

# Clasificación de sentimientos – Base de datos IMBD

Por Juan Pablo Bernal Lafarga – A01742342

## BOW (Bag Of Words)

Accuracy				
Models/Parameters	Logistic Regression	Random Forest	Support Vector Classifier	Gradient Boosting Machines
maxtokens = 50, maxtokenlen = 20	70%	68.33%	67.66%	66.66%
maxtokens = 100, maxtokenlen = 100	72.66%	72.83%	72.66%	71%
maxtokens = 200, maxtokenlen = 100	80.33%	79.83%	79.16%	73.66%

## TF-IDF

Accuracy				
Models/Parameters	Logistic Regression	Random Forest	Support Vector Classifier	Gradient Boosting Machines
maxtokens = 50, maxtokenlen = 20	70.5%	65.16%	70.5%	65.33%
maxtokens = 100, maxtokenlen = 100	76.16%	70%	75.5%	69.5%
maxtokens = 200, maxtokenlen = 100	82%	77.33%	82.66%	74.66%

Los cuadros de color verde son los mejores scores por fila. El modelo más exacto según la configuración de parámetros.