

# Natural Language Processing

Le traitement automatique du langage (NLP pour Natural Language Processing) applique des modèles de machine learning (ML) au texte et à la langue. Utiliser les algorithmes ML pour comprendre ce qui est dit dans les mots parlés et écrits est l'objet du traitement du langage naturel. Chaque fois que vous dictez quelque chose dans votre iPhone / appareil Android qui est ensuite converti en texte, c'est un algorithme NLP en action.

Vous pouvez également utiliser NLP dans une révision de texte pour prédire si la révision est bonne ou mauvaise. Vous pouvez utiliser NLP sur un article pour prédire certaines catégories d'articles que vous essayez de segmenter. Vous pouvez utiliser NLP sur un livre pour prédire le genre du livre. Et cela peut aller plus loin, vous pouvez utiliser NLP pour construire un traducteur automatique ou un système de reconnaissance vocale, et dans ce dernier exemple, vous utilisez des algorithmes de classification pour classer le langage. Parlant des algorithmes de classification, la plupart des algorithmes NLP sont des modèles de classification, et ils incluent la régression logistique, Naive Bayes, CART qui est un modèle basé sur les arbres de décision, l'entropie maximale liée aux arbres de décision, les modèles de Markov cachés.

Un modèle très connu en NLP est le modèle Bag of Words. C'est un modèle utilisé pour pré-traiter les textes à classer avant d'ajuster les algorithmes de classification sur les observations contenant les textes.

Dans cette partie, vous comprendrez et apprendrez comment:

1. Nettoyer les textes pour les préparer aux modèles ML,
2. Créer un modèle Bag of Words,
3. Appliquez des modèles ML à ce modèle Bag of Words.