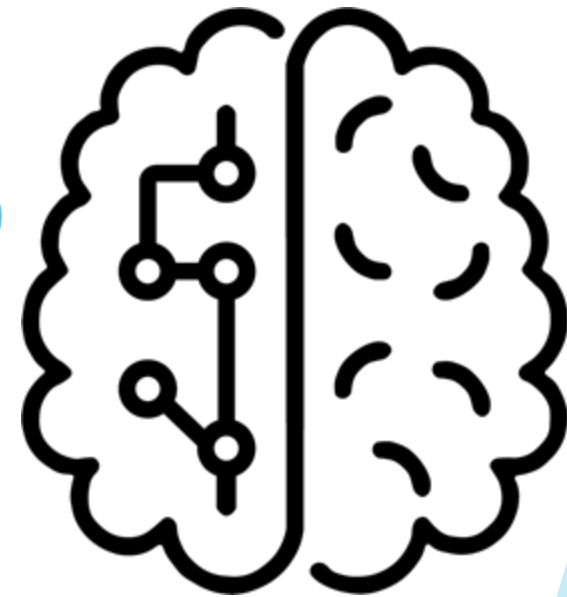



Le Traitement du langage naturel (NLP)

AIForYou - Morgan Gautherot



Chaîne You Tube



AI FOR YOU

Alforyou - Morgan Gautherot
2,2 k abonnés

PERSONNALISER LA CHAÎNE GÉRER LES VIDÉOS

ACCUEIL VIDÉOS PLAYLISTS COMMUNAUTÉ CHAÎNES À PROPOS

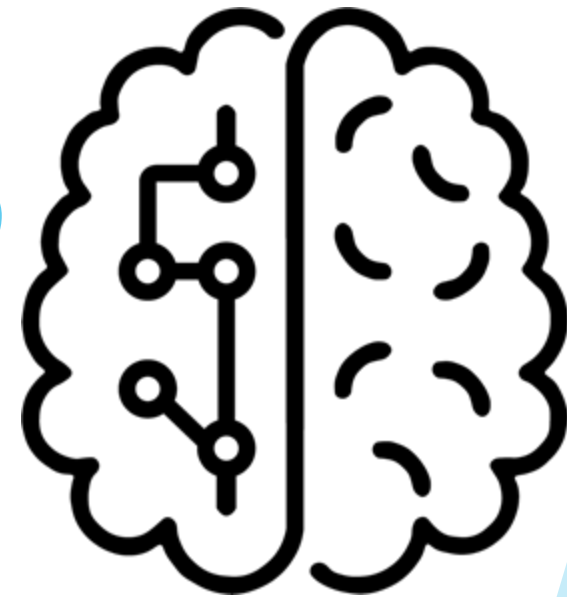
Les modèles d'arbres de décision ▶ TOUT REGARDER

Dans cette série de vidéos, vous allez découvrir les algorithmes de machine learning à base d'arbres de décision. Vous allez enfin comprendre ce qui se cache derrière les arbres de décision...

Formation Machine Learning : Les modèles d'arbres	Formation Machine Learning : Les modèles d'arbres	Formation Machine Learning : Les modèles d'arbres	Formation Machine Learning : Les modèles d'arbres	Formation Machine Learning : Les modèles d'arbres	Formation Machine Learning : Les modèles d'arbres
1. Introduction aux modèles d'arbres de décision	2. Tout savoir sur les arbres de décision	3. Implémentation des arbres de régression from scratch	4. Implémentation des arbres de classification from scratch	5. Entraîner un modèle d'arbres de classification avec Scikit-Learn	6. Entraîner un modèle d'arbres de régression avec Scikit-Learn
3:37	16:58	15:46	14:55	17:15	15:03
Formation de machine learning sur les modèles...	Tout savoir sur les arbres de décision - LES MODELES...	Coder un algorithme d'arbre de décision from scratch - ...	Coder un algorithme d'arbre de décision from scratch - ...	Entraîner un modèle d'arbre de décision en Python avec...	Entraîner un modèle d'arbre de décision en Python avec...
Alforyou - Morgan Gautherot 509 vues • il y a 1 mois	Alforyou - Morgan Gautherot 22 k vues • il y a 1 an	Alforyou - Morgan Gautherot 1,7 k vues • il y a 1 an	Alforyou - Morgan Gautherot 1,2 k vues • il y a 1 an	Alforyou - Morgan Gautherot 1,5 k vues • il y a 1 an	Alforyou - Morgan Gautherot 1,1 k vues • il y a 1 an

Le Traitement du langage naturel (NLP)

