



Rapport Technique sur l'analyse de sentiment

Sommaire :

Qu'est-ce que l'analyse de sentiment ? ..... page 2

Comment fonctionne l'analyse des sentiments ? ... page 3

Pourquoi analyser les sentiments ?..... page 3

Quand utiliser l'analyse des sentiments ? ..... page 3

Cahier des charges ..... page 4

La solution proposée ..... page 6

Next steps & pistes d'amélioration ..... page 7

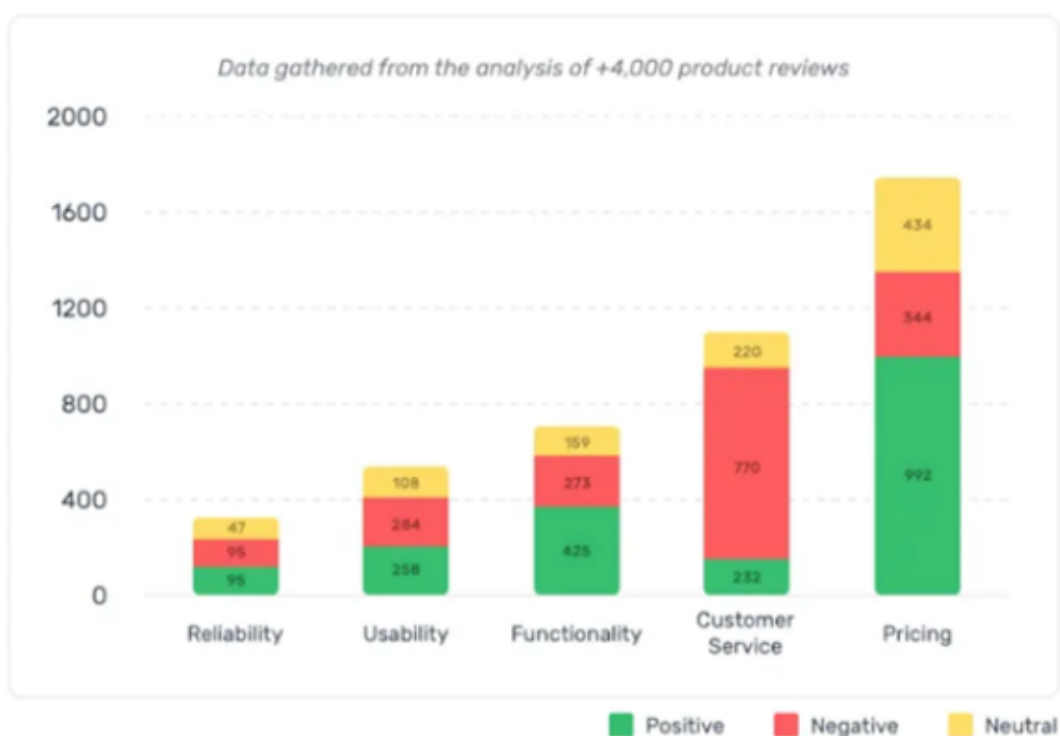
Planning ..... page 9

### Qu'est-ce que l'analyse de sentiment ?

L'**analyse de sentiment** est le mécanisme qui permet de détecter l'opinion, le jugement et l'émotion qui se cache derrière le langage naturel.

Lorsqu'une personne laisse un avis en ligne, rédige un commentaire sur un produit ou répond à une étude de marché, son appréciation est forcément teintée par des sentiments, qu'ils soient positifs, négatifs ou neutres.

En effet, lorsqu'un consommateur exprime son avis à travers un texte rédigé, celui-ci peut être transformé en données catégorielles (positif ou négatif). Une fois cette catégorisation effectuée, l'entreprise peut générer des rapports pour obtenir une vue d'ensemble des ressentis et opinions de ses clients.



Source : <https://datascience.eu/fr/>

## Comment fonctionne l'analyse des sentiments ?

Il existe trois **méthodologies de base pour analyser les sentiments**. Chacune répond à un objectif différent :

- .. **Détection de l'émotion** : analyser le texte et trier les données d'opinion à partir de data objectives (transformer « l'hôtel est situé dans le centre » en « positif » par exemple)
- .. **Catégorisation du sentiment** : détecter si l'appréciation est positive, négative ou neutre
- .. **Polarité** : dans le cas où le texte contient plusieurs opinions ou des appréciations mitigées, être en mesure d'évaluer cette polarité évite que le commentaire soit classé dans la mauvaise catégorie.

## Pourquoi analyser les sentiments ?

L'**analyse de sentiment est très utile dans le cas où l'on est face à une grande quantité de données**. Les **outils de sentiment analysis** permettent d'établir des rapports précis, en très peu de temps.

