



PROJET DE FIN D'ANNÉE

SPÉCIALITÉ : GÉNIE LOGICIEL

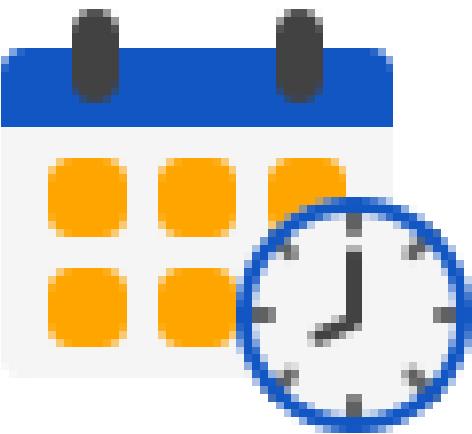
MISE EN PLACE D'UN SYSTÈME DE
RECOMMANDATION HOTELIÈRE
basé sur l'analyse sentimentale

RÉALISÉ PAR :

EYA RAOUINE
FATMA HAMADA
WIEM BEN
MLOUKA

RÉALISÉ PAR :

MME SANA
HAMDI

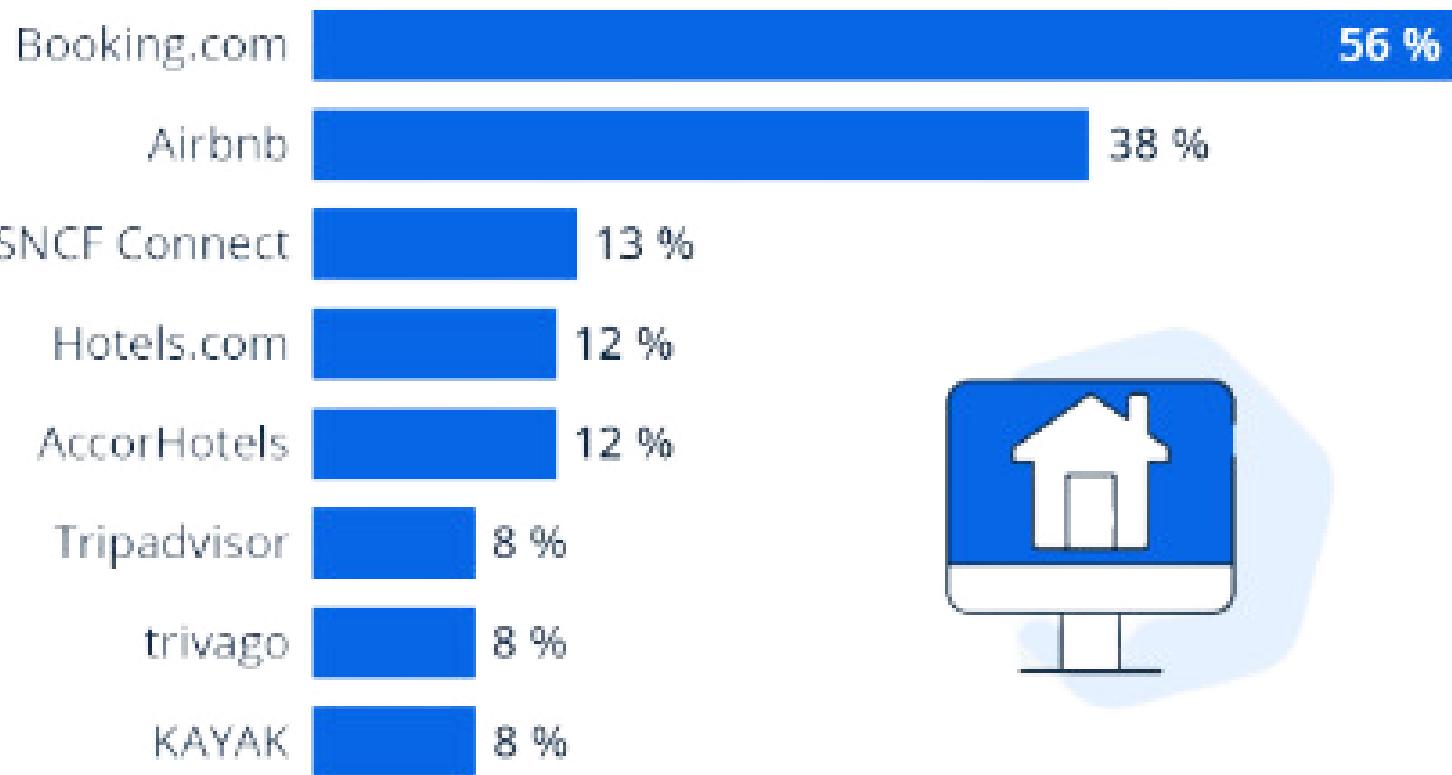


PLAN



- 1** Contexte du projet
- 2** Méthodologie et architecture
- 3** Collecte et Préparation de données
- 4** Modélisation et évaluation
- 5** Démonstration
- 6** Conclusion et perspectives

INTRODUCTION



**E-TOURISME: LES SITES LES PLUS POPULAIRES EN FRANCE
1000 ADULTES INTERROGÉ ENTRE JANVIER ET DECEMBRE 2023**



SYSTÈMES DE NOTATION PAR ÉTOILE



CLASSIFICATION DES AVIS EN OPINIONS POSITIVES, NÉGATIVES ET NEUTRES



APPROCHE basée sur les aspects de l'analyse sentimentale





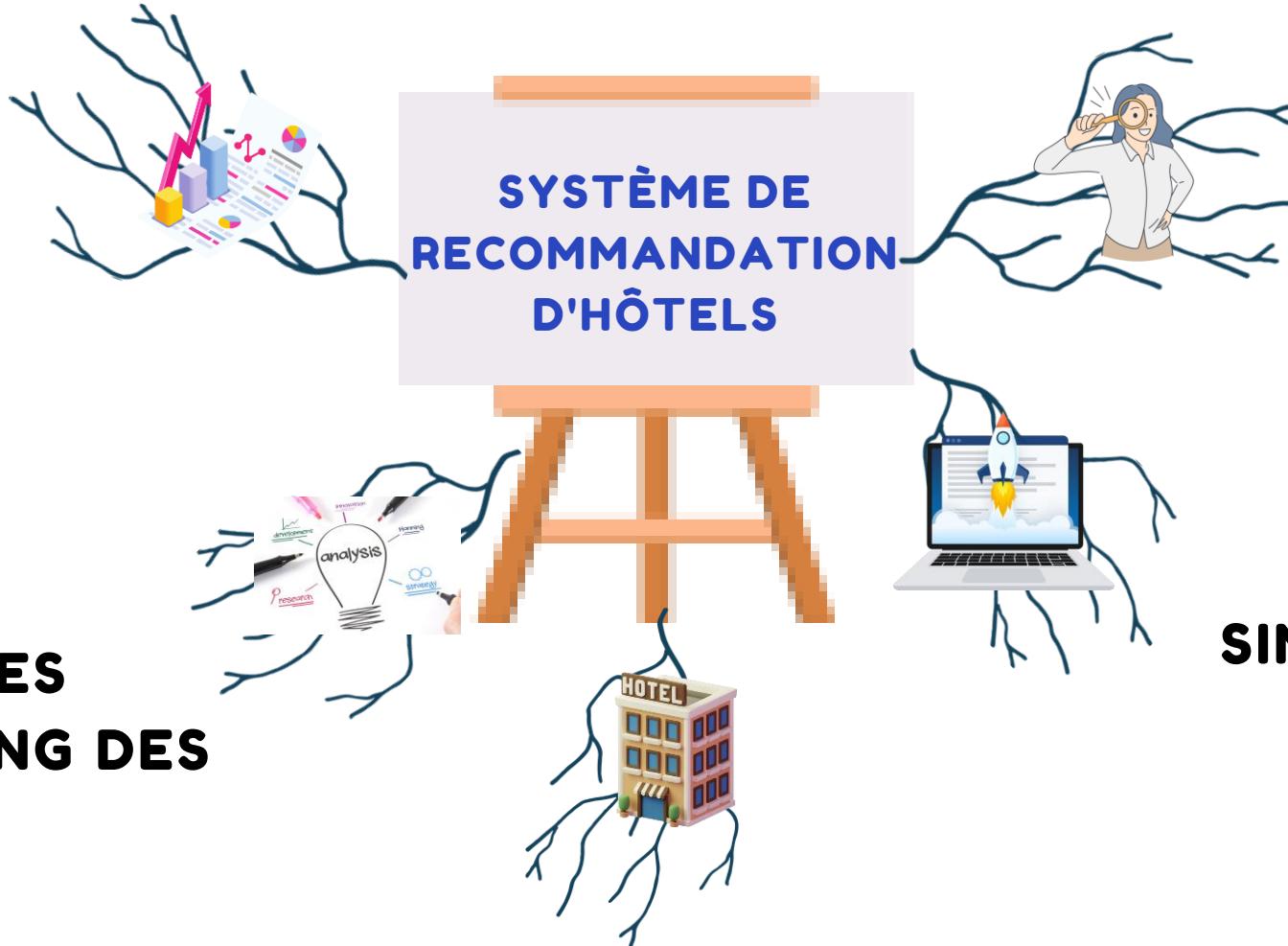
1

CONTEXTE DU PROJET

CONTEXTE GÉNÉRAL

**PLEINE CROISSANCE DE
L'INDUSTRIE HÔTELIÈRE**

**EXPLOITER L'ANALYSE DES
SENTIMENTS ET LE CLUSTERING DES
UTILISATEURS POUR**



**BESOIN DE RECOMMANDATIONS
PERSONNALISÉES**

**SIMPLIFIER ET OPTIMISER LEUR
RECHERCHE D'HÔTELS.**

**PROPOSER DES SUGGESTIONS
D'HÔTELS PERTINENTES ET
SATISFAISANTES.**

PROBLÉMATIQUE

**COMMENT EXPLOITER EFFICACEMENT LES
DONNÉES?**

RÉSEAUX
SOCIAUX

PLATEFORMES
DE RÉSERVATION
EN LIGNE

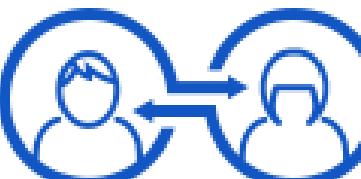


COMMENT PRENDRE EN COMPTE ?

LES INTÉRÊTS
PERSONNELS DES
UTILISATEURS

LES AVIS DES
AUTRES
CLIENTS

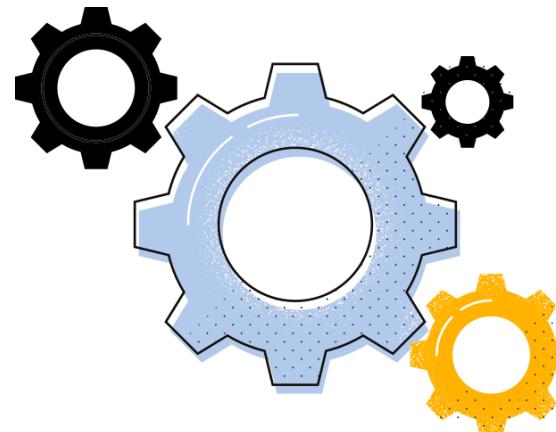
**COMMENT FAIRE LE MATCHING ET LA
RECOMMANDATION ?**



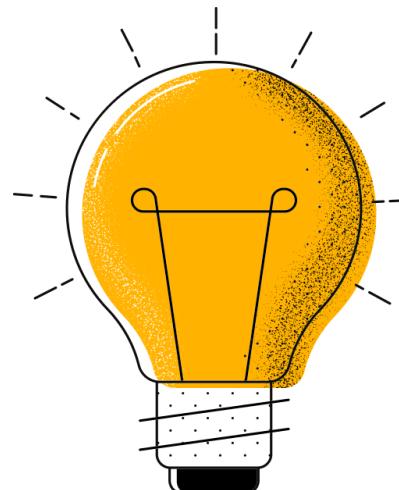
POUR PROPOSER
DES HÔTELS
PERTINENTS

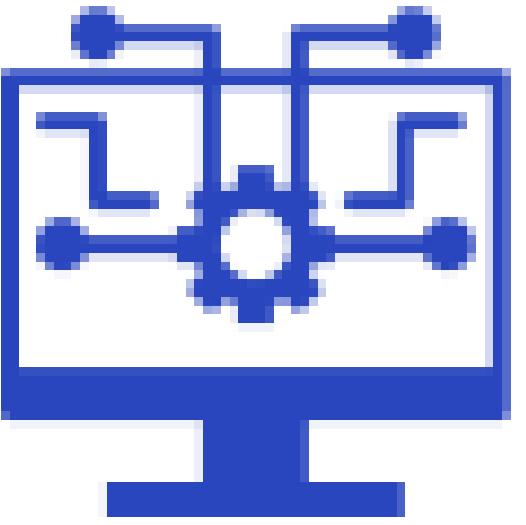


OBJECTIFS DU PROJET



- EXTRACTION ET ANALYSE DES PROFILS DES UTILISATEURS
- CLUSTERING DES UTILISATEURS PAR INTÉRÊTS
- ANALYSE SENTIMENTALE DES COMMENTAIRES D'HÔTELS
- MATCHING DES UTILISATEURS
- RECOMMANDATION D'HÔTELS