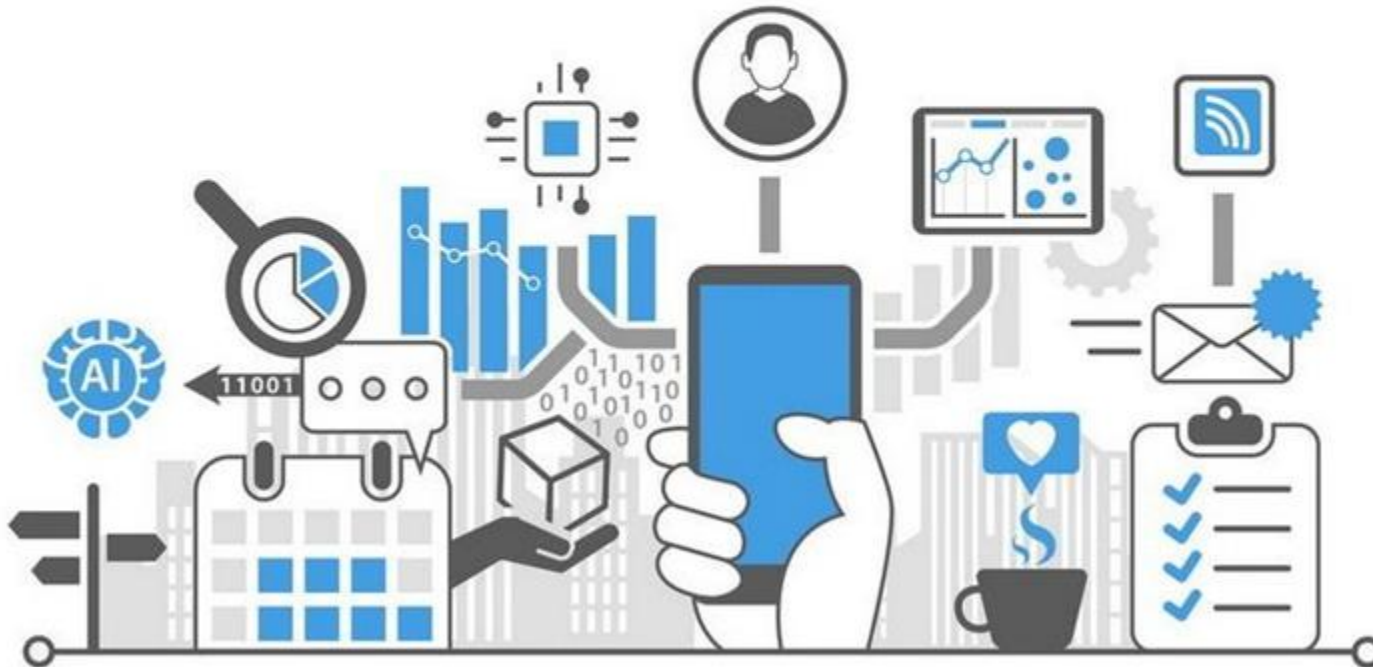
AIForYou - Morgan Gautherot



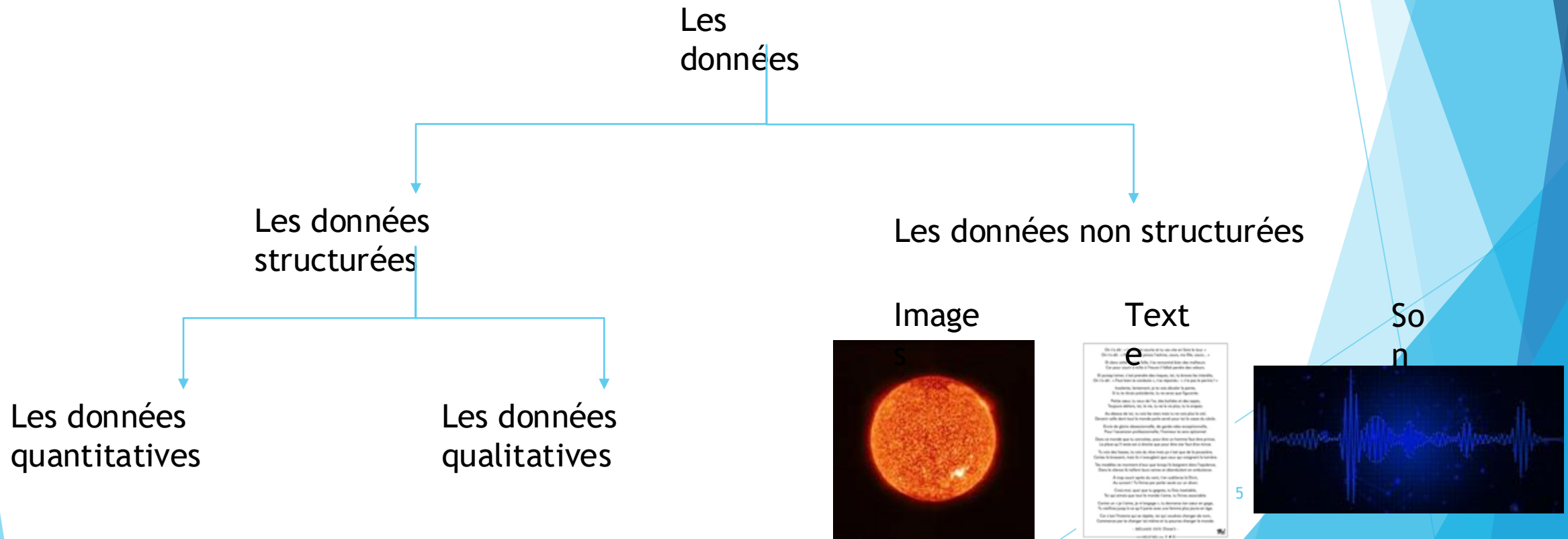
Le Big Data

Seulement 21% de ces données sont structurées !

