

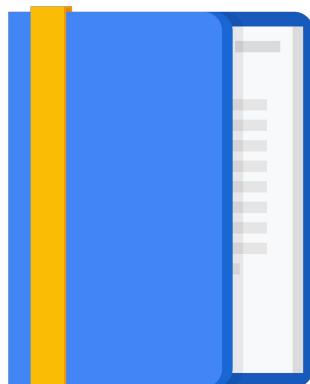


Prédefinir des API
de modèles pour les
données non structurées

Programme

Les données non structurées
sont difficiles

API de ML pour enrichir les données



Comment extraire des données utiles de cela ?



ou



 Google Cloud

<https://pixabay.com/fr/photos/newspaper-old-1925-daily-newspaper-2608394/>

Exemples de cas d'utilisation réels de données non structurées

Fixation du prix des véhicules grâce à un modèle d'image personnalisé



Acheminement des e-mails des clients via l'API de traitement du langage naturel (NLP)



Utiliser l'API Vision pour rechercher du texte dans des mèmes



Créer une expérience d'achat à l'aide de Dialogflow



Uniqlo (qui se prononce uni-clo) a conçu un chatbot d'achat à l'aide d'une interface utilisateur de Dialogflow. Dialogflow est une entreprise appartenant à Google, spécialisée dans la création d'interfaces basées sur le machine learning telles que les chatbots intelligents.

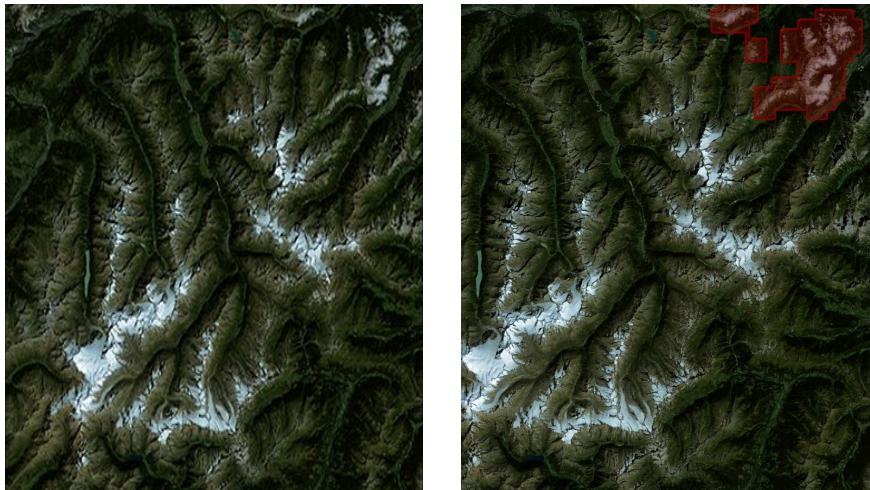
Étudions quelques autres cas d'utilisation du machine learning dans le monde de l'entreprise.

[Ressource]

Témoignages de clients lors de la conférence Google Next

https://youtu.be/BwWg_HVfsM?t=4m41s

Nuages ou montagnes enneigées ?



 Google Cloud

Si vous étiez le groupe Airbus, vous pourriez recourir au machine learning pour différencier des nuages d'un manteau neigeux. Si, comme moi, vous n'arrivez pas à les distinguer, les nuages sont surlignés en rouge dans la partie supérieure de l'image de droite.

<https://cloud.google.com/blog/products/gcp/google-cloud-machine-learning-now-open-to-all-with-new-professional-services-and-education-programs>

Vide ou plein ?

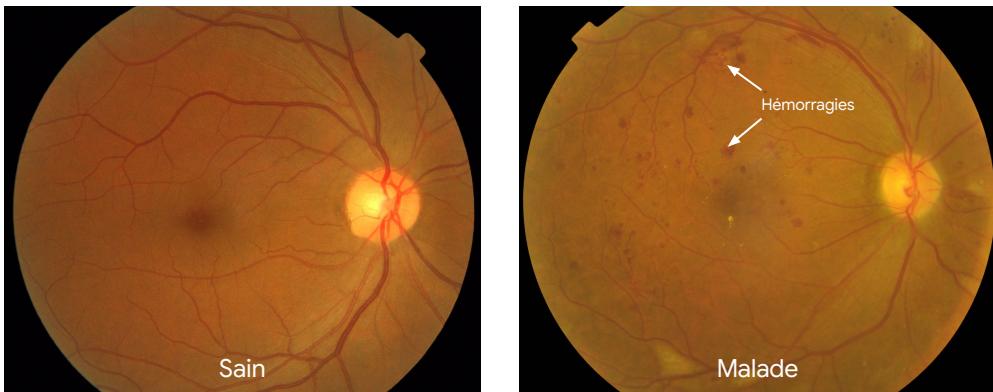


 Google Cloud

Si vous étiez une société de prévisions économiques, vous pourriez effectuer le suivi de la flotte mondiale de navires porte-conteneurs via l'imagerie satellitaire. En effet, la connaissance de la quantité de marchandises transportées pourrait vous aider à améliorer la précision de vos prévisions économiques avec plusieurs jours ou mois d'avance sur les chiffres officiels.

