

Mise en place d'un système de notation par aspects pour les articles de E-commerce en se basant sur les avis en français des consommateurs

Mathias Chastan¹, Rania Zaatour¹, Cheik Ouedraogo¹, Karim Baati¹

Teolia Consulting

8 rue de l'Hôtel de Ville, 92200 Neuilly-sur-Seine, France

mathias.chastan@teolia.fr, rania.zaatour@teolia.fr, cheik.ouedraogo@teolia.fr, karim.baati@teolia.fr

2024

Résumé

Un très grand volume de données d'avis sur les articles consommés sont mis en ligne chaque jour. Ce sont des informations sur des produits ainsi que sur leurs différentes caractéristiques spécifiques. A l'encontre de la majorité des travaux de recherche qui se sont penchés sur l'évaluation globale des produits à partir des avis collectés, nous proposons, dans cet article, un système de notation par aspects. Le système suggéré s'appuie sur une analyse poussée d'avis rédigés en français, et permet d'extraire les différentes propriétés des articles évalués ainsi que l'opinion utilisateur afférente à chaque propriété. La validation technique de ce système est opérée sur des avis relatifs à des articles issus d'un site de E-commerce français.

Mots-clés

Traitement automatique du langage naturel, Analyse sentimentale par aspects, Avis produits, Notation, Extraction d'aspects

Abstract

A very large volume of review data on consumed articles is put online every day. These data provide information about products as well as their various specific characteristics. In contrast to the majority of research works that have focused on the global evaluation of products based on the collected reviews, we propose, in this paper, an aspect-based scoring system. This latter is based on a thorough analysis of reviews written in French. It allows to extract the different properties of assessed items as well as the user opinion related to each of these properties. The technical validation of this system is performed on reviews related to the products suggested by a French E-commerce website.

Keywords

Natural language processing, Aspect based sentiment analysis, Product reviews, Scoring, Aspect mining

1 Introduction

Avec l'expansion des plateformes de e-commerce, le shopping en ligne est devenu une méthode incontournable de consommation. De très grands volumes de données générées par les utilisateurs sont mis en ligne chaque jour. Il s'agit d'informations sur les préférences et les opinions des utilisateurs au regard de différents produits, souvent dirigées vers des caractéristiques spécifiques de ces articles. Bien que ces textes regorgent d'informations, ils sont souvent non structurés et demandent un effort d'analyse de la part des utilisateurs ou des enseignes, afin de parvenir à extraire l'information dont chacun a besoin.

Le domaine de l'intelligence artificielle [3, 4] a vu émerger ces derniers temps de nouvelles techniques de traitement du langage naturel (en anglais *Natural Language Processing* (NLP)) qui permettent d'assurer l'analyse sentimentale du langage humain. En d'autres termes, ces techniques permettent d'extraire le ressenti global d'un utilisateur sur un sujet à partir d'un message oral ou écrit qu'il a produit.

Dans le cas d'avis soumis en ligne, ces méthodes servaient, dans un premier temps, à assigner un score global à chaque produit évalué. Néanmoins, avec le temps, il s'est avéré que le score global est réducteur, étant donné qu'il ne prend pas en compte les opinions sur les différentes propriétés de l'article évalué, et que ces opinions peuvent être d'intensité et de polarité différentes. En effet, un client peut trouver que la qualité d'un produit est satisfaisante mais que son prix est trop élevé. Dans ce cadre, des techniques d'analyse sentimentale plus poussées permettent de répondre à cette problématique en effectuant une analyse par aspects. Le travail présenté dans cet article se joint à ces techniques et propose un système d'analyse sentimentale par aspects qui s'assoit sur les avis en français soumis en ligne. Le système suggéré fait usage de différents modèles de NLP pré-entraînés ainsi qu'un algorithme de règles pour extraire et noter automatiquement la satisfaction client au regard des différentes propriétés d'un article.

La suite de cet article est scindée en 5 sections. La section 2 porte sur une synthèse des travaux antérieurs ayant traité la

même problématique. Ensuite, dans la section 3, nous présentons notre système tout en détaillant les modèles ainsi que les données utilisés. Quant à la section 4, elle est dédiée à la discussion et aux applications pratiques de notre système. Finalement, la section 5 conclut ce travail et en dégage des perspectives pour des travaux futurs.

2 Etat de l'art

Une première remarque que l'on peut soulever en évoquant les travaux antérieurs en lien avec l'analyse sentimentale des avis en ligne, est que bon nombre de travaux ont visé l'attribution de scores globaux pour les produits évalués. [7, 18]. Par opposition à ces études, le travail suggéré dans cet article fait appel à l'analyse sentimentale en vue d'assigner des scores par aspects.

Un aspect est défini comme toute caractéristique ou propriété d'un produit pouvant influencer le choix d'un client potentiel (e.g., qualité, confort, couleur, prix) [22]. Un sentiment est défini comme une connaissance ou une impression qu'un client a sur un produit.

L'analyse sentimentale a pour objectif de déterminer l'opinion ou l'impression qui se trouve derrière le langage naturel. En se fondant sur cette définition, nous pouvons définir l'analyse sentimentale par aspects comme le processus qui permet de déterminer l'opinion ou l'impression associée à chaque caractéristique d'un produit. Elle est généralement fondée sur trois étapes : (1) l'extraction des aspects et des mots sentiments, (2) le regroupement des aspects, et (3) la notation des mots sentiments.