BIG DATA



Les catégories de données



Le texte dans le big data



Google



Gmail



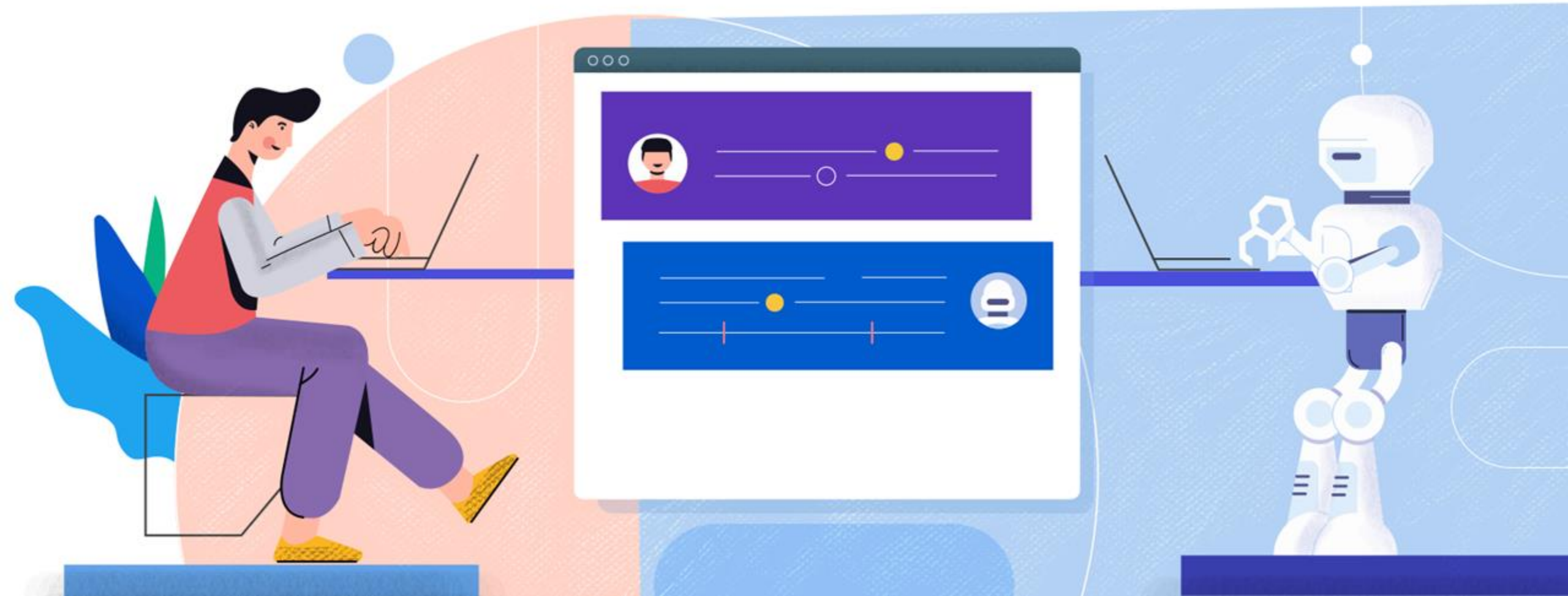
Langage informatique



Langage naturel



Natural Language Processing (NLP)



Text
e

• MÉLANIE DITE DIAM'S •

—www.MELBYMEL.com



Données exploitables

[illegible]

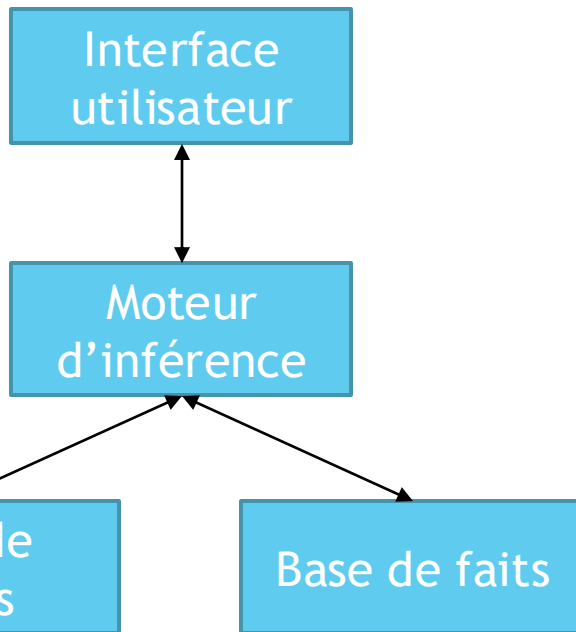
Modèle

S

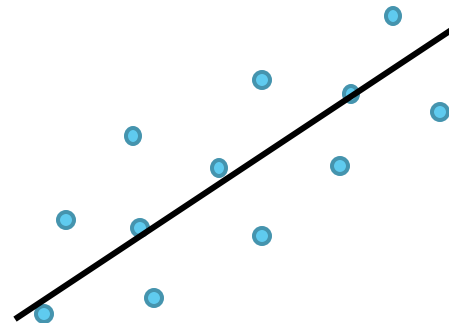
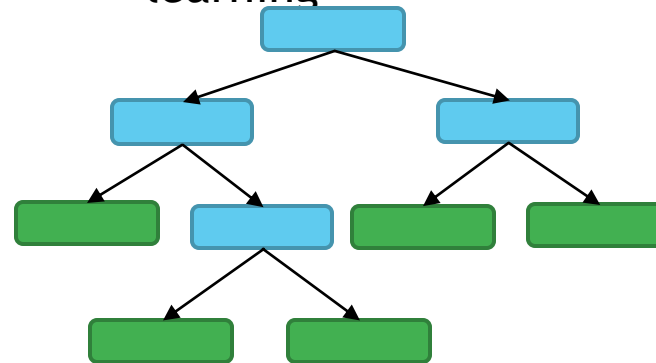


Evolution des méthodologies

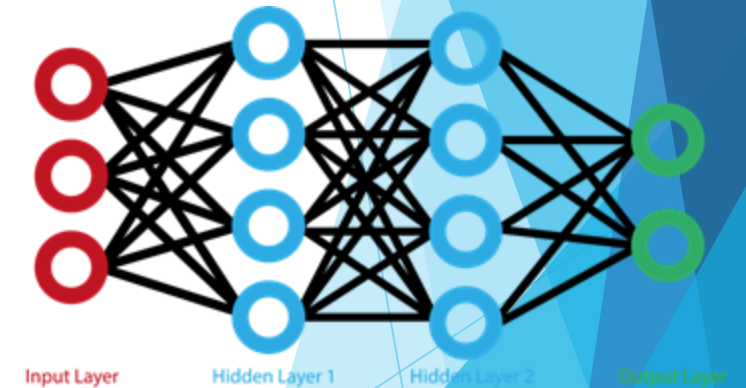
Systeme expert





Machine learning



Modèle s



Traduction automatique

 [Traducteur](#) [Linguee](#) [DeepL Pro](#) [Blog](#) [Infos](#) 

Traduire de **n'importe quelle langue** ▼

Saisissez votre texte, collez-le
ou déposez votre document.