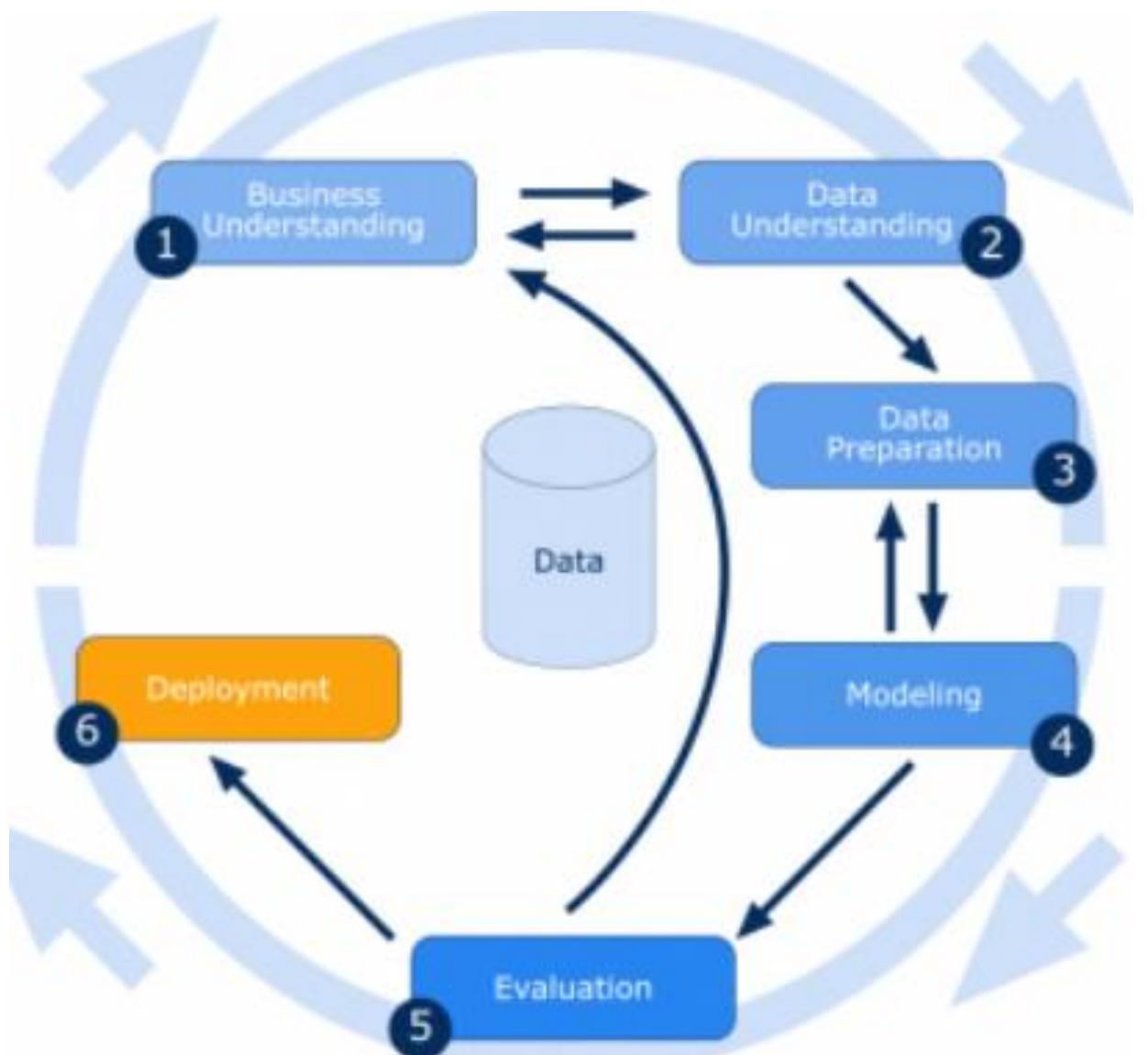




2

MÉTHODOLOGIE ET ARCHITECTURE

MÉTHODOLOGIE ADOPTÉE :CRISP-DM



- 1. COMPRENDRE LES OBJECTIFS ET LES BESOINS DU PROJET.**
- 2. EXPLORER ET ÉVALUER LA QUALITÉ DES DONNÉES**
- 3. NETTOYER, TRANSFORMER ET PRÉPARER LES DONNÉES POUR L'ANALYSE**
- 4. CONSTRUIRE DES MODÈLES PRÉDICTIFS OU DES ALGORITHMES D'APPRENTISSAGE AUTOMATIQUE**
- 5. ÉVALUER LA PERFORMANCE DES MODÈLES ET LES AJUSTER SI NÉCESSAIRE.**
- 6. DÉPLOYER ET SURVEILLER LES MODÈLES DANS UN ENVIRONNEMENT OPÉRATIONNEL .**

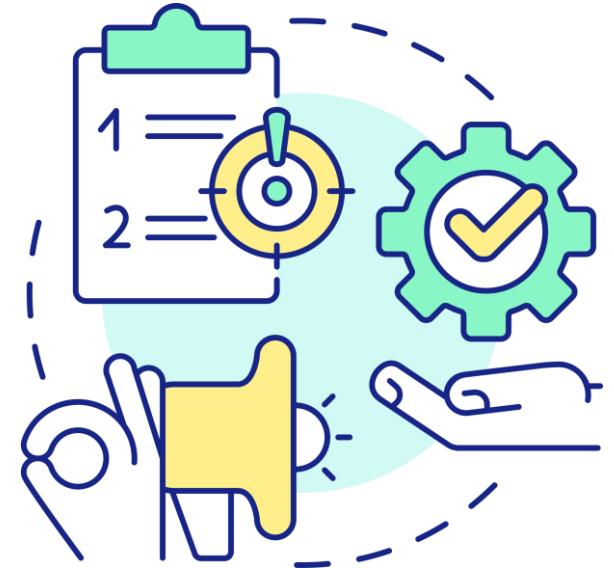
TYPE DE SYSTÈME DE RECOMMANDATION CHOISI

SYSTÈME DE
RECOMMANDATION HYBRIDE

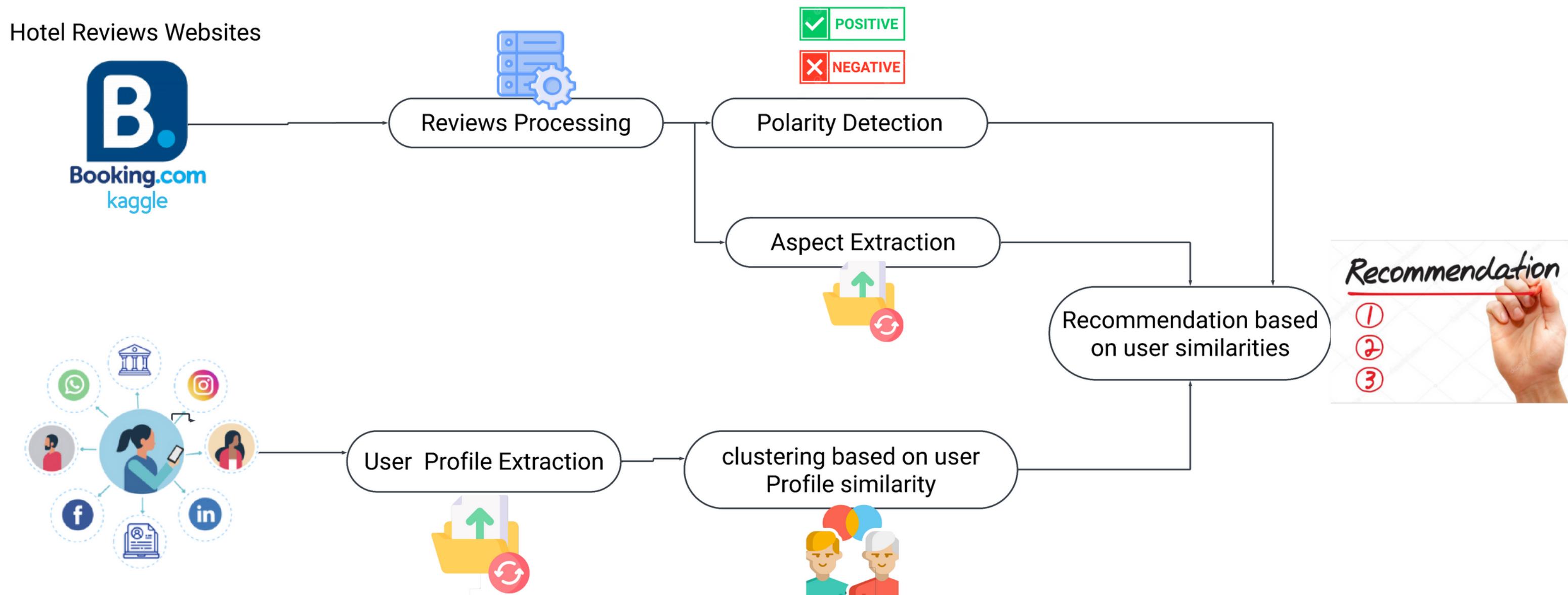


BASÉ SUR LE CONTENU : PREND EN COMPTE LE PROFIL DE L'UTILISATEUR

COLLABORATIF : PREND EN COMPTE LES PRÉFÉRENCES DES UTILISATEURS SIMILAIRES



VUE GLOBALE DE L'ARCHITECTURE PROPOSÉE

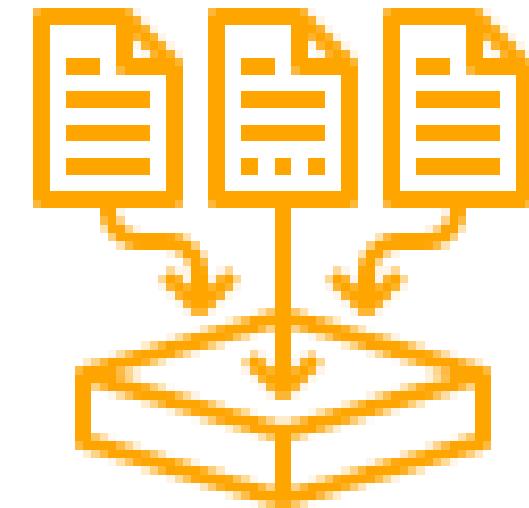




3

COMPRÉHENSION ET PRÉPARATION DE DONNÉES

COMPRÉHENSION DES DONNÉES



2 SOURCES DE DONNÉES :

1

DATASET CONTENANT DES PROFILS D'UTILISATEURS INSTAGRAM

2

DATASET CONTENANT DES AVIS SUR DES HÔTELS



1

DATASET PROFILS D'UTILISATEURS INSTAGRAM



**CONTIENT LES
PROFILS
D'UTILISATEURS
SUR INSTAGRAM**

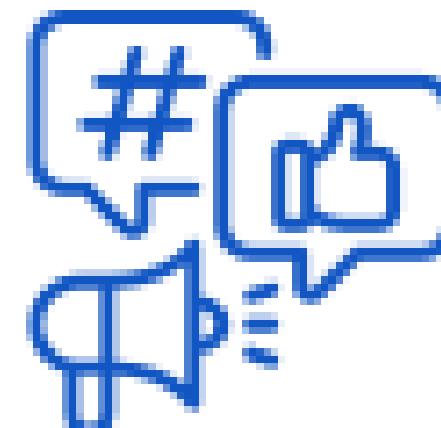
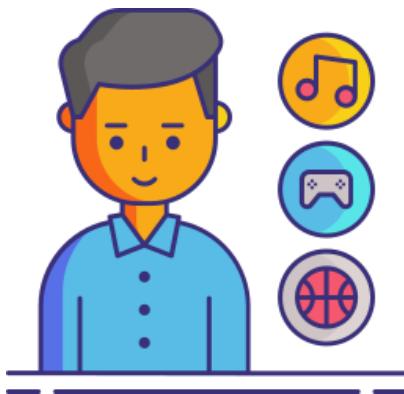
**COMPREND 884
INSTANCES**