1. L'extraction des aspects et des dépendances consiste à extraire les caractéristiques mentionnées par les consommateurs dans leurs avis en ligne et les sentiments qui y sont associés. Le mot sentiment est le mot qui décrit l'opinion du consommateur sur une caractéristique. Ces aspects et mots sentiments sont cachés dans les phrases et pour les retrouver un *Part Of Speech tagger (PoS tagger)* est utilisé. Un *PoS tagger* permet d'extraire la nature d'un mot dans une phrase. Pour ce faire, quatre bibliothèques sont principalement utilisées, à savoir « NLTK »¹, « Stanford Core », « NLP suite », et « Spacy ». Plus de notion sur ces bibliothèques sont disponibles dans [1].
2. Le regroupement des aspects permet de regrouper sous un seul groupe, les mots désignant le même aspect. Ce regroupement prend en compte la sémantique de l'aspect dans l'avis client. Cette étape requiert l'utilisation d'un modèle de plongement lexical tel que « Word2vec » [9].
3. La dernière étape vise à conférer un score de sentiment aux mots sentiments. Pour cela, il faut tout d'abord trouver la définition du mot par rapport au contexte. Elle est réalisée à travers une analyse sentimentale dont l'objectif est de juger le sentiment présent dans un écrit.

Il existe de nombreuses manières qui permettent d'effectuer une analyse sentimentale. Par exemple, dans [6], Banerjee et Pedersen font appel à l'algorithme « Lesk » et à ses variantes pour désambiguïser le sens des mots, pour ensuite déterminer le score de sentiment à partir de « SentiWordNet »², un jeu de données qui associe un score de sentiment à chaque définition d'un mot [5].

Il est également possible d'utiliser d'autres modèles d'analyse sentimentale tels que celui de Google pour calculer le score et la polarité d'un sentiment. Le modèle Google est avantageux puisqu'il détecte le contexte de la phrase de manière automatique et existe en langue française. C'est pourquoi nous l'avons considéré pour la mise en place de notre système.

Plusieurs articles décrivant des méthodes d'analyse sentimentale par aspects ont suivi les étapes précitées. A titre d'exemple, dans [22], les auteurs ont fait usage de l'analyse sentimentale afin d'analyser des avis sur des téléphones mobiles, dans [16], la méthode suggérée recommande des restaurants en fonction de l'analyse sentimentale des avis par aspects, et les travaux présentés dans [15] et [20] traitent des données de e-commerce. Les auteurs de [21] ont élaboré une méthode qui intègre la dépendance entre la cible (que nous nommons dans notre article "aspect") et la catégorie d'aspect (que nous appelons famille d'aspect ou classe d'aspect) dans la classification des sentiments. A l'encontre de la méthode suggérée dans [21], l'approche adoptée par notre article se limite à la classification des sentiments basée sur la classe d'aspect et le mot-sentiment.

D'un autre côté, la méthode exposée dans le cadre de nos travaux se positionne dans le contexte de recherches antérieures telles que résumées dans les articles [8] [13]. Ces articles examinent diverses techniques utilisées pour la préparation des données d'entrée [8] et l'extraction des aspects [13] ainsi que la classification des sentiments et l'évaluation des résultats. Pour la préparation des données d'entrée, plusieurs éléments peuvent être extraits tels que le contexte, l'aspect, la catégorie d'aspect et le mot-sentiment. Les techniques employées pour cela incluent le Part-of-speech tagging et le dependency parsing [8]. Concernant la classification des sentiments, nous avons opté pour le modèle de machine learning de Google dans le cadre de ce travail, mais d'autres modèles auraient également pu être envisagés. Des approches alternatives, telles que les modèles basés sur les connaissances ou hybrides [8], auraient pu donner des résultats comparables. L'article [8] suggère également différentes mesures de performance qui peuvent être utilisées pour évaluer les résultats. Dans le cadre de notre article, nous avons choisi d'utiliser l'erreur quadratique moyenne, bien que d'autres mesures auraient pu être prises en considération.

1. Natural Language Toolkit : <https://www.nltk.org/>

2. <https://github.com/aesuli/SentiWordNet/blob/master/papers/LREC06.pdf>

3 Présentation du système

En dépit du fait que plusieurs travaux antérieurs aient proposé des méthodes performantes pour l'analyse sentimentale par aspects, nous avons été contraints de ne pas considérer bon nombre d'entre elles vu qu'elles sont pré-entraînées sur des textes en anglais.

En effet, pour faire usage de ces méthodes, il faut passer par une étape onéreuse de traduction. Par ailleurs, un retour en arrière consistant à récupérer les résultats en français ne sera pas possible avec de telles méthodes.

Ainsi, le système que nous proposons est fondé sur l'utilisation de différents algorithmes gratuits ainsi qu'un algorithme payant de Google pour la notation des sentiments. La figure 1 présente le modèle conceptuel du système.

Dans les sous sections suivantes, nous détaillons les données utilisées ainsi que les différents modèles composant notre système, à savoir le modèle d'extraction des aspects et des mots sentiments, le modèle de regroupement des aspects, et le modèle de calcul du score associé aux aspects.