(français, anglais, allemand, espagnol, italien, néerlandais,
polonais)

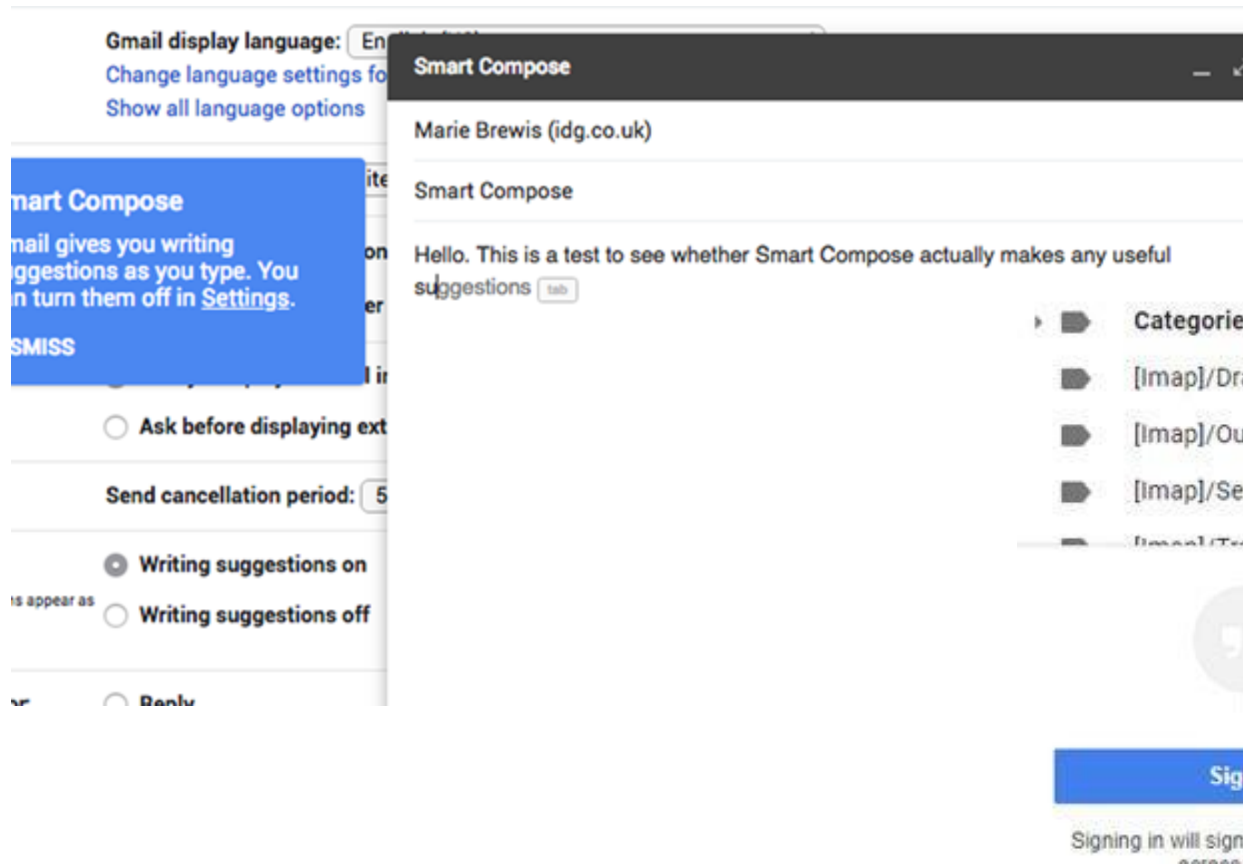
⤶ Traduire le document

>

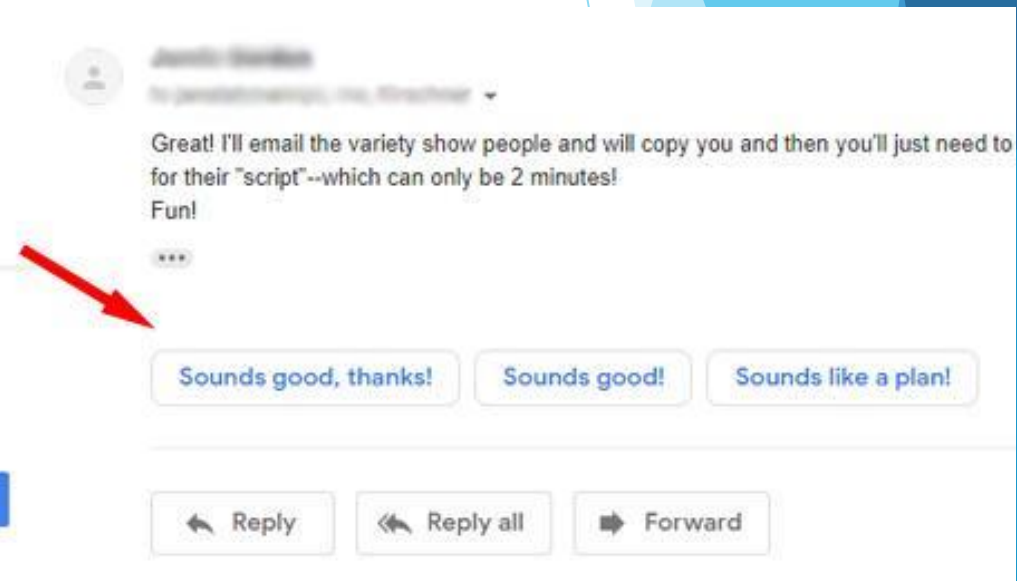