**COMPREND 6
ATTRIBUTS**

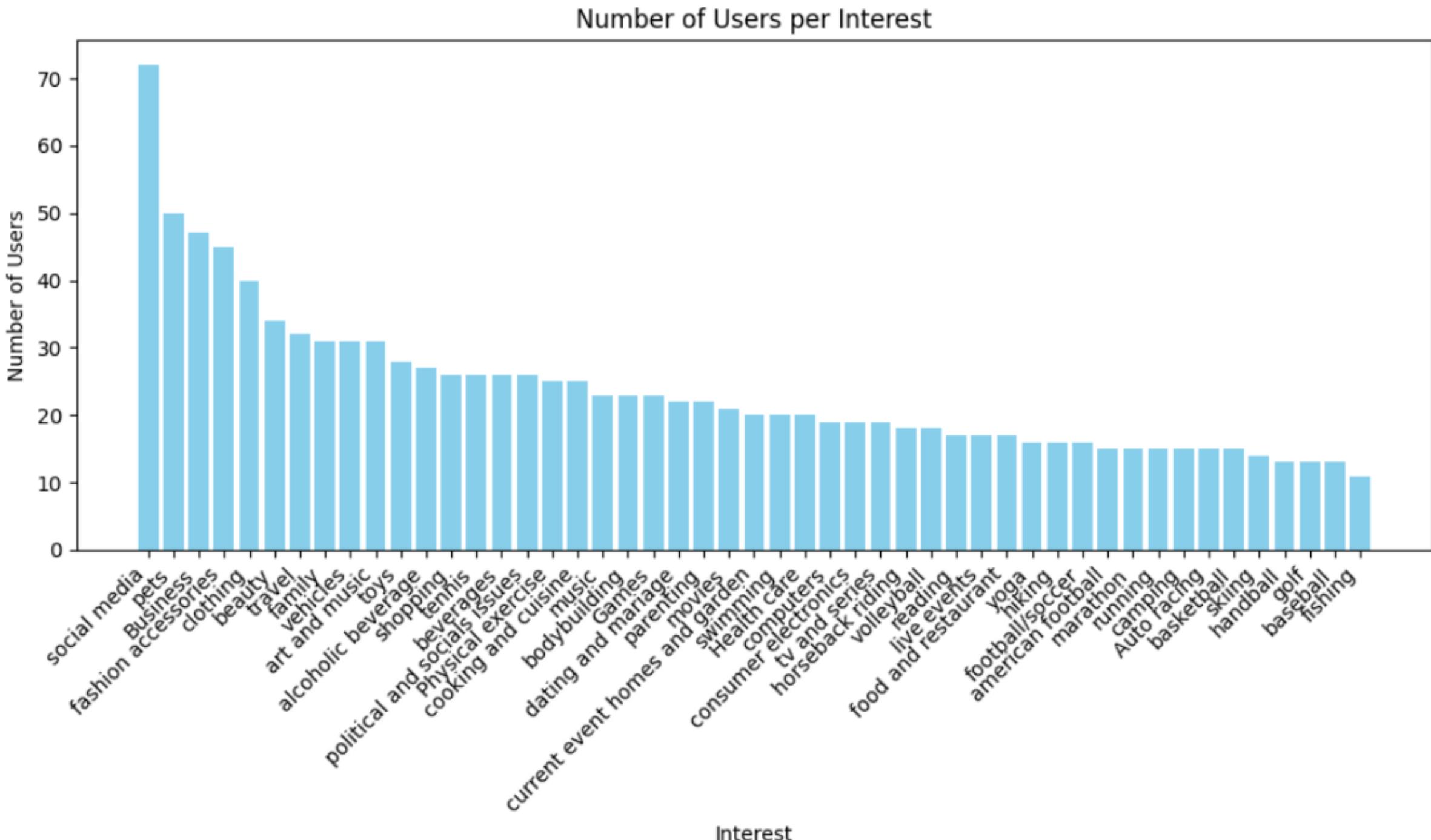
ATTRIBUTS DETAILLÉS



- **USERNAME** : NOM DE L'UTILISATEUR
- **USER ID** : IDENTIFIANT UNIQUE DE L'UTILISATEUR
- **FOLLOWERS COUNT** : NOMBRE DE FOLLOWERS DE L'UTILISATEUR
- **POSTS COUNT** : NOMBRE DE PUBLICATIONS DE L'UTILISATEUR
- **POSTS** : DÉTAILS DES PUBLICATIONS DE L'UTILISATEUR
- INTERESTS** : CENTRES D'INTÉRÊT DE L'UTILISATEUR



NOMBRE D'UTILISATEURS PAR INTÉRET



2

DATASET AVIS HOTELS



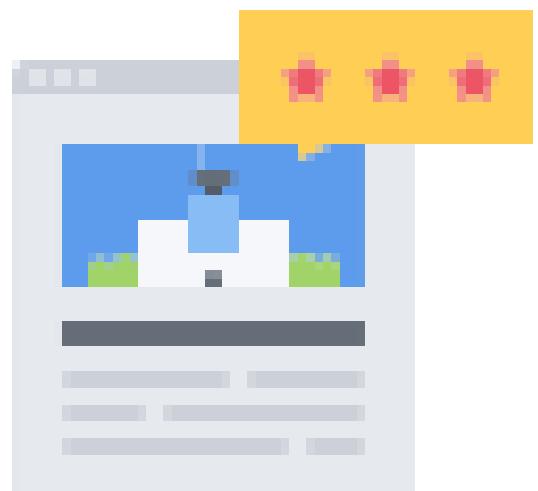
CONTIENT LES
AVIS
D'UTILISATEURS
SUR DIFFERENTS
HOTELS, EXTRAITES
DU SITE
BOOKING.COM DE
LA PLATEFORME
KAGGLE

COMPREND 26675
INSTANCES

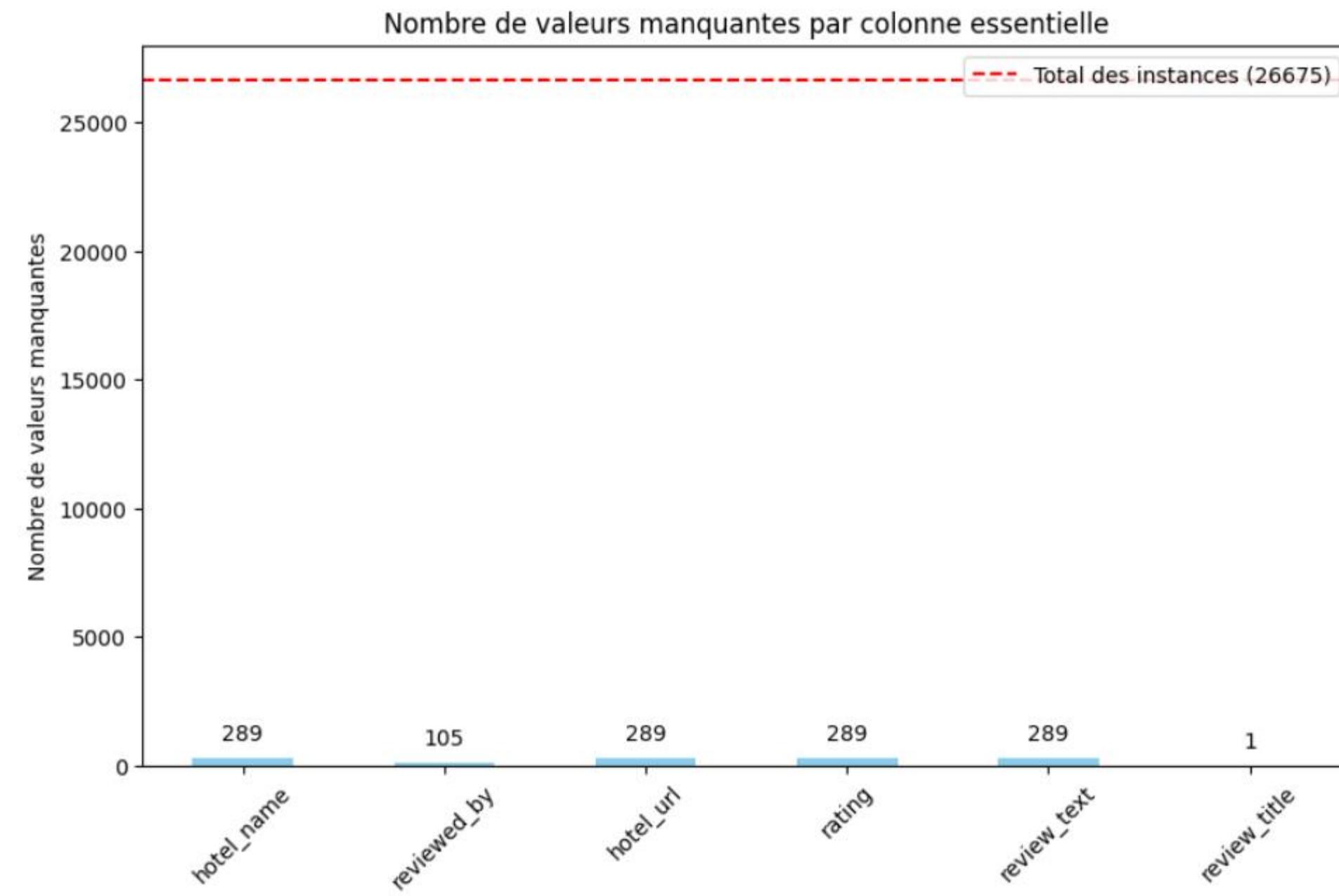
COMPREND 16
ATTRIBUTS

ATTRIBUTS DETAILLÉS

- REVIEW TITLE : TITRE DE L'AVIS
- REVIEWED AT : DATE DE L'AVIS
- REVIEWED BY : AUTEUR DE L'AVIS
- IMAGES : IMAGES INCLUSES DANS L'AVIS
- CRAWLED AT : DATE DE COLLECTE DES DONNEES
- URL : URL DE L'AVIS
- HOTEL NAME : NOM DE L'HOTEL
- HOTEL URL : URL DE L'HOTEL
- AVG RATING : NOTE MOYENNE DE L'HOTEL
- NATIONALITY : NATIONALITÉ DE L'AUTEUR DE L'AVIS
- RATING : NOTE DE L'AVIS
- REVIEW TEXT : TEXTE DE L'AVIS
- RAW REVIEW TEXT : TEXTE BRUT DE L'AVIS
- TAGS : TAGS ASSOCIES A L'AVIS
- META : METADONNEES



NOMBRE DE VALEURS MANQUANTES



EXEMPLES DE VALEURS DE LA COLONNE “REVIEWS TEXT”

There are no comments available for this review

Location

Nothing,\n\nEverything

Everything

Nothing ,\n\nEverything

...

Little crowded when there was a wedding reception ,\n\nProfessional and friendly staff, very comfortable accommodation, great breakfast and excellent location. The bed is huge!

PRÉPARATION DES DONNÉES

1

DATASET PROFILS D'UTILISATEURS INSTAGRAM



Input	Output
Business and Industry: [social media, Business], family and relationships: [family, parenting], fitness and wellness: [Physical exercise], shopping and fashion: [clothing, fashion accessories], sports: [football/soccer]	social media , Business family, parenting, Physical exercise ,clothing, fashion accessories, football/soccer

PRÉPARATION DES DONNÉES

2

DATASET AVIS HOTELS

FILTRAGE DES COLONNES REQUISES

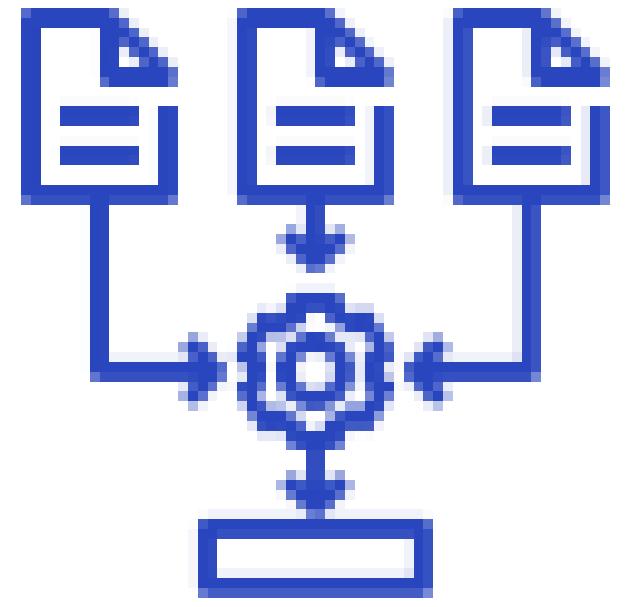
AUTEUR DE L'AVIS, NOM DE L'HOTEL, URL DE L'HÔTEL, TITRE DE L'AVIS, TEXTE DE L'AVIS, NOTE ATTRIBUÉE A L'HÔTEL ET TAGS ASSOCIÉS A L'AVIS

TRAITEMENT DES AVIS ET NORMALISATION DU TEXTE

CONVERSION EN MINUSCULES, SUPPRESSION DES ESPACES, SUPPRESSION DES MOTS "EVERYTHING" ET "NOTHING"