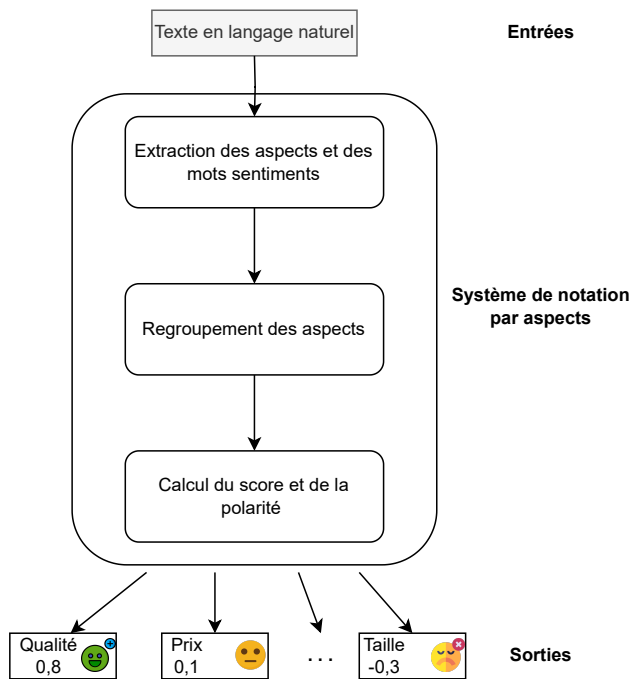


FIGURE 1 – MODÈLE CONCEPTUEL DU SYSTÈME.

3.1 Présentation du corpus de données utilisé

Nous présentons dans cette section les corpus de données existants et celui que nous avons utilisé pour la mise en place de notre système.

3.1.1 Corpus de données existant

Comme présenté dans la section 2, des solutions ont été proposées pour la notation des avis clients en s'appuyant sur des corpus données déjà annotés. Nous pouvons citer le corpus de données « *Amazon Reviews : Unlocked Mobile* » proposé par [17] et repris dans les travaux de [22] pour noter les avis clients sur les téléphones. Ce corpus contient plus

de 400 milles avis clients en anglais, suivant une annotation binaire où 1 correspond à un avis positif et 0 à un avis négatif. Nous pouvons également citer le corpus de données proposé par [5] où l'auteur a manuellement noté les avis clients sur les restaurants suivant un certain nombre d'aspects (e.g., la nourriture, l'environnement, le service). Malheureusement ces données ne sont pas en libre accès.

A notre connaissance, la plupart des corpus de données annotés, comme ceux que nous venons de présenter, permettant d'entraîner des modèles de notation automatique des avis clients qui sont soit en anglais, soit pas en libre accès. D'où l'intérêt de créer notre propre corpus de données en français qui nous servira également pour évaluer les performances de notre système.

3.1.2 Corpus de données utilisé

Les données utilisées pour développer et évaluer notre système sont les avis clients d'un échantillon de 81 produits d'un site de E-commerce français. Les produits sélectionnés pour notre cas d'usage sont des robes. Les 113 premiers avis ont été étiquetés à la main pour pouvoir tester et évaluer nos modèles. Pour ce faire, nous avons lu chaque avis, avons identifié les aspects importants, et avons ensuite attribué un score dans l'intervalle $[-1, 1]$ à chaque aspect identifié en fonction de sa perception de l'expérience client.

Par exemple, sur l'avis client « super jolie mais taille court. », nous avons identifié deux aspects : la « robe » avec le mot sentiment « jolie », et la « taille » avec le mot sentiment « court ». Il leur a ensuite attribué les scores respectifs de 0,9 et de $-0,3$.

Au terme de cette phase d'étiquetage de notre jeu de données, nous avons identifié les 6 principales caractéristiques suivantes, 160 fois :

- robe
- taille
- coupe
- qualité
- tissu
- matière

Quant au modèle de vectorisation, il a été entraîné avec des corpus de données composés de texte qui représentent des actualités en ligne et qui sont en libre accès sur ³.

3.2 Modèle d'extraction des aspects et des mots sentiments

Pour effectuer l'analyse sentimentale par aspects, nous avons commencé par extraire les aspects et les mots sentiments des avis. Pour rappel, un aspect est une caractéristique du produit cible qui est dans notre cas d'étude une robe. Des caractéristiques possibles de ce produit sont la qualité, la coupe, le tissu, le prix et la couleur. Pour la mise en place du modèle, ces caractéristiques doivent être définies en amont.

D'un autre côté, les mots sentiments sont les termes qui portent sur le jugement subjectif du créateur de l'avis (e.g., belle, mauvaise, satisfait, etc.). La positivité ou la négativité

3. <https://wortschatz.uni-leipzig.de/en/download/French>

du sentiment associé à ce type de terme peut être généralement notée.

Le modèle d'extraction des aspects repose sur l'analyse des dépendances syntaxique des mots d'un avis et sur un ensemble de règles métiers sous la forme d'un algorithme.

3.2.1 Analyse de dépendances syntaxiques et étiquetage des avis

L'étiquetage morpho-syntaxique a pour objectif d'attribuer une étiquette grammaticale à chaque mot d'une phrase. L'étiquette grammaticale, aussi appelée *Part of Speech* (PoS), est le rôle du mot dans la phrase. Par ailleurs, l'analyse de dépendances permet de déterminer les relations syntaxiques entre les mots.

La revue [1] détaille différentes solutions existantes d'étiquetage morpho-syntaxique et d'analyse de dépendances. Ce travail a mis en avant la pertinence de la bibliothèque Python gratuite « Spacy »⁴ pour une meilleure précision globale. De ce fait, nous avons opté pour cette bibliothèque afin d'aborder la phase d'étiquetage et d'analyse des dépendances syntaxiques.

Spacy utilise des règles spécifiques à chaque langue pour déterminer le rôle de chaque mot. De plus, un système de transition non monotonique [12] basé sur l'analyse de dépendances pseudo-projectives [14] est invoqué afin de déterminer les liens syntaxiques.