Traduire en **français** ▼

Génération de textes

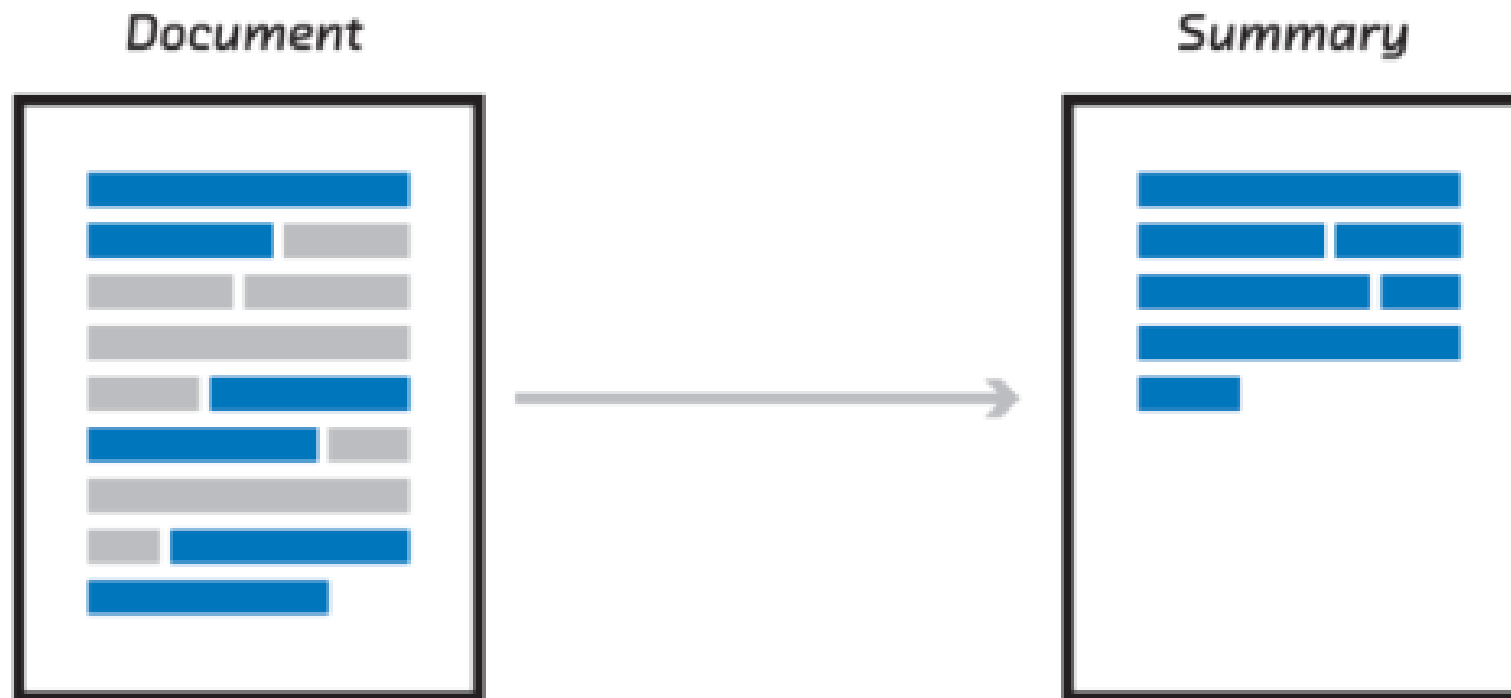
Auto complétion



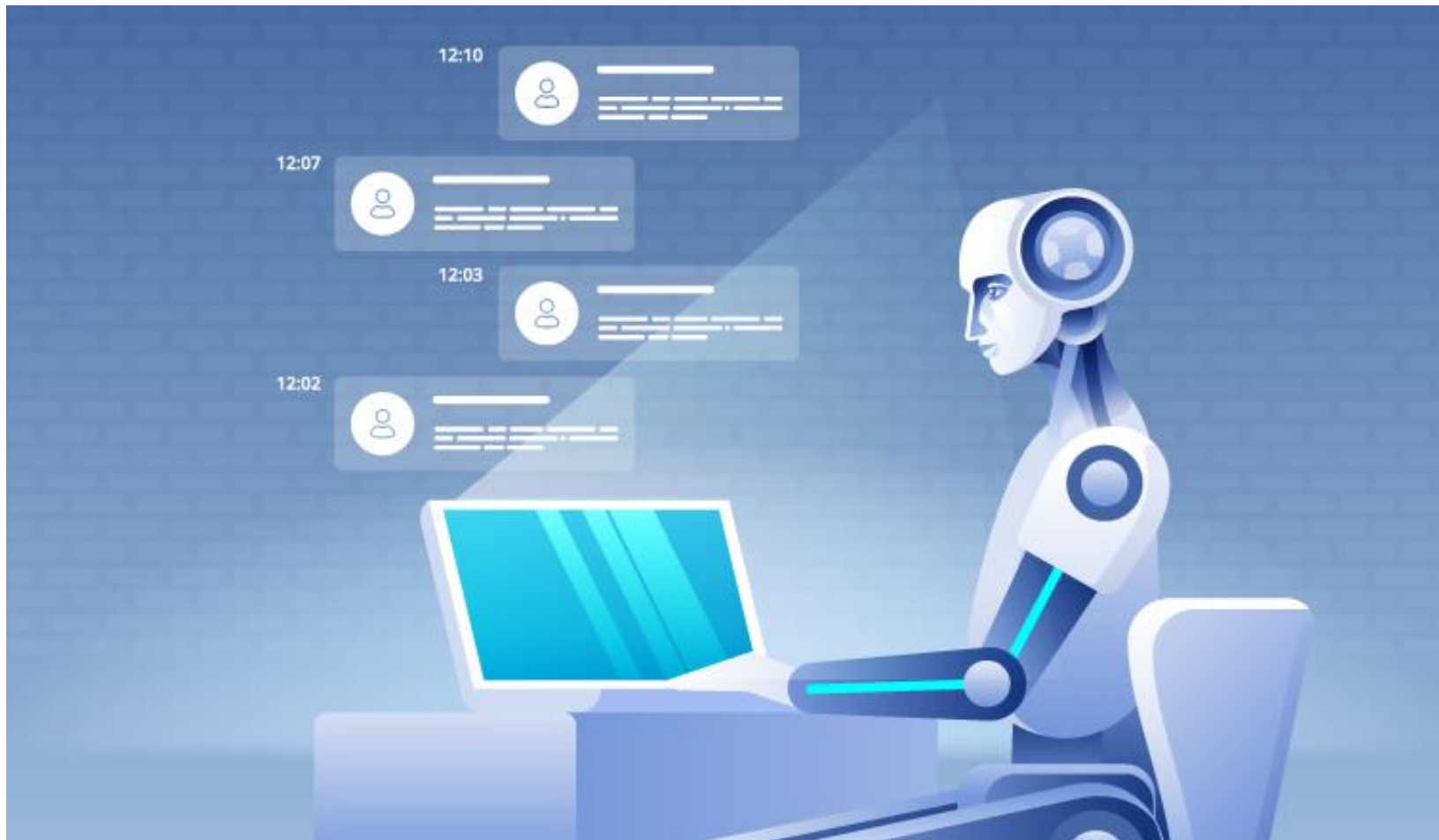
Suggestion de réponse