COMBINAISON DU TITRE ET DU TEXTE DE L'AVIS

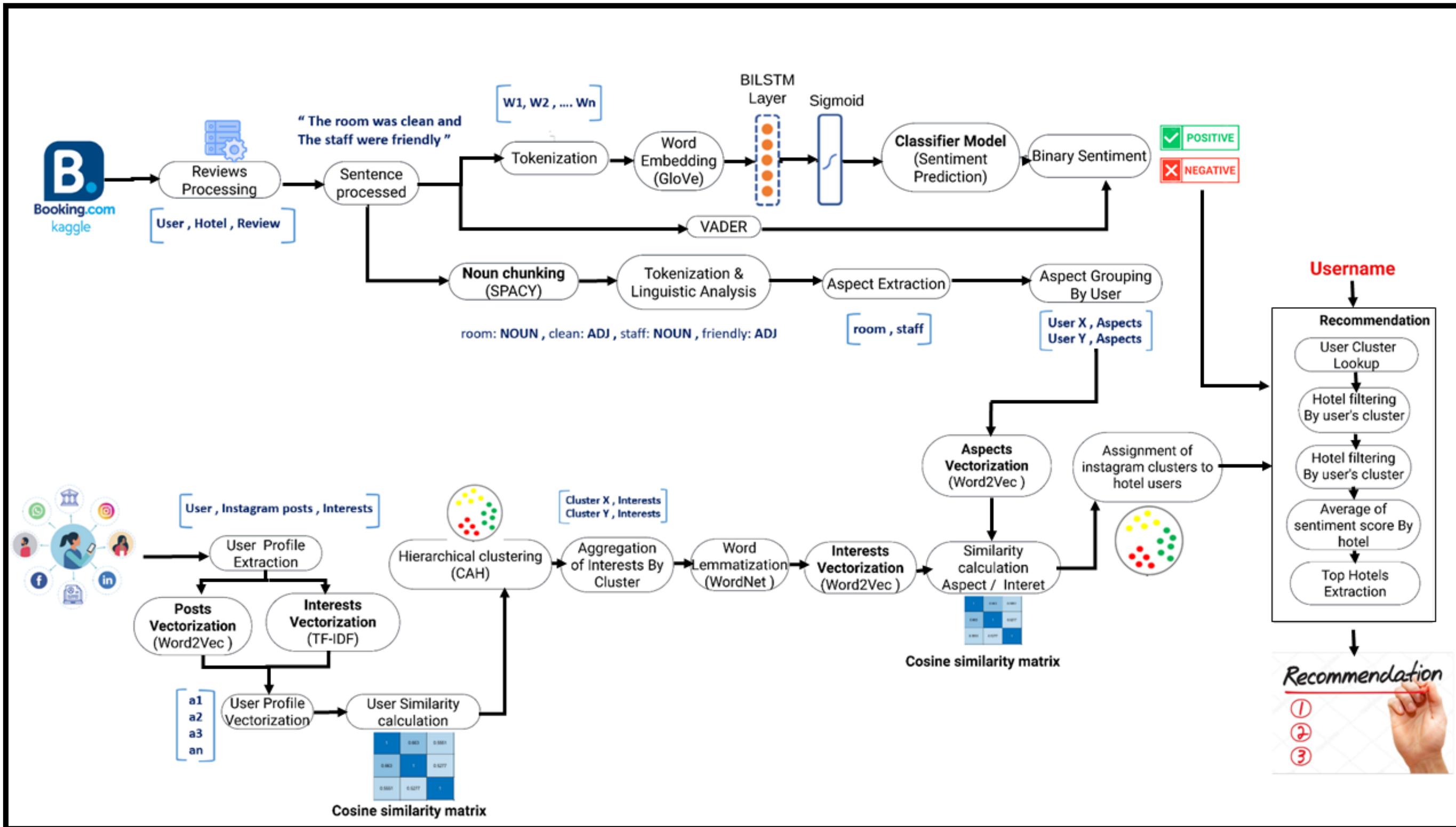
SUPPRESSION DES VALEURS MANQUANTES



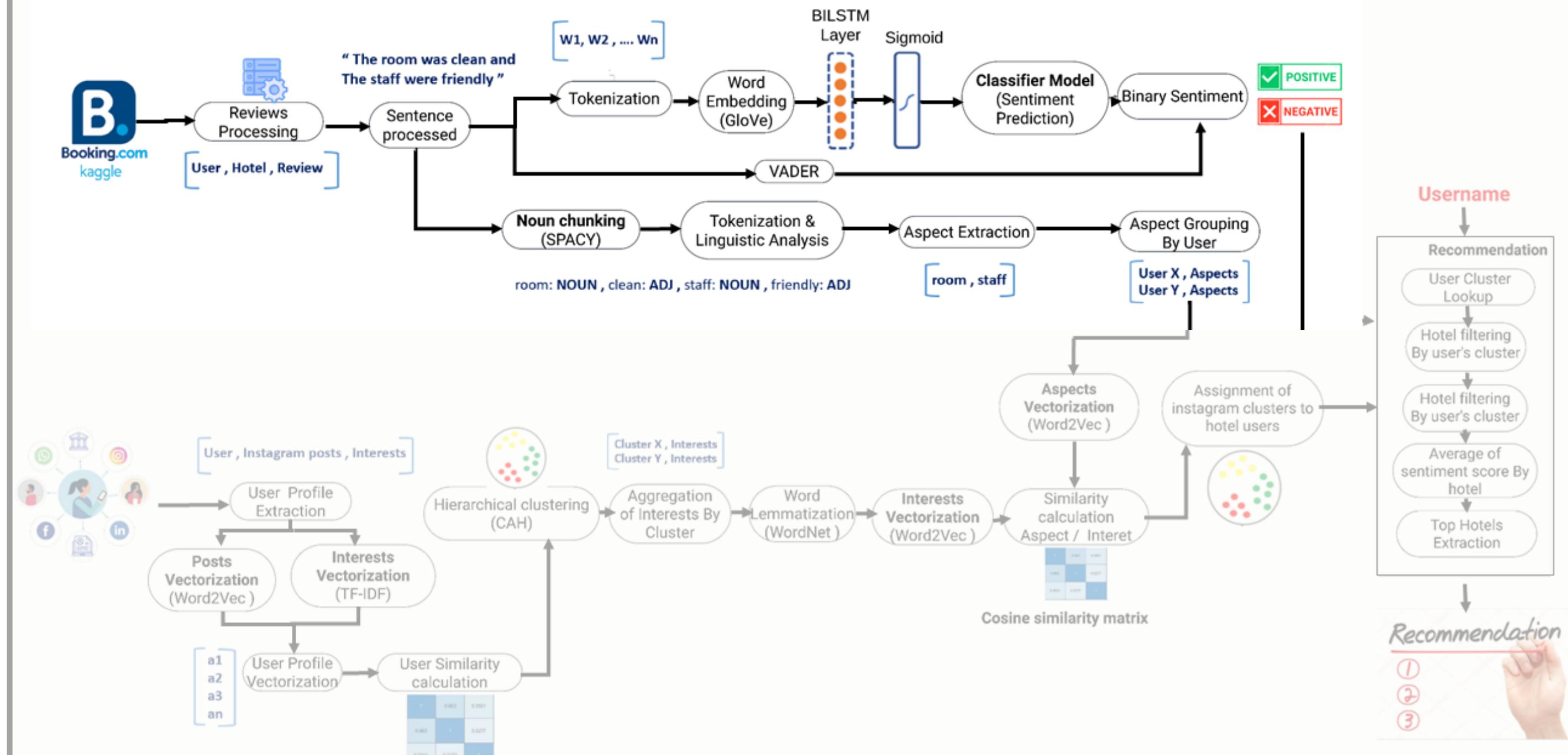
4

MODÉLISATION ET ÉVALUATION

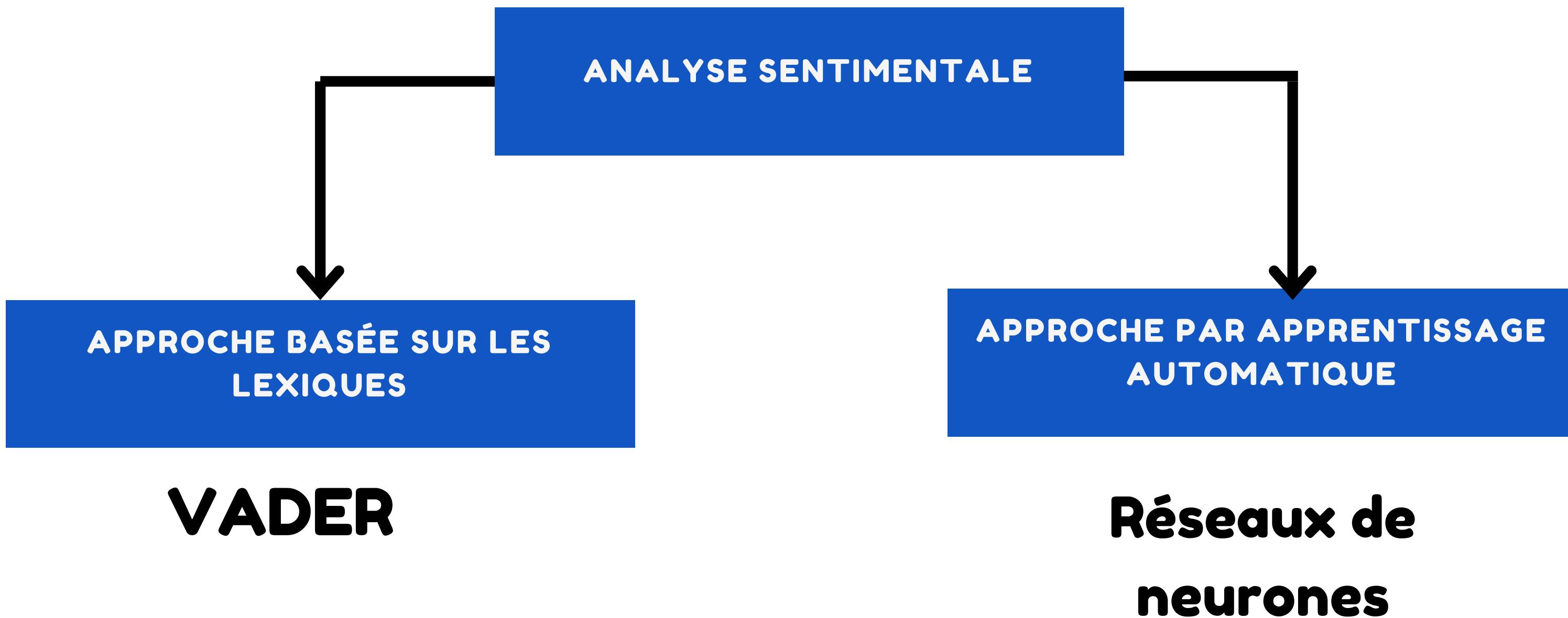
Architecture détaillée de la solution proposée