3.2.2 Algorithme d'extraction

L'algorithme d'extraction permet d'extraire les aspects et les mots sentiments dans un avis client. Pour ce faire, un groupe de règles fondé sur du *PoS tagging* et les dépendances entre les différents mots a été défini. Notre attention s'est portée sur les rôles syntaxique des aspects et les mots sentiments dans les avis clients et les dépendances entre ces derniers.

Le tableau 1 présente quelques exemples de rôles syntaxiques considérés lors de la mise en place de notre algorithme. Cette logique nous permet de considérer un avis comme une suite de *PoS* et de dépendances.

TABLE 1 – EXEMPLE DE DÉFINITION D'ACRONYMES D'ÉTIQUETAGES ET DE DÉPENDANCES SYNTAXIQUES

Acronyme	Signification
ADV	Adverbe
ADJ	Adjectif
NOUN	Nom
CCONJ	conjonction de coordination
amod	Modificateur d'adjectif/ d'un nom
root	Racine de la phrase

Pour illustrer le processus de mise en place de nos règles d'extraction, prenons l'étiquetage morpho-syntaxique associé à l'avis client suivant : « super jolie mais taille court ». Le *PoS tagging* de cet avis est le suivant :

```

super (ADV) : advmod  —>  jolie (ADJ)
jolie (ADJ) : root    —>  jolie (ADJ)
mais (CCONJ) : cc     —>  taille (NOUN)
taille (NOUN) : conj  —>  jolie (ADJ)
court (ADJ) : amod    —>  taille (NOUN)

```

Pour cet exemple nous pouvons remarquer que le *PoS tagging* associé au mot *jolie* est (ADJ) et sa dépendance est « root ». Le tag « root » indique que l'adjectif est directement en racine de la phrase et fait donc référence au produit (robe). Une de nos règles métiers mises en place va extraire *robe* comme aspect (grâce au « root ») et « *jolie* » comme son mot sentiment. Dans la suite de cet avis le mot *court* dont le tag est (ADJ) a une dépendance de type « amod » avec *taille* dont le tag est (NOUN). Cette dépendance de tag conduit également à une autre règle d'extraction.

En analysant les différents avis clients présents dans notre corpus de données, nous avons établi au total 41 règles métiers qui permettent d'extraire les aspects et les mots sentiments. Ces règles ont été formalisées sous la forme d'un algorithme dont un aperçu est présenté par l'algorithme 1, où w^i désigne un mot présent dans l'avis client. La première condition de l'algorithme 1 correspond à l'extraction de l'aspect *robe* et du mot sentiment *jolie* et la deuxième condition permet l'extraction de l'aspect *taille* et du mot sentiment *court*.

Algorithme 1 : Algorithme d'extraction des aspects et des sentiments

Entrées : les mots w^i des avis clients

Sorties : aspect, sentiment

initialisation

Pour chaque w^i faire ;

si $PoS(w^i) = ADJ$ **alors**

si $dep(w^i) = root$ **alors**

$aspect \leftarrow robe;$

$sentiment \leftarrow w^i;$

finsi

sinon

si $dep(w) = amod$ and $PoS(w+1) = NOUN$ **alors**

$aspect \leftarrow w^{i+1};$

$sentiment \leftarrow w^i$

finsi

finsi

L'un des enjeux majeurs auxquels fait face notre algorithme est la difficulté à couvrir un maximum de type d'avis. En effet, les avis des utilisateurs en ligne sont souvent non structurés et syntaxiquement incorrects ce qui compromet souvent l'efficacité d'un tel algorithme. Dans le cas de notre système, une liste de types d'avis et de leurs constructions syntaxiques a été réalisée. Grâce à ce travail préalable, il a été possible de couvrir une vingtaine de types d'avis. Une liste des aspects (robe, taille, matière, coupe, qualité, couleur, tissu) à extraire a été également définie puis utilisée comme paramètre de l'algorithme.

4. Spacy: <https://spacy.io/>

3.2.3 Évaluation du modèle d'extraction d'aspect

Pour évaluer notre modèle d'extraction d'aspect nous utilisons le jeu de données de test présenté dans la sous-section 3.1.2. Pour rappel, ce jeu de données de test est composé de 113 avis clients qui ont été labélisés à la main suivant la méthode présentée dans la sous-section 3.1.2. Dans cette collection de donnée, nous avons pu extraire 6 aspects (robe, taille, matière, coupe, qualité, tissu) 160 fois.

Pour l'évaluation du modèle d'extraction des aspects, nous avons utilisé la justesse globale (OA⁵) et la justesse moyenne (AA⁶) du modèle comme métriques. Selon [11], la justesse globale mesure la proportion de prédictions correctes dans l'ensemble de données de test du modèle. Elle représente le rapport entre le nombre de prédictions correctes et le nombre total d'échantillons d'entrées. Quant à la justesse moyenne, elle permet de mesurer la justesse de prédiction par classe. Les équations 1 et 2 présentent les formules de calcul respectives de la justesse globale et de la justesse moyenne, où # indique le nombre.

$$OA = \frac{\# \text{ d'aspects extraits correctement}}{\# \text{ d'aspects dans le jeu de données}} \quad (1)$$

$$AA = \sum_{i=1}^n \frac{\# \text{ d'aspects extraits correctement de la classe } i}{\# \text{ d'aspects dans la classe } i} \quad (2)$$

En assimilant nos opérations d'extractions à des prédictions, notre modèle présente une justesse globale OA de 1,14 et une justesse moyenne AA de 0,44. La valeur de OA supérieure à 1 indique que le modèle a extrait plus d'aspects lors de la prédiction que celle annotés. En effet l'utilisation de l'algorithme d'extraction sur le jeu de données nous retourne 185 aspects. Cependant la valeurs de AA qui mesure la justesse du modèle par aspect, nous montre que le modèle peine à retrouver les aspects que nous avons défini. Une analyse qualitative de l'extraction automatique montre que des mots tel que « décollété » ou « textile » ont été extraits comme aspects. Même si ces mots ne sont pas les aspects que nous avons définis, elles s'y rapprochent. Par exemple « textile » est proche de « tissu ». D'où la nécessité de mettre en place un modèle de regroupement des aspects qui est présenté dans la sous-section suivante.

