

XLNET

XLNet est un modèle de langage qui utilise des unités Transformer et qui nécessite un pré-entraînement intensif avant son implémentation pour des tâches particulières. Plus spécifiquement, XLNet est un modèle autorégressif (AR) généralisé pour la compréhension de langage naturel où chaque mot dépend de tous les mots précédents. Il est généralisé parce qu'il capture le contexte de façon bidirectionnelle au moyen d'un mécanisme de modélisation par permutation de langage ('permutation language modeling' (PLM)) en entraînant un modèle AR sur toutes les permutations possibles de mots dans une phrase tout en conservant l'ordre des mots dans la séquence originale.

XLNet maximise la vraisemblance logarithmique attendue sur toutes les permutations possibles de la séquence de mots. Chaque position apprend alors à utiliser l'information contextuelle de toutes les positions, ce qui capture le contexte bidirectionnel.

Les modèles XLNet de base et large ont les mêmes hyper paramètres d'architecture que BERT, ce qui donne des modèles de même grandeur.

Résultats sur la compréhension en lecture

Modèle	Précision de course	SQuAD1.1 EM	SQuAD2.0 EM
BERT-Grand	72,0	84.1	78,98
Base XLNet			80.18
XLNet-Large	81,75	88,95	86.12

Résultats sur la classification de texte

Modèle	IMDB	Yelp-2	Yelp-5	DBpédia	Amazone-2	Amazone-5
BERT-Grand	4.51	1,89	29.32	0,64	2.63	34.17
XLNet-Large	3,79	1,55	27,80	0,62	2.40	32.26

Les chiffres ci-dessus sont des taux d'erreur.

<https://github.com/zihangdai/xlnet/>

https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/26062/Archambault_Jean_2021_memoire.pdf?sequence=2