<https://pixabay.com/fr/container-ship-container-ship-port-2786842/>

Diagnostiquer la rétinopathie diabétique



Les images médicales sont propices aux innovations. Par exemple, vous pourriez assurer le diagnostic précoce de [problèmes de santé](#) tels qu'une rétinopathie diabétique, afin d'en faciliter le traitement et d'éviter tout risque de cécité.

<https://ai.googleblog.com/2016/11/deep-learning-for-detection-of-diabetic.html>

Pour les tâches ML courantes, envisagez d'utiliser des API préformées

Créer un modèle personnalisé



Cloud TPU



Compute Engine



Cloud Dataproc



Kubernetes Engine



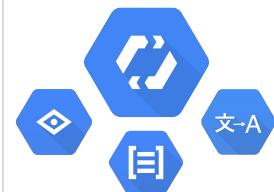
Plateforme Cloud AI



BigQuery ML

Créer un modèle personnalisé (sans code)

AutoML



Invoquer un modèle préformé



Cloud API de traduction



Cloud API Vision



Cloud API Speech



Cloud Video Intelligence API



API de prévention contre la perte de données



API de synthèse vocale Cloud



Cloud Natural API de langue



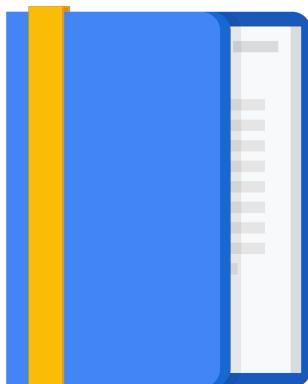
Dialogflow



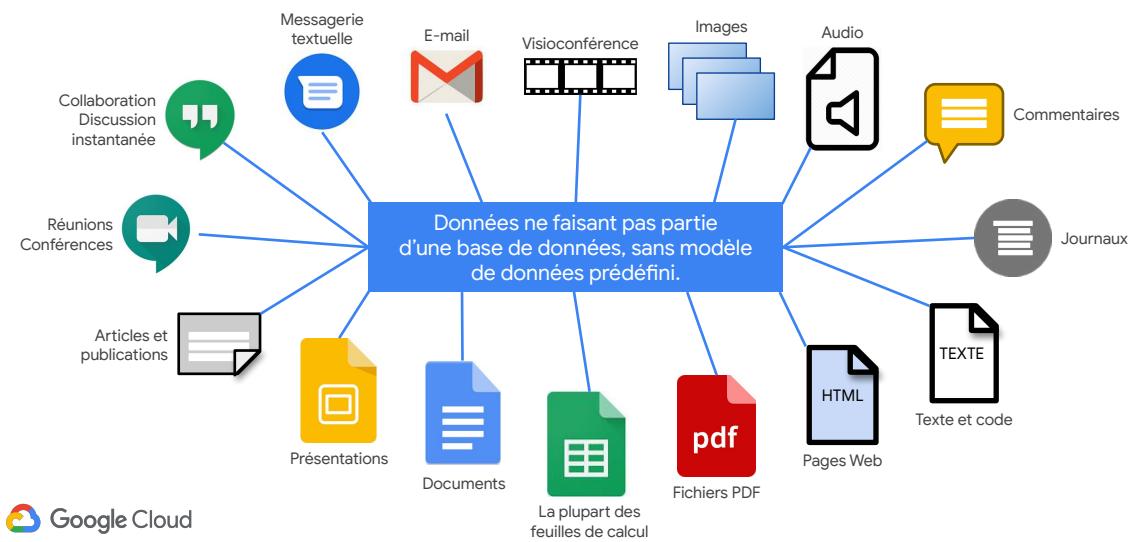
Programme

Les données non structurées
sont difficiles

API de ML pour enrichir les données

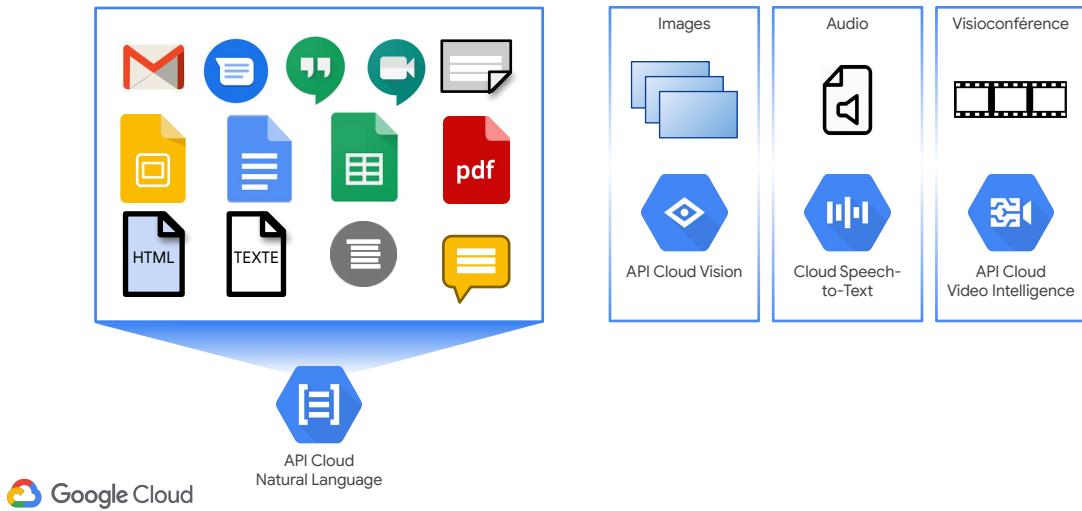


La plupart des données commerciales sont des données non structurées, et sont essentiellement du texte



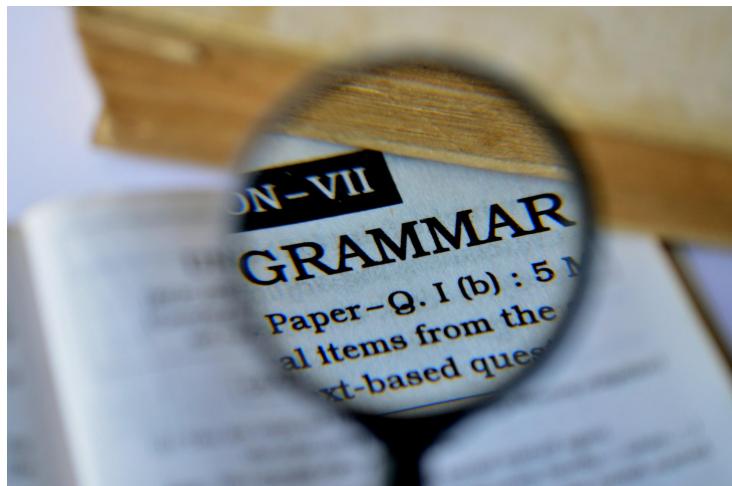
Que sont les données structurées ? Des données d'une base de données, dans un fichier JSON et dans certaines feuilles de calcul

Vous traitez des données non structurées en les identifiant avec l'IA