3.3 Regroupement des aspects

3.3.1 Le modèle de regroupement des aspects

Comme montré dans la section précédente, le nombre de caractéristiques mentionnées dans les avis au regard d'un produit donné est généralement trop élevé pour réaliser une analyse pertinente. Les synonymes et les aspects similaires doivent être ainsi regroupés. Par exemple, pour les avis sur les robes, les aspects « couleur » et « teinte » doivent être associés. Pour ce faire nous utilisons le plongement lexical. Le plongement lexical ou vectorisation des mots est une technique d'apprentissage qui permet de représenter chaque mot par un vecteur de nombres réels. Cette technique a été

décrite en détails dans [10]. L'avantage d'une telle représentation est qu'elle permet d'établir la proximité contextuelle des mots [15]. En effet, des mots sémantiquement proches auront des représentations vectorielles proches.

Des modèles tels que « Word2vec » [9] ou « BERT » [2] permettent de réaliser le plongement lexical des mots en anglais. Cependant, à notre connaissance, il n'existe pas de modèle gratuit pré-entraîné en français. Par conséquent, nous avons créé notre propre modèle de plongement lexical. Pour ce faire, nous avons utilisé le modèle « Word2vec » que nous avons ré-entraîné avec un corpus de données en français. Comme indiqué dans la section 3.1.2, le corpus de données utilisé pour l'entraînement est composé de sept millions de phrases issues de l'actualité française. Voici un exemple de phrase type de ce corpus : « 20h30 : devant les députés socialistes Après les journaux télévisés, le premier ministre, Myriam El Khomri et Emmanuel Macron se sont retrouvés à l'Assemblée nationale devant les députés du groupe socialiste pour un séminaire ».

L'environnement mis en place pour l'apprentissage est une machine virtuelle Google Cloud à 36 Go de RAM. Le modèle « Word2vec » a été paramétré sur la version « CBOW » avec une taille de vecteur de 300. Le modèle ainsi entraîné est utilisé pour créer la plongement lexical des aspects extraits automatiquement, et par la suite, calculer sa similarité avec un aspect défini au préalable.

La figure 2 présente le principe de fonctionnement du regroupement des aspects.

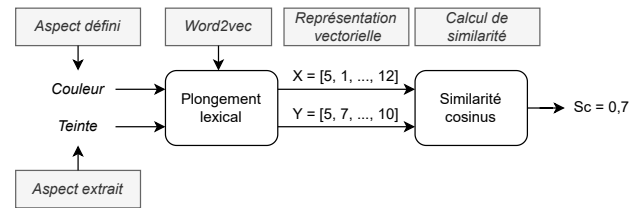


FIGURE 2 – PROCESSUS DE REGROUPEMENT DES ASPECTS.

La fonction de similarité offerte par le modèle « Word2vec » est basée sur le calcul de la similarité cosinus (S_c) entre deux vecteurs. L'équation 3 présente la formule utilisée pour calculer la S_c entre deux vecteurs X et Y [19].

$$S_c(X, Y) = \frac{X^T Y}{\|X\| \|Y\|} \quad (3)$$

où

$$\|X\| = \sqrt{\sum_{i=1}^n x_i^2} \quad (4)$$

et

$$\|Y\| = \sqrt{\sum_{i=1}^n y_i^2} \quad (5)$$

5. Overall accuracy
6. Average Accuracy

Lors de l'entraînement du modèle, nous avons procédé au nettoyage du corpus de texte afin d'améliorer sa qualité, et ce en supprimant la ponctuation. Pour rappel, notre modèle permet de mesurer la proximité entre deux aspects en produisant un score de similarité dans l'intervalle $[-1, 1]$. Pour utiliser le modèle dans notre système, il faut mettre en place un score de similarité minimum α à partir duquel nous pouvons considérer deux mots comme synonymes afin de pouvoir les regrouper.

3.3.2 Choix du seuil de similarité α

Dans l'optique de déterminer le score minimum de similarité α à partir duquel nous pouvons regrouper deux aspects, nous avons réalisé une étude empirique avec notre jeu de données de test présenté dans la sous-section 3.1.2. Nous avons appliqué à notre jeu de données le modèle de regroupement des aspects en faisant varier la valeur de α entre 0, 1 et 1, avec un pas de 0, 1.

Nous considérons le regroupement d'aspects comme un problème de classification multiple, les classes étant les aspects que nous souhaitons regrouper autour des aspects principaux définis dans la sous-section 3.1.2. Ceci nous permet d'évaluer les performances obtenues avec les différentes valeurs de α pour ensuite ne retenir que celle qui nous permet d'avoir les meilleurs résultats. Le tableau 2 présente la valeur de *AA* obtenue avec les différentes valeurs de α .