1. Analyse sentimentale des avis sur les hôtels



1. Analyse sentimentale des avis sur les hôtels

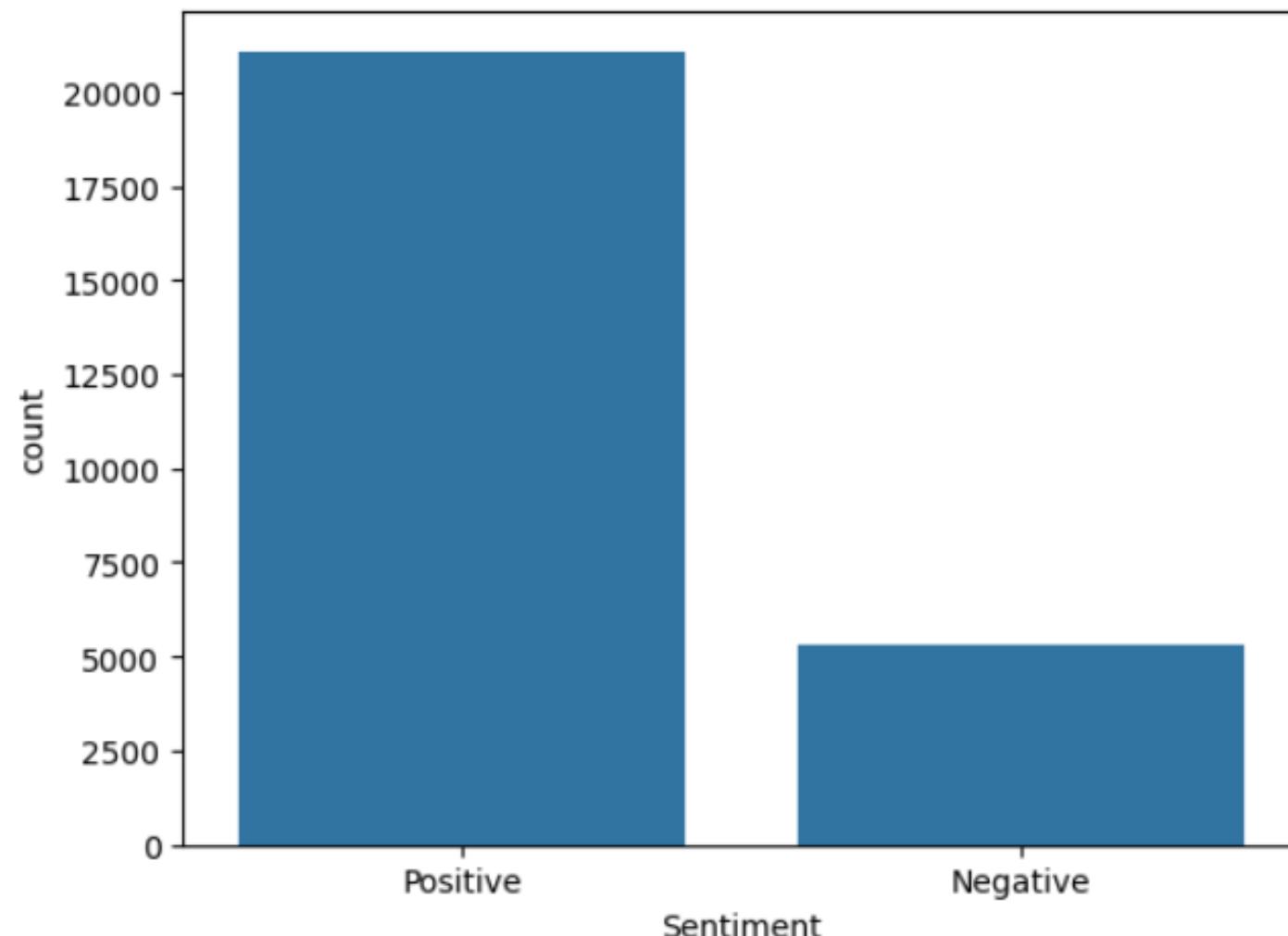


1. Analyse sentimentale des avis sur les hôtels

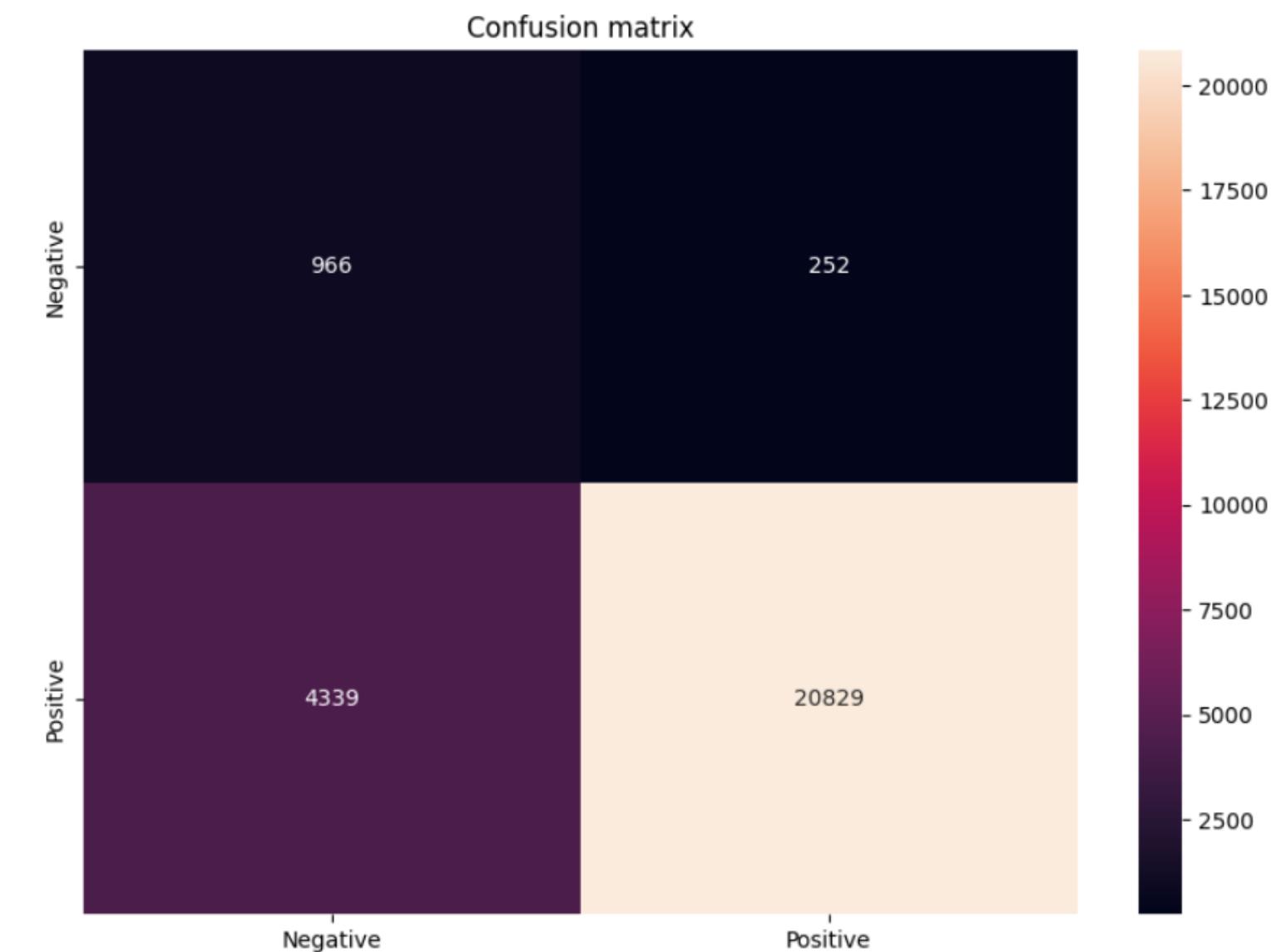
Approche basée sur les lexiques : VADER

Répartition des avis positifs et négatifs

```
Positive    21081  
Negative   5305  
Name: count, dtype: int64
```



Matrice de confusion



1. Analyse sentimentale des avis sur les hôtels

Approche basée sur les lexiques : VADER

	Precision
Negative	0.18
Positive	0.99
Accuracy Totale	0.83

1. Analyse sentimentale des avis sur les hotels

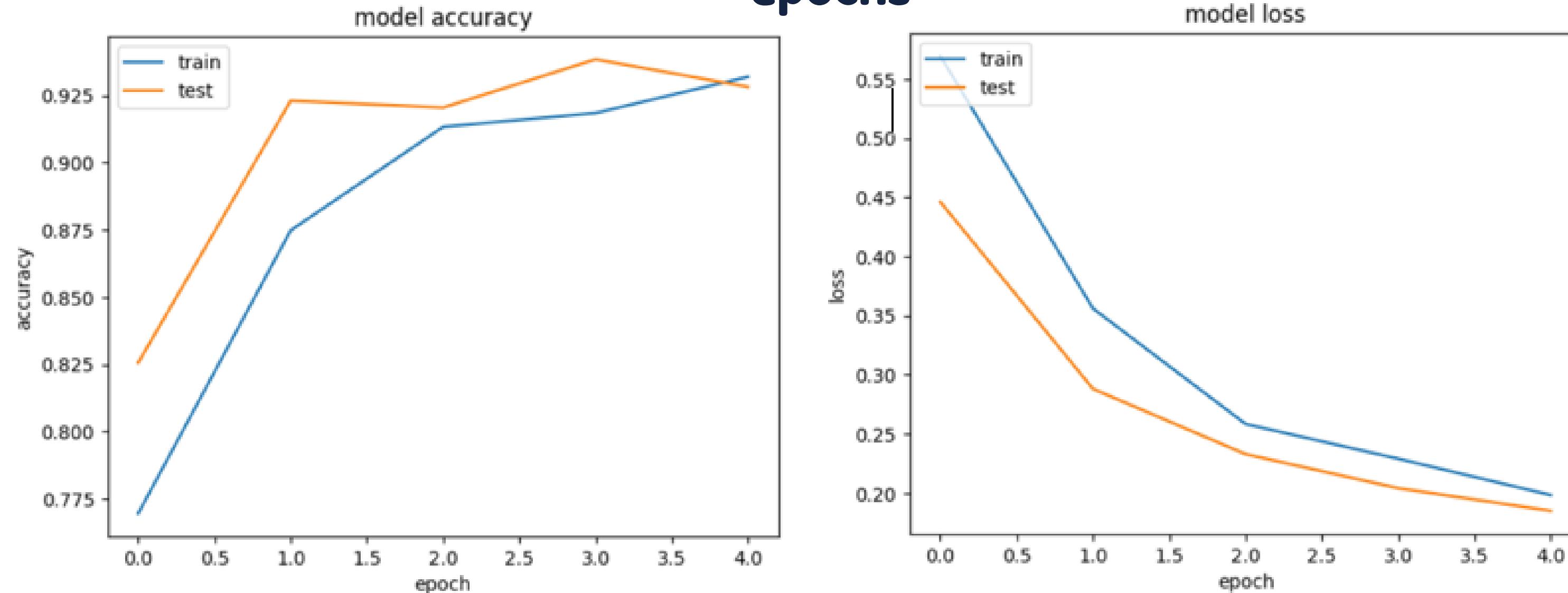
Approche basée sur les réseaux de neurones

Modèle	Train Accuracy	Test Accuracy
LSTM Simple	54%	56%
LSTM + Word2Vec	56%	75%
BILSTM	99.17%	93.19%
BILSTM + GLOVE	93.20%	93.44%

1. Analyse sentimentale des avis sur les hôtels

Approche basée sur les réseaux de neurones

Evolution du modèle en fonction des
epochs



1. Analyse sentimentale des avis sur les hôtels

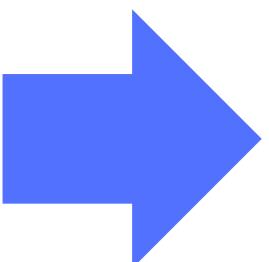
Approche basée sur les réseaux de neurones

Prédiction

	reviewed_by	hotel_name	final_review	Label_Sentiment	predicted_score	predicted_sentiment
0	Kyrylo	Villa Pura Vida	exceptional. was perfect! quite, cozy place t...	Positive	0.999794	Positive
1	Dimitri	Villa Pura Vida	i highly recommend this b&bl! we enjoyed it a l...	Positive	0.995007	Positive
2	Virginia	Hydro Palace Apartment	exceptional. it was just what we wanted for a ...	Positive	0.999318	Positive

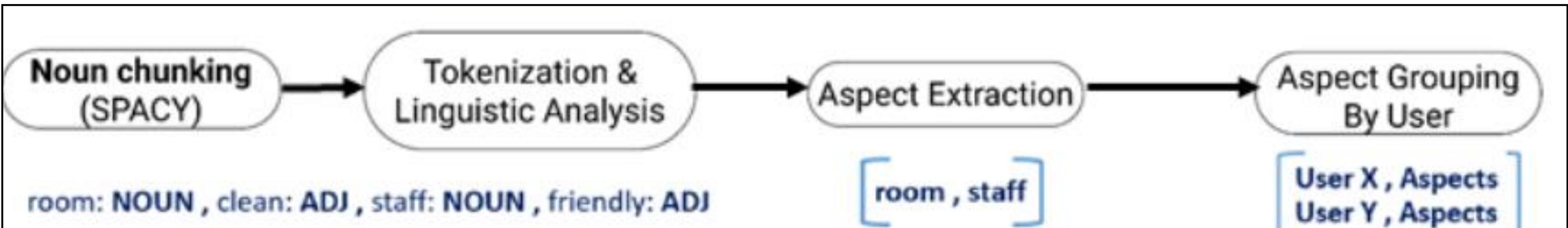
Extraction des aspects à partir des avis

Bibliothèque : spaCy



Identifier les structures grammaticales et les dépendances syntaxiques dans le texte

"The room was clean and
The staff were friendly"