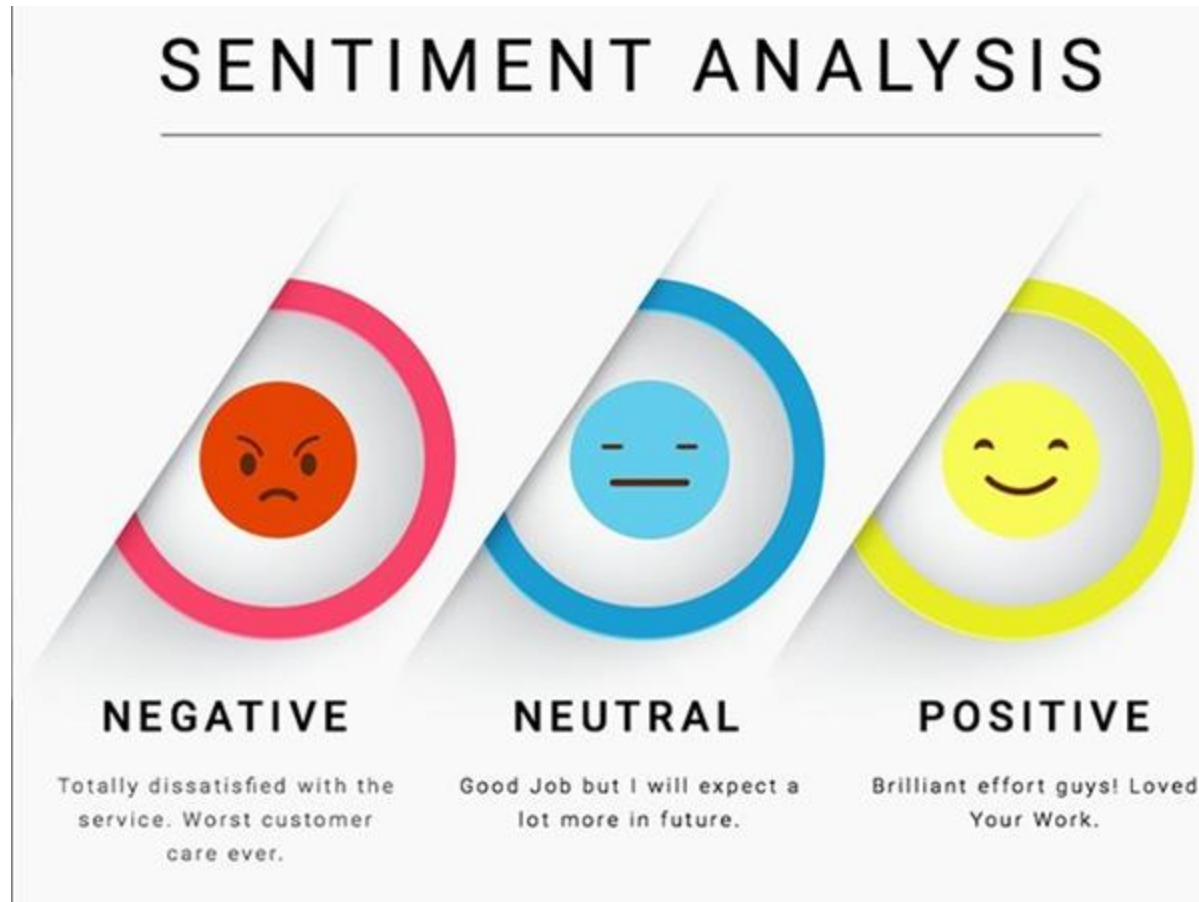
Résumé automatique



Chatbot



Analyse du sentiment



Sommaire

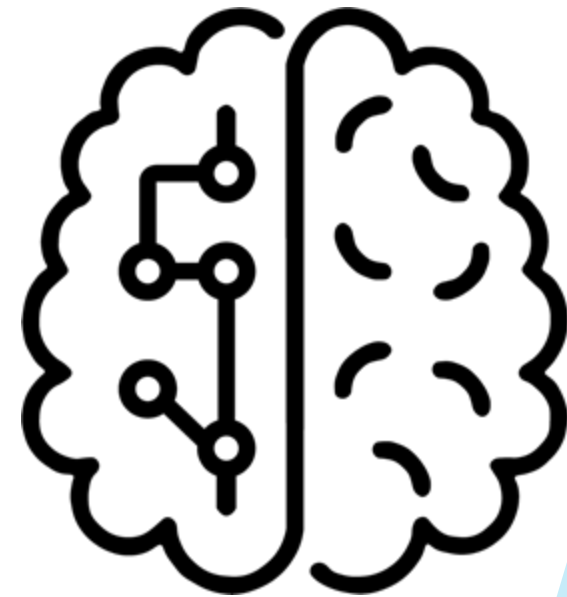
Partie 1 : Prétraiter les données textuelles

