5%
SVM		73%
Nu SVM		73%
Ada		70%
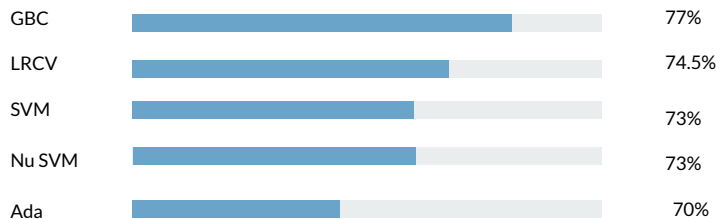
Testing accuracy

76%

Confusion matrix

	precision	recall	f1-score	support
-1	0.76	0.80	0.78	35
1	0.63	0.57	0.60	21
micro avg	0.71	0.71	0.71	56
macro avg	0.69	0.69	0.69	56
weighted avg	0.71	0.71	0.71	56

https://github.com/BYSD/Kernel-Methodes