

## 1 - Quiz

**A quoi servent la lemmatisation et la stemmification ?**



A standardiser le vocabulaire



A normaliser la norme des vecteurs de représentation



A augmenter la taille du vocabulaire



Ce sont des algorithmes de classification



2 - True or false

**Les vecteurs TF-IDF contiennent beaucoup de 0**



True



False



### 3 - Quiz

**La taille d'un vecteur BoW est typiquement de l'ordre de ...**



100



10000



$10e^6$







10



#### 4 - Quiz

**Pour fixer mes hyper-paramètres, je regarde les métriques de performances sur ...**

	Le jeu d'entrainement	✗
	le jeu de validation	✓
	Le jeu de test	✗
	un échantillon du jeu d'entrainement	✗

5 - True or false

**La LDA est un algorithme de classification de texte ?**



True



False



## 6 - Quiz

**Comment sont modélisés les topics dans la LDA ?**



Par une distribution sur l'ensemble des documents



Par une liste de mots caractéristiques



Par un mot clé



Par une distribution sur le vocabulaire



## 7 - Quiz

**Dans la LDA, le nombre de topics est ...**



Fixé à postériori quand on a appris les paramètres du modèle



Un hyper-paramètre, il est fixé au début



Obtenu à la fin, c'est un output du modèle



Fixé lors de l'apprentissage sur les données



## 8 - Quiz

**Les paramètres de la LDA peuvent être appris à l'aide de**



Le sampling de Gibbs



La back propagation du gradient



L'inférence variationnelle



Une génération aléatoire





9 - True or false

**La LDA prend en input un unique document**



True



False