Partie 2 : Transformer les données textuelles

Partie 3 : Modélisation des données textuelles

Normalisation des données textuelles

AIForYou - Morgan Gautherot



Exploration d'un corpus de texte

- Fréquence des mots
- La combinaison des mots
- La richesse du vocabulaire employé
- Le distribution du register
- ...



tonceurbat

Tu me donnes envie de faire table rase du passé, et de t'aider à faire table rase du tien. Tu me donnes envie d'être pour toi ce que je n'ai jamais su être pour quelqu'un d'autre. Tu me donnes envie de t'offrir mon cœur, mes ambitions, mon amour sur un plateau d'argent. Tu me donnes envie de t'aimer à en faire trembler le monde entier. Tu me donnes envie de parler de sentiments et de bonheur tout en conjuguant mes verbes au présent. Tu me donnes envie de devenir quelqu'un qui t'apportera fierté et douceur. Tu me donnes envie de décrocher la Lune, de croire en l'impossible. Tu me donnes envie de vivre, bordel. Parce que tu es l'étincelle au fond de mes yeux, mon sourire terriblement sincère. Parce que je ne suis plus seule désormais, tu es là. Et c'est si précieux.

Le supermarché

Sophie fait les courses au supermarché avec son neveu Clément. Après avoir garé sa voiture sur le parking, elle a pris un chariot dans lequel Clément s'est installé. Tous deux pénètrent maintenant dans le magasin. Les clients sont nombreux à cause des promotions et des soldes. Sophie sort sa liste des courses pour ne rien oublier. Elle regarde les marques et compare les prix. Elle choisit des produits frais, des conserves et du poisson surgelé qu'elle met dans son chariot. Clément est turbulent, il s'agite sur son siège, il ne cesse de réclamer des bonbons, des biscuits, du chocolat, des jouets. Sophie le grande gement puis elle regarde sa montre, il est déjà seize heures. Elle doit se hâter de rentrer car ce soir elle est invitée avec Jean à dîner chez la maman de Clément, sa sœur Lucie. Sophie termine les courses, elle se dirige vers la caisse où elle attend calmement son tour. À la caisse, elle place ses achats sur le tapis roulant puis elle règle sa note avec sa carte bancaire. À présent, elle regagne le parking, cherche quelques instants sa voiture. Elle la repère au fond de l'allée puis se dirige vers son véhicule. Elle ouvre le coffre et y met les courses puis elle installe Clément dans son siège pour enfant. Elle part retrouver Jean et se préparer pour la soirée.

100 Collections.com

006

L'AMITIÉ

L'amitié est un repère, une ligne d'horizon. Une lumière et une porte secrète vers les émotions... Elle nous aide par temps durs, elle fait rire, elle rassure. Elle ne juge pas, elle accepte, elle est là... Tantôt près, tantôt loin, mais toujours à distance de coeurs, notre amitié c'est du bonheur. Notre amitié, c'est une plante que je cultive avec grands soins, le trésor de mon petit jardin. Une pierre précieuse qui m'éclaire dans le noir, et quelqu'un qui connaît toute mon histoire... J'avance sur mon chemin, tu avances sur le tien. L'horizon devant nous est infini, tout comme notre amitié, merci.

On t'a dit : « La vie est courte et tu vas vite en faire le tour. »
On t'a dit : « Ne cours jamais l'achine, cours, ma fille, cours... »
Et dans cette course folle, t'as rencontré bien des malheurs.
Car pour courir à mille à l'heure il fallait perdre des valeurs.
Et puisqu'aimer, c'est prendre des risques, toi, tu braves les intempéries.
On t'a dit : « Faut bien te conduire », t'as répondu : « J'ai pas le permis ! »
Insolente, lentement, je te vois dévaler la pente,
Si tu te réveils présidentielle, tu ne seras que figurante.
Petite sauteuse, tu veux de l'or, des bolides et des sapas.
Toujours dehors, toi, la vie, tu ne la vis plus, tu la snapas.
Aurais-tu de toi, tu vois les stars mais tu ne vois plus le ciel.
Devenir celle dont tout le monde parle serait pour toi le casse du siècle.
Ence de gloire éphémère, de grande-née exceptionnelle,
Pour l'ascension professionnelle, l'honneur te sera optionnel.
Dans ce monde que tu convoites, pour être un homme faut être prince,
La place qu'il reste est si étroite que pour être star faut être mince.
Tu vois des fautes, tu vois du rêve mais ça n'est que de la poussière.
Certes ils brillent, mais ils n'irradient que ceux qui originent la lumière.
Tes modèles ne montent d'eux que lorsqu'ils baignent dans l'opulence,
Dans le silence ils taillent leurs veines et déboulent en ambulance.
À trop courir après du vent, t'en oublies le Ouis,
Au suivant ! Tu finiras par parler seule sur un divan.
Crois-moi, quel que tu gagnes, tu finis instable,
Toi qui aimais que tout le monde t'aime, tu finiras associable.
Contraire un « je t'aime, je m'engage », tu donneras ton cœur en gage,
Tu vieilliras jusqu'à ce qu'il porte avec une femme plus jeune en âge.
Car c'est l'histoire qui ne régit, toi qui voudrais changer de nom,
Commence par te changer toi-même et tu pourras changer le monde.