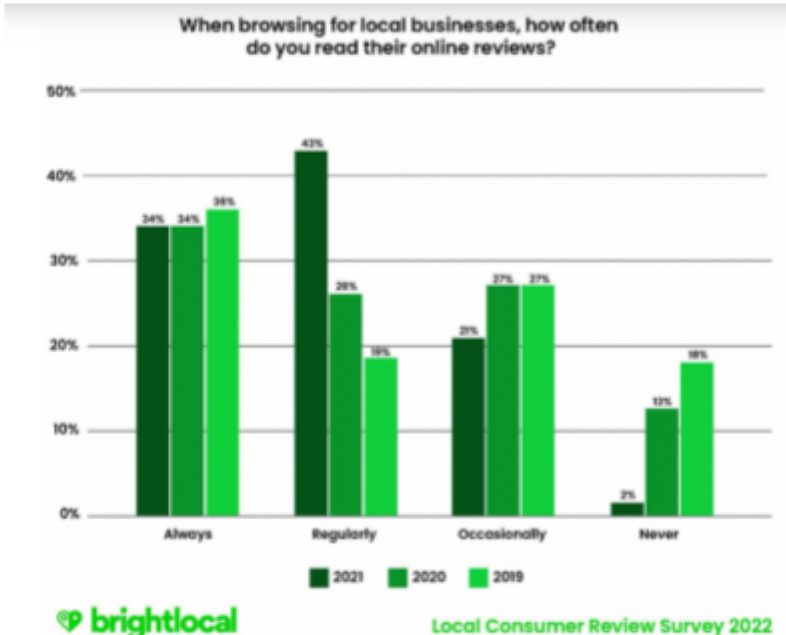
Grâce aux scores obtenus, les entreprises sont en mesure de comprendre **quel type d'émotion leur marque suscite** auprès de leurs clients, et à partir de ces informations, l'entreprise peut mettre en place un plan d'action pour améliorer l'expérience des clients qui ont laissé des commentaires négatifs.

## Quand utiliser l'analyse des sentiments ?

La **sentiment analysis** peut aider les entreprises à savoir comment leurs produits sont perçus à un instant T. Voici quelques exemples d'**applications où l'analyse de sentiments** prend tout son sens :

- .. **Surveiller des médias sociaux**
- .. **Analyser des réponses d'une enquête**
- .. **Analyser les feedbacks des collaborateurs**
- .. **Identifier et visualiser des tendances** géographiques ou sociodémographiques relatives aux émotions de la clientèle
- .. **Faire une recherche concurrentielle** en comparant votre image et le ressenti des consommateurs par rapport à vos concurrents

## Cahier des charges :

<p>Contexte :</p>	<p>Les achats se faisant de plus en plus sur Internet, et avec le contexte du Covid accélérant ce processus, nous observons que le nombre de personnes lisant les commentaires en ligne ont augmenté :</p> <div><p>When browsing for local businesses, how often do you read their online reviews?</p><table><thead><tr><th>Frequency</th><th>2021</th><th>2020</th><th>2019</th></tr></thead><tbody><tr><td>Always</td><td>34%</td><td>34%</td><td>36%</td></tr><tr><td>Regularly</td><td>43%</td><td>26%</td><td>18%</td></tr><tr><td>Occasionally</td><td>21%</td><td>27%</td><td>27%</td></tr><tr><td>Never</td><td>2%</td><td>13%</td><td>19%</td></tr></tbody></table><p>brightlocal Local Consumer Review Survey 2022</p></div>	Frequency	2021	2020	2019	Always	34%	34%	36%	Regularly	43%	26%	18%	Occasionally	21%	27%	27%	Never	2%	13%	19%
Frequency	2021	2020	2019																		
Always	34%	34%	36%																		
Regularly	43%	26%	18%																		
Occasionally	21%	27%	27%																		
Never	2%	13%	19%																		
	<p>Les commentaires sont de plus en plus lues. Il est donc important d'analyser les commentaires négatifs pour pouvoir apporter des actions correctives. La société Zalanda veut pouvoir analyser l'avis de ces clients afin de mettre en place des stratégies concernant les produits vendus qui pourrait s'avérer défectueux ou non plaisant à ces consommateurs.</p>																				
<p>Fonctionnalités:</p>	<ul style="list-style-type: none"><li>- Création d'un serveur accueillant chaque commentaire.</li><li>- Commentaires analysés, et classifiés selon leur importance et rapidité d'être examiné.</li><li>- Retour automatique des commentaires aux services concernés.</li></ul>																				
<p>Livrables :</p>	<ul style="list-style-type: none"><li>- Serveur pouvant accueillir 5000 commentaires (se nettoie et accepte d'autres commentaires lorsque le modèle a fini de les classifier).</li><li>- Modèle (Solution Intelligence Artificielle) pouvant traiter les commentaires.</li><li>- Interface utilisateur. Permet de tester le modèle et montre le fonctionnement du modèle.</li></ul>																				

<b>Spécificité du serveur :</b>	<ul style="list-style-type: none"> <li>- Capacité maximum de 5000 commentaires.</li> <li>- Quand un commentaire a été traité/étudié, le commentaire est enlevé du serveur et un nouveau plus récent est ajouté.</li> </ul>
<b>Spécificité du modèle :</b>	<ul style="list-style-type: none"> <li>- Permet de savoir si le commentaire est positif ou négatif. Un score est obtenu pour l'intensité du sentiment.</li> <li>- Il est possible de créer une échelle de priorité en fonction de l'intensité des messages</li> </ul>

### La solution proposée :

Afin de pouvoir traiter chacun des commentaires et construire un modèle nous avons dû transformer ces valeurs textuelles en valeurs numériques.