Extraction des aspects à partir des avis

Liste des aspects par utilisateur

2 Virginia room, ,windows,rooms , bathrooms ,shower,roo...

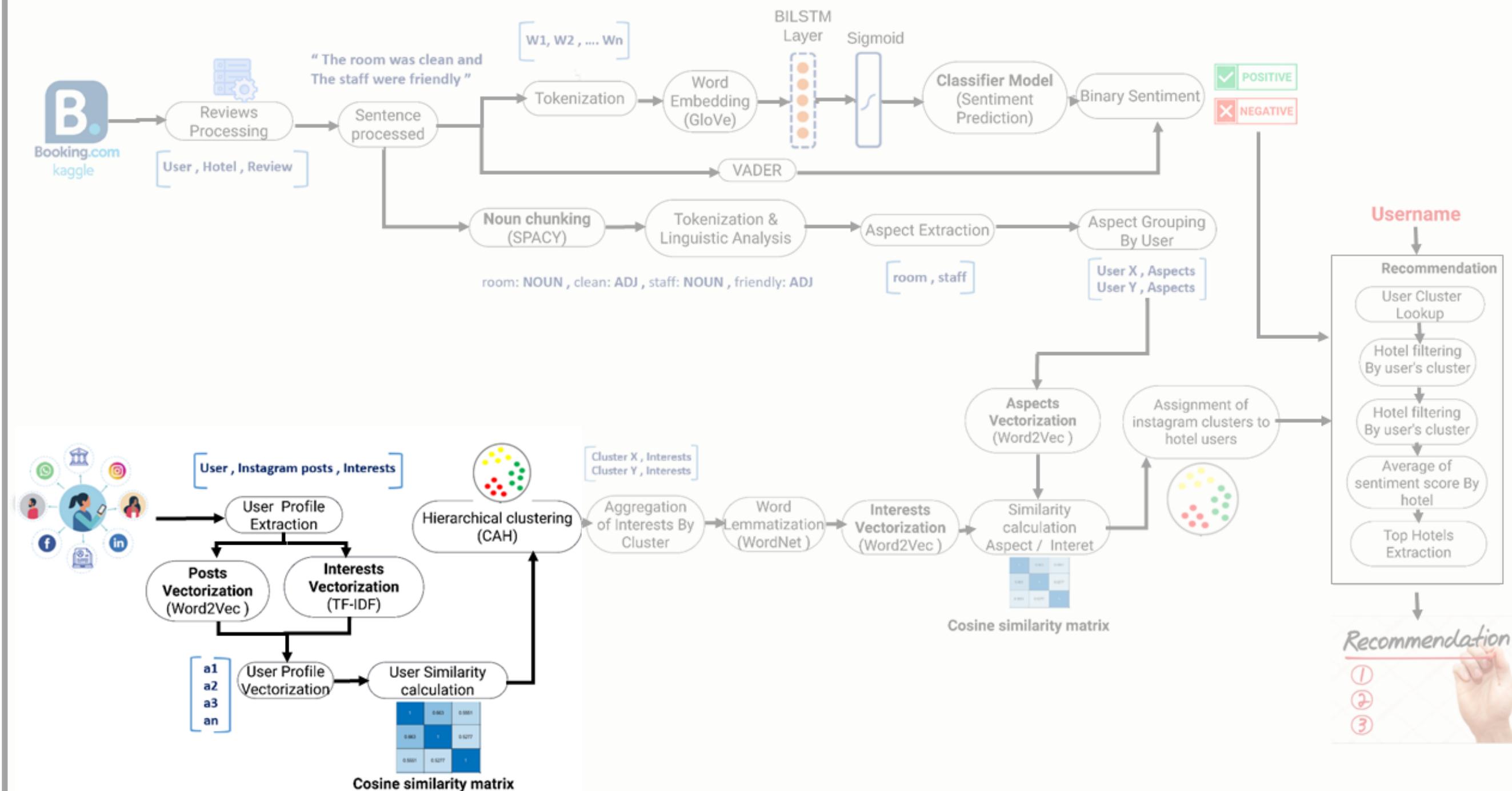
3 Kannan house,stay

4 Sue apartment, reception,water,beds,area , recepti...

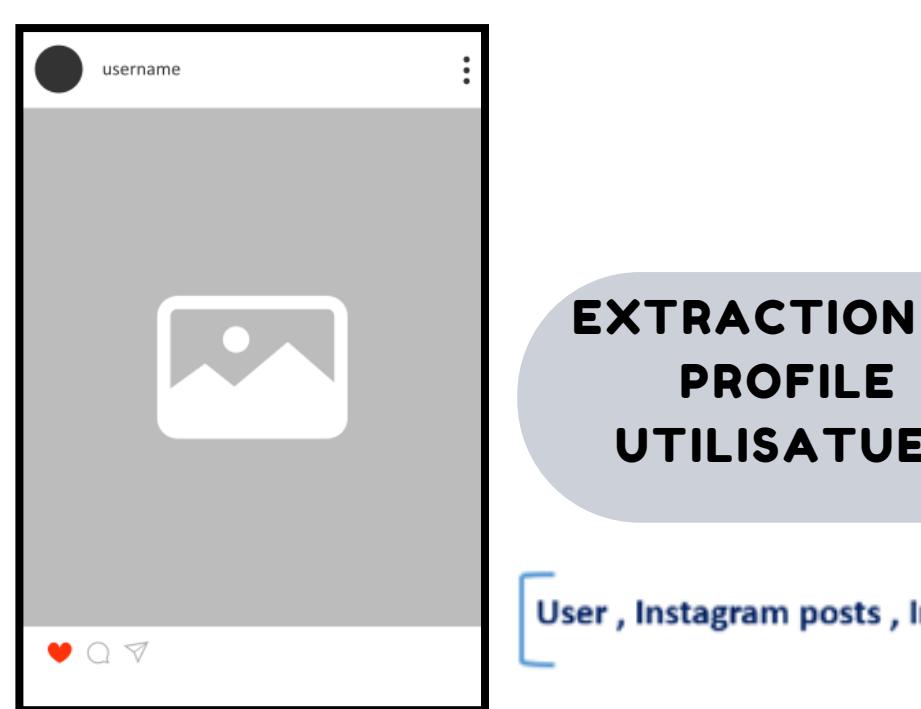
...

26369 Aynur room

2. Clustering des utilisateurs en fonction de leurs profils



2. Clustering des utilisateurs en fonction de leurs profils



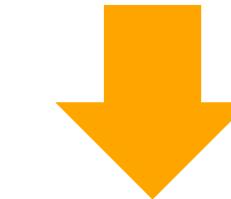
VECTORISATION DES CAPTIONS WORD2VEC

VECTORISATION DES INTERETS TF-IDF

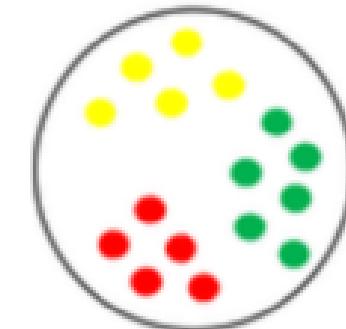
VECTEUR DU PROFILE UTILISATEUR

Matrice de similarité cosinus entre les profils utilisateurs

1	0.663	0.5551
0.663	1	0.5277
0.5551	0.5277	1

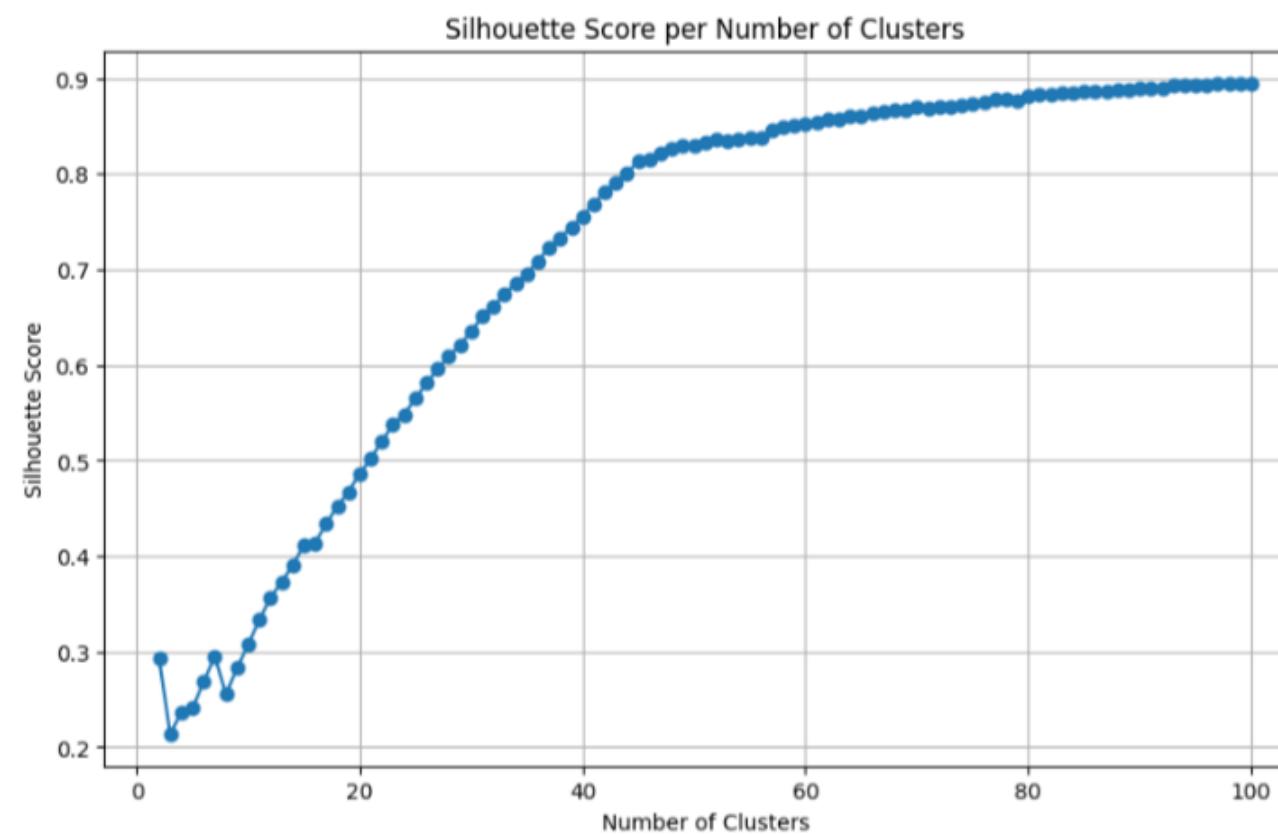


Clustering hiérarchiques des utilisateurs (CAH)