— MÉLANIE DITE DIAMÉ —

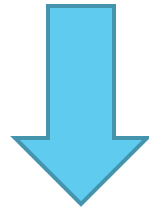
www.melanie-dite-diamé.com



La tokenization

Transformer un texte en vecteur de mots

« Le machine learning est un apprentissage par minimisation de l'erreur ! »



['Le', 'machine', 'learning', 'est', 'un', 'apprentissage', 'par',
'minimisation', 'de', 'l'erreur', '!']

La normalisation

Suppression de la ponctuation, enlever les apostrophes, mettre en minuscule.

['Le', 'machine', 'learning', 'est', 'un', 'apprentissage', 'par',
'minimisation', 'de', 'l'erreur', '!']

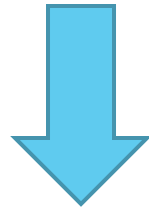


['le', 'machine', 'learning', 'est', 'un', 'apprentissage', 'par',
'minimisation', 'de', 'le', 'erreur']

Les stop words

Les mots de liaison qui n'apporte pas de sens mais du bruit

['Le', 'machine', 'learning', 'est', 'un', 'apprentissage', 'par',
'minimisation', 'de', 'l'erreur', '!']

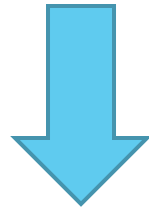


['machine', 'learning', 'apprentissage', 'minimisation',
'erreur']

Doublons

Plusieurs mots avec le même sens mais une orthographe différentes.

« Julien est plus grand que toi, mais tu es plus grande que Sophie.
Julien et Sophie sont plus grands que Marie. »

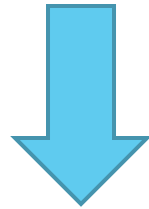


['grand', 'grande', 'grands']

Lemmatisation

Transformer un mot en son sens radical (lemme)

« Julien est plus grand que toi, mais tu es plus grande que Sophie.
Julien et Sophie sont plus grands que Marie. »



['grand', 'grand', 'grand']

Stemming

Supprimer la fin des mots

« Julien est plus grand que toi, mais tu es plus grande que Sophie.
Julien et Sophie sont plus grands que Marie. »



['gran', 'gran', 'gran']

Prétraitements essentiels

- ▶ Tokenization
- ▶ Normalisation
- ▶ Suppression des stop words
- ▶ Lemmatisation / Stemming

Collocation

Lier les mots qui apparaissent souvent ensemble

« Je lis souvent le New-York times »



« New-York » +

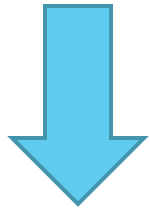
« times »

« New-York times »

Part of speech tagging

Ajouter des précision sur le mot

« J'aime chevaucher des chevaux »



Chevaux → Chevaux/NOM

Chevaucher → Chevaucher/VER:infi

Natural Language Toolkit (NLTK)

NLTK 3.5 documentation

[NEXT](#) | [MODULES](#) | [INDEX](#)

Natural Language Toolkit

NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to [over 50 corpora and lexical resources](#) such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active [discussion forum](#).

Thanks to a hands-on guide introducing programming fundamentals alongside topics in computational linguistics, plus comprehensive API documentation, NLTK is suitable for linguists, engineers, students, educators, researchers, and industry users alike. NLTK is available for Windows, Mac OS X, and Linux. Best of all, NLTK is a free, open source, community-driven project.

NLTK has been called “a wonderful tool for teaching, and working in, computational linguistics using Python,” and “an amazing library to play with natural language.”

[Natural Language Processing with Python](#) provides a practical introduction to programming for language processing. Written by the creators of NLTK, it guides the reader through the fundamentals of writing Python programs, working with corpora, categorizing text, analyzing linguistic structure, and more. The online version of the book has been updated for Python 3 and NLTK 3. (The original Python 2 version is still available at http://nltk.org/book_1ed.)


Some simple things you can do with NLTK

TABLE OF CONTENTS




[NLTK News](#)[Installing NLTK](#)[Installing NLTK Data](#)[Contribute to NLTK](#)[FAQ](#)[Wiki](#)[API](#)[HOWTO](#)


SEARCH


Google colab

 Bienvenue dans Colaboratory


FichierModifierAffichageInsérerExécutionOutilsAide

 Partager


 Sommaire


 Premiers pas

Science des données

 Machine learning








Autres ressources


 Exemples de machine learning

 Section

+ Code + Texte Copier sur Drive

Connecter Modification

 Qu'est-ce que Colaboratory ?

Colaboratory, souvent raccourci en "Colab", vous permet d'écrire et d'exécuter du code Python dans votre navigateur. Il offre les avantages suivants :

- Aucune configuration requise
- Accès gratuit aux GPU
- Partage facile

Que vous soyez **étudiant**, **data scientist** ou **chercheur en IA**, Colab peut vous simplifier la tâche. Regardez la [présentation de Colab](#) pour en savoir plus ou commencez tout de suite.

▼ Premiers pas

Le document que vous consultez n'est pas une page Web statique, mais un environnement interactif appelé **notebook Colab**, qui vous permet d'écrire et d'exécuter du code.

Voici par exemple une **cellule de code** avec un bref script en Python qui calcule une valeur, l'enregistre dans une variable et imprime le résultat :

```
[ ] seconds_in_a_day = 24 * 60 * 60
    seconds_in_a_day
```

86400

Pour exécuter le code dans la cellule ci-dessus, sélectionnez-le en cliquant dessus, puis cliquez sur le bouton de lecture à gauche du code ou utilisez le raccourci clavier Commande/Ctrl+Entrée. Pour modifier le code, cliquez simplement sur la cellule.

Notebook 1 - Se familiariser avec NLTK

Objectif :

- ▶ Utiliser Python pour travailler des données textuelles
- ▶ Apprendre à utiliser NLTK pour transformer un texte
- ▶ Implémenter un détecteur de langue avec les stopwords