Grâce à des techniques de **tokenisation** de **padding** et de **vectorisation**, nos données ont pu être exploitées par un modèle.

Plusieurs modèles ont été ensuite explorés, en **machine learning** et en **deep learning** nous garderons après analyse celui présentant les meilleurs résultats.

Des premières itérations en machine learning, concernant des modèles en classification, ont donné des résultats satisfaisants.

Nous avons retenu un modèle de régression logistique avec un score en recall de 0.86.

Notre modèle de deep learning a donné un score inférieur à celui de notre modèle de machine learning, cet axe sera donc écarté.

Après entraînement du modèle retenu, il est maintenant possible de prédire la satisfaction ou le mécontentement d'un utilisateur grâce aux commentaires laissés sur le site. Un taux de confiance sera également donné. **Si le commentaire est négatif** à plus de 50%, cela montrera une plus grande gravité dans le mécontentement du consommateur. Si ce taux reste autour de 0 cela montre un commentaire complètement neutre.

Pour utiliser l'application, il suffit d'écrire votre commentaire dans l'encadré prévu à cet effet. Puis de cliquer sur le bouton en dessous.

Type your review here !

Love the top design, not so much the fabric. i should've read the top was 100% polyester. returning an

Am I giving it a good review, or a bad one ?

IT may be a bad review

Not Recommended with %27.71

Vous obtiendrez comme réponse si le commentaire est plutôt positif ou négatif. Dans l'exemple du dessus, Le commentaire ne recommande pas l'article avec une valeur de 27%

Type your review here !

I ordered this dress in 0p since i am 5ft. it fits great its not too short just above the knee.

Am I giving it a good review, or a bad one ?

It seems to be a good review !

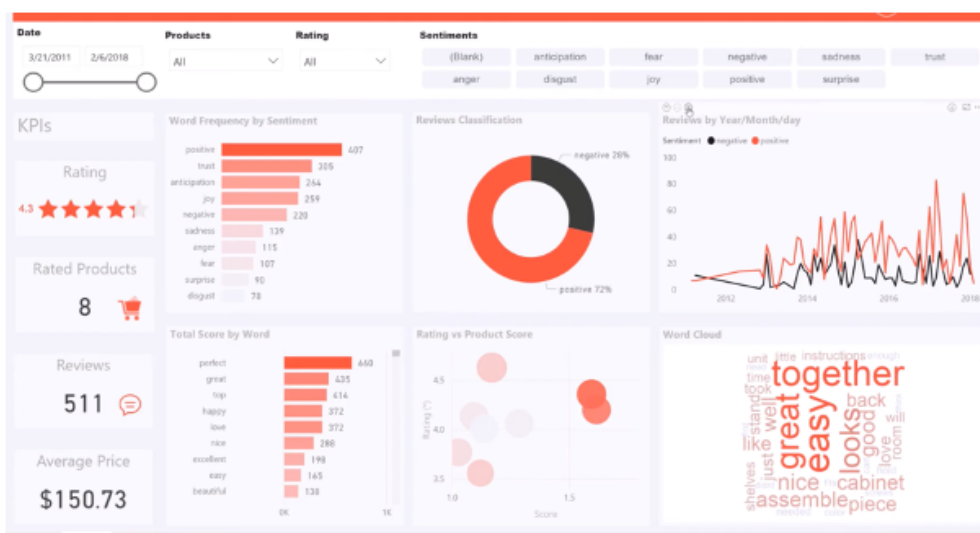
Recommended with %78.77

Dans cet exemple, le commentaire est recommandé à un taux de 79 %

## Next steps & pistes d'amélioration

Comme demandé, nous avons exécuté notre mission dans le respect du cahier des charges et des consignes communiquées par votre société **Zalanda**. Cependant nous avons identifié quelques axes d'amélioration qui permettraient à votre entreprise de faire évoluer ce projet pour le rendre encore plus complet et qu'il puisse répondre totalement à votre problématique.

- **EDA (exploratory data analysis)** : création d'un Dashboard Sentiment Analysis, avec une analyse plus détaillée des profils de vos clients. Ce qui pourrait vous aider à cibler davantage le profil des clients « pas satisfaits » pour pouvoir cibler votre plan d'action. Exemple de dashboard en sentiment analysis :







## Planning de la mission

	Mercredi 12/01	Jeudi 13/01	Vendredi 14/01	Mercredi 19/01	Jeudi 20/01	Vendredi 21/01
EDA						
Préprocessing + Text mining						
Models machine learning						
Models Deep learning						
Déploiement streamlit						
Présentation Canva						