

Testing ADABOOST and Bagging

Para esta tarea se utilizaron las características extra extraídas en la entrega previa:

- Cantidad de menciones que contiene el tweet (numérica).
- Es retweet o no (nominal).
- Cantidad de ligas que contiene el tweet (numérica).
- Cantidad de palabras que contiene el tweet (numérica).
- 4 atributos de personalidad provistos por la api de indico.

Bagging

Se probó Bagging con diferentes clasificadores, Decision Stump, REPTree, Linnear Knn y Random Forest. Se varió también el tamaño de la bolsa.

[Bagging20BSP100lwREPTree-18F.csv](#)

0.69043

3 hours ago by [Carlos A](#)

Bagging 20 Bag Size Percent, 100 Iterations with REPTree: 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[Bagging10BSP100lwDS-18F.csv](#)

0.61928

3 hours ago by [Carlos A](#)

Bagging 20 Bag Size Percent, 100 Iterations with Decision Stump: 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[Bagging25PBS10lwKNN-18F.csv](#)

0.64871

3 hours ago by [Carlos A](#)

Bagging 20 Bag Size Percent, 10 Iterations with best Knn: 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[Bagging50PBS10lwBRF-18F.csv](#)

0.72644

3 hours ago by [Carlos A](#)

Bagging 50 Bag Size Percent, 10 Iterations with best random forest: 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[Bagging20PBS10lwBRF-18F.csv](#)

0.65735

4 hours ago by [Carlos A](#)

Bagging 20 Bag Size Percent, 10 Iterations with best random forest: 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

ADA Boost

Se probó ADA Boost con dos clasificadores: Decision Stump y REPTree. Los resultados fueron muy similares.

[ADA1000wDS-18F.csv](#)

0.68414

3 hours ago by [Carlos A](#)

ADABOOST 1000 Iterations with Decision Stump: 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[ADA10000lwDS-18F.csv](#)

0.68085

3 hours ago by [Carlos A](#)

ADABOOST 10000 Iterations with Decision Stump : 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[ADA10000lwREPTree.csv](#)

0.68003

3 hours ago by [Carlos A](#)

ADABOOST 10000 Iterations with REPTree : 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

[ADA1000lwREPTree.csv](#)

0.68003

3 hours ago by [Carlos A](#)

ADABOOST 1000 Iterations with REPTree : 5 sentiment 5 emotion 4 personality, number of mentions, retweet and count of links and word

Conclusiones

Los resultados obtenidos mejoran la eficacia con respecto a los obtenidos por los clasificadores sólo, sin embargo, son inferiores a los obtenidos haciendo Resampling y aplicando sensibilidad al costo.