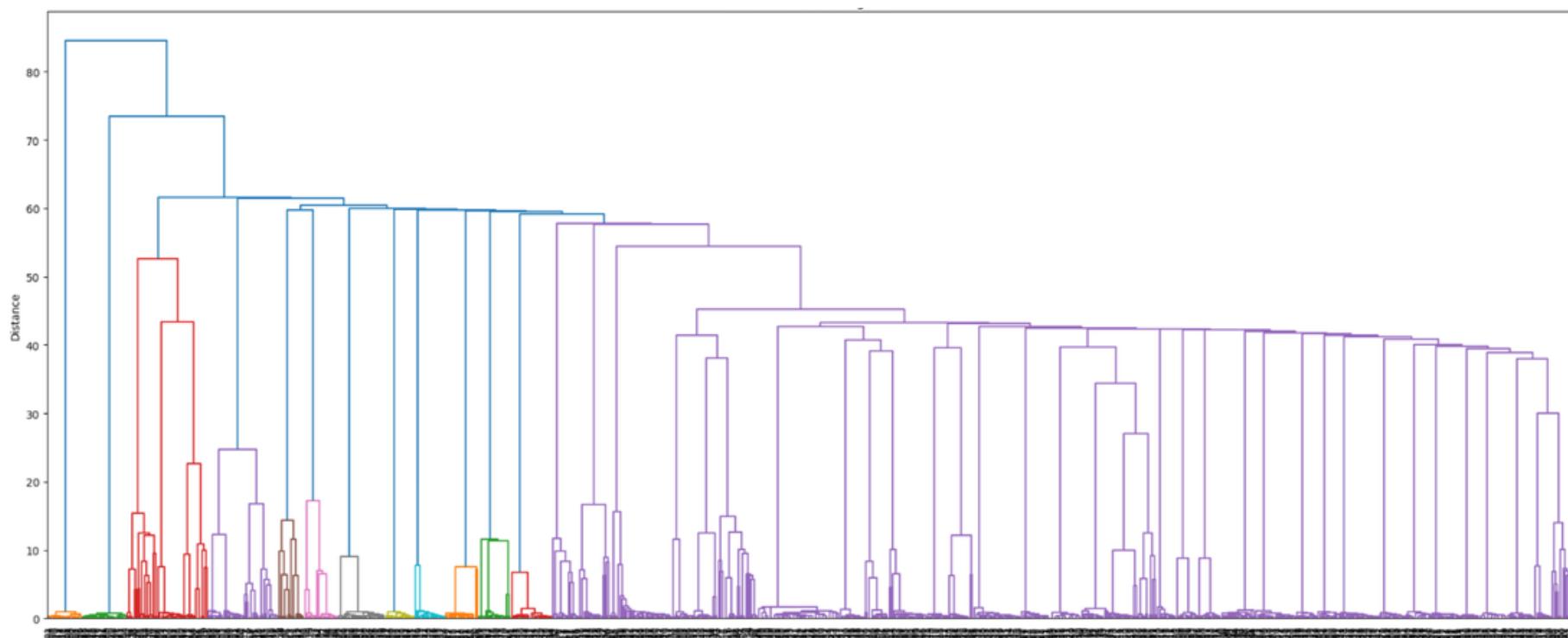


2. Clustering des utilisateurs en fonction de leurs profils

évolution du score de silhouette en fonction du nombre de cluster



Dendrogramme du clustering hiérarchique



Nombre idéale de cluster = 45

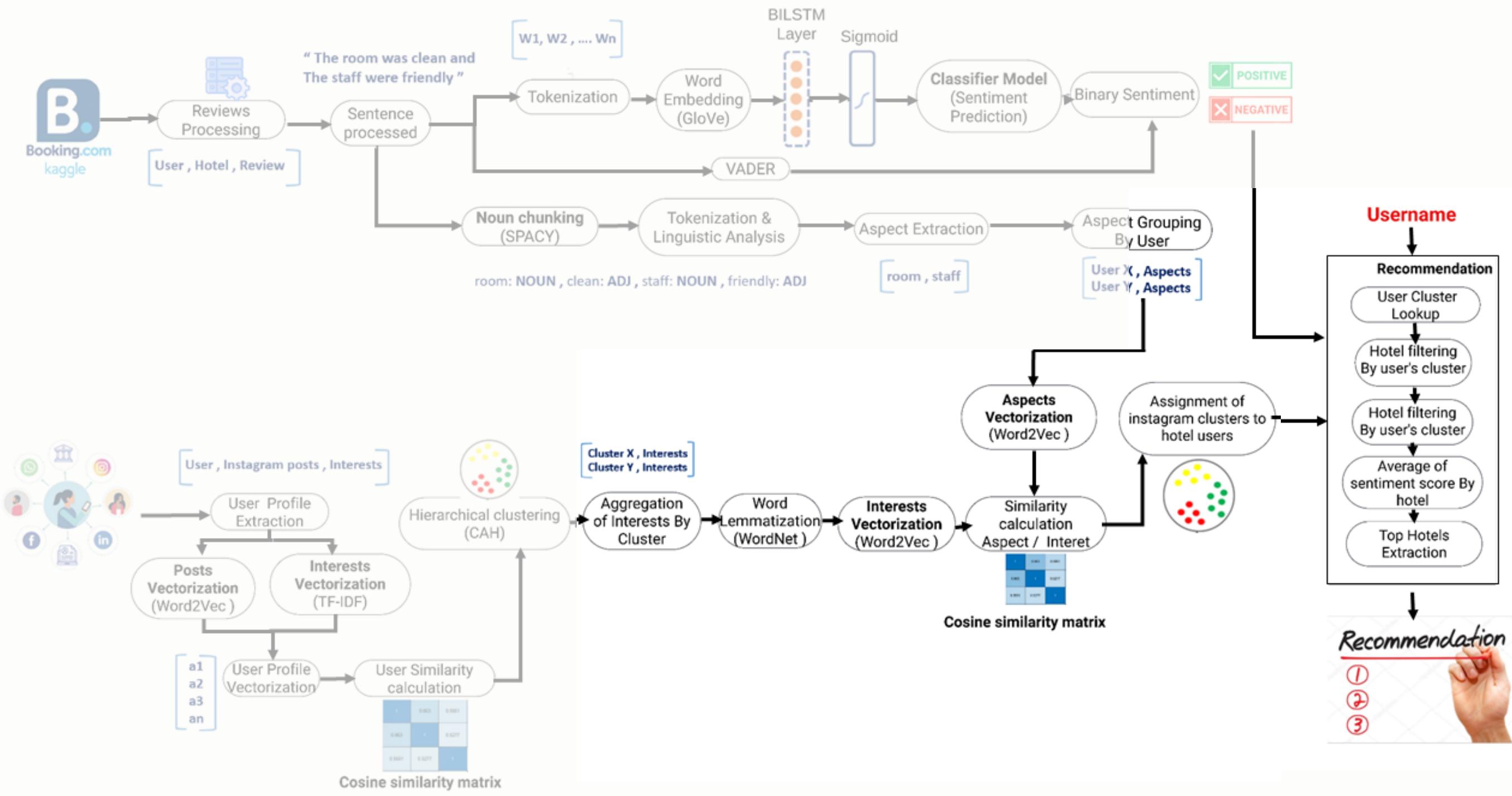
2. Clustering des utilisateurs en fonction de leurs profils

les utilisateurs du cluster 5

Users in the Cluster 5:

```
['Cristiano Ronaldo', 'David Beckham', 'Álvaro Morata', 'Álvaro  
Morata', 'Leo Messi', 'Jack Grealish', 'Rúben Gato Dias PT',  
'Nike', 'Real Madrid C.F.', 'FC Bayern', 'UEFA Champions League',  
'Kyle Walker', 'Premier League', 'LIVE HERE WE GO 🎉']
```

3. Matching & Recommendation



3. Matching & Recommandation

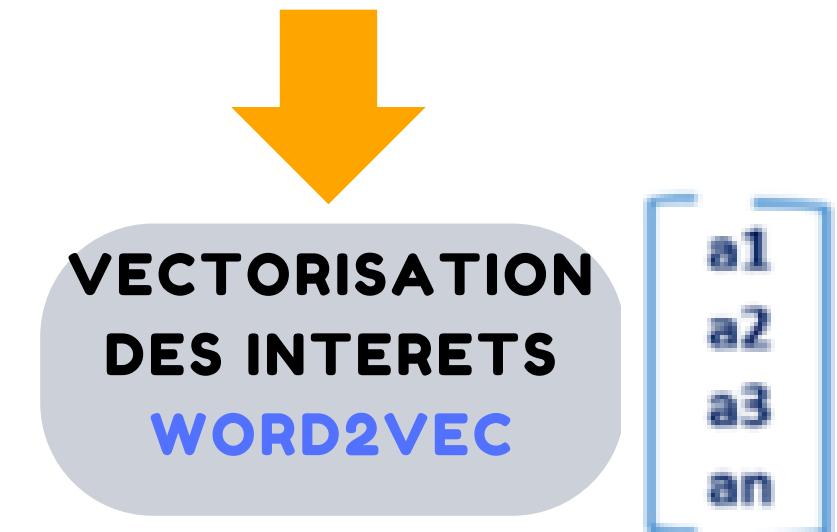
Agrégation des intérêts de chaque cluster

	cluster	interests_text
0	1	current event homes and garden
1	2	political and social issues
2	3	running, Physical exercise, social media, pare...
3	4	family, american football
4	5	basketball, dating and marriage, beauty, Physic...
5	6	shopping, social media, fashion accessories, c...
6	7	vehicles, Physical exercise, social media, Aut...
7	8	bodybuilding, Health care, parenting, social m...
8	9	alcoholic beverage, beverages
9	10	food and restaurant



Expansion des synonymes des aspects et des intérêts

	reviewed_by	aspects	expanded_aspects
0	Kyrylo	Solo traveller,Stayed 1 night,,Business trip,J...	{Jnr, persist, stayed, „ stick, I, take, stay...
1	Dimitri	Deluxe Suite,Executive Double Room with Sofa B...	{residue, persist, ii, jazz, state, get_Jaid, ...
2	Virginia	Couple,room Business trip,location Bus...	{Little_Joe, persist, ii, placement, state, bi...
3	Kannan	Solo traveller,stay,Business trip,Stayed 4 nig...	{Little_Joe, Jnr, persist, stick, „ stayed, t...
4	Sue	Couple,owner Leisure trip,Standard Queen Room,...	{Little_Joe, persist, ii, avail, drinking, pla...



3. Matching & Recommandation

VECTORISATION
DES INTERETS
WORD2VEC

[
a1
a2
a3
an
]

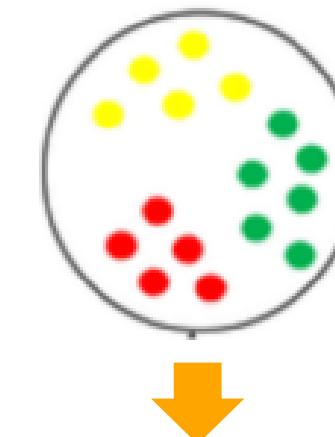
VECTORISATION
DES ASPECTS
WORD2VEC

[
a1
a2
a3
an
]

Matrice de similarité cosinus
entre aspects et interets

1	0.663	0.5551
0.663	1	0.5277
0.5551	0.5277	1

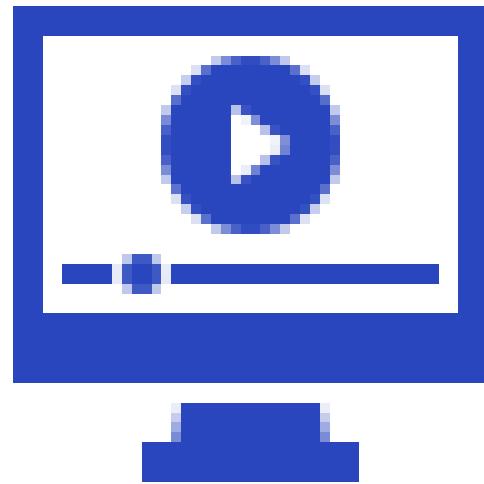
Affectation des clusters aux utilisateurs qui
ont séjourné dans des hôtels



	username	aspects	assigned_cluster	hotel_name
0	Kyrylo	Solo traveller,Stayed 1 night,Business trip,J...	25	Villa Pura Vida
1	Dimitri	Deluxe Suite,Executive Double Room with Sofa B...	19	Villa Pura Vida
2	Dimitri	Deluxe Suite,Executive Double Room with Sofa B...	19	Novotel Gent Centrum
3	Dimitri	Deluxe Suite,Executive Double Room with Sofa B...	19	Hotel Dolce La Hulpe Brussels
4	Dimitri	Deluxe Suite,Executive Double Room with Sofa B...	19	B&B Le flaneur
...
26381	Aynur	Stayed 1 night,Couple,room,Double Room,Submitt...	25	Hotel Du Congres
26382	Chrizzer	Studio (2 Adults),Stayed 1 night,Family with ...	44	Aparthotel Adagio Access Bruxelles Europe Apar...
26383	Olya	Solo traveller,place,Stayed 3 nights,Double Ro...	25	Logies Windsor
26384	Subodh	Solo traveller,Standard Double Room,,Stayed 2 ...	2	Hotel Ter Eist
26385	Zhl	Couple,Stayed 2 nights,Standard Double or Twin...	2	NH Brussels Carrefour de L'Europe
26386 rows × 6 columns				

3. Matching & Recommandation





5

DÉMONSTRATION

Technologies utilisées



Python pour la modélisation



Framework Django Pour Le Développement



Flask pour les Requêtes de Recommandation REST API

Interface Principale :

liste des Hôtels

[Acceuil](#) [test](#) [A propos](#) [Connexion](#)

Système De Recommandation Hôtelière

[AVOIR DES RECOMMANDATIONS](#)**Villa Pura Vida**

Sint-Denijsestraat 191C, 8500
Kortrijk, Belgium

[VOIR HOTEL](#)**Hydro Palace Apartment**

Zeedijk 250 bus 17, Mariakerke,
8400 Ostend, Belgium

[VOIR HOTEL](#)**Love Nest Versailles - ...**

39 Rue des Croisiers, 5000 Namur,
Belgium

[VOIR HOTEL](#)

Interface Principale : Accès directe à Booking.com

Accueil test A propos Connexion

Système De Recommandation Hôtelière

AVOIR DES RECOMMANDATIONS



a Vida
estraat 191C, 8500
lgium



Hydro Palace Apartment
Zeedijk 250 bus 17, Mariakerke,
8400 Ostend, Belgium

VOIR HOTEL



Love Nest Versailles - ...
39 Rue des Croisiers, 5000 Namur,
Belgium

VOIR HOTEL

https://www.booking.com/hotel/be/villa-pura-vida.fr.html?label=gen173nr-1FCAdoFUIPdmhsbGtchVvys12aWRhSDNYBGhiAEImAEI

Booking.com

TND  Ajoutez votre établissement

Hébergements Vols Voitures de location Attractions Taxis aéroport

Courtrai Date d'arrivée — Date de départ 2 adultes · 0 enfant · 1 chambre

Accueil > Hôtels > B&B/Chambres d'hôtes : toutes les options > Belgique > Flandre-Occidentale > Courtrai > Offres à l'établissement Villa Pura Vida (Belgique)

Vue d'ensemble Infos et tarifs Équipements Règles de la maison À savoir

 Villa Pura Vida

Sint-Denijsestraat 191C, 8500 Courtrai, Belgique – Excellent emplacement - voir la carte



Interface :

Authentification des

Utilisateurs

Système De Recommandation Hôtelière

AVOIR DES RECOMMANDATIONS

Saisir Votre Nom d'utilisateur



Nom d'utilisateur:

Heleen

Mot de passe :

....

