

4) Analyse des performances des modèles ..

Méthodes	Classifieurs	Performances	Explications
TF-IDF	MultinomialNB	Matrice de confusion [[30 78] [1 91]] accuracy : 0.60 recall: 69% bcp de faux négatif	Nous notons beaucoup de faux positifs ce qui ne donne pas un meilleur score. les notes attribuées sur Amazons surtout les critiques négatives ne sont généralement pas à un contenu négatif
	LogisticRegression	score de :0.665 Matrice de confusion: [45 63] [4 88]	les notes attribuées sur Amazons surtout les critiques négatives ne sont généralement pas à un contenu négatif
Word2Vec	LogisticRegression	Score : 0.47 matrice de confusion [0 53] [0 47] recall = 1 (faux négatif)	Qui est du à la notation des commentaires et aussi aux similarités qui ne sont pas conformes. Une réduction de dimension a été fait ce qui a amélioré juste un peu le classifieur avec un score de 60%
CAMEMBERT	LogisticRegression	score : 0.72 matrice de confusion: [49 25] [10 41]	Un score amélioré grâce à Camembert.. Des faux négatifs qui ont largement baissé..

Durant ce projet de classification, j'ai implémenter des modèles de classifications dans le but de classer des commentaires tirés sur soit positif ou négatif.

Titre du livre: Les cinq blessures qui empêchent d'être soi-même"

Lien

:https://www.amazon.fr/cinq-blessures-emp%C3%A0chent-d%C3%A0tre-soi-m%C3%A0me/product-reviews/2266229486/ref=cm_cr_ar_p_d_paging_btm_2?ie=UTF8&pageNumber=1

Les performances sont différentes et des écarts sont notés.