Cette étude empirique nous a permis de conclure que la valeur $\alpha = 0, 4$ est la valeur optimale avec laquelle le modèle se trompe le moins sur le regroupement des aspects. Par la suite, nous avons opté pour cette valeur dans la mise en place du système.

TABLE 2 – PERFORMANCE DE LA MÉTHODE EN FONCTION DU PARAMÈTRE DE SIMILARITÉ α

Similarité α	AA
1	0.620
0,9	0.625
0,8	0.626
0,7	0.626
0,6	0.626
0,5	0.626
0,4	0.74
0,3	0.589
0,2	0.589
0,1	0.589

3.4 Calcul du score et de la polarité d'un sentiment

3.4.1 Description du modèle d'analyse des sentiments

Le modèle d'analyse de sentiments a pour objectif d'affecter un score compris entre -1 et 1 à un aspect en fonction du mot sentiment qui lui est associé. Pour trouver le score de sentiment, nous avons fait le choix d'utiliser le modèle « *sentiment analysis* » de Google.

Notons, dans ce cadre, qu'il existe tout un système Google dédié à l'analyse sentimentale par aspect mais, pour des raisons de coût et d'efficacité, nous nous sommes limités à la composante de ce système qui est dédiée à l'analyse sentimentale.

Le modèle « *sentiment analysis* » prend comme entrée un texte et fournit en sortie un score $\beta \in [-1, 1]$. Notre calcul de score et de polarité est fondé sur ce modèle. Dans notre cas d'étude, nous utilisons comme entrée du modèle « *sentiment analysis* » l'aspect et le mot sentiment que nous extrayons de l'avis client grâce à nos modèles présentés dans la section 3.2. La sortie de notre modèle est un score exprimant la polarité associée à l'aspect qui a été précédemment extrait.

3.4.2 Illustration et évaluation du modèle d'analyse de sentiments

Pour illustrer le fonctionnement et évaluer les performances du modèle d'analyse de sentiments, nous utilisons le même jeu de données de tests que celui présenté dans la sous-section 3.3.2.

Le tableau 4 présente l'étiquetage et la notation effectués manuellement sur les avis présentés dans le tableau 3. La colonne *score* représente le score affecté à l'aspect en s'appuyant sur son mot sentiment.

TABLE 3 – EXEMPLES D'AVIS

Id	Avis
1	Très belle robe mais pas une grande qualité
2	La taille est bonne par contre attention mesdames si vous êtes grande ! Elle est hyper courte ... J'ai du la renvoyer car ça faisait vraiment trop court. Damage

TABLE 4 – ÉTIQUETAGE À LA MAIN

Id	Aspect	Mot-sentiment	Score
1	robe	belle	0,9
1	qualité	grande	-0,6
2	taille	bonne	0,8
2	robe	courte	-0,6

A partir de ces exemples, nous remarquons que le premier avis est un avis facile pour l'analyse sentimentale. En effet, la construction syntaxique de la phrase est juste et précise. En revanche, le deuxième avis est plus complexe au regard de notre modèle vu que la phrase n'est ni en français correct ni précise.

Dans le cas de l'étiquetage à la main, l'auteur est capable de comprendre que le premier avis exprime deux polarité d'opinions différentes. L'opinion positive est sur la robe en général et l'opinion négative est sur la qualité. L'auteur est aussi capable d'estimer que l'opinion positive sur la robe est plus forte que celle négative sur la qualité et ce, au travers de l'usage du superlatif « très ». Quant au deuxième avis, il est plus difficile à analyser. En effet, deux opinions opposées sont exprimées sur ce qui semble être la même

caractéristique. L’opinion négative « hyper courte » semble porter sur la robe dans le cas où la personne est de grande taille.

Le tableau 5 présente les résultats de l’analyse sentimentale associée aux deux exemples du tableau 3.

TABLE 5 – PRÉDICTIONS DE LA MÉTHODE PROPOSÉE

Id	Aspect	Mot-sentiment	Score
1	robe	belle	0,9
1	qualité	grande	-0,4
2	taille	bonne	0,9
2	robe (mesdames)	grande	-0,6

Le système est capable d’extraire les aspects et sentiments et de les noter correctement pour le premier avis. Cependant, pour le deuxième exemple, la solution proposée fonctionne moins bien. En effet, le modèle d’analyse sentimentale est programmée pour sélectionner seulement le sentiment, s , le plus fort sur un même aspect :

$$s = \max(|score(mot^i)|) \quad (6)$$

D’un autre côté, on s’aperçoit, par le biais des exemples, que le mot « mesdames » a été extrait comme un synonyme de robe par le modèle d’extraction des aspects. Cette erreur provient du modèle de regroupement « Word2Vec ». Ceci s’explique par le fait que, dans le jeu de données d’entraînement les mots « mesdames » et « robes » doivent être souvent associées. L’aspect robe aurait dû être extrait pour « elle est hyper courte » et non pas pour « mesdames si vous êtes grande ! ». Une des limites de l’analyse d’avis consommateurs en ligne est que ces avis peuvent être syntaxiquement faux, ce qui rend leur analyse très difficile, voire impossible.

Nous avons utilisé l’erreur quadratique moyenne (EQM), exprimée par l’équation 7, pour évaluer la performance de l’analyse de sentiment de notre jeu de test, où \hat{y}_i sont les scores prédits par le modèle, y_i sont les scores attribuées par l’auteur, et n représente le nombre d’aspects.

$$EQM = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n} \quad (7)$$

L’erreur quadratique moyenne permet de mesurer les écarts entre les valeurs réellement observées y_i et les valeurs prédites par le modèle \hat{y}_i . Plus elle est proche de 0 plus le modèle est précis.