Nous allons nous concentrer sur l'API Cloud Natural Language.
Vous pouvez y accéder de plusieurs manières, mais si vous avez déjà des données que vous migrez de Hadoop vers Cloud Dataproc, vous pouvez les rendre disponibles pour analyse en utilisant l'API Cloud Natural Language.

L'analyse syntaxique fournit des informations linguistiques



 Google Cloud

L'analyse syntaxique décompose le texte en une série de jetons (généralement, des mots et des phrases) et fournit des informations sur la structure interne du jeton (morphologie) et son rôle dans la phrase (syntaxe). Elle peut étiqueter un jeton comme nom ou verbe, singulier ou pluriel, à la première ou la deuxième personne, au masculin, au féminin ou neutre, et fournit des informations grammaticales telles que la casse, le temps, l'humeur et la voix.

Au moment de la rédaction du présent document, elle prend en charge 10 langues. Voir la documentation en ligne pour la prise en charge linguistique actuelle.

<https://pixabay.com/photos/grammar-magnifier-magnifying-glass-389907/>

L'analyse par entité reconnaît de nombreux termes utiles



Emplacements et repères



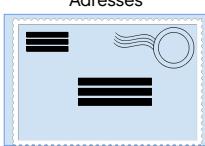
Les personnes



Œuvres d'art et travaux artistiques



Produits de grande consommation



Adresses



Nombres

Prix

Dates

Tout nombre

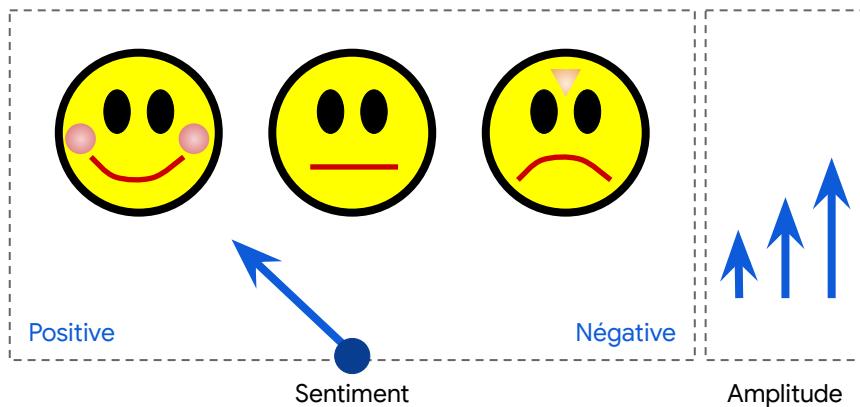


les noms propres tels que les personnalités publiques, les lieux importants

Personne, lieu, organisation, événement, œuvre d'art, produit de consommation, numéro de téléphone, adresse, date, n'importe quel numéro, prix

