PROJECT 5

ANALYSIS

* Are the TSNE graphs correct ?

SUPERVISED

* Can I use XGBoost instead of RF ?Same result less parameters to optimize.
* Should I include a confusion matrix ?
* Si trop de temps a train tu pourrais splitter en 40-30-30 et utiliser un set de validation.

MLOPS

* What should be unit tested here ? The words embedding functions ?
  + Test ta methode de clean et preprocessing.
  + Bert et USE tu peux comparer les shapes avant apres.

Check what matrice creuse est pour prez.