Nous avons obtenu une $EQM = 0,19$ pour notre jeu de données de test. Ce qui nous permet d’affirmer que nos résultats sont prometteurs.

4 Discussion et applications pratiques du système

Dans cet article, nous proposons un système de notation des avis clients par aspects pour les produits des sites de e-commerce afin d’aider les enseignes à améliorer leur offres de produits et de services, et également guider les clients potentiels à affiner leurs choix lors de leurs achats.

Quoiqu’il existe des systèmes adressant la même problématique que la nôtre, ces derniers tendent à fournir une note globale sur un avis client, et les rares proposant des notes par aspects sont destinés aux avis clients rédigés en anglais uniquement [20].

Le système que nous proposons est destiné aux avis clients rédigés en français et est capable d’extraire automatiquement les aspects importants ayant été prédéfinis au préalable. Les enjeux majeurs actuels autour de la satisfaction client, dont les attentes sont de plus en plus hautes, et la concurrence forte entre les différentes enseignes de commerce montrent la nécessité d’un tel système.

Notre système est fondé sur des techniques de traitement automatique du langage naturel et sur de l’analyse sentimentale. Il est composé de trois modèles principaux : (a) un modèle d’extraction des aspects d’un produit ainsi que des mots sentiments reflétant l’opinion sur ces caractéristiques, (b) un modèle de regroupement d’aspects qui permet de regrouper les caractéristiques extraites d’un avis dans des classes, et (c) un modèle de calcul du score et de la polarité d’un sentiment qui permet d’attribuer un score entre -1 et 1 à un aspect extrait.

L’ensemble de ces modèles a été implémenté et évalué à partir d’un jeu de données de tests qui a été décrit dans la sous-section 3.1.2. Les résultats de l’évaluation des modèles sur notre cas d’usage (notation des robes) montrent des résultats prometteurs pour une future industrialisation.

Un cas d’application pratique de notre système a été effectué sur un jeu de données composé de 1944 avis clients suivants 77 types de robes. La figure 3 illustre la sortie de notre système sous forme de diagramme radars pour 3 robes suivant 8 caractéristiques, à savoir la matière, la robe, la couleur, la taille, la longueur, le tissu, et la coupe. Les notes pour chaque aspect représentent la moyenne en fonction des différentes occurrences présentées par la figure 4.

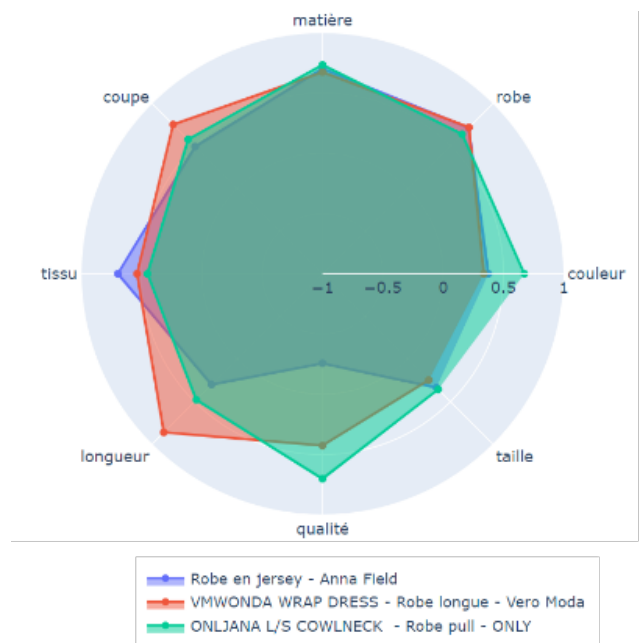


FIGURE 3 – DIAGRAMME RADAR DES ROBES.

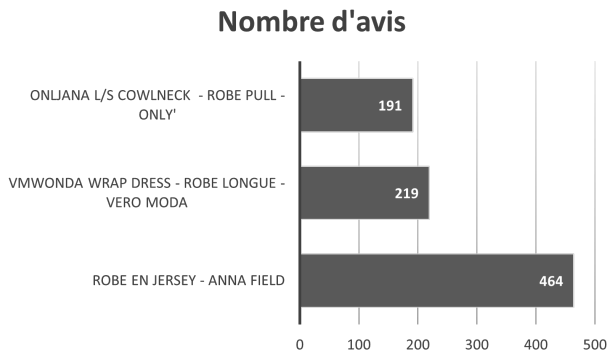


FIGURE 4 – OCCURRENCES DES ROBES DANS LES AVIS.

Ce diagramme radar offre une visibilité sur les différentes notes associées aux principales caractéristiques des robes dans les avis clients. Il représente un bon indicateur sur lequel les clients potentiels pourraient se baser pour effectuer leur choix ou encore les enseignes pour améliorer un des aspects sur lequel ils sont moins compétitifs.

5 Conclusion, limites et perspectives

Le système de notation par aspects présenté dans cet article est un système prometteur qui permet de traiter des textes en français sans avoir recours à un traducteur.

Deux populations peuvent être visées par ce système : les futurs consommateurs afin de pouvoir affiner leurs choix de produits selon leurs préférences, et les enseignes des produits notés afin d'améliorer leurs offres en considérant les différents aspects auxquels les consommateurs se sont intéressés.