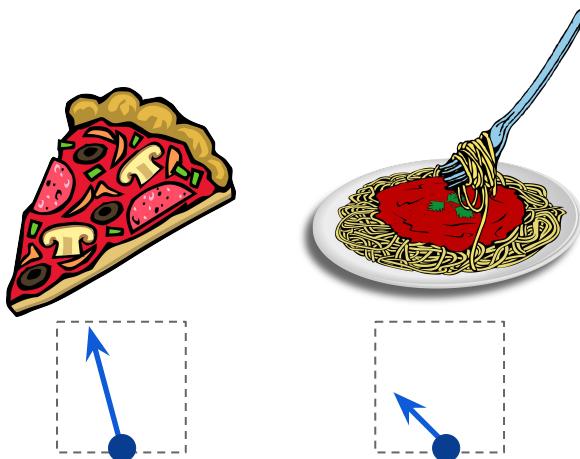
- <https://pixabay.com/fr/illustrations/males-3d-model-isolated-3d-model-2373729/>
- <https://pixabay.com/fr/photos/art-creative-hand-painting-paper-1868562/>
- <https://pixabay.com/fr/photos/eiffel-tower-france-paris-landscape-975004/>
- <https://pixabay.com/fr/vectors/toast-toaster-breakfast-kitchen-154616/>
- <https://pixabay.com/fr/vectors/postage-stamp-postal-stamp-1479775/>
- <https://pixabay.com/fr/vectors/iphone-cell-phone-phone-160307/>
- <https://pixabay.com/fr/vectors/tag-luggage-label-blank-price-309129/>

L'analyse de sentiment identifie les émotions dans le texte



L'analyse de sentiments identifie l'opinion émotionnelle dans l'attitude d'un écrivain. Elle se présentait sous la forme d'une note numérique et d'une magnitude pour l'intensité du sentiment. Elle n'identifie pas d'émotions spécifiques, mais les regroupe en émotions généralement positives ou négatives ou neutres. Ainsi, "triste" et "fâché" sont tous deux négatifs, et "drôle" et "heureux" sont tous deux positifs. Comme il s'agit de valeurs continues, vous devez définir vos propres seuils, qui sont adaptés à votre application. Par exemple, une magnitude inférieure à 10 % n'est peut-être pas un sentiment assez fort, ou une magnitude comprise entre -1,0 et +1,0 n'exprime pas clairement une émotion.

Analyse combinée des entités et des sentiments



L'analyse de sentiments des entités combine à la fois l'analyse des entités et l'analyse des sentiments et tente de déterminer le sentiment (positif ou négatif) exprimé à propos des entités dans le texte. Le sentiment d'entité est représenté par des notes numériques et de magnitude et est déterminé pour chaque mention d'une entité. Ces notes sont ensuite agrégées en une note globale de sentiment et de magnitude pour une entité.

<https://pixabay.com/fr/vectors/pizza-food-slice-cheese-mushroom-23477/>
<https://pixabay.com/fr/vectors/essen-food-noodle-plate-spaghetti-1293911/>

Classification du contenu



Sports



Aliments



Littérature

Plus de 600 catégories



Au moment de la rédaction du présent document, le contenu est classé en 620 catégories.

<https://cloud.google.com/fr/natural-language/docs/categories>

<https://pixabay.com/fr/photos/salad-fruits-berries-healthy-2756467/>

<https://pixabay.com/fr/illustrations/games-sport-sports-competition-70833/>

<https://pixabay.com/fr/photos/book-read-old-literature-pages-1659717/>



Atelier

Utilisation de l'API Natural Language pour classifier du texte non structuré

Objectifs

- Créer une requête de l'API Natural Language
- Invoquer l'API via curl
- Utiliser la fonctionnalité de classification du texte de l'API NL
- Utiliser la classification de texte pour comprendre un ensemble de données sur des articles d'actualité

Utiliser l'API Natural Language pour classifier du texte non structuré

<https://googlestaging.qwiklabs.com/labs/32334/>

Résumé du module

- Les données non structurées sont omniprésentes... et très difficiles à exploiter
- Heureusement, GCP dispose de nombreuses API pour rendre les données non structurées utiles