ANNULER

AVOIR DES RECOMMANDATIONS



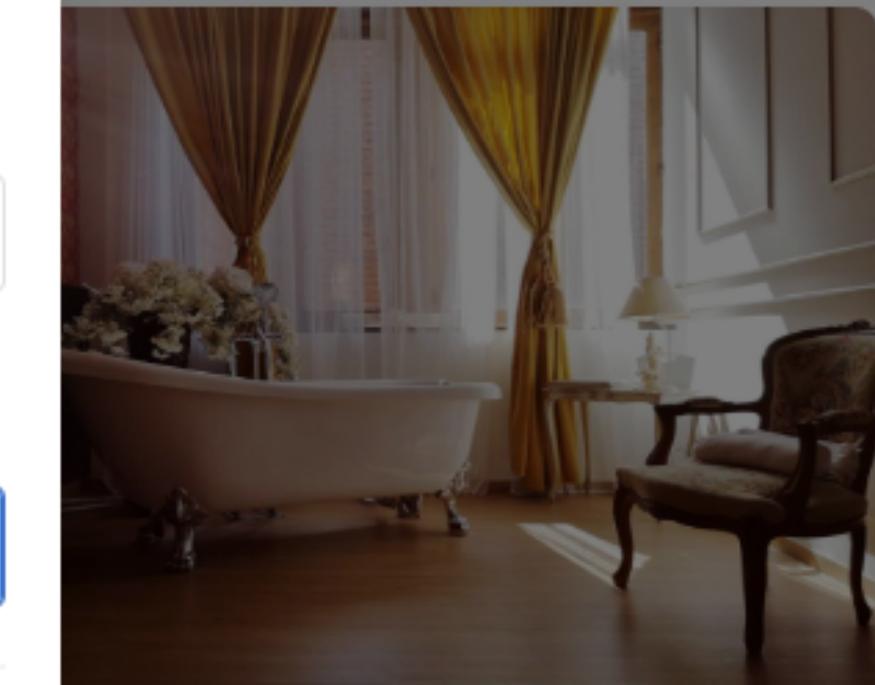
Villa Pura Vida

Sint-Denijsestraat 191C, 8500
Kortrijk, Belgium

VOIR HOTEL

8400 Ostend, Belgium

VOIR HOTEL



Nest Versailles - ...

Le des Croisiers, 5000 Namur,
Belgium

VOIR HOTEL

Interface: Recommandations Personnalisées

Système De Recommandation Hotelière

AVOIR DES RECOMMANDATIONS

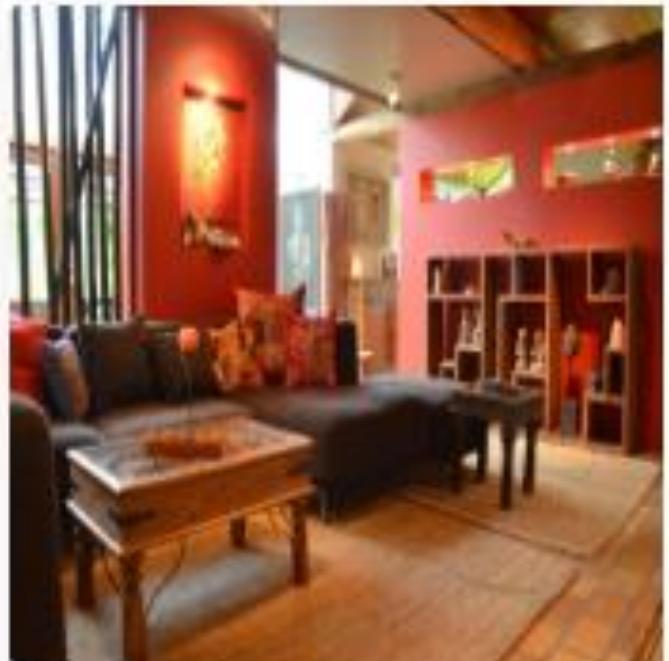


Heleen

Profile Utilisateur : breakfast,rain,tents,tent,setting

Interets: ['health care']

Recommandés pour vous



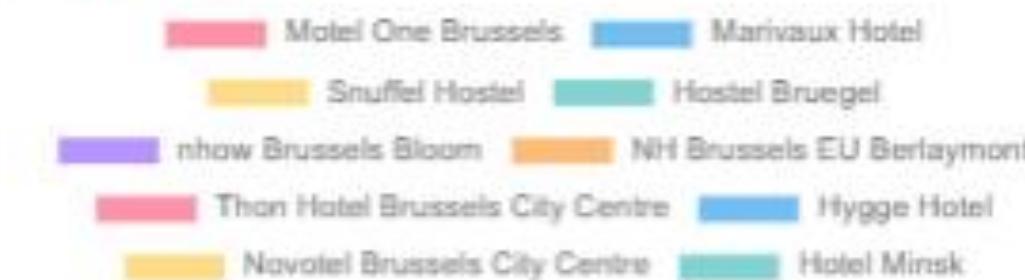
Les Bambous 9.3

7 Thier du Pierreux, 6980 La Roche-en-Ardenne, Belgium

Aspects: 'kitchen', 'location', 'bathroom'

VISITER L'HOTEL

Top 10 Hôtels Aimés par des Utilisateurs Similaires



Top 5 Utilisateurs Similaires

1. Adriana

Exemple d'utilisateur 1 “HELEN”

Centres d'interets : “health care”

Système De Recommandation Hôtelière

AVOIR DES RECOMMANDATIONS

 Heleen

Profile Utilisateur : breakfast,rain,tents,tent,setting

Interets: ['health care']

Recommandés pour vous

 **Les Bambous** 9.3
7 Thier du Pierreux, 6980 La Roche-en-Ardenne, Belgium
Aspects: 'kitchen', 'location', 'bathroom'
[VISITER L'HOTEL](#)

 **Le dernier métro** 8.1
Rue de l'église, 96, Sint-Pieters-Woluwe / Woluwe-Saint-Pierre, 1150 Brussels, Belgium
Aspects: 'location', 'wife', 'location', 'host', 'parking', 'tv', 'owner', 'host', 'breakfast', 'toilet', 'thing', 'host'
[VISITER L'HOTEL](#)

 **Hotel V butanezai tibu**

Top 10 Hôtels Aimés par des Utilisateurs Similaires

Hôtel	Score
Motel One Brussels	9.3
Marivaux Hotel	9.2
Snuffel Hostel	9.1
Hostel Bruegel	9.0
nhow Brussels Bloom	8.9
NH Brussels EU Berlaymont	8.8
Than Hotel Brussels City Centre	8.7
Hygge Hotel	8.6
Novotel Brussels City Centre	8.5
Hotel Minsk	8.4



Top 5 Utilisateurs Similaires

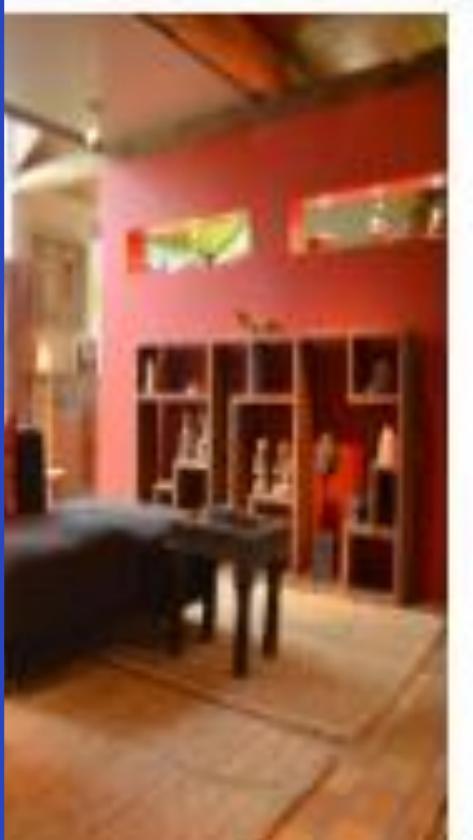
1. Adriana
staff,hotel , ,day , water, room, novotel, service
2. Volodymyr_p
shops, price, beds, balconies, ventilation, sealine, locals, check, reception
3. Projjal
owners, restaurants
4. Lidia



Heleen

Profile Utilisateur : breakfast,rain,tents,tent,setting

Interets: ['health care']



Les Bambous

9.3

7 Thier du Pierreux, 6980 La Roche-en-Ardenn

Aspects: 'kitchen', 'location', 'bathroom'

VISITER L'HOTEL



Le dernier métro

8.1

Rue de l'église, 96, Sint-Pieters-Woluwe / W
1150 Brussels, Belgium

Aspects: 'location', 'wife', 'location', 'host', 'p
'host', 'breakfast', 'toilet', 'thing', 'host'

VISITER L'HOTEL



Exemple d'utilisateur 2 “OLYA”

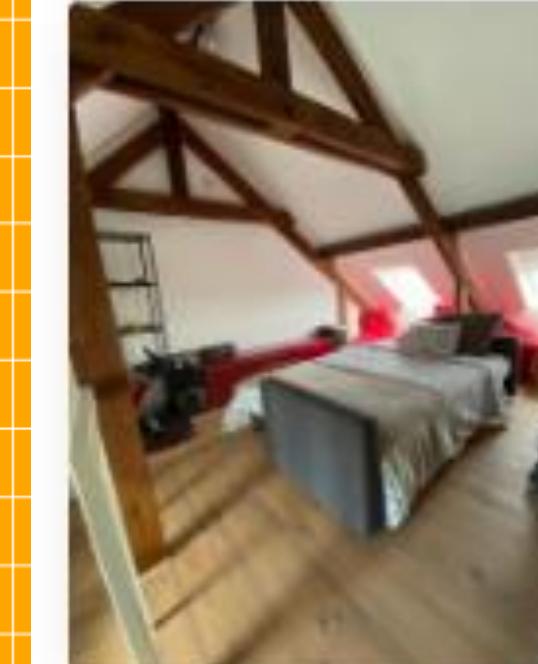
Centres d'interets : “ event homes and garden ”

Olya

Profile Utilisateur : place

Interets: [event homes and garden]

Recommandés pour vous

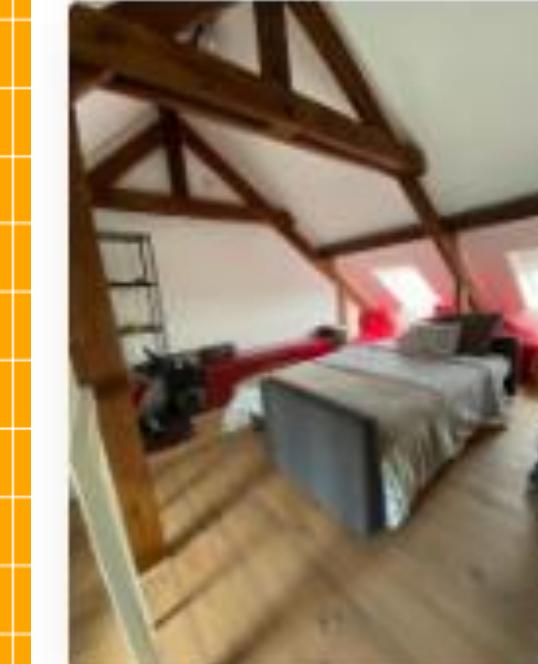
 appartement stylé dans maison bourgeoise

appartement stylé dans maison bourgeoise 10.0

Pas de localisation disponible

Aspects: 'house', 'bonus', 'carrefour'

[VISITER L'HOTEL](#)

 B&B Duplex apartment in old mansion 9.1

41 Toekomststraat, 3500 Hasselt, Belgium

Aspects: 'place', 'cosy', 'rooms', 'breakfast', 'hosts', 'sylvie'

[VISITER L'HOTEL](#)

Top 5

1. Jip

2. X

3. Mirald
staff,n



Olya

Profile Utilisateur : place

Interets: [event homes and garden]

Recommandés pour vous



ManavaHotel 9.2

Route de Liers 114, 4041 Herstal, Belgium

Aspects: 'hotel', 'location', 'room', 'people', 'owner',
'service'

VISITER L'HOTEL



appartement stylé dans maison bourgeoise

Pas de localisation disponible

Aspects: 'house', 'bonus', 'carrefour'

VISITER L'HOTEL



6

CONCLUSION ET PERSPECTIVES

CONCLUSION

- PRÉCISION DE L'ANALYSE SENTIMENTALE DE 93,44% ATTEINTE GRÂCE À L'UTILISATION DU MODÈLE BILSTM ET DES EMBEDDINGS GLOVE.
- RECOMMANDATIONS PERSONNALISÉES BASÉES SUR LES PROFILS ET INTÉRÊTS DES UTILISATEURS, GARANTISSANT UNE CORRESPONDANCE OPTIMALE ENTRE LES HÔTELS SUGGÉRÉS ET LES PRÉFÉRENCES INDIVIDUELLES.
- UN SCORE DE SILOUHETTE ELEVÉ DE L'ORDRE DE 0,82
- INTERFACE UTILISATEUR CONVIVIALE DÉVELOPPÉE OFFRANT UNE EXPERIENCE CONVIVIALE DE RECOMMANDATIONS PERSONNALISÉES.

PERSPECTIVES



ÉVALUER LE SYSTÈME À GRANDE ÉCHELLE

ENRICHIR LE SYSTÈME DE RECOMMANDATION

**EXPLORER D'AUTRES TECHNIQUES DE
MODÉLISATION**

**AMÉLIORER L'ANALYSE SENTIMENTALE ET L'EXTRACTION
D'ASPECTS
(APPRENTISSAGE MULTITÂCHE)**



MERCI POUR VOTRE
ATTENTION



RECOMMENDED