Cependant, la performances encourageante du système proposé ne doit pas être un frein pour des améliorations futures. En effet, pour améliorer cette solution, un modèle de notation des sentiments en français pourra être entraîné. Cela permettra à notre système d'être entièrement indépendant au regard de tout modèle payant. L'utilisation du modèle Google a un coût d'environ 0,0010€ par 1000 caractères. C'est le seul modèle payant de notre algorithme. En outre, il est possible d'augmenter le nombre de types d'avis couverts par la méthode ce qui entraîne une véritable valeur commerciale pour le système suggéré.

Finalement, nous pourrions également envisager des perspectives dont l'objectif sera de comparer la méthode présentée dans cet article avec les méthodes récentes d'IA Générative (GPT, BARD, GROK, MISTRAL) en utilisant les jeux de données annotés SemEval.

Remerciements

Merci à notre entreprise, Teolia⁷, de nous avoir soutenu dans ce projet de recherche. Un grand merci, en particulier, à tous les consultants Data de Teolia qui ont participé au développement des différents algorithmes.

7. <https://www.teolia.fr/>

Références

- [1] Fouad Nasser A Al Omran and Christoph Treude. Choosing an nlp library for analyzing software documentation : a systematic literature review and a series of experiments. In *2017 IEEE/ACM 14th international conference on mining software repositories (MSR)*, pages 187–197. IEEE, 2017.
- [2] Shivaji Alaparathi and Manit Mishra. Bert : A sentiment analysis odyssey. *Journal of Marketing Analytics*, 9(2) :118–126, 2021.
- [3] Karim Baati. Hybridization of adaboost with random forest for real-time prediction of online shoppers' purchasing intention. In *Hybrid Intelligent Systems : 20th International Conference on Hybrid Intelligent Systems (HIS 2020), December 14-16, 2020*, pages 234–241. Springer, 2021.
- [4] Karim Baati and Mouad Mohsil. Real-time prediction of online shoppers' purchasing intention using random forest. In *Artificial Intelligence Applications and Innovations : 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part I 16*, pages 43–51. Springer, 2020.
- [5] Stefano Baccianella, Andrea Esuli, Fabrizio Sebastiani, et al. Sentiwordnet 3.0 : an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec*, volume 10, pages 2200–2204, 2010.
- [6] Satanjeev Banerjee and Ted Pedersen. An adapted lesk algorithm for word sense disambiguation using wordnet. In *Computational Linguistics and Intelligent Text Processing : Third International Conference, CILCling 2002 Mexico City, Mexico, February 17–23, 2002 Proceedings 3*, pages 136–145. Springer, 2002.
- [7] Bhavana R Bhamare, P Jeyanthi, and R Subhashini. Aspect level sentiment analysis approaches. In *2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, pages 1–5. IEEE, 2019.
- [8] Gianni Brauwere and Flavius Frasinca. A survey on aspect-based sentiment classification. *ACM Computing Surveys*, 55(4) :1–37, 2022.
- [9] Kenneth Ward Church. Word2vec. *Natural Language Engineering*, 23(1) :155–162, 2017.
- [10] Yoav Goldberg and Omer Levy. word2vec explained : deriving mikolov et al.'s negative-sampling word-embedding method, 2014.
- [11] Ahmed Taha Haouari, Labiba Souici-Meslati, Fadila Atil, and Djamel Meslati. Empirical comparison and evaluation of artificial immune systems in inter-release software fault prediction. *Applied Soft Computing*, 96 :106686, 2020.
- [12] Matthew Honnibal and Mark Johnson. An improved non-monotonic transition system for dependency parsing. In *Proceedings of the 2015 conference on empirical*

rical methods in natural language processing, pages 1373–1378, 2015.

- [13] Ambreen Nazir, Yuan Rao, Lianwei Wu, and Ling Sun. Issues and challenges of aspect-based sentiment analysis : A comprehensive survey. *IEEE Transactions on Affective Computing*, 13(2) :845–863, 2020.
- [14] Joakim Nivre, Johan Hall, Jens Nilsson, Gülşen Eryiğit, and Svetoslav Marinov. Labeled pseudo-projective dependency parsing with support vector machines. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, pages 221–225, 2006.
- [15] Debanjan Paul, Sudeshna Sarkar, Muthusamy Chelliah, Chetan Kalyan, and Prajit Prashant Sinai Nadkarni. Recommendation of high quality representative reviews in e-commerce. In *Proceedings of the eleventh ACM conference on recommender systems*, pages 311–315, 2017.
- [16] IKCU Perera and HA Caldera. Aspect based opinion mining on restaurant reviews. In *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCI)*, pages 542–546. IEEE, 2017.
- [17] M Rathan, Vishwanath R Hulipalled, KR Venugopal, and LM Patnaik. Consumer insight mining : aspect based twitter opinion mining of mobile phone reviews. *Applied Soft Computing*, 68 :765–773, 2018.
- [18] Kim Schouten and Flavius Frasinca. Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 28(3) :813–830, 2015.
- [19] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern recognition*. Elsevier, 2006.
- [20] Satuluri Vanaja and Meena Belwal. Aspect-level sentiment analysis on e-commerce data. In *2018 International conference on inventive research in computing applications (ICIRCA)*, pages 1275–1279. IEEE, 2018.
- [21] Hai Wan, Yufei Yang, Jianfeng Du, Yanan Liu, Kunxun Qi, and Jeff Z Pan. Target-aspect-sentiment joint detection for aspect-based sentiment analysis. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 9122–9129, 2020.
- [22] Ye Yiran and Sangeet Srivastava. Aspect-based sentiment analysis on mobile phone reviews with lda. In *Proceedings of the 2019 4th International Conference on Machine Learning Technologies*, pages 101–